MATHEMATICS OF COMPUTATION
Volume 72, Number 241, Pages 355–373
S 0025-5718(02)01417-5
Article electronically published on February 22, 2002

ORTHOGONAL POLYANALYTIC POLYNOMIALS AND NORMAL MATRICES

MARKO HUHTANEN

Abstract. The Hermitian Lanczos method for Hermitian matrices has a well-known connection with a 3-term recurrence for polynomials orthogonal on a discrete subset of $\mathbb{R}$. This connection disappears for normal matrices with the Arnoldi method. In this paper we consider an iterative method that is more faithful to the normality than the Arnoldi iteration. The approach is based on enlarging the set of polynomials to the set of polyanalytic polynomials. Denoting by d the number of elements computed so far, the arising scheme yields a recurrence of length bounded by $\sqrt{8d}$ for polyanalytic polynomials orthogonal on a discrete subset of $\mathbb{C}$. Like this slowly growing length of the recurrence, the method preserves, at least partially, the properties of the Hermitian Lanczos method. We employ the algorithm in least squares approximation and bivariate Lagrange interpolation.

1. Introduction

The set of normal matrices $\mathcal{N} \subset \mathbb{C}^{n\times n}$, i.e., those $N \in \mathbb{C}^{n\times n}$ that satisfy the algebraic condition

$$NN^* - N^*N = 0, \tag{1.1}$$

is a rich class of matrices particularly well suited to numerical computations. This is largely due to the fact that normal matrices are unitarily diagonalizable. However, many algorithms are aimed exclusively at Hermitian matrices, that is, for a subset of $\mathcal{N}$. The most famous among them is probably the Hermitian Lanczos method yielding a unitary similarity transformation tridiagonalizing a given Hermitian matrix. At the same time there arises a connection with a 3-term recurrence for discrete orthogonal polynomials; see, e.g., [12].

If one does not consider the Toeplitz decomposition [17, 18], all the interesting properties of the Hermitian Lanczos method vanish for non-Hermitian normal matrices, as then the Arnoldi method [1] is the scheme yielding a unitary similarity. With this similarity transformation the tridiagonal structure cannot be preserved anymore, as the process yields (full) Hessenberg matrices in general. Simultaneously, the 3-term recurrence relation disappears for the resulting sequence of orthogonal polynomials as the length of the recurrence grows linearly with the iteration number.

Received by the editor September 12, 2000 and, in revised form, March 1, 2001.
2000 *Mathematics Subject Classification.* Primary 42C05; Secondary 15A57.
Key words and phrases. Normal matrix, orthogonal polynanalytic polynomial, slowly growing length of the recurrence, least squares approximation, bivariate Lagrange interpolation.
This work was supported by the Academy of Finland and the Alfred Kordelin Foundation.

In this paper we construct an appropriate structure around the condensed form for normal matrices of Elsner and Ikramov [10] to partially preserve the properties of the Hermitian Lanczos method. We obtain an iterative scheme for normal matrices which can be used to generate orthogonal functions with better recurrence properties than does the Arnoldi method. The set of polynomials $\mathcal{P}$ is simply not a sufficiently large class of functions to this end. An appropriate enlargement of $\mathcal{P}$ is the set of polyanalytic polynomials. Denoting by $\mathcal{P}_k$ the set of polynomials of degree k at most, polyanalytic polynomials are functions of the form

$$p(z) = \sum_{j=0}^{k} h_j(z)\overline{z}^j, \tag{1.2}$$

with $h_j \in \mathcal{P}_{k-j}$ and $k \in \mathbb{N}_0 = \{0\} \cup \mathbb{N}$. If, for some j, we have $\deg(h_j) = k - j$ for p in (1.2), then the degree of p is defined to be k. We use the notation $\mathcal{PP}_k$ for polyanalytic polynomials of degree k at most, and we set $\mathcal{PP} = \bigcup_{k\in\mathbb{N}_0} \mathcal{PP}_k$. Because of the commutativity relation (1.1), $p(N)$ is well-defined for any $p \in \mathcal{PP}$ by identifying z and $\overline{z}$ with N and N^* respectively.

Polyanalytic polynomials of the form $z^j\overline{z}^l$ are called polyanalytic monomials. To get a method with properties resembling more the Hermitian Lanczos than the Arnoldi method, we set an order $>$ among them as follows. Let $z^{j_1}\overline{z}^{l_1}$ and $z^{j_2}\overline{z}^{l_2}$ be two polyanalytic monomials. If $j_1 + l_1 > j_2 + l_2$, then $z^{j_1}\overline{z}^{l_1} > z^{j_2}\overline{z}^{l_2}$. If $j_1 + l_1 = j_2 + l_2$ and $j_1 > j_2$, then $z^{j_1}\overline{z}^{l_1} > z^{j_2}\overline{z}^{l_2}$. Using this order, we define the minimal polyanalytic polynomial of $N \in \mathcal{N}$ to be the monic, i.e., its leading term is $z^j\overline{z}^l$, of the least possible degree annihilating N. A property of the minimal polyanalytic polynomial of N is that its zero set contains the spectrum of N yielding, thereby, a spectral exclusion set for N. Furthermore, by employing this order, an Arnoldi type of iterative method can be introduced for normal matrices.

Recall that in the Arnoldi method, for a matrix $A \in \mathbb{C}^{n\times n}$ and a vector $\hat{q}_0 \in \mathbb{C}^n$, a monic polynomial $P_k(z) = z^k - p(z)$ of degree k is computed realizing

$$\|A^k\hat{q}_0 - p(A)\hat{q}_0\| = \min_{\hat{p}\in\mathcal{P}_{k-1}} \|A^k\hat{q}_0 - \hat{p}(A)\hat{q}_0\| \tag{1.3}$$

while, simultaneously, an orthonormal basis of $\mathbb{C}^n$ is generated. Now, for a normal $N \in \mathbb{C}^{n\times n}$ and for $j + l = k$, we set analogously

$$\|N^jN^{*l}\hat{q}_0 - p(N)\hat{q}_0\| = \min_{\hat{p}<z^j\overline{z}^l} \|N^jN^{*l}\hat{q}_0 - \hat{p}(N)\hat{q}_0\| \tag{1.4}$$

by using the order introduced. The corresponding monic polyanalytic polynomial of interest equals $P_{j,l}(z) = z^j\overline{z}^l - p(z)$, yielding a sequence $P_{j,l} \in \mathcal{PP}_k$, with $j + l = k$ for $k \in \mathbb{N}_0$.

The computation of the $P_{j,l}$'s can be carried out iteratively by generating an orthonormal basis of $\mathbb{C}^n$. This is achieved by multiplying an already computed vector with either N or N^* in a particular order. Then a new vector is obtained by orthogonalizing the vector obtained against the vectors computed so far. This yields a realization of the canonical form for N of Elsner and Ikramov. Simultaneously, by identifying multiplications by N and N^* with z and $\overline{z}$ respectively, the process gives rise to a sequence of orthogonal polyanalytic polynomials. The length of the recurrence to this end is not fixed three, nor does it grow linearly with the iteration number as is the case with the Hermitian Lanczos and the Arnoldi method respectively. Denoting by d the number of the orthogonal polyanalytic polynomials computed so far, the length of the recurrence is bounded by $\sqrt{8d}$; see Theorem 3.8

for the exact statement. Although a very modest growth, this is the worst case estimate. If the spectrum of N is a subset of an algebraic curve of degree k, then the length of the recurrence reaches at most $2k$.

The case of the spectrum of N being a subset of an algebraic curve of degree k, for a reasonably small k, is of particular interest on its own. Namely then, after at most $\frac{(k+1)(k+2)}{2}$ steps, zero is attained in the minimization problem (1.4) and the process has yielded the minimal polyanalytic polynomial of N. However, as opposed to the Arnoldi method, the described orthogonalization process is designed to continue beyond having zero in (1.4). This is realized in such a way that none of the generated orthogonal polyanalytic polynomials vanish on the eigenvalues of N. More precisely, the algorithm discards elements annihilating the spectrum giving rise, in this manner, to a subclass of $\mathcal{PP}$. To give an example, if N is Hermitian, then the scheme reduces to the Hermitian Lanczos method yielding discrete orthogonal polynomials satisfying a 3-term recurrence relation. This property can be employed in a number of classical problems of numerical analysis. In this paper we consider least squares approximation and bivariate Lagrange interpolation. As opposed to univariate interpolation, it is well-known that bivariate interpolation is a far more complicated problem; see, e.g., [3, 4, 25] and references therein.

Start with n interpolation nodes belonging to $\mathbb{R}^2$. By identifying them with the corresponding points of $\mathbb{C}$, form a diagonal matrix $N \in \mathcal{N}$ with the interpolation nodes on the diagonal. Then generate with N a sequence of orthogonal polynanalytic polynomials as just described. Computing Fourier coefficients simultaneously yields least squares approximations in a straightforward manner to any set of values associated with the interpolation nodes. If n is very large, then this may suffice before completing the iteration process. After n steps the method yields an interpolant whose uniqueness in the generated subclass of polyanalytic polynomials is guaranteed by the fact that the algorithm does not generate nonzero elements of $\mathcal{PP}$ vanishing on the interpolation nodes. Like being of minimal degree, this subclass has all the relevant properties needed in bivariate Lagrange interpolation.

The paper is organized as follows. In Section 2 we introduce the minimal polyanalytic polynomial of a normal matrix. Based on a local version of this, in Section 3 we consider an Arnoldi type of minimization problem and the corresponding iterative scheme for normal matrices. This gives rise to a realization of the canonical form for normal matrices of Elsner and Ikramov. The realization is carried out in such a way that there arises a connection with a slowly growing recurrence for orthogonal polyanalytic polynomials. This is considered in Section 4, and in Section 5 we apply the algorithm to least squares approximation and bivariate interpolation.

2. The minimal polyanalytic polynomial of a normal matrix

As described in the introduction, the Hermitian Lanczos method cannot be executed with a general normal matrix N. Then a standard approach is to employ the Arnoldi method, which, being a general method, does not use the normality in any reasonable way. As a consequence, no short term recurrence relation for discrete orthogonal polynomials arises. A way to circumvent this problem is to consider an enlargement of the set of polynomials to have an appropriate structure that is more faithful to the normality. Namely, since N commutes with its adjoint, the set of polynomials is not the largest class of functions that can be associated with $\mathcal{N}$. Identifying multiplications by N and N^* with z and $\overline{z}$ respectively renders the

usage of the polyanalytic polynomials (1.2) equally natural. The set $\mathcal{PP}$ is a subset of polyanalytic functions extensively studied in [2] by Balk. Although $\mathcal{P} \subset \mathcal{PP}$, these two classes of functions obviously have quite different properties.

In contrast to polynomials, equations involving polyanalytic polynomials can have none, a discrete set, or an infinite number of solutions.

Example 2.1. By the fundamental theorem of algebra, for any non-constant $p \in \mathcal{P}$, the equation $p(z) = 0$ has at least one solution and the number of solutions is less than or equal to $\deg(p)$, the degree of p. Consider the polyanalytic polynomial $q(z) = z\overline{z} + 1$. Obviously, the equation $q(z) = 0$ has no solutions. On the other hand, if $w(z) = z\overline{z} - 1$, then the solution set of $w(z) = 0$ is the unit circle, i.e., a continuum. Another example of a polyanalytic polynomial having a large zero-set is $v(z) = z + \overline{z}$, which annihilates the whole imaginary axis.

Among the polyanalytic monomials we use the order defined in the introduction. Then the leading term $\mathrm{LT}(p)$ of a polyanalytic polynomial p is defined in the natural way. Let us illustrate this with an example.

Example 2.2. For $p(z) = i\overline{z}^3 + z\overline{z}^2 + (1-i)z^2 - 3\overline{z}$ the leading term is $\mathrm{LT}(p) = z\overline{z}^2$.

With an order on $\mathcal{PP}$ we define $p \in \mathcal{PP}$ to be monic if $\mathrm{LT}(p) = z^j\overline{z}^l$, that is, if the leading term of p is a polyanalytic monomial. For instance, p in Example 2.2 is a monic polyanalytic polynomial. Recall that the minimal polynomial of (a not necessarily normal) $A \in \mathbb{C}^{n\times n}$ is the monic polynomial of least degree annihilating A. For normal matrices it is completely natural to define the following.

Definition 2.3. A minimal polyanalytic polynomial of $N \in \mathcal{N}$ is a monic $p \in \mathcal{PP}$ of least possible degree annihilating N.

Example 2.4. Assume $N \in \mathbb{C}^{n\times n}$ is unitary with at least 4 separate eigenvalues. Then a minimal polyanalytic polynomial of N is $p(z) = z\overline{z} - 1$. On the other hand, if N is Hermitian, then $p(z) = z - \overline{z}$ is a minimal polyanalytic polynomial of N.

A minimal polyanalytic polynomial is unique as well as unitarily invariant.

Proposition 2.5. *Every $N \in \mathcal{N}$ possesses a unique minimal polyanalytic polynomial. Moreover, unitarily similar $N_1, N_2 \in \mathcal{N}$ have the same minimal polyanalytic polynomial.*

Proof. Let $p(z) = z^j\overline{z}^l +$ *lower terms* and $q(z) = z^j\overline{z}^l +$ *lower terms* be two minimal polyanalytic polynomials of $N \in \mathcal{N}$. Then subtracting them yields either a zero polyanalytic polynomial or a smaller nonzero polyanalytic polynomial than p or q. In the latter case this $p-q$ would yield an annihilating polyanalytic polynomial of N as well, which contradicts the minimality of p and q.

If N_1 and N_2 are unitarily similar, say $N_1 = UN_2U^*$, then

$$p_2(N_1) = Up_2(N_2)U^* = 0,$$

where p_2 is the minimal polyanalytic polynomial of N_2. Thus the minimal polyanalytic polynomial of N_2 is also the minimal polyanalytic polynomial of N_1, and conversely. □

An algorithm for finding the minimal polyanalytic polynomial of N will be presented in the following section.

We denote the minimal polyanalytic polynomial of a given $N \in \mathcal{N}$ by $p_{j,l}$, where $\mathrm{LT}(p_{j,l}) = z^j\overline{z}^l$. The following is obvious.

Proposition 2.6. *Let $p_{j,l}$ be the minimal polyanalytic polynomial of $N \in \mathcal{N}$. Then $q_{j,l}(z) := \overline{p_{j,l}(\overline{z})}$ is the minimal polyanalytic polynomial of N^*.*

The degree of $p_{j,l}$ can be bounded as follows, where we denote by $\deg(A)$ the degree of the minimal polynomial of a square matrix A.

Theorem 2.7. *Let $p_{j,l}$ be the minimal polyanalytic polynomial of $N \in \mathcal{N}$. Then $\deg(p_{j,l}) \leq \sqrt{2\deg(N)}$.*

Proof. See the proof of Theorem 3.1 in Section 3. □

By $\sigma(A)$ we denote the spectrum of a square matrix A. The zero set of the minimal polynomial of a matrix equals its spectrum, whereas with the minimal polyanalytic polynomial we only have an inclusion.

Theorem 2.8. *Assume $N \in \mathcal{N}$ and $p \in \mathcal{PP}$. Then $\sigma(N) \subset \{z \in \mathbb{C} : |p(z)| \leq \|p(N)\|\}$.*

Proof. A simple way to see this is to recall that, since N is normal, $N^* = q(N)$ for a polynomial q. Therefore, inserting this into $p(N) \equiv p(N, N^*)$ in place of N^*, we have $\|p(N, N^*)\| = \max_{\lambda \in \sigma(N)} |p(\lambda, q(\lambda))|$, which yields the claim. □

See also [21, Section 2.10] for this concept when p is polynomial.

Since for normal matrices the norm equals the specral radius, we have the following.

Corollary 2.9. *Assume $N \in \mathcal{N}$ and $p \in \mathcal{PP}$. Then the boundary of $\{z \in \mathbb{C} : |p(z)| \leq \|p(N)\|\}$ contains an eigenvalue of N.*

Corollary 2.10. *Assume $p_{j,l}$ is the minimal polyanalytic polynomial of $N \in \mathcal{N}$. Then the eigenvalues of N are contained in the zero-set of $p_{j,l}$.*

Consequently, consider those $N \in \mathcal{N}$ whose eigenvalues lie on a given algebraic curve. Then the polyanalytic polynomial corresponding to the algebraic curve annihilates N regardless of the number of the distinct eigenvalues of N. In contrast, the degree of the minimal polynomial of a normal matrix N is always the cardinality of $\sigma(N)$.

Example 2.11. Assume the spectrum of a normal matrix N is contained in the parabola $y = x^2$, i.e., in the zero set of $p(z) = \frac{1}{2i}(z - \overline{z}) - \frac{1}{4}(z^2 + 2z\overline{z} + \overline{z}^2)$. Thus the minimal polyanalytic polynomial of N is $p_{2,0}(z) = z^2 + 2z\overline{z} + \overline{z}^2 + 2iz - 2i\overline{z}$ as long as $\sigma(N)$ has at least 5 separate points.

Example 2.12. Assume the spectrum of a normal matrix N is contained in the set defined by the limacon $r = 1 - \cos(\theta)$. Then

$$p_{2,2}(z) = z^2\overline{z}^2 - z^2\overline{z} - z\overline{z}^2 + \frac{1}{4}z^2 - \frac{1}{2}z\overline{z} - \frac{1}{4}\overline{z}^2$$

is the minimal polyanalytic polynomial of N (as long as $\sigma(N)$ has sufficiently many points).

Thus, finding the minimal polyanalytic polynomial of $N \in \mathcal{N}$ yields a spectral exclusion set after solving for the zero set. Recall that, in Bauer's terminology, an exclusion set for the eigenvalues of N contains no eigenvalues, hence, its complement contains all the eigenvalues of N; see, e.g., [16, Chapter 19]. As we have seen in Examples 2.11 and 2.12, this can be a continuum, so that the difference can be

genuine. This does not mean that the information cannot be far more precise. In Examples 2.11 and 2.12 the imaginary part of the minimal polyanalytic polynomial vanishes. Obviously, if the minimal polyanalytic polynomial has linearly independent real and imaginary parts, then the intersection of their zero sets contains the spectrum.

Example 2.13. Let $N = \mathrm{diag}(1, 0, i) \in \mathbb{C}^{3\times 3}$. The minimal polyanalytic polynomial of N is $p_{0,2}(z) = \overline{z}^2 - (\frac{1}{2} + \frac{i}{2})z + (\frac{-1}{2} + \frac{i}{2})\overline{z} = x^2 - y^2 - x + y - 2xyi$. Its real part can be factored as $(x - y)(x + y - 1)$, for which the zero set is the union of the lines $y = x$ and $y = -x + 1$. The zero set of the imaginary part is clearly the union of the real and imaginary axes. Intersecting these yields the set $\{0, 1, i\}$, which equals $\sigma(N)$. Thus, *exact* information was obtained by considering second degree bivariate polynomials. The degree of the minimal polynomial of N is 3.

Denote by $\Gamma(N)$ the zero set of the minimal polyanalytic polynomial of $N \in \mathcal{N}$. First, $\Gamma(N)$ cannot have interior points in $\mathbb{C}$. Although the length of $\Gamma(N)$ may be infinite, a relevant portion of it can be bounded as follows.

Theorem 2.14. *Let $N \in \mathcal{N}$ and assume a disk D of radius R contains $\sigma(N)$. Let $p_{j,l} = \Re(p_{j,l}) + i\Im(p_{j,l})$ be the minimal polyanalytic polynomial of N, and let d be the smallest strictly positive value of* $\deg(\Re(p_{j,l}))$ *and* $\deg(\Im(p_{j,l}))$. *Then the length of $\Gamma(N) \cap D$ is at most $2\pi Rd$.*

Proof. This is a consequence of Poincaré's formula [24], proved in [5] by Borwein, which we apply to the case $\Gamma(N) \cap D$. Since the degree of an algebraic curve annihilating $\sigma(N)$ is d, the claim follows. □

Note the significance of the square root in Theorem 2.7 giving $d \leq \sqrt{2n}$, which is a very modest growth (as a function of the dimension n) for the length of $\Gamma(N) \cap D$.

3. An Arnoldi type of minimization problem for normal matrices

The classical Arnoldi method for eigenvalue approximation starts from the construction of a Krylov subspace and the corresponding orthonormal basis with (a not necessarily normal) $A \in \mathbb{C}^{n\times n}$ and a vector $\hat{q}_0 \in \mathbb{C}^n$. Then a monic polynomial P_k of degree k is computed with the property that $\|P_k(A)\hat{q}_0\|$ is minimized over all monic polynomials of degree k, that is,

$$\|A^k\hat{q}_0 - p(A)\hat{q}_0\| = \min_{\hat{p}\in\mathcal{P}_{k-1}} \|A^k\hat{q}_0 - \hat{p}(A)\hat{q}_0\|. \tag{3.1}$$

Sometimes $P_k(z) = z^k - p(z)$ is called the Arnoldi polynomial of degree k of A at $\hat{q}_0$. The underlying idea of the Arnoldi method is that P_k yields, in a sense, an approximation to the minimal polynomial of A. Since the zero set of the minimal polynomial equals the spectrum of A, the zero set of this approximation is used in eigenvalue approximation. For more information concerning the computational aspects of the Arnoldi method for eigenvalue problems, see, e.g., [23].

An order among the polyanalytic monomials is needed for defining an Arnoldi type of iterative method for normal matrices such that the normality is genuinely employed. This is achieved by considering a minimization problem analogous to (3.1) as follows. For a normal N and for $j + l = k$ compute

$$\|N^jN^{*l}\hat{q}_0 - p(N)\hat{q}_0\| = \min_{\hat{p}<z^j\overline{z}^l} \|N^jN^{*l}\hat{q}_0 - \hat{p}(N)\hat{q}_0\|, \tag{3.2}$$

for a vector $\hat{q}_0 \in \mathbb{C}^n$. Here $\hat{p} < z^j\overline{z}^l$ means that the leading term of $\hat{p}$ is strictly less than $z^j\overline{z}^l$. The resulting monic polyanalytic polynomial equals $P_{j,l}(z) = z^j\overline{z}^l - p(z)$, whose zero set can be taken to approximate the spectrum of N by the same reasoning as P_k can be taken in (3.1).

Since N is normal, $\|Nx\| = \|N^*x\|$ holds for all $x \in \mathbb{C}^n$. Consequently, the chosen order does not have any effect on the behavior of the minima (3.2). More precisely, we could have defined $z^{j_1}\overline{z}^{l_1} < z^{j_2}\overline{z}^{l_2}$ instead, when $j_1 + l_1 = j_2 + l_2$ and $j_1 > j_2$. The numbers (3.2) would still be the same.

Let us now demonstrate how to compute (3.2) as well as the $P_{j,l}$'s in practice. To this end the multiplications by N and N^* need to be ordered appropriately to take all the vectors of the form $N^jN^{*l}\hat{q}_0$, for $j, l \in \mathbb{N}_0$, into account. Also, the corresponding orthogonalization process must be realized in a numerically stable way while, simultaneously, keeping track of the resulting polyanalytic polynomials. For that purpose, with a $\hat{q}_0 \in \mathbb{C}^n$ of unit length, form a table

$$
\begin{array}{cccccccccccccccc}
I & \overbrace{N^*\ N} & & \overbrace{N^*\ N\ N} & & & \overbrace{N^*\ N\ N\ N} & & & & \overbrace{N^*\ N\ N\ N\ N} & & & & & \cdots \\
\hat{q}_0 & \hat{q}_0 & \hat{q}_0 & \hat{q}_1 & \hat{q}_1 & \hat{q}_2 & \hat{q}_3 & \hat{q}_3 & \hat{q}_4 & \hat{q}_5 & \hat{q}_6 & \hat{q}_6 & \hat{q}_7 & \hat{q}_8 & \hat{q}_9 & \cdots \\
\hat{q}_0 & \hat{q}_1 & \hat{q}_2 & \hat{q}_3 & \hat{q}_4 & \hat{q}_5 & \hat{q}_6 & \hat{q}_7 & \hat{q}_8 & \hat{q}_9 & \hat{q}_{10} & \hat{q}_{11} & \hat{q}_{12} & \hat{q}_{13} & \hat{q}_{14} & \cdots
\end{array}
\tag{3.3}
$$

which is read from left to right. The first row indicates the multiplying matrix. The second row indicates the numbering of the vector which is multiplied by the matrix in the first row above. And the third row indicates, after orthogonalizing against all the previous vectors in the third row and scaling by its length, the numbering of the resulting new vector. The purpose of the over-braces is to make the proposed rule more clear, as there are several ways to arrange the orthogonalizations. For $k \geq 1$ each "cycle" indicated by the over-braces consists of $k+1$ vectors. The rule is such that, to obtain the k^{th} cycle, the first vector produced in the $(k-1)^{th}$ cycle is multiplied once by N^*. Then all the vectors of the $(k-1)^{th}$ cycle are multiplied once by N starting from the first vector. Consequently, after k cycles have been completed there are $\frac{(k+1)(k+2)}{2}$ vectors in all.

With the ordering (3.3) all the powers $N^jN^{*l}\hat{q}_0$ get covered in a numerically stable way, and we have obtained an Arnoldi type of iterative method for normal matrices. We stress that we have not ruled out the possibility of having zero vectors among $\hat{q}_j$'s. For instance, if N is Hermitian, then many of these vectors equal zero. In this manner we can readily keep track of the generated polyanalytic polynomials. Obviously, if each cycle produces $k+1$ new linearly independent vectors, then the number of cycles is bounded according to $\frac{(k+1)(k+2)}{2} \leq n$.

A pseudo-code for the process (3.3) for computing the orthogonalizations as well as all the relevant quantities is below. This particular implementation is from [19].

ALGORITHM 1.
Assume $N \in \mathcal{N}$, and $\hat{q}_0 \in \mathbb{C}^n$, is of unit length.
for $k = 1 : K$
  Define $k_0 \equiv (k-2)(k-1)/2$, $k_1 \equiv (k-1)k/2$, $k_2 \equiv k(k+1)/2$
  $q_{k_2} = N^*\,\hat{q}_{k_1}$
  for $s = k_0 : k_2 - 1$
    $\alpha_s^{k_1} = (q_{k_2}, \hat{q}_s), \quad q_{k_2} = q_{k_2} - \alpha_s^{k_1}\hat{q}_s$
  end
  $\alpha_{k_2}^{k_1} = \|q_{k_2}\|$

if $\alpha_{k_2}^{k_1} > 0$ **then**
  $\hat{q}_{k_2} = (1/\alpha_{k_2}^{k_1})\, q_{k_2}$
else
  $\hat{q}_{k_2} = 0$
for $l = k_1 : k_2 - 1$
  $q_{k+l+1} = N\,\hat{q}_l$
  for $s = l - k + 1 : k + l$
    $\beta_s^l = (q_{k+l+1}, \hat{q}_s), \quad q_{k+l+1} = q_{k+l+1} - \beta_s^l \hat{q}_s$
  end
  $\beta_{k+l+1}^l = \|q_{k+l+1}\|$
  if $\beta_{k+l+1}^l > 0$ **then**
    $\hat{q}_{k+l+1} = (1/\beta_{k+l+1}^l)\, q_{k+l+1}$
  else
    $\hat{q}_{k+l+1} = 0$
end
end

We state a number of basic properties of the orthogonalization process (3.3) and the corresponding Algorithm 1. First of all, the reason for arranging the orthogonalizations as suggested is to be able to control the leading terms of the polyanalytic polynomials yielding $\hat{q}_j$'s.

Theorem 3.1. *Assume $N \in \mathcal{N}$ and $\hat{q}_0 \in \mathbb{C}^n$, and define $l(k,s) = \frac{k(k+1)}{2} + s$. Then the process (3.3) yields $\hat{q}_{l(k,s)} = p_{s,k-s}(N)\hat{q}_0$ with $\mathrm{LT}(p_{s,k-s}) = c_{l(k,s)} z^s \overline{z}^{k-s}$ and $c_{l(k,s)} \in \mathbb{C}$ for $k \in \mathbb{N}$ and $0 \le s \le k$.*

Proof. We identify multiplications by N and N^* with z and $\overline{z}$ respectively. Then, by using the ordering of the table (3.3), the recurrence can be written as

$$p_{0,0} = 1 \tag{3.4}$$

$$\alpha_1^0 p_{0,1} = \overline{z} p_{0,0} - (\overline{z} p_{0,0}, p_{0,0}) p_{0,0}, \tag{3.5}$$

$$\beta_2^0 p_{1,0} = z p_{0,0} - (z p_{0,0}, p_{0,1}) p_{0,1} - (z p_{0,0}, p_{0,0}) p_{0,1}, \tag{3.6}$$

$$\begin{aligned} \alpha_3^1 p_{0,2} = {} & \overline{z} p_{0,1} - (\overline{z} p_{0,1}, p_{1,0}) p_{1,0} \\ & - (\overline{z} p_{0,1}, p_{0,1}) p_{0,1} - (\overline{z} p_{0,1}, p_{0,0}) p_{0,0}, \end{aligned} \tag{3.7}$$

$$\begin{aligned} \beta_4^1 p_{1,1} = {} & z p_{0,1} - (z p_{0,1}, p_{0,2}) p_{0,2} - (z p_{0,1}, p_{1,0}) p_{1,0} \\ & - (z p_{0,1}, p_{0,1}) p_{0,1} - (z p_{0,1}, p_{0,0}) p_{0,0}, \end{aligned} \tag{3.8}$$

$$\begin{aligned} \beta_5^2 p_{2,0} = {} & z p_{1,0} - (z p_{1,0}, p_{1,1}) p_{1,1} - (z p_{1,0}, p_{0,2}) p_{0,2} \\ & - (z p_{1,0}, p_{1,0}) p_{1,0} - (z p_{1,0}, p_{0,1}) p_{0,1} - (z p_{1,0}, p_{0,0}) p_{0,0}, \end{aligned} \tag{3.9}$$

etc., where the constants $\alpha_j^l, \beta_j^l \in \mathbb{R}$ equal the norms of the right-hand sides (the choice of the indices will become clear later in this section). Here we have used the notation

$$(p, q) := (p(N)\hat{q}_0, q(N)\hat{q}_0) \tag{3.10}$$

for the inner product involving polyanalytic polynomials p and q in N applied to $\hat{q}_0$. Now, assume the claim is true in the $(k-1)^{th}$ cycle. Then, by induction, we obtain for the leading term

$$\mathrm{LT}(\alpha^{k-1}_{\frac{k(k+1)}{2}} p_{0,k}) = \mathrm{LT}(\overline{z} p_{0,k-1} + \ \textit{lower terms}) = c_{\frac{k(k+1)}{2}} \overline{z}^k$$

with $c_{\frac{k(k+1)}{2}} \in \mathbb{C}$ as long as $\alpha^{k-1}_{\frac{k(k+1)}{2}} \neq 0$. If $\alpha^{k-1}_{\frac{k(k+1)}{2}} = 0$ we make the convention that $c_{\frac{k(k+1)}{2}} = 0$. Similarly for $s = 1, \ldots, k$ we have

$$\mathrm{LT}(\beta^s_{\frac{k(k+1)}{2}+s} p_{s,k-s}) = \mathrm{LT}(z p_{s,k-1-s} + \ \textit{lower terms}) = c_{\frac{k(k+1)}{2}+s} z^s \overline{z}^{k-s}$$

with $c_{\frac{k(k+1)}{2}+s} \in \mathbb{C}$ as long as $\beta^s_{\frac{k(k+1)}{2}+s} \neq 0$. If $\beta^s_{\frac{k(k+1)}{2}+s} = 0$ we make the convention that $c_{\frac{k(k+1)}{2}+s} = 0$.

This also proves Theorem 2.7. Namely, assume $\hat{q}_0$ is supported by every spectral subspace of N. Then the first coefficient c_j that equals zero implies that the minimal polyanalytic polynomial of N has been found. As long as all the c_j's are nonzero, the vectors $\hat{q}_j$ produced so far are orthonormal. For $1, \ldots, n$ denote by

$$\mathcal{K}_k(N; \hat{q}_0) = \mathrm{span}\{\hat{q}_0, N\hat{q}_0, \ldots, N^{k-1}\hat{q}_0\} \tag{3.11}$$

the Krylov subspaces of N at $\hat{q}_0$. As N is normal, the dimension of $\mathcal{K}_n(N; \hat{q}_0)$ equals $\deg(N)$. Since each $\hat{q}_j \in \mathcal{K}_n(N; \hat{q}_0)$, there can be at most $\deg(N)$ linearly independent vectors $\hat{q}_j$. Therefore, after finishing the k^{th} cycle for which $\frac{(k+1)(k+2)}{2} > \deg(N)$, a linearly dependent vector must have occurred in the orthogonalization process (3.3). To this there corresponds an annihilating polyanalytic polynomial and, in particular, $k \leq \sqrt{2\deg(N)}$. □

Note that we allow $c_{k(k+1)/2+s} = 0$. Obviously, for $c_{k(k+1)/2+s} \neq 0$, the polyanalytic polynomial $P_{s,k-s}(z) = \frac{1}{c_{k(k+1)/2+s}} p_{s,k-s}(z)$ realizes (3.2).

If there exists an algebraic curve of degree k, with $\frac{(k+1)(k+2)}{2} < \deg(N)$, containing the spectrum of N, then this can be detected with Algorithm 1 as follows. By a generic $\hat{q}_0 \in \mathbb{C}^n$ for $N \in \mathcal{N}$ we mean a vector that is supported by every spectral subspace of N.

Corollary 3.2. *Let $\hat{q}_0 \in \mathbb{C}^n$ be generic for $N \in \mathcal{N}$ with the minimal polyanalytic polynomial $p_{j,l}$. Then, for $k \geq j + l$, we have $\hat{q}_{\frac{k(k+1)}{2}+s} = 0$ for $j \leq s \leq k - l$.*

Example 3.3. Assume N is unitary with the spectrum having sufficiently many separate points. Then the minimal polyanalytic polynomial of N is $p_{1,1}(z) = z\overline{z} - 1$. Now, corresponding to $k = j + l = 2$, we have $\frac{2(2+1)}{2} + s = 4$ with $s = 1$. Thus, $\hat{q}_4 = 0$ is the first zero vector of the process (3.3). Then for $k = 3$ we have $\hat{q}_7 = \hat{q}_8 = 0$ corresponding to $\frac{3(3+1)}{2} + s$ with $1 \leq s \leq 3 - 1$.

Remark 1. If the spectrum of $N \in \mathcal{N}$ belongs to an algebraic curve of low degree, then this curve is found after a very moderate number of steps with Algorithm 1. Since this yields a spectral exclusion set, the information is exact. With the Arnoldi method, accurate spectral information is never achieved before the step number equals the cardinality of $\sigma(N)$ (unless the starting vector is exceptional).

There is a convenient way of visualizing what happens in Corollary 3.2 when a $p \in \mathcal{PP}_k \backslash \{0\}$ of low degree vanishes on $\sigma(N)$ causing a portion of the vectors $\hat{q}_j$ to

equal zero. To this end the two-dimensional table

$$\begin{array}{ccccc} 1 & \overline{z} & \overline{z}^2 & \overline{z}^3 & \cdots \\ z & z\overline{z} & z\overline{z}^2 & z\overline{z}^3 & \cdots \\ z^2 & z^2\overline{z} & z^2\overline{z}^2 & z^2\overline{z}^3 & \cdots \\ z^3 & z^3\overline{z} & z^3\overline{z}^2 & z^3\overline{z}^3 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{array} \tag{3.12}$$

of the polyanalytic monomials is useful, where monomials of the same degree lie on the same anti-diagonal. According to Theorem 3.1, the ordering (3.3) of the orthogonalizations means that, at each step, the leading terms corresponding to the polyanalytic polynomials yielding $\hat{q}_j$'s are obtained by moving downwards on the corresponding anti-diagonal.

Example 3.4. This is Example 3.3, continued. We already noticed in Example 3.3 that the vector $\hat{q}_4$ will be linearly dependent on the previous vectors $\hat{q}_0, \ldots, \hat{q}_3$ when Algorithm 1 is executed. By multiplying both sides of the equation $p_{1,1}(z) = 0$ by z and $\overline{z}$ we obtain two equations

$$z^2\overline{z} - z = 0 \text{ and } z\overline{z}^2 - \overline{z} = 0, \tag{3.13}$$

respectively. This means that in the 3^{rd} cycle $\hat{q}_7$ and $\hat{q}_8$ will be linearly dependent on the basis vectors computed so far, and Algorithm 1 yields zero vectors at the corresponding steps. To illustrate how this extends, we form a table

$$\begin{array}{ccccc} 1 & \overline{z} & \overline{z}^2 & \overline{z}^3 & \cdots \\ z & \underline{z\overline{z}} & \underline{z\overline{z}^2} & \underline{z\overline{z}^3} & \cdots \\ z^2 & \underline{z^2\overline{z}} & \underline{z^2\overline{z}^2} & \underline{z^2\overline{z}^3} & \cdots \\ z^3 & \underline{z^3\overline{z}} & \underline{z^3\overline{z}^2} & \underline{z^3\overline{z}^3} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{array} \tag{3.14}$$

by underlining the monomials that can be expressed in terms of lower degree polyanalytic polynomials on $\sigma(N)$. The leading terms of the non-vanishing polyanalytic polynomials are not underlined. In particular, the number of new nonzero vectors in every cycle remains constant whereas the number of zero vectors increases linearly.

A straightforward application of the proof of Theorem 3.1 yields the following.

Corollary 3.5. *Let $N \in \mathcal{N}$ and $\hat{q}_0 \in \mathbb{C}^n$. Then* $\operatorname{span}\{\hat{q}_0, \ldots, \hat{q}_{\frac{(k+1)(k+2)}{2}-1}\} = \operatorname{span}_{p\in\mathcal{PP}_k}\{p(N)\hat{q}_0\}$.

Proof. For $1 \le t \le k$ we have $\hat{q}_{\frac{t(t+1)}{2}} = p_{s,t-s}(N)\hat{q}_0$, with $p_{s,t-s}$ having the leading term $\operatorname{LT}(p_{s,t-s}) = c_{\frac{t(t+1)}{2}+s} z^s \overline{z}^{t-s}$ for $0 \le s \le t$. This covers all the polyanalytic monomials of degree less than or equal to k. Now, $c_{\frac{t(t+1)}{2}+s} = 0$ if and only if $\hat{q}_{\frac{t(t+1)}{2}} = p(N)\hat{q}_0$ with a $p \in \mathcal{PP}$ having a lower order leading term than $p_{s,t-s}$. □

Corollary 3.6. *Let $\hat{q}_0 \in \mathbb{C}^n$ be generic for $N \in \mathcal{N}$, and assume no element of $\mathcal{PP}_k\backslash\{0\}$ vanishes on $\sigma(N)$. Then $\hat{q}_0, \ldots, \hat{q}_{\frac{(k+1)(k+2)}{2}-1}$ are linearly independent.*

Krylov subspaces were defined in (3.11).

Proposition 3.7. *Assume $N \in \mathcal{N}$ and $\hat{q}_0 \in \mathbb{C}^n$. Then the number of nonzero vectors generated by the process (3.3) equals* $\dim(\mathcal{K}_n(N; \hat{q}_0))$.

Proof. Clearly $\mathcal{K}_n(N;\hat{q}_0) \subset \operatorname{span}_{p\in\mathcal{PP}_{n-1}}\{p(N)\hat{q}_0\}$. On the other hand, since N is normal, we have $N^* = p(N)$ for a polynomial p, so that $\operatorname{span}_{p\in\mathcal{PP}_{n-1}}\{p(N)\hat{q}_0\} \subset \mathcal{K}_n(N;\hat{q}_0)$. □

By Corollary 3.5 the vectors $\hat{q}_j$ have the property that, after $k-1$ cycles,

$$\operatorname{span}\{\hat{q}_0,\dots,\hat{q}_{\frac{k(k+1)}{2}-1}\} = \operatorname{span}_{p\in\mathcal{PP}_{k-1}}\{p(N)\hat{q}_0\}. \tag{3.15}$$

Then the k^{th} cycle is obtained by multiplying, for $\frac{(k-1)k}{2} \le s \le \frac{k(k+1)}{2}-1$, the vectors $\hat{q}_s$ by either N^* or N. However, the inner products

$$(\hat{q}_j, N\hat{q}_s) = (N^*\hat{q}_j, \hat{q}_s) \text{ and } (\hat{q}_j, N^*\hat{q}_s) = (N\hat{q}_j, \hat{q}_s) \tag{3.16}$$

are zero for $\hat{q}_j \in \operatorname{span}_{p\in\mathcal{PP}_{k-3}}\{p(N)\hat{q}_0\}$. The property that the new iterates spanning the k^{th} cycle are orthogonal against the previous span, modulo a small portion of vectors, has been employed by Elsner and Ikramov [10] to compute condensed forms for normal matrices. All in all, (3.16) means that, to finish the k^{th} cycle, explicit orthogonalization is made against $(k-1)+k+(k+1) = 3k$ vectors at most, which therefore equals the number of vectors that needs to be stored. This is an overestimate, and the precise statement resulting from the ordering of the orthogonalizations according to (3.3) is as follows.

Theorem 3.8. *Let $\hat{q}_0 \in \mathbb{C}^n$ be generic for $N \in \mathcal{N}$ and assume no element of $\mathcal{PP}_k\backslash\{0\}$ vanishes on $\sigma(N)$. Then, to finish the k^{th} cycle, the length of the recurrence for computing $\hat{q}_{\frac{k(k+1)}{2}+s}$ is $2k$, for $s = 0,\dots,k$.*

Proof. In the beginning of the k^{th} cycle there are $(k-1)+k = 2k-1$ stored vectors. Then, to generate $\hat{q}_{\frac{k(k+1)}{2}+2}$, orthogonalize $N\hat{q}_{\frac{(k-1)k}{2}+1}$ against the vectors generated at the $(k-2)^{th}$ and $(k-1)^{th}$ cycle as well as against $\hat{q}_{\frac{k(k+1)}{2}}$ and $\hat{q}_{\frac{k(k+1)}{2}+1}$, i.e., in all against $2k+1$ vectors. However, the first vector of the $(k-2)^{th}$ cycle is already orthogonal against $N\hat{q}_{\frac{(k-1)k}{2}+1}$, since

$$(N\hat{q}_{\frac{(k-1)k}{2}+1}, \hat{q}_{\frac{(k-2)(k-1)}{2}}) = (\hat{q}_{\frac{(k-1)k}{2}+1}, N^*\hat{q}_{\frac{(k-2)(k-1)}{2}}) \tag{3.17}$$

and $N^*\hat{q}_{\frac{(k-2)(k-1)}{2}}$ is a linear combination of the vectors $\hat{q}_j$, with $0 \le j \le \frac{(k-1)k}{2}$. Thus, orthogonalization needs to be made only against $2k$ vectors. Similarly, to compute $\hat{q}_{\frac{k(k+1)}{2}+3}$ the first and second vectors of the $(k-2)^{th}$ cycle are already orthogonal against $N\hat{q}_{\frac{(k-1)k}{2}+2}$, since

$$(N\hat{q}_{\frac{(k-1)k}{2}+2}, \hat{q}_{\frac{(k-2)(k-1)}{2}}) = (\hat{q}_{\frac{(k-1)k}{2}+2}, N^*\hat{q}_{\frac{(k-2)(k-1)}{2}}) \tag{3.18}$$

and

$$(N\hat{q}_{\frac{(k-1)k}{2}+2}, \hat{q}_{\frac{(k-2)(k-1)}{2}+1}) = (\hat{q}_{\frac{(k-1)k}{2}+2}, N^*\hat{q}_{\frac{(k-2)(k-1)}{2}+1}). \tag{3.19}$$

Now (3.17) is zero by the same argument as (3.18) was, and (3.19) is zero by the following arguments. By Theorem 3.1, the leading term of $\hat{q}_{\frac{(k-2)(k-1)}{2}+1}$ is $z\overline{z}^{(k-2)-1}$ multiplied by a constant. Therefore the leading term of $N^*\hat{q}_{\frac{(k-2)(k-1)}{2}+1}$ is $z\overline{z}^{(k-1)-1}$ multiplied by a constant. Consequently, $N^*\hat{q}_{\frac{(k-2)(k-1)}{2}+1}$ is a linear combination of the vectors $\hat{q}_j$, with $0 \le j \le \frac{(k-1)k}{2}+1$, and thus (3.18) is zero. This reasoning extends throughout the k^{th} cycle in the sense that when $\hat{q}_{\frac{k(k+1)}{2}+s}$ is computed with

TABLE 1. The dimension d of the subspace versus $\sqrt{8d}$, a bound to the number of the stored vectors.

The dimension of the subspace $= d$	The number of the stored vectors $\leq \sqrt{8d}$
30	16
50	20
100	29
200	40
1000	90
10000	283
100000	895

$2 \leq s \leq k$, then orthogonalizations against $\hat{q}_j$ with $0 \leq j \leq \frac{(k-2)(k-1)}{2} + s - 2$ are redundant. □

Remark 2. To generate $\hat{q}_j$ in the k^{th} cycle, we thus need to orthogonalize only against $2k$ vectors computed most recently. Note that Algorithm 1 is implemented by using this property. Denoting by $d = \frac{(k+2)(k+1)}{2}$ the number of vectors at the end of the k^{th} cycle, this means that the number of the stored vectors $2k+1$ is bounded by $\sqrt{8d}$. This is a very modest growth, as shown in Table 1.

Remark 3. If the minimal polyanalytic polynomial of $N \in \mathcal{N}$ is of degree k, then zero vectors will occur among $\hat{q}_j$'s according to Corollary 3.2. As their number increases in every cycle, these zero vectors should not be saved in practice. As a result, the need for storage and the length of the recurrence do not increase, but remain fixed at $2k$ from there on.

With the ordering (3.3) of the orthogonalization process the canonical form $\hat{Q}^* N \hat{Q}$ of Elsner and Ikramov has the nonzero structure

$$
\begin{bmatrix}
\beta_{\mathbf{0}}^{\mathbf{0}} & \beta_0^1 & & & & & & \\
\beta_1^0 & \beta_{\mathbf{1}}^{\mathbf{1}} & \beta_1^2 & \beta_1^3 & & & & \\
\beta_2^0 & \beta_2^1 & \beta_{\mathbf{2}}^{\mathbf{2}} & \beta_2^3 & \beta_2^4 & & & \\
 & \beta_3^1 & \beta_3^2 & \beta_{\mathbf{3}}^{\mathbf{3}} & \beta_3^4 & \beta_3^5 & \beta_3^6 & \\
 & \beta_4^1 & \beta_4^2 & \beta_4^3 & \beta_{\mathbf{4}}^{\mathbf{4}} & \beta_4^5 & \beta_4^6 & \\
 & & \beta_5^2 & \beta_5^3 & \beta_5^4 & \beta_{\mathbf{5}}^{\mathbf{5}} & \beta_5^6 & \\
 & & & \beta_6^3 & \beta_6^4 & \beta_6^5 & \beta_{\mathbf{6}}^{\mathbf{6}} & \\
 & & & \beta_7^3 & \beta_7^4 & \beta_7^5 & \beta_7^6 & \cdots \\
 & & & & \beta_8^4 & \beta_8^5 & \beta_8^6 & \\
 & & & & & \beta_9^5 & \beta_9^6 & \\
 & & & & & & \beta_{10}^6 & \\
 & & & & & & \beta_{11}^6 & \\
 & & & \vdots & & & &
\end{bmatrix}
\tag{3.20}
$$

under the assumptions that the starting vector is generic and no $p \in \mathcal{PP}_k \backslash \{0\}$ vanishes on $\sigma(N)$ having cardinality n. This is the table (3.3) written in matrix form for the action of N. The first column results from the first cycle, i.e., how N maps $\hat{q}_0$. Then the second and third columns result from the second cycle, i.e., how N maps $\hat{q}_1$ and $\hat{q}_2$. Then the 4^{rd}, 5^{th} and 6^{th} columns result from the third cycle,

i.e., how N maps $\hat{q}_3$, $\hat{q}_4$ and $\hat{q}_5$; and so on. Note that the k^{th} cycle gives rise to k new columns in (3.20). For clarity, the diagonal has been bold faced.

A portion of the computed inner products with Algorithm 1 are actually redundant. Namely, since we also compute at the beginning of each cycle how N^* maps the first vector of the preceding cycle, we could employ this information to recover the corresponding row in the canonical form (3.20). For the sake of clarity we do not consider a minimum complexity version of Algorithm 1.

We illustrate with an example how the minima (3.2) are realized while computing the canonical form of Elsner and Ikramov with Algorithm 1.

Example 3.9. Assume we have

$$N = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & i & 0 \\ 0 & 0 & 0 & -i \end{bmatrix} \text{ and } \hat{q}_0 = \frac{1}{2}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}. \tag{3.21}$$

Then a simple computation yields

$$\hat{q}_1 = \frac{1}{2}\begin{bmatrix} 1 \\ -1 \\ -i \\ i \end{bmatrix},\ \hat{q}_2 = \frac{1}{2}\begin{bmatrix} 1 \\ -1 \\ i \\ -i \end{bmatrix} \text{ and } \hat{q}_3 = \frac{1}{2}\begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix} \tag{3.22}$$

with $\alpha_0^0 = \alpha_0^1 = \alpha_1^1 = \alpha_2^1 = 0$ and $\alpha_1^0 = \alpha_3^1 = 1$, and

$$\hat{Q}^* N \hat{Q} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}. \tag{3.23}$$

Thus, the polyanalytic polynomials realizing (3.2) are $P_{0,0}(z) = 1$, $P_{0,1}(z) = \overline{z}$, $P_{1,0}(z) = z$, $P_{0,2}(z) = \overline{z}^2$ and $P_{1,1}(z) = z\overline{z} - 1$. Obviously $P_{1,1}$ is also the minimal polyanalytic polynomial of N.

4. Discrete orthogonal polyanalytic polynomials

Generating polynomials orthogonal with respect to a measure on a subset of $\mathbb{R}$ is a classical problem; see, e.g., [11, 12, 9, 22]. See also [7] for analytic polynomials orthogonal with respect to a measure on a discrete subset of $\mathbb{C}$. For orthogonality on an algebraic curve, see [6] and references therein. For polynomials of two variables orthogonal with respect to a measure on a plane region, see [28, 27, 26] and references therein.

The process described in the previous section yields orthogonal polyanalytic polynomials on discrete subsets of $\mathbb{C}$ with respect to the following measure. Let $N = U\Lambda U^*$ be a diagonalization of a normal $N \in \mathbb{C}^{n\times n}$ by a unitary matrix U. Denote by $q_1, \ldots, q_n$ the columns of U and by $\lambda_1, \ldots, \lambda_n$ the corresponding eigenvalues. Without loss of generality, assume that the eigenvalues of N are distinct. For a given vector $\hat{q}_0 \in \mathbb{C}^n$ of unit length we define an inner product on $\mathcal{PP}$ via

$$(p, q) = (p(N)\hat{q}_0, q(N)\hat{q}_0) = \sum_{j=1}^{n} |(\hat{q}_0, q_j)|^2 p(\lambda_j)\overline{q(\lambda_j)}, \tag{4.1}$$

for polyanalytic polynomials p and q. In particular, for $j = 1, \ldots, n$, attaching to λ_j the mass $m_j = |(\hat{q}_0, q_j)|^2$ yields a discrete measure on the spectrum of N.

With this inner product and by using the ordering of the table (3.3), a recurrence for polyanalytic polynomials is obtained as initialized in (3.4)-(3.9). Collecting the coefficients from Algorithm 1 yields

$$p_{0,0}(z) = 1, \tag{4.2}$$

$$p_{0,1}(z) = \frac{1}{\alpha_1^0}\overline{z}p_{0,0}(z) - \frac{\alpha_0^0}{\alpha_1^0}p_{0,0}(z), \tag{4.3}$$

$$p_{1,0}(z) = \frac{1}{\beta_2^0}zp_{0,0}(z) - \frac{\beta_1^0}{\beta_2^0}p_{0,1}(z) - \frac{\beta_0^0}{\beta_2^0}p_{0,0}(z), \tag{4.4}$$

$$p_{0,2}(z) = \frac{1}{\alpha_3^1}\overline{z}p_{0,1}(z) - \frac{\alpha_2^1}{\alpha_3^1}p_{1,0}(z) - \frac{\alpha_1^1}{\alpha_3^1}p_{0,1}(z) - \frac{\alpha_0^1}{\alpha_3^1}p_{0,0}(z), \tag{4.5}$$

$$p_{1,1}(z) = \frac{1}{\beta_4^1}zp_{0,1}(z) - \frac{\beta_3^1}{\beta_4^1}p_{0,2}(z) - \frac{\beta_2^1}{\beta_4^1}p_{1,0}(z) - \frac{\beta_1^1}{\beta_4^1}p_{0,1}(z) - \frac{\beta_0^1}{\beta_4^1}p_{0,0}(z), \tag{4.6}$$

and so on. If the denominator is zero, i.e., if the corresponding vector $\hat{q}_j$ generated with Algorithm 1 equals zero, then the arising polyanalytic polynomial is set identically zero. As in the previous section, the sum $k = j + l$ of the sub-indices of $p_{j,l}$ refers to the degree of the (nonzero) polyanalytic polynomials.

The length of the recurrence for computing these elements has a very modest growth, being, according to Theorem 3.8, at most $2k$ during the k^{th} cycle. Already for $k \geq 3$ the recurrence will be shorter than the number of the computed orthogonal polyanalytic polynomials.

Theorem 4.1. *Assume $N \in \mathcal{N} \subset \mathbb{C}^{n\times n}$ has n distinct eigenvalues and $\hat{q}_0 \in \mathbb{C}^n$ is generic for N. Then Algorithm 1 produces n orthonormal polyanalytic polynomials on $\sigma(N)$ with respect to the inner product (4.1).*

Proof. According to the proof of Theorem 3.1, the nonzero $p_{s,k-s}$ are orthogonal with respect to the inner product (4.1).

To see that Algorithm 1 yields n mutually orthogonal polyanalytic polynomials, consider (3.15). Then

$$\mathcal{K}_k(N;\hat{q}_0) = \text{span}\{\hat{q}_0, N\hat{q}_0, \dots, N^{k-1}\hat{q}_0\} \subset \text{span}_{p\in\mathcal{PP}_{k-1}}\{p(N)\hat{q}_0\}$$

and $\mathcal{K}_k(N;\hat{q}_0) = \mathbb{C}^n$ by the genericity assumption. Thus the process (3.3) yields a basis of $\mathbb{C}^n$, and the claim follows because to each nonzero $\hat{q}_j$ there corresponds a nonzero polyanalytic polynomial. □

This orthonormal set of polyanalytic polynomials is obtained without resorting to the polyanalytic monomial basis $\text{span}\{z^j\overline{z}^l\}_{j,l\in\mathbb{N}_0}$ at any point. This is a significant feature of the approach, as, for well-known reasons [13], starting from the monomial basis can be numerically very unstable.

Obviously, since the spectrum of N is finite, the orthonormal polyanalytic polynomials of Theorem 4.1 could be converted into orthonormal polynomials by replacing $\overline{z}$ with a polynomial in z. If we considered the analogous problem on a continuum, then this would no longer be possible.

Corollary 4.2. *Denote by $p_1, \dots, p_n$ the orthonormal polyanalytic polynomials of Theorem 4.1 and by $\mathcal{PP}_{\sigma(N)}$ those $p \in \mathcal{PP}\backslash\{0\}$ that vanish on $\sigma(N)$. Then* $\text{span}\{p_k\} \cap \mathcal{PP}_{\sigma(N)} = \emptyset$.

Proof. Assume $p \in \mathcal{PP}\backslash\{0\}$ vanishes on $\sigma(N)$. Then, if $p = \sum_{j=1}^{k} \alpha_j p_j$, for $\alpha_j \in \mathbb{C}$ with $1 \leq j \leq k \leq n$, we would have $\| \sum_{j=1}^{k} \alpha_j p_j(N)\hat{q}_0\| = 0$. Due to the linear independence of p_j's, this forces $\alpha_j = 0$ for $1 \leq j \leq k$, and the claim follows. □

Corollary 4.3. *Under the assumptions of Theorem 4.1, if N is Hermitian, then Algorithm 1 produces n Lanczos polynomials.*

Proof. According to Corollary 3.2, only the vectors $\hat{q}_0$, $\hat{q}_1$, $\hat{q}_3$, $\hat{q}_6$, $\hat{q}_{10}$ ... are nonzero. These correspond to the Lanczos polynomials, and the claim follows. □

Remark 4. Under the assumptions of Corollary 4.3, the highest degree polyanalytic polynomial among the computed orthonormal (polyanalytic) polynomials is of degree $n-1$. Since the dimension of $\mathcal{PP}_{n-1}$ is $\frac{n(n+1)}{2}$, the obtained set of orthonormal functions clearly spans only a small subspace of $\mathcal{PP}_{n-1}$. This is always the case if the minimal polyanalytic polynomial of N is of low degree. Let us illustrate this with another example.

Example 4.4. Let $N = \mathrm{diag}(1, -1, i, -i, e^{i\frac{\pi}{4}}, e^{-i\frac{\pi}{4}}, e^{i\frac{3\pi}{4}}, e^{-i\frac{3\pi}{4}}) \in \mathbb{C}^{8\times 8}$ and $\hat{q}_0 = \frac{1}{2\sqrt{2}}(1, \ldots, 1)$, i.e., the starting vector has equal weights at the eigenvalues. Since N is unitary, the unit circle is an algebraic curve of degree 2 containing the spectrum of N. For this N and $\hat{q}_0$ Algorithm 1 produces zero vectors $\hat{q}_4 = \hat{q}_7 = \hat{q}_8 = 0$, so that $\mathbb{C}^8 = \{\hat{q}_0, \hat{q}_1, \hat{q}_2, \hat{q}_3, \hat{q}_5, \hat{q}_6, \hat{q}_9, \hat{q}_{10}\}$. The corresponding orthogonal polyanalytic polynomials are $p_{0,0}(z) = 1$, $p_{0,1}(z) = \overline{z}$, $p_{1,0}(z) = z$, $p_{0,2}(z) = \overline{z}^2$, $p_{2,0}(z) = z^2$, $p_{0,3}(z) = \overline{z}^3$, $p_{3,0}(z) = z^3$ and $p_{0,4}(z) = \overline{z}^4$, which span only a small portion of $\mathcal{PP}_4$.

Orthogonal polynomials satisfying a 3-term recurrence with respect to a measure on $\mathbb{R}$ give rise to Jacobi matrices; see, e.g., [11]. In the same way a discrete measure on $\mathbb{C}$ gives rise to a unique matrix with the sparsity pattern (3.20).

5. Least squares approximation and bivariate Lagrange interpolation

Vandermonde-like systems arise when polynomials (of one complex variable) are used to approximate analytic functions on subsets of the complex plane; see, e.g., [14, 22]. However, in numerous problems the function $f : \mathbb{C} \to \mathbb{C}$, with $f(x, y) = f_1(x, y) + if_2(x, y)$, being approximated is not analytic. Real problems belong to this category, that is, when $g : \mathbb{R}^2 \to \mathbb{R}$ is identified with a function $f : \mathbb{C} \to \mathbb{C}$ by setting $f(x, y) = g(x, y) + i0$. For these types of problems, using (analytic) polynomials is obviously not the most natural choice. On the other hand, it is well-known that bivariate least squares approximation and interpolation cannot be handled with a straightforward extension of the univariate techniques; see, e.g., [20, 3, 4, 25].

Denote by $\mathcal{P}_k(\mathbb{R}^2)$ the set of bivariate polynomials of degree k at most. Regarding the uniqueness of multivariate Lagrange interpolation, there arise the following problems. First, $\mathcal{P}_k(\mathbb{R}^2)$ is $\frac{(k+2)(k+1)}{2}$ dimensional, so that exactly $\frac{(k+2)(k+1)}{2}$ interpolation nodes seem to be needed. However, even if this was the case, the uniqueness does not follow. Namely, if the interpolation nodes belong to an algebraic curve of degree less than k, then the solution is not unique, as the equation defining the algebraic curve can be added to an interpolant without affecting the interpolation. To avoid these problems, different ways of constructing polynomial

subspaces (depending on the number and the location of the nodes) have been suggested; see [3, 25] and references therein. Some of these construction are quite elaborate and somewhat difficult to follow.

By using orthogonal polyanalytic polynomials we can compute least squares approximations as well as Lagrange interpolants, due to the fact that Algorithm 1 generates a set of orthogonal polyanalytic polynomials regardless of how the interpolation nodes are located. If the nodes do lie on a low dimensional algebraic curve, then the corresponding vanishing elements of $\mathcal{PP}$ are discarded by the process. Also, through attaining zero with (3.3), the algebraic curve the nodes belong to can be readily detected. In particular, since in this manner the original problem is completely converted into a problem of numerical linear algebra, a large variety of numerically stable techniques become available.

To this end, let $\mathcal{X}_n$ be a given ordered set of interpolation nodes of $\mathbb{R}^2$ of cardinality n and let $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{R}^n$ (or $\mathbb{C}^n$). We identify the elements of $\mathcal{X}_n$ with points of $\mathbb{C}$ in the standard manner, and set $N \in \mathbb{C}^{n\times n}$ to be a diagonal matrix having the points of $\mathcal{X}_n$ on the diagonal. Let $\hat{q}_0 \in \mathbb{C}^n$ be a "weight" vector of unit length with nonzero elements. With this notation, $p \in \mathcal{PP}$ is called an interpolating polyanalytic polynomial for the pair $(\mathcal{X}_n, \alpha)$ if it solves the corresponding interpolation problem.

Theorem 5.1. *Assume $\mathcal{X}_n \subset \mathbb{C}$, $\alpha \in \mathbb{R}^n$, and let $N \in \mathbb{C}^{n\times n}$ be a diagonal matrix corresponding to $\mathcal{X}_n$. If $\hat{q}_0 \in \mathbb{C}^n$ is of unit length with nonzero elements and $p_1, \ldots, p_n$ are the orthogonal polyanalytic polynomials of Theorem 4.1, then there exists $(c_1, \ldots, c_n) \in \mathbb{C}^n$ such that $\sum_{j=1}^n c_j p_j$ is an interpolating polyanalytic polynomial for the pair $(\mathcal{X}_n, \alpha)$.*

Proof. This follows directly from the fact that $p_1, \ldots, p_n$ correspond to an orthonormal basis of $\mathbb{C}^n$ with respect to the inner product (4.1). Thus, representing α in this basis yields $(c_1, \ldots, c_n) \in \mathbb{C}^n$. □

With this solution we can readily find an interpolating bivariate polynomial.

Corollary 5.2. *The real part of $\sum_{j=1}^n c_j p_j$ solves the corresponding real interpolation problem.*

Proof. Since $\alpha \in \mathbb{R}^n$, the interpolant $\sum_{j=1}^n c_j p_j$ attains real values at the nodes $\mathcal{X}_n$. Consequently, removing its imaginary part does not change the interpolation. □

We denote by $\mathcal{PP}(\mathcal{X}_n)$ the subspace of $\mathcal{PP}$ obtained from the span of the orthogonal polyanalytic polynomials of Theorem 4.1. Let us now consider an example.

Example 5.3. We want to find a bivariate polynomial attaining at the nodes

$$\mathcal{X}_8 = \{(x_j, y_j) \in \mathbb{R}^2 : (x_j, y_j) = (\cos(2\pi j/8), \sin(2\pi j/8)), \text{ for } j = 0, \ldots, 7\} \tag{5.1}$$

the values $(-1)^j$. If we choose to use equal weights at every node (x_j, y_j), then we have computed all the necessary quantities in Example 4.4. The corresponding "Vandermonde-like" linear system is

$$\begin{bmatrix} p_{0,0}(x_0, y_0) & \cdots & p_{0,4}(x_0, y_0) \\ \vdots & \vdots & \vdots \\ p_{0,0}(x_7, y_7) & \cdots & p_{0,4}(x_7, y_7) \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_8 \end{bmatrix} = \begin{bmatrix} 1 \\ \vdots \\ -1 \end{bmatrix}, \tag{5.2}$$

giving $c = (0, \dots, 0, 1)$. Thus, the interpolating polyanalytic polynomial equals $p_{0,4}(z) = \overline{z}^4 = x^4 - 6x^2y^2 + y^4 - 4xy(x^2 - y^2)i \in \mathcal{PP}(\mathcal{X}_8)$. Taking its real part yields an interpolating bivariate $p(x, y) = x^4 - 6x^2y^2 + y^4 \in \mathcal{P}_4(\mathbb{R}^2)$.

Remark 5. The computation of the interpolant is, of course, not realized by writing the linear system as was done in (5.2). This would mean saving all the computed orthogonal polyanalytic polynomials, which is not necessary because of the slowly growing length of the recurrence. Instead, the interpolant is found by computing the Fourier coefficients of $(\alpha_1, \dots, \alpha_n)$ with respect to the generated basis while the iteration proceeds.

Remark 6. In practice the interpolation nodes often do lie on an algebraic curve of low degree. To give a trivial example, $\mathbb{R}$ is an algebraic curve of degree 1 in $\mathbb{R}^2$, and then, by Corollary 4.3, the proposed interpolation scheme reduces to the Hermitian Lanczos method. More interesting examples arise in FEM problems in 2 dimensions or when solving problems in 3-dimensional domains by the boundary integral equation method. Then the nodes can be located, for example, on the edges of triangles. The edges of a single triangle belong to a union of 3 lines, which is an algebraic curve of degree 3. Thus, for a single triangle the length of the recurrence for computing the interpolant is 6, regardless of the number of interpolation nodes on the edges.

Remark 7. This process yields least squares approximants (with respect to the norm (4.1)) during the iteration. This is of interest if the number of interpolation nodes is very large and if sufficient accuracy is attained before obtaining the exact interpolant. Bivariate approximants are obtained by taking the real part of the computed best approximation with respect to the norm (4.1). Obviously this is always an error-decreasing operation. It is not linear, as we are mapping from a complex subspace to the set of bivariate polynomials.

Example 5.4. We illustrate how taking the real part decreases the error. Assume we are interpolating the function $f(x, y) = 2x$ at the nodes given in (5.1). Let the weights again be equal. We know that the solution is the real part yielded by the Fourier coefficients $c = (0, 1, 1, 0, 0, 0, 0, 0)$. The approximations are $p^0(z) = 0$, $p^1(z) = \overline{z}$ and $p^2(z) = \overline{z} + z$, so that the real parts of these are $\hat{p}^0(z) = 0$, $\hat{p}^1(z) = x$ and $\hat{p}^2(z) = 2x$.

In addition to the orthogonality of the computed functions and the slowly growing length of the recurrence, the proposed scheme has all the further properties that the Lagrange interpolation is desired to have; see [3, 25]. We state these for $\mathcal{PP}(\mathcal{X}_n)$.

Proposition 5.5. *The Lagrange interpolation problem is uniquely solvable with respect to $\mathcal{X}_n$ in $\mathcal{PP}(\mathcal{X}_n)$.*

For the definition of a minimal degree interpolation space, see [3, 4].

Proposition 5.6. *The subspace $\mathcal{PP}(\mathcal{X}_n)$ is of minimal degree with respect to $\mathcal{X}_n$.*

Since in practice least squares approximation and interpolation are realized in finite precision, two issues need to be addressed. First, if the number of nodes $\mathcal{X}_n$ is very large, then saving a small, although increasing, portion of vectors will likely result in a loss of orthogonality. However, since we have transformed our

problem into a problem of linear algebra, this can be remedied, at least partially, by employing either full or selective orthogonalization methods with Algorithm 1. These tools are now standard in numerical linear algebra and therefore quite well understood; see, e.g., [8] and references therein. Of course, even if this can improve approximations dramatically, the work and storage will increase accordingly.

An equally problematic issue in finite precision multivariate interpolation arises when there is a need to decide whether the nodes lie on a low dimensional algebraic curve or not. In our approach this can be done by monitoring the size of (3.2). More precisely, if (3.2) is tiny, then one needs to judge whether to set the corresponding vector identically zero and, thus, discard the corresponding polyanalytic polynomial from the basis.

As a final comment, in practical implementation there arises the question whether it is always reasonable to strive for minimal degree interpolation. If the nodes $\mathcal{X}_n$ lie almost on a low degree algebraic curve, then large coefficients will be introduced into the interpolation process. For obvious reasons this is not desirable. One option is to discard the associated vector and the corresponding polyanalytic polynomial and continue to expand the basis from the step that follows. A drawback of this is that then the inner products in (3.16) will be nonzero and, as a result, a linear growth of the length of the recurrence and storage must be accepted.

6. Conclusions

In this paper we have considered an iterative method for normal matrices that partially preserves the properties of the Hermitian Lanczos method for generating discrete orthogonal polynomials. This is based on employing both N and N^* in a particular order during the iteration. With the algorithm a sequence of orthogonal polyanalytic polynomials can be computed such that the length of the recurrence to this end is bounded by $\sqrt{8d}$, where d denotes the number of elements computed so far. The scheme extends the standard Hermitian Lanczos method, as for a Hermitian matrix N the method yields orthogonal polynomials satisfying a 3-term recurrence. We have applied the algorithm to least squares approximation and bivariate interpolation.

Acknowledgment

I am very grateful to Professor Carl de Boor for his very useful comments on an earlier version of this paper.

References

1. W.E. Arnoldi, *The principle of minimized iterations in the solution of the matrix eigenvalue problem,* Quart. Apll. Math., **9**, (1951), 17-29. MR **13:**163e
2. M.B. Balk, *Polyanalytic functions,* Wiley/VCH, Weinh., 1991. MR **93k:**30076
3. C. de Boor, *Polynomial interpolation in several variables,* Studies in Computer Science, R. DeMillo and J. R. Rice (eds.), Plenum Press, New York, (1994), pp. 87-119.
4. C. de Boor and A. Ron, *Computational aspects of polynomial interpolation in several variables,* Math. Comp., **58**, (1992), 705-727. MR **92i:**65022
5. P. Borwein, *The arc length of the lemniscate* $\{|p(z)| = 1\}$, Proc. Amer. Math. Soc., **123**, (1995), 797-799. MR **95d:**31001
6. C. Brezinski, *Formal orthogonality on an algebraic curve,* Annals of Numer. Math., **2**, (1995), 21-33. MR **97a:**42017

7. A. Bultheel and M. Van Barel, *Vector orthogonal polynomials and least squares approximation,* SIAM J. Matrix Anal. Appl., **16**, (1995), 863-885. MR **96h:**65060
8. J. Demmel, *Applied Numerical Linear Algebra,* SIAM, Philadelphia, 1997. MR **98m:**65001
9. S. Elhay, G. Golub and J. Kautsky, *Updating and downdating of orthogonal polynomials with data fitting applications,* SIAM J. Matrix Anal. Appl., **12**, 1991, 327-353. MR **91m:**65054
10. L. Elsner and KH.D. Ikramov, *On a condensed form for normal matrices under finite sequence of elementary similarities,* Lin. Alg. Appl., **254**, (1997), 79-98. MR **98b:**15011
11. W. Gautschi, *On generating orthogonal polynomials,* SIAM J. Sci. Comp., **3** (1982), 289-317. MR **84e:**65022
12. W. Gautschi, *Orthogonal Polynomials: Applications and Computations,* Acta Numerica 1996, Cambridge University Press, (1996), 45-119. MR **99i:**65019
13. W. Gautschi and G. Inglese, *Lower bounds for the condition number of Vandermonde matrices,* Numer. Math., **52** (1988), 241-250. MR **89b:**65108
14. G. H. Golub and C. F. Van Loan, *Matrix Computations,* The John Hopkins University Press, Baltimore and London, the 3^{rd} ed., 1996. MR **97g:**65006
15. R.A. Horn and C.R. Johnson, *Topics in Matrix Analysis,* Cambridge Univ. Press, 1991. MR **92e:**15003
16. A.S. Householder, *Lectures on Numerical Algebra,* Mathematical Association of America, Buffalo, N.Y., 1972. MR **53:**11952
17. M. Huhtanen, *A stratification of the set of normal matrices,* SIAM J. Matrix Anal. Appl., to appear.
18. M. Huhtanen, *A Hermitian Lanczos method for normal matrices,* SIAM J. Matrix Anal. Appl., to appear.
19. M. Huhtanen and R. M. Larsen, *Exclusion and inclusion regions for the eigenvalues of a normal matrix*, Stanford University, SCCM-01-02 report, 2001.
20. G. G. Lorentz and R. A. Lorentz, *Bivariate Hermite interpolation and applications to algebraic geometry,* Numer. Math. **57**, (1990), 669-680. MR **92f:**41004
21. O. Nevanlinna, *Convergence of Iterations for Linear Equations,* Lectures in Mathematics ETH Zürich, Birkäuser Verlag, Basel, 1993. MR **94h:**65055
22. L. Reichel, *Fast QR decomposition of Vandermonde-like matrices and polynomial least squares approximation*, SIAM J. Matrix Anal. Appl., **12**, (1991), 552-564. MR **92a:**65126
23. Y. Saad, *Numerical Methods for Large Eigenvalue Problems,* Halstead Press, NY, 1992. MR **93h:**65052
24. L. A. Santalo, *Integral geometry and geometric probability,* Addison-Wesley, Reading, MA, 1976. MR **55:**6340
25. T. Sauer, *Polynomial interpolation of minimal degree*, Numer. Math., **78**, (1997), 59-85. MR **99f:**41042
26. P. K. Suetin, *Orthogonal polynomials in two variables*, Anal. Meth. and Spec. Funct., **3**, Gordon and Breach Sci. Publ., Amsterdam, 1999. MR **2000h:**42020
27. I. G. Sprinkhuizen-Kuyper, *Orthogonal polynomials in two variables. A further analysis of the polynomials orthogonal over a region bounded by two lines and a parabola,* SIAM J. Math. Anal., **7**, (1976), 501-518. MR **54:**3278
28. M.J. Zygmunt, *Recurrence formula for polynomials of two variables, orthogonal with respect to a rotation invariant measure,* Constr. Approx., **15** (1999), 301-309. MR **2000b:**33008

SCCM program, Computer Science Department, Stanford University, Stanford, California 94305

Current address: Department of Mathematics, MIT, 77 Massachusetts Avenue, Cambridge, Massachusetts 01239

E-mail address: Marko.Huhtanen@hut.fi