

GAUSS QUADRATURE FOR QUASI-DEFINITE LINEAR FUNCTIONALS

STEFANO POZZA*, MIROSLAV S. PRANIĆ†, AND ZDENĚK STRAKOŠ‡

Abstract. The Gauss quadrature can be formulated as a method for approximating positive definite linear functionals. Its mathematical context is extremely rich, with orthogonal polynomials, continued fractions and Padé approximation on one (functional analytic or approximation theory) side, and with the method of moments, (real) Jacobi matrices, spectral decompositions, and the Lanczos method on the other (algebraic) side. The quadrature concept can therefore be developed using many different ways. After a brief review of the mathematical interconnections in the positive definite case, this paper will investigate the question of a meaningful generalization of the Gauss quadrature for approximation of linear functionals which are not positive definite. For that purpose we use the algebraic approach, and, in order to build up the main ideas, recall the existing results presented in literature. Along the way we refer to the associated results expressed through the language of rational approximations. As the main result we present the form of the generalized Gauss quadrature and prove that the quasi-definiteness of the underlying linear functional represents the necessary and sufficient condition for its existence.

Key words. Quasi-definite linear functionals, Gauss quadrature, orthogonal polynomials, Hankel determinants, complex Jacobi matrices, matching moments

1. Introduction. Let \mathcal{L} be a *linear* functional on the space \mathcal{P} of polynomials with generally complex coefficients, $\mathcal{L} : \mathcal{P} \rightarrow \mathbb{C}$. The functional \mathcal{L} is fully determined by its values on monomials, called moments,

$$\mathcal{L}(x^\ell) = m_\ell, \quad \ell = 0, 1, \dots \quad (1.1)$$

Development of Gauss quadrature is linked with the concept of *orthogonal polynomials*.

DEFINITION 1.1. *A sequence of polynomials π_0, π_1, \dots satisfying the conditions*

1. $\deg(\pi_j) = j$ (π_j is of degree j , $\pi_0 \neq 0$),
2. $\mathcal{L}(\pi_i \pi_j) = 0$, $i < j$,
3. $\mathcal{L}(\pi_j^2) \neq 0$,

is called a sequence of orthogonal polynomials with respect to the linear functional \mathcal{L} .

The conditions 2. - 3. are equivalent to

$$\mathcal{L}(p\pi_j) = 0, \quad \forall p \in \mathcal{P}_{j-1}, \quad \text{and} \quad \mathcal{L}(p\pi_j) \neq 0, \quad \text{if } \deg(p) = j,$$

where $\mathcal{P}_{j-1} \subset \mathcal{P}$ is the subspace of polynomials of degree at most $j-1$. Providing that they exist, $\pi_j(x)$, $j = 1, 2, \dots$, are uniquely determined up to nonzero multiplicative constants.

In the classical case (see [21], [40], [11], [12] and [63]) the functional \mathcal{L} is positive definite, i.e., its value is positive for every not identically zero polynomial which is

*Department of Mathematics, University of Bologna, Piazza di Porta San Donato 5, 40126 Bologna, Italy. E-mail: stefano.pozza@unibo.it. Research supported in part by INdAM, GNCS (Gruppo Nazionale per il Calcolo Scientifico).

†Department of Mathematics and Informatics, University of Banja Luka, Faculty of Science, M. Stojanovića 2, 51000 Banja Luka, Bosnia and Herzegovina. E-mail: pranic77m@yahoo.com. Research supported in part by the Ministry of Science and Technology of R. Srpska and by the Serbian Ministry of Education and Science.

‡Faculty of Mathematics and Physics, Charles University in Prague, Sokolovská 83, 186 75 Praha 8, Czech Republic E-mail: strakos@karlin.mff.cuni.cz. Research supported by the ERC-CZ project LL1202.

non-negative for all real arguments; see, e.g., [10, Definition 3.1]. Then there exists a non-decreasing positive distribution function μ defined on the real axis having finite limits at $\pm\infty$ such that we can define the value \mathcal{L} for any continuous real function f as the Riemann-Stieltjes integral

$$\mathcal{L}(f) = \int f(x) d\mu(x);$$

see [10, Chapter II, Section 3]. We can consider for every integer n its unique approximation by the n -node interpolatory quadrature that has the maximal algebraic degree of exactness $2n - 1$. This gives the well-known *Gauss quadrature rule*.

The classical theory of Gauss quadrature can be found in many books; see, for example, [66, Chapters III and XV], [10, Chapter I, Section 6], [23], [22, Chapter 3.2], [45, Section 3.2]. It can be described via orthogonal polynomials or Jacobi matrices which store the coefficients of the associated three-term recurrences for *orthonormal* polynomials. The Jacobi matrices are therefore typically defined as real, tridiagonal, symmetric matrices with positive sub-diagonals (in general, depending on the orthogonal polynomials normalization, the tridiagonal matrices storing the recurrence coefficients are quasi-symmetric; see [70, pp. 335-336]). The n -th Jacobi matrix J_n is thus determined by the first $2n$ moments of the distribution function μ . Moreover, $m_0 \mathbf{e}_1^T (J_n)^i \mathbf{e}_1$ is equal to the moment m_i for $i = 0, \dots, 2n - 1$, which is known as *the moment matching property*.

In this paper we consider the generalization of the Gauss quadrature satisfying the following basic properties:

- G1: The n -node Gauss quadrature attains the maximal algebraic degree of exactness $2n - 1$, i.e., it is exact for all polynomials of degree at most $2n - 1$.
- G2: The n -node Gauss quadrature is well-defined and it is unique. Moreover, the Gauss quadratures with a smaller number of nodes also exist and they are unique.
- G3: The Gauss quadrature of a function f can be written in the form $m_0 \mathbf{e}_1^T f(J_n) \mathbf{e}_1$, where J_n is the Jacobi matrix containing the coefficients from the three-term recurrence relation for *orthonormal* polynomials associated with \mathcal{L} ; $m_0 = \mathcal{L}(x^0)$.

They are naturally valid for the Gauss quadrature approximations of positive definite functionals (for the definition of functions of matrices refer, e.g., [36]). The question arises how far we can go with generalization of the Gauss quadrature satisfying the conditions G1–G3 as an approximation for an arbitrary linear functional. The presented paper answers this. It turns out that the proposed generalization of the Gauss quadrature is equivalent to that presented in a bit different setting of functionals defined by integrals based on a complex measure by Milovanović et al. [51]. That paper focuses primarily on questions of convergence in the operator setting, it points out also some earlier developments. The focus of our paper is different.

Section 2 briefly reviews the positive definite case and points out various important interconnections. In Section 3 we recall properties of (formal) orthogonal polynomials. Here and elsewhere in the paper we recall many results present in literature, with the classical monograph by Szegő [66] and the beautiful summary of the general theory by Chihara [10] as basic references. As pointed out in [45, Section 3.4], Jacobi matrices represent a cornerstone binding the approximation theory and functional analytic approach with the algebraic approach using matrices, spectral

theory, matrix moments, and the Lanczos method. We will base our further exposition on complex Jacobi matrices and describe their spectral properties in Section 4. Section 5 is devoted to the moment matching property for complex Jacobi matrices and quasi-definite linear functionals. Known generalizations of the Gauss quadrature using restrictive assumptions are summarized in Section 6. Section 7 presents the n -weight generalization of the Gauss quadrature satisfying the conditions G1-G3. It is proved that quasi-definiteness of the underlying linear functional (see Definition 3.1 below) represents the necessary and sufficient condition for its existence. If the linear functional is not quasi-definite, then an analogous generalization cannot be established. Demonstration of this fact concludes the paper.

2. Gauss quadrature for positive definite linear functionals. In the Introduction we have considered, as in [10], infinite sequences of orthogonal polynomials. Throughout the paper we describe linear functionals on finite dimensional spaces of polynomials and characterize them by finite sequences of Hankel determinants. Therefore we will work with finite sequences of orthogonal polynomials, distribution functions with finite numbers of the points of increase etc. Such approach is convenient, e.g., for linear functionals associated with finite dimensional problems such as in Krylov subspace methods; see [45]. For extension to infinite dimensional problems we refer, e.g., to [10, Chapter II, Section 3, in particular Theorems 2.2 and 3.1], [47, Chapter 5].

Since the linear functionals on the space of polynomials are fully determined by their moments, it is convenient to consider the Hankel matrices of moments M_j and their determinants (called also Hankel determinants) Δ_j (see (1.1)),

$$M_j = \begin{bmatrix} m_0 & m_1 & \dots & m_j \\ m_1 & m_2 & \dots & m_{j+1} \\ \vdots & \vdots & \ddots & \vdots \\ m_j & m_{j+1} & \dots & m_{2j} \end{bmatrix}, \quad \Delta_j = \begin{vmatrix} m_0 & m_1 & \dots & m_j \\ m_1 & m_2 & \dots & m_{j+1} \\ \vdots & \vdots & \ddots & \vdots \\ m_j & m_{j+1} & \dots & m_{2j} \end{vmatrix}. \quad (2.1)$$

They were used in the related context by many authors; see, e.g., the seminal paper by Stieltjes that established the analytic theory of continued fractions published in 1894 [63]; see also, e.g., [42] and [10, Chapter I].

Positive definite linear functionals recalled in the Introduction can with no loss of generality be restricted to polynomials of a real variable and equivalently defined in the following way (see, e.g., [10, Chapter I, Theorem 3.4]):

DEFINITION 2.1. *The linear functional \mathcal{L} is positive definite on the subspace of polynomials of a real variable of degree most k if $m_s \in \mathbb{R}$ for $s = 0, \dots, 2k$ and $\Delta_j > 0$ for $j = 0, \dots, k$, i.e., if its first $2k + 1$ moments are real and all left principal submatrices of M_k are positive definite.*

Then we can define the following inner product on the subspace of polynomials of a real variable of degree at most k

$$(p, q) = \mathcal{L}(p(x)\overline{q(x)}) \quad (2.2)$$

(see [10, Remark on pp. 16-17, Chapter 1]) and write the symmetric positive definite Hankel matrix M_k as the Gram matrix of the monomials $1, x, \dots, x^k$,

$$M_k = \begin{bmatrix} (1, 1) & (1, x) & \dots & (1, x^k) \\ (x, 1) & (x, x) & \dots & (x, x^k) \\ \vdots & \vdots & \ddots & \vdots \\ (x^k, 1) & (x^k, x) & \dots & (x^k, x^k) \end{bmatrix}. \quad (2.3)$$

Considering the Cholesky factorization (see, e.g., [27, Theorem 4.2.5])

$$M_k = L_k L_k^T, \quad L_k^{-1} M_k L_k^{-T} = I_{k+1},$$

where I_{k+1} denotes the identity matrix, it can be seen that the entries of the rows 1 to $k+1$ of the lower tridiagonal matrix L_k^{-1} are equal to the coefficients of the orthonormal polynomials $\tilde{p}_j, j = 0, 1, \dots, k$ associated with the inner product (2.2). The roots of the polynomial \tilde{p}_k determined by the $(k+1)$ th row of the matrix L_k^{-1} give the real and distinct nodes $\lambda_1, \dots, \lambda_k$ of the associated Gauss quadrature

$$\mathcal{L}(f) = \sum_{j=1}^k \omega_j f(\lambda_j) + R_k(f), \quad (2.4)$$

where the last term $R_k(f)$ stands for the quadrature error. The weights $\omega_1, \dots, \omega_k$ are determined in the standard interpolatory quadrature way; see [45, Section 3.2, relation (3.2.5)]. The nodes and weights determine the associated distribution function

$$\mu_k(\lambda) = \begin{cases} 0 & \text{if } \lambda < \lambda_1 \\ \sum_{j=1}^n \omega_j & \text{if } \lambda_n \leq \lambda < \lambda_{n+1}, \quad n = 1, \dots, k-1 \\ \sum_{j=1}^k \omega_j & \text{if } \lambda_k \leq \lambda \end{cases} ;$$

see [63], [45, Section 3.1]. Clearly, the orthonormal polynomials are all real and, with no loss of generality, the linear functional \mathcal{L} can be restricted to the space of real polynomials of a real variable. Using the Gauss quadrature, the inner product (2.2) can then be written as the Riemann-Stieltjes integral

$$(p, q) = \mathcal{L}(p(x)q(x)) = \sum_{j=1}^k \omega_j p(\lambda_j)q(\lambda_j) = \int p(x)q(x) d\mu_k(x). \quad (2.5)$$

Since the nodes of the quadrature, i.e., the points of increase of the distribution μ_k , are the roots of the polynomial \tilde{p}_k ,

$$\int \tilde{p}_k d\mu_k(x) = \int (\tilde{p}_k)^2 d\mu_k(x) = 0.$$

Therefore the integral representation (2.5) holds for all polynomials p and q of a real variable satisfying $p(x)q(x) \in \mathcal{P}_{2k-1}$ but it does not hold whenever the product polynomial $p(x)q(x)$ is of degree $2k$.

The comprehensive description can be found in [45, Section 3.6] with references to many results presented in literature including the works of Gordon [28], Golub and Welsch [26], Wilf [69] and Mysovskih [52].

In the rest of this section we will very briefly recall some closely related mathematical developments. Details, historical comments and numerous references to original works can be found in [45, Chapter 5].

The *monic* orthogonal counterparts π_j of $\tilde{p}_j, j = 0, 1, \dots, k$ (from now on π is always used for monic orthogonal polynomials) satisfy the three-term recurrence relation ([66, Theorem 3.2.1] and [10, Theorem 4.1])

$$\pi_j(x) = (x - \delta_{j-1})\pi_{j-1}(x) - \eta_{j-1}\pi_{j-2}(x), \quad j = 1, 2, \dots, k, \quad (2.6)$$

where we set $\eta_0 = m_0$, $\delta_0 = m_1/m_0$, $\pi_{-1}(x) = 0$, $\pi_0(x) = 1$, with the other coefficients defined as

$$\delta_{j-1} = \frac{\mathcal{L}(x\pi_{j-1}^2)}{\mathcal{L}(\pi_{j-1}^2)}, \quad \eta_{j-1} = \frac{\mathcal{L}(\pi_{j-1}^2)}{\mathcal{L}(\pi_{j-2}^2)} > 0, \quad j = 2, \dots, k. \quad (2.7)$$

The orthonormal polynomials associated with (2.6) are unique up to multiplication by (-1) ; typically the leading coefficient is fixed to be positive. The particular polynomials \tilde{p}_j determined by the Cholesky factorization of the matrix of moments can be expressed as

$$\tilde{p}_j(x) = \frac{\pi_j(x)}{\sqrt{\mathcal{L}(\pi_j^2)}} = \frac{\pi_j(x)}{\sqrt{\eta_0\eta_1 \dots \eta_j}}, \quad j = 0, 1, \dots, k. \quad (2.8)$$

The three-term recurrence relation for the orthonormal polynomials $\tilde{p}_0, \dots, \tilde{p}_n$, $n \leq k$, can be written in the matrix form

$$x \begin{bmatrix} \tilde{p}_0(x) \\ \tilde{p}_1(x) \\ \vdots \\ \tilde{p}_{n-1}(x) \end{bmatrix} = J_n \begin{bmatrix} \tilde{p}_0(x) \\ \tilde{p}_1(x) \\ \vdots \\ \tilde{p}_{n-1}(x) \end{bmatrix} + \sqrt{\eta_n} \begin{bmatrix} 0 \\ 0 \\ \vdots \\ \tilde{p}_n(x) \end{bmatrix}, \quad (2.9)$$

where J_n is the (real) tridiagonal symmetric matrix with positive subdiagonal entries, called the Jacobi matrix

$$J_n = \begin{bmatrix} \delta_0 & \sqrt{\eta_1} & & & \\ \sqrt{\eta_1} & \delta_1 & \sqrt{\eta_2} & & \\ & \sqrt{\eta_2} & \delta_2 & \ddots & \\ & & & \ddots & \sqrt{\eta_{n-1}} \\ & & & & \sqrt{\eta_{n-1}} & \delta_{n-1} \end{bmatrix}. \quad (2.10)$$

From (2.9) we see that the zeros λ_j , $j = 1, \dots, n$, of \tilde{p}_n are the eigenvalues of J_n , with

$$\mathbf{w}_j = [\tilde{p}_0(\lambda_j), \tilde{p}_1(\lambda_j), \dots, \tilde{p}_{n-1}(\lambda_j)]^T, \quad j = 1, \dots, n, \quad (2.11)$$

the associated non-normalized eigenvectors. Moreover, we see that the first entry of every eigenvector of J_n is different from zero.

The zeros λ_j , $j = 1, \dots, n$, of \tilde{p}_n , i.e., the nodes of the n -node Gauss quadrature are the eigenvalues of J_n . Moreover, it can be proved (see [26], [69, Sections 2.6 and 2.9], and [45, Section 3.4.1]) that the corresponding weights ω_j are given by the squared first entries of the normalized eigenvectors of J_n multiplied by m_0 . Thus the moments of the positive definite functional \mathcal{L} are given by

$$\mathcal{L}(x^i) = m_0 \mathbf{e}_1^T (J_n)^i \mathbf{e}_1, \quad i = 0, \dots, 2n-1, \quad (2.12)$$

and the Gauss quadrature of a function f can be written in the form

$$\sum_{j=1}^n \omega_j f(\lambda_j) = m_0 \mathbf{e}_1^T f(J_n) \mathbf{e}_1, \quad (2.13)$$

where the equality follows from the spectral decomposition of the Jacobi matrix J_n . This means that the properties G1–G3 from the Introduction are satisfied and, using the associated Jacobi matrices, Gauss quadrature can be for positive definite linear functionals expressed in an algebraic way.

In literature one can find several different definitions of Jacobi matrices. Most frequently, a Jacobi matrix is defined as a real, symmetric, tridiagonal matrix with positive elements on the super-diagonal ([1, p. 2], [13, p. 72], [25, p. 13], [45, p. 30]). Jacobi matrices are important objects both in the field of matrix computations (approximating eigenvalues and eigenvectors or solving linear algebraic systems) and in approximation theory (approximating functions and integrals). They were named after Carl Gustav Jacob Jacobi (1804-1851), one of the most prolific mathematician of the 19th century. In [39] he showed that using a linear transformation with determinant equal to ± 1 it is possible to reduce any quadratic form with n variables into a particular quadratic form defined by $2n - 1$ coefficients, now expressed in terms of the $n \times n$ Jacobi matrix. The first Jacobi matrix appeared probably on page 202 of [33]. In this work Toeplitz and Hellinger studied the relationship between quadratic forms with infinitely many unknowns and the analytic theory of continued fractions by Stieltjes [63]. A review of the general theory of the unitary analogue of Jacobi matrices, the CMV matrices, can be found in [60]. For a detailed history of Jacobi matrices we refer to [45, Section 3.4.3].

Any real $n \times n$ Jacobi matrix J_n can be orthogonally diagonalized, i.e.,

$$J_n W = W \text{diag}(\lambda_1, \dots, \lambda_n),$$

where $\lambda_1, \dots, \lambda_n$ are the eigenvalues of the matrix J_n and $W = [\mathbf{w}_1, \dots, \mathbf{w}_n]$ is an orthogonal matrix whose columns are the normalized eigenvectors of J_n , $W^T W = W W^T = I$.

THEOREM 2.2 (see, e.g., [45, Theorem 3.4.3 on p. 114] and the references given there). *The following properties hold for every (real) Jacobi matrix:*

1. *Eigenvalues are real and distinct;*
2. *The first and the last component of each of its eigenvectors are nonzero.*

Let J_1, \dots, J_k be Jacobi matrices such that J_i is the leading principal $i \times i$ submatrix of J_k for $i = 1, \dots, k - 1$. From now on J_1, \dots, J_k will always denote the described sequence of Jacobi matrices.

THEOREM 2.3 (Interlacing Property, [45, Theorem 3.3.1, p. 92 and Remark 3.4.4, p. 115] and the references given there). *Let J_1, \dots, J_k be the (real) Jacobi matrices as described above. Let n and ℓ be integers with $n+1 \leq \ell \leq k$ and let $\lambda_i^{(n)}$ for $i = 1, \dots, n$ be the eigenvalues of J_n . Then at least one of the eigenvalues of J_ℓ is contained in any of the $n + 1$ open intervals*

$$\left(-\infty, \lambda_1^{(n)}\right), \left(\lambda_1^{(n)}, \lambda_2^{(n)}\right), \dots, \left(\lambda_{n-1}^{(n)}, \lambda_n^{(n)}\right), \left(\lambda_n^{(n)}, +\infty\right).$$

As a trivial consequence we get the strict interlacing property for the eigenvalues of two subsequent Jacobi matrices J_n and J_{n+1} and, equivalently, the strict interlacing property of the roots of two consecutive orthogonal polynomials.

We recall that Jacobi matrices are linked with Krylov subspace methods via the (Hermitian) Lanczos algorithm. This method was introduced by Lanczos in [43, 44] (for a description of the method and its properties we refer, for example, to [18, Section 4], [50], [25, Section 4.1], and [45, Section 2.4.1]). Given a Hermitian matrix A and a starting vector \mathbf{v}_1 such that $\|\mathbf{v}_1\| = \sqrt{\mathbf{v}_1^* \mathbf{v}_1} = 1$, the n -th iteration of the

(Hermitian) Lanczos algorithm gives an orthogonal matrix $V_n = [\mathbf{v}_1, \dots, \mathbf{v}_n]$ such that

$$J_n = V_n^* A V_n$$

where J_n is an $n \times n$ (real) Jacobi matrix. The matrices J_1, \dots, J_{n-1} obtained by the previous iterations are the leading principal submatrices of J_n .

The polynomials $\tilde{p}_0, \dots, \tilde{p}_{n-1}$ whose three-term recurrence relation is associated with J_n satisfy $\mathbf{v}_{i+1} = \tilde{p}_i(A)\mathbf{v}_1$ for $i = 0, \dots, n-1$. Moreover, they are orthonormal with respect to the functional \mathcal{L} given by $\mathcal{L}(x^i) = \mathbf{v}_1^* A^i \mathbf{v}_1 = \mathbf{e}_1^T (J_n)^i \mathbf{e}_1$ for $i = 0, 1, \dots, 2n-1$ that can be seen as the Riemann-Stieltjes integral with respect to a piecewise constant distribution function; see, e.g., [18, Example 2.1.2-b, p. 23] and, in particular, the remarkable description in the paper by Hestenes and Stiefel on the method of conjugate gradients [35, Section 14]. The columns of V_n form a basis for the Krylov subspace

$$\mathcal{K}_n(A, \mathbf{v}_1) = \text{span}\{\mathbf{v}_1, A\mathbf{v}_1, A^2\mathbf{v}_1, \dots, A^{n-1}\mathbf{v}_1\}.$$

Any (real) Jacobi matrix J_n is a result of the n -th iteration of the (Hermitian) Lanczos algorithm. In fact, it is sufficient to apply the algorithm to the matrix J_n itself with the initial vector \mathbf{e}_1 , the first vector of the canonical basis.

Orthogonal polynomials and Jacobi matrices are closely related to continued fractions, a subject of an exceptionally long history and of applications much beyond the context of our paper, and to the more recent Padé approximations; see, e.g., the seminal analytic paper by Stieltjes from 1984 [63], monographs [68, 5, 4, 8, 41] with references to very many original sources including the contributions of Euler, Chebyshev and Markov, and a sample of papers linking the approximation theory with the algebraic approach [30, 29, 15]. Our exposition uses the algebraic way and therefore we will not go further into a detailed description of Padé approximation. We find, however, useful to recall several related results on continued fractions, with referring for a detailed exposition to [45, Sections 3.3 and 3.9].

Using the coefficients of the recurrence (2.6) for the monic orthogonal polynomials, we can define for $n = 1, 2, \dots$ the n th convergent $\mathcal{F}_n(\lambda)$ of a *continued fraction*,

$$\mathcal{F}_n(x) = \frac{1}{x - \delta_0 - \frac{\eta_1}{x - \delta_1 - \frac{\eta_2}{\ddots \frac{\eta_{n-1}}{x - \delta_{n-2} - \frac{\eta_{n-1}}{x - \delta_{n-1}}}}} = \frac{\chi_n(x)}{\pi_n(x)}. \quad (2.14)$$

The n th convergent $\mathcal{F}_n(x)$ is the rational function with the denominator given by the monic orthogonal polynomial $\pi_n(x)$ of degree n . The numerator $\chi_n(x)$ is of degree $n-1$ and it satisfies for $n = 2, 3, \dots$ the same recurrence (2.6) as $\pi_n(x)$, starting from $\chi_0(x) = 0, \chi_1(x) = 1$. Continued fractions are directly related to Gauss quadrature and moments. Indeed, the partial fraction decomposition of $\mathcal{F}_n(x)$ reveals the nodes and weights of the corresponding Gauss quadrature,

$$\mathcal{F}_n(x) = \frac{\chi_n(x)}{\pi_n(x)} = \sum_{j=1}^n \frac{\omega_j}{x - \lambda_j}, \quad (2.15)$$

where $\lambda_1, \dots, \lambda_n$ and $\omega_1, \dots, \omega_n$ are the (distinct) nodes and weights of the n -node Gauss quadrature (2.4) associated with the distribution function $\mu_n(x)$ determined by the positive definite linear functional \mathcal{L} . Moreover, expanding $\mathcal{F}_n(x)$ around infinity gives

$$\mathcal{F}_n(x) = \sum_{\ell=1}^{2n} \frac{m_{\ell-1}}{x^\ell} + \mathcal{O}\left(\frac{1}{x^{2n+1}}\right) \quad (2.16)$$

with the moments of the functional \mathcal{L} as the expansion coefficients. In model reduction of linear dynamical systems developed in the second half of the 20th century this is called *minimal partial realization*, but it has been used already by Christoffel in 1858 (see [11]) with many related results by Chebyshev, Markov, Stieltjes and others; see [45, Sections 3.3.2, 3.3.5, 3.3.6, and 3.9.1].

If we change the variable in (2.16) by $x = 1/u$, we have that around zero

$$\frac{1}{u} \mathcal{F}_n\left(\frac{1}{u}\right) = m_0 + \sum_{\ell=1}^{2n-1} m_\ell u^\ell + \mathcal{O}(u^{2n}), \quad (2.17)$$

which shows that $\mathcal{F}_n(1/u)/u$ written in the form of a rational function is $[n-1, n]$ Padé approximant¹ for the (formal) power series

$$f(u) = m_0 + \sum_{\ell=1}^{\infty} m_\ell u^\ell.$$

For further details on Padé approximants we refer, e.g. to [46, Chapter VIII], [8, Chapter 5], and [41, Section 6.2].

We will now turn into generalizations of Gauss quadrature for linear functionals that are *not* positive definite. In the positive definite case the approaches using matrices of moments, interpolatory quadrature and orthogonal polynomials, Jacobi matrices and matching moments matrix representations, continued fractions (and related Padé approximations), are all mathematically equivalent and any of them can be used for describing the Gauss quadrature. In the general case the situation is different. Since (formal) orthogonal polynomials can be easily extended for moment matrices that are strongly regular, we will base our exposition on generalized Jacobi matrices that store the coefficients of the associated three term recurrences. The next two sections will summarize the known results on existence of orthogonal polynomials and on complex generalizations of Jacobi matrices.

3. Orthogonality and quasi-definite linear functionals. We start with description of quasi-definite linear functionals. Some results we refer to in this section are stated in a part of the literature only for the polynomials of a real variable. They obviously hold also when the variable is complex.

DEFINITION 3.1. *A linear functional \mathcal{L} for which the first $k+1$ Hankel determinants are nonzero, i.e., $\Delta_j \neq 0$ for $j = 0, 1, \dots, k$, is called quasi-definite on the space of polynomials \mathcal{P}_k of degree at most k .*

For the existence of orthogonal polynomials in the sense of Definition 1.1, the positive definiteness of \mathcal{L} is not needed; quasi-definiteness is necessary and sufficient.

¹The first number is degree of a nominator and the second number is degree of a denominator.

THEOREM 3.2 ([10, Chapter I, Theorem 3.1], [46, Chapter VII, Theorem 1]). A sequence $\{\pi_j\}_{j=0}^k$ of orthogonal polynomials with respect to \mathcal{L} exists if and only if \mathcal{L} is quasi-definite on \mathcal{P}_k .

Note that if \mathcal{L} is quasi-definite on \mathcal{P}_k then $\mathcal{L}(\pi_j^2) \neq 0$ for $j = 0, \dots, k$. Unlike in the positive-definite case, for \mathcal{L} quasi-definite the coefficients of the associated orthogonal polynomials are not necessarily real, the coefficients in the three-term recurrence relation are, in general, complex, and zeros of the orthogonal polynomials can be complex and multiple.

In general, any sequence of orthogonal polynomials p_0, p_1, \dots satisfies the three term recurrence relationship of the form

$$\beta_j p_j(x) = (x - \alpha_{j-1})p_{j-1}(x) - \gamma_{j-1}p_{j-2}(x), \quad \text{for } j = 1, 2, \dots, \quad (3.1)$$

where we set $\gamma_0 = 0$, $p_{-1}(x) = 0$, $p_0(x) = c$ (c is a given complex number different from zero) and

$$\alpha_{j-1} = \frac{\mathcal{L}(xp_{j-1}^2)}{\mathcal{L}(p_{j-1}^2)}, \quad \beta_j = \frac{\mathcal{L}(xp_{j-1}p_j)}{\mathcal{L}(p_j^2)}, \quad \gamma_{j-1} = \frac{\mathcal{L}(xp_{j-2}p_{j-1})}{\mathcal{L}(p_{j-2}^2)}, \quad (3.2)$$

(see [66, Theorem 3.2.1], [10, p. 19], [5, Theorem 2.4]). We remark that all denominators in (3.2) are nonzero. For monic orthogonal polynomials we have $p_0(x) = 1$ and $\beta_j = 1$, $j = 1, 2, \dots$; the formulas (3.1) and (3.2) reduce to (2.6) and (2.7) with $\alpha_{j-1} = \delta_{j-1}$ and $\gamma_{j-1} = \eta_{j-1}$. However, for a quasi definite linear functional the coefficients δ_{n-1} and η_{n-1} in (2.6) have complex values and η_{n-1} is nonzero. We can define orthonormal polynomials \tilde{p}_j as is done in (2.8) in the positive definite case.²

Providing that p_0, p_1, \dots, p_n exist, all coefficients β_0, \dots, β_n and $\gamma_1, \dots, \gamma_{n-1}$ are different from zero. The recurrences (3.1) can be written in the matrix form

$$x \begin{bmatrix} p_0(x) \\ p_1(x) \\ \vdots \\ p_{n-1}(x) \end{bmatrix} = T_n \begin{bmatrix} p_0(x) \\ p_1(x) \\ \vdots \\ p_{n-1}(x) \end{bmatrix} + \beta_n \begin{bmatrix} 0 \\ 0 \\ \vdots \\ p_n(x) \end{bmatrix}. \quad (3.3)$$

Now, T_n is a tridiagonal complex matrix

$$T_n = \begin{bmatrix} \alpha_0 & \beta_1 & & & \\ \gamma_1 & \alpha_1 & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & & \ddots & \beta_{n-1} \\ & & & \gamma_{n-1} & \alpha_{n-1} \end{bmatrix}.$$

On the other hand, as shown in [10, Chapter I, Theorem 4.4], in the survey [48, Theorem 2.14] and firstly for the positive definite case by Favard in [17], if we consider any sequence of polynomials satisfying

$$b_j p_j(x) = (x - a_{j-1})p_{j-1}(x) - c_{j-1}p_{j-2}(x), \quad j = 1, 2, \dots, \quad (3.4)$$

²This means that for a sequence of monic orthogonal polynomials π_0, \dots, π_k , there are 2^{k+1} associated sequences of orthonormal polynomials that differ in executing the complex square roots of the individual normalization coefficients. In order to avoid ambiguity, we take always the principal value of the complex square root, i.e., we consider $\arg(\sqrt{c}) \in (-\pi/2, \pi/2]$.

where

$$p_{-1}(x) = 0, p_0(x) = c, c_0 = 0, a_j, b_j, c_j, c \in \mathbb{C}, b_j, c_j, c \neq 0,$$

then there exists a quasi-definite linear functional \mathcal{L} such that p_0, p_1, \dots , are orthogonal polynomials with respect to \mathcal{L} .³ In other words, providing that $c, b_j, c_j \neq 0$, polynomials generated by (3.4) are always orthogonal polynomials. In addition, they are orthonormal if and only if $c_j = b_j$ and p_0 is such that $\mathcal{L}(p_0^2) = 1$.

This also means that for any irreducible tridiagonal matrix T_n , i.e., a tridiagonal matrix without any zero components on the sub- and super-diagonal, there exists a linear functional \mathcal{L} quasi-definite on \mathcal{P}_{n-1} such that T_n is determined by the first $2n$ moments of \mathcal{L} . As shown, e.g. in [3, proof of Theorem 2.3], two irreducible tridiagonal matrices T_n and \widehat{T}_n are determined by the first $2n$ moments of the same linear functional if and only if they are diagonally similar, i.e., if $T_n = D^{-1}\widehat{T}_n D$, where D is an invertible diagonal matrix. Or, equivalently, if and only if

- $\alpha_i = \widehat{\alpha}_i$ for $i = 0, \dots, n-1$;
- $\beta_i \gamma_i = \widehat{\beta}_i \widehat{\gamma}_i$ for $i = 1, \dots, n-1$,

where the elements of \widehat{T}_n are marked with a hat.

Any irreducible tridiagonal matrix is diagonally similar to a complex symmetric tridiagonal matrix (for the quasi-symmetric real case see [70, pp. 335-336]). Analogously to the nonuniqueness of the sequences of orthonormal polynomials mentioned above, there exist 2^{n-1} different tridiagonal symmetric matrices J_n determined by the moments m_0, \dots, m_{2n-1} . In fact, two *symmetric* irreducible tridiagonal matrices J_n and \widehat{J}_n are determined by the first $2n$ moments of a linear functional if and only if they have the same diagonal and $\beta_i = \pm \widehat{\beta}_i$ for $i = 1, \dots, n-1$.

4. Complex generalization of Jacobi matrices. For quasi-definite functionals J_n is complex symmetric but generally not Hermitian. We therefore find it useful to recall some classical results on Jacobi matrices and investigate its complex generalization.

Other definitions of Jacobi matrices can be found in [20, Vol. 2, p. 99] (a real tridiagonal matrix), [38, p. 86] (a tridiagonal matrix with a real diagonal and such that the product of the corresponding elements of the sub- and super-diagonal is non-negative), [34, p. 103] (a symmetric irreducible tridiagonal matrix with a complex diagonal). In this paper we use the definition by Beckermann from the paper about spectral properties of *complex Jacobi matrices* [2].

DEFINITION 4.1. *A square complex matrix is called Jacobi matrix if it is tridiagonal, irreducible and symmetric.*

Probably the first study of a class of this kind of matrices appeared in [68, p. 226], where Wall investigated the convergence of complex Jacobi continued fractions (J-fractions). We remark that a (complex) Jacobi matrix is Hermitian if and only if it is real.

4.1. Complex tridiagonal matrices. The relevant part of the first statement of Theorem 2.2 corresponds to the following classical result.

THEOREM 4.2 (see, e.g. [54, Lemma 7.7.1] and [27, Theorem 7.4.4]). *Every tridiagonal matrix $T_n \in \mathbb{C}^{n \times n}$ with nonzero elements on its super-diagonal (or sub-diagonal) is non-derogatory, i.e., its eigenvalues have geometric multiplicity 1.*

³Such a linear functional will be constructed for a finite number of recurrence steps in the proof of Theorem 5.1.

Proof. Let λ be an eigenvalue of the tridiagonal matrix T_n with the nonzero super-diagonal (the other case is analogous). Deleting the first column and the last row of $T_n - \lambda I$ gives a lower triangular non-singular matrix. Thus, the null space of $T_n - \lambda I$ has dimension 1 because its rank is not smaller than $n - 1$. \square

COROLLARY 4.3. *Every tridiagonal matrix $T_n \in \mathbb{C}^{n \times n}$ with nonzero elements on its super-diagonal (or sub-diagonal) is diagonalizable if and only if it has distinct eigenvalues.*

Given a tridiagonal matrix $T_n \in \mathbb{C}^{n \times n}$ with nonzero elements on its super-diagonal, $p_0(x) = c \neq 0$, we can define by (3.1) a sequence of polynomials p_0, \dots, p_n . Letting

$$\mathbf{z}(\xi) = \begin{bmatrix} p_0(\xi) \\ p_1(\xi) \\ \vdots \\ p_{n-1}(\xi) \end{bmatrix}, \quad (4.1)$$

by (3.3) we get

$$(T_n - \xi I)\mathbf{z}(\xi) = \beta_n p_n(\xi) \mathbf{e}_n.$$

Now, similarly to what is done in [16], differentiating j times the previous equation we obtain

$$(T_n - \xi I)\mathbf{z}^{(j)}(\xi) = j\mathbf{z}^{(j-1)}(\xi) + \beta_n p_n^{(j)}(\xi) \mathbf{e}_n, \quad j = 1, 2, \dots$$

Denoting

$$\mathbf{w}_0(\xi) = \mathbf{0}, \quad \mathbf{w}_1(\xi) = \mathbf{z}(\xi), \quad \mathbf{w}_{j+1}(\xi) = \frac{1}{j} \mathbf{w}_j'(\xi) = \frac{1}{j!} \mathbf{z}^{(j)}(\xi), \quad j = 1, 2, \dots, \quad (4.2)$$

we get

$$(T_n - \xi I)\mathbf{w}_{j+1}(\xi) = \mathbf{w}_j(\xi) + \frac{1}{j!} \beta_n p_n^{(j)}(\xi) \mathbf{e}_n, \quad \text{where } j = 0, 1, \dots$$

Notice that the first j elements of $\mathbf{w}_{j+1}(\xi)$ are zero since p_0, \dots, p_{j-1} have degree lower than j , while the j -th element is nonzero since $p_j(\xi)$ has degree j . Hence $\mathbf{w}_1, \dots, \mathbf{w}_n$ are linearly independent. If λ is a root of $p_n(\xi)$ with multiplicity s , then

$$(T_n - \lambda I)\mathbf{w}_{j+1}(\lambda) = \mathbf{w}_j(\lambda) \text{ for } j = 0, \dots, s-1. \quad (4.3)$$

Therefore $\mathbf{w}_1(\lambda)$ is the eigenvector and $\mathbf{w}_j(\lambda)$ for $j = 2, \dots, s$ are the *generalized eigenvectors* (called also principal vectors or Jordan canonical vectors) of T_n corresponding to λ . We summarize the previous development in the following proposition. An equivalent result was presented in the lecture of Ilse Ipsen at the ILAS 2005 conference.

PROPOSITION 4.4. *Let $T_n \in \mathbb{C}^{n \times n}$ be a tridiagonal matrix with nonzero elements on its super-diagonal. Moreover, let p_0, \dots, p_{n-1} be obtained by (3.1) with $p_0(x) = c \neq 0$ and the coefficients given by T_n . Let λ be an eigenvalue of the algebraic multiplicity s and $\mathbf{w}_{j+1}(\lambda)$ be the corresponding generalized eigenvectors satisfying $(T_n - \lambda I)\mathbf{w}_{j+1}(\lambda) = \mathbf{w}_j(\lambda)$, $j = 1, \dots, s-1$, with $\mathbf{w}_1 = \mathbf{z}(\lambda)$ from (4.1). Then*

$$\mathbf{w}_j(\lambda) = \frac{1}{(j-1)!} \begin{bmatrix} \mathbf{0}_{j-1} \\ p_{j-1}^{(j-1)}(\lambda) \\ \vdots \\ p_{n-1}^{(j-1)}(\lambda) \end{bmatrix}, \quad j = 2, \dots, s,$$

where $\mathbf{0}_\ell$ is the zero vector of length ℓ .

As for the generalization of the second statement of Theorem 2.2, we can give the following results. The formula (4.1) for $\mathbf{z}(\lambda)$ shows that for any complex tridiagonal matrix with nonzero elements on its super-diagonal the first elements of its eigenvectors are nonzero. In order to prove the same for the last eigenvector elements, we must prove that $p_{n-1}(\lambda) \neq 0$, i.e., that the eigenvalues of T_n and T_{n-1} are distinct. If λ is a root of both the orthogonal polynomials p_n and p_{n-1} , then by (3.1) it is also a root of p_{n-2} . By induction we conclude that $p_0 = 0$, that is a contradiction.

4.2. Complex symmetric matrices. Unlike real symmetric matrices, complex symmetric matrices may not be diagonalizable. This fact is linked with the existence (in the complex field) of *isotropic* vectors. An *isotropic* vector is a vector \mathbf{x} such that $\mathbf{x}^T \mathbf{x} = 0$ and $\mathbf{x} \neq 0$ (for example $(1, i)^T$). In [14, Theorem 3] Craven proved the following theorem.

THEOREM 4.5. *If A is a complex symmetric matrix, then the following statements are equivalent:*

1. *There exists a (complex) nonsingular matrix V such that $V^{-1} = V^T$ and $V^T A V$ is a diagonal matrix;*
2. *Every eigenspace of A has a basis $\mathbf{v}_1, \dots, \mathbf{v}_s$ without isotropic vectors and such that $\mathbf{v}_i^T \mathbf{v}_j = 0$ for $i \neq j$.*

Moreover, [59, theorems 1 and 3] give the following equivalence.

THEOREM 4.6. *A singular complex symmetric matrix contains an isotropic vector in its null space if and only if the trace of its adjoint vanishes.*

We will use it in the following lemma; we refer again to the lecture of Ipsen at the ILAS 2005, p. 26.

LEMMA 4.7. *Let λ be an eigenvalue of a complex Jacobi matrix J and \mathbf{v} an associated eigenvector. Then, \mathbf{v} is isotropic if and only if λ has algebraic multiplicity greater than 1.*

Proof. Given a matrix $A(\xi)$ depending on a parameter ξ , Jacobi's formula states that

$$\frac{d}{d\xi} \det A(\xi) = \text{tr}(\text{adj}(A(\xi)) \frac{dA(\xi)}{d\xi}),$$

where adj is the adjoint matrix (sometimes it is used the term adjugate to avoid confusion with the Hermitian adjoint); for a proof see, e.g., [49, Theorem 1 at p. 149]. If $A(\xi) = \xi I - J$, then the previous formula becomes

$$\frac{d}{d\xi} \phi(\xi) = \text{tr}(\text{adj}(\xi I - J)),$$

where, ϕ is the characteristic polynomial of J . Let λ be an eigenvalue of J , then $(\lambda I - J)$ is a complex symmetric matrix such that

$$\det(\lambda I - J) = 0 \text{ and } \text{tr}(\text{adj}(\lambda I - J)) = \phi'(\lambda).$$

Since $\phi'(\lambda) = 0$ if and only if the algebraic multiplicity of λ is greater than 1, by Theorem 4.6 the eigenspace of J corresponding to λ contains an isotropic vector if and only if the algebraic multiplicity of λ is greater than 1. Since by Theorem 4.2 any complex Jacobi matrix is non-derogatory, the proof is finished. \square

We will summarize the situation in the following proposition.

PROPOSITION 4.8. *If J is a Jacobi matrix, then the following properties are equivalent:*

1. J is diagonalizable;
2. There exist a (complex) nonsingular matrix V such that $V^{-1} = V^T$ and $V^T J V$ is a diagonal matrix;
3. None of the eigenvectors of J is isotropic.

Proof. The second and the third properties are equivalent by Theorem 4.5. Obviously the second one implies the first one. So it remains to prove that if J is diagonalizable, then no eigenvector is isotropic. Since J is non-derogatory, using Lemma 4.7 finishes the proof. \square

In the rest of this section we recall the relationship between Jacobi matrices and non-Hermitian Lanczos algorithm (for details we refer to [5, Section 2.7.2], [25, Section 4.2], [45, Section 2.4.2] and [57, Chapter 7]). The input of the algorithm consists of a matrix A and two vectors $\mathbf{w}_1, \mathbf{v}_1$ such that $\|\mathbf{v}_1\| = 1$ and $\mathbf{w}_1^* \mathbf{v}_1 = 1$. Assuming that the algorithm does not breakdown before the n -th iteration, we obtain as the result of the first n iterations the matrices $V_n = [\mathbf{v}_1, \dots, \mathbf{v}_n]$ and $W_n = [\mathbf{w}_1, \dots, \mathbf{w}_n]$ whose columns form bases of the Krylov subspaces $\mathcal{K}_n(A, \mathbf{v}_1)$ and $\mathcal{K}_n(A^*, \mathbf{w}_1)$ respectively with the biorthogonality property $W_n^* V_n = I$. The associated tridiagonal matrix T_n

$$T_n = W_n^* A V_n$$

is not, in general, symmetric, and therefore it does not represent a Jacobi matrix. Let p_0, \dots, p_{n-1} be the sequence of polynomials determined by T_n ; see Section 3. Then they are orthogonal with respect to the quasi-definite functional \mathcal{L} such that

$$\mathcal{L}(x^i) = \mathbf{w}_1^* A^i \mathbf{v}_1 = \|\mathbf{w}_1\| \|\mathbf{v}_1\| \mathbf{e}_1^T (T_n)^i \mathbf{e}_1, \quad \text{for } i = 0, 1, \dots, 2n-1.$$

Furthermore, $\mathbf{w}_{i+1} = p_i(A^*) \mathbf{w}_1$ and $\mathbf{v}_{i+1} = p_i(A) \mathbf{v}_1$ for $i = 0, \dots, n-1$. The Lanczos vectors can be normalized in different ways. In particular, they can be normalized in such a way that the matrix T_n is complex symmetric and therefore a Jacobi matrix.

We have assumed no breakdown of the non-Hermitian Lanczos algorithm in steps 1 through n . The complicated issues related to the breakdown are outlined, e.g., in [45, p. 33], with the detailed exposition presented in [55, 6, 7, 31, 53, 32].

5. Moment Matching Property for Jacobi matrices. If the values of the linear functional on monomials are defined by $\mathcal{L}(x^i) = \mathbf{v}^* A^i \mathbf{v}$, $i = 0, 1, \dots$, where A is a Hermitian matrix and \mathbf{v} is a nonzero vector, then the associated orthogonal polynomials are given by the Lanczos (Stieltjes) algorithm; see, e.g., [25, Chapter 7] or [45, Section 3.5]). Then (recall (2.12) and (2.13) in Section 2 above)

$$\mathcal{L}(x^i) = \mathbf{v}^* A^i \mathbf{v} = \|\mathbf{v}\|^2 \mathbf{e}_1^T (J_n)^i \mathbf{e}_1 = m_0 \mathbf{e}_1^T (J_n)^i \mathbf{e}_1, \quad i = 0, 1, \dots, 2n-1,$$

where J_n is the Jacobi matrix associated with the first n steps of the Lanczos process. Using the Vorobyev method of moments [67, in particular Chapter III], this property can be easily extended, assuming existence of the first n steps of the non-Hermitian Lanczos process, to a general complex matrix A ; see [64]. Here we will prove an analogous property for Jacobi matrices determined by quasi-definite linear functionals.

THEOREM 5.1 (Moment Matching Property). *Let \mathcal{L} be a quasi-definite linear functional on \mathcal{P}_n and let J_n be the Jacobi matrix of coefficients from the recurrence relations for orthogonal polynomials with respect to \mathcal{L} ; see (3.1) – (3.3) and the discussion at the end of Section 3. Then*

$$\mathcal{L}(x^i) = m_0 \mathbf{e}_1^T (J_n)^i \mathbf{e}_1, \quad i = 0, \dots, 2n-1, \quad (5.1)$$

where $m_0 = \mathcal{L}(x^0)$.

Proof. We first prove that p_0, \dots, p_{n-1} are orthogonal with respect to the functional $\tilde{\mathcal{L}}$ defined by

$$\tilde{\mathcal{L}}(x^i) = m_0 \mathbf{e}_1^T (J_n)^i \mathbf{e}_1,$$

with $m_0 = 1/p_0^2$.⁴ For $i = 0, \dots, n-1$ the $(i+1)$ -st entry of the vector $(J_n)^i \mathbf{e}_1$ is nonzero. Moreover, for $i = 0, \dots, n-2$, the entries $i+2, \dots, n$ of $(J_n)^i \mathbf{e}_1$ are zero. Hence the canonical basis $\mathbf{e}_1, \dots, \mathbf{e}_k$ is an orthonormal basis of $\mathcal{K}_k(J_n, \mathbf{e}_1)$, $k = 1, \dots, n$, i.e., $\mathbf{e}_k = \tilde{p}_{k-1}(J_n) \mathbf{e}_1$ for some polynomial \tilde{p}_{k-1} of degree $k-1$.

The polynomials $\hat{p}_{k-1} = \tilde{p}_{k-1}/\sqrt{m_0}$, $k = 1, \dots, n$, are orthonormal with respect to $\tilde{\mathcal{L}}$. Indeed,

$$\tilde{\mathcal{L}}(\hat{p}_i \hat{p}_j) = m_0 \mathbf{e}_1^T \hat{p}_i(J_n) \hat{p}_j(J_n) \mathbf{e}_1 = \mathbf{e}_i^T \mathbf{e}_j.$$

From $\mathbf{e}_1 = \tilde{p}_0(J_n) \mathbf{e}_1$ we get $\tilde{p}_0 = 1$, i.e., $\hat{p}_0 = 1/\sqrt{m_0} = p_0$. Finally, we prove that $\hat{p}_k = p_k$ for $k = 1, \dots, n-1$. Notice that

$$\tilde{\mathcal{L}}(x \hat{p}_i \hat{p}_j) = m_0 \mathbf{e}_1^T \hat{p}_i(J_n) J_n \hat{p}_j(J_n) \mathbf{e}_1 = (J_n)_{i,j}.$$

Therefore, by (2.9) we see that the coefficients from the three-term recurrence relation for $x \hat{p}_0, \dots, x \hat{p}_{n-1}$ are the same as those for $x p_0, \dots, x p_{n-1}$.

It remains to prove that

$$\mathcal{L}(x^i) = \tilde{\mathcal{L}}(x^i) \text{ for } i = 0, \dots, 2n-1. \quad (5.2)$$

The proof uses an analogy of the property of the matrix of moments for the positive definite case recalled in Section 2. Let m_i and \tilde{m}_i be the moments of, respectively, \mathcal{L} and $\tilde{\mathcal{L}}$, let $M_{n-1} = (m_{i+j})_{i,j=0,\dots,n-1}$ (see (2.1)) and U_{n-1} be the upper triangular matrix with the elements in the $(j+1)$ -th column equal to the coefficients of p_j . Then $U_{n-1}^T M_{n-1} U_{n-1} = I_n$ since the polynomials are orthogonal with respect to \mathcal{L} . Repeating the argument for $\tilde{M}_{n-1} = (\tilde{m}_{i+j})_{i,j=0,\dots,n-1}$ gives

$$U_{n-1}^T M_{n-1} U_{n-1} = U_{n-1}^T \tilde{M}_{n-1} U_{n-1}.$$

Since the diagonal elements of U_{n-1} are nonzero, U_{n-1} is invertible and we get $M_{n-1} = \tilde{M}_{n-1}$. Finally, by (3.2) we have

$$\frac{\mathcal{L}(x p_{n-1}^2)}{\mathcal{L}(p_{n-1}^2)} = \alpha_{n-1} = \frac{\tilde{\mathcal{L}}(x p_{n-1}^2)}{\tilde{\mathcal{L}}(p_{n-1}^2)}$$

and therefore $m_{2n-1} = \tilde{m}_{2n-1}$. \square

As remarked at the end of Section 3, a different normalization of the orthogonal polynomials is associated with a tridiagonal matrix $T_n = D J_n D^{-1}$. Hence, the statement of Theorem 5.1 remains valid for any tridiagonal matrix T_n associated with a sequence of orthogonal polynomials defined by the functional \mathcal{L} .

⁴This way we actually prove the finite dimensional counterpart of Favard Theorem (see Section 3).

6. Quasi-definite linear functionals and Gauss quadrature under restrictive assumptions. As recalled in the Introduction and in Section 2, *positive definite* linear functionals lead to Gauss quadrature with the properties G1 – G3. Since the degree of exactness is larger than $n - 1$, the Gauss quadrature (2.4) is interpolatory quadrature, i.e., the weights ω_i satisfy

$$\omega_i = \mathcal{L}(\ell_i), \quad i = 1, \dots, n, \quad (6.1)$$

where $\ell_i(x)$ is the *Lagrange interpolation polynomial*, defined as

$$\ell_i(x) = \frac{(x - \lambda_1) \dots (x - \lambda_{i-1})(x - \lambda_{i+1}) \dots (x - \lambda_n)}{(\lambda_i - \lambda_1) \dots (\lambda_i - \lambda_{i-1})(\lambda_i - \lambda_{i+1}) \dots (\lambda_i - \lambda_n)} = \frac{\pi_n(x)}{(x - \lambda_i)\pi'_n(\lambda_i)}.$$

The nodes of the Gauss quadrature are the eigenvalues of J_n and the weights are equal to the squared first entries of the associated normalized eigenvectors of J_n multiplied by m_0 ; see Section 2.

We will now revisit the situation for the functional \mathcal{L} that is only quasi-definite. We start with the usual form of an n -node quadrature (cf. (2.4))

$$\mathcal{L}(f) = \sum_{i=1}^n \omega_i f(\lambda_i) + R_n(f), \quad (6.2)$$

where the nodes $\lambda_1, \dots, \lambda_n$ are complex and distinct and the last term stands for the quadrature error.

THEOREM 6.1. *The quadrature (6.2) is exact for every f from \mathcal{P}_{2n-1} if and only if it is interpolatory and the polynomial*

$$\varphi_n(x) = \prod_{i=1}^n (x - \lambda_i) \quad (6.3)$$

satisfies $\mathcal{L}(\varphi_n p) = 0$ for every $p \in \mathcal{P}_{n-1}$.

Proof. Assume that (6.2) is exact for every f from \mathcal{P}_{2n-1} . Then for every $p \in \mathcal{P}_{n-1}$ we get $R_n(\varphi_n p) = 0$ and therefore, using $\varphi_n(\lambda_i) = 0$ for $i = 1, \dots, n$,

$$\mathcal{L}(\varphi_n p) = \sum_{i=1}^n \omega_i \varphi_n(\lambda_i) p(\lambda_i) = 0.$$

Inversely, assume $\mathcal{L}(\varphi_n p) = 0$ for all $p \in \mathcal{P}_{n-1}$. Since any $f \in \mathcal{P}_{2n-1}$ can be written in the form $f(x) = \varphi_n(x)q(x) + r(x)$ for some q and r from \mathcal{P}_{n-1} , $\mathcal{L}(f) = \mathcal{L}(r)$. An interpolatory quadrature on n nodes must have algebraic degree of exactness at least $n - 1$. Therefore $\mathcal{L}(r) = \sum_{i=1}^n \omega_i r(\lambda_i)$. The standard argument $r(\lambda_i) = f(\lambda_i)$ for $i = 1, \dots, n$ completes the proof. \square

So, the quadrature rule (6.2) has the properties G1 and G2 if and only if the following conditions simultaneously hold:

1. There exists a sequence of orthogonal polynomials p_0, \dots, p_n with respect to the linear functional \mathcal{L} (i.e., \mathcal{L} is quasi-definite on \mathcal{P}_n);
2. Zeros of the individual polynomials p_j , $j = 1, \dots, n$, in the sequence are distinct; i.e., the matrices J_j , $j = 1, \dots, n$, are diagonalizable.

The standard argument then shows that the quadrature rule (6.2) can be expressed in the form $m_0 \mathbf{e}_1^T f(J_n) \mathbf{e}_1$, i.e., the property G3 is generalized in a straightforward way; see [23, p. 153], [58, p. 267-268]. Quadrature (6.2) was considered (to our

knowledge) for the first time by Gragg in [29] for real valued linear functionals. A generalization for complex valued functionals was considered by Saylor and Smolarski in [58] ([65, p. 6 Theorem 2] also implicitly considered, without further elaboration on it, a quadrature formula with complex nodes). Due to the assumption on distinct roots (see Property 2 above) this construction is restrictive.

Indeed, if \mathcal{L} is quasi-definite on \mathcal{P}_k , then the orthogonal polynomials in the sequence p_1, \dots, p_k can have multiple zeros. Hence it can happen that for some values ℓ , $\ell \leq k$, the ℓ -point interpolatory quadrature defined by

$$\mathcal{L}(f) \approx \sum_{i=1}^{\ell} \omega_i f(\lambda_i)$$

cannot be properly defined (i.e., it represents an interpolatory quadrature on strictly less than ℓ distinct points) and it cannot achieve the algebraic degree of exactness $2\ell - 1$. This is illustrated in the following example.

Example 1. Consider the linear functional \mathcal{L} defined by a sequence of moments with the first seven terms given by

$$1, 3, 8, 20, 52, 156, i.$$

Then \mathcal{L} is quasi-definite on \mathcal{P}_3 , since

$$\Delta_0 = 1, \quad \Delta_1 = -1, \quad \Delta_2 = -4, \quad \Delta_3 = 2128 - 4i.$$

The associated monic orthogonal polynomials are

$$\pi_0 = 1, \quad \pi_1(x) = x - 3, \quad \pi_2(x) = x^2 - 4x + 4, \quad \pi_3(x) = x^3 - 7x^2 + 20x - 24.$$

The zeros of π_2 are $\lambda_1 = \lambda_2 = 2$, which means that the 2-node quadrature (6.2) which is exact on \mathcal{P}_3 does not exist. However, the zeros of π_3 are $\lambda_1 = 3$, $\lambda_2 = 2 - 2i$ and $\lambda_3 = 2 + 2i$, which means that there exists the 3-node quadrature (6.2) which is exact on \mathcal{P}_5 . The corresponding Jacobi matrix is

$$J_3 = \begin{bmatrix} 3 & i & 0 \\ i & 1 & 2i \\ 0 & 2i & 3 \end{bmatrix}.$$

The matrix J_3 is diagonalizable, whereas its leading principal 2×2 submatrix is not.

7. Gauss quadrature for general quasi-definite linear functionals. In order to avoid restrictions to quasi-definite linear functionals that produce diagonalizable Jacobi matrices, and allow full generality, we have to modify the quadrature concept presented in relation (6.2). In particular, we will consider, in agreement with [51, Section 2, p. 122]⁵, the *n-weight quadrature formula*

$$\mathcal{L}(f) = \sum_{i=1}^{\ell} \sum_{j=0}^{s_i-1} \omega_{i,j} f^{(j)}(\lambda_i) + R_n(f), \quad (7.1)$$

where $n = s_1 + \dots + s_\ell$. Note that the rule (6.2) is the special case of the rule (7.1) when $\ell = n$ and $s_1 = \dots = s_n = 1$. So we generalize the rule (6.2) in the way that

⁵As mentioned in the Introduction, the focus of the presented paper is different from [51].

incorporates, in addition to the functional values $f(\lambda_1), \dots, f(\lambda_\ell)$, also the values of the derivatives of f at the points $\lambda_1, \dots, \lambda_\ell$. The generalization (7.1) therefore requires more smoothness of the argument function f in $\mathcal{L}(f)$ to be approximated. The following theorems justify the given construction and tell us how to choose the values of s_1, \dots, s_ℓ when we want to achieve the maximal degree of exactness.

THEOREM 7.1. *Let \mathcal{L} be an arbitrary linear functional on \mathcal{P} . The quadrature (7.1) is exact for every f from \mathcal{P}_{2n-1} if and only if it is exact on \mathcal{P}_{n-1} and the polynomial*

$$\varphi_n(x) = (x - \lambda_1)^{s_1} (x - \lambda_2)^{s_2} \dots (x - \lambda_\ell)^{s_\ell} \quad (7.2)$$

satisfies $\mathcal{L}(\varphi_n p) = 0$ for every $p \in \mathcal{P}_{n-1}$.

Proof. Following [62] we consider for each root λ_i , $i = 1, \dots, \ell$, of φ_n the following s_i polynomials of degree $n - 1$

$$h_{i,j}(x) = \frac{(x - \lambda_i)^j}{j!} \left\{ \sum_{\nu=0}^{s_i-1-j} \frac{(x - \lambda_i)^\nu}{\nu!} \left(\frac{1}{g_i(x)} \right)^{(\nu)} \Big|_{x=\lambda_i} \right\} g_i(x), \quad (7.3)$$

$$j = 0, 1, \dots, s_i - 1,$$

where $g_i(x) = \prod_{\substack{t=1 \\ t \neq i}}^{\ell} (x - \lambda_t)^{s_t}$. As proved in [61, Section 3], from (7.3) we obtain

$$h_{i,j}^{(t)}(\lambda_k) = 1 \quad \text{for } \lambda_k = \lambda_i \text{ and } t = j,$$

$$h_{i,j}^{(t)}(\lambda_k) = 0 \quad \text{for } \lambda_k \neq \lambda_i \text{ or } t \neq j,$$

where $k = 1, 2, \dots, \ell$, and $t = 0, 1, \dots, s_i - 1$. Defining the generalized (Hermite) interpolating polynomial (see [61])

$$h_{n-1}(x) = \sum_{i=1}^{\ell} \sum_{j=0}^{s_i-1} f^{(j)}(\lambda_i) h_{i,j}(x),$$

we get that the formula (7.1) is exact for any polynomial f of degree at most $n - 1$ if and only if

$$\mathcal{L}(f) = \sum_{i=1}^{\ell} \sum_{j=0}^{s_i-1} w_{i,j} f^{(j)}(\lambda_i) = \sum_{i=1}^{\ell} \sum_{j=0}^{s_i-1} \mathcal{L}(h_{i,j}) f^{(j)}(\lambda_i),$$

i.e., if and only if the weights of the quadrature (7.1) are given by

$$\omega_{i,j} = \mathcal{L}(h_{i,j}).$$

The rest of the proof is fully analogous to the proof of Theorem 6.1. \square

We say that the n -weight quadrature (7.1) is unique if the nodes and the weights are uniquely determined by \mathcal{L} and n .

THEOREM 7.2. *Let \mathcal{L} be an arbitrary linear functional on \mathcal{P} . The n -weight quadrature (7.1) of degree of exactness at least $2n - 1$ exists and is unique if and only if the n -th Hankel determinant (2.1) is nonzero, i.e., $\Delta_{n-1} \neq 0$.*

Proof. Theorem 7.1 says that the n -weight interpolatory quadrature (7.1) is of degree of exactness at least $2n - 1$ if and only if the monic polynomial

$$\varphi_n(x) = x^n + c_{n-1}x^{n-1} + \dots + c_1x + c_0$$

given by (7.2) is orthogonal to the space \mathcal{P}_{n-1} . The conditions $\mathcal{L}(x^j\varphi_n) = 0$, $j = 0, \dots, n-1$, are satisfied if and only if the linear system

$$\begin{bmatrix} m_0 & m_1 & \dots & m_{n-1} \\ m_1 & m_2 & \dots & m_n \\ \vdots & \vdots & \ddots & \vdots \\ m_{n-1} & m_n & \dots & m_{2n-2} \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_{n-1} \end{bmatrix} = \begin{bmatrix} -m_n \\ -m_{n+1} \\ \vdots \\ -m_{2n-1} \end{bmatrix} \quad (7.4)$$

has a unique solution. \square

Finally, the following theorem gives the condition under which the degree of exactness of (7.1) is exactly $2n - 1$ (i.e., it does not exceed $2n - 1$). This issue has no counterpart in the positive-definite case where the n -node Gauss quadrature cannot have algebraic degree of exactness larger than $2n - 1$.

THEOREM 7.3. *Let \mathcal{L} be an arbitrary linear functional on \mathcal{P} and let the n -weight quadrature (7.1) has degree of exactness at least $2n-1$. Then the degree of exactness of the quadrature (7.1) is (exactly) $2n-1$ if and only if the $(n+1)$ -st Hankel determinant (2.1) is nonzero, i.e., $\Delta_n \neq 0$.*

Proof. Since the n -weight quadrature (7.1) has degree of exactness at least $2n - 1$, the polynomial φ_n given by (7.2) is orthogonal to \mathcal{P}_{n-1} . Moreover, φ_n is orthogonal to \mathcal{P}_n if and only if $\mathcal{L}(\varphi_n^2) = 0$ in which case the degree of exactness of (7.1) is at least $2n$. Thus we conclude that the quadrature (7.1) has degree of exactness larger than $2n - 1$ if and only if $\mathcal{L}(\varphi_n x^j) = 0$ for $j = 0, \dots, n$, i.e., if and only if there is a vector $[c_0, \dots, c_{n-1}, 1]^T$ such that

$$\begin{bmatrix} m_0 & m_1 & \dots & m_n \\ m_1 & m_2 & \dots & m_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ m_n & m_{n+1} & \dots & m_{2n} \end{bmatrix} \begin{bmatrix} c_0 \\ \vdots \\ c_{n-1} \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \quad (7.5)$$

in which case $\Delta_n = 0$. \square

COROLLARY 7.4. *The quadrature rule (7.1) has the properties G1 and G2 if and only if \mathcal{L} is quasi-definite on \mathcal{P}_n .*

Proof. The n -weight quadrature (7.1) is unique and of an algebraic degree of exactness $2n - 1$ if and only if both Δ_{n-1} and Δ_n are nonvanishing. The property G2 requires the same for all j -weight quadratures with $j = 1, \dots, n - 1$, and thus all Hankel determinants Δ_j , $j = 0, \dots, n$ have to be nonvanishing; i.e., \mathcal{L} has to be quasi-definite on \mathcal{P}_n . \square

We say that a function f is *defined on the spectrum of the given matrix J* if for every eigenvalue λ_i of J there exist $f^{(j)}(\lambda_i)$ for $j = 0, 1, \dots, s_i - 1$, where s_i is the order of the largest Jordan block of J in which λ_i appears. Let Λ be a Jordan block of J of the size s corresponding to the eigenvalue λ . The matrix function $f(\Lambda)$ is then

defined as

$$f(\Lambda) = \begin{bmatrix} f(\lambda) & \frac{f'(\lambda)}{1!} & \frac{f^{(2)}(\lambda)}{2!} & \cdots & \frac{f^{(s-1)}(\lambda)}{(s-1)!} \\ 0 & f(\lambda) & \frac{f'(\lambda)}{1!} & \cdots & \frac{f^{(s-2)}(\lambda)}{(s-2)!} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \frac{f'(\lambda)}{1!} \\ 0 & \cdots & \cdots & 0 & f(\lambda) \end{bmatrix}.$$

Denoting

$$J = W \text{diag}(\Lambda_1, \dots, \Lambda_\nu) W^{-1}$$

the Jordan normal form of J , the matrix function $f(J)$ is defined by

$$f(J) = W \text{diag}(f(\Lambda_1), \dots, f(\Lambda_\nu)) W^{-1};$$

for further details on matrix functions we refer to [36]. Let J_n be an $n \times n$ Jacobi matrix, with λ_i its eigenvalues of the algebraic multiplicities s_i , $i = 1, \dots, \ell$. By Theorem 4.2 the matrix J_n is non-derogatory. Denoting the first row of W as

$$\mathbf{w}^T = [w_{1,0}, \dots, w_{1,s_1-1}, w_{2,0}, \dots, w_{2,s_2-1}, \dots, w_{\ell,0}, \dots, w_{\ell,s_\ell-1}],$$

and the first column of W^{-1} as

$$\hat{\mathbf{w}} = [\hat{w}_{1,0}, \dots, \hat{w}_{1,s_1-1}, \hat{w}_{2,0}, \dots, \hat{w}_{2,s_2-1}, \dots, \hat{w}_{\ell,0}, \dots, \hat{w}_{\ell,s_\ell-1}]^T,$$

we get

$$\begin{aligned} \mathbf{e}_1^T f(J_n) \mathbf{e}_1 &= \mathbf{e}_1^T W \text{diag}(f(\Lambda_1), \dots, f(\Lambda_\ell)) W^{-1} \mathbf{e}_1 \\ &= \mathbf{w}^T \text{diag}(f(\Lambda_1), \dots, f(\Lambda_\ell)) \hat{\mathbf{w}} \\ &= \sum_{i=1}^{\ell} [w_{i,0}, \dots, w_{i,s_i-1}] f(\Lambda_i) [\hat{w}_{i,0}, \dots, \hat{w}_{i,s_i-1}]^T. \end{aligned}$$

Using (4.1) and Proposition 4.4, the first elements of the columns of the matrix W are zero except for the columns that are eigenvectors of J_n . Therefore, the individual terms in the previous sum can be rewritten as

$$[w_{i,0}, 0, \dots, 0] \begin{bmatrix} f(\lambda_i) & \frac{f'(\lambda_i)}{1!} & \cdots & \frac{f^{(s_i-1)}(\lambda_i)}{(s_i-1)!} \\ 0 & f(\lambda_i) & \cdots & \frac{f^{(s_i-2)}(\lambda_i)}{(s_i-2)!} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & f(\lambda_i) \end{bmatrix} \begin{bmatrix} \hat{w}_{i,0} \\ \hat{w}_{i,1} \\ \vdots \\ \hat{w}_{i,s_i-1} \end{bmatrix}.$$

Hence we have

$$\mathbf{e}_1^T f(J_n) \mathbf{e}_1 = \sum_{i=1}^{\ell} \sum_{j=0}^{s_i-1} \frac{w_{i,0} \hat{w}_{i,j}}{j!} f^{(j)}(\lambda_i) = \sum_{i=1}^{\ell} \sum_{j=0}^{s_i-1} \tilde{\omega}_{i,j} f^{(j)}(\lambda_i), \quad (7.6)$$

with

$$\tilde{\omega}_{i,j} = \frac{w_{i,0} \hat{w}_{i,j}}{j!}, \quad \text{for } i = 1, \dots, \ell, \quad j = 0, \dots, s_i - 1.$$

Using $\omega_{i,j} = m_0 \tilde{\omega}_{i,j}$ in (7.6) we get

$$m_0 \mathbf{e}_1^T f(J_n) \mathbf{e}_1 = \sum_{i=1}^{\ell} \sum_{j=0}^{s_i-1} \omega_{i,j} f^{(j)}(\lambda_i). \quad (7.7)$$

Now we are able to state and prove the following corollary.

COROLLARY 7.5. *The quadrature rule (7.1) having the properties G1 and G2 satisfies also the property G3.*

Proof. The right-hand side of (7.7) is of the form (7.1). It remains to prove that the weights $\omega_{i,j}$ are indeed equal to $\mathcal{L}(h_{i,j})$, where polynomials $h_{i,j}$ are defined by (7.3); see the proof of Theorem 7.1. Since the quadrature (7.1) satisfies the properties G1 and G2, by Corollary 7.4 the functional \mathcal{L} is quasi-definite on \mathcal{P}_n . Using Theorem 5.1, its values on monomials x^i must then be equal for $i = 0, 1, \dots, 2n-1$ to the right-hand side of (7.7) with $f(\lambda)$ replaced by the same monomials. Consequently, the right-hand side of (7.7) represents a quadrature with the algebraic degree at least $2n-1$. From uniqueness it must be equal to the quadrature (7.1) with the weights $\mathcal{L}(h_{i,j})$ and the proof is finished. \square

We will summarize the characterization of the n -weight quadrature formula (7.1) in terms of the associated Jacobi matrix as a theorem.

THEOREM 7.6. *Let \mathcal{L} be a linear functional on \mathcal{P} and $m_0 = \mathcal{L}(x^0)$. There exists a Jacobi matrix J_n of dimension n such that*

$$\mathcal{L}(x^i) = m_0 \mathbf{e}_1^T (J_n)^i \mathbf{e}_1, \quad \text{for } i = 0, \dots, 2n-1, \quad (7.8)$$

$$\mathcal{L}(x^{2n}) \neq m_0 \mathbf{e}_1^T (J_n)^{2n} \mathbf{e}_1, \quad (7.9)$$

if and only if \mathcal{L} is quasi-definite on \mathcal{P}_n .

Proof. Let J_n be the Jacobi matrix satisfying (7.8) and (7.9). Then, by (7.7) there exists the n -weight quadrature (7.1) whose degree of exactness is (exactly) $2n-1$. By Theorem 7.3 it follows that $\Delta_n \neq 0$. To prove that \mathcal{L} is quasi-definite on \mathcal{P}_n it remains to prove that \mathcal{L} is quasi-definite on \mathcal{P}_{n-1} . As shown in the proof of Theorem 5.1, the polynomials p_0, \dots, p_{n-1} associated with the three-term recurrence relation whose coefficients are given by J_n are orthonormal with respect to the linear functional

$$\tilde{\mathcal{L}}(f) = m_0 \mathbf{e}_1^T f(J_n) \mathbf{e}_1, \quad \text{for } f \in \mathcal{P}.$$

They are also orthonormal with respect to \mathcal{L} by (7.8), which means, using Theorem 3.2, that \mathcal{L} is quasi-definite on \mathcal{P}_{n-1} . The statement in the opposite direction follows directly by corollaries 7.4 and 7.5. \square

The presented construction (7.1) and the statements proved throughout this section show that it is possible to construct the n -weight quadrature (7.1) having the properties G1–G3 of the classical Gauss quadrature whenever the linear functional is quasi-definite. In order to avoid confusion, it should be stressed that the Gauss quadrature proposed in this paper (the quadrature (7.1) satisfying G1–G3) is different from the Gauss quadrature with multiple nodes considered in [9] and [56], and later in [24]. The latter assumes *positive-definite linear functionals* and its degree of exactness is equal to

$$(\text{the number of weights}) + (\text{the number of nodes}) - 1.$$

The Gauss quadrature proposed in this paper is constructed for *quasi-definite linear functionals* and it has the degree of exactness

$$2 \times (\text{the number of weights}) - 1$$

that is larger than in the previous case.

We conclude this Section by another interesting link between n -weight quadrature (7.1) and Padé approximants. Hankel determinants Δ_j , $j = 0, 1, \dots$, coincide with super-diagonal entries $c_{j,j+1}$ in the so-called c -table for the power series $f(u) = \sum_{k=0}^{\infty} m_k u^k$; see, e.g., [30]. The conditions $\Delta_{n-1} = c_{n-1,n} \neq 0$ and $\Delta_n = c_{n,n+1} \neq 0$ provide that the denominator $q(x)$ of the $[n-1, n]$ Padé approximant $r(x) = p(x)/q(x)$ is of degree n and that the Maclaurin expansion of $r(u)$ agrees with $f(u)$ exactly through the power u^{2n-1} . It further means that around infinity

$$\frac{1}{x} r\left(\frac{1}{x}\right) = \sum_{k=0}^{2n-1} \frac{m_k}{x^{k+1}} + \mathcal{O}\left(\frac{1}{x^{2n+1}}\right). \quad (7.10)$$

Expanding every summand in the partial fraction decomposition

$$\frac{1}{x} r\left(\frac{1}{x}\right) = \sum_{i=1}^{\ell} \sum_{j=0}^{s_i-1} \frac{A_{i,j}}{(x - \alpha_i)^j} \quad (7.11)$$

around infinity, we get that

$$\frac{1}{x} r\left(\frac{1}{x}\right) = \sum_{k=0}^{\infty} \left(\sum_{i=1}^{\ell} \sum_{j=0}^{s_i-1} \frac{A_{i,j}}{j!} \{(x^k)^{(j)}\}_{x=\alpha_i} \right) \frac{1}{x^{k+1}}. \quad (7.12)$$

Comparing (7.10) and (7.12) we get that

$$\sum_{i=1}^{\ell} \sum_{j=0}^{s_i-1} \frac{A_{i,j}}{j!} \{(x^k)^{(j)}\}_{x=\alpha_i} = m_k, \quad k = 0, \dots, 2n-1,$$

which means that $A_{i,j}/(j!)$ are the weights $\omega_{i,j}$ and α_i are the nodes λ_i in the unique n -weight quadrature having degree of exactness (exactly) $2n-1$. We remark that in [51, pp. 124–125] this relation with Padé approximant was used to define the nodes λ_i and the weights $\omega_{i,j}$ when the linear functional is an integral with respect to some complex measure.

8. Conclusion. In Section 7 we saw that quasi-definiteness of \mathcal{L} is not only sufficient but also necessary condition for the n -weight quadrature (7.1) to have all three properties G1, G2 and G3. For *non-definite linear functionals* all three properties cannot hold.

Let \mathcal{L} be a linear functional such that the n -th Hankel determinant (2.1) is equal to zero, i.e., $\Delta_{n-1} = 0$. Using Theorem 7.2, the n -weight quadrature (7.1) having degree of exactness at least $2n-1$ either does not exist (the system (7.4) has no solution), or there are infinitely many of them (the system (7.4) has infinitely many solutions). Thus the property G2 cannot be satisfied. If there exist infinitely many n -weight quadratures (7.1), then Δ_n must also be equal to zero. Indeed, using (7.4), the first n rows of the matrix of the system (7.5) are linearly dependent. Hence by Theorem 7.3 the degree of exactness of the n -weight quadratures (7.1) is then at least $2n$ and the property G1 is not satisfied as well.

Furthermore, assuming that n is the smallest index such that $\Delta_{n-1} = 0$, there exists an unique $(n-1)$ -weight quadrature Q_{n-1} of the form (7.1) having degree of exactness at least $2n-3$. However, by Theorem 7.3 it does not satisfy the property

G1 since its degree of exactness is larger than $2n - 3$. In the quasi-definite case the degree of exactness is uniquely determined; see theorems 7.2 and 7.3. With the $(n - 1)$ -weight quadrature Q_{n-1} the situation is different. If we only know the moments m_0, \dots, m_{2n-2} , then we cannot determine the degree of exactness of Q_{n-1} . Indeed, if $Q_{n-1}(x^{2n-1}) \neq m_{2n-1}$, then the degree of exactness is $2n - 2$. However, if $Q_{n-1}(x^{2n-1}) = m_{2n-1}$, then the degree of exactness of Q_{n-1} is at least $2n - 1$, and so on. The following example demonstrates this fact.

Example 2. Consider the linear functional \mathcal{L} from Example 1 in Section 6 defined by a sequence of moments with the first seven terms given by

$$1, 3, 8, 20, 52, 156, i,$$

which is quasi-definite on \mathcal{P}_3 . We saw that the 2-node quadrature (6.2) of degree of exactness 3 does not exist since the zeros of π_2 are $x_1 = x_2 = 2$. Instead of the 2-node quadrature (6.2) we can use a 2-weight quadrature of the form (7.1), i.e., $A_1 f(2) + A_2 f'(2)$. Since $\Delta_1 \neq 0$, by Theorem 7.2 the nonlinear system $A_1 z^j + j A_2 z^{j-1} = m_j$ for monomials $1, z, z^2$ and z^3 , i.e.,

$$\begin{aligned} A_1 \cdot 1 + A_2 \cdot 0 &= 1 \\ A_1 z + A_2 \cdot 1 &= 3 \\ A_1 z^2 + 2A_2(z) &= 8 \\ A_1 z^3 + 3A_2(z^2) &= 20 \end{aligned}$$

has a unique solution (in \mathbb{C}). Indeed, from the first two equations we get $A_1 = 1$ and $A_2 = 3 - z$. Putting the latter in the third equation we get $z^2 - 6z + 8 = 0$, and thus $z = 2$ or $z = 4$. The values $A_1 = 1, A_2 = 1, z = 2$ satisfy the fourth equation, while the values $A_1 = 1, z = 4$ and $A_2 = -1$ do not. Moreover, since $\Delta_2 \neq 0$ by Theorem 7.3 the quadrature $f(2) + f'(2)$ has degree of exactness 3. Its degree of exactness would be higher if and only if $m_4 = 2^4 + 4 \cdot 2^3 = 48$. In this case we would have $\Delta_2 = 0$, i.e., \mathcal{L} would not be quasi-definite on \mathcal{P}_2 . If $m_5 = 2^5 + 5 \cdot 2^4 = 112$, then the quadrature $f(2) + f'(2)$ would have degree of exactness at least 5.

In the Introduction we wondered how far we can go with generalization of the Gauss quadrature as an approximant for an arbitrary linear functional. We suggest that any (generalization of the) Gauss quadrature should have the properties G1–G3. In this sense, the quasi-definiteness of the linear functional represents the *necessary and sufficient condition* for the existence of the Gauss quadrature. The n -weight Gauss Quadrature (7.1) for linear functionals that are quasi-definite on \mathcal{P}_n gives the maximal possible extension of this concept.

Generalization of the Gauss quadrature to complex plane can be linked with many topics and developed in different ways. As pointed by the referees, there is a link with Prony's method for exponential sums, Gauss-Konrod quadrature, etc. Such links cannot be covered within a single paper; we hope to return to them elsewhere.

Acknowledgment. We are grateful to the referees for the insightful reports and for many comments and suggestions that substantially helped in preparing the revised version of the text.

REFERENCES

- [1] N.I. Akhiezer, The Classical Moment Problem and some Related Questions in Analysis, Oliver & Boyd, Edinburgh, 1965.

- [2] B. Beckermann, Complex Jacobi matrices, *J. Comput. Appl. Math.* 127 (2001) 17–65.
- [3] B. Beckermann, V. Kaliaguine, The Diagonal of the Padé Table and the Approximation of the Weyl Function of Second-Order Difference Operators, *Constr. Approx.* 13 (1997) 481–510.
- [4] C. Brezinski, History of Continued Fractions and Padé Approximants, vol. 12 of Springer Series in Computational Mathematics, Springer-Verlag, Berlin, 1991.
- [5] C. Brezinski, Padé-type approximation and general orthogonal polynomials, International series of numerical mathematics, Birkhäuser, 1980.
- [6] C. Brezinski, M. Redivo Zaglia, H. Sadok, Avoiding breakdown and near-breakdown in Lanczos type algorithm, *Numer. Algorithms* 1 (1991) 261–284.
- [7] C. Brezinski, H. Sadok, Avoiding breakdown in the CGS algorithm, *Numer. Algorithms* 1 (1991) 199–206.
- [8] A. Bultheel and M. Van Barel, Linear Algebra, Rational Approximation and Orthogonal Polynomials, vol. 6 of Studies in Computational Mathematics, North-Holland Publishing Co., Amsterdam, 1997.
- [9] L. Chakalov, General quadrature formulae of Gaussian type, *Bulgar. Akad. Nauk Izv. Mat. Inst.* 1 (1954) 67–84.
- [10] T.S. Chihara, An introduction to orthogonal polynomials, Gordon and Breach, New York, 1978.
- [11] E.B. Christoffel, Über die Gaußsche Quadratur und eine Verallgemeinerung derselben, *J. Reine Angew. Math.* 55 (1858) 61–82. Reprinted in: *Gesammelte mathematische Abhandlungen*, vol. 1, B.G. Teubner, Leipzig, 1910, pp. 65–87.
- [12] E.B. Christoffel, Sur une classe particulière de fonctions entières et de fractions continues, *Ann. Mat. Pura Appl.* 8 (1877) 1–10. Reprinted in: *Gesammelte mathematische Abhandlungen*, vol. 2, B.G. Teubner, Leipzig, 1910, pp. 42–50.
- [13] T.C. Chu Moody, G. H. Golub, Inverse Eigenvalue Problems, Theory, Algorithms, and Applications, Oxford University Press, New York, 2005.
- [14] B.D. Craven, Complex symmetric matrices, *J. Aust. Math. Soc.* 10 (1969) 341–354.
- [15] G. Cybenko, Restriction of normal operators, Padé approximation and autoregressive time series, *SIAM Math. Anal.* 15 (1984) 753–767.
- [16] D.Ž. Djoković, Eigenvectors obtained from the adjoint matrix, *Aequat. Math.* 2 (1969) 94–97.
- [17] J. Favard, Sur les polynomes de Tchebicheff, *C. R. Acad. Sci., Paris*, 200 (1935) 2052–2053.
- [18] B. Fischer, Polynomial Based Iteration Methods for Symmetric Linear Systems, Wiley-Teubner Series Advances in Numerical Mathematics, John Wiley & Sons Ltd., Chichester, 1996.
- [19] R.W. Freund, M. Hochbruck, Gauss quadratures associated with the Arnoldi process and the Lanczos algorithm, in: M.S. Moonen, G.H. Golub, B.L.R. De Moor (Eds.), *Linear Algebra for Large Scale and Real-Time Application*, Kluwer, Dordrecht, The Netherlands, 1993, pp. 377–380.
- [20] F.R. Gantmakher, The Theory of Matrices 1,2, Chelsea Publishing Co., New York, 1959.
- [21] C.F. Gauss, Methodus nova integralium valores per approximationem inveniendi, *Commentationes Societatis Regiae Scientiarum Gottingensis* (1814) 39–76. Reprinted in: *Werke*, vol. 3, Göttingen, 1876, pp. 163–196.
- [22] W. Gautschi, Numerical Analysis, 2nd ed., Birkhäuser, Boston, 2012.
- [23] W. Gautschi, Orthogonal Polynomials: Computation and Approximation, Oxford University Press, Oxford, 2004.
- [24] G.H. Golub, J. Kautsky, Calculation of Gauss quadratures with multiple free and fixed knots, *Numer. Math.* 41 (1983) 147–163.
- [25] G.H. Golub, G.A. Meurant, Matrices, Moments, and Quadrature with Applications, Princeton University Press, Princeton, N.J., 2010.
- [26] G.H. Golub, J.H. Welsch, Calculation of Gauss quadrature rules, *Math. Comput.* 23 (1969) 221–230; addendum, *ibid.*, 23 (1969) A1–A10.
- [27] G.H. Golub and C.F. Van Loan, Matrix Computations, Johns Hopkins Studies in the Mathematical Sciences, Johns Hopkins University Press, Baltimore, MD, third ed., 1996.
- [28] R.G. Gordon, Error bounds in equilibrium statistical mechanics, *J. Math. Phys.* 9 (1968) pp. 655–663.
- [29] W.B. Gragg, Matrix interpretations and applications of the continued fraction algorithm, *Rocky Mountain J. Math.* 4 (1974) 213–225.
- [30] W.B. Gragg, The Padé table and its relation to certain algorithms of numerical analysis, *SIAM Review* 14 (1972) 1–62.
- [31] M.H. Gutknecht, A completed theory of the unsymmetric Lanczos process and related algorithms. I, *SIAM J. Matrix Anal. Appl.* 13 (1992) 594–639.
- [32] M.H. Gutknecht, A completed theory of the unsymmetric Lanczos process and related algorithms. II, *SIAM J. Matrix Anal. Appl.* 15 (1994) 15–58.
- [33] E.D. Hellinger, O. Toeplitz, Zur Einordnung der Kattenbruchtheorie in die Theorie der

- quadratischen Formen von unendlichvielen Veränderlichen, *J. Reine Angew. Math.* 144 (1914) 212–238.
- [34] E.D. Hellinger, H.S. Wall, Contributions to the Analytic Theory of Continued Fractions and Infinite Matrices, *Ann. Math.* 44 (1943) 103–127.
- [35] M.R. Hestenes, E. Stiefel, Methods of conjugate gradients for solving linear systems, *J. Research Nat. Bur. Standards* 49 (1952) 409–436.
- [36] N.J. Higham, *Functions of Matrices: Theory and Computation*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2008.
- [37] R.A. Horn, C.R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, UK, 1985.
- [38] A.S. Householder, *The Theory of Matrices in Numerical Analysis*, Dover, New York, 1964.
- [39] C.G.J. Jacobi, Über die Reduction der quadratischen Formen auf die kleinste Anzahl Glieder, *J. Reine Angew. Math.* 39 (1850) 290–292. Reprinted in: *Gesammelte Werke*, vol. 6, Reimer, Berlin, 1891, pp. 318–320.
- [40] C.G.J. Jacobi, Ueber Gauss neue Methode, die Werthe der Integrale näherungsweise zu finden, *J. Reine Angew. Math.* 1 (1826) 301–308. Reprinted in: *Gesammelte Werke*, vol. 6, Reimer, Berlin, 1891, pp. 3–11.
- [41] S. Khrushchev, *Orthogonal Polynomials and Continued Fractions. From Euler’s Point of View*, vol. 122 of *Encyclopedia of Mathematics and its Applications*, Cambridge University Press, Cambridge, 2008.
- [42] P. Kravanja, M. Van Barel, *Computing the zeros of analytic functions. Lecture Notes in Mathematics 1727*. Springer, Berlin, 2000.
- [43] C. Lanczos, An iteration method for the solution of the eigenvalue problem of linear differential and integral operators, *J. Research Nat. Bur. Standards* 45 (1950) 225–282.
- [44] C. Lanczos, Solution of systems of linear equations by minimized iterations, *J. Research Nat. Bur. Standards* 49 (1952) 33–53.
- [45] J. Liesen, Z. Strakoš, *Krylov Subspace Methods: Principles and Analysis*, Oxford University Press, Oxford, 2013.
- [46] L. Lorentzen, H. Waadeland, *Continued fractions with applications*, North-Holland, Amsterdam, 1992.
- [47] J. Málek, Z. Strakoš, *Preconditioning and the Conjugate Gradient Method in the Context of Solving PDEs*, *SIAM Spotlights*, SIAM, Philadelphia, 2015.
- [48] F. Marcelán, R. Álvarez-Nodarse, On the “Favard theorem” and its extension, *J. Comput. Appl. Math.* 127 (2001) 231–254.
- [49] J.R. Magnus, H. Neudecker, *Matrix Differential Calculus with Applications in Statistics and Econometrics*, J. Wiley & Sons Inc., Chichester, 1988.
- [50] G. Meurant, Z. Strakoš, The Lanczos and conjugate gradient algorithms in finite precision arithmetic, *Acta Numer.* 15 (2006) 471–542.
- [51] G.V. Milovanović, A.S. Cvetković, Complex Jacobi matrices and quadrature rules, *Filomat* 17 (2003) 117–134.
- [52] I.P. Mysovskih, On the construction of cubature formulas with the smallest number of nodes, *Dokl. Akad. Nauk SSSR* 178 (1968) 1252–1254.
- [53] B.N. Parlett, Reduction to tridiagonal form and minimal realizations, *SIAM J. Matrix Anal. Appl.* 13 (1992) 567–593.
- [54] B.N. Parlett, *The Symmetric Eigenvalue Problem*, Prentice-Hall, Englewood Cliffs, 1980.
- [55] B.N. Parlett, D.R. Taylor, Z.A. Liu, A look-ahead Lanczos algorithm for unsymmetric matrices, *Math. Comp.* 44 (1985) 105–124.
- [56] T. Popoviciu, Sur une généralisation de la formule d’integration numérique de Gauss, *Acad. R. P. Romine Fil. Iași Stud. Cerc. Ști.* 6 (1955) 29–57.
- [57] Y. Saad, *Iterative Methods for Sparse Linear Systems*, SIAM, Philadelphia, 2003.
- [58] P.E. Saylor, D.C. Smolarski, Why Gauss quadrature in the complex plane?, *Numer. Algorithms* 26 (2001) 251–280.
- [59] N.H. Scott, A Theorem on Isotropic Null Vectors and Its Application to Thermoelasticity, *Proceedings: Mathematical and Physical Sciences* 440 (1993) 431–442.
- [60] B. Simon, CMV matrices: five years after, *J. Comput. Appl. Math.* 208 (2007) 120–154.
- [61] A. Spitzbart, A Generalization of Hermite’s Interpolation Formula, *Amer. Math. Monthly* 67(1) (1960) 42–46.
- [62] D.D. Stancu, A.H. Stroud, Quadrature formulas with simple gaussian nodes and multiple fixed nodes, *Math. Comput.* 17(84) (1963) 384–394.
- [63] T.J. Stieltjes, Recherches sur les fractions continues, *Ann. Fac. Sci. Toulouse Sci. Math. Sci. Phys.* 8 (1894) J. 1–122. Reprinted in: *Oeuvres*, vol. 2, P. Noordhoff, Groningen, 1918, pp. 402–566. English translation, Investigation on continued fractions, in: *Thomas Jan Stieltjes, Collected Papers*, vol. 2, Springer-Verlag, Berlin, 1993, pp. 609–745.

- [64] Z. Strakoš, Model reduction using the Vorobyev moment problem, *Numer. Algorithms* 51 (2009) 363–379.
- [65] A.H. Stroud, D. Secrest, *Gaussian Quadrature Formulas*, Prentice-Hall, Englewood Cliffs (NJ), 1966.
- [66] G. Szegő, *Orthogonal polynomials*, Amer. Math. Soc. Colloquium Publications, 23, New York, 1939.
- [67] Y.V. Vorobyev, *Method of moments in applied mathematics*, Translated from the Russian by Bernard Seckler, Gordon and Breach Science Publishers, New York, 1965.
- [68] H.S. Wall, *Analytic Theory of Continued Fractions*, Chelsea Pub. Co., Bronx, N.y., 1948.
- [69] H. Wilf, *Mathematics for the Physical Sciences*, J. Wiley & Sons Inc., London, 1962.
- [70] J.H. Wilkinson, *The algebraic eigenvalue problem*, Clarendon Press, Oxford, 1965. Reprinted, 1988.