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 approximation of positive definite linear functionals. The underlying theory connects several classical topics including orthogonal polynomials and (real) Jacobi matrices. In this paper we investigate the problem of generalizing the concept of Gauss quadrature for approximation of linear functionals which are not positive definite. We show that the concept can be generalized to quasi-definite functionals and based on a close relationship with orthogonal polynomials and complex Jacobi matrices.

Key words. Quasi-definite linear functional, Gauss quadrature, orthogonal polynomials, complex Jacobi matrices

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

$$\mathcal{L}(x^k) = m_k, \quad k = 0, 1, \dots . \tag{1.1}$$

We say that $\pi_0, \pi_1, \dots, \pi_k$ is the sequence of *formal orthogonal polynomials* if

$$\pi_j \in \mathcal{P}_j \quad \text{and} \quad \mathcal{L}(p\,\pi_j) = 0, \quad \forall p \in \mathcal{P}_{j-1}, \quad \text{for } j = 1, 2, \dots, k,$$

where $\mathcal{P}_j$ is the space of polynomials of degree at most j.

In the classical case (see [17], [33], [8], [9] and [50]) $\mathcal{L}$ is the Riemann, the weighted Riemann or the more general Riemann-Stieltjes integral with respect to a non-decreasing distribution function μ defined on the real axis having finite limits at $\pm\infty$ and infinitely many points of increase. Since μ is of bounded variation, the integral $\int f \mathrm{d}\mu$ exists for every continuous function f. Moreover, under these assumptions, and with $\mathcal{L}(f) = \int f \mathrm{d}\mu$, formal orthogonal polynomials π_j have some additional properties: they exist, they are unique up to a nonzero constant factor, they satisfy three-term recurrence relation, π_j is of degree j and $\mathcal{L}(\pi_j^2) \neq 0$ for $j = 0, 1, \dots$ These properties allow us to approximate the integral $\int f \mathrm{d}\mu$ for every integer n by the unique n-node quadrature that has algebraic degree of exactness $2n-1$, the well-known *Gauss quadrature rule*.

The classical theory of Gauss quadrature can be found in many books; see, for example, [52, Chapters III and XV], [7, Chapter I, Section 6], [19], [18, Chapter 3.2], [36, Section 3.2]. It can be described via orthogonal polynomials or Jacobi matrices which store the coefficients of the associated three-term recurrences and are typically defined as real, tridiagonal, symmetric matrices with positive sub-diagonals (here the last property is not essential; see [56, pp. 335-336]). The n-th Jacobi matrix J_n is

*Department of Mathematics, University of Padova, Via Trieste 63, 35121, Padova, Italy. E-mail: spozza@math.unipd.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 Serbian Ministry of Education and Science (Research Project: “Methods of numerical and nonlinear analysis with applications” (#174002)).

‡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.

thus determined by the first $2n$ moments of the measure μ. Moreover, the $(1,1)$ entry of the i-th power of J_n, $\mathbf{e}_1^T(J_n)^i\,\mathbf{e}_1$, is equal to the moment m_i for $i = 0, \ldots, 2n-1$, which is known as *the moment matching property.*

In this paper we want to see how far we can go with generalization of the Gauss quadrature as an approximant for an arbitrary linear functional. We will show that it is possible to define Gauss quadrature whenever the corresponding sequence of formal orthogonal polynomials satisfies some particular properties that naturally hold in the classical case.

In Section 2 we recall properties of (formal) orthogonal polynomials. As a basic reference we consider, besides the classical monograph by Szegö [52], the beautiful summary of the general theory by Chihara [7]. Section 3 describes the spectral properties of complex Jacobi matrices. In Section 4 we give a proof of the moment matching property for complex Jacobi matrices and quasi-definite linear functionals. Known results generalizing the Gauss quadrature under restrictive assumptions are summarized in Section 5. The results described in Sections 3 and 4 lead to the fully general extension of the *n-node Gauss quadrature rule* in Sections 6 and 7. This rule extends the main properties of the classical Gauss quadrature to the case when the functional $\mathcal{L}$ is quasi-definite. If the linear functional is not quasi-definite, then an analogous generalization can not be established. Demonstration of this fact concludes the paper.

2. Orthogonality and quasi-definite linear functionals. The term *orthogonal polynomials* is usually used for polynomials orthogonal with respect to an inner product (i.e., with respect to an integral with the positive Borel measure). More generally, given the linear functional $\mathcal{L}$, we will use the following definition.

Definition 2.1. *A sequence of polynomials $\pi_0, \pi_1, \ldots, \pi_k$ satisfying the conditions*

1. *deg$(\pi_j) = j$ (π_j is of degree j),*
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 the following:

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

Moreover, $\pi_0(x) \neq 0$ and, providing that they exist, $\pi_n(x)$, $n = 1, 2, \ldots, k$, are uniquely determined up to a nonzero constant factor.

The question of existence of orthogonal polynomials is considered, for example, in [7, Chapter I]; for the case of classical orthogonal polynomials see also Theorem 2.1.1 and pages 24 and 25 of [52]. In the following we will use the Hankel determinants Δ_j of the matrices of moments (see (1.1)),

$$\Delta_j = \begin{vmatrix} m_0 & m_1 & \ldots & m_j \\ m_1 & m_2 & \ldots & m_{j+1} \\ \vdots & \vdots & \ddots & \vdots \\ m_j & m_{j+1} & \ldots & m_{2j} \end{vmatrix}. \tag{2.1}$$

Definition 2.2. *A linear functional $\mathcal{L}$ for which the first $k+1$ Hankel determinants are nonzero, i.e., $\Delta_j \neq 0$ for $j = 0, 1, \ldots, k$, is called quasi-definite on the space of polynomials $\mathcal{P}_k$ of degree at most k.*

In the same spirit we define positive definite functionals (see, e.g., [7, Chapter I, Theorem 3.4]).

DEFINITION 2.3. *The linear functional $\mathcal{L}$ is said to be positive definite on $\mathcal{P}_k$ if $m_s \in \mathbb{R}$ for $s = 0, \ldots, 2k$ and $\Delta_j > 0$ for $j = 0, \ldots, k$.*

The following statements are equivalent (see Chapter I Theorem 3.4 and Chapter II Section 3 in [7]):

- $\mathcal{L}$ is positive definite on $\mathcal{P}_k$.
- $\mathcal{L}(p) > 0$ for every nonzero and nonnegative real polynomial from $\mathcal{P}_k$.
- There exists a positive non-decreasing distribution function μ supported on the real axis such that $\mathcal{L}(p) = \int p(x) \mathrm{d}\mu(x)$ for all p from $\mathcal{P}_{2k}$.

Hence, the classical theory of orthogonal polynomials concerns positive definite linear functionals. For the existence of orthogonal polynomials in the sense of Definition 2.1, the positive definiteness of $\mathcal{L}$ is, however, not needed.

THEOREM 2.4. *[7, Chapter I, Theorem 3.1] 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 $\mathcal{P}_k$.*

It is also shown in [52, Section 3.2] and [7, Section 4] that *monic* orthogonal polynomials (from now on π is always used for monic orthogonal polynomials) with respect to a quasi-definite linear functional have the following properties:

- They satisfy the three-term recurrence relation ([52, Theorem 3.2.1] and [7, Theorem 4.1])

$$\pi_n(x) = (x - \delta_{n-1})\pi_{n-1}(x) - \eta_{n-1}\pi_{n-2}(x), \quad n = 1, 2, \ldots \tag{2.2}$$

  where we set $\eta_0 = m_0$, while the other elements are defined as

$$\delta_{n-1} = \frac{\mathcal{L}(x\pi_{n-1}^2)}{\mathcal{L}(\pi_{n-1}^2)}, \quad \eta_{n-1} = \frac{\mathcal{L}(\pi_{n-1}^2)}{\mathcal{L}(\pi_{n-2}^2)} \neq 0, \quad \pi_{-1}(x) = 0, \, \pi_0(x) = 1;$$

- They satisfy the Christoffel-Darboux identitites ([52, Theorem 3.2.2] and [7, theorem 4.5 and 4.6])

$$\sum_{j=0}^{n} \frac{\pi_j(x)\pi_j(y)}{K_j} = \frac{1}{K_n} \frac{\pi_{n+1}(x)\pi_n(y) - \pi_{n+1}(y)\pi_n(x)}{x - y}, \quad n \geq 0, \tag{2.3}$$

$$\sum_{j=0}^{n} \frac{\pi_j^2(x)}{K_j} = \frac{1}{K_n} [\pi_{n+1}'(x)\pi_n(x) - \pi_{n+1}(x)\pi_n'(x)], \quad n \geq 0, \tag{2.4}$$

  with $K_j = \mathcal{L}(\pi_j^2) = \eta_0\eta_1 \ldots \eta_j$, $j = 0, \ldots, n$.

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.

Normalizing polynomials π_j we can get the sequence of orthonormal polynomials $\tilde{p}_j$. They are unique up to multiplication by (-1), and one particular sequence within the whole family can be expressed as

$$\tilde{p}_j(x) = \frac{\pi_j(x)}{\sqrt{\mathcal{L}(\pi_j^2)}} = \frac{\pi_j(x)}{\sqrt{\eta_0\eta_1 \cdots \eta_j}}, \; j = 0, 1, \ldots, k\,, \tag{2.5}$$

where we take $arg(\sqrt{c}) \in (-\pi/2, \pi/2]$, i.e., consider the principal value of the square root. This means that if there exists a sequence of monic orthogonal polynomials

$\pi_0, \ldots, \pi_k$, then there are 2^{k+1} associated sequences of orthonormal polynomials which differ in executing the complex square roots of the individual coefficients $\eta_0, \ldots, \eta_k$. The three-term recurrence relation for orthonormal polynomials $\tilde{p}_0, \ldots, \tilde{p}_n$, $n \leq k$, can be written as

$$x \begin{bmatrix} \tilde{p}_0(x) \\ \tilde{p}_1(x) \\ \vdots \\ \tilde{p}_{n-1}(x) \end{bmatrix} = J_n \begin{bmatrix} \tilde{p}_0(x) \\ \tilde{p}_1(x) \\ \vdots \\ \tilde{p}_{n-1}(x) \end{bmatrix} + \sqrt{\eta_n} \begin{bmatrix} 0 \\ 0 \\ \vdots \\ \tilde{p}_n(x) \end{bmatrix}, \tag{2.6}$$

where J_n is the (complex) tridiagonal symmetric matrix

$$J_n = \begin{bmatrix} \delta_0 & \sqrt{\eta_1} & & & \\ \sqrt{\eta_1} & \delta_1 & \sqrt{\eta_2} & & \\ & \sqrt{\eta_2} & \delta_2 & \ddots & \\ & & \ddots & \ddots & \sqrt{\eta_{n-1}} \\ & & & \sqrt{\eta_{n-1}} & \delta_{n-1} \end{bmatrix}. \tag{2.7}$$

From (2.6) we see that the zeros λ_i, $i = 1, \ldots, n$, of $\tilde{p}_n$ are the eigenvalues of J_n, with

$$\mathbf{w_i} = [\tilde{p}_0(\lambda_i),\ \tilde{p}_1(\lambda_i),\ \ldots,\ \tilde{p}_{n-1}(\lambda_i)]^T,\ i = 1, \ldots, n, \tag{2.8}$$

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

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

$$\beta_n p_n(x) = (x - \alpha_{n-1}) p_{n-1}(x) - \gamma_{n-1} p_{n-2}(x), \quad \text{for} \quad n = 1, 2, \ldots, \tag{2.9}$$

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

$$\alpha_{n-1} = \frac{\mathcal{L}(x p_{n-1}^2)}{\mathcal{L}(p_{n-1}^2)},\ \beta_n = \frac{\mathcal{L}(x p_{n-1} p_n)}{\mathcal{L}(p_n^2)},\ \gamma_{n-1} = \frac{\mathcal{L}(x p_{n-2} p_{n-1})}{\mathcal{L}(p_{n-2}^2)}, \tag{2.10}$$

(see [52, Theorem 3.2.1], [7, p. 19], [3, Theorem 2.4]). Providing that $p_0, p_1, \ldots, p_n$ exist, all coefficients $\beta_0, \ldots, \beta_n$ and $\gamma_1, \ldots, \gamma_{n-1}$ are different from zero. The recurrences (2.9) can be written in the matrix form

$$x \begin{bmatrix} p_0(x) \\ p_1(x) \\ \vdots \\ p_{n-1}(x) \end{bmatrix} = T_n \begin{bmatrix} p_0(x) \\ p_1(x) \\ \vdots \\ p_{n-1}(x) \end{bmatrix} + \beta_n \begin{bmatrix} 0 \\ 0 \\ \vdots \\ p_n(x) \end{bmatrix}. \tag{2.11}$$

Now, T_n is a tridiagonal complex matrix

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

On the other hand, as shown in [7, Chapter I, Theorem 4.4], in the survey [37, Theorem 2.14] and firstly for the positive definite case by Favard in [13], if we consider any sequence of polynomials satisfying

$$b_n p_n(x) = (x - a_{n-1}) p_{n-1}(x) - c_{n-1} p_{n-2}(x), \ n = 1, 2, \ldots, \tag{2.12}$$

where

$$p_{-1}(x) = 0, \ p_0(x) = c, \ c_0 = 0, \ a_n, b_n, c_n, c \in \mathbb{C}, \ b_n, c_n, c \neq 0,$$

then there exists a quasi-definite linear functional $\mathcal{L}$ such that $p_0, p_1, \ldots,$ are orthogonal polynomials with respect to $\mathcal{L}$. In other words, providing that $c, b_n, c_n \neq 0$, polynomials generated by (2.12) are always orthogonal polynomials. In addition, they are orthonormal if and only if $c_n = b_n$ and p_0 is such that $\mathcal{L}(p_0^2) = 1$.

This also means that for any tridiagonal matrix T_n 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 $\mathcal{L}$. The following proposition clarifies the nonuniqueness of determining T_n from the moments of $\mathcal{L}$.

PROPOSITION 2.5. *Let T_n and $\widehat{T}_n$ be two tridiagonal matrices without zero components on the sub- and super-diagonal. Then, T_n and $\widehat{T}_n$ are determined by the first $2n$ moments of the same linear functional which is quasi-definite on $\mathcal{P}_{n-1}$ if and only if T_n and $\widehat{T}_n$ are similar matrices such that $T_n = D^{-1}\widehat{T}_n D$, where D is an invertible diagonal matrix.*

Proof. The proof uses formula (2.11) and the observation that two sets of polynomials $p_0, \ldots, p_{n-1}$ and $\hat{p}_0, \ldots, \hat{p}_{n-1}$ are orthogonal with respect to the same linear functional if and only if

$$\begin{bmatrix} \hat{p}_0(x) \\ \vdots \\ \hat{p}_{n-1}(x) \end{bmatrix} = D \begin{bmatrix} p_0(x) \\ \vdots \\ p_{n-1}(x) \end{bmatrix},$$

where D is an invertible diagonal matrix.

We first assume that T_n and $\widehat{T}_n$ are two matrices determined by the same moments of the linear functional $\mathcal{L}$ quasi-definite on $\mathcal{P}_{n-1}$. The matrices T_n and $\widehat{T}_n$ determine two sequences of orthogonal polynomials that we name respectively $p_0, \ldots, p_{n-1}$ and $\hat{p}_0, \ldots, \hat{p}_{n-1}$. Using the recurrence relation (2.9) we can define the polynomial

$$q_n = (x - \alpha_{n-1}) p_{n-1} - \gamma_{n-1} p_{n-2}, \tag{2.13}$$

and analogously the polynomial $\hat{q}_n$. The recurrence relation (2.11) for the polynomials $\hat{p}_0, \ldots, \hat{p}_{n-1}$ and $\hat{q}_n$ then gives

$$xD \begin{bmatrix} p_0(x) \\ \vdots \\ p_{n-1}(x) \end{bmatrix} = \widehat{T}_n D \begin{bmatrix} p_0(x) \\ \vdots \\ p_{n-1}(x) \end{bmatrix} + \begin{bmatrix} 0 \\ \vdots \\ \hat{q}_n(x) \end{bmatrix}. \tag{2.14}$$

Hence, we obtain that $T_n = D^{-1}\widehat{T}_n D$ and $q_n = \hat{q}_n / d_n$, with d_n the last diagonal element of D.

Vice versa, putting $T_n = D^{-1}\widehat{T}_n D$ in (2.11) and multiplying from the left by D we get (2.14) which means that we obtain two sequences of orthogonal polynomials such that $[\hat{p}_0, \ldots, \hat{p}_{n-1}]^T = D[p_0, ..., p_{n-1}]^T$. $\square$

REMARK 2.6. *Using (2.13) and (2.11) we get that* q_n *has degree* n *and its zeros are the eigenvalues of* T_n *(analogously to the positive definite linear functional case in [36, Sections 3.2.1 and 3.4.1]). Moreover,* q_n *is orthogonal to* $\mathcal{P}_{n-1}$.
In the following the elements of $\widehat{T}_n$ are marked with a hat.

COROLLARY 2.7. *Let* T_n *and* $\widehat{T}_n$ *be two tridiagonal matrices without zero components on the sub- and super-diagonals.* T_n *and* $\widehat{T}_n$ *are determined by the first* $2n$ *moments of the same linear functional if and only if*

- $\alpha_i = \hat{\alpha}_i$ *for* $i = 0, \ldots, n-1$;
- $\beta_i\gamma_i = \hat{\beta}_i\hat{\gamma}_i$ *for* $i = 1, \ldots, n-1$.

Proof. By Proposition 2.5 we know that T_n and $\widehat{T}_n$ are determined by the first $2n$ moments of the same linear functional if and only if $T_n = D^{-1}\widehat{T}_nD$, with $D = \mathrm{diag}(d_1, \ldots, d_n)$ an invertible diagonal matrix. We first assume that $T_n = D^{-1}\widehat{T}_nD$. Comparing the corresponding entries of matrices T_n and $D^{-1}\widehat{T}_nD$ we get $\alpha_i = \hat{\alpha}_i$, for $i = 0, \ldots, n-1$, as well as $\gamma_i = (d_i/d_{i+1})\hat{\gamma}_i$ and $\beta_i = (d_{i+1}/d_i)\hat{\beta}_i$ for $i = 1, \ldots, n-1$. Thus we see that $\gamma_i\beta_i = \hat{\gamma}_i\hat{\beta}_i$ for $i = 1, \ldots, n-1$.

Vice versa, if $\alpha_i = \hat{\alpha}_i$ for $i = 0, \ldots, n-1$ and $\beta_i\gamma_i = \hat{\beta}_i\hat{\gamma}_i$ for $i = 1, \ldots, n-1$, then the diagonal matrix $D = \mathrm{diag}(d_1, \ldots, d_n)$ whose elements are $d_1 = 1$ and

$$d_i = \frac{\beta_1\beta_2\cdots\beta_{i-1}}{\hat{\beta}_1\hat{\beta}_2\cdots\hat{\beta}_{i-1}} = \frac{\hat{\gamma}_1\hat{\gamma}_2\cdots\hat{\gamma}_{i-1}}{\gamma_1\gamma_2\cdots\gamma_{i-1}}, \qquad \text{for } i = 2, \ldots, n,$$

gives $T_n = D^{-1}\widehat{T}_nD$. □

In addition, every tridiagonal matrix T_n with nonzero entries on sub- and super-diagonal is similar to a *complex tridiagonal symmetric matrix* J_n. Indeed, the diagonal matrix $D = \mathrm{diag}(d_1, \ldots, d_n)$ whose elements are

$$d_1 = 1, \quad d_i = \left(\frac{\gamma_1\gamma_2\ldots\gamma_{i-1}}{\beta_1\beta_2\ldots\beta_{i-1}}\right)^{1/2}, \qquad \text{for } i = 2, \ldots, n, \tag{2.15}$$

provides the desired similarity transformation. This result is well known in the case of positive definite linear functional, for which J_n is a real tridiagonal symmetric matrix ([56, pp. 335-336]). However, we remark that in the case of quasi-definite linear functionals the matrix J_n is, in general, complex. Hence J_n is symmetric but it may not be a Hermitian matrix.

Finally, Corollary 2.7 implies (analogously to the nonuniqueness of the sequences of orthonormal polynomials mentioned above) that there exist 2^{n-1} different tridiagonal symmetric matrices J_n determined by the moments $m_0, \ldots, m_{2n-1}$. In fact, two symmetric tridiagonal matrices J_n and $\widehat{J}_n$ with nonzero entries on the sub-diagonal (or super-diagonal) are determined by the first $2n$ moments of a linear functional if and only if they have the same diagonal and $\beta_i = \pm\hat{\beta}_i$ for $i = 1, \ldots, n-1$.

3. Complex generalization of Jacobi matrices. For the positive definite linear functionals the matrix J_n given above is often called Jacobi matrix. Since for quasi-definite functionals J_n is complex symmetric but generally not Hermitian, we find useful to recall some classical results on Jacobi matrices and investigate its complex generalization.

In literature one can find several different definitions of Jacobi matrices. Most frequently, Jacobi matrix is defined as a real, symmetric, tridiagonal matrix with positive elements on the super-diagonal ([1, p. 2], [10, p. 72], [21, p. 13], [36, 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 [32] 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 [26]. 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 [50]. A review of the general theory of the unitary analogue of Jacobi matrices, the CMV matrices, can be found in [47]. For a detailed history of Jacobi matrices we refer to [36, Section 3.4.3].

Other definitions of Jacobi matrices can be found in [16, Vol. 2, p. 99] (a real tridiagonal matrix), [31, 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 nonnegative), [27, p. 103] (a tridiagonal symmetric matrix with a complex diagonal and with nonzero real elements on the sub- and super-diagonal). In this paper we use the definition by Beckermann from the paper about spectral properties of *complex Jacobi matrices* [2].

DEFINITION 3.1. *A square complex matrix is called* Jacobi matrix *if it is tridiagonal, symmetric and has no zero elements on its sub- and super-diagonal.*

Probably the first study of a class of this kind of matrices appeared in [54, 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.

3.1. Real Jacobi matrices. Here we summarize some well-known properties of real Jacobi matrices. Any real $n \times n$ Jacobi matrix J_n can be orthogonally diagonalized, i.e.,

$$J_n W = W \operatorname{diag}(\lambda_1, \ldots, \lambda_n),$$

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

THEOREM 3.2 (see, e.g., [36, Theorem 3.4.3 on p. 114]). *The following properties stand 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 $J_1, \ldots, J_n$ be Jacobi matrices such that J_i is the leading principal $i \times i$ submatrix of J_n for $i = 1, \ldots, n-1$. From now on $J_1, \ldots, J_n$ will always denote the described sequence of Jacobi matrices.

THEOREM 3.3 (Interlacing Property, [36, Theorem 3.3.1, p. 92 and Remark 3.4.4, p. 115]). *Let $J_1, \ldots, J_n$ be real Jacobi matrices as described above. Let k and ℓ be integers with $k+1 \leq \ell \leq n$ and let $\lambda_i^{(k)}$ for $i = 1, \ldots, k$ be the eigenvalues of J_k. Then at least one of the eigenvalues of J_ℓ is contained in any of the $k+1$ open intervals*

$$\left(-\infty, \lambda_1^{(k)}\right), \left(\lambda_1^{(k)}, \lambda_2^{(k)}\right), \ldots, \left(\lambda_{k-1}^{(k)}, \lambda_k^{(k)}\right), \left(\lambda_k^{(k)}, +\infty\right).$$

As a trivial consequence we get the strict interlacing property for the eigenvalues of two subsequent Jacobi matrices J_k and J_{k+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 [34, 35] (for a description of the method and its properties we refer, for example, to [14, Section 4], [21, Section 4.1], [36, Section 2.4.1] and [39]). 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 a $n \times n$ 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., [14, 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 [28, Section 14]. The columns of V_n form a basis for the Krylov subspace

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

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.

In the following subsections we will discuss possible extensions of the properties given above to complex Jacobi matrices. In particular, we will describe some properties of the associated Jordan canonical form.

3.2. Complex tridiagonal matrices. The first statement of Theorem 3.2 is reduced in the following straightforward way.

THEOREM 3.4. *Every tridiagonal matrix $T \in \mathbb{C}^{n\times n}$ with nonzero elements on its super-diagonal (or sub-diagonal) is non-derogatory, i.e., its eigenvalues have geometric multiplicity* 1.

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

COROLLARY 3.5. *Every tridiagonal matrix $T \in \mathbb{C}^{n\times n}$ with nonzero elements on its super-diagonal (or sub-diagonal) is diagonalizable if and only if it has distinct eigenvalues.*

It is well known that the eigenvectors corresponding to the eigenvalues of the geometric multiplicity one can be expressed using the adjoint matrix (sometimes it is used the term adjugate to avoid confusion with the Hermitian adjoint). Indeed, if λ is an eigenvalue with geometric multiplicity one, then $\mathrm{rank}(A - \lambda I) = n - 1$. This implies that $\mathrm{adj}(A - \lambda I)$ is not identically zero, i.e., it has at least one nonzero column, say $\mathrm{adj}(A - \lambda I)\mathbf{e}_i$. Using (for later convenience we consider any $\xi \in \mathbb{C}$)

$$(A - \xi I)\,\mathrm{adj}(A - \xi I) = \det(A - \xi I)\, I,$$

the i-th column gives

$$(A - \xi I)\mathbf{z}(\xi) = \det(A - \xi I)\mathbf{e}_i,$$

where $\mathbf{z}(\xi) = \mathrm{adj}(A - \xi I)\mathbf{e}_i$. For $\xi = \lambda$ we get $(A - \lambda I)\mathbf{z}(\lambda) = \mathbf{0}$ which finishes the argument. Please notice that the same eigenvector (apart from the normalization) is given by any nonzero column of $\mathrm{adj}(A - \lambda I)$.

Now, as done in [12], differentiating j times $(A - \xi I)\mathbf{z}(\xi) = \det(A - \xi I)\mathbf{e}_i$, we obtain

$$(A - \xi I)\mathbf{z}^{(j)}(\xi) = j\mathbf{z}^{(j-1)}(\xi) + \frac{d^j}{d\xi^j}\det(A - \xi I)\mathbf{e}_i.$$

Denoting

$$\mathbf{w}_0(\xi) = \mathbf{0}, \ \ \mathbf{w}_1(\xi) = \mathbf{z}(\xi), \ \ \mathbf{w}_{j+1}(\xi) = \frac{1}{j}\mathbf{w}_j'(\xi) = \frac{1}{j!}\mathbf{z}^{(j)}(\xi), \ j = 1, 2, \ldots, \qquad (3.1)$$

we get

$$(A - \xi I)\mathbf{w}_{j+1}(\xi) = \mathbf{w}_j(\xi) + \frac{1}{j!}\frac{d^j}{d\xi^j}\det(A - \xi I)\mathbf{e}_i, \text{ where } j = 0, 1, \ldots.$$

If λ is an eigenvalue with geometric multiplicity 1 and algebraic multiplicity s, then this gives

$$(A - \lambda I)\mathbf{w}_{j+1}(\lambda) = \mathbf{w}_j(\lambda) \text{ for } j = 0, \ldots, s - 1. \qquad (3.2)$$

Therefore $\mathbf{w}_1(\lambda)$ is the eigenvector and $\mathbf{w}_j(\lambda)$ for $j = 2, \ldots, s$ are the *generalized eigenvectors* of A (Jordan canonical vectors of A) corresponding to λ. Since $\mathbf{w}_1(\lambda) \neq \mathbf{0}$, (3.2) implies that $\mathbf{w}_2(\lambda), \ldots, \mathbf{w}_s(\lambda)$ are also nonzero vectors.

Now, let $\mathbf{z}(\xi) = \mathrm{adj}(T_n - \xi I)\mathbf{e}_n$, where T_n is a tridiagonal matrix of dimension $n \times n$. Direct computation shows that

$$\mathbf{z}(\xi) = \begin{bmatrix} \beta_1 \cdots \beta_{n-1} \\ -\beta_2 \cdots \beta_{n-1}\phi_1(\xi) \\ \vdots \\ (-1)^{n-2}\beta_{n-1}\phi_{n-2}(\xi) \\ (-1)^{n-1}\phi_{n-1}(\xi) \end{bmatrix}, \qquad (3.3)$$

where $\beta_1, \ldots, \beta_{n-1}$ are the elements of the super-diagonal and $\phi_i(\xi) = \det(T_i - \xi I)$, with T_i the i-th leading principal submatrix of T_n. This result was shown for Hermitian tridiagonal matrices by Wilkinson in [56, Chapter 5, Section 48]. Providing that T_n has no zeros on its super- and sub-diagonal, $\phi_i(\xi) = (-1)^i\pi_i(\xi)$, $i = 1, \ldots, n$, where $\pi_1, \ldots, \pi_n$ is the sequence of monic orthogonal polynomials corresponding to T_n. The following property was presented in the lecture of Ipsen at the ILAS 2005 conference.

Proposition 3.6. *Let $T_n \in \mathbb{C}^{n\times n}$ be a tridiagonal matrix with nonzero elements on its super-diagonal. Let λ be an eigenvalue of algebraic multiplicity s and $\mathbf{w}_{j+1}(\lambda)$, for $j = 1, \ldots, s - 1$, the corresponding generalized eigenvectors satisfying $(T_n - \lambda I)\mathbf{w}_{j+1}(\lambda) = \mathbf{w}_j(\lambda)$ (with $\mathbf{w}_0 = \mathbf{0}$, $\mathbf{w}_1 = \mathbf{z}(\lambda)$ from (3.3)). Then we can give* *the following explicit formulation*

$$\mathbf{w}_j(\lambda) = \frac{1}{(j-1)!}\begin{bmatrix} \mathbf{0}_{j-1} \\ \beta_j \cdots \beta_{n-1} \\ (-1)^j \beta_{j+1} \cdots \beta_{n-1} \phi_j^{(j-1)}(\lambda) \\ \vdots \\ (-1)^{n-2} \beta_{n-1} \phi_{n-2}^{(j-1)}(\lambda) \\ (-1)^{n-1} \phi_{n-1}^{(j-1)}(\lambda) \end{bmatrix}, \quad j = 2, \ldots, s,$$

where $\mathbf{0}_\ell$ is the zero vector of length ℓ, $\beta_1, \ldots, \beta_{n-1}$ are the elements of the super-diagonal of T_n and $\phi_i(\lambda) = det(T_i - \lambda I)$, with T_i the i-th leading principal submatrix of T_n.

Proof. Since $\beta_1, \beta_2, \ldots, \beta_{n-1} \neq 0$, every eigenvector of T_n corresponding to the eigenvalue λ can be expressed as a nonzero multiple of $\mathbf{z}(\lambda)$ from (3.3). Using (3.1) we obtain the form of $\mathbf{w}_j(\lambda)$ in the statement. □

As for the generalization of the second statement of Theorem 3.2, we can give the following results. The formula (3.3) 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 $\phi_{n-1}(\lambda) \neq 0$, i.e., that the eigenvalues of T_n and T_{n-1} are distinct. Using the standard argument, if λ is a root of both the orthogonal polynomials ϕ_n and ϕ_{n-1}, then by (2.9) it is also a root of ϕ_{n-2}. Hence, by induction, $\phi_0 = 0$, that is a contradiction.

3.3. 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 [11, Theorem 3] Craven proved the following theorem.

THEOREM 3.7. *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^TAV is a diagonal matrix;*
2. *Every eigenspace of A has a basis $\mathbf{v}_1, \ldots, \mathbf{v}_s$ without isotropic vectors and such that $\mathbf{v}_i^T\mathbf{v}_j = 0$ for $i \neq j$.*

Moreover, [46, theorems 1 and 3] give the following equivalence.

THEOREM 3.8. *A singular symmetric matrix contains an isotropic vector in its null space if and only if the trace of its adjugate vanishes.*

We will use it in the following lemma.

LEMMA 3.9. *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

$$\frac{d}{d\xi}\mathrm{det}A(\xi) = \mathrm{tr}(\mathrm{adj}(A(\xi)\frac{dA(\xi)}{d\xi}));$$

for a proof see, e.g., [38, Theorem 1 at p. 149]. If $A(\xi) = \xi I - J$, then the previous formula becomes

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

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

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

Since $\phi'(\lambda) = 0$ if and only if the algebraic multiplicity of λ is greater than 1, by Theorem 3.8 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 3.4 any complex Jacobi matrix is non-derogatory, the proof is finished. □

We will summarize the situation in the following proposition.

PROPOSITION 3.10. *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 $V^T JV$ is a diagonal matrix;*
3. *None of the eigenvectors of J is isotropic.*

Proof. The second and the third properties are equivalent by Theorem 3.7. 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 3.9 finishes the proof. □

In the rest of this section we recall the relationship between Jacobi matrices and non-Hermitian Lanczos algorithm (for details we refer to [3, Section 2.7.2], [21, Section 4.2], [36, Section 2.4.2] and [44, 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 2. 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 symmetric and therefore a Jacobi matrix; see (2.15).

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 [36, p. 33], with the detailed exposition presented in [42, 4, 5, 24, 40, 25].

4. 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., [21, Chapter 7] or [36, Section 3.5]). Then

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

where J_n is the Jacobi matrix associated with the first n steps of the Lanczos process. Using the Vorobyev method of moments [53, 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 [51]. Here we will prove an analogous property for Jacobi matrices determined by quasi-definite linear functionals.

THEOREM 4.1. *[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 (2.6). Then*

$$\mathcal{L}(x^i) = m_0\, \mathbf{e}_1^T (J_n)^i\, \mathbf{e}_1, \quad i = 0, \ldots, 2n-1, \tag{4.1}$$

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

We will give the proof of this theorem using two lemmas.

LEMMA 4.2. *The polynomials $p_0, \ldots, p_{n-1}$ associated with the three-term recurrence relation whose coefficients are given by the Jacobi matrix J_n are orthonormal with respect to the functional $\widetilde{\mathcal{L}}$ defined by*

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

with $m_0 = 1/p_0^2$.

Proof. Let J_n be the Jacobi matrix associated with the polynomials $p_0, \ldots, p_{n-1}$. We note that for $i = 0, \ldots, n-1$ the $(i+1)$-st entry of the vector $(J_n)^i \mathbf{e}_1$ is nonzero. Moreover, for $i = 0, \ldots, n-2$, the entries $i+2, \ldots, n$ of $(J_n)^i \mathbf{e}_1$ are zero. Hence the canonical basis $\mathbf{e}_1, \ldots, \mathbf{e}_k$ is an orthonormal basis of $\mathcal{K}_k(J_n, \mathbf{e}_1)$, $k = 1, \ldots, n$, 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 $\hat{p}_{k-1} = \tilde{p}_{k-1}/\sqrt{m_0}$, $k = 1, \ldots, n$, are orthonormal with respect to $\widetilde{\mathcal{L}}$. Indeed,

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

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 $k = 1, \ldots, n-1$. Notice that

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

Therefore, by (2.6) we see that the coefficients from the three-term recurrence relation for $x\hat{p}_0, \ldots, x\hat{p}_{n-1}$ are the same as those for $xp_0, \ldots, xp_{n-1}$. □

The following lemma gives the remaining part of the proof of Theorem 4.1.

LEMMA 4.3. *Let $\mathcal{L}$ and $\widetilde{\mathcal{L}}$ be linear functionals such that there exists a sequence of polynomials p_i for $i = 0, \ldots, n-1$ that are orthogonal with respect to both $\mathcal{L}$ and $\widetilde{\mathcal{L}}$. Let $\mathcal{L}(x^0) = \widetilde{\mathcal{L}}(x^0)$. Then*

$$\mathcal{L}(x^i) = \widetilde{\mathcal{L}}(x^i) \text{ for } i = 0, \ldots, 2n-1. \tag{4.2}$$

Proof. We prove it by induction. Using (2.10),

$$\frac{\mathcal{L}(xp_0^2(x))}{\mathcal{L}(p_0^2(x))} = \alpha_0 = \frac{\widetilde{\mathcal{L}}(xp_0^2(x))}{\widetilde{\mathcal{L}}(p_0^2(x))}, \qquad \text{i.e.,} \qquad \frac{m_1}{m_0} = \frac{\widetilde{m}_1}{\widetilde{m}_0}.$$

Since we have assumed $m_0 = \widetilde{m}_0$, we conclude $m_1 = \widetilde{m}_1$. Let $m_i = \widetilde{m}_i$ for $i = 0, \ldots, 2k-3$. Using (2.10) we have

$$\frac{\mathcal{L}(xp_{k-1}(x)p_{k-2}(x))}{\mathcal{L}(p_{k-2}^2(x))} = \gamma_{k-1} = \frac{\widetilde{\mathcal{L}}(xp_{k-1}(x)p_{k-2}(x))}{\widetilde{\mathcal{L}}(p_{k-2}^2(x))}.$$

Rewriting

$$xp_{k-1}(x)p_{k-2}(x) = \sum_{i=0}^{2k-2} a_i x^i \quad \text{and} \quad p_{k-2}^2(x) = \sum_{i=0}^{2k-4} b_i x^i,$$

the induction assumptions gives $m_{2k-2} = \widetilde{m}_{2k-2}$. Repeating the same argument with the coefficient α_{k-1} finishes the proof. □

A different normalization of the orthogonal polynomials is associated with a tridiagonal matrix T_n such that $J_n = D^{-1}T_nD$, with D the diagonal matrix with elements given in (2.15) (see also Proposition 2.5). Hence, the statement of Theorem 4.1 remains valid for any tridiagonal matrix T_n associated with a sequence of orthogonal polynomials defined by the functional $\mathcal{L}$.

5. Quasi-definite linear functionals and Gauss quadrature under restrictive assumptions. *Positive definite* linear functionals naturally lead to the well-known Gauss quadrature. Let us recall some of its basic properties:

- G1: The n-node Gauss quadrature attains the maximal algebraic degree of exactness $2n-1$.
- G2: The n-node Gauss quadrature is well-defined and it is unique. Naturally, the Gauss quadratures with a smaller number of nodes also exist and they are unique.
- G3: The Gauss quadrature of the 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 $\mathcal{L}$; $m_0 = \mathcal{L}(x^0)$.

Since the degree of exactness is larger than $n-1$, the Gauss quadrature is interpolatory quadrature, i.e., the weights ω_i satisfy

$$\omega_i = \mathcal{L}(\ell_i), \ i = 1, \dots, n, \tag{5.1}$$

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 given as the squared first entries of the associated normalized eigenvectors of J_n multiplied by m_0; see, e.g., [55, Sections 2.5 and 2.9] and [22].

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

$$\mathcal{L}(f) = \sum_{i=1}^{n} \omega_i f(\lambda_i) + R_n(f), \tag{5.2}$$

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

Theorem 5.1. *The quadrature (5.2) is exact for every f from $\mathcal{P}_{2n-1}$ if and only if it is interpolatory and the polynomial*

$$\varphi_n(x) = \prod_{i=1}^{n}(x-\lambda_i) \tag{5.3}$$

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

Proof. Assume that (5.2) 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 $i = 1, ..., n$,

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

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 $i = 1, \ldots, n$ completes the proof. □

Theorem 5.1 can be interpreted in the following way. Let the monic polynomial φ_n of degree n that is orthogonal to the space $\mathcal{P}_{n-1}$ with respect to the linear functional $\mathcal{L}$ quasi-definite on $\mathcal{P}_n$ has n distinct roots $\lambda_1, \ldots, \lambda_n$. Then the interpolatory quadrature with the nodes $\lambda_1, \ldots, \lambda_n$ has algebraic degree of exactness at least $2n-1$. So, the quadrature rule (5.2) has the properties G1 and G2 if and only if the following conditions simultaneously hold:

1. There exists a sequence of orthogonal polynomials $p_0, \ldots, p_n$ 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, $j = 1, \ldots, n$, in the sequence are distinct; i.e., the matrices J_j, $j = 1, \ldots, n$, are diagonalizable.

The standard argument then shows that the quadrature rule (5.2) 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 [19, p. 153], [45, p. 267-268]. Quadrature (5.2) was considered (to our knowledge) for the first time by Gragg in [23] for real valued linear functionals. A generalization for complex valued functionals was considered by Saylor and Smolarski in [45]. 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 $p_1, \ldots, p_k$ *can have multiple zeros.* Hence it can happen that for some values ℓ, $\ell \le k$, 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 (5.2) 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 (5.2) 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.

6. 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 (5.2). In particular, we will consider the *n-weight quadrature formula*

$$\mathcal{L}(f) = \sum_{i=1}^{\ell} \sum_{j=0}^{s_i-1} \omega_{i,j}\, f^{(j)}(\lambda_i) + R_n(f), \tag{6.1}$$

where $n = s_1 + \ldots + s_\ell$. Note that the rule (5.2) is the special case of the rule (6.1) when $\ell = n$ and $s_1 = \ldots = s_n = 1$. So we generalize the rule (5.2) in the way that incorporates, in addition to the functional values $f(\lambda_1), \ldots, f(\lambda_\ell)$, also the values of the derivatives of f at the points $\lambda_1, \ldots, \lambda_\ell$. The generalization (6.1) 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 $s_1, \ldots, s_\ell$ when we want to achieve the maximal degree of exactness.

THEOREM 6.1. *Let $\mathcal{L}$ be an arbitrary linear functional on $\mathcal{P}$. The quadrature (6.1) 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*

$$\varphi_n(x) = (x - \lambda_1)^{s_1}(x - \lambda_2)^{s_2} \ldots (x - \lambda_\ell)^{s_\ell} \tag{6.2}$$

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

Proof. Following [49] we consider for each root λ_i, $i = 1, \ldots, \ell$, of φ_n the following s_i polynomials of degree $n - 1$

$$h_{i,j}(x) = \frac{(x-\lambda_i)^j}{j!} \left\{ \sum_{\nu=0}^{s_i-1-j} \frac{(x-\lambda_i)^\nu}{\nu!} \left(\frac{1}{g_i(x)} \right)^{(\nu)} \Bigg|_{x=\lambda_i} \right\} g_i(x), \quad j = 0, 1, \ldots, s_i - 1, \tag{6.3}$$

where $g_i(x) = \prod_{\substack{t=1 \\ t \neq i}}^{\ell} (x - \lambda_t)^{s_t}$. As proved in [48, Section 3], from (6.3) 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 $k = 1, 2, \ldots, \ell$, and $t = 0, 1, \ldots, s_i - 1$. Defining the generalized (Hermite) interpolating polynomial (see [48])

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

we get that the formula (6.1) 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 (6.1) are given by

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

The rest of the proof is fully analogous to the proof of Theorem 5.1. □

We say that the n-weight quadrature (6.1) is unique if the nodes and the weights are uniquely determined by $\mathcal{L}$ and n.

THEOREM 6.2. *Let $\mathcal{L}$ be an arbitrary linear functional on $\mathcal{P}$. The n-weight quadrature (6.1) 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 6.1 says that the n-weight interpolatory quadrature (6.1) is of degree of exactness at least $2n-1$ if and only if the monic polynomial

$$\varphi_n(x) = x^n + c_{n-1}x^{n-1} + \ldots + c_1 x + c_0$$

given by (6.2) is orthogonal to the space $\mathcal{P}_{n-1}$. The conditions $\mathcal{L}(x^j\varphi_n) = 0$, $j = 0, \ldots, n-1$, are satisfied if and only if the linear system

$$\begin{bmatrix} m_0 & m_1 & \ldots & m_{n-1} \\ m_1 & m_2 & \ldots & m_n \\ \vdots & \vdots & \ddots & \vdots \\ m_{n-1} & m_n & \ldots & m_{2n-2} \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_{n-1} \end{bmatrix} = \begin{bmatrix} -m_n \\ -m_{n+1} \\ \vdots \\ -m_{2n-1} \end{bmatrix} \tag{6.4}$$

has a unique solution, which gives the statement. □

Finally, the following theorem gives the condition under which the degree of exactness of (6.1) 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 6.3. *Let $\mathcal{L}$ be an arbitrary linear functional on $\mathcal{P}$ and let the n-weight quadrature (6.1) has degree of exactness at least $2n-1$. Then the degree of exactness of the quadrature (6.1) 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 (6.1) has degree of exactness at least $2n-1$, the polynomial φ_n given by (6.2) 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 (6.1) is at least $2n$. Thus we conclude that the quadrature (6.1) has degree of exactness larger than $2n-1$ if and only if $\mathcal{L}(\varphi_n x^j) = 0$ for $j = 0, \ldots, n$, i.e., if and only if there is a vector $[c_0, \ldots, c_{n-1}, 1]^T$ such that

$$\begin{bmatrix} m_0 & m_1 & \ldots & m_n \\ m_1 & m_2 & \ldots & m_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ m_n & m_{n+1} & \ldots & m_{2n} \end{bmatrix} \begin{bmatrix} c_0 \\ \vdots \\ c_{n-1} \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \tag{6.5}$$

in which case $\Delta_n = 0$. ∎

COROLLARY 6.4. *The quadrature rule (6.1) has the properties G1 and G2 if and only if $\mathcal{L}$ is quasi-definite on $\mathcal{P}_n$.*

Proof. The n-weight quadrature (6.1) is unique and of 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 $\mathcal{P}_n$. ∎

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 exists $f^{(j)}(\lambda_i)$ 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 (see [29]). 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

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

Denoting

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

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

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

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 3.4 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

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

Using (3.3) and Proposition 3.6, 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

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

Hence we have

$$\mathbf{e}_1^T f(J_n)\,\mathbf{e}_1 = \sum_{i=1}^{\ell}\sum_{j=0}^{s_i-1} \frac{w_{i,0}\hat{w}_{i,j}}{j!}\, f^{(j)}(\lambda_i) = \sum_{i=1}^{\ell}\sum_{j=0}^{s_i-1} \tilde{\omega}_{i,j}\, f^{(j)}(\lambda_i), \tag{6.6}$$

with

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

Using $\omega_{i,j} = m_0\tilde{\omega}_{i,j}$ in (6.6) we get

$$m_0\mathbf{e}_1^T f(J_n)\,\mathbf{e}_1 = \sum_{i=1}^{\ell}\sum_{j=0}^{s_i-1} \omega_{i,j}\, f^{(j)}(\lambda_i). \tag{6.7}$$

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

Corollary 6.5. *The quadrature rule (6.1) having the properties G1 and G2 satisfies also the property G3.*

Proof. The right-hand side of (6.7) is of the form (6.1). 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 (6.3); see the proof of Theorem 6.1. Since the quadrature (6.1) satisfies the properties G1 and G2, by Corollary 6.4 the functional $\mathcal{L}$ is quasi-definite on $\mathcal{P}_n$. Using Theorem 4.1, its values on monomials x^i must then be equal for $i = 0, 1, \ldots, 2n-1$ to the right-hand side of (6.7) with $f(\lambda)$ replaced by the same monomials. Consequently, the right-hand side of (6.7) represents a quadrature with the algebraic degree at least $2n-1$. From uniqueness it must be equal to the quadrature (6.1) with the weights $\mathcal{L}(h_{i,j})$ and the proof is finished. ☐

We will summarize the characterization of the n-weight quadrature formula (6.1) in terms of the associated Jacobi matrix as a theorem.

Theorem 6.6. *Let $\mathcal{L}$ be an arbitrary linear functional on $\mathcal{P}$ and $m_0 = \mathcal{L}(x^0)$. There exists a Jacobi matrix J_n of dimension n such that*

$$\mathcal{L}(x^i) = m_0\,\mathbf{e}_1^T (J_n)^i\,\mathbf{e}_1, \qquad \textit{for } i = 0,\ldots,2n-1, \tag{6.8}$$

$$\mathcal{L}(x^{2n}) \neq m_0\,\mathbf{e}_1^T (J_n)^{2n}\,\mathbf{e}_1, \tag{6.9}$$

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

Proof. Let J_n be the Jacobi matrix satisfying (6.8) and (6.9). Then, by (6.7) there exists the n-weight quadrature (6.1) whose degree of exactness is (exactly) $2n-1$. By Theorem 6.3 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 $\mathcal{P}_{n-1}$. By Lemma 4.2 the polynomials $p_0,\ldots,p_{n-1}$ associated with the three-term recurrence relation whose coefficients are given by J_n are orthonormal with respect to the linear functional

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

They are also orthonormal with respect to $\mathcal{L}$ by (6.8), which means, using Theorem 2.4, that $\mathcal{L}$ is quasi-definite on $\mathcal{P}_{n-1}$. The statement in the opposite direction follows directly by corollaries 6.4 and 6.5. ☐

The presented construction (6.1) and the statements proved throughout this section show that it is possible to construct the n-weight quadrature (6.1) 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 (6.1) satisfying G1–G3) is different from the Gauss quadrature with multiple nodes considered in [6] and [43], and later in [20]. 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.

7. Conclusion. In Section 6 we saw that quasi-definitness of $\mathcal{L}$ is not only sufficient but also necessary condition for the n-weight quadrature (6.1) 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 6.2, the n-weight quadrature (6.1) having degree of exactness at least $2n-1$ either does not exist (the system (6.4) has no solution), or there are infinitely many of them (the system (6.4) has infinitely many solutions). Thus the property G2 cannot be satisfied. If there exist infinitely many n-weight quadratures (6.1), then Δ_n must also be equal to zero. Indeed, using (6.4), the first n rows of the matrix of the system (6.5) are linearly dependent. Hence by Theorem 6.3 the degree of exactness of the n-weight quadratures (6.1) 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 (6.1) having degree of exactness at least $2n-3$. However, by Theorem 6.3 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 6.2 and 6.3. With the $(n-1)$-weight quadrature Q_{n-1} the situation is different. If we only know the moments $m_0, \ldots, m_{2n-2}$, 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 5 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 (5.2) 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 (5.2) we can use a 2-weight quadrature of the form (6.1), i.e., $A_1 f(2) + A_2 f'(2)$. Since $\Delta_1 \neq 0$, by Theorem 6.2 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 6.3 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 (6.1) for linear functionals that are quasi-definite on $\mathcal{P}_n$ gives the maximal possible extension of this concept.

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] C. Brezinski, Padè-type approximation and general orthogonal polynomials, International series of numerical mathematics, Birkhäuser, 1980.

[4] C. Brezinski, M. Redivo Zaglia, H. Sadok, Avoiding breakdown and near-breakdown in Lanczos type algorithm, Numer. Algorithms 1 (1991) 261–284.

[5] C. Brezinski, H. Sadok, Avoiding breakdown in the CGS algorithm, Numer. Algorithms 1 (1991) 199–206.

[6] L. Chakalov, General quadrature formulae of Gaussian type, Bulgar. Akad. Nauk Izv. Mat. Inst. 1 (1954) 67–84.

[7] T.S. Chihara, An introduction to orthogonal polynomials, Gordon and Breach, New York, 1978.

[8] E.B. Christoffel, Über die Gaußische 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.

[9] 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.

[10] T.C. Chu Moody, G. H. Golub, Inverse Eigenvalue Problems, Theory, Algorithms, and Applications, Oxford University Press, New York, 2005.

[11] B.D. Craven, Complex symmetric matrices, J. Aust. Math. Soc. 10 (1969) 341–354.

[12] D.Ž. Djoković, Eigenvectors obtained from the adjoint matrix, Aequat. Math. 2 (1969) 94–97.

[13] J. Favard, Sur les polynomes de Tchebicheff, C. R. Acad. Sci., Paris, 200 (1935) 2052–2053.

[14] B. Fischer, Polynomial Based Iteration Methods for Symmetric Linear Systems, Wiley-Teubner Series Advances in Numerical Mathematics, John Wiley & Sons Ltd., Chichester, 1996.

[15] R.W. Freund, M. Hochbruck, Gauss quadratures associated with the Arnoldi process and the Laczos 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.

[16] F.R. Gantmakher, The Theory of Matrices 1,2, Chelsea Publishing Co., New York, 1959.

[17] 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.

[18] W. Gautschi, Numerical Analysis, 2nd ed., Birkhäuser, Boston, 2012.

[19] W. Gautschi, Orthogonal Polynomials: Computation and Approximation, Oxford University Press, Oxford, 2004.

[20] G.H. Golub, J. Kautsky, Calculation of Gauss quadratures with multiple free and fixed knots, Numer. Math. 41 (1983) 147–163.

[21] G.H. Golub, G.A. Meurant, Matrices, Moments, and Quadrature with Applications, Princeton

University Press, Princeton, N.J., 2010.
[22] G.H. Golub, J.H. Welsch, Calculation of Gauss quadrature rules, Math. Comput. 23 (1969) 221–230.
[23] W.B. Gragg, Matrix interpretations and applications of the continued fraction algorithm, Rocky Mountain J. Math. 4 (1974) 213–225.
[24] M.H. Gutknecht, A completed theory of the unsymmetric Lanczos process and related algorithms. I, SIAM J. Matrix Anal. Appl. 13 (1992) 594–639.
[25] M.H. Gutknecht, A completed theory of the unsymmetric Lanczos process and related algorithms. II, SIAM J. Matrix Anal. Appl. 15 (1994) 15–58.
[26] 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.
[27] E.D. Hellinger, H.S. Wall, Contributions to the Analytic Theory of Continued Fractions and Infinite Matrices, Ann. Math. 44 (1943) 103–127.
[28] M.R. Hestenes, E. Stiefel, Methods of conjugate gradients for solving linear systems, J. Research Nat. Bur. Standards 49 (1952) 409–436.
[29] N.J. Higham, Functions of Matrices: Theory and Computation. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2008.
[30] R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge University Press, Cambridge, UK, 1985.
[31] A.S. Householder, The Theory of Matrices in Numerical Analysis, Dover, New York, 1964.
[32] 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.
[33] 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.
[34] 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.
[35] C. Lanczos, Solution of systems of linear equations by minimized iterations, J. Research Nat. Bur. Standards 49 (1952) 33–53.
[36] J. Liesen, Z. Strakoš, Krylov Subspace Methods: Principles and Analysis, Oxford University Press, Oxford, 2013.
[37] F. Marcelán, R. Álvarez-Nodarse, On the "Favard theorem" and its extension, J. Comput. Appl. Math. 127 (2001) 231–254.
[38] J.R. Magnus, H. Neudecker, Matrix Differential Calculus with Applications in Statistics and Econometrics, J. Wiley & Sons Inc., Chichester, 1988.
[39] G. Meurant, Z. Strakoš, The Lanczos and conjugate gradient algorithms in finite precision arithmetic, Acta Numer. 15 (2006) 471–542.
[40] B.N. Parlett, Reduction to tridiagonal form and minimal realizations, SIAM J. Matrix Anal. Appl. 13 (1992) 567–593.
[41] B.N. Parlett, The Symmetric Eigenvalue Problem, Prentice-Hall, Englewood Cliffs, 1980.
[42] B.N. Parlett, D.R. Taylor, Z.A. Liu, A look-ahead Lanczos algorithm for unsymmetric matrices, Math. Comp. 44 (1985) 105–124.
[43] T. Popoviciu, Sur une généralisation de la formule d'integration numérique de Gauss, Acad. R. P. Romîne Fil. Iaşi Stud. Cerc. Şti. 6 (1955) 29–57.
[44] Y. Saad, Iterative Methods for Sparse Linear Systems, SIAM, Philadelphia, 2003.
[45] P.E. Saylor, D.C. Smolarski, Why Gauss quadrature in the complex plane?, Numer. Algorithms 26 (2001) 251–280.
[46] N.H. Scott, A Theorem on Isotropic Null Vectors and Its Application to Thermoelasticity, Proceedings: Mathematical and Physical Sciences 440 (1993) 431–442.
[47] B. Simon, CMV matrices: five years after, J. Comput. Appl. Math. 208 (2007) 120–154.
[48] A. Spitzbart, A Generalization of Hermite's Interpolation Formula, Amer. Math. Monthly 67(1) (1960) 42–46.
[49] D.D. Stancu, A.H. Stroud, Quadrature formulas with simple gaussian nodes and multiple fixed nodes, Math. Comput. 17(84) (1963) 384–394.
[50] 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.
[51] Z. Strakoš, Model reduction using the Vorobyev moment problem, Numer. Algorithms 51 (2009) 363–379.
[52] G. Szegö, Orthogonal polynomials, Amer. Math. Soc. Colloquium Publications, 23, New York,

1939.
[53] Y.V. Vorobyev, Method of moments in applied mathematics, Translated from the Russian by Bernard Seckler, Gordon and Breach Science Publishers, New York, 1965.
[54] H.S. Wall, Analytic Theory of Continued Fractions, Chelsea Pub. Co., Bronx, N.y., 1948.
[55] H. Wilf, Mathematics for the Physical Sciences, J. Wiley & Sons Inc., London, 1962.
[56] J.H. Wilkinson, The algebraic eigenvalue problem, Clarendon Press, Oxford, 1988.