Electronic Transactions on Numerical Analysis.
Volume 00, pp. 1–xx, 0000.

ISSN 1068–9613.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

THE LANCZOS ALGORITHM AND COMPLEX GAUSS QUADRATURE*

STEFANO POZZA†, MIROSLAV S. PRANIƆ, AND ZDENĚK STRAKOŠ§

Abstract. Gauss quadrature can be naturally generalized to approximate quasi-definite linear functionals where the interconnections with (formal) orthogonal polynomials, (complex) Jacobi matrices and Lanczos algorithm are analogous to those in the positive definite case. In this survey we review these relationships with giving references to literature that presents them in several related contexts. In particular, the existence of the n-weight (complex) Gauss quadrature corresponds to successfully performing the first n steps of the Lanczos algorithm for generating the biorthogonal bases of the two associated Krylov subspaces. The Jordan decomposition of the (complex) Jacobi matrix can be explicitly expressed in terms of the Gauss quadrature nodes and weights and the associated orthogonal polynomials. Since the output of the Lanczos algorithm can be made real whenever the input is real, the value of the Gauss quadrature is a real number whenever all relevant moments of the quasi-definite linear functional are real.

Key words. quasi-definite linear functionals, Gauss quadrature, formal orthogonal polynomials, complex Jacobi matrices, matching moments, Lanczos algorithm.

AMS subject classifications. 65D15, 65D32, 65F10, 47B36

1. Introduction.

The presented survey examines the interconnection between the Gauss quadrature for quasi-definite linear functionals and the Lanczos algorithm for generating the biorthogonal bases of the two associated Krylov subspaces.

We first briefly recall basic results on quasi-definite linear functionals and formal orthogonal polynomials; see, e.g., the summary in Chihara [7] and in the literature given below. As described in [12, Introduction], the term *formal orthogonal polynomials* was chosen in order to avoid the ambiguity of the term *general orthogonal polynomials* (used, e.g., in [2]) since the latter term has often appeared in literature regarding positive definite linear functional. Sometimes (as in [7]) *orthogonal polynomials* is used instead of *formal orthogonal polynomials*, i.e., the meaning of the simpler term is extended beyond the classical setting with a positive definite linear functional and a Riemann-Stieltjes integral with a non-decreasing distribution function; see, e.g., [55], [22], and [41, Section 3.3]. Since no confusion can arise, in what follows we will use this simplified terminology.

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

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

---

*Version March 15, 2018

†Faculty of Mathematics and Physics, Charles University, Sokolovská 83, 186 75 Praha 8, Czech Republic. Associated to Istituto di scienza e tecnologie dell'informazione - Consiglio Nazionale delle Ricerche, Via Giuseppe Moruzzi 1, Pisa, Italy. E-mail: `pozza@karlin.mff.cuni.cz`. This work has been supported by Charles University Research program UNCE/SCI/023, and 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: `miroslav.pranic@pmf.unibl.org`. Research supported in part by the Ministry of Science and Technology of R. Srpska and by the Serbian Ministry of Education and Science. The survey was prepared while on stay at Charles University.

§Faculty of Mathematics and Physics, Charles University, Sokolovská 83, 186 75 Praha 8, Czech Republic. E-mail: `strakos@karlin.mff.cuni.cz`. Supported by the Grant Agency of the Czech Republic under the contract No. 18–12719S.

with the associated Hankel determinants

$$\Delta_j = \begin{vmatrix} m_0 & m_1 & \dots & m_j \\ m_1 & m_2 & \dots & m_{j+1} \\ \vdots & \vdots & & \vdots \\ m_j & m_{j+1} & \dots & m_{2j} \end{vmatrix}, \quad j = 0, 1, \dots. \tag{1.2}$$

Hankel matrices have been used in the related contexts throughout more than a century by many authors; see, e.g., the seminal paper by Stieltjes [53, Sections 8–11, p. 624–630], [7, Chapter I], [27, Section 2], and [12, Chapter 1]. The linear functional (1.1) is generally determined by an infinite sequence of moments. This survey, however, considers linear functionals on finite-dimensional spaces of polynomials which are characterized by finite sequences of Hankel determinants (1.2). This approach is appropriate for linear functionals associated with finite-dimensional Krylov subspace methods; see [41]. For the infinite-dimensional problems, we refer, e.g., to [7, Chapter II, Section 3, in particular Theorem 3.1] and for the relationship to infinite dimensional Krylov subspace methods, e.g., to [57], [28] and [43, Chapter 5] that contain many references to original works.

In this survey we focus on quasi-definite linear functionals. Linear functionals that are not quasi-definite are, apart from several remarks, beyond the scope of this survey. For results in this more general setting we refer an interested reader to [12].

DEFINITION 1.1 (cf. [7, Chapter I, Definition 3.1, Definition 3.2 and Theorem 3.4]). *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 with complex coefficients $\mathcal{P}_k$ of degree at most k. In particular, if $\mathcal{L}$ has real moments $m_0, \dots, m_{2k}$ and $\Delta_j > 0$ for $j = 0, 1, \dots, k$ we will say that the linear functional is* positive definite *on $\mathcal{P}_k$.*

In the sequel we use for simplicity the term *quasi-definite linear functional* (*positive definite linear functional*) for linear functionals that are quasi-definite (positive definite) on the space of polynomials of sufficiently large degree. A quasi-definite linear functional can be associated with a sequence of orthogonal polynomials uniquely determined up to multiplicative constants.

DEFINITION 1.2. *Polynomials $p_0, p_1, \dots$ satisfying the conditions*

1. *$deg(p_j) = j$ (p_j is of degree j),*
2. *$\mathcal{L}(p_i\, p_j) = 0$, $i < j$,*
3. *$\mathcal{L}(p_j^2) \neq 0$,*

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

Orthogonal polynomials such that $\mathcal{L}(p_j^2) = 1$ are known as *orthonormal* polynomials. Proof of the following classical result can be found, e.g., in [7, Chapter I, Theorem 3.1], [42, Chapter VII, Theorem 1].

THEOREM 1.3. *A sequence $\{p_j\}_{j=0}^{k}$ of orthogonal polynomials with respect to $\mathcal{L}$ exists if and only if $\mathcal{L}$ is quasi-definite on $\mathcal{P}_k$.*

A sequence of orthogonal polynomials $p_0, p_1, \dots$ satisfies the three-term recurrences of the form

$$\delta_j p_j(\lambda) = (\lambda - \alpha_{j-1}) p_{j-1}(\lambda) - \gamma_{j-1} p_{j-2}(\lambda), \quad \text{for} \quad j = 1, 2, \dots, \tag{1.3}$$

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

$$\alpha_{j-1} = \frac{\mathcal{L}(\lambda p_{j-1}^2)}{\mathcal{L}(p_{j-1}^2)},\ \delta_j = \frac{\mathcal{L}(\lambda p_{j-1} p_j)}{\mathcal{L}(p_j^2)},\ \gamma_{j-1} = \frac{\mathcal{L}(\lambda p_{j-2} p_{j-1})}{\mathcal{L}(p_{j-2}^2)},$$

(see [55, Theorem 3.2.1], [7, p. 19], [2, Theorem 2.4]). If the first $n+1$ polynomials $p_0, p_1, \ldots, p_n$ exist, then all $\delta_1, \ldots, \delta_n$ and $\gamma_1, \ldots, \gamma_{n-1}$ are different from zero. The recurrence (1.3) for the first $n+1$ polynomials can be written in the matrix form

$$\lambda \begin{bmatrix} p_0(\lambda) \\ p_1(\lambda) \\ \vdots \\ p_{n-1}(\lambda) \end{bmatrix} = T_n \begin{bmatrix} p_0(\lambda) \\ p_1(\lambda) \\ \vdots \\ p_{n-1}(\lambda) \end{bmatrix} + \delta_n \begin{bmatrix} 0 \\ 0 \\ \vdots \\ p_n(\lambda) \end{bmatrix}, \tag{1.4}$$

where T_n is the irreducible tridiagonal complex matrix

$$T_n = \begin{bmatrix} \alpha_0 & \delta_1 & & \\ \gamma_1 & \alpha_1 & \ddots & \\ & \ddots & \ddots & \delta_{n-1} \\ & & \gamma_{n-1} & \alpha_{n-1} \end{bmatrix}.$$

We say that T_n is determined by the first $2n$ moments $m_0, m_1, \ldots, m_{2n-1}$ of $\mathcal{L}$. The $(2n+1)$st moment m_{2n} present in (1.2) for $j = n$ affects only the value of δ_n. Its value must assure that $\Delta_n \neq 0$; otherwise $\mathcal{L}(p_n^2) = 0$ and therefore p_n is not orthogonal polynomial with respect to $\mathcal{L}$.

A linear functional quasi-definite on $\mathcal{P}_n$ determines a family of irreducible tridiagonal matrices that are diagonally similar where this diagonal similarity is equivalent to rescaling the sequence of orthogonal polynomials. Any irreducible tridiagonal matrix is diagonally similar to a *symmetric* irreducible tridiagonal matrix, called *complex Jacobi matrix*. The properties of complex Jacobi matrices are summarized, e.g., in [49, Section 4]. Here we recall the following result that is valid for any tridiagonal matrix T_n associated with a sequence (1.4) of orthogonal polynomials determined by a quasi-definite linear functional (see [49, Section 5]).

THEOREM 1.4 (Matching moment property). *Let $\mathcal{L}$ be a quasi-definite linear functional on $\mathcal{P}_n$ and let T_n be given by (1.4). Then*

$$\mathcal{L}(\lambda^i) = m_0 \, \mathbf{e}_1^T (T_n)^i \, \mathbf{e}_1, \quad i = 0, \ldots, 2n-1. \tag{1.5}$$

A proof for the matching moment property was given in [17, Theorem 2] for the linear functionals defined by

$$\mathcal{L}(\lambda^i) = \mathbf{w}^* A^i \mathbf{v}, \quad \text{for} \quad i = 0, 1, 2, \ldots, \tag{1.6}$$

with A a complex matrix and $\mathbf{w}, \mathbf{v}$ vectors; cf. also [11, Theorem 1]. In [54] it was obtained using the Vorobyev method of moments (see [57, in particular Chapter III]). The class of non quasi-definite linear functionals of the kind (1.6) is treated in [29, Theorem 2.10]. We point out that assuming real moments (with the extension to complex moments being straightforward), the matching moment properties in [17], [49] and [29] can be derived from Theorem 5 of [27] where this issue is related to the *minimal partial realization* problem.

A *partial realization of the order* $2n$ of a sequence of moments $m_0, m_1, \ldots$ is the triplet $\{\mathbf{w}, A, \mathbf{v}\}$ where A is a matrix and $\mathbf{w}, \mathbf{v}$ are vectors such that

$$\mathbf{w}^* A^i \mathbf{v} = m_i, \quad \text{for} \quad i = 0, \ldots, 2n-1. \tag{1.7}$$

The solutions with the smallest dimension are known as *minimal partial realizations of the order* $2n$; see, e.g., [26], [37] and [27]. The moments $m_0, \ldots, m_{2n-1}$ define the linear functional $\mathcal{L}$ on $\mathcal{P}_{2n-1}$. If $\mathcal{L}$ is quasi-definite, then by Theorem 1.4 the triplet $A = T_n$, $\mathbf{w} = \mathbf{e}_1$, and $\mathbf{v} = m_0\mathbf{e}_1$ gives a solution of the minimal partial realization problem (1.7); cf. [27, Theorem 5]. Therefore, as beautifully presented by Gragg and Lindquist in [27] for real moments, the matching moment property connects the minimal partial realization problem with orthogonal polynomials, Jacobi matrices, Lanczos algorithm, continued fractions, and other related topics. The generalization to the case of complex moments is straightforward. For $\mathcal{L}$ positive definite, the concept equivalent to the minimal partial realization is present (without using the name) in the papers by Chebyshev from 1855–1859 [5, 6] and Christoffel from 1858 [8]; cf. the comment in [4, p. 23]. An instructive description can be found in the seminal paper by Stieltjes on continued fractions published in 1894 [53, Sections 7–8, p. 623–625, and Section 51, p. 688–690]; see also [41, Section 3.9.1], the survey by Gautschi [21] and the references therein.

On the other hand, as shown in [7, Chapter I, Theorem 4.4], in the survey [44, Theorem 2.14] and firstly for the positive definite case by Favard in [13], for any sequence of polynomials satisfying

$$d_j p_j(\lambda) = (\lambda - a_{j-1})p_{j-1}(\lambda) - c_{j-1}p_{j-2}(\lambda),\ j = 1, 2, \ldots, \tag{1.8}$$

where

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

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 $c, d_j, c_j \neq 0$, polynomials generated by (1.8) are always orthogonal polynomials. In addition, they are orthonormal if and only if $c_j = d_j$ and p_0 is such that $\mathcal{L}(p_0^2) = 1$.

This also means that for any irreducible tridiagonal matrix T_n, 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 [1, 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

$$\alpha_i = \widehat{\alpha}_i, \quad i = 0, \ldots, n-1, \tag{1.9}$$

and

$$\delta_i\, \gamma_i = \widehat{\delta}_i\, \widehat{\gamma}_i, \quad i = 1, \ldots, n-1, \tag{1.10}$$

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

The matching moment property in Theorem 1.4 can also be interpreted as matrix formulation of a generalized Gauss quadrature for approximation of quasi-definite linear functionals; see [45, 49]. Moreover, given the matrix A and the vectors $\mathbf{v}$ and $\mathbf{w}$ with the associated quasi-definite linear functional defined by (1.6), the matrix T_n can be determined assuming no breakdown by the non-Hermitian Lanczos algorithm. Therefore the non-Hermitian Lanczos algorithm can be linked with Gauss quadrature; see [17, Theorem 2].

A linear functional (1.1) with real moments can be naturally restricted to the space of polynomials with real coefficients $\mathcal{R} \subset \mathcal{P}$. If $\mathcal{L}$ is quasi-definite, we can construct real monic polynomials orthogonal with respect to $\mathcal{L}$ with the corresponding real tridiagonal matrix T_n satisfying the matching moment property (1.5). In Chapter 5 of the book [12] published in 1983 Draux introduced a generalization of the Gauss quadrature formula for

ETNA
Kent State University and
Johann Radon Institute (RICAM)

approximating the real-valued linear functionals satisfying for quasi-definite functionals (1.5). The associated results in [45, 49], obtained independently of [12], can be considered as straightforward generalizations to the complex case. Some results in [49] do not have, however, a straightforward real setting analogue in [12]. This holds, e.g., for the concept of orthonormal polynomials that can have even for real quasi-definite functionals complex coefficients.

The paper is organized as follows. In Section 2 we recall the link of the Lanczos algorithm for generating biorthonormal bases for the spaces

$$\text{span}\{\mathbf{v}, A\mathbf{v}, \dots, A^{n-1}\mathbf{v}\} \quad \text{and} \quad \text{span}\{\mathbf{w}, A^*\mathbf{w}, \dots, (A^*)^{n-1}\mathbf{w}\}$$

to the Stieltjes procedure for generating orthonormal polynomials. If n is the maximal number of steps that can be performed in the Lanczos algorithm without breakdown, then there exists no complex Gauss quadrature in the sense of [45, 49] for approximating the functional (1.6) with more than n weights. This is presented in Section 3. Section 4 shows that the rows of the matrix W^{-1} in the Jordan decomposition $J_n = W\,\Lambda\,W^{-1}$ of the complex Jacobi matrix J_n can be expressed as a linear combination of some particular generalized eigenvectors of J_n. The coefficients in these linear combinations are the Gauss quadrature weights. Section 5 focuses to quasi-definite functionals with real moments. Then the value of the Gauss quadrature is a real number. Using a proper rescaling, the Lanczos algorithm involves only computations with real numbers. We conclude with some remarks on the non quasi-definite case.

Throughout the survey we deal with mathematical relationships between quantities that are determined exactly. Since the effects of rounding errors to computations using short recurrences are substantial, the results of this survey cannot be applied to finite precision computations without a thorough analysis. Such analysis is out of the scope of this survey. As in the positive definite case, however, understanding of the relationship assuming exact computation is a prerequisite for any further investigation.

2. Orthogonal polynomials and the Lanczos algorithm. Let A be a square complex matrix and let $\mathbf{v}$ be a complex vector of the corresponding dimension. The nth *Krylov subspace* generated by A and $\mathbf{v}$ is defined by

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

or, equivalently,

$$\mathcal{K}_n(A, \mathbf{v}) = \{p(A)\mathbf{v} : p \in \mathcal{P}_{n-1}\},$$

where $\mathcal{P}_{n-1}$ is the subspace of the polynomials of degree at most $n-1$ with complex coefficients. The basic facts about Krylov subspaces had been formulated by Gantmacher in 1934; see [19]. In particular, there exists a uniquely defined integer $d = d(A, \mathbf{v})$, called *the grade of* $\mathbf{v}$ *with respect to* A, so that the vectors $\mathbf{v}, \dots, A^{d-1}\mathbf{v}$ are linearly independent and the vectors $\mathbf{v}, \dots, A^{d-1}\mathbf{v}, A^d\mathbf{v}$ are linearly dependent. Clearly there exists a polynomial $p_d(\lambda)$ of degree d, called the minimal polynomial of $\mathbf{v}$ with respect to A, such that $p_d(A)\mathbf{v} = 0$. The other facts about Krylov subspaces can be found elsewhere; see, e.g., [41, Section 2.2].

For the given complex matrix A and $\mathbf{v} \neq 0, \mathbf{w} \neq 0$ complex vectors, consider the linear functional on the space of polynomials with complex coefficients $\mathcal{P}$ (see (1.6))

$$\mathcal{L}(p) = \mathbf{w}^* p(A)\mathbf{v}. \tag{2.1}$$

Since for any polynomial $p \in \mathcal{P}$ we get

$$p(A)^* = \bar{p}(A^*),$$

with $\bar{p}$ the polynomial whose coefficients are the conjugates of the coefficients of p, given $p, q \in \mathcal{P}_{n-1}$ we have

$$\mathcal{L}(pq) = \mathbf{w}^* q(A) p(A) \mathbf{v} = \widehat{\mathbf{w}}^* \widehat{\mathbf{v}},$$

with $\widehat{\mathbf{v}} = p(A)\mathbf{v} \in \mathcal{K}_n(A, \mathbf{v})$ and $\widehat{\mathbf{w}} = \bar{q}(A^*)\mathbf{w} \in \mathcal{K}_n(A^*, \mathbf{w})$. We give the proof of the following elementary fact for completeness.

THEOREM 2.1. *The linear functional $\mathcal{L}$ defined by (2.1) determines a sequence of orthogonal polynomials $p_0, \ldots, p_{n-1}$ if and only if there exist bases $\mathbf{v}_0, \ldots, \mathbf{v}_{\ell-1}$ of $\mathcal{K}_\ell(A, \mathbf{v})$ and $\mathbf{w}_0 \ldots, \mathbf{w}_{\ell-1}$ of $\mathcal{K}_\ell(A^*, \mathbf{w})$, $\ell = 1, \ldots, n$, satisfying the biorthogonality condition*

$$\mathbf{w}_i^* \mathbf{v}_j = 0 \ \ \textit{for } i \neq j, \quad \textit{and} \quad \mathbf{w}_i^* \mathbf{v}_i \neq 0, \ \ i, j = 0, \ldots, n-1. \tag{2.2}$$

Proof. Given polynomials $p_0, \ldots, p_{n-1}$ orthogonal with respect to $\mathcal{L}$, the vectors $\mathbf{v}_j = p_j(A)\mathbf{v}$ ($j = 0, \ldots, n-1$) form the basis for $\mathcal{K}_n(A, \mathbf{v})$, vectors $\mathbf{w}_i = \bar{p}_i(A^*)\mathbf{w}$ ($i = 0, \ldots, n-1$) form the basis for $\mathcal{K}_n(A^*, \mathbf{w})$, and

$$\mathbf{w}_i^* \mathbf{v}_j = \mathcal{L}(p_i p_j), \quad i, j = 0, \ldots, n-1,$$

satisfy the biorthogonality condition (2.2). On the other hand, let $\mathbf{v}_j = p_j(A)\mathbf{v}$ and $\mathbf{w}_i = \bar{q}_i(A^*)\mathbf{w}$ satisfy

$$\mathbf{w}_i^* \mathbf{v}_j = 0 \ \ \text{for } i \neq j, \quad \text{and} \quad \mathbf{w}_i^* \mathbf{v}_i \neq 0, \ \ i, j = 0, \ldots, n-1,$$

and p_j and q_i are polynomials of degree j and i, respectively. It means that the polynomial p_i is orthogonal to the polynomials $q_0, q_1, \ldots, q_{i-1}$, and therefore also to polynomials $p_0, p_1, \ldots, p_{i-1}$. The polynomial p_i is not orthogonal to q_i, and thus $\mathcal{L}(p_i^2) \neq 0$. □

We denote $\widetilde{p}_0, \ldots, \widetilde{p}_{n-1}$ the sequence of orthonormal polynomials with respect to $\mathcal{L}$. They satisfy the three-term recurrences (cf. (1.3))

$$\beta_j \widetilde{p}_j(\lambda) = (\lambda - \alpha_{j-1}) \widetilde{p}_{j-1}(\lambda) - \beta_{j-1} \widetilde{p}_{j-2}(\lambda), \qquad j = 1, 2, \ldots, n-1, \tag{2.3}$$

with $\widetilde{p}_{-1} = 0$, $\widetilde{p}_0 = 1/\sqrt{m_0}$, and

$$\alpha_{j-1} = \mathcal{L}(\lambda \widetilde{p}_{j-1}^2), \ \ \beta_{j-1} = \mathcal{L}(\lambda \widetilde{p}_{j-2} \widetilde{p}_{j-1}). \tag{2.4}$$

Note that $\beta_j = \sqrt{\mathcal{L}(\widehat{p}_j^2)}$, with

$$\widehat{p}_j(\lambda) = (\lambda - \alpha_{j-1}) \widetilde{p}_{j-1}(\lambda) - \beta_{j-1} \widetilde{p}_{j-2}(\lambda). \tag{2.5}$$

Algorithm 2.2 generates the sequence of the first n orthonormal polynomials $\widetilde{p}_j$, $j = 0, \ldots, n-1$, using the formulas (2.3) and (2.4). In order to avoid ambiguity, we take always the principal value of the complex square root, i.e., we consider $arg(\sqrt{c}) \in (-\pi/2, \pi/2]$. For positive definite functionals this algorithm is known as the Stieltjes procedure [52]. Then the coefficients β_j, $j = 1, \ldots, n-1$, are positive. The monograph by Gautschi [22] can serve as a valuable source of related results as well as of historical information.

The Lanczos algorithm (introduced in [39] and [40]) gives the matrix formulation of the Stieltjes procedure; for details we refer to [2, Section 2.7.2], [31, 32, 33], [51, Chapter 7], [24, Chapter 4], [41, Section 2.4]. Indeed, with

$$\mathbf{v}_j = \widetilde{p}_j(A)\mathbf{v}, \quad \mathbf{w}_j = \overline{\widetilde{p}_j}(A^*)\mathbf{w}, \quad j = 0, \ldots, n-1,$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

ALGORITHM 2.2 (Stieltjes Procedure).
Input: linear functional $\mathcal{L}$ quasi-definite on $\mathcal{P}_{n-1}$.
Output: polynomials $\widetilde{p}_0, \ldots, \widetilde{p}_{n-1}$ orthonormal with respect to $\mathcal{L}$.

$$
\begin{aligned}
&\textit{Initialize: } \widetilde{p}_{-1} = 0, \beta_0 = \sqrt{m_0} = \sqrt{\mathcal{L}(\lambda^0)}, \widetilde{p}_0 = 1/\beta_0. \\
&\textit{For } j = 1, 2, \ldots, n-1 \\
&\qquad \alpha_{j-1} = \mathcal{L}(\lambda \widetilde{p}_{j-1}^2(\lambda)), \\
&\qquad \widehat{p}_j(\lambda) = (\lambda - \alpha_{j-1})\widetilde{p}_{j-1}(\lambda) - \beta_{j-1}\widetilde{p}_{j-2}(\lambda), \\
&\qquad \beta_j = \sqrt{\mathcal{L}(\widehat{p}_j^2)}, \\
&\qquad \widetilde{p}_j(\lambda) = \widehat{p}_j(\lambda)/\beta_j, \\
&end.
\end{aligned}
$$

ALGORITHM 2.3 (Lanczos algorithm).
Input: complex matrix A, two complex vectors $\mathbf{v}, \mathbf{w}$ such that $\mathbf{w}^*\mathbf{v} \neq 0$.
Output: vectors $\mathbf{v}_0, \ldots, \mathbf{v}_{n-1}$ that span $\mathcal{K}_n(A, \mathbf{v})$ and vectors $\mathbf{w}_0, \ldots, \mathbf{w}_{n-1}$ that span $\mathcal{K}_n(A^*, \mathbf{w})$, satisfying the biorthogonality conditions (2.2).

$$
\begin{aligned}
&\textit{Initialize: } \mathbf{v}_{-1} = \mathbf{w}_{-1} = 0,\ \beta_0 = \sqrt{\mathbf{w}^*\mathbf{v}} \\
&\qquad\qquad \mathbf{v}_0 = \mathbf{v}/\beta_0,\ \mathbf{w}_0 = \mathbf{w}/\bar{\beta}_0. \\
&\textit{For } j = 1, 2, \ldots, n-1 \\
&\qquad \alpha_{j-1} = \mathbf{w}_{j-1}^* A \mathbf{v}_{j-1}, \\
&\qquad \widehat{\mathbf{v}}_j = A\mathbf{v}_{j-1} - \alpha_{j-1}\mathbf{v}_{j-1} - \beta_{j-1}\mathbf{v}_{j-2}, \\
&\qquad \widehat{\mathbf{w}}_j = A^*\mathbf{w}_{j-1} - \bar{\alpha}_{j-1}\mathbf{w}_{j-1} - \bar{\beta}_{j-1}\mathbf{w}_{j-2}, \\
&\qquad \beta_j = \sqrt{\widehat{\mathbf{w}}_j^* \widehat{\mathbf{v}}_j}, \\
&\qquad \textit{if } \beta_j = 0 \textit{ then stop}, \\
&\qquad \mathbf{v}_j = \widehat{\mathbf{v}}_j/\beta_j, \\
&\qquad \mathbf{w}_j = \widehat{\mathbf{w}}_j/\bar{\beta}_j, \\
&end.
\end{aligned}
$$

we have for $j = 1, \ldots, n-1$

$$\alpha_{j-1} = \mathcal{L}(\lambda \widetilde{p}_{j-1}^2) = \mathbf{w}^* \widetilde{p}_{j-1}(A) A \widetilde{p}_{j-1}(A) \mathbf{v} = \mathbf{w}_{j-1}^* A \mathbf{v}_{j-1}.$$

Since $\beta_j^2 = \mathcal{L}(\widehat{p}_j^2(\lambda))$ with the polynomial $\widehat{p}_j$ defined by (2.5), we get

$$\beta_j = \sqrt{\mathbf{w}^* \widehat{p}_j(A) \widehat{p}_j(A) \mathbf{v}} = \sqrt{\widehat{\mathbf{w}}_j^* \widehat{\mathbf{v}}_j}, \quad j = 1, \ldots, n-1.$$

The vectors $\mathbf{v}_0, \ldots, \mathbf{v}_{n-1}$ satisfy the three-term recurrences (2.3)

$$\beta_j \mathbf{v}_j = (A - \alpha_{j-1})\mathbf{v}_{j-1} - \beta_{j-1}\mathbf{v}_{j-2}, \quad \text{for } j = 1, \ldots, n-1.$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Since $\mathbf{w}_j = \overline{\widetilde{p}_j}(A^*)\,\mathbf{w}$,

$$\bar{\beta}_j \mathbf{w}_j = (A^* - \bar{\alpha}_{j-1})\mathbf{w}_{j-1} - \bar{\beta}_{j-1}\mathbf{w}_{j-2}, \qquad \text{for } j = 1, \dots, n-1.$$

The resulting form of the Lanczos algorithm is given as Algorithm 2.3; see, e.g., [10, 9]. The matrices $V_n = [\mathbf{v}_0, \dots, \mathbf{v}_{n-1}]$ and $W_n = [\mathbf{w}_0, \dots, \mathbf{w}_{n-1}]$ satisfy

$$AV_n = V_n J_n + \widehat{\mathbf{v}}_n \mathbf{e}_n^T,$$
$$A^* W_n = W_n J_n^* + \widehat{\mathbf{w}}_n \mathbf{e}_n^T,$$

with $\mathbf{e}_n$ the nth vector of the canonical basis, J_n the complex Jacobi matrix associated with the polynomials $\widetilde{p}_0, \dots, \widetilde{p}_{n-1}$,

$$J_n = \begin{bmatrix} \alpha_0 & \beta_1 & & \\ \beta_1 & \alpha_1 & \ddots & \\ & \ddots & \ddots & \beta_{n-1} \\ & & \beta_{n-1} & \alpha_{n-1} \end{bmatrix}, \tag{2.6}$$

and α_{n-1}, $\widehat{\mathbf{v}}_n$, $\widehat{\mathbf{w}}_n$ are determined at the step n of the Lanczos algorithm*. The biorthogonality conditions (2.2) then give

$$W_n^* V_n = I_n,$$
$$W_n^* A V_n = J_n,$$

where I_n is the identity matrix of dimension n. Algorithm 2.3 can be seen as a tool for restriction of A to the Krylov subspace $\mathcal{K}_n(A, \mathbf{v})$ with the subsequent projection orthogonal to $\mathcal{K}_n(A^*, \mathbf{w})$. The reduced operator on $\mathcal{K}_n(A, \mathbf{v})$ then can be expressed via the complex Jacobi matrix J_n. Lanczos algorithm 2.3 is based on orthonormal polynomials. Obviously, any other scaling of orthogonal polynomials can be used, i.e., the Lanczos algorithm can be based on any sequence of orthogonal polynomials associated with the linear functional (2.1).

Recall that if $\mathcal{L}$ is quasi-definite on $\mathcal{P}_{n-1}$, then $\beta_j = \sqrt{\mathcal{L}(\widehat{p}_j^2)}$ must be different from zero for $j = 1, \dots, n-1$. Therefore no breakdown can occur in the first $n-1$ steps of the Lanczos algorithm. There is a breakdown at the step n if and only if $\beta_n = 0$. This can happen in two cases:

1. one of the vectors $\widehat{\mathbf{v}}_n$ and $\widehat{\mathbf{w}}_n$ is the zero vector,
2. $\widehat{\mathbf{v}}_n \neq \mathbf{0}$ and $\widehat{\mathbf{w}}_n \neq \mathbf{0}$, but $\widehat{\mathbf{w}}_n^* \widehat{\mathbf{v}}_n = 0$.

In the first case, either $\mathcal{K}_n(A, \mathbf{v})$ is A-invariant or $\mathcal{K}_n(A^*, \mathbf{w})$ is A^*-invariant. This is known as *lucky breakdown* (or *benign breakdown*) because the computation of an invariant subspace is often a desirable result; see, e.g., [46, Section 5] and [25, Section 10.5.5]. The second case is known as *serious breakdown*; for further details we refer to [50], [36, p. 34], [56, Chapter IV], [47], [46, Section 7], and [31, 32, 33]. The previous development is summarized in the following Theorem, cf. also [3, 46].

THEOREM 2.4. *Let $A \in \mathbb{C}^{N\times N}$, $\mathbf{v} \in \mathbb{C}^N$ and $\mathbf{w} \in \mathbb{C}^N$ be the input for the Lanczos algorithm, let $m_k = \mathbf{w}^* A^k \mathbf{v}$, and let Δ_k be the corresponding Hankel determinants (1.2) for $k = 0, 1, \dots$. There are no breakdowns at the first $n-1$ steps of the Lanczos algorithm if and only if*

$$\prod_{k=0}^{n-1} \Delta_k \neq 0. \tag{2.7}$$

*The coefficient α_{n-1} present in J_n and the vectors $\widehat{\mathbf{v}}_n$ and $\widehat{\mathbf{w}}_n$ are well defined even in the case of breakdown at the step n.

There is a breakdown at the subsequent step n if and only if, in addition to (2.7), $\Delta_n = 0$. In other words, the Lanczos algorithm has a breakdown at the step n if and only if the linear functional (2.1) is quasi-definite on $\mathcal{P}_{n-1}$, but not on $\mathcal{P}_n$.

If the matrix A is Hermitian, $\mathbf{v} = \mathbf{w} \neq 0$, and $d = d(A, \mathbf{v})$ is the grade of $\mathbf{v}$ with respect to A, then the moments of $\mathcal{L}$ defined by (2.1) are real and there exists the non-decreasing distribution function μ supported on the real axis having d points of increase such that $\mathcal{L}$ can be represented by the Riemann-Stieltjes integral

$$\mathcal{L}(p) = \int_{\mathbb{R}} p(\lambda)\, \mathrm{d}\mu(\lambda), \qquad \text{for } p \in \mathcal{P}\,;$$

see, e.g., [24, Section 7.1] and [41, Section 3.5]. Then $\mathcal{L}$ is a *positive definite linear functional* on $\mathcal{P}_{d-1}$, the corresponding Hankel determinants Δ_j, $j = 0, \dots, d-1$, are positive and $\Delta_d = 0$; see, e.g., [7, Chapter I, Definition 3.1 and Theorem 3.4] and [49, Section 2].

3. The Gauss quadrature and the Lanczos algorithm. Consider a non-decreasing distribution function $\mu(\lambda)$ on $\mathbb{R}$ having finite limits at $\pm\infty$ and infinitely many points of increase. If all the moments of the Riemann-Stieltjes integral

$$m_i = \int_{\mathbb{R}} \lambda^i\, \mathrm{d}\mu(\lambda), \quad i = 0, 1, \dots$$

exist and are finite, then we can define the positive definite linear functional on the space of polynomials with real coefficients $\mathcal{L} : \mathcal{R} \to \mathbb{R}$ as

$$\mathcal{L}(p) = \int_{\mathbb{R}} p(\lambda)\, \mathrm{d}\mu(\lambda), \quad p \in \mathcal{R}\,. \tag{3.1}$$

Then the Gauss quadrature is given by the unique n-node quadrature formula which matches the first $2n$ moments of the Riemann-Stieltjes integral (3.1). The classical results on the Gauss quadrature can be found in many books; see, e.g., [55, Chapters III and XV], [7, Chapter I, Section 6]; [22, Section 1.4], [23, Chapter 3.2], [41, Section 3.2]. The 1981 survey by Gautschi [21] contains many results as well as historical comments of the matter. In this section we present results about the extension of the Gauss quadrature for the approximation of quasi-definite linear functionals $\mathcal{L} : \mathcal{P} \to \mathbb{C}$ with generally complex moments

$$m_i = \mathcal{L}(\lambda^i), \quad i = 0, 1, \dots\,.$$

We recall the definition of *matrix function*; for more information including equivalence to the other definitions of matrix function see, e.g., [34]. A function f is *defined on the spectrum of the given matrix A* if for every eigenvalue λ_i of A 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 A in which λ_i appears. Let Λ be a Jordan block of A of the size s corresponding to the eigenvalue λ. The matrix function $f(\Lambda)$ is then defined as

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

Denoting

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

the Jordan decomposition of A, the *matrix function* $f(A)$ is defined by

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

Given a linear functional $\mathcal{L}$ on the space of sufficiently smooth functions, consider the quadrature of the form (see [12, Chapter 5], [45, Section 2], and [49, Section 7])

$$\mathcal{L}(f) \approx \mathcal{G}_n(f) = \sum_{i=1}^{\ell} \sum_{j=0}^{s_i-1} \omega_{i,j} \, f^{(j)}(\lambda_i), \quad n = s_1 + \cdots + s_\ell, \tag{3.2}$$

with $\omega_{i,j}$ the weights, λ_i the nodes, and s_i the multiplicity of the node λ_i. Notice that the number of *different* nodes in (3.2) is equal to ℓ, and ℓ can be less than n. If we count the multiplicities, then the number of nodes is equal to n, that is also the number of weights in (3.2). In order to avoid ambiguity, we refer to (3.2) as the *n-weight quadrature*, instead of the n-point or n-node quadrature as is usually done. For any choice of (different) nodes λ_i, $i = 1, \dots, \ell$, and their multiplicities s_i, such that $s_1 + \cdots + s_\ell = n$, it is possible to achieve that the quadrature (3.2) is exact for any f from $\mathcal{P}_{n-1}$. As shown in [12, Theorem 5.1] or in the proof of Theorem 7.1 in [49], it is necessary and sufficient to take

$$\omega_{i,j} = \mathcal{L}(h_{i,j}), \tag{3.3}$$

where $h_{i,j}$ are polynomials from $\mathcal{P}_{n-1}$ such that

$$\begin{aligned} h_{i,j}^{(t)}(\lambda_k) &= 1 \quad \text{for } \lambda_k = \lambda_i \text{ and } t = j, \\ h_{i,j}^{(t)}(\lambda_k) &= 0 \quad \text{for } \lambda_k \neq \lambda_i \text{ or } t \neq j, \end{aligned}$$

with $k = 1, 2, \dots, \ell$, and $t = 0, 1, \dots, s_i - 1$. In this case we say that the quadrature (3.2) is interpolatory, since it can be obtained by applying the linear functional $\mathcal{L}$ to the generalized (Hermite) interpolating polynomial for the function f at the nodes λ_i of the multiplicities s_i.

In [49], it is referred to (3.2) as the n-weight Gauss quadrature approximating the linear functionals $\mathcal{L}$ on the space of polynomials $\mathcal{P}$ if and only if the following three properties are satisfied.

- G1: the n-weight 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-weight Gauss quadrature is well-defined and it is unique. Moreover, Gauss quadratures with a smaller number of weights also exist and they are unique.
- G3: the Gauss quadrature of a function f can be written as the quadratic form $m_0 \, \mathbf{e}_1^T f(J_n) \mathbf{e}_1$, where J_n is the complex Jacobi matrix containing the coefficients from the three-term recurrences for orthonormal polynomials associated with $\mathcal{L}$; $m_0 = \mathcal{L}(\lambda^0)$.

In what follows we will refer to this quadrature as *complex Gauss quadrature*. We will, however, use the adjective *complex* only when it is necessary to emphasize the difference with respect to the standard n-node Gauss quadrature described at the beginning of this section.

The property G3 assumes existence of the first n orthonormal polynomials with respect to $\mathcal{L}$, i.e., by Theorem 1.3, it considers only quasi-definite linear functionals on $\mathcal{P}_n$. Naturally, we can state the following theorem. The detailed proof and discussion can be found, e.g., in [49, Section 7, in particular Corollaries 7.4 and 7.5].

THEOREM 3.1. *Let $\mathcal{L}$ be a linear functional on $\mathcal{P}$. There exists the n-weight complex Gauss quadrature, i.e., the quadrature (3.2) having properties G1, G2 and G3, if and only if $\mathcal{L}$ is quasi-definite on $\mathcal{P}_n$.*

The nodes λ_i, $i = 1, \ldots, \ell$, of the n-weight Gauss quadrature (3.2), and their multiplicities s_i, $s_1 + \cdots + s_\ell = n$, coincide with:

- the roots of the n-degree orthonormal polynomial $\widetilde{p}_n$ with respect to $\mathcal{L}$ with their corresponding multiplicities;
- the eigenvalues of the complex Jacobi matrix J_n with their corresponding algebraic multiplicities;

see, e.g., [49, Theorem 7.1 and the discussion on p. 21–22]. The weights are given by (3.3). Theorem 3.1 says that the definition of the complex Gauss quadrature (3.2) satisfying G1–G3 cannot be used for non quasi-definite linear functionals. A slightly different definition for an arbitrary real-valued linear functional defined on $\mathcal{R}$ was given in [12, Section 5]; Draux considers Gauss quadrature (3.2) having a maximal possible degree of exactness (which is $2n-1$ in the quasi-definite case).

The property G3 is actually a consequence of the properties G1 and G2 [49, Corollary 7.5]. We formulated it explicitly in order to stress the link of the complex Gauss quadrature with complex Jacobi matrices. Complex Gauss quadrature (3.2) for a quasi-definite linear functional $\mathcal{L} : \mathcal{P} \to \mathbb{C}$ is associated with a complex Jacobi matrix J_n, which is unique, providing that the arguments of the off-diagonal complex entries are in $(-\pi/2, \pi/2]$. Moreover, by Favard Theorem (see Section 1), any complex Jacobi matrix determines the Gauss quadrature for some quasi-definite linear functional. The setting in [12] considers the Gauss quadrature for real linear functionals on the space of polynomials with real coefficients $\mathcal{L} : \mathcal{R} \to \mathbb{R}$ and therefore the link with complex Jacobi matrices (i.e., symmetric irreducible tridiagonal matrices; see Section 1) is not given there.

If the linear functional quasi-definite on $\mathcal{P}_n$ is given by (2.1), then the associated complex Jacobi matrix (2.6) can be constructed by performing n steps of the Algorithm 2.3; see Section 2. The property G3 then presents Lanczos algorithm as a matrix formulation of the Gauss quadrature (see [17, in particular Theorem 2]). Analogous arguments for the block Lanczos algorithm can be found, e.g., in [14, Section 3].

The same can be stated for any linear functional $\mathcal{L}$ quasi-definite on $\mathcal{P}_n$. Given the numbers $m_0, m_1, \ldots, m_{2n}$ such that the Hankel determinant Δ_j is nonzero for $j = 0, 1, \ldots, n$ (see (1.2)), there always exist a square matrix A and vectors $\mathbf{v}$ and $\mathbf{w}$ such that

$$\mathbf{w}^* A^k \mathbf{v} = m_k, \quad k = 0, \ldots, 2n.$$

For instance, take $A \in \mathbb{C}^{2n+1 \times 2n+1}$ and $\mathbf{v}, \mathbf{w} \in \mathbb{C}^{2n+1}$ as

$$A = \begin{bmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} m_0 \\ m_1 \\ \vdots \\ m_{2n} \end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Then the first $2n+1$ moments of $\mathcal{L}$ and the first $2n+1$ moments of the functional $\widetilde{\mathcal{L}}(f) = \mathbf{w}^* f(A)\mathbf{v}$ are equal and $\widetilde{\mathcal{L}}$ is quasi-definite on $\mathcal{P}_n$. Moreover, the n-weight Gauss quadrature for $\mathcal{L}$ can be identified with $m_0 \mathbf{e}_1^T f(J_n) \mathbf{e}_1$, where J_n is the complex Jacobi matrix obtained at the step n of the Algorithm 2.3 with the input A, $\mathbf{v}$ and $\mathbf{w}$. Therefore any complex Gauss quadrature given by G1–G3 can be constructed by the Lanczos algorithm.

We remark that if $\mathcal{L}$ is quasi-definite on $\mathcal{P}_{n-1}$ but it is not quasi-definite on $\mathcal{P}_n$, then the Lanczos algorithm has a breakdown at the step n; see Theorem 2.4. However, the nth step of Algorithm 2.3 still gives the complex Jacobi matrix J_n related to the recurrences of the n orthonormal polynomials $\widetilde{p}_0, \dots, \widetilde{p}_{n-1}$. The quadrature rule $\mathcal{L}(f) \approx m_0 \mathbf{e}_1^T f(J_n)\mathbf{e}_1$ is not the complex Gauss quadrature since its degree of exactness is larger than $2n-1$, i.e.,

$$\mathcal{L}(\lambda^k) = m_0 \mathbf{e}_1^T (J_n)^k \mathbf{e}_1, \quad k = 0, \dots, j,$$

where $j \geq 2n$; see [49, Sections 7 and 8]. However, since Draux considers in [12] Gauss quadrature (3.2) having a maximal possible degree of exactness, the property G3 formulates Gauss quadrature in the sense of [12] (in the real setting).

4. Jordan decomposition of complex Jacobi matrices. Let J_n be an arbitrary $n \times n$ complex Jacobi matrix. Then there exists a linear functional $\mathcal{L}$ quasi-definite on $\mathcal{P}_n$ such that J_n contains the coefficients from the three-term recurrences for orthonormal polynomials $\widetilde{p}_j$, $j = 0, \dots, n$, associated with $\mathcal{L}$. J_n is a non-derogatory matrix (see, e.g., [49, Section 4]), i.e., it has ℓ distinct eigenvalues $\lambda_1, \dots, \lambda_\ell$, all having the geometric multiplicity 1. We write its Jordan decomposition as

$$J_n = W \operatorname{diag}(\Lambda_1, \dots, \Lambda_\ell) W^{-1}, \tag{4.1}$$

where Λ_i is the Jordan block of dimension s_i associated with the eigenvalue λ_i, $i = 1, \dots, \ell$. For any $t = 1, \dots, n$ there is exactly one integer i between 1 and ℓ, and exactly one integer j between 0 and $s_i - 1$, such that $t = s_1 + \dots + s_{i-1} + j + 1$ (here, for $i = 1$, $s_0 \equiv 0$). In other words, fixed t uniquely determines i and j, and vice versa, fixed i and j uniquely determine t. The tth column $\mathbf{w}_{t(i,j)}$ of W can be written as (see [48, p. 274], [38, Lemma 2], and [49, Proposition 4.4])

$$\mathbf{w}_{t(i,j)} = \frac{1}{j!} \begin{bmatrix} \mathbf{0}_j \\ \widetilde{p}_j^{(j)}(\lambda_i) \\ \vdots \\ \widetilde{p}_{n-1}^{(j)}(\lambda_i) \end{bmatrix}, \tag{4.2}$$

where $\mathbf{0}_j$ is the zero vector of length j. The next theorem, which can also be derived, considering the extension to complex linear functionals and Favard Theorem, from the formulas on p. 277 of [48], gives the explicit formula for the rows of W^{-1}.

THEOREM 4.1. *Let $J_n = W \operatorname{diag}(\Lambda_1, \dots, \Lambda_\ell) W^{-1}$ be the Jordan decomposition of an $n \times n$ complex Jacobi matrix J_n. Let $\mathcal{L}$ be the quasi-definite linear functional on $\mathcal{P}_n$ such that J_n contains the coefficients from the three-term recurrences for the orthonormal polynomials $\widetilde{p}_0, \dots, \widetilde{p}_n$ with respect to $\mathcal{L}$, and let $\sum_{i=1}^{\ell} \sum_{j=0}^{s_i-1} \omega_{i,j} f^{(j)}(\lambda_i)$ be the Gauss quadrature for $\mathcal{L}$ defined by (3.2) and (3.3). Then the rth row $\mathbf{v}_{r(i,j)}^T$ of W^{-1},*

$$\mathbf{v}_{r(i,j)}^T = \mathbf{e}_{r(i,j)}^T W^{-1}, \quad r = s_1 + \dots + s_{i-1} + j + 1 \quad (s_0 \equiv 0 \text{ for } i = 1),$$

has the following representation

$$\mathbf{v}_{r(i,j)} = \sum_{\nu=j}^{s_i-1} \nu! \, \omega_{i,\nu} \, \mathbf{w}_{t(i,\nu-j)}, \tag{4.3}$$

with $\mathbf{w}_{t(i,\nu-j)}$ defined by (4.2).

Proof. Let V be the $n \times n$ matrix with the rows $\mathbf{v}_{r(i,j)}$, $r = 1, \dots, n$, given by (4.3). We will show that $WV = I_n$, i.e., $V = W^{-1}$. Denote the kth row of W by $\mathbf{a}_k^T$, and the mth column of V by $\mathbf{b}_m$ and prove that

$$\mathbf{a}_k^T \mathbf{b}_m = \mathcal{L}(\widetilde{p}_{k-1} \widetilde{p}_{m-1}).$$

By (4.2) the qth element of $\mathbf{a}_k$ is

$$a_{k,q} = \frac{\widetilde{p}_{k-1}^{(j)}(\lambda_i)}{j!}, \quad q = s_0 + s_1 + \ldots + s_{i-1} + j + 1,$$

where for $k-1 < j$ we have $\widetilde{p}_{k-1}^{(j)}(\lambda_i) = 0$. Using (4.3), the qth element of $\mathbf{b}_m$ is

$$b_{m,q} = \sum_{\nu=j}^{s_i-1} \nu!\, \omega_{i,\nu} \frac{\widetilde{p}_{m-1}^{(\nu-j)}(\lambda_i)}{(\nu-j)!} = j! \sum_{\nu=j}^{s_i-1} \binom{\nu}{j} \omega_{i,\nu}\, \widetilde{p}_{m-1}^{(\nu-j)}(\lambda_i).$$

Thus we get, by rearranging the order of summations,

$$\begin{aligned} \sum_{q=1}^{n} a_{k,q} b_{m,q} &= \sum_{q=1}^{n} \sum_{\nu=j}^{s_i-1} \binom{\nu}{j} \omega_{i,\nu}\, \widetilde{p}_{m-1}^{(\nu-j)}(\lambda_i) \widetilde{p}_{k-1}^{(j)}(\lambda_i) \\ &= \sum_{i=1}^{\ell} \sum_{j=0}^{s_i-1} \omega_{i,j} \sum_{u=0}^{j} \binom{j}{u} \widetilde{p}_{m-1}^{(j-u)}(\lambda_i) \widetilde{p}_{k-1}^{(u)}(\lambda_i) \\ &= \sum_{i=1}^{\ell} \sum_{j=0}^{s_i-1} \omega_{i,j} (\widetilde{p}_{m-1}\widetilde{p}_{k-1})^{(j)}(\lambda_i) = \mathcal{L}(\widetilde{p}_{k-1}\widetilde{p}_{m-1}), \end{aligned}$$

which gives the result. $\square$

The weights $\omega_{i,j}$ defined by (3.3) of the Gauss quadrature in Theorem 4.1 can be expressed by the matrix W and its inverse; see [38, Equations (8) and (11)].

REMARK 4.2. The fact that a complex Jacobi matrix J_n is symmetric is associated with the requirement $WV = I_n$ and therefore the orthogonal polynomials $\widetilde{p}_j$, $j = 0, \ldots, n$, being orthonormal. The previous development can be easily modified for the Jordan decomposition $T_n = W \mathrm{diag}(\Lambda_1, \ldots, \Lambda_\ell) W^{-1}$ of an arbitrary irreducible tridiagonal matrix T_n. The representation (4.2) of the columns of W will then use the orthogonal polynomials p_j satisfying the three-term recurrences with the coefficients given by T_n (see, e.g., [49, Proposition 4.4]),

$$\mathbf{w}_{t(i,j)} = \frac{1}{j!} \begin{bmatrix} \mathbf{0}_j \\ p_j^{(j)}(\lambda_i) \\ \vdots \\ p_{n-1}^{(j)}(\lambda_i) \end{bmatrix}. \tag{4.4}$$

The matrix V with the rows defined by (4.3) satisfies

$$WV = \mathrm{diag}(\mathcal{L}(p_0^2), \ldots, \mathcal{L}(p_{n-1}^2)),$$

i.e.,

$$W^{-1} = V \mathrm{diag}(1/\mathcal{L}(p_0^2), \ldots, 1/\mathcal{L}(p_{n-1}^2)).$$

The rows of W^{-1} can then be written as

$$\mathbf{v}_{r(i,j)} = \sum_{\nu=j}^{s_i-1} \nu!\, \omega_{i,\nu}\, \widetilde{\mathbf{w}}_{t(i,\nu-j)}, \tag{4.5}$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

with

$$\widetilde{\mathbf{w}}_{t(i,j)} = \frac{1}{j!}\begin{bmatrix} \mathbf{0}_j \\ p_j^{(j)}(\lambda_i)/\mathcal{L}(p_j^2) \\ \vdots \\ p_{n-1}^{(j)}(\lambda_i)/\mathcal{L}(p_{n-1}^2) \end{bmatrix}; \tag{4.6}$$

cf. [48] where the real monic orthogonal polynomials are considered.

5. The Gauss quadrature for linear functionals with real moments. Let us now focus on a quasi-definite linear functional $\mathcal{L} : \mathcal{P} \to \mathbb{C}$ which has real moments $m_j = \mathcal{L}(\lambda^j)$, for $j = 0, 1, \dots$. Restricting $\mathcal{L}$ to the space of polynomials with real coefficients $\mathcal{R}$ gives a real-valued linear functional. We can still use the *complex Gauss quadrature* $\mathcal{G}_n$ described in Section 3 to approximate $\mathcal{L}$ and its restriction to $\mathcal{R}$. At first glance, the idea of approximating such a functional by the quadrature with complex nodes and weights does not seem attractive. As we will see, however, the value of $\mathcal{G}_n(f)$ is, for suitable f, always a real number.

As presented above, in [12, Chapter 5] Draux defined a slightly different Gauss quadrature for arbitrary real-valued linear functional defined on the space of polynomials with real coefficients $\mathcal{R}$. Using Draux definition based on the maximal degree of exactness, it is possible to approximate real-valued linear functionals which are not quasi-definite, which means that, in general, Draux quadrature does not satisfy the properties G1–G3 in Section 3. If $\mathcal{L}$ is a linear functional with real moments quasi-definite on the space of polynomials with real coefficients, then the complex Gauss quadrature $\mathcal{G}_n$ is equal to the n-weight quadrature defined by Draux. In general, we have the following statement.

THEOREM 5.1. *Let $\mathcal{L}$ be a quasi-definite linear functional on $\mathcal{P}_n$ whose moments $m_0, \dots, m_{2n-1}$ are real, and let $\mathcal{G}_n$ be the associated Gauss quadrature* (3.2)*,*

$$\mathcal{G}_n(f) = \sum_{i=1}^{\ell}\sum_{j=0}^{s_i-1} \omega_{i,j}\, f^{(j)}(\lambda_i).$$

Then the following holds:

1. *The nodes λ_i, $i = 1, \dots, \ell$, are real or appear in complex conjugate pairs, i.e., for any $\lambda_i \notin \mathbb{R}$ with multiplicity s_i there is a node $\lambda_m = \overline{\lambda}_i$ with the same multiplicity.*
2. *For any $\lambda_i \in \mathbb{R}$ we have $\omega_{i,j} \in \mathbb{R}$, $j = 0, \dots, s_i - 1$. If $\lambda_i \notin \mathbb{R}$ and $\lambda_m = \overline{\lambda}_i$, then $\omega_{m,j} = \overline{\omega}_{i,j}$ for $j = 0, \dots, s_i - 1$.*
3. *If f is a real-valued function satisfying $f^{(j)}(\bar{\lambda}_i) = \overline{f^{(j)}(\lambda_i)}$ for $i = 1, \dots, \ell$ and $j = 0, \dots, s_i - 1$, then $\mathcal{G}_n(f)$ is a real number.*

Proof. The monic orthogonal polynomials $\pi_0, \pi_1, \dots, \pi_n$ associated with $\mathcal{L}$ satisfy

$$\pi_j(\lambda) = (\lambda - \alpha_{j-1})\pi_{j-1}(\lambda) - \eta_{j-1}\pi_{j-2}(\lambda), \quad j = 1, 2, \dots, n,$$

with $\alpha_0 = m_1/m_0, \pi_{-1}(\lambda) = 0$, $\pi_0(\lambda) = 1$, and

$$\alpha_{j-1} = \frac{\mathcal{L}(\lambda\pi_{j-1}^2)}{\mathcal{L}(\pi_{j-1}^2)}, \quad \eta_{j-1} = \frac{\mathcal{L}(\pi_{j-1}^2)}{\mathcal{L}(\pi_{j-2}^2)}, \quad j = 2, \dots, n\,.$$

The moments of $\mathcal{L}$ are real, which implies that $\alpha_{j-1}, \eta_{j-1} \in \mathbb{R}$ for $j = 2, \dots, n$, and the polynomials π_j, $j = 0, \dots, n$ have real coefficients. Since the roots of π_n are the nodes $\lambda_1, \dots, \lambda_\ell$ with the corresponding multiplicities $s_1, \dots, s_\ell$, we have proved the first statement.

Let T_n be the tridiagonal matrix associated with $\pi_0, \dots, \pi_n$. Then T_n is real and has the eigenvalues $\lambda_1, \dots, \lambda_\ell$ with the multiplicities $s_1, \dots, s_\ell$. We will prove the second statement by induction on j, using the Jordan decomposition

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

with (4.4), (4.5), and (4.6). If λ_i is not real, then there exists the eigenvalue $\lambda_m = \overline{\lambda_i}$, with $s_m = s_i$. Since $\pi_k(\bar{\lambda}) = \overline{\pi_k(\lambda)}$ for $k = 0, \dots, n$, then

$$\mathbf{w}_{t(i,j)} = \overline{\mathbf{w}_{u(m,j)}}, \quad \widetilde{\mathbf{w}}_{t(i,j)} = \overline{\widetilde{\mathbf{w}}_{u(m,j)}}, \quad j = 0, \dots, s_i - 1.$$

Fix $j = s_i - 1 = s_m - 1$ as the base case of the inductive proof. Then expression (4.5) gives

$$(\mathbf{v}_{r(i,s_i-1)})^T = (s_i - 1)!\, \omega_{i,s_i-1} (\widetilde{\mathbf{w}}_{t(i,0)})^T,$$

$$(\mathbf{v}_{q(m,s_m-1)})^T = (s_i - 1)!\, \omega_{m,s_m-1} (\widetilde{\mathbf{w}}_{t(i,0)})^*.$$

Using $(\mathbf{v}_{r(i,s_i-1)})^T \mathbf{w}_{r(i,s_i-1)} = 1$ and $(\mathbf{v}_{q(m,s_m-1)})^T \overline{\mathbf{w}_{r(i,s_i-1)}} = 1$ with the two previous equations, it follows that

$$\frac{1}{\omega_{i,s_i-1}} = (s_i - 1)!\, (\widetilde{\mathbf{w}}_{t(i,0)})^T \mathbf{w}_{r(i,s_i-1)} \quad \text{and} \quad \frac{1}{\omega_{m,s_m-1}} = (s_i - 1)!\, \overline{(\widetilde{\mathbf{w}}_{t(i,0)})^T \mathbf{w}_{r(i,s_i-1)}}\,.$$

Hence $\omega_{i,s_i-1} = \overline{\omega}_{m,s_m-1}$, which finishes the initial step. Let us fix j between 0 and $s_i - 2$ and let $\omega_{i,k} = \overline{\omega}_{m,k}$, $k = j+1, \dots, s_i - 1$, be the inductive assumptions. Then $(\mathbf{v}_{t(i,j)})^T \mathbf{w}_{t(i,j)} = 1$ and (4.5) give

$$\sum_{\nu=j}^{s_i-1} \nu!\, \omega_{i,\nu}\, (\widetilde{\mathbf{w}}_{r(i,\nu-j)})^T \mathbf{w}_{t(i,j)} = 1.$$

The first summand on the left-hand side of the previous equation can be written as

$$\begin{aligned} j!\, \omega_{i,j}\, (\widetilde{\mathbf{w}}_{r(i,0)})^T \mathbf{w}_{t(i,j)} &= 1 - \sum_{\nu=j+1}^{s_i-1} \nu!\, \omega_{i,\nu}\, (\widetilde{\mathbf{w}}_{r(i,\nu-j)})^T \mathbf{w}_{t(i,j)} \\ &= 1 - \sum_{\nu=j+1}^{s_i-1} \nu!\, \overline{\omega}_{m,\nu}\, \overline{(\widetilde{\mathbf{w}}_{q(m,\nu-j)})^T}\; \overline{\mathbf{w}_{u(m,j)}} \\ &= j!\, \overline{\omega_{m,j}\, (\widetilde{\mathbf{w}}_{q(m,0)})^T \mathbf{w}_{u(m,j)}} \\ &= j!\, \overline{\omega}_{m,j}\, (\widetilde{\mathbf{w}}_{r(i,0)})^T \mathbf{w}_{t(i,j)}. \end{aligned}$$

Therefore $\omega_{i,j} = \overline{\omega}_{m,j}$ for $j = 0, \dots, s_i - 1$. If, on the other hand, λ_i is real, then an analogous induction gives $\omega_{i,j} \in \mathbb{R}$, $j = 0, \dots, s_i - 1$. In this case, the vectors $\mathbf{w}_{t(i,j)}$ and $\widetilde{\mathbf{w}}_{t(i,j)}$ are real, which finishes the proof of the second part of the statement.

Finally, if f is a real-valued function satisfying $f^{(j)}(\bar{\lambda}_i) = \overline{f^{(j)}(\lambda_i)}$ for $i = 1, \dots, \ell$ and $j = 0, \dots, s_i - 1$, then $\mathcal{G}_n(f)$ is real by construction. □

As shown in the proof of Theorem 5.1, if $\mathcal{L}$ is a linear functional with real moments quasi-definite on $\mathcal{P}_n$, then there exists a irreducible real tridiagonal matrix T_n associated with the monic orthogonal polynomials $\pi_1, \dots, \pi_n$. Therefore by (1.9) all the tridiagonal matrices determined by a quasi-definite linear functional with real moments have real numbers on the

ETNA
Kent State University and
Johann Radon Institute (RICAM)

ALGORITHM 5.2 (Lanczos algorithm in the real number setting).
Input: real matrix A, two real vectors $\mathbf{v}, \mathbf{w}$ such that $\mathbf{w}^*\mathbf{v} \neq 0$.
Output: vectors $\mathbf{v}_0, \ldots, \mathbf{v}_{n-1}$ that span $\mathcal{K}_n(A, \mathbf{v})$ and vectors $\mathbf{w}_0, \ldots, \mathbf{w}_{n-1}$ that span $\mathcal{K}_n(A^*, \mathbf{w})$, satisfying the biorthogonality conditions (2.2).

$$
\begin{aligned}
&\textit{Initialize: } \mathbf{v}_{-1} = \mathbf{w}_{-1} = 0,\ \gamma_0 = 0,\ \widehat{s} = 1,\ s = 1,\\
&\qquad\qquad \mathbf{v}_0 = \mathbf{v}/||\mathbf{v}||,\ \mathbf{w}_0 = \mathbf{w}/(\mathbf{w}^*\mathbf{v}_0).\\
&\textit{For } j = 1, 2, \ldots, n\\
&\qquad \alpha_{j-1} = s \cdot \mathbf{w}_{j-1}^* A \mathbf{v}_{j-1},\\
&\qquad \widehat{\mathbf{v}}_j = A\mathbf{v}_{j-1} - \alpha_{j-1}\mathbf{v}_{j-1} - \gamma_{j-1}\mathbf{v}_{j-2},\\
&\qquad \widehat{\mathbf{w}}_j = A^*\mathbf{w}_{j-1} - \alpha_{j-1}\mathbf{w}_{j-1} - \gamma_{j-1}\mathbf{w}_{j-2},\\
&\qquad s = \textit{sign}\,(\widehat{\mathbf{w}}_j^*\widehat{\mathbf{v}}_j),\\
&\qquad \textit{if } s = 0 \textit{ then stop},\\
&\qquad \delta_j = \sqrt{|\widehat{\mathbf{w}}_j^*\widehat{\mathbf{v}}_j|},\\
&\qquad \gamma_j = s \cdot \widehat{s} \cdot \delta_j,\\
&\qquad \widehat{s} = s,\\
&\qquad \mathbf{v}_j = \widehat{\mathbf{v}}_j/\delta_j,\\
&\qquad \mathbf{w}_j = \widehat{\mathbf{w}}_j/\delta_j,\\
&\textit{end}.
\end{aligned}
$$

main diagonal. Moreover, by (1.10) the elements at the super-diagonal of the corresponding (complex symmetric) Jacobi matrix are either real or pure imaginary. Notice that a complex Jacobi matrix J_n is real if and only if it is determined by a linear functional positive definite on $\mathcal{P}_n$; see, e.g., [44, Theorem 2.14].

The previous discussion can now be applied to the Lanczos algorithm with a real input. For the given real matrix A and $\mathbf{v} \neq 0, \mathbf{w} \neq 0$ real vectors, the moments of the linear functional $\mathcal{L} : \mathcal{P} \to \mathbb{C}$ defined by

$$\mathcal{L}(p) = \mathbf{w}^* p(A)\mathbf{v}, \quad p \in \mathcal{P} \tag{5.1}$$

are real. The output after n steps of the Lanczos algorithm is real if and only if the algorithm is based on orthogonal polynomials satisfying the three-term recurrences with real coefficients. Since Algorithm 2.3 is based on orthonormal polynomials, its n steps cannot result in a real output unless the functional (5.1) is positive definite on $\mathcal{P}_n$. If this assumption cannot be used, the output of the Lanczos algorithm is real providing that the algorithm is based on monic orthogonal polynomials. However, in this case there is no further rescaling of the vectors $\widehat{\mathbf{v}}_j$ and $\widehat{\mathbf{w}}_j$, $j = 0, 1, \ldots$. If the rescaling of the vectors $\widehat{\mathbf{v}}_j, \widehat{\mathbf{w}}_j$ is required (for any reason), then one can use the following modification; cf. [30, Section 2, in particular equation (2.21a)]. The polynomials $p_0 = \widetilde{p}_0, \ldots, p_{j-1} = \widetilde{p}_{j-1}$ are constructed by Algorithm 2.2 as long as they have real coefficients, i.e., as long as $\mathcal{L}(\widehat{p}_k^2)$, $k = 0, \ldots, j-1$, is positive. When $\mathcal{L}(\widehat{p}_j^2)$ is negative, then we rescale $\widehat{p}_j$ in the following way:

$$\delta_j = \sqrt{|\mathcal{L}(\widehat{p}_j^2)|}, \qquad p_j = \frac{\widehat{p}_j}{\delta_j}.$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Thus we get the sequence of orthogonal polynomials such that $\mathcal{L}(p_j^2)$ is either 1 or -1. The other coefficients from the three-term recurrences are also real. They are given by

$$\gamma_j = \frac{\mathcal{L}(\lambda p_{j-1} p_j)}{\mathcal{L}(p_{j-1}^2)} = \frac{\mathcal{L}(p_j^2)}{\mathcal{L}(p_{j-1}^2)} \delta_j = \begin{cases} \delta_j, & \text{if} \quad \mathcal{L}(p_{j-1}^2) \cdot \mathcal{L}(p_j^2) = 1 \\ -\delta_j, & \text{if} \quad \mathcal{L}(p_{j-1}^2) \cdot \mathcal{L}(p_j^2) = -1, \end{cases}$$

$$\alpha_j = \frac{\mathcal{L}(\lambda p_j^2)}{\mathcal{L}(p_j^2)} = \begin{cases} \mathcal{L}(\lambda p_j^2), & \text{if} \quad \mathcal{L}(p_j^2) = 1 \\ -\mathcal{L}(\lambda p_j^2), & \text{if} \quad \mathcal{L}(p_j^2) = -1. \end{cases}$$

The resulting form of the Lanczos algorithm involving only real number computations is given as Algorithm 5.2; see, e.g., Algorithm 1 with equation (2.21a) in [30]. The tridiagonal matrix $T_n = W_n^* A V_n$ obtained by the first n iterations of the algorithm has sub- and super-diagonal elements such that $\delta_j = \gamma_j$ or $\delta_j = -\gamma_j$, for $j = 1, \ldots, n-1$.

6. Conclusion. The survey presents in the comprehensive form the Lanczos algorithm as a matrix representation of the complex Gauss quadrature, with pointing out many related results published in various contexts previously. The weights $\omega_{i,j}$ of the Gauss quadrature (3.2) appear in the representation (4.3) of the rows of W^{-1} from the Jordan decomposition (4.1) of the corresponding complex Jacobi matrix. When the moments of the quasi-definite linear functional approximated by the Gauss quadrature $\mathcal{G}_n$ are real, the non-real nodes and weights of $\mathcal{G}_n$ come in the conjugate pairs. Therefore the value of $\mathcal{G}_n(f)$ is a real number whenever the real-valued function f satisfies $f^{(j)}(\bar{\lambda}_i) = \overline{f^{(j)}(\lambda_i)}$ for $i = 1, \ldots, \ell$ and $j = 0, \ldots, s_i - 1$. This property is linked with the fact that if the input is real, then the Lanczos algorithm with an appropriate rescaling can be performed in the real number setting.

If the linear functional $\mathcal{L}$ is not quasi-definite on $\mathcal{P}_n$, then the maximal algebraic degree of exactness of the n-weight quadrature (3.2) is not given a priori (see Section 3). The well-known Theorem 1.3 shows that it is not possible to define a sequence of n orthogonal polynomials for a linear functional which is not quasi-definite on $\mathcal{P}_n$ (it should be recalled that throughout the paper, as pointed out at the beginning of Section 1, the term orthogonal polynomials covers also the widely used term formal orthogonal polynomials). Therefore it is not trivial to extend the Gauss quadrature and the Lanczos algorithm to the case of a non quasi-definite linear functional. In order to extend the discussed results to the non quasi-definite case, it is required to define a sequence of polynomials $q_0, q_1, \ldots, q_n$ satisfying some relaxed orthogonality conditions; see, e.g., [12, Chapter 1]. These polynomials satisfy short recurrences that generalize the three-term recurrences (1.3) (see, e.g., [26, p. 222–223], Remark 1.2 in [12, p. 71] and Theorem 2 in [27]). The polynomials q_j, $j = 0, \ldots, n$, determine the Gauss quadratures with at most n weights as defined in [12, Chapter 5] for the case of real-valued linear functionals, and they are at the basis of the look-ahead strategies for the Lanczos algorithm; see, e.g., [15, 18, 16, 32] and [35, Section 6.3]. Moreover, the matching moment property for arbitrary linear functionals is also related to the minimal partial realization problem for a general sequence of moments; see [27, Section 3]. Assuming real moments (with the extension to complex moments being straightforward) the results about the Gauss quadrature for an arbitrary linear functional, and about minimal partial realization of a general sequence of moments were published in the same year (1983) by Draux [12, Chapter 5] and by Gragg and Lindquist [27]. We remark that the Gauss quadrature from [12] and the minimal partial realization described in [27] are equivalent. Further connections between Gauss quadrature for arbitrary linear functionals on the space of polynomials with complex coefficients, the look-ahead Lanczos algorithm, and the minimal partial realization problem will be considered elsewhere.

REFERENCES

[1] B. BECKERMANN AND V. KALIAGUINE, *The diagonal of the Padé table and the approximation of the Weyl function of the second-order difference operators*, Constr. Approx., 13 (1997), pp. 481–510.

[2] C. BREZINSKI, *Padé-type approximation and general orthogonal polynomials*, Internat. Ser. Numer. Math., Birkhäuser, 1980.

[3] C. BREZINSKI, M. R. ZAGLIA, AND H. SADOK, *A breakdown-free Lanczos type algorithm for solving linear systems*, Numer. Math., 63 (1992), pp. 29–38.

[4] 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.

[5] P. CHEBYSHEV, *Sur les fractions continues*, (1855). Reprinted in Oeuvres I, 11 (Chelsea, New York, 1962), pp. 203–230.

[6] ———, *Le développement des fonctions à une seule variable*, (1859). Reprinted in Oeuvres I, 19 (Chelsea, New York, 1962), pp. 501–508.

[7] T. S. CHIHARA, *An Introduction to Orthogonal Polynomials*, vol. 13 of Mathematics and its Applications, Gordon and Breach Science Publishers, New York, 1978.

[8] E. B. CHRISTOFFEL, *Über die Gaußische Quadratur und eine Verallgemeinerung derselben*, J. Reine Angew. Math., 55 (1858), pp. 61–82. Reprinted in Gesammelte mathematische Abhandlungen I (B. G. Teubner, Leipzig, 1910), pp. 65–87.

[9] J. CULLUM, W. KERNER, AND R. WILLOUGHBY, *A generalized nonsymmetric Lanczos procedure*, Comput. Phys. Commun., 53 (1989), pp. 19–48.

[10] J. CULLUM AND R. A. WILLOUGHBY, *A practical procedure for computing eigenvalues of large sparse nonsymmetric matrices*, in North-Holland Mathematics Studies, vol. 127, Elsevier, 1986, pp. 193–240.

[11] G. CYBENKO, *An explicit formula for Lanczos polynomials*, Linear Algebra Appl., 88 (1987), pp. 99–115.

[12] A. DRAUX, *Polynômes Orthogonaux Formels*, vol. 974 of Lecture Notes in Math. , Springer-Verlag, Berlin, 1983.

[13] J. FAVARD, *Sur les polynomes de Tchebicheff*, C. R. Acad. Sci. Paris, 33 (1935), pp. 2052–2053.

[14] C. FENU, D. MARTIN, L. REICHEL, AND G. RODRIGUEZ, *Block Gauss and Anti-Gauss Quadrature with Application to Networks*, SIAM J. Matrix Anal. Appl., 34 (2013), pp. 1655–1684.

[15] R. W. FREUND, *The Look-Ahead Lanczos Process for Large Nonsymmetric Matrices and Related Algorithms*, in Linear Algebra for Large Scale and Real-Time Applications, M. S. Moonen, G. H. Golub, and B. L. R. de Moor, eds., vol. 232 of NATO ASI Series, Kluwer Acad. Publ., Dordrecht, 1993, pp. 137–163.

[16] R. W. FREUND, M. H. GUTKNECHT, AND N. M. NACHTIGAL, *An implementation of the look-ahead Lanczos algorithm for non-Hermitian matrices*, SIAM J. Sci. Comput., 14 (1993), pp. 137–158.

[17] R. W. FREUND AND M. HOCHBRUCK, *Gauss Quadratures Associated With the Arnoldi Process and the Lanczos Algorithm*, in Linear Algebra for Large Scale and Real-Time Applications, M. S. Moonen, G. H. Golub, and B. L. R. de Moor, eds., vol. 232 of NATO ASI Series, Kluwer Acad. Publ., 1993, pp. 377–380.

[18] R. W. FREUND AND H. ZHA, *A look-ahead algorithm for the solution of general Hankel systems*, Numer. Math., 64 (1993), pp. 295–321.

[19] F. R. GANTMACHER, *On the algebraic analysis of Krylov's method of transforming the secular equation*, Trans. Second Math. Congress, II (1934), pp. 45–48. In Russian. Title translation as in [20].

[20] ———, *The Theory of Matrices. Vols. 1, 2*, Chelsea Publishing Co., New York, 1959.

[21] W. GAUTSCHI, *A survey of Gauss-Christoffel quadrature formulae*, in E. B. Christoffel (Aachen/Monschau, 1979), Birkhäuser, Basel, 1981, pp. 72–147.

[22] ———, *Orthogonal Polynomials: Computation and Approximation*, Numer. Math. Sci. Comput. , Oxford University Press, New York, 2004.

[23] ———, *Numerical Analysis*, Birkhäuser Boston, 2012.

[24] G. H. GOLUB AND G. MEURANT, *Matrices, Moments and Quadrature with Applications*, Princeton Ser. Appl. Math., Princeton University Press, Princeton, NJ, 2010.

[25] G. H. GOLUB AND C. F. VAN LOAN, *Matrix Computations*, Johns Hopkins Stud. Math. Sci., Johns Hopkins University Press, Baltimore, MD, fourth ed., 2013.

[26] W. B. GRAGG, *Matrix interpretations and applications of the continued fraction algorithm*, Rocky Mountain J. Math., 4 (1974), pp. 213–225.

[27] W. B. GRAGG AND A. LINDQUIST, *On the partial realization problem*, Linear Algebra Appl., 50 (1983), pp. 277–319.

[28] A. GÜNNEL, R. HERZOG, AND E. SACHS, *A note on preconditioners and scalar products in Krylov subspace methods for self-adjoint problems in Hilbert space*, Elect. Trans. Numer. Anal., 41 (2014), pp. 13–20.

[29] H. GUO AND R. A. RENAUT, *Estimation of $u^T f(A) v$ for large-scale unsymmetric matrices*, Numer. Linear Algebra Appl., 11 (2004), pp. 75–89.

[30] M. H. GUTKNECHT, *The unsymmetric Lanczos algorithms and their relations to Padé approximation, continued fractions and the QD algorithm*, in Proceedings of the Copper Mountain Conference on Iterative Methods, 1990.

[31] ———, *A completed theory of the unsymmetric Lanczos process and related algorithms. I*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 594–639.

[32] ———, *A completed theory of the unsymmetric Lanczos process and related algorithms. II*, SIAM J. Matrix Anal. Appl., 15 (1994), pp. 15–58.

[33] ———, *The Lanczos Process and Padé approximation*, in Proceedings of the Cornelius Lanczos International Centenary Conference (Raleigh, NC, 1993), SIAM, Philadelphia, PA, 1994, pp. 61–75.

[34] N. J. HIGHAM, *Functions of Matrices. Theory and Computation*, SIAM, Philadelphia, PA, 2008.

[35] M. HOCHBRUCK, *The Padé Table and its Relation to Certain Numerical Algorithms*, Habilitation thesis, Universität Tübingen, (1996).

[36] A. S. HOUSEHOLDER AND F. L. BAUER, *On certain methods for expanding the characteristic polynomial*, Numer. Math., 1 (1959), pp. 29–37.

[37] R. E. KALMAN, *On partial realizations, transfer functions, and canonical forms*, Acta Polytech. Scand. Math. Comput. Sci. Ser., 31 (1979), pp. 9–32.

[38] J. KAUTSKÝ, *Matrices related to interpolatory quadratures*, Numer. Math., 36 (1980/81), pp. 309–318.

[39] 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), pp. 255–282.

[40] ———, *Solution of systems of linear equations by minimized iterations*, J. Research Nat. Bur. Standards, 49 (1952), pp. 33–53.

[41] J. LIESEN AND Z. STRAKOŠ, *Krylov subspace methods: principles and analysis*, Numer. Math. Sci. Comput., Oxford University Press, Oxford, 2013.

[42] L. LORENTZEN AND H. WAADELAND, *Continued Fractions with Applications*, vol. 3 of Stud. Comput. Math., North-Holland Publishing Co., Amsterdam, 1992.

[43] J. MALEK AND Z. STRAKOS, *Preconditioning and the Conjugate Gradient Method in the Context of Solving PDEs:*, SIAM Spotlights, SIAM, Philadelphia, 2015.

[44] F. MARCELLÁN AND R. ÁLVAREZ-NODARSE, *On the "Favard theorem" and its extensions*, J. Comput. Appl. Math., 127 (2001), pp. 231–254.

[45] G. V. MILOVANOVIĆ AND A. S. CVETKOVIĆ, *Complex Jacobi Matrices and quadrature rules*, Filomat, 17 (2003), pp. 117–134.

[46] B. N. PARLETT, *Reduction to tridiagonal form and minimal realizations*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 567–593.

[47] B. N. PARLETT, D. R. TAYLOR, AND Z. A. LIU, *A look-ahead Lanczos algorithm for unsymmetric matrices*, Math. Comp., 44 (1985), pp. 105–124.

[48] M. PIÑAR AND V. RAMÍREZ, *Matrix interpretation of formal orthogonal polynomials for nondefinite functionals*, J. Comput. Appl. Math., 18 (1987), pp. 265–277.

[49] S. POZZA, M. S. PRANIĆ, AND Z. STRAKOŠ, *Gauss quadrature for quasi-definite linear functionals*, IMA J. Numer. Anal., 37 (2017), pp. 1468–1495.

[50] H. RUTISHAUSER, *Beiträge zur Kenntnis des Biorthogonalisierungs-Algorithmus von Lanczos*, Z. Angew. Math. Physik, 4 (1953), pp. 35–56.

[51] Y. SAAD, *Iterative Methods for Sparse Linear Systems*, SIAM, Philadelphia, PA, second ed., 2003.

[52] T. J. STIELTJES, *Quelques recherches sur la théorie des quadratures dites mécaniques*, Ann. Sci. École Norm. Sup. (3), 1 (1884), pp. 409–426. Reprinted in Oeuvres I (P. Noordhoff, Groningen, 1914), pp. 377–394.

[53] ———, *Recherches sur les fractions continues*, Ann. Fac. Sci. Toulouse Sci. Math. Sci. Phys., 8 (1894), pp. J. 1–122. Reprinted in Oeuvres II (P. Noordhoff, Groningen, 1918), pp. 402–566. English translation *Investigations on continued fractions* in Thomas Jan Stieltjes, Collected Papers, Vol. II (Springer-Verlag, Berlin, 1993), pp. 609–745.

[54] Z. STRAKOŠ, *Model reduction using the Vorobyev moment problem*, Numer. Algorithms, 51 (2009), pp. 363–379.

[55] G. SZEGÖ, *Orthogonal Polynomials*, vol. XXIII of Amer. Math. Soc. Colloq. Publ., American Mathematical Society, New York, 1939.

[56] D. R. TAYLOR, *Analysis of the look ahead Lanczos algorithm*, PhD thesis, University of California, Berkeley, 1982.

[57] Y. V. VOROBYEV, *Methods of Moments in Applied Mathematics*, Translated from the Russian by Bernard Seckler, Gordon and Breach Science Publishers, New York, 1965.