Numer. Math. 57:607–624

Numerische Mathematik

Modified Moments for Indefinite Weight Functions

Gene H. Golub[1] **and Martin H. Gutknecht**[2]

[1] Computer Science Department, Stanford University, Stanford, CA 94305, USA
[2] Interdisciplinary Project Center for Supercomputing, ETH Zurich, ETH-Zentrum, CH-8092 Zurich, Switzerland

Received July 7, 1989 / January 8, 1990

Dedicated to R.S. Varga on the occasion of his sixtieth birthday

Summary. The problem of generating the recurrence coefficients of orthogonal polynomials from the moments or from modified moments of the weight function is well understood for positive weight distributions. Here we extend this theory and the basic algorithms to the case of an indefinite weight function. While the generic indefinite case is formally not much different from the positive definite case, there exist nongeneric degenerate situations, and these require a different more complicated treatment. The understanding of these degenerate situations makes it possible to construct a stable approximate solution of an ill-conditioned problem.

The application to adaptive iterative methods for linear systems of equations is anticipated.

Subject Classifications: AMS(MOS) 30E05; 41A21, 65F10.

1 Modified moments and matrix relations associated to a measure with a full set of formal orthogonal polynomials (generic case)

Let λ be any complex measure defined on a subset Ω of the complex plane $\mathbf{C}$ with the property that the *moments*

$$\mu_n := \varphi(z^n) := \int_\Omega z^n d\lambda(z) \quad (n \in \mathbf{N} := \{0, 1, 2, \ldots\}) \tag{1.1}$$

all exist and such that there exists a full set of monic *formal orthogonal polynomials* π_n ($n \in \mathbf{N}$) so that π_n has exact degree n and

---

Offprint requests to: G.H. Golub (USA, Canada), M.H. Gutknecht (other countries)

(1.2) $$\varphi(\pi_m\pi_n) = \int_\Omega \pi_m(z)\pi_n(z)d\lambda(z) = \begin{cases} 0 & \text{if } n \neq m, \\ \sigma_{n,n} \neq 0 & \text{if } n = m. \end{cases}$$

This implies in particular that

(1.3) $$\varphi(p\pi_n) = 0 \text{ for any } p \in \mathscr{P}_{n-1}.$$

φ is a linear functional defined by (1.1) on the monomials and extended to the set of all polynomials by linearity. Note that the bilinear form that could be generated from (1.2) by linearity is in general not a positive definite inner product, even if the measure is positive real, since the polynomial values may be complex, but the overbar indicating the complex conjugate values is missing. If the measure is real and has real support ("real case"), then the polynomials π_n can be chosen to be real-valued on $\mathbf{R}$.

Under the above assumption the polynomials satisfy a *three-term recurrence*

(1.4) $$\pi_{n+1}(z) = (z - \alpha_n)\pi_n(z) - \beta_n\pi_{n-1}(z) \quad (n \in \mathbf{N}),$$

with $\pi_{-1}(z) :\equiv 0$, $\pi_0(z) :\equiv 1$. We could assign to β_0 an arbitrary value, but it is useful to set $\beta_0 := \mu_0$, cf. [4]. In general, $\alpha_n, \beta_n \in \mathbf{C}$, but in the "real case" $\alpha_n, \beta_n \in \mathbf{R}$.

In addition, consider a second sequence $\{\tau_n\}$ of monic polynomials with τ_n having exact degree n. They satisfy in general a recurrence of the form

(1.5) $$\tau_{n+1}(z) = z\tau_n(z) - \sum_{k=0}^{n} \tau_{k,n}\tau_k(z) \quad (n \in \mathbf{N}),$$

but we are particularly interested in the case where they too fulfill a three-term recurrence, say,

(1.6) $$\tau_{n+1}(z) = (z - \alpha'_n)\tau_n(z) - \beta'_n\tau_{n-1}(z) \quad (n \in \mathbf{N}),$$

with $\tau_{-1}(z) :\equiv 0$, $\tau_0(z) :\equiv 1$. To be more specific, we are, e.g., interested in the case where τ_n is a suitably shifted and normalized nth Chebyshev polynomial.

To link the two sets of polynomials we consider the *modified moments*

(1.7) $$\nu_n := \int_\Omega \tau_n(z)d\lambda(z) \quad (n \in \mathbf{N})$$

(the integrals of the second set of polynomials with respect to the measure of the first set of polynomials) and the quantities

(1.8) $$\sigma_{m,n} := \varphi(\tau_m\pi_n) = \int_\Omega \tau_m(z)\pi_n(z)d\lambda(z) \quad (m, n \in \mathbf{N}).$$

Note that

(1.9) $$\sigma_{m,0} = \nu_m \quad (m \in \mathbf{N}),$$

that the definition of $\sigma_{n,n}$ is in accordance with the one in (1.2) since $\varphi((\tau_n - \pi_n)\pi_n) = 0$ due to $\tau_n - \pi_n \in \mathscr{P}_{n-1}$, and that in view of (1.3)

(1.10) $$\sigma_{m,n} = 0 \quad \text{if} \quad m < n.$$

Next let us introduce matrix notation: The infinite row vectors

$$\mathbf{p} := [\pi_0, \pi_1, \pi_2, \ldots], \quad \mathbf{t} := [\tau_0, \tau_1, \tau_2, \ldots], \tag{1.11}$$

have polynomials as their elements, and $\mathbf{p}(z)$, $\mathbf{t}(z)$ will denote the corresponding vectors of values at z. The matrices

$$H := \begin{bmatrix} \alpha_0 & \beta_1 & & \\ 1 & \alpha_1 & \beta_2 & \\ & 1 & \alpha_2 & \ddots \\ & & \ddots & \ddots \end{bmatrix}, \quad T := \begin{bmatrix} \tau_{0,0} & \tau_{0,1} & \tau_{0,2} & \cdots \\ 1 & \tau_{1,1} & \tau_{1,2} & \cdots \\ & 1 & \tau_{2,2} & \cdots \\ & & \ddots & \ddots \end{bmatrix}, \tag{1.12}$$

and

$$S := \begin{bmatrix} \sigma_{0,0} & & & \\ \sigma_{1,0} & \sigma_{1,1} & & \\ \sigma_{2,0} & \sigma_{2,1} & \sigma_{2,2} & \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \quad D := \begin{bmatrix} \sigma_{0,0} & & & \\ & \sigma_{1,1} & & \\ & & \sigma_{2,2} & \\ & & & \ddots \end{bmatrix}. \tag{1.13}$$

are also infinite. H is tridiagonal, T is upper Hessenberg (and under assumption (1.6) also tridiagonal), S is nonsingular lower triangular, and D is diagonal. (Note that S is the transposed of Gautschi's corresponding matrix Σ [4]. In the case $\tau_m(z) = z^m$ the elements of this matrix were already used by Chebyshev [1].

With this notation the recurrence relations (1.4) and (1.5) become

$$z\mathbf{p}(z) = \mathbf{p}(z)H, \quad z\mathbf{t}(z) = \mathbf{t}(z)T, \tag{1.14}$$

while (1.2) and (1.8) turn into

$$\varphi(\mathbf{p}^T\mathbf{p}) = D, \quad \varphi(\mathbf{t}^T\mathbf{p}) = S. \tag{1.15}$$

Here, φ is supposed to be applied element by element to the rank-one matrices $\mathbf{p}^T\mathbf{p}$ and $\mathbf{t}^T\mathbf{p}$.

From (1.14) and (1.15) we get a matrix relation connecting H, T, and S:

$$SH = \varphi(\mathbf{t}^T\mathbf{p})H = \varphi(\mathbf{t}^T\mathbf{p}H) = \varphi(\mathbf{t}^T z\mathbf{p}) = \varphi(T^T\mathbf{t}^T\mathbf{p}) = T^T\varphi(\mathbf{t}^T\mathbf{p}) = T^T S,$$

i.e.,

$$SH = T^T S. \tag{1.16}$$

Since H, T, and S are tridiagonal, upper Hessenberg, and lower triangular, respectively, the matrix products in (1.16) are lower Hessenberg.

Clearly, (1.16) implies for $L := S^{-1}$ that

$$HL = LT^T, \tag{1.17}$$

which is again an equality between two lower Hessenberg matrices. Likewise, if $\tilde{\mathbf{t}} := [\tilde{\tau}_0, \tilde{\tau}_1, \tilde{\tau}_2, \ldots]$ denotes yet another sequence of monic polynomials and $z\tilde{\mathbf{t}} = \tilde{\mathbf{t}}\tilde{T}$ is the matrix form of its recurrence, we get by the argument that lead to (1.16) for the new matrix $W := \varphi(\mathbf{t}^T\tilde{\mathbf{t}})$ the relation

$$W\tilde{T} = T^T W. \tag{1.18}$$

Note that W contains in its first row the modified moments of the $\tilde{\tau}_n$, in its first column those of the τ_n. The special choice $\tilde{\tau}_n = \pi_n$ has the effect that $W = S$ is lower triangular.

Next, in analogy to (1.11) we let

$$\mathbf{z} := [\zeta_0, \zeta_1, \zeta_2, \ldots] \quad \text{with} \quad \zeta_n(z) := z^n \tag{1.19}$$

be the row vector containing all monomials. The *moment matrix*, which is a Hankel matrix, can then be written

$$M := [\mu_{m+n}]_{m,n=0}^{\infty} = \varphi(\mathbf{z}^T \mathbf{z}), \tag{1.20}$$

while the *"formal" Gramian matrix* of the polynomials τ_n is defined by

$$N := [\varphi(\tau_m \tau_n)]_{m,n=0}^{\infty} = \varphi(\mathbf{t}^T \mathbf{t}). \tag{1.21}$$

Clearly, there are unit upper triangular matrices Z and R containing the coefficients of the monomials ζ_n and the polynomials τ_n, respectively, in terms of the π_n:

$$\mathbf{z}(z) = \mathbf{p}(z) Z, \tag{1.22}$$

$$\mathbf{t}(z) = \mathbf{p}(z) R. \tag{1.23}$$

Inserting (1.22) into (1.20) and making use of (1.15) yields

$$M = Z^T D Z. \tag{1.24}$$

On the other hand, from (1.15) and (1.23) we get

$$S = R^T D, \tag{1.25}$$

which means that S is obtained by scaling the columns of R^T. Inserting (1.23) and this into (1.21) leads to

$$N = R^T D R = S R. \tag{1.26}$$

Consequently, SR is the LU decomposition (with unit upper triangular matrix) of N. If the diagonal matrix D has a positive diagonal, $Z^T D Z$ and $R^T D R$ are the normalized Cholesky decompositions (with unit triangular matrices) of M and N, respectively. Otherwise, they are (symmetric) LDU decompositions. Our assumptions imply that these decompositions exist. (If D has a positive diagonal, π_n can be renormalized by $\varphi(\pi_n^2) = 1$, so that $D = I$, $S = R^T$, and $N = R^T R = S S^T$, which is then a true Cholesky decomposition. Some of the other formulas then also change slightly; in particular, H becomes symmetric [3, 8].

Finally, from (1.14) and (1.23) there follows that

$$HR = RT. \tag{1.27}$$

This formula is similar to, but different from (1.16) and (1.17), the latter two being equivalent to $H^T DR = DRT$. Consequently,

$$H^T D = DH. \tag{1.28}$$

Moreover, for $\hat{L} := R^{-T}$, (1.27) yields

$$H^T \hat{L} = \hat{L} T^T. \tag{1.29}$$

2 Algorithms for the generic case

In typical applications the polynomials τ_n are known, but the orthogonal polynomials π_n are unknown, and the aim is to compute their recurrence coefficients α_n and β_n from some other information, as, e.g., the moments μ_n (if $\tau_n = \zeta_n$), the modified moments ν_n, or the formal Gramian N. In most cases the problem of determining the recurrence coefficients from the moments μ_n is highly ill-conditioned and therefore prone to roundoff error blow-up. In the case of a positive weight function it has been shown by many experiments and by Gautschi's detailed analysis [3–6], that using modified moments ν_n or the formal Gramian N instead of the moments is usually much more stable. One must expect that often the same is true in the indefinite case, although ill-conditioned examples for the modified process are also known [4].

In this section we discuss a number of algorithms for computing the recurrence coefficients α_n and β_n, i.e., the matrix H, from some of the above mentioned other data. We also discuss how, conversely, the modified moments are obtained from these recurrence coefficients. The first algorithm concerns a partial task applicable in several situations, namely the computation of H from S.

(i) *Computation of H from the diagonal and codiagonal of S.* From (1.16) one sees that $H = S^{-1} T^T S$ is determined by S and T, but from the same relation we can easily derive the well known fact that it suffices to know the diagonal and the first codiagonal of S.

Equating in (1.16) the (m, n)-element yields (with $\sigma_{m,-1} := \sigma_{-1,n} := 0$)

$$\sigma_{m,n+1} + \alpha_n \sigma_{m,n} + \beta_n \sigma_{m,n-1} = \sigma_{m+1,n} + \sum_{k=n}^{m} \tau_{k,m} \sigma_{k,n} \quad (m, n \geq 0). \tag{2.1}$$

If T is tridiagonal, as in (1.6), this reduces to

$$\sigma_{m,n+1} + \alpha_n \sigma_{m,n} + \beta_n \sigma_{m,n-1} = \sigma_{m+1,n} + \alpha'_m \sigma_{m,n} + \beta'_m \sigma_{m-1,n} \quad (m, n \geq 0). \tag{2.1'}$$

For $m < n - 1$ both sides of (2.1) are zero, for $m = n - 1$ one has

$$\beta_n \sigma_{n-1,n-1} = \sigma_{n,n} \quad (n \geq 1), \tag{2.2}$$

while for $m = n$ one obtains

$$\alpha_n \sigma_{n,n} + \beta_n \sigma_{n,n-1} = \sigma_{n+1,n} + \tau_{n,n} \sigma_{n,n} \quad (n \geq 0). \tag{2.3}$$

Given, say, the first $\overline{m}+1$ elements on the diagonal and the first codiagonal of S, (2.2) and (2.3) allow us to compute $\alpha_0, \ldots, \alpha_{\overline{m}}$ and $\beta_1, \ldots, \beta_{\overline{m}}$ recursively. By substituting β_n in (2.3) according to (2.2) one gets the following well known procedure [4, 13, 16]. For $n = 0, 1, \ldots, \overline{m}$:

$$\beta_n := \frac{\sigma_{n,n}}{\sigma_{n-1,n-1}} \quad (n \neq 0). \tag{2.4a}$$

$$\alpha_n := \tau_{n,n} + \frac{\sigma_{n+1,n}}{\sigma_{n,n}} - \frac{\sigma_{n,n-1}}{\sigma_{n-1,n-1}}. \tag{2.4b}$$

(ii) *The modified Chebyshev algorithm for generating S recursively from the modified moments* ν_m. If the modified moments, i.e., the first column of S are given, then, using an idea that goes back to Chebyshev, one can build up S simultaneously with the computation (2.4) by using relation (2.1) once again. Starting with (1.9),

$$\sigma_{m,0} := \nu_m, \quad m = 0, 1, \ldots, 2\overline{m}+1, \tag{2.5a}$$

S is built up recursively by applying formula (2.1) solved for $\sigma_{m,n+1}$: For $n = 0, 1, \ldots, \overline{m}-1$ one computes simultaneously β_n and α_n according to (2.4) and

$$\begin{aligned} \sigma_{m,n+1} &:= \sigma_{m+1,n} - \alpha_n \sigma_{m,n} - \beta_n \sigma_{m,n-1} + \sum_{k=n}^{m} \tau_{k,m} \sigma_{k,n} \\ & m = n+1, n+2, \ldots, 2\overline{m}-n. \end{aligned} \tag{2.5b}$$

(Again, the sum reduces to two terms if T is tridiagonal.) Given the $2\overline{m}+2$ data $\nu_0, \ldots, \nu_{2\overline{m}+1}$, one can in this way compute $2\overline{m}+2$ coefficients $\alpha_0, \ldots, \alpha_{\overline{m}}$ and $\beta_0, \ldots, \beta_{\overline{m}}$, where, as before, $\beta_0 := \nu_0 = \sigma_{0,0}$; cf. Fig. 1.

This modified Chebyshev algorithm is due to Sack and Donovan [13] and Wheeler [16]. Usually it is only formulated for tridiagonal T, but already Sack and Donovan pointed out that it generalizes easily to an upper Hessenberg matrix. If the modified moments are replaced by the ordinary moments, one has $\tau_{m,m} = 1$, $\tau_{k,m} = 0$ ($\forall k \neq m$), and the algorithm specializes to the classical Chebyshev algorithm. For a positive measure λ both algorithms have been extensively discussed and analyzed by Gautschi [3–6]. Note that the modified moments, with which one starts, are the entries of the first columns of both S and N, since $\tau_0(z) = \zeta_0(z) \equiv 1$.

(ii') *The inverse modified Chebyshev algorithm for computing S from H*. If, on the other hand, the recurrence coefficients α_n, β_n (including $\beta_0 = \nu_0$), and $\tau_{m,n}$ ($m \leq n$) are known, say, for $n = 0, \ldots, \overline{m}$, the formulas (2.4) and (2.5) can be inverted in order to compute the diagonal and the first codiagonal of S, then to build up the whole matrix S from these elements, and, finally, to compute the modified moments $\nu_0, \ldots, \nu_{2\overline{m}+1}$: For $n = 0, 1, \ldots, \overline{m}$,

$$\sigma_{n,n} := \beta_n \sigma_{n-1,n-1} \quad (n \neq 0), \tag{2.6a}$$

$$\sigma_{n+1,n} := (\alpha_n - \tau_{n,n})\sigma_{n,n} + \beta_n \sigma_{n,n-1}, \tag{2.6b}$$

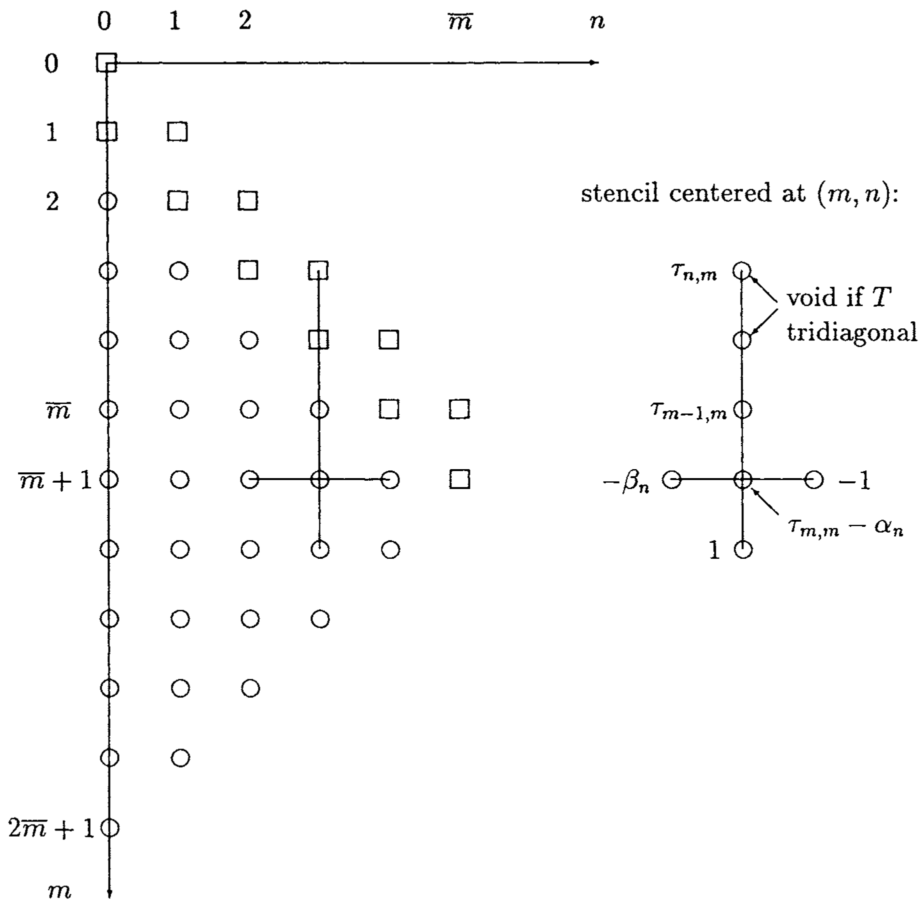

Fig. 2. The modified Chebyshev algorithm for generating S from the modified moments ν_m according to (2.5). The elements marked by a square are used to build up H according to (2.4).

and, for $n = 0, 1, \ldots, \min\{m-1, 2\overline{m}-m\}$ and $m = 1, 2, \ldots, 2\overline{m}$,

$$\sigma_{m+1,n} := \sigma_{m,n+1} + \alpha_n \sigma_{m,n} + \beta_n \sigma_{m,n-1} - \sum_{k=n}^{m} \tau_{k,m} \sigma_{k,n}. \tag{2.6c}$$

Note that $\sigma_{0,0} = \beta_0$ is used in (2.6a). Note also that (2.6a) and (2.6b) can be executed independently of (2.6c).

It should be mentioned that the *qd algorithm* of Rutishauser [12] can fulfill the same two tasks as the unmodified Chebyshev algorithm and its inverse: Either construct $\{\mu_n\}_{n=0}^{2\overline{m}+1}$ from $\{\alpha_n, \beta_n\}_{n=0}^{\overline{m}}$, or the reverse. However, as has been noted by Wheeler [16], the qd algorithm is more likely to break down due to degeneracies; in fact, it requires $O(\overline{m}^2)$ nonvanishing Hankel determinants, while the Chebyshev algorithm requires only $O(\overline{m})$ such determinants, cf. Gragg [9, p. 214]. Furthermore, Sack and Donovan [13] point out that the qd algorithm involves $O(\overline{m}^2)$ divisions, as opposed to the $2\overline{m}+1$ needed here.

(iii) *LU decomposition of N into* $N = SR$, cf. (1.26) [3, 8]. From an $(\overline{m}+2)\times(\overline{m}+1)$ principal submatrix of N one obtains an $(\overline{m}+2)\times(\overline{m}+1)$ principal submatrix of S, and from the latter one can either compute $\alpha_0,\ldots,\alpha_{\overline{m}}$ and $\beta_0,\ldots,\beta_{\overline{m}}$ by using (2.4) or the additional modified moments $\nu_{\overline{m}+2},\ldots,\nu_{2\overline{m}+1}$. This is a straightforward procedure, but its cost is $O(\overline{m}^3)$ even if T is tridiagonal. Moreover, it is a disadvantage that the full matrix N is required. (As shown by Gautschi [3], it is, however, possible to build up the symmetric matrix N from its first row and column.) The following algorithm, which makes use of the LU decomposition and the recurrence (2.6c), only requires to know the diagonal and the first codiagonal of N.

(iv) *Recursive generation of S from the diagonal and the subdiagonal of N.* Let $N = (\nu_{m,n})_{m,n=0}^{\infty}$, and note that $N = SR = SD^{-1}S^T$ by (1.25) and (1.26). Assuming $\nu_{m,m}$ and $\nu_{m+1,m}$ are known for $m = 0,\ldots,\overline{m}$, one can find the leading $(\overline{m}+2)\times(\overline{m}+1)$ submatrix of S by making use of (2.4), (2.6c), and the above formula $N = SD^{-1}S^T$: Once $\sigma_{k,k}$ and $\sigma_{m,k}$ are known for $k = 0,\ldots,m-1$, the latter formula yields $\sigma_{m,m}$. Then, assuming the α_k and β_k are also known, (2.6c) allows us to find $\sigma_{m+1,k}$, $k = 0,\ldots,m-1$, and with those numbers, the above formula then yields $\sigma_{m+1,m}$. Finally, (2.4) gives the new β_m and α_m. Hence, one can build up S from its diagonal and its first codiagonal.

In summary, we compute for $m = 0,1,\ldots,\overline{m}$:

$$\sigma_{m,m} := \nu_{m,m} - \sum_{k=0}^{m-1} \sigma_{m,k}^2 \sigma_{k,k}^{-1}, \tag{2.7a}$$

$$\sigma_{m+1,n} := \sigma_{m,n+1} + \alpha_n \sigma_{m,n} + \beta_n \sigma_{m,n-1} - \sum_{k=n}^{m} \tau_{k,m}\sigma_{k,n}, \quad n = 0,\ldots,m-1, \tag{2.7b}$$

$$\sigma_{m+1,m} := \nu_{m+1,m} - \sum_{k=0}^{m-1} \sigma_{m,k}\sigma_{m+1,k}\sigma_{k,k}^{-1}, \tag{2.7c}$$

and β_m and α_m according to (2.4).

Moreover, (2.6c) can afterwards be executed for $m = \overline{m}+1,\ldots,2\overline{m}$ for building up the lower half of the triangle in Fig. 1 if needed. Hence, again, the $2m+2$ data $\{\nu_{m,m},\nu_{m+1,m}\}_{m=0}^{\overline{m}}$ yield $2\overline{m}+2$ coefficients $\alpha_0,\ldots,\alpha_{\overline{m}}$ and $\beta_0,\ldots,\beta_{\overline{m}}$ and $2m+2$ modified moments $\nu_0,\ldots,\nu_{2\overline{m}+1}$. Note that although one may only be interested in the diagonal and the subdiagonal of S one has to build up the whole matrix since in (2.7a) and (2.7c) other elements of the matrix are needed.

This algorithm is particularly useful in applications where the parameters of an iterative method for solving a linear system are adapted according to approximate spectrum information derived from the residuals that have been generated in previous steps [7]. The residuals $r_0, r_1,\ldots,r_{\overline{m}}$ allow us to compute exactly the $(\overline{m}+2)\times(\overline{m}+1)$ principal submatrix of N, since $\nu_{m,n} = (r_m, r_n)$. By algorithms (iii) and (iv) one can then compute a total of $2\overline{m}+2$ recurrence coefficients, while the modified Chebyshev algorithm would only yield half as many. Algorithm (iv) has over algorithm (iii) the important advantage that only two successive residual vectors need to be stored.

It should be mentioned at this point, however, that when the τ_n are shifted and scaled Chebyshev polynomials (as, e.g., in the Chebyshev iteration method), the relation $\cos(k+l)\theta = 2\cos k\theta \cos l\theta - \cos(k-l)\theta$ allows one to compute v_{2m} and v_{2m+1} from $v_{m,m}$ and $v_{m+1,m}$, see [7]. Hence, in this case, the above mentioned advantage of this algorithm becomes inessential.

(v) *Computation of H from the diagonal and codiagonal of S^{-1}.* Let $L := [\lambda_{m,n}]_{m,n=0}^{\infty} := S^{-1}$. By equating the (m,n)-element in (1.17), one obtains in complete analogy to (2.1)

$$\lambda_{m,n-1} + \sum_{k=n}^{m} \lambda_{m,k}\tau_{n,k} = \lambda_{m-1,n} + \alpha_m \lambda_{m,n} + \beta_{m+1}\lambda_{m+1,n}\,. \tag{2.8}$$

Here, both sides are again zero if $m < n-1$. For $m = n-1$ and $m = n$ there follows, by solving for α_n and β_n:

$$\beta_n := \frac{\lambda_{n-1,n-1}}{\lambda_{n,n}} \quad (n \geq 1)\,, \tag{2.9a}$$

$$\alpha_n := \tau_{n,n} + \frac{\lambda_{n,n-1}}{\lambda_{n,n}} - \frac{\lambda_{n+1,n}}{\lambda_{n+1,n+1}}\,. \tag{2.9b}$$

These formulas were used by Gautschi [3] and on Gautschi's proposal also in the published version of Golub and Welsch [8] (but there in the unmodified case only) for deriving the formulas (2.4). This is astonishing, since our direct derivation of (2.4) is as easy as the one of (2.9). (In the preprint of [8] determinantal relationships were used instead for the derivation.) Formulas (2.9) could be applied after an "inverse LU decomposition" of N, $LN\hat{L}^T = I$, where $\hat{L}^T := R^{-1} = L^T D$ is unit upper triangular. However, (2.8) and (2.9) do not lead to a "Chebyshev-type algorithm", where H can be built up concurrently with L. But the following approach, although very similar, does.

(vi) *A modified "Chebyshev-type algorithm" for computing H from N.* Let now $\hat{L} := [\hat{\lambda}_{m,n}]_{m,n=0}^{\infty} := DL = DS^{-1} = R^{-T}$. The matrix relation (1.29) yields

$$\hat{\lambda}_{m,n-1} + \sum_{k=n}^{m} \hat{\lambda}_{m,k}\tau_{n,k} = \beta_m \hat{\lambda}_{m-1,n} + \alpha_m \hat{\lambda}_{m,n} + \hat{\lambda}_{m+1,n}\,. \tag{2.10}$$

Here, the formulas obtained by setting $m = n$ and $m = n+1$ can be simplified by noticing that $\hat{\lambda}_{n,n} = 1$ and then solved for α_n and β_n. But again they do not yet lead to a "Chebyshev-type algorithm". However, from (1.26) we know that $\hat{L}N = S^T$, so that

$$\sigma_{m,m} = \sum_{k=0}^{m} \hat{\lambda}_{m,k} v_{m,k}, \qquad \sigma_{m+1,m} = \sum_{k=0}^{m} \hat{\lambda}_{m,k} v_{m+1,k}\,. \tag{2.11}$$

Therefore, if $\hat{\lambda}_{l,k}$ is known for $k \leq l \leq m$, then (2.11) and (2.4) allow us to compute α_m and β_m, and then (2.10) yields $\hat{\lambda}_{m+1,n}$ $(n = 0, 1, \ldots, m)$. Hence, here one can indeed build up $\hat{L}$ row by row simultaneously with H^T by an $O(m^2)$ process. The version for ordinary moments of this algorithm appears in disguised form in Wall [15, pp. 196–200] and, explicitely, in Gragg [9, p. 215], but its origin lies probably further back. In contrast to the modified Chebyshev algorithm the Gramian N is required, and not only its first column.

(vii) *Transforming modified moments into another set of modified moments.* Relations (2.1), (2.8) and (2.10) are all based on matrix identities in which an unknown matrix (S, L or $\hat{L}$) is multiplied from left and right, respectively, by two other matrices, namely the known matrix T^T and the matrix H (or its transposed) that has to be determined. In each case, the first unknown matrix is triangular and the two others are unit Hessenberg. But, in the same way one can use (1.18), $W\tilde{T} = T^T W$, to derive a recurrence for building up the full matrix $W = [\omega_{m,n}]_{m,n=0}^{\infty}$ from its first column containing the modified moments of the set $\{\tau_n\}$. In fact, (1.18) yields

$$\omega_{m,n+1} + \sum_{k=0}^{n} \tilde{\tau}_{k,n}\omega_{m,k} = \omega_{m+1,n} + \sum_{k=0}^{m} \tau_{k,m}\omega_{k,n}\,,$$

which corresponds to a stencil that goes in the mth row from 0 to $n+1$ and in nth column from 0 to $m+1$. Given $\omega_{m,0} = \nu_m$, $m = 0, \ldots, \overline{m}$, this stencil allows us to compute the triangle $\omega_{m,n}$, $m + n \leq \overline{m}$, and, hence, in particular, $\omega_{0,n} = \tilde{\nu}_n$, $n = 0, \ldots, \overline{m}$, the first $\overline{m}+1$ modified moments of the set $\{\tilde{\tau}_n\}$. This procedure is due to Wheeler [16] who proposed it in the case of tridiagonal matrices T and $\tilde{T}$.

3 Matrix relations for the general case

Let now φ be an arbitrary complex linear functional defined on the linear space of all polynomials, and let $\mu_n := \varphi(\zeta_n) := \varphi(z^n)$ ($n \in \mathbf{N}$) be the associated *moments.* (Hence, φ, which is uniquely determined by these moments, may or may not be given by a measure λ as in (1.1).) There exists a finite or infinite sequence $\{n_j\}_{j=0}^{J}$ of indices (with $0 = n_0 < n_1 < n_2 < \ldots$ and $J \leq \infty$) for which a *regular formal orthogonal polynomial of the first kind (regular FOP1)* π_{n_j} exists, i.e., a uniquely determined monic polynomial of degree n_j with the property (1.3) [2, 11, 14]. A full sequence $\{\pi_n\}$ of monic FOP1s is obtained by defining

$$\pi_n(z) := z^{n-n_j}\pi_{n_j}(z) \text{ if } n_j < n < n_{j+1}. \tag{3.1}$$

(If $J < \infty$, i.e., if there is only a finite number of regular FOP1s, we set $n_{J+1} := \infty$.) The more general definition

$$\tilde{\pi}_n(z) := \omega_n(z)\pi_{n_j}(z) \text{ if } n_j \leq n < n_{j+1}, \tag{3.1'}$$

in which ω_n is an arbitrary monic polynomial of degree $n - n_j$ yields an *equivalent* sequence of FOP1s. We let also

$$h_j := n_{j+1} - n_j, \quad h'_j := \left\lfloor \frac{h_j - 1}{2} \right\rfloor \quad (0 \leq j \leq J \leq \infty). \tag{3.2}$$

($h_J = h'_J = \infty$ if $J < \infty$.) Then, for $n_j < n \le n_j + h'_j$, the polynomials π_n and $\tilde{\pi}_n$ still satisfy (1.3), but they are no longer uniquely characterized by this property. Moreover, $\varphi(\pi_n \pi_n) = \varphi(\tilde{\pi}_n \tilde{\pi}_n) = 0$ for these *singular* FOP1s. In contrast, for $n_j + h'_j < n < n_{j+1}$, π_n and $\tilde{\pi}_n$ no longer satisfy (1.3), and we call them therefore *deficient* FOP1s.

The orthogonality properties of the sequence $\{\pi_n\}$ and an immediate consequence concerning the moment matrix are summarized in the following matrix result [2, 9, 10, 11]:

Theorem 3.1. *Let* $\mathbf{p} := [\pi_0, \pi_1, \pi_2, \ldots]$ *as in (1.11). Then*

$$\varphi(\mathbf{p}^T \mathbf{p}) = D := \text{block diag } [D_0, D_1, D_2, \ldots] \tag{3.3}$$

is a block diagonal matrix with square blocks of size h_j $(j = 0, 1, \ldots, J)$ *which are right lower triangular Hankel matrices*

$$D_j := \begin{bmatrix} & & & & \delta_j \\ & & & \delta_j & \star \\ & & \cdot^{\cdot^{\cdot}} & \cdot^{\cdot^{\cdot}} & \vdots \\ & \delta_j & \cdot^{\cdot^{\cdot}} & & \star \\ \delta_j & \star & \cdots & \star & \star \end{bmatrix}, \tag{3.4}$$

with $\delta_j := \varphi(z^{h_j - 1} \pi_{n_j}^2)$. *If* $J < \infty$, D_J *is the infinite zero matrix and* $\delta_J := 0$.

Moreover, if $\mathbf{z}(z) = \mathbf{p}(z)Z$ *as in (1.22), and if* $M = \varphi(\mathbf{z}^T \mathbf{z})$ *denotes the moment matrix, then*

$$M = Z^T D Z \tag{3.5}$$

is a block LDU decomposition (with unit upper triangular Z).

For any equivalent sequence $\{\tilde{\pi}_n\}$ *of FOP1s there holds* $\tilde{\mathbf{p}}(z) = \mathbf{p}(z) W$ *with a block diagonal matrix* W, *whose diagonal blocks* W_j *have the same size as those of D and are unit upper triangular matrices. Conversely, any such matrix* W *defines an equivalent sequence of FOP1s. For such a sequence there holds* $\varphi(\tilde{\mathbf{p}}^T \tilde{\mathbf{p}}) = \tilde{D} = W^T D W$, *where* $\tilde{D}$ *has the same structure as D. There is a particular equivalent sequence* $\{\hat{\pi}_n\}$ *for which the diagonal blocks* $\tilde{D}_j$ *of* $\tilde{D}$ *become the antidiagonal matrices* $\hat{D}_j$ *whose antidiagonal elements are* δ_j.

Another result we need to cite is the one on the recurrence relations for the sequence $\{\pi_n\}$. We give it also directly in matrix form [2, 9, 10, 11]:

Theorem 3.2. *There holds* $z\mathbf{p}(z) = \mathbf{p}(z)H$, *where* H *is an infinite irreducible block tridiagonal upper Hessenberg matrix,*

$$H := \begin{bmatrix} A_0 & B_1 & & \\ C_0 & A_1 & B_2 & \\ & C_1 & A_2 & \ddots \\ & & \ddots & \ddots \end{bmatrix}, \tag{3.6}$$

in which, for $0 \le i < J$, A_i *is a companion matrix of order* h_i,

$$(3.7)\qquad A_i := \begin{bmatrix} 0 & & & \alpha_{0,i} \\ 1 & \ddots & & \vdots \\ & \ddots & 0 & \alpha_{h_i-2,i} \\ & & 1 & \alpha_{h_i-1,i} \end{bmatrix} = \begin{bmatrix} 0\cdots 0 & \\ & a_i \\ I & \end{bmatrix},$$

while, for $0 < i < J$, B_i *and* C_{i-1} *are rectangular rank-one matrices of size* $h_{i-1} \times h_i$ *and* $h_i \times h_{i-1}$, *respectively,*

$$(3.8)\qquad B_i := \begin{bmatrix} 0 & \cdots & 0 & \beta_i \\ 0 & \cdots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & 0 \end{bmatrix}, \quad C_{i-1} := \begin{bmatrix} 0 & \cdots & 0 & 1 \\ 0 & \cdots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & 0 \end{bmatrix},$$

with $\beta_i \neq 0$. *If* $J < \infty$, A_J *is the infinite forward shift matrix, while* $B_J = C_{J-1}^T = 0$ *is the* $h_{J-1} \times \infty$ *zero matrix.*

For an equivalent sequence of FOP1s, $\tilde{\mathbf{p}}(z) = \mathbf{p}(z)W$, *there holds* $z\tilde{\mathbf{p}}(z) = \tilde{\mathbf{p}}(z)\tilde{H}$, *where* $\tilde{H} = W^{-1}HW$ *is a block tridiagonal upper Hessenberg matrix with the same off-diagonal blocks* B_j *and* C_j *as* H *and with diagonal blocks* $\tilde{A}_j = W_j^{-1}A_jW_j$ *which are unit upper Hessenberg.*

Again we consider an arbitrary other sequence $\{\tau_n\}$ of monic polynomials, as in (1.5), and the corresponding matrix $S = [\sigma_{m,n}]$, cf. (1.8). Setting $\mathbf{t} := (\tau_0, \tau_1, \tau_2, \ldots)$ we have $z\mathbf{t}(z) = \mathbf{t}(z)T$ as in (1.14) and $\mathbf{t}(z) = \mathbf{p}(z)R$ with a unit upper triangular marix R as in (1.23). The matrix T may be an arbitrary upper unit Hessenberg matrix, although in practice we are still most interested in the case where T is tridiagonal. The matrix relations (1.16), (1.25), (1.26), and (1.27) remain valid:

$$(3.9)\qquad S = R^TD, \quad N = R^TDR = SR, \quad SH = T^TS, \quad HR = RT.$$

The same set of equations, with tildes everywhere except on N and T, holds for any equivalent set of FOP1s. (However, in (1.17) and (1.29) L and $\hat{L}$ may in general not exist, so that one would have to turn to corresponding relations with nonsingular matrices of finite order. We will not explore these relations further here, however.) Let us partition the matrix R into blocks of size $h_i \times h_j$ $(i, j = 0, \ldots, J)$, i.e., exactly in the same way D and H are partitioned: $R = (R_{i,j})_{i,j=0}^J$, where $R_{i,j}$ is a $h_i \times h_j$ submatrix. Since R is unit upper triangular and D is block diagonal, $S = R^TD$ is block lower triangular, and its diagonal blocks $S_{j,j}$ are products of a unit lower triangular submatrix $R_{j,j}^T$ and a matrix D_j from (3.4). Hence, $S_{j,j}$ is (as D_j) a right lower triangular matrix, and it has the same constant antidiagonal elements $\delta_j \neq 0$ as D_j $(j < J)$; if $J < \infty$, the last diagonal block $S_{J,J}$ is the infinite zero matrix:

$$(3.10)\quad S = \begin{bmatrix} S_{0,0} & & & \\ S_{1,0} & S_{1,1} & & \\ S_{2,0} & S_{2,1} & S_{2,2} & \\ \vdots & & & \ddots \end{bmatrix} \quad \text{with} \quad S_{j,j} = \begin{bmatrix} & & & & \delta_j \\ & & & \delta_j & \star \\ & & \iddots & \iddots & \vdots \\ & \delta_j & \iddots & & \star \\ \delta_j & \star & \cdots & \star & \star \end{bmatrix}.$$

Next, also T is partitioned in the same way, and we write it as

$$
\tag{3.11} T = \begin{bmatrix} T_{0,0} & T_{0,1} & T_{0,2} & \cdots \\ C_0 & T_{1,1} & T_{1,2} & \cdots \\ & C_1 & T_{2,2} & \cdots \\ & & \ddots & \ddots \end{bmatrix} .
$$

Here C_i is the same matrix as in (3.8), and $T_{i,i}$ is unit upper Hessenberg but in general not a companion matrix. If the polynomials τ_n satisfy a three-term recurrence, $T_{i,i}$ is tridiagonal, $T_{i-1,i}$ has its only nonzero element in its lower left corner, and $T_{i-1,j} = 0$ for $j > i$.

4 Algorithms for the general case

Our aim is again to compute the recurrence coefficients, i.e., the matrix H, of the unknown polynomials π_n. The polynomials τ_n are assumed to be known along with some additional information, e.g., the formal Gramian N or the modified moments $\nu_n = \nu_{n,0}$. Results for the moment matrix M and the unmodified moments are included as the special case $N := M$, $\mathbf{t} := \mathbf{z}$.

Again we start by deriving formulas for computing H from S and T. Then, we discuss methods for computing S.

(i) *Computation of H from the block diagonal and block codiagonal of S.* The relation $SH = T^T S$ translates now in analogy to (2.1) into

$$
\tag{4.1} S_{i,j+1}C_j + S_{i,j}A_j + S_{i,j-1}B_j = C_i^T S_{i+1,j} + T_{i,i}^T S_{i,j} + \sum_{l=j}^{i-1} T_{l,i}^T S_{l,j} \quad (i, j \geq 0),
$$

with $S_{i,-1} := 0$. For $i < j-1$ both sides are again zero, for $i = j-1$ there follows

$$
\tag{4.2} S_{j-1,j-1}B_j = C_{j-1}^T S_{j,j} \quad (j \geq 1),
$$

and for $i = j$ one has

$$
\tag{4.3} S_{j,j}A_j + S_{j,j-1}B_j = C_j^T S_{j+1,j} + T_{j,j}^T S_{j,j} \quad (j \geq 0).
$$

Since $S_{j,j}$ is nonsingular for $j < J$, one obtains by multiplication from the left with $S_{j,j}^{-1}$ indeed formulas generalizing (2.4). (These formulas, with tildes, hold also when equivalent FOP1s are used.) However, since we know the structure of A_j and B_j, these formulas contain a lot of redundant information. All that has to be determined are the order h_j of A_j, its last column a_j, and the nontrivial element β_j of B_j, cf. (3.7) and (3.8). It is easily verified that $S_{j,j}^{-1}$ is left upper triangular and that its antidiagonal elements are all δ_j^{-1}. Therefore, from (4.2) one obtains

$$
\tag{4.4} \beta_j = \delta_{j-1}^{-1}\delta_j .
$$

(4.3) is more complicated. We partition $S_{j,j-1}$ and $S_{j,j}$ according to

$$(4.5) \qquad S_{j,j-1} = \begin{bmatrix} & \star & \cdots & \star & \sigma'_{j-1} \\ s'_j & \star & \cdots & \star & \star \\ & \vdots & & \vdots & \vdots \\ & \star & \cdots & \star & \star \end{bmatrix}, \qquad S_{j,j} = \begin{bmatrix} & & \star & s_j \\ & \cdot^{\cdot^{\cdot}} & \vdots & \\ \star & \cdots & \star & \end{bmatrix},$$

and we denote the last standard basis vector in $\mathbf{C}^{h_j}$ by e_{h_j}. (If $h_{j-1} = 1$, σ'_{j-1} is the first component of s'_j.) Then, due to the special structure of C_j^T and B_j, (4.3) yields

$$(4.6) \qquad a_j = S_{j,j}^{-1} \left[T_{j,j}^T s_j + e_{h_j} \sigma'_j - s'_j \beta_j \right] .$$

Moreover, if we return to the single elements of $S = (\sigma_{m,n})$, $H = (\gamma_{m,n})$, and $T = (\tau_{m,n})$, the identity $SH = T^T S$ still yields, in analogy to (2.1),

$$(4.7) \qquad \sigma_{m,n+1} + \sum_{l=0}^{n} \sigma_{m,n-l} \gamma_{n-l,n} = \sigma_{m+1,n} + \sum_{l=0}^{m} \tau_{m-l,m} \sigma_{m-l,n} \,.$$

Here, if $n_j \le n = n_j + k < n_{j+1} - 1$, then $\gamma_{n-l,n} = 0$ for all $l \ne -1$ and $\sigma_{m-l,n} = 0$ if $m - l < n_{j+1} - k - 1$:

$$(4.8a) \qquad \sigma_{m,n+1} = \sigma_{m+1,n} + \sum_{l=0}^{m+k-n_{j+1}+1} \tau_{m-l,m} \sigma_{m-l,n} \,.$$

Hence, in this case, the left branch of the stencil in Fig. 1 is missing, and, unless T is tridiagonal, the upper branch extends up to the first nonvanishing element in the nth column of S; this element lies on the antidiagonal of $S_{j,j}$, cf. Fig. 2. On the other hand, if $n = n_{j+1} - 1$,

$$(4.8b) \quad \sigma_{m,n_{j+1}} + \sum_{l=1}^{h_j} \sigma_{m,n_{j+1}-l} \alpha_{h_j-l,j} + \sigma_{m,n_{j-1}} \beta_j = \sigma_{m+1,n_{j+1}-1} + \sum_{k=n_j}^{m} \tau_{k,m} \sigma_{k,n_{j+1}-1} \,.$$

This stencil has a left branch, which, in general, is longer than the one in Fig. 1.

(ii) *The non-generic modified Chebyshev algorithm for generating S recursively from the modified moments ν_m.* Assume $\nu_0, \ldots, \nu_{2\overline{m}+1}$ are given. Starting with (2.5a) one can again build up the elements $\sigma_{m,n}$ in the triangle $m + n \le 2\overline{m} + 1$ of S from left to right, using now (4.4), (4.6), (4.8a), and (4.8b). Note that the results of (4.4) and (4.6) are only needed in (4.8b), i.e., for computing $\sigma_{m,n_{j+1}}$, where $m \ge n_{j+1}$. At this moment, $S_{j,j}$ (containing s_j), $S_{j,j-1}$ (containing s'_j) and the first row of $S_{j+1,j}$ (containing σ'_j) have all been computed. Also, the size h_j of $S_{j,j}$ and the constant value of its antidiagonal elements are known.

(ii') *The non-generic inverse modified Chebyshev algorithm for computing S from H.* Given the first block $S_{0,0}$ of S and the matrices H and T with the recurrence coefficients one can again build up S from top to bottom by capitalizing upon the relation $SH = T^T S$. The structure of S, which is determined by the diagonal blocks $S_{j,j}$, is known from H. The antidiagonal

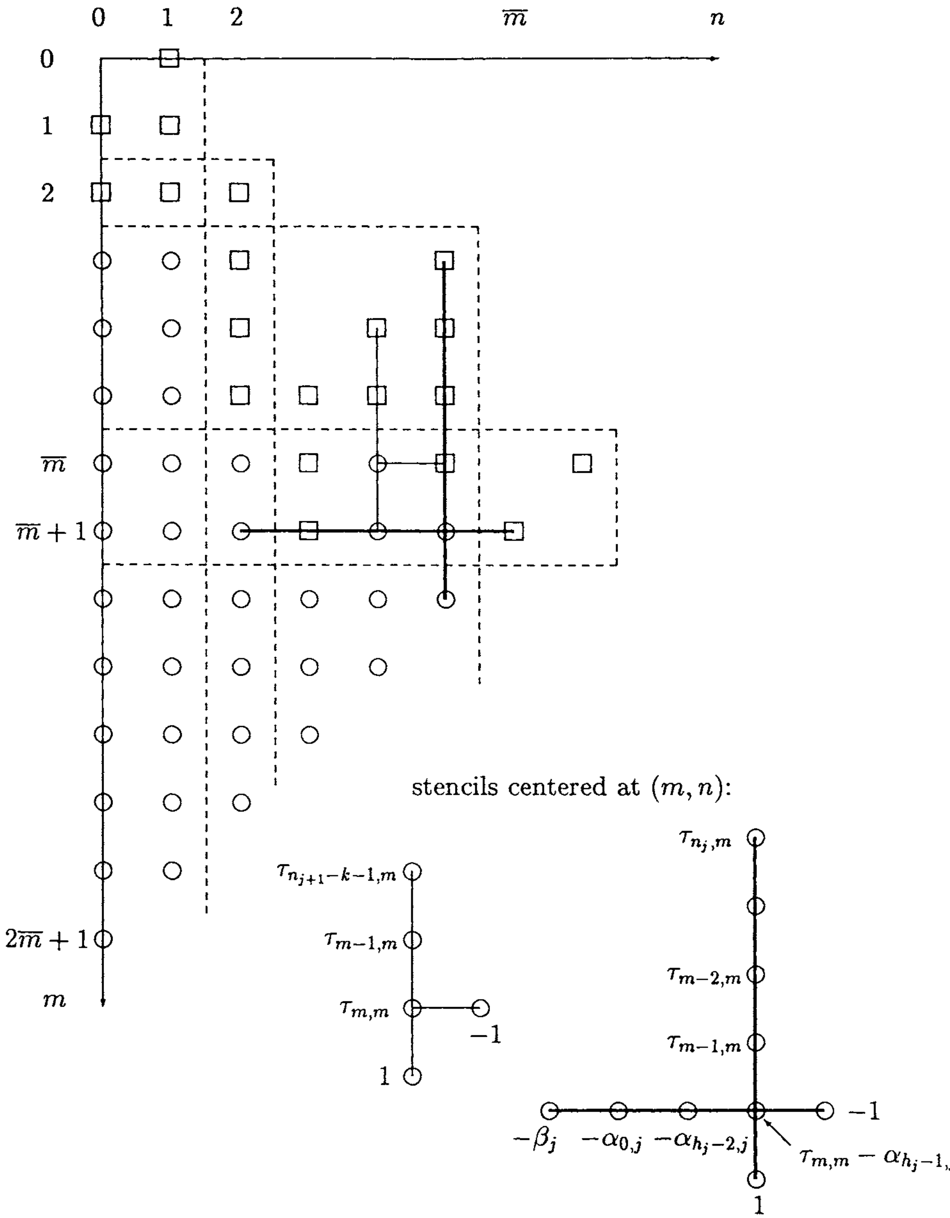

Fig. 1. The non-generic modified Chebyshev algorithm with the stencils (4.8a) and (4.8b) (bold). The elements marked by a square are used to build up H according to (4.4) and (4.6).

elements δ_j of S_j are recursively obtained from (4.4): $\delta_j = \delta_{j-1}\beta_j$ $(j = 1, 2, \ldots J - 1)$. Then (4.8a) and (4.8b), both solved for $\sigma_{m+1,n}$, are used to proceed downward. One must be aware that these two relations hold for all $m \geq 0$, hence also if some of the occuring elements belong to blocks $S_{i,j}$ with $i < j$ (i.e., to blocks above the block diagonal) and are therefore known to be zero. In particular, if $n_j \le m < n_{j+1}$ and $n = n_{j+1} - 1$ (i.e., if $\sigma_{m,n}$ belongs to the last column of $S_{j,j}$), then $\sigma_{m,n_{j+1}} = 0$ in (4.8b).

(iii) *Block LU decomposition* $N = SR$: In the block LU decomposition of N there is some freedom left in factorizing the block pivots $N_{j,j}^{(j)} = N_{j,j} - \sum_{k=0}^{j-1} S_{j,k}R_{k,j}$, which in view of $N_{j,j}^{(j)} = S_{j,j}R_{j,j}$ are known to be right lower triangular. Here, in order to obtain S, R, and, subsequently, by (4.4) and (4.6), the matrix H with the structure (3.6)–(3.8) one would need $S_{j,j}$ lower right triangular and $R_{j,j}$ unit upper (right) triangular, and additionally, one would have to aim for the companion matrix structure of the diagonal blocks A_j of H. It seems to be too complicated to attain all these goals at once, hence, we suggest to aim first at a matrix $\tilde{H}$ of an equivalent set of FOP1s $\tilde{\pi}_n$.

We recall that $\tilde{\mathbf{p}} = \mathbf{p}W$, where W is block diagonal with unit upper triangular blocks W_j of order h_j. These blocks W_j can otherwise be chosen arbitrarily. Since the related matrices $\tilde{S}, \tilde{R}, \tilde{H}$ satisfy

$$\tag{4.9} \tilde{S} = SW, \quad \tilde{R} = W^{-1}R, \quad \tilde{H} = W^{-1}HW,$$

$N = \tilde{S}\tilde{R}$ is also a block LU decomposition of N and the diagonal blocks $\tilde{A}_j$ of the resulting $\tilde{H}$ are similar to those of H: $\tilde{A}_j = W_j^{-1}A_jW_j$. By choosing $W_j = S_{j,j}^{-1}\hat{E}_j$, where $\hat{E}_j$ is any matrix of the form (3.4), we obtain $\tilde{S}_{j,j} = \hat{E}_j$. In particular, we can choose $\tilde{S}_{j,j}$ antidiagonal:

$$\tag{4.10} \tilde{S}_{j,j} = \hat{E}_j := \begin{bmatrix} 0 & \cdots & 0 & \delta_j \\ \vdots & \cdot^{\cdot^{\cdot}} & \cdot^{\cdot^{\cdot}} & 0 \\ 0 & \cdot^{\cdot^{\cdot}} & \cdot^{\cdot^{\cdot}} & \vdots \\ \delta_j & 0 & \cdots & 0 \end{bmatrix}.$$

Hence, we suggest to compute the block LU decomposition $N = \tilde{S}\tilde{R}$ of N that is characterized by a splitting of the block pivots $N_{j,j}^{(j)} = N_{j,j} - \sum_{k=0}^{j-1} \tilde{S}_{j,k}\tilde{R}_{k,j}$ into $N_{j,j}^{(j)} = \hat{E}_j\tilde{R}_{j,j}$. This is just an antidiagonal scaling of the block pivot and yields a unit upper triangular $\tilde{R}_{j,j}$.

Once the decomposition $N = \tilde{S}\tilde{R}$ has been computed, the formulas (4.4) and (4.3) allow us to find $\tilde{B}_{j-1} = B_{j-1}$ and $\tilde{A}_j$. For the typically small matrices $\tilde{A}_j$ one can finally determine the similar companion matrices A_j and the corresponding similarity transformation matrices W_j.

An alternative is to find first instead of $\tilde{S}$ the matrix $\hat{S} := \varphi(\mathbf{t}^T\hat{\mathbf{p}})$ that belongs to the equivalent set of FOP1s $\hat{\pi}_n$ with the property that

$$\tag{4.11} \varphi(\hat{\mathbf{p}}^T\hat{\mathbf{p}}) = \hat{D} := \text{block diag}\,[\hat{E}_0, \hat{E}_1, \hat{E}_2, \ldots]$$

with the antidiagonal blocks $\hat{E}_j$ of (4.10). In view of $\hat{D}^T = \hat{D}$ there holds $N = \hat{S}\hat{D}^{-1}\hat{S}^T$ as in the generic case, and there follows that

$$\tag{4.12} \hat{S}_{j,j}\hat{E}_j^{-1}\hat{S}_{j,j}^T = N_{j,j}^{(j-1)} := N_{j,j} - \sum_{k=0}^{j-1} \hat{S}_{j,k}\hat{E}_k^{-1}\hat{S}_{k,j}^T.$$

Hence, here the block pivot $N_{j,j}^{(j-1)}$, which is symmetric and right lower triangular with antidiagonal elements δ_j, must be factorized into a right lower triangular matrix $\hat{S}_{j,j}$ with the same antidiagonal, the antidiagonal matrix $\hat{E}_j^{-1}$, and the transposed of $\hat{S}_{j,j}$. It is shown in [11] that one can further require that $\hat{S}_{j,j}$ is antisymmetric (i.e., symmetric with respect to the antidiagonal), that the decomposition is then uniquely determined, and that $\hat{S}_{j,j}$ can be found in a recursive process starting from its diagonal.

(iv) *Recursive generation of S from the block diagonal and the block subdiagonal of N.* One might expect that the corresponding algorithm (2.7) for the generic case is easily replaced by its block version. However, serious obstacles appear. For let us assume that the blocks $S_{j,k}$, $k = 0, \ldots, j-1$, of S are known, as well as all blocks $S_{l,k}$ with $l < j$. Then the pivot $N_{j,j}^{(j-1)} := \sum_{k=0}^{j-1} S_{j,k} S_{k,k}^{-1} S_{k,j}^T$ can be computed, but once again, as in (i), the factorization of $N_{j,j}^{j-1}$ into $S_{j,j} D_j S_{j,j}^T$ such that the corresponding block A_j in H has companion matrix structure seems not to be possible at this moment. If it were, we could then compute the first rows of the blocks $S_{j+1,k}$, $k = 0, \ldots, j$, except the last element in the first row of $S_{j+1,j}$ if $h_j > 0$, by applying (4.8a) and (4.8b). Then this last element could be obtained in a fashion analoguous to (2.7c) and the nontrivial elements of B_{j-1} and A_j would result from (4.4) and (4.6). Finally, (4.8a) and (4.8b) could again be used to find $S_{j+1,k}$, $k = 0, \ldots, j$, and $N_{j+1,j+1}^{(j)}$; the structure of the latter allows one to determine h_{j+1}.

Since in this procedure the correct factorization of the block pivot is unknown, one might try to build up $\tilde{S}$ or $\hat{S}$ instead, as in (iii). But then the determination of $\tilde{A}_j$ or $\hat{A}_j$ according to (4.3) would require the whole block $S_{j+1,j}$, which, on the other hand, cannot be determined from (4.8) unless $\tilde{A}_j$ or $\hat{A}_j$, respectively, are known! Although there may be a way out of this circle, it is likely that the procedure becomes very complicate.

Remark. There is some overlap between our Sections 1 and 2 and the dissertation of Mark Kent from Stanford University, which has been worked out in the same time period. We would like to point out that neither did Mark Kent have any preliminary version of our paper nor did we make use of any preliminary text from his thesis.

References

1. Chebyshev, P.: Sur l'interpolation par la méthode des moindres carrés. Mém. Acad. Impér. des Sciences St. Pétersbourg, série 7, **1**, 1–24 (1859)
2. Draux, A.: Polynômes Orthogonaux Formels-Applications. LNM Vol. 974, Berlin Heidelberg New York: Springer, 1983
3. Gautschi, W.: On the construction of Gaussian quadrature rules from modified moments. Math. Comp. **24**, 245–260 (1970)
4. Gautschi, W.: On generating orthogonal polynomials, SIAM J. Sci. Stat. Comput., **3**, 289–317 (1982)
5. Gautschi, W.: Questions of numerical conditions related to polynomials, In: G. H. Golub (ed.) Studies in Numerical Analysis, 140–177. Mathematical Association of America, 1984
6. Gautschi, W.: On the sensitivity of orthogonal polynomials to perturbations in the moments. Numer. Math. **48**, 369–382 (1986)

7. . Golub, G.H. and Kent, M.D.: Estimates of eigenvalues for iterative methods. Math. Comp. (to appear)
8. Golub, G.H. and Welsch, J.H.: Calculation of Gauss quadrature rules. Math. Comp. **23**, 221–230 (1969)
9. Gragg, W.B.: Matrix interpretations and applications of the continued fraction algorithm, Rocky Mountain J. Math. **4**, 213–225 (1974)
10. Gragg, W.B. and Lindquist, A.: On the partial realization problem. Linear Algebra Applics. **50**, 277–319 (1983)
11. Gutknecht, M.H.: A completed theory for the Lanczos process and related algorithms (in preparation)
12. Rutishauser, H.: Der Quotienten-Differenzen-Algorithmus. Mitteilungen aus dem Institut für angewandte Mathematik, Nr.7, Basel Stuttgart: Birkhäuser, 1957
13. Sack, R.A. and Donovan, A.F.: An algorithm for Gaussian quadrature given modified moments, Numer. Math. **18**, 465–478 (1972)
14. Struble, G.W.: Orthogonal polynomials: variable-signed weight functions. Numer. Math. **5**, 88–94 (1963)
15. Wall, H.S.: Analytic Theory of Continued Fractions. New York: D. Van Nostrand Company, 1948.
16. Wheeler, J.C.: Modified moments and Gaussian quadratures, J. Math. **4**, 287–295 (1974)

This article was processed by the author
using the Springer-Verlag TEX NumMat macro package 1990.