A COMPLETED THEORY OF THE UNSYMMETRIC LANCZOS PROCESS AND RELATED ALGORITHMS, PART II*

MARTIN H. GUTKNECHT†


Abstract. This paper is a continuation of Part I [M. H. Gutknecht, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 594–639], where the theory of the “unsymmetric” Lanczos biorthogonalization (BO) algorithm and the corresponding iterative method BIORES for non-Hermitian linear systems was extended to the nongeneric case. The analogous extension is obtained here for the biconjugate gradient (or BIOMIN) method and for the related BIODIR method. Here, too, the breakdowns of these methods can be cured. As a preparation, mixed recurrence formulas are derived for a pair of sequences of formal orthogonal polynomials belonging to two adjacent diagonals in a nonnormal Padé table, and a matrix interpretation of these recurrences is developed. This matrix interpretation leads directly to a completed formulation of the progressive qd algorithm, valid also in the case of a nonnormal Padé table. Finally, it is shown how the cure for exact breakdown can be extended to near-breakdown in such a way that (in exact arithmetic) the well-conditioned formal orthogonal polynomials and the corresponding Krylov space vectors do not depend on the threshold specifying the near-breakdown.

Key words. Lanczos algorithm, biconjugate gradient algorithm, BIOMIN, BIODIR, breakdown, formal orthogonal polynomial, recurrence, Padé approximation, staircase, quotient difference algorithm, qd algorithm

AMS subject classifications. 65F10, 30E05, 41A21, 65F15

Introduction. In Part I [13] we derived a number of basic results on sequences of formal orthogonal polynomials of the first and second kind (FOP1s and FOP2s, respectively). Given a linear functional \( \Phi : \mathcal{P} \to \mathbb{C} \) defined on the space \( \mathcal{P} \) of complex polynomials by

\[
\Phi_\ell(x^k) := \phi_{k+\ell} \quad (k \in \mathbb{N}),
\]

there is a finite or infinite sequence \( \{n_j\}_{j=0}^J \) of indices with \( 0 =: n_0 < n_1 < n_2 < \cdots < n_J \) if \( J < \infty \) for which a regular (monic) FOP1 \( P_{nj} := P_{i;n_j} \) exists. By definition, these are those values of the index \( n \) for which a unique monic polynomial \( P_n := P_{i;n} \) of exact degree \( n \) satisfying

\[
\Phi_\ell(pP_n) = 0 \quad (\forall p \in \mathcal{P}_{n-1})
\]

exists. These indices are also characterized by the nonsingularity of the \( n \times n \) moment matrix

\[
M_n := M_{i;n} := \begin{bmatrix}
\phi_0 & \phi_1 & \cdots & \phi_{i+n-1} \\
\phi_1 & \phi_2 & \cdots & \phi_{i+n} \\
\vdots & \vdots & \ddots & \vdots \\
\phi_{i+n-1} & \phi_{i+n} & \cdots & \phi_{i+2n-2}
\end{bmatrix}.
\]
Starting from these regular FOPls $P_n$, we have obtained a full sequence $\{P_n\}_{n=0}^{\infty}$ of monic FOPls by setting

(1.28) \[ P_n(z) := W_{n-n_j}(z)P_{n_j}(z) \quad \text{if} \quad n_j \leq n < n_{j+1} =: n_j + h_j. \]

Here $W_{n-n_j}$ could be an arbitrary monic polynomial of exact degree $n - n_j$, but in view of actual implementations we have primarily considered the case where $W_{n-n_j}$ is the $n - n_j$ element of a fixed sequence $\{W_m\}$ of monic polynomials satisfying a three-term recurrence

(2.10) \[ W_{m+1}(z) = (z - \alpha_m^W) W_m(z) - \beta_m^W W_{m-1}(z) \quad (m \in \mathbb{N}) \]

(with $W_0(z) := 1$, $W_{-1}(z) := 0$, $\beta_0^W := 0$).

The FOPls $P_n$ of (1.28) satisfy the formal orthogonality conditions

(2.1) $\Phi_l(pP_n) = 0 \quad (\forall p \in \mathcal{P}_{n-1})$, \quad $\Phi_l(z^\hat{n}P_n) \neq 0$, \quad where $\hat{n} := n_j + n_{j+1} - n - 1$.

In particular, when $P_n$ is regular, $n = n_j$, then

(2.2) \[ \hat{n} = n_j + h_j - 1 = n_{j+1} - 1. \]

Equivalently, assuming $n_i \leq n' \leq n_{i+1}$ and $n_j \leq n \leq n_{j+1}$, we can write

(2.5) \[ \Phi_l(P_{n'}, P_n) = 0 \quad \text{if} \quad i \neq j \quad \text{or} \quad i = j \quad \text{and} \quad n' + n < n_j + n_{j+1} - 1, \]

(2.6) \[ \Phi_l(P_{n'}, P_n) = \delta_j \neq 0 \quad \text{if} \quad i = j \quad \text{and} \quad n' + n = n_j + n_{j+1} - 1, \]

where $\delta_j$ is independent of $n - n_j$ and $n' - n_j$.

The formula (2.10) implies that the nonregular FOPls $P_n$ ($n_j < n < n_{j+1}$) can be generated according to

(2.11) \[ P_{n+1}(z) = (z - \alpha_{n-n_j}^W) P_n(z) - \beta_{n-n_j}^W P_{n-1}(z), \quad n_j \leq n \leq n_{j+1} - 2. \]

(Likewise, any other recurrence for $\{W_m\}$ leads to one for those $P_n$.) Less trivial is the fact that the orthogonality relation (2.1) allows us to establish for the regular FOPls a three-term recurrence

(2.17) \[ P_{n_j+1}(z) = (W_{h_j}(z) - a_j(z))P_{n_j}(z) - \beta_j P_{n_j-1}(z), \quad j = 0, \ldots, J - 1, \]

with a monic polynomial coefficient $W_{h_j} - a_j \in \mathcal{P}_{h_j}$ (hence, $a_j \in \mathcal{P}_{h_j-1}$) and a scalar coefficient $\beta_j \in \mathbb{C}$. (The initial values are: $P_{n-1}(z) := 0$, $P_{n_0}(z) := 1$, $\beta_0 := 0$.) For the coefficient $\beta_j$ there is an explicit formula and for the coefficients of $a_j(z) = \sum_{s=0}^{h_j-1} \alpha_{s,j} W_s(z)$ we have found a recursive formula based on the solution of a lower triangular system:

(2.23a) \[ \beta_j \Phi_l(P_{n_j-1}P_{n_j-1}) = \Phi_l(zP_{n_j-1}P_{n_j+1-1}), \]

(2.23b) \[ \Phi_l(a_jP_{n_j+k}P_{n_j}) = \Phi_l(zP_{n_j+k}P_{n_j+1-1}) - \alpha_{h_{j-1}}^W \Phi_l(P_{n_j+k}P_{n_j+1-1}) - \beta_{h_{j-1}}^W \Phi_l(P_{n_j+k}P_{n_j+1-2}), \quad k = 0, \ldots, h_j - 1. \]

After introducing the infinite row vector $p := [P_0, P_1, \ldots]$, we can express the recurrences (2.11) and (2.17) as

(3.11) \[ zp(z) = p(z)\mathbf{H}, \]
where the Gragg matrix $H$ is an infinite block tridiagonal unit upper Hessenberg matrix

$$
H := \begin{bmatrix}
A_0 & B_1 & C_0 & A_1 & B_2 & \cdots \\
C_0 & A_1 & C_1 & A_2 & \cdots & \cdots \\
& & & & & (B_J) \\
& & & & & (C_{J-1}) \\
& & & & & (A_J)
\end{bmatrix}.
$$

Under the assumption (2.10), the diagonal blocks $A_j$ are $h_j \times h_j$ matrices containing on the diagonal and the first superdiagonal the coefficients $\alpha_m^W$ and $\beta_m^W$ from (2.10), and in the last column the coefficients $\alpha_{j,i}$ of the polynomial $a_j$. The off-diagonal blocks $C_j$ and $B_j$ are zero except for the element in the upper right corner, which is 1 in $C_j$ and $\beta_j$ in $B_j$. If $J < \infty$, $B_J$ is the $h_{J-1} \times \infty$ zero matrix, and $A_j$ is the infinite tridiagonal matrix $T_W$ representing the recurrence (2.10) in the form $zw(z) w(z) T_w$ (where $w := [W_0, W_1, \ldots]$).

By using matrix notation, we can express the orthogonality properties (2.5)–(2.6) in compact form:

$$
(3.22) \Phi_l(p^T p) = D,
$$

where $D$ is a block diagonal matrix whose blocks $D_j$ are $h_j \times h_j$ lower right triangular matrices with all antidiagonal elements equal to $\delta_j$.

The FOP1s $P_n$ and the associated FOP2s $Q_n$ ($n \in \mathbb{N}$) are essentially the denominators and the numerators, respectively, of the proper parts of the Padé approximant lying on the $l$th diagonal of the Padé table of the formal Laurent series

$$
(1.25) f(z) = \sum_{k=-\infty}^{\infty} \phi_k z^k.
$$

More exactly, the $(m, n) := (l + n - 1, n)$ Padé approximant of $f$ is equal to

$$
r_{m,n}(z) := \sum_{k=-\infty}^{m-n} \phi_k z^k + z^{m-n} \frac{Q_n(z^{-1})}{P_n(z^{-1})},
$$

cf. (1.21), (1.22), and (1.34). The rational function $r_{m,n}$ is the $(m, n)$ entry of the Padé table.

An important feature of the Padé table is its block structure: Identical entries occur in finite or infinite square blocks, cf. Corollary 1.6. The regular FOP1s belong to entries on the first row or the first column of such a square block. This Block Structure Theorem is important in this second part, where we now consider pairs of sequences of FOP1s, $(P_n)_{n=0}^{\infty} := (P_{1:n})_{n=0}^{\infty}$ and $(P'_n)_{n=0}^{\infty} := (P_{1+n})_{n=0}^{\infty}$, which belong to two adjacent diagonals of the Padé table. We mark the quantities corresponding to the second sequence by a prime, writing for example $M'_n$, $H'$, $p'$. We also set $\Phi := \Phi_l$ and $\Phi' := \Phi_{l+1}$. Note that this usage of primes differs from the one in Part I, where they indicated quantities belonging to the FOP2s, the polynomials of the second kind.

In §5 we define a new sequence of regular FOP1s whose elements are alternatively taken from the two above-mentioned sequences and belong to all distinct Padé approximants that lie on the two diagonals. We call the corresponding sequence of
Padé approximants a block staircase sequence. In analogy to the three-term recurrence (2.17) for the regular FOP1s that belong to one diagonal, we derive a pair of three-term recurrence formulas (with one polynomial coefficient in each) for the new sequence of FOP1s. Since FOP1s from both diagonals appear in these formulas, we call them mixed recurrences. Two equivalent but different matrix formulations for them are given in §6. Actually, these matrix formulations involve all the polynomials from the two sequences \( \{P_n\}_{n=0}^{\infty} \) and \( \{P'_n\}_{n=0}^{\infty} \) and not just the regular ones.

By eliminating either the first or the second sequence, we rediscover in §7 the matrix formulations of the separate recurrences for the second and the first sequence, respectively, which are both of the type discussed in Part I, i.e., they are determined by Gragg matrices of the form (3.6). It turns out that the Gragg matrix of the second sequence is obtained from the one of the first sequence by executing one step of a block LR algorithm, i.e., we have to compute a particular block LU decomposition and then multiply the factors together in reverse order. The factors, which are block bidiagonal (but none of which is chosen with unit block diagonal), are exactly the matrices that describe the mixed recurrences of the block staircase. This block LR algorithm generalizes Rutishauser’s LR algorithm for tridiagonal matrices, and hence also his (equivalent) qd algorithm [23]. It is the key to a nongeneric progressive qd algorithm, which, in contrast to the classical (generic) progressive qd algorithm, never breaks down in exact arithmetic.

In §4 of Part I we applied the results on (diagonal) sequences of FOP1s to the unsymmetric Lanczos process. Let \( A : \mathcal{H} \to \mathcal{H} \) be a bounded linear operator mapping a separable real or complex Hilbert space into itself. The standard inner product in \( \mathcal{H} \) is denoted by \( \langle ., . \rangle \), but we use instead a formal inner product \( \langle ., . \rangle_B \) defined by \( \langle y, x \rangle_B := \langle y, Bx \rangle \), which is induced by another bounded linear operator \( B : \mathcal{H} \to \mathcal{H} \) that commutes with \( A \). (The cases of practical interest are \( B = I \), \( B = A \), and \( B = A^{-1} \).) Orthogonality with respect to this indefinite inner product is referred to as formal orthogonality. Associated with \( A \), \( x_0 \), \( y_0 \), and this inner product are the Schwarz constants or moments

\[
\phi_k := \langle y_0, A^k x_0 \rangle_B := \langle y_0, B A^k x_0 \rangle \quad (k \in \mathbb{N}).
\]

The link to the above-described theory of FOPs is based on the identification of these moments with the values that the linear functional \( \Phi = \Phi_0 \) of (1.1) takes on the monomials.

Starting from \( A \), \( x_0 \), \( y_0 \), the classical (generic) Lanczos biorthogonalization (BO) algorithm [19], [15], [11] generates the two sequences \( \{x_n\}_{n=0}^{\nu-1} \) and \( \{y_n\}_{n=0}^{\nu-1} \) such that for \( n = 0, 1, \ldots, \nu - 1 \)

\[
\begin{align*}
(4.6a) \quad x_n & \in \mathcal{X}_{n+1} := \text{span} \{ x_0, A x_0, A^2 x_0, \ldots, A^n x_0 \}, \\
(4.6b) \quad y_n & \in \mathcal{X}_{n+1} := \text{span} \{ y_0, A^H y_0, (A^H)^2 y_0, \ldots, (A^H)^n y_0 \},
\end{align*}
\]

and

\[
(4.7) \quad \langle y_m, x_n \rangle_B \begin{cases} = 0 & \text{if } m \neq n, \\ \neq 0 & \text{if } m = n. \end{cases}
\]

In view of (4.6), \( x_n \) must be equal to a polynomial in \( A \) times \( x_0 \), and \( y_n \) must be equal to a polynomial in \( A^H \) times \( y_0 \). From the orthogonality condition (4.7) and the uniqueness of the regular FOPs it follows easily that actually

\[
(4.14) \quad x_n = P_n(A)x_0 \Gamma_n, \quad y_n = P'_n(A^H)y_0 \tilde{\Gamma}_n,
\]
where $P_n$ is the monic regular FOP1 of degree $n$, $\overline{P_n}$ is the polynomial with the complex conjugate coefficients, and $\Gamma_n$ and $\overline{\Gamma_n}$ are scale factors. As we know, such a regular FOP1 need not exist, and that is when (4.7) no longer holds and the generic BO algorithm breaks down. Our remedy for this breakdown was to use (4.14), with $P_n$ being any FOP1 of degree $n$ (regular or not). The nongeneric BO algorithm (Algorithm 1) was then obtained by translating the recurrences for the FOP1s via (4.14) to recurrences for $x_n$ and $y_n$. In these recurrences the scale factors $\Gamma_n$ and $\overline{\Gamma_n}$ are replaced by the relative scale factors

\begin{equation}
\gamma_{n,i} := \frac{\Gamma_n}{\overline{\Gamma_{n-i}}}, \quad \overline{\gamma}_{n,i} := \frac{\Gamma_n}{\overline{\Gamma_{n-i}}} \quad (n \in \mathbb{N}, i \in \mathbb{N}).
\end{equation}

The application of the BO algorithm to solving linear systems of equations $Ax = b$ is based on defining a sequence of approximants $z_n$ in such a way that $x_n$ is the residual vector for $z_n$,

\begin{equation}
x_n = b - Az_n, \quad n = 0, 1, 2, \ldots,
\end{equation}

(normalized BIORES algorithm) or such that $x_n$ is the residual of $z_n$ in a system with scaled right-hand side

\begin{equation}
x_n = b\rho_n - Az_n, \quad n = 0, 1, 2, \ldots,
\end{equation}

(unnormalized BIORES algorithm).

The BO algorithm terminates when $x_n = 0$ or $y_n = 0$. But while the generic BO algorithm (and thus also BIORES) breaks down seriously whenever $(y_n, x_n)_B = 0$, our generalization fails only if the inner product vanishes for all $n$ beyond some bound $n_f$. This is then called an incurable breakdown [22], [21]. Unfortunately, in order to detect such an incurable breakdown, we theoretically have to work with exact arithmetic and to iterate until $n$ reaches the rank of $A$. When the algorithm is applied to solving a linear system, the hope is that in theory $x_n = 0$ for some $n$, and that in practice the residual $x_n$ is sufficiently small even much earlier. There is, however, the additional difficulty that $y_n = 0$ causes the algorithm to stop, and this requires that we find a nonzero replacement for $y_n$ that is orthogonal to $K_n$ and can be used to proceed.

Here, we discuss in §§8 three iterative linear system solvers that are closely related to the unnormalized nongeneric BIORES algorithm and have in fact the same breakdown behavior. (For the generic versions this is not true [11].) The first two, normalized and unnormalized nongeneric BIOMIN, are extensions of the well-known and highly successful biconjugate gradient (BCG) method [20], [7], [11]. The third is (normalized) nongeneric BIODIR, which, independently, has also been developed by Joubert [17], [18]. These three methods generate relevant subsequences of essentially the same sequences of approximants $z_n$, residuals $x_n$, and corresponding B-biorthogonal vectors $y_n$ as nongeneric BIORES; but additionally they produce two BA-biorthogonal sequences $\{u_n\}$ and $\{v_n\}$. The elements of the first serve as direction vectors, i.e., they specify the direction of the correction for $z_n$. While the vectors $x_n$ and $y_n$ correspond according to (4.14) to a diagonal sequence of FOP1s, the vectors $u_n$ and $v_n$ correspond similarly to the FOP1s on the adjacent diagonal. And while nongeneric BIORES is based on the recurrence for the first block diagonal sequence of FOP1s, nongeneric BIOMIN is based on the recursion for the block staircase sequence, and nongeneric BIODIR makes use of the one for the block diagonal sequence on the adjacent diagonal.

Finally, in §§9 and 10, we present a theoretically clean approach to treating near-breakdowns. Of course, such an approach is of great importance in practice, where
the occurrence of exact breakdown is very unlikely, but near-breakdown may cause severe numerical effects. We formulate this theory for polynomials, but the application to the above-mentioned iterative linear solvers is straightforward. (The actual implementation still requires a careful treatment of many nontrivial details. This is the subject of joint work with Roland Freund and Noël Nachtigal [8].)

This treatment of near-breakdown is based on defining appropriate clusters or blocks of polynomials in such a way that formal orthogonality is maintained between the blocks, but not within the blocks. The first polynomial in each block is still a regular FOP1, and, moreover, it is well conditioned if the blocks are suitably chosen. To construct these polynomials (or the corresponding sequences of Krylov space vectors) we apply a block orthogonalization process, which is just the appropriate generalization of the Gram–Schmidt process.2 The resulting algorithm, which is described here in terms of polynomials and in [8] in terms of Krylov space vectors, can be considered as a generalization of the nongeneric BO algorithm of §4 and of the similar algorithms proposed by Parlett, Taylor, and Liu [22], [21], [24] and by Boley et al. [1]. Although mainly exact breakdowns were considered in [22], [24], we suggest applying Parlett’s adjective “look-ahead,” which is by now well established, to the near-breakdown versions of all of the above-mentioned algorithms, while the adjective “nongeneric” should be reserved for the versions curing exact breakdown.

In Parlett, Taylor, and Liu [22] the discussion was actually restricted to 2 × 2 blocks. Several options of block LDU decomposition of the moment matrix were considered for this case, but the resulting generalizations of the Lanczos algorithm differ in detail considerably from the proposals made here, even for exact breakdown. In particular, the left Lanczos vectors are chosen differently; thus, in the relations (4.14) the polynomial $P_n$ is there in general not the complex conjugate polynomial of $P_n$. Moreover, as we will see in §9, the above-mentioned “appropriate generalization” of the Gram–Schmidt process for near-breakdowns is not the straightforward one, which would not yield a “block three-term” recurrence. Finally, here we not only treat block diagonal sequences (§9), but also block staircase sequences (§10), which present some additional difficulties.

Upon revision of this paper we learned of nongeneric algorithms developed by Hegediüs [14] for applying conjugate gradients to a particular indefinite problem. From his treatment one must conclude that he was probably also aware of the possibility of developing some of the nongeneric algorithms given here; but he did not specify them.

5. **Block staircase sequences and corresponding recurrences.** In this section we consider two adjacent sequences of FOP1s, \( \{P_n\} := \{P_{i:n}\} \) and \( \{P'_n\} := \{P_{i+1:n}\} \), and their associated sequences of FOP2s, \( \{Q_n\} := \{Q_{i:n}\} \) and \( \{Q'_n\} := \{Q_{i+1:n}\} \), and derive a recurrence for a particular sequence formed alternatively from regular elements of \( \{P_n\} \) and \( \{P'_n\} \). We denote those elements of this sequence that are taken from \( \{P_n\} \) by \( P_{n,j}^r \), and those taken from \( \{P'_n\} \) by \( P_{n,j}^l \). The index sequences

\[
\{n_j^r\}_{j=0}^{\infty} \quad \text{and} \quad \{n_j^l\}_{j=0}^{\infty}
\]

still indicate the degrees of the corresponding polynomials and are subsequences of the two index sequences \( \{n_j\} \) and \( \{n_j^l\} \), respectively, that belong to the regular FOP1s in

---

2 We use the notion “block orthogonal” here, although there is a danger of confusion with the different meaning the word “block” has in “block Lanczos” or “block Gram–Schmidt.”
Fig. 3. A block staircase sequence. The elements of the staircase are marked by *, while the other regular elements on the two adjacent diagonals are marked by o. The notation $n_j^\wedge \vee$ means $n_j^\wedge = n_j^{\vee}$.

$\{P_n\}$ and $\{P'_n\}$, respectively. These subsequences are chosen such that $n_0^\wedge := 0$ and

\begin{equation}
(5.1) \quad n_j^\wedge \leq n_j^{\vee} \quad (j = 0, \ldots, J^{\vee}), \quad n_j^{\vee} < n_{j+1}^\wedge \quad (j = 0, \ldots, J^{\wedge} - 1),
\end{equation}

and such that they contain as many indices as possible; in the cases where the interlacing condition (5.1) does not determine $n_j^{\vee}$ or $n_{j+1}^\wedge$ uniquely, we make the last of the choices allowed by (5.1). From the Block Structure Theorem 1.6 for the Padé table, it is seen that such ambiguities occur in connection with blocks that contain elements from both $\{P_n\}$ and $\{P'_n\}$, i.e., which are intersected both by the upper and the lower diagonal (on which $m - n = l$ and $m - n = l + 1$, respectively), cf. Fig. 3. In such cases, either the first column or the first row of the block contains regular elements out of both $\{P_n\}$ and $\{P_{l+1,n}\}$, and then the lower or the right, respectively,
of these elements is dropped, while the other becomes an element of \( \{P_{n_j^\dagger}\} \) or \( \{P_{n_j'}\} \), respectively. Note that
\[
(5.2) \quad J^\dagger \leq J^\wedge \leq J^\dagger + 1 \leq \infty,
\]
i.e., there can be at most one more marked element on the upper diagonal, and the number of elements may be infinite.

Let us also set
\[
(5.3) \quad m_j^\dagger := n_j^\dagger + l - 1, \quad m_j^\wedge := n_j^\wedge + l,
\]
\[
(5.4) \quad h_j^\dagger := n_j^\dagger - n_j^\wedge + 1 = m_j^\dagger - m_j^\wedge, \quad h_j^\wedge := n_j^\wedge + 1 - n_j^\dagger = m_j^\wedge + 1 - m_j^\dagger + 1.
\]
Both \( h_j^\dagger \) and \( h_j^\wedge \) are \( \geq 1 \). If \( J^\dagger, J^\wedge < \infty \), then \( h_j^\dagger, h_j^\wedge := h_j^\wedge := n_j^\wedge := n_j^\wedge + 1 := \infty \).

The index pair sequences \( \{(m_j^\dagger, n_j^\dagger)\}_{j=0}^{\infty} \) and \( \{(m_j^\wedge, n_j^\wedge)\}_{j=0}^{\infty} \) define the *block staircase sequence* for the two adjacent diagonals of the Padé table. (For the more general situation of the Newton–Padé table, block staircase sequences have been introduced in [12].) Note that the index pairs specify for each block an entry of minimum degrees \( m \) and \( n \) on one of the two adjacent diagonals, but do not indicate the upper left corner of the block.

The orthogonality result of Theorem 2.1 yields the following lemma for such block staircase sequences.

**Lemma 5.1.** The following formal orthogonality properties hold:

\[
(5.5a) \quad \Phi(pP_{n_j^\dagger}) = 0 \quad (\forall p \in \mathcal{P}_{n_j^\dagger + h_j^\dagger - 2} = \mathcal{P}_{n_j^\wedge - 1})
\]
\[
(5.5b) \quad \delta_j^\dagger := \Phi(z^{n_j^\wedge} P_{n_j^\dagger}) \neq 0,
\]

and

\[
(5.6a) \quad \Phi(pP_{n_j'}) = 0 \quad (\forall p \in \mathcal{P}_{n_j^\wedge + h_j' - 2} = \mathcal{P}_{n_j^\wedge + 1 - 2})
\]
\[
(5.6b) \quad \delta_j^\wedge := \Phi(z^{n_j^\wedge - 1} P_{n_j'}) \neq 0.
\]

**Proof.** Apply (2.1) and (2.2) to the current situation and note that the regular elements following \( P_{n_j^\dagger} \) and \( P_{n_j'} \) on the same diagonal have the indices
\[
(5.7) \quad n_j^\dagger + 1 = n_j^\wedge + h_j^\dagger \quad \text{and} \quad n_j^\wedge + 1 = n_j^\wedge + h_j^\dagger,
\]
respectively, independently of whether this following regular element is part of the block staircase or not. This is due to the fact that, in case of a dropped element, the index of the next element differs from that of the dropped one only by 1, as can be seen from Fig. 3. □

**Corollary 5.2.** The following formal orthogonality properties hold:

\[
(5.8a) \quad \Phi'(pP_{n_j^\dagger}) = 0 \quad (\forall p \in \mathcal{P}_{n_j^\dagger + h_j^\dagger - 3} = \mathcal{P}_{n_j^\wedge - 2})
\]
\[
(5.8b) \quad \delta_j^\dagger := \Phi'(z^{n_j^\wedge - 1} P_{n_j^\dagger}) \neq 0,
\]

and

\[
(5.9a) \quad \Phi(zP'_{n_j'}) = 0 \quad (\forall p \in \mathcal{P}_{n_j^\wedge + h_j' - 2} = \mathcal{P}_{n_j^\wedge + 1 - 2})
\]
\[
(5.9b) \quad \delta_j^\wedge := \Phi(z^{n_j^\wedge + 1} P_{n_j'}) \neq 0.
\]
Proof. In the same way that (1.2) was written as the homogeneous linear system (1.6), (5.5a) can be written as

$$\sum_{j=0}^{n} \phi_{i+j-n} \pi_{i,j,n} = 0, \quad i = l + n, \ldots, l + 2n + h - 2,$$

where $n := n^\wedge, h := h^\wedge$. When the first equation is deleted, this system represents (5.8a). Both (5.5b) and (5.8b) express that the “next equation after (5.10),” with $i := l + 2n + h - 1$, is not homogeneous.

Similarly, (5.6a) and (5.9a) are represented by the same homogeneous linear system, and (5.6b) and (5.9b) yield the same strictly inhomogeneous equation. □

Lemma 5.1 and Corollary 5.2 allow us now to establish recurrence formulas for block staircase sequences. By analogy to the derivation of the recurrence formula for diagonal sequences in §2, we start from representations of $P_{n^{\wedge}+1}$ and $P'_{n^{\wedge}+1}$ in terms of the previous elements of the staircase. Clearly, in view of (5.4), there are polynomials $t_{s,j}^\wedge \in \mathcal{P}_{h_{s}^\wedge -1}$ and $t_{s,j}^\wedge \in \mathcal{P}_{h_{s}^\wedge -2}$ ($s = 0, 1, \ldots, j$) such that

$$P_{n^{\wedge}+1}^\wedge(z) = zW_{h_{s}^\wedge -1}(z)P_{n_{j}^\wedge}(z) - \sum_{s=0}^{j} \left[ z^{t_{s,j}^\wedge}(z)P'_{n_{j}^\wedge}(z) + t_{s,j}^\wedge(z)P_{n_{j}^\wedge}(z) \right].$$

(If $h_{s}^\wedge = 1$, we set $t_{s,j}^\wedge := 0.$) We multiply this relation with the monomials of degree at most $n_{j}^\wedge + h_{s}^\wedge - 1 = n_{j}^\wedge + 1 - 1$. Each of these monomials can be written as either $z^{n_{i}^\wedge + k}$ with $0 \leq k \leq h_{s}^\wedge - 2$, $0 \leq i \leq j$, or $z^{n_{j}^\wedge + k}$ with $0 \leq k \leq h_{s}^\wedge - 1$, $0 \leq i \leq j$. (Note that the range of $k$ in the first loop is empty if $h_{s}^\wedge = 1$, in which case the multiplier $z_{n_{i}^\wedge + k}$ is not used.) Then, we apply $\Phi$ to the resulting $n_{j}^\wedge + 1$ relations in order to obtain a linear system of $n_{j}^\wedge + 1$ equations for the polynomials $t_{s,j}^\wedge$ and $t_{s,j}^\wedge$, $s = 0, \ldots, j$. In view of (5.4), these polynomials have a total of $n_{j}^\wedge$ coefficients. In view of (5.8a), all expressions $\Phi(z^{n_{j}^\wedge + k}P_{n_{j}^\wedge}^\wedge)$ and $\Phi(z^{n_{j}^\wedge + k}P'_{n_{j}^\wedge}^\wedge)$ vanish, so that $P_{n_{j}^\wedge + 1}^\wedge$ does not appear in the system. In order to verify the structure of this system, we list a number of results following directly from (5.4), (5.5), and (5.9); $k$ is always assumed to lie in the given ranges.

$$\Phi(z^{n_{j}^\wedge + k}W_{h_{j}^\wedge -1}P'_{n_{j}^\wedge}) = 0 \quad \text{if} \quad i \leq j.$$ (Here, for the case $i = j - 1$, we have used that $n_{j-1}^\wedge + k \leq n_{j-1}^\wedge + h_{j-1}^\wedge - 2 = n_{j-1}^\wedge - 1 \leq n_{j}^\wedge - 2 \leq n_{j}^\wedge - 2$.)

$$\Phi(z^{n_{j}^\wedge + k + 1}W_{h_{j}^\wedge -1}P'_{n_{j}^\wedge}) \begin{cases} = 0 & \text{if} \quad i \leq j - 1, \\ \neq 0 & \text{if} \quad i = j, \quad k = 0. \end{cases}$$ (Here, we have used that $n_{j-1}^\wedge + h_{j-1}^\wedge + h_{j}^\wedge - 2 = n_{j}^\wedge + h_{j}^\wedge - 2 = n_{j}^\wedge - 1$.)

$$\Phi(z^{n_{j}^\wedge + k + 1}t_{s,j}^\wedge P_{n_{j}^\wedge}) = 0 \quad \text{if} \quad i \leq s;$$

$$\Phi(z^{n_{j}^\wedge + k + 1}t_{s,j}^\wedge P'_{n_{j}^\wedge}) \begin{cases} = 0 & \text{if} \quad i < s, \\ = 0 & \text{if} \quad i = s, \quad k + \partial t_{s,j}^\wedge < h_{s}^\wedge - 1, \\ \neq 0 & \text{if} \quad i = s, \quad k + \partial t_{s,j}^\wedge = h_{s}^\wedge - 1; \end{cases}$$
By capitalizing on these formulas, the above-mentioned linear system (which was obtained by multiplying (5.11), for \( i = 0, 1, \ldots, j \), with \( z^{n_i + k} \) \( (k = 0, \ldots, h_i^\wedge - 2) \) and \( z^{n_i + k} \) \( (k = 0, \ldots, h_i^\wedge - 1) \), and then applying \( \Phi \)) reduces to the following set of equations:

\[
(5.14a) \quad \sum_{s=0}^{i-1} \Phi(z^{n_i + k + 1} t_{s,j}^\wedge P_{n_j}^\wedge) + \sum_{s=0}^{i} \Phi(z^{n_i + k} t_{s,j}^\wedge P_{n_j}^\wedge) = 0, \\
\quad k = 0, \ldots, h_i^\wedge - 2; \ i = 0, \ldots, j; \\
(5.14b) \quad \sum_{s=0}^{i} \Phi(z^{n_i + k + 1} t_{s,j}^\wedge P_{n_j}^\wedge) + \sum_{s=0}^{i} \Phi(z^{n_i + k} t_{s,j}^\wedge P_{n_j}^\wedge) = \begin{cases} 0 & \text{if } i \leq j - 1, \\
\quad \Phi(z^{n_j + k + 1} W_{h_j^\wedge - 1} P_{n_j}^\wedge) \neq 0 & \text{if } i = j, \ k = 0, \\
\quad \Phi(z^{n_j + k + 1} W_{h_j^\wedge - 1} P_{n_j}^\wedge) & \text{if } i = j, \\
\quad k = 0, \ldots, h_i^\wedge - 1; \ i = 0, \ldots, j. \end{cases}
\]

This system is of triangular structure, like the one for diagonal sequences, which consists of (2.13) and (2.14). Again, except for the last few equations, the system is homogeneous and we can conclude that "most" of the unknown polynomials are zero. In fact, let us assume that \( j > 0 \) and, at first, that \( t_{0,j}^\wedge \neq 0 \). Then, if \( \partial t_{0,j}^\wedge > 0 \), the equation with \( i = 0 \) and \( k = h_0^\wedge - \partial t_{0,j}^\wedge - 1 \) in (5.14a), which is homogeneous but contains in view of (5.13c) exactly one nonzero term, yields a contradiction. Likewise, if \( \partial t_{0,j}^\wedge = 0 \), we let \( k = 0 \) and use (5.14b), (5.13b), and (5.13d) to obtain a contradiction. Next, suppose that \( t_{0,j}^\wedge = t_{1,j}^\wedge = \cdots = t_{i,j}^\wedge = 0 \), but \( t_{i,j}^\wedge \neq 0 \) for some \( i \leq j - 1 \). Then, (5.14b) and (5.13b) with \( k = h_i^\wedge - \partial t_{i,j}^\wedge - 1 \in \{1, 2, \ldots, h_i^\wedge - 1\} \) lead again to a contradiction; hence \( t_{i,j}^\wedge = 0 \). If \( t_{0,j}^\wedge = t_{i,j}^\wedge = \cdots = t_{i-1,j}^\wedge = 0 \) for some \( i \leq j - 1 \), we can likewise conclude from (5.14a) and (5.13c) or from (5.14b), (5.13b), and (5.13d) that \( t_{i,j}^\wedge = 0 \). In summary, if \( j > 0 \), there holds

\[
(5.15) \quad t_{0,j}^\wedge = t_{1,j}^\wedge = \cdots = t_{j-2,j}^\wedge = t_{j-1,j}^\wedge = t_{j,j}^\wedge = 0.
\]

By the same arguments we conclude from (5.14a) and (5.13c) that \( \partial t_{j,j}^\wedge \leq 0 \); from (5.14b), (5.13b), and (5.13d), by choosing \( i = j \) and \( k = 0 \), we conclude that

\[
(5.16) \quad t_{j,j}^\wedge (z) \equiv: \varphi_j^\wedge \neq 0,
\]

the constant \( \varphi_j^\wedge \) being given by

\[
(5.17a) \quad \varphi_j^\wedge = \delta_j^\wedge / \delta_j^\wedge, \quad \text{where } \delta_j^\wedge = \Phi(z^{n_j} P_{n_j}^\wedge), \ \delta_j^\wedge = \Phi(z^{n_j+1} P_{n_j}^\wedge).
\]

For \( e_j^\wedge := t_{j,j}^\wedge \) we get from (5.14b) (with \( i = j \)) the additional \( h_j^\wedge - 1 \) equations

\[
(5.17b) \quad \Phi(z^{n_j + k + 1} e_j^\wedge P_{n_j}^\wedge) + \varphi_j^\wedge \Phi(z^{n_j + k} P_{n_j}^\wedge) = \Phi(z^{n_j + k + 1} W_{h_j^\wedge - 1} P_{n_j}^\wedge), \\
\quad k = 1, \ldots, h_j^\wedge - 1.
\]
If the polynomial $e_j^\nu$ is expressed in powers of $z$ or in terms of the polynomials $W_m$, it follows from (5.9b) that the matrix of the resulting linear system for the coefficients is right lower triangular and regular, since its antidiagonal elements are all equal to $\delta_j^\nu \neq 0$. In case of the monomial basis, the matrix is Hankel.

Summarizing, we have shown so far that the general representation (5.11) reduces actually to a mixed three-term recurrence formula

$$P_{n_j+1}^\nu(z) = [zW_{h_j^\nu-1}(z) - ze_j^\nu(z)]P_{n_j}^\nu(z) - \varphi_j^\nu P_{n_j}^\nu(z), \quad j = 0, 1, \ldots, J^\nu - 1.$$  \hspace{1cm}(5.18)

In a completely analogous manner we can derive a mixed three-term recurrence formula for computing $P_{n_j}^\nu$. We start from the representation

$$P_{n_j}^\nu(z) = W_{h_j^\nu-1}(z)P_{n_j}^\nu(z) - \sum_{s=0}^{j} t_{s,j}^\nu(z)P_{n_s}^\nu(z) - \sum_{s=0}^{j-1} t_{s,j}^\nu(z)P_{n_s}^\nu(z)$$

with new polynomials $t_{s,j}^\nu \in P_{h_j^\nu-2}$ $(s = 0, 1, \ldots, j)$ and $t_{s,j}^\nu \in P_{h_j^\nu-1}$ $(s = 0, 1, \ldots, j - 1)$. (If $h_j^\nu = 1$, we set $t_{s,j}^\nu := 0$.)

This time, we multiply this relation with the monomials $zn_i^{\nu+k}$ $(1 \leq k \leq h_j^\nu - 1, \, 0 \leq i \leq j)$ and $zn_i^{\nu+k}$ $(1 \leq k \leq h_j^\nu, \, 0 \leq i \leq j - 1)$, which together are all the monomials with degrees between 1 and $n_j^\nu + h_j^\nu - 1 = n_j^\nu$. Again, we then apply $\Phi$ to both sides. Due to (5.9a), $P_{n_j}^\nu$ does not appear in the resulting linear system of $n_j^\nu$ equations, and in view of (5.4), the total number of coefficients of $t_{s,j}^\nu$ and $t_{s,j}^\nu$ in (5.19) is also $n_j^\nu$. To simplify the system we need the following formulas, which are analogous to (5.12) and (5.13). (Note that the ranges of $k$ and the maximum degrees of $\partial t_{s,j}^\nu$ and $\partial t_{s,j}^\nu$ have changed.)

$$(5.20a) \quad \Phi(z^{n_i^{\nu+k}}W_{h_j^\nu-1}P_{n_j}^\nu) = 0 \text{ if } i \leq j - 1;$$  \hspace{1cm}

$$(5.20b) \quad \Phi(z^{n_i^{\nu+k}}W_{h_j^\nu-1}P_{n_j}^\nu) \left\{ \begin{array}{l} = 0 \text{ if } i < j - 1, \\ = 0 \text{ if } i = j - 1, \, k \leq h_j^\nu - 1 - 1, \\ \neq 0 \text{ if } i = j - 1, \, k = h_j^\nu - 1 - 1; \end{array} \right.$$  \hspace{1cm}(5.21a)

$$(5.21b) \quad \Phi(z^{n_i^{\nu+k}}t_{s,j}^\nu P_{n_j}^\nu) \left\{ \begin{array}{l} = 0 \text{ if } i < s, \\ = 0 \text{ if } i = s, \, k + \partial t_{s,j}^\nu < h_s^\nu, \\ \neq 0 \text{ if } i = s, \, k + \partial t_{s,j}^\nu = h_s^\nu; \end{array} \right.$$  \hspace{1cm}(5.21b)

$$(5.21c) \quad \Phi(z^{n_i^{\nu+k}}t_{s,j}^\nu P_{n_s}^\nu) \left\{ \begin{array}{l} = 0 \text{ if } i < s, \\ = 0 \text{ if } i = s, \, k + \partial t_{s,j}^\nu < h_s^\nu - 1, \\ \neq 0 \text{ if } i = s, \, k + \partial t_{s,j}^\nu = h_s^\nu - 1; \end{array} \right.$$  \hspace{1cm}(5.21c)

$$(5.21d) \quad \Phi(z^{n_i^{\nu+k}}t_{s,j}^\nu P_{n_s}^\nu) = 0 \text{ if } i < s.$$
By applying these formulas, we can rewrite the linear system as

\[(5.22a) \quad \sum_{s=0}^{i} \Phi(z^{n_j+k} t_{s,j} P_{n_j}) + \sum_{s=0}^{i} \Phi(z^{n_j+k} t_{s,j} P_{n_j}) = \begin{cases} 0 & \text{if } i < j - 1, \\ 0 & \text{if } i = j - 1, k \leq h_{j-1}^\gamma - 1, \\ \Phi(z^{n_j-1+k} W_{h_{j-1}^\gamma} P_{n_j}) & \text{if } i = j - 1, k = h_{j-1}^\gamma - 1, \\ k = 1, \ldots, h_i^\gamma; i = 0, \ldots, j - 1; \end{cases}\]

\[(5.22b) \quad \sum_{s=0}^{i-1} \Phi(z^{n_j+k} t_{s,j} P_{n_j}) + \sum_{s=0}^{i} \Phi(z^{n_j+k} t_{s,j} P_{n_j}) = \begin{cases} 0 & \text{if } i \leq j - 1, \\ \Phi(z^{n_j+k} W_{h_{j-1}^\gamma} P_{n_j}) & \text{if } i = j, \\ k = 1, \ldots, h_i^\gamma - 1; i = 0, \ldots, j. \end{cases}\]

Again, the system has a triangular structure. For let us assume that \( j > 0 \) and \( t_{0,j}^\gamma \neq 0 \). Then (5.21c) and (5.22b) with \( i = 0 \) and \( k = h_{0,j}^\gamma - \delta t_{0,j}^\gamma - 1 \in \{1, 2, \ldots, h_{0,j}^\gamma - 1\} \) yield a contradiction, showing that \( t_{0,j}^\gamma = 0 \). By the same argument, \( t_{i,j}^\gamma = t_{i-1,j}^\gamma = \cdots = t_{i-j+1,j}^\gamma = 0 \) (for some \( i \leq j - 1 \)) implies that \( t_{i,j}^\gamma = 0 \). On the other hand, assuming \( t_{i,j}^\gamma = t_{i-1,j}^\gamma = \cdots = t_{0,j}^\gamma = 0, t_{i,j}^\gamma \neq 0 \) (for some \( i \leq j - 1 \)), we conclude from (5.21b) and (5.22a), by choosing \( k = h_{j}^\gamma - \delta t_{i,j}^\gamma \in \{1, 2, \ldots, h_j^\gamma\} \) there, that \( t_{i,j}^\gamma = 0 \), unless \( i = j - 1 \) and \( h_{j-1}^\gamma \neq 0 \). In the latter case there holds

\[(5.23) \quad t_{j-1,j}^\gamma(z) \equiv \varphi_j^\gamma \neq 0,\]

where the constant \( \varphi_j^\gamma \) is given by

\[(5.24a) \quad \varphi_j^\gamma = \delta_j^\gamma / \delta_{j-1}^\gamma, \quad \text{where } \delta_{j-1}^\gamma = \Phi(z^{n_j} P_{n_j}^\gamma - 1), \quad \delta_j^\gamma = \Phi(z^{n_j} P_{n_j}^\gamma - 1 P_{n_j}^\gamma).\]

Finally, choosing \( i = j \) in (5.22b) yields a linear system of \( h_j^\gamma - 1 \) equations for \( e_j^\gamma := t_{i,j}^\gamma:\)

\[(5.24b) \quad \Phi(z^{n_j+k} e_j^\gamma P_{n_j}^\gamma) + \varphi_j^\gamma \Phi(z^{n_j+k} P_{n_j}^\gamma - 1) = \Phi(z^{n_j+k} W_{h_{j-1}^\gamma} P_{n_j}^\gamma), \quad k = 1, \ldots, h_j^\gamma - 1.\]

Here also, if \( e_j^\gamma \) is expressed in a basis of polynomials with ascending degrees, the coefficient matrix of this system is right lower triangular and regular, with antidiagonal elements \( \delta_j^\gamma \neq 0 \). If the monomial basis is used, the matrix is again Hankel.

Summarizing, we get the following theorem.

**Theorem 5.3.** The regular FOP1s \( P_{n_j}^\gamma, j = 0, \ldots, J^\gamma, \) and \( P_{n_j}^\gamma, j = 0, \ldots, J^\gamma, \) of the block staircase sequence starting at \( (0, n_0^\gamma) = (0, 1) \) satisfy a pair of mixed three-term recurrence formulas:

\[(5.25a) \quad P_{n_j}^\gamma(z) = [W_{h_j}^\gamma - 1(z) - e_j^\gamma(z)] P_{n_j}^\gamma(z) - \varphi_j^\gamma P_{n_j}^\gamma - 1(z), \quad j = 0, 1, \ldots, J^\gamma,\]

\[(5.25b) \quad P_{n_j}^\gamma(z) = [W_{h_j}^\gamma - 1(z) - \varphi_j^\gamma(z)] P_{n_j}^\gamma(z) - \varphi_j^\gamma P_{n_j}^\gamma(z), \quad j = 0, 1, \ldots, J^\gamma - 1,\]

with initial values \( P_{n_{j-1}}^\gamma(z) \equiv 0, P_{n_j}^\gamma(z) \equiv 1, \varphi_0^\gamma := 0. \) \( \{W_m\}_{m=0}^\infty \) is an arbitrary prescribed sequence of monic polynomials of respective degree \( m \); \( \{\varphi_j^\gamma\}_{j=0}^{J^\gamma-1} \) and \( \{\varphi_j^\gamma\}_{j=0}^{J^\gamma-1} \)
are sequences of uniquely determined nonzero complex constants, each of which is
given by (5.24a) or (5.17a), respectively; and \( \{e_j^\wedge\} \) and \( \{e_j^\vee\} \) are sequences of
complex polynomials of respective degree \( \delta e_j^\wedge \leq h_j^\wedge - 2 \) and \( \delta e_j^\vee \leq h_j^\vee - 2 \), each of which
is the unique solution of a linear system (5.24b) or (5.17b), respectively. \( e_j^\wedge(z) \equiv 0 \)
if \( h_j^\wedge = 1 \), and \( e_j^\vee(z) \equiv 0 \) if \( h_j^\vee = 1 \). The integers \( h_j^\wedge \) and \( h_j^\vee \) are determined by (5.5)
and (5.9), respectively, i.e.,

\[
(5.26a) \quad h_j^\wedge := \min\{k \in \mathbb{N}^+; \Phi(z^{n_j^\wedge + k - 1}P_{n_j^\wedge}) \neq 0\} = \min\{k \in \mathbb{N}^+; \Phi(W_{k-1}(P_{n_j^\wedge})^2) \neq 0\},
\]

\[
(5.26b) \quad h_j^\vee := \min\{k \in \mathbb{N}^+; \Phi(z^{n_j^\vee + k}P_{n_j^\vee}) \neq 0\} = \min\{k \in \mathbb{N}^+; \Phi(zW_{k-1}(P_{n_j^\vee})^2) \neq 0\}.
\]

Similar recurrences, which, however, involve all regular FOP1s on both diagonals,
have been given by Draux [6, pp. 394–398].

As in §2 we could state as corollaries the special results obtained for the cases
where the polynomials \( W_m \) satisfy a three-term recurrence and where the polynomials
\( e_j^\wedge \) and \( e_j^\vee \) are expressed as linear combinations of these \( W_m \). At this point, the
formulation of these results is left to the reader, but in §6 we will give their matrix
formulation.

Theorem 5.3 allows us to construct recursively the sequence \( P_{n_j^\wedge}, P_{n_j^\vee}, P_{n_j^\wedge}, P_{n_j^\vee}, P_{n_j^\wedge}, \ldots \)
that contains all essentially distinct regular FOP1s on two adjacent diagonals.
However, as is apparent from Fig. 3, each of the sequences \( \{P_{n_j^\wedge}\}_{j=0}^{\infty} \) and \( \{P_{n_j^\vee}\}_{j=0}^{\infty} \) in
general does not contain all regular FOP1s of the corresponding diagonal. If \( n_j^\vee > n_j^\wedge \),
then \( P_{n_j^\wedge} \) is also a regular element of the lower \( (\vee) \) diagonal, and if \( n_{j+1}^\wedge > n_j^\wedge + 1 \),
then \( zP_{n_j^\wedge} \) is also a regular element of the upper \( (\wedge) \) diagonal. To obtain two full
sequences of FOP1s, we use additionally the definition (1.28):

\[
(5.27a) \quad P_n(z) := \begin{cases} 
W_{n-n_j^\wedge}(z)P_{n_j^\wedge}(z) & \text{if } n_j^\wedge \leq n \leq n_j^\vee \quad (= n_j^\wedge + h_j^\wedge - 1), \\
zW_{n-n_j^\vee-1}(z)P_{n_j^\vee}(z) & \text{if } n_j^\vee < n < n_{j+1}^\vee \quad (= n_j^\vee + h_j^\vee); 
\end{cases}
\]

\[
(5.27b) \quad P'_n(z) := \begin{cases} 
W_{n-n_j^\wedge}(z)P_{n_j^\wedge}(z) & \text{if } n_j^\wedge \leq n \leq n_j^\vee \quad (= n_j^\wedge + h_j^\wedge - 1), \\
W_{n-n_j^\vee}(z)P_{n_j^\vee}(z) & \text{if } n_j^\vee \leq n < n_{j+1}^\vee \quad (= n_j^\vee + h_j^\vee). 
\end{cases}
\]

Then, clearly,

\[
(5.28a) \quad P'_n(z) = P_n(z) \quad \text{if } n_j^\wedge \leq n < n_j^\vee, \\
(5.28b) \quad P_n(z) = zP'_n(z) \quad \text{if } n_j^\vee < n < n_{j+1}^\vee.
\]

Equations (5.27) and (5.28) allow us to modify the recurrences (5.25) in several ways.
If the three-term recurrence (2.10) is assumed to hold, we have moreover

\[
(5.29a) \quad P_{n+1}(z) = (z - \alpha_{n-n_j^\wedge})P_n(z) - \beta_{n-n_j^\wedge}P_{n-1}(z), \quad n_j^\wedge \leq n \leq n_j^\vee - 1,
\]

\[
(5.29b) \quad P'_{n+1}(z) = (z - \alpha_{n-n_j^\vee})P'_n(z) - \beta_{n-n_j^\wedge}P'_{n-1}(z), \quad n_j^\vee \leq n \leq n_{j+1}^\vee - 2.
\]

Here, by (5.28), the terms \( zP_n(z) \) and \( zP'_n(z) \) on the right-hand side can be replaced
by \( zP'_n(z) \) and \( P_{n+1}(z) \), respectively:

\[
(5.30a) \quad zP_n(z) = P_{n+1}(z) + \alpha_{n-n_j^\wedge}P_n(z) + \beta_{n-n_j^\wedge}P_{n-1}(z), \quad n_j^\wedge \leq n \leq n_j^\vee - 1,
\]

\[
(5.30b) \quad P_{n+1}(z) = P'_{n+1}(z) + \alpha_{n-n_j^\vee}P'_n(z) + \beta_{n-n_j^\vee}P'_{n-1}(z), \quad n_j^\vee \leq n \leq n_{j+1}^\vee - 2.
\]
It is also worth noting that the terms $e_j^\gamma(z) P_{n_j^\gamma}(z)$ and $z e_j^\gamma(z) P'_{n_j^\gamma}(z)$, which appear in (5.25) and are only nonzero if $h_j^\gamma > 1$ and $h_j^\gamma > 1$, respectively, are, according to (5.28), equal to $e_j^\hat{\gamma}(z) P'_{n_j^\gamma}(z)$ and $e_j^\gamma(z) P_{n_j^\gamma + 1}(z)$, respectively. In view of the definitions (5.27), the terms $\hat{W}_{n_j^\gamma - 1}(z) P_{n_j^\gamma}(z)$ and $W_{n_j^\gamma - 1}(z) P'_{n_j^\gamma}(z)$ can be written as $P_{n_j^\gamma}(z)$ and $P'_{n_j^\gamma + 1}(z)$, so that the recurrences (5.25) become

\[
\begin{align*}
(5.31a) & \quad P_{n_j^\gamma}(z) = P_{n_j^\gamma}(z) - e_j^\gamma(z) P'(n_j^\gamma - 1)(z), \quad j = 0, 1, \ldots, J^\gamma, \\
(5.31b) & \quad P_{n_j^\gamma + 1}(z) = z P_{n_j^\gamma + 1}(z) - e_j^\gamma(z) P'_{n_j^\gamma + 1}(z) - e_j^\gamma(z) P_{n_j^\gamma + 1}(z), \quad j = 0, 1, \ldots, J^\gamma - 1.
\end{align*}
\]

The notation $P_n^{(l)}$ indicates here and below that one can use either $P_n$ or $P_n'$.

Finally, in view of the later application to the nongeneric BCG process, we give the analog of Theorem 2.9. By an argument based on the orthogonality relations (5.5) and (5.9), which is analogous to the one for establishing (2.22), and by using (5.27), we obtain:

\[
\begin{align*}
(5.32a) & \quad \varphi_j^\gamma = \delta_j^\gamma / \delta_j^\gamma - 1, \quad \text{where} \quad \delta_j^\gamma = \Phi(z P_{n_j^\gamma - 1} P'_{n_j^\gamma - 1}), \quad \delta_j^\hat{\gamma} = \Phi(P_{n_j^\gamma} P_{n_j^\gamma}), \\
(5.32b) & \quad \Phi(e_j^\gamma P_{n_j^\gamma + k} P_{n_j^\gamma}) = \Phi(P_{n_j^\gamma + k} P_{n_j^\gamma}), \quad k = 1, \ldots, h_j^\gamma - 1; \\
(5.32c) & \quad \varphi_j^\gamma = \delta_j^\gamma / \delta_j, \quad \text{where} \quad \delta_j = \Phi(P_{n_j^\gamma} P_{n_j^\gamma}), \quad \delta_j^\gamma = \Phi(z P_{n_j^\gamma + 1} P_{n_j^\gamma}), \\
(5.32d) & \quad \Phi(z e_j^\gamma P_{n_j^\gamma + k} P_{n_j^\gamma + 1}) = \Phi(z P_{n_j^\gamma + k} P_{n_j^\gamma + 1}), \quad k = 1, \ldots, h_j^\gamma - 1.
\end{align*}
\]

If (2.10) holds, we may insert $P_{n_j^\gamma}$ and $P_{n_j^\gamma + 1}$ into (5.32b) and (5.32d), according to (5.29). We may also use (5.28b), although, in contrast to the situation in §2, there is no need to do that:

\[
\begin{align*}
(5.32e) & \quad \Phi(e_j^\hat{\gamma} P_{n_j^\gamma + k} P_{n_j^\gamma}) = \Phi(P_{n_j^\gamma + k}(z P_{n_j^\gamma - 1} - \alpha_{n_j^\gamma - 1} P_{n_j^\gamma - 1} - \beta_{n_j^\gamma - 2} P_{n_j^\gamma - 2})), \\
& \quad k = 1, \ldots, h_j^\gamma - 1, \\
(5.32f) & \quad \Phi(e_j^\gamma P_{n_j^\gamma + k} P_{n_j^\gamma + 1}) = \Phi(P_{n_j^\gamma + k}(z P_{n_j^\gamma + 1} - \alpha_{n_j^\gamma - 1} P_{n_j^\gamma + 1} - \beta_{n_j^\gamma - 2} P_{n_j^\gamma + 2})), \\
& \quad k = 1, \ldots, h_j^\gamma - 1.
\end{align*}
\]

**Theorem 5.4.** The linear systems (5.24) and (5.17) for computing the polynomials $e_j^\gamma$ and the constants $\varphi_j^\gamma$, $\varphi_j^\hat{\gamma}$ can be replaced by the equivalent system (5.32a)–(5.32d) consisting of single equations for $\varphi_j^\gamma$ and $\varphi_j^\hat{\gamma}$, and of a right lower triangular systems for the coefficients of $e_j^\gamma$ and $e_j^\hat{\gamma}$. If (2.10) holds, we may replace (5.32b) by (5.32e), and (5.32d) by (5.32f).

6. Matrix interpretations of staircase recurrences. In this section we give a matrix formulation of Theorem 5.3 on the mixed three-term recurrences for block staircase sequences of orthogonal polynomials. It is analogous to the matrix formulation in Theorem 3.1 for the diagonal three-term recurrence (Theorem 2.7 and Corollary 2.8). Whereas a block tridiagonal matrix $H$ emerged there, we first find here two block diagonal matrices, $E^\gamma$ and $E^\hat{\gamma}$, and a lower and an upper block bidiagonal matrix, $F^\gamma$ and $F^\hat{\gamma}$, respectively. They give rise to two further block bidiagonal matrices $G^\gamma := F^\gamma (E^\gamma)^{-1}$ and $G^\hat{\gamma} := F^\hat{\gamma} (E^\hat{\gamma})^{-1}$, which turn out to be block LU factors of $H$, but also block UL factors of another matrix $H'$ of the same structure, which belong to the functional $\Phi^\gamma$ instead of to $\Phi$. 

Again, we assume in view of the later application to the BCG method that the polynomial basis \( \{ W_m \} \) satisfies the three-term recurrence (2.10) whose matrix formulation is given by (3.1)–(3.3). Then Theorem 5.3, the definitions (5.27), and the relations (5.28) and (5.30) easily yield the following theorem.

**Theorem 6.1.** Let \( \{ P_n \} \) and \( \{ P^\prime_n \} \) be the sequences of FOPls corresponding to the functionals \( \Phi \) and \( \Phi' \), respectively, uniquely specified by (5.27), where \( \{ W_m \} \) is a sequence satisfying the three-term recurrence (2.10) (possibly with \( \alpha^W_m = \beta^W_m = 0, \forall m \in \mathbb{N} \)). Define the infinite row vectors

\[
\mathbf{p} := [P_0, P_1, \ldots], \quad \mathbf{p'} := [P'_0, P'_1, \ldots],
\]

and, for finite or infinite value of \( J^\wedge \), the infinite matrices

\[
E^\wedge := \text{block diag} \left[ E^\wedge_0, E^\wedge_1, \ldots, E^\wedge_{J^\wedge} \right], \quad E^\wedge := \text{block diag} \left[ E^\wedge_0, E^\wedge_1, \ldots, E^\wedge_{J^\wedge} \right],
\]

with square blocks \( E^\wedge_j \) and \( E^\wedge_j \) of order \( h^\wedge_j + h^\wedge_j - 1 \) that for \( j = 0, \ldots, J^\wedge - 1 \) are given by

\[
E^\wedge_j := \begin{bmatrix}
I^\wedge_j & 0 & 0 \\
0 & I^\wedge_j & -e^\wedge_j \\
0^T & 0^T & 1
\end{bmatrix}, \quad E^\wedge_j := \begin{bmatrix}
I^\wedge_j & -e^\wedge_j & 0 \\
0^T & 1 & 0^T \\
0 & 0 & I^\wedge_j
\end{bmatrix}.
\]

Here, \( I^\wedge_j \) and \( I^\wedge_j \) are the unit matrices of order \( h^\wedge_j - 1 \) and \( h^\wedge_j - 1 \), respectively, and the row vectors

\[
e^\wedge_j := [e^\wedge_{0,j}, e^\wedge_{1,j}, \ldots, e^\wedge_{h^\wedge_j-2,j}]^T, \quad e^\wedge_j := [e^\wedge_{0,j}, e^\wedge_{1,j}, \ldots, e^\wedge_{h^\wedge_j-2,j}]^T
\]

contain the coefficients of the polynomials

\[
e^\wedge_j(z) = \sum_{i=0}^{h^\wedge_j-2} e^\wedge_{i,j} W_i(z), \quad e^\wedge_j(z) = \sum_{i=0}^{h^\wedge_j-2} e^\wedge_{i,j} W_i(z)
\]

from Theorem 5.3, expressed in terms of the basis \( \{ W_m \} \). If \( h^\wedge_j = 1 \) or \( h^\wedge_j = 1 \), the rows and columns containing \( I^\wedge_j \) and \( I^\wedge_j \), respectively, are missing in \( E^\wedge_j \) and \( E^\wedge_j \).

If \( J^\wedge < \infty \) (and thus \( J^\wedge < \infty \) also), then \( E^\wedge_{J^\wedge} := I \) (the infinite unit matrix); if \( J^\wedge = J^\wedge - 1 \), \( E^\wedge_{J^\wedge} := I \) also, while, if \( J^\wedge = J^\wedge - 1 \), then \( E^\wedge_{J^\wedge} \) has the same structure as in (6.3), with \( I^\wedge_{J^\wedge} := I \).

Also define the infinite block bidiagonal matrices

\[
F^\wedge := \begin{bmatrix}
F^\wedge_0 & F^\wedge_1 & \cdots & F^\wedge_{J^\wedge} \\
L^\wedge_0 & F^\wedge_1 & \cdots & F^\wedge_{J^\wedge} \\
L^\wedge_1 & F^\wedge_1 & \cdots & F^\wedge_{J^\wedge} \\
\vdots & \vdots & \ddots & \vdots \\
L^\wedge_{J^\wedge - 1} & \cdots & \cdots & F^\wedge_{J^\wedge}
\end{bmatrix}, \quad F^\wedge := \begin{bmatrix}
L^\wedge_0 & F^\wedge_1 & F^\wedge_2 & \cdots & F^\wedge_{J^\wedge} \\
L^\wedge_1 & L^\wedge_2 & \cdots & \cdots & \cdots \\
\vdots & \vdots & \ddots & \cdots & \cdots \\
\vdots & \vdots & \cdots & \vdots & \cdots \\
L^\wedge_{J^\wedge} & \cdots & \cdots & \cdots & \cdots
\end{bmatrix},
\]

with square blocks \( F^\wedge_j \) and \( L^\wedge_j \) of order \( h^\wedge_j + h^\wedge_j - 1 \), whose upper left corner is the \((n^\wedge_j, n^\wedge_j)\)-entry\(^3\) of \( F^\wedge \) and \( F^\wedge \), respectively, and which for \( j = 0, 1, \ldots, J^\wedge - 1 \) are

---

\(^{3}\) The entries in the upper left corner of \( F^\wedge \) and \( F^\wedge \) are considered as \((0,0)\)-entries. The entries in the subblocks are in this text identified by their indices in the full matrix, e.g., the \((n^\wedge_j, n^\wedge_j)\) entry of \( F^\wedge_j \) is the same as the \((n^\wedge_j, n^\wedge_j)\) entry of \( F^\wedge \). Likewise, the \( n^\wedge_j \)th row (column) of \( F^\wedge_j \) is the \( n^\wedge_j \)th row (column) of \( F^\wedge \).
given as follows: If \( h_j^\gamma > 1 \) and \( h_j^\nu > 1 \),

\[
(6.7a) \quad F_j^\gamma := \begin{bmatrix}
T_j^\gamma & 0 & f_j^\gamma \\
O & 0^T & 0 \\
O & I_j^\gamma & 0
\end{bmatrix}, \quad L_j^\gamma := \begin{bmatrix}
I_j^\gamma & 0 & 0 \\
0 & c_j^\gamma & T_j^\nu \\
0^T & 0 & I_j^\nu
\end{bmatrix};
\]

whereas, if \( h_j^\gamma = 1 \) and \( h_j^\nu > 1 \),

\[
(6.7b) \quad F_j^\gamma := \begin{bmatrix}
0^T & \varphi_j^\gamma \\
I_j^\gamma & 0
\end{bmatrix}, \quad L_j^\gamma := \begin{bmatrix}
c_j^\gamma & T_j^\nu \\
0 & I_j^\nu
\end{bmatrix};
\]

and, if \( h_j^\gamma > 1 \) and \( h_j^\nu = 1 \),

\[
(6.7c) \quad F_j^\gamma := \begin{bmatrix}
T_j^\gamma & f_j^\gamma \\
I_j^\gamma & 0
\end{bmatrix}, \quad L_j^\gamma := \begin{bmatrix}
I_j^\gamma & 0 \\
0^T & 1
\end{bmatrix},
\]

which reduces even further to

\[
(6.7d) \quad F_j^\gamma := [\varphi_j^\gamma], \quad L_j^\gamma := [1],
\]

if both \( h_j^\gamma = 1 \) and \( h_j^\nu = 1 \). If \( T_m^W \) denotes the tridiagonal matrix of order \( m + 1 \) with entries \( \beta_k^W \), \( \alpha_k^W \) and 1 in its \((k+1)\)th column, cf. (3.2), then the blocks \( F_j^\gamma, F_j^\nu, L_j^\gamma, \) and \( L_j^\nu \) contain

\[
(6.8a) \quad T_j^\gamma := T_{h_j^\gamma-2}^W, \quad T_j^\nu := T_{h_j^\nu-2}^W,
\]

\[
(6.8b) \quad f_j^\nu := [\varphi_j^\nu, 0, \ldots, 0]^T \in C^{h_j^\nu-1}, \quad f_j^\gamma := [\varphi_j^\gamma, 0, \ldots, 0]^T \in C^{h_j^\gamma-1},
\]

\[
(6.8c) \quad c_j^\nu := [1, 0, \ldots, 0]^T \in C^{h_j^\nu-1}, \quad c_j^\gamma := [1, 0, \ldots, 0]^T \in C^{h_j^\gamma-1},
\]

\[
(6.8d) \quad l_j^{\nu T} := [0, \ldots, 0, 1] \in C^{h_j^\nu-1}, \quad l_j^\gamma := [0, \ldots, 0, 1] \in C^{h_j^\gamma-1}.
\]

If \( J^\nu = J^\gamma - 1 < \infty \),

\[
(6.8e) \quad F_j^{\nu \wedge} := T^W, \quad L_j^{\nu \wedge} := I,
\]

while, if \( J^\nu = J^\gamma < \infty \),

\[
(6.8f) \quad F_j^{\gamma \wedge} := \begin{bmatrix}
T_j^{\gamma \wedge} & 0 \\
I_j^{\gamma T} & 0^T
\end{bmatrix}, \quad L_j^{\gamma \wedge} := \begin{bmatrix}
I_j^{\gamma \wedge} & 0 & 0 \\
0 & c_j^{\gamma \wedge} & T^W
\end{bmatrix},
\]

where \( c_j^{\gamma \wedge} = [1, 0, 0, \ldots]^T \) has infinitely many components.

The off-diagonal blocks \( F_j^{\gamma \wedge} \) and \( L_j^{\gamma \wedge} \) are rank-one matrices of size \((h_j^{\gamma - 1} + h_j^{\gamma - 1} - 1) \times (h_j^{\gamma + h_j^{\nu} - 1}) \) and \((h_j^{\gamma + h_j^{\nu} - 1} - 1) \times (h_j^{\gamma} + h_j^{\nu} - 1) \), respectively. Each has a single nonzero element, namely, the \((n_j^{\gamma - 1}, n_j^{\gamma})\) entry \( \varphi_j^{\gamma \wedge} \) of \( F_j^{\gamma \wedge} \) and the \((n_j^{\gamma - 1}, n_j^{\gamma + 1})\) entry 1 of \( F_j^{\nu \wedge} \), which lies in the upper right corner of \( L_j^{\gamma \wedge} \). If \( j < J^\gamma \) and \( h_j^{\gamma - 1}, h_j^{\gamma + h_j^{\nu} - 1}, h_j^{\gamma}, h_j^{\gamma + 1}, h_j^{\nu} > 1 \), then \( F_j^{\gamma \wedge} \) and \( L_j^{\gamma \wedge} \) have the structure

\[
(6.9) \quad F_j^{\gamma} := \begin{bmatrix}
O & 0 & O \\
O & f_j^{\gamma} & O \\
0^T & 0 & 0^T
\end{bmatrix}, \quad L_j^{\gamma} := \begin{bmatrix}
O & O & c_j^{\gamma} \\
0^T & 0^T & 0 \\
O & O & 0
\end{bmatrix},
\]
with diagonal blocks of order \((h_{j-1}^\wedge - 1) \times (h_{j-1}^\wedge - 1), (h_j^\wedge - 1) \times 1, 1 \times (h_j^\wedge - 1),\) and 
\((h_{j+1}^\wedge - 1) \times (h_{j+1}^\wedge - 1), 1 \times (h_{j+1}^\wedge - 1), (h_j^\wedge - 1) \times 1,\) respectively. If one or several of these sizes are zero, the structure is again modified, but the nonzero entry of \(L_j^\wedge\) is always in the upper right corner; the one of \(F_j^\wedge\) is in the first column if \(h_j^\wedge = 1,\) in the last column if \(h_j^\wedge = 1,\) and in the last row if \(h_{j-1}^\wedge = 1.\) If \(J^\wedge = J^\wedge - 1 < \infty, F_j^\wedge = 0;\) and if \(J^\wedge = J^\wedge < \infty, F_j^\wedge\) has the same structure as in (6.9), but the last block column is infinitely wide.

In terms of the above-defined quantities, the mixed recurrences (5.25) for the sequences \(\{P_n^\wedge\}\) and \(\{P'_n^\wedge\}\) can be written as

\[
\begin{align*}
\text{(6.10a)} & \quad p(z)E^\wedge = p'(z)F^\wedge, \\
\text{(6.10b)} & \quad zp'(z)E^\wedge = p(z)F^\wedge.
\end{align*}
\]

Likewise, if \(E_{[n]}^\wedge, E^\wedge_{[n]}, F^\wedge_{[n]}, F^\wedge_{[n]}\) denote the principal submatrices of order \(n + 1\) of \(E^\wedge, E^\wedge, F^\wedge,\) and \(F^\wedge,\) respectively, and if

\[
\begin{align*}
\text{(6.11)} & \quad p_n := [P_0, P_1, \ldots, P_n], \quad p'_n := [P'_0, P'_1, \ldots, P'_n],
\end{align*}
\]

then we have

\[
\begin{align*}
\text{(6.12a)} & \quad p_n(z)E^\wedge_{[n]} = p'_n(z)F^\wedge_{[n]}, \\
\text{(6.12b)} & \quad zp'_n(z)E^\wedge_{[n]} = p_n(z)F^\wedge_{[n]} + [0, \ldots, 0, P_{n+1}(z)].
\end{align*}
\]

Remarks. (i) \(E^\wedge\) and \(E^\wedge\) are unit upper triangular and block diagonal, and the same holds for their inverses, which have the blocks

\[
\begin{align*}
\text{(6.13)} & \quad (E_j^\wedge)^{-1} = \begin{bmatrix} I_j^\wedge & 0 & 0 \\
0 & I_j^\wedge & e_j^\wedge \\
0^T & 0^T & 1 \end{bmatrix}, \quad (E_j^\wedge)^{-1} = \begin{bmatrix} I_j^\wedge & e_j^\wedge & 0 \\
0^T & 1 & 0^T \\
0 & 0 & I_j^\wedge \end{bmatrix},
\end{align*}
\]

differing from \(E_j^\wedge\) and \(E^\wedge\) by the missing minus sign in front of \(e_j^\wedge\) and \(e_j^\wedge\) only.

(ii) \(F^\wedge\) is unit upper Hessenberg, and \(F^\wedge\) is unit upper triangular.

(iii) If \(h_j^\wedge = h_j^\wedge = 1 (\forall j), E^\wedge\) and \(E^\wedge\) are the infinite unit matrix, and \(F^\wedge\) and \(F^\wedge\) are lower and upper bidiagonal,

\[
\begin{align*}
\text{(6.14)} & \quad F^\wedge = \begin{bmatrix} \varphi_1^\wedge & \varphi_2^\wedge & \cdots \\
1 & 1 & \varphi_2^\wedge \\
& & \ddots \end{bmatrix}, \quad F^\wedge = \begin{bmatrix} 1 & \varphi_1^\wedge \\
& 1 & \varphi_2^\wedge \\
& & \ddots \end{bmatrix}.
\end{align*}
\]

By using (6.13), we can readily turn the relations (6.10) into formally explicit formulas for \(p\) and \(p',\) which, however, are used in an implicit way as recurrences for the elements of \(p'\) and \(p,\) respectively, appearing on the right-hand side.

**Corollary 6.2.** Under the assumptions of Theorem 6.1, there holds

\[
\begin{align*}
\text{(6.15a)} & \quad p(z) = p'(z) G^\wedge \\
\text{(6.15b)} & \quad zp'(z) = p(z) G^\wedge,
\end{align*}
\]
where

\[
(6.16a) \quad G^\gamma = F^\gamma (E^\gamma)^{-1} := \begin{bmatrix}
G_0^\gamma & G_1^\gamma & G_2^\gamma \\
L_0^\gamma & L_1^\gamma & \cdots \\
& \ddots & \\
L_{j_\wedge^\gamma - 1}^\gamma & & G_{j_\wedge^\gamma}^\gamma
\end{bmatrix}
\]

and

\[
(6.16b) \quad G^\wedge = F^\wedge (E^\wedge)^{-1} := \begin{bmatrix}
G_0^\wedge & F_1^\wedge & F_2^\wedge \\
G_1^\wedge & \ddots & \\
& \ddots & F_{j_\wedge^\wedge}^\wedge \\
G_{j_\wedge^\wedge}^\wedge
\end{bmatrix}
\]

have the same off-diagonal blocks as \(F^\gamma\) and \(F^\wedge\). The diagonal blocks \(G_j^\gamma\) and \(G_j^\wedge\), \(j = 0, 1, \ldots, J_\wedge^\gamma - 1\), are defined as follows: If \(h_j^\gamma > 1\) and \(h_j^\wedge > 1\),

\[
(6.17a) \quad G_j^\gamma := \begin{bmatrix}
T_j^\gamma & 0^T & f_j^\gamma \\
0 & 0 & 0 \\
0 & I_j^\gamma & e_j^\gamma
\end{bmatrix}, \quad G_j^\wedge := \begin{bmatrix}
I_j^\wedge & e_j^\wedge & 0 \\
0 & e_j^\wedge & 0 \\
0 & 0 & I_j^\wedge^T
\end{bmatrix};
\]

whereas, if \(h_j^\wedge = 1\) and \(h_j^\gamma > 1\),

\[
(6.17b) \quad G_j^\gamma := \begin{bmatrix}
0^T & \varphi_j^\gamma & f_j^\gamma \\
0 & I_j^\gamma & e_j^\gamma
\end{bmatrix}, \quad G_j^\wedge := \begin{bmatrix}
I_j^\wedge^T & e_j^\wedge \\
0 & e_j^\wedge & 0
\end{bmatrix} = L_j^\wedge;
\]

if \(h_j^\wedge > 1\) and \(h_j^\gamma = 1\),

\[
(6.17c) \quad G_j^\gamma := \begin{bmatrix}
T_j^\gamma & f_j^\gamma \\
0 & e_j^\gamma
\end{bmatrix} = F_j^\gamma, \quad G_j^\wedge := \begin{bmatrix}
I_j^\wedge & e_j^\wedge \\
0 & 1
\end{bmatrix};
\]

and, if both \(h_j^\wedge = 1\) and \(h_j^\gamma = 1\),

\[
(6.17d) \quad G_j^\gamma := [\varphi_j^\gamma] = F_j^\gamma, \quad G_j^\wedge := [1] = L_j^\wedge.
\]

If \(J^\gamma = J_\wedge^\gamma - 1 < \infty\),

\[
(6.17e) \quad G_{j_\wedge^\gamma}^\gamma := F_{j_\wedge^\gamma}^\gamma = T^W, \quad G_{j_\wedge^\gamma}^\wedge := L_{j_\wedge^\gamma}^\wedge = I,
\]

while, if \(J^\gamma = J_\wedge^\gamma < \infty\),

\[
(6.17f) \quad G_{j_\wedge^\gamma}^\wedge := F_{j_\wedge^\gamma}^\wedge := \begin{bmatrix}
T_{j_\wedge^\gamma}^\wedge & O \\
I_{j_\wedge^\gamma}^\wedge & 0^T
\end{bmatrix}, \quad G_{j_\wedge^\gamma}^\wedge := \begin{bmatrix}
I_{j_\wedge^\gamma}^\wedge & e_{j_\wedge^\gamma}^\wedge \\
O & e_{j_\wedge^\gamma}^\wedge & T^W
\end{bmatrix}.
\]

Note that the two relations (6.15) are just the recurrences (5.30) and (5.31) expressed in matrix notation.
Table 1

Computation of $H$ from $G^v$ and $G^\wedge$ by equating the elements of $H$ with those of $G^v G^\wedge$.

<table>
<thead>
<tr>
<th>$h_{j-1}^\wedge$</th>
<th>$h_j^\wedge$</th>
<th>$h_j^v$</th>
<th>$n_{i-1}$</th>
<th>$n_i$</th>
<th>$n_{i+1}$</th>
<th>$n_{i+2}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$&gt; 1$</td>
<td>$&gt; 1$</td>
<td>$&gt; 1$</td>
<td>$n_{j-1}^\wedge + 1$</td>
<td>$n_j^\wedge$</td>
<td>$n_j^v + 1$</td>
<td>$n_{j+1}^\wedge$</td>
</tr>
<tr>
<td>$= 1$</td>
<td>$&gt; 1$</td>
<td>$&gt; 1$</td>
<td>$n_j^\wedge + 1$</td>
<td>$n_j^\wedge = n_j^v + 1$</td>
<td>$n_j^\wedge + 1$</td>
<td>$n_{j+1}^\wedge$</td>
</tr>
<tr>
<td>$= 1$</td>
<td>$= 1$</td>
<td>$&gt; 1$</td>
<td>$n_j^\wedge = n_j^v + 1$</td>
<td>$n_{j+1}^\wedge = n_{j+1}^v + 1$</td>
<td>$n_{j+1}^\wedge$</td>
<td></td>
</tr>
<tr>
<td>$= 1$</td>
<td>$= 1$</td>
<td>$&gt; 1$</td>
<td>$n_j^\wedge = n_j^\wedge + 1 = n_j^v$</td>
<td>$n_{j+1}^\wedge = n_{j+1}^v + 1$</td>
<td>$n_{j+1}^\wedge$</td>
<td></td>
</tr>
<tr>
<td>$= 1$</td>
<td>$= 1$</td>
<td>$= 1$</td>
<td>$n_j^\wedge = n_j^\wedge + 1$</td>
<td>$n_{j+1}^\wedge = n_{j+1}^v + 1$</td>
<td>$n_{j+1}^\wedge$</td>
<td></td>
</tr>
<tr>
<td>$= 1$</td>
<td>$= 1$</td>
<td>$= 1$</td>
<td>$n_j^\wedge = n_j^\wedge + 1$</td>
<td>$n_{j+1}^\wedge = n_{j+1}^v + 1$</td>
<td>$n_{j+1}^\wedge$</td>
<td></td>
</tr>
</tbody>
</table>

7. Matrix relations between diagonal and staircase recurrences: The nongeneric qd algorithm. From the two relations (6.15) we can eliminate either $p$ or $p'$ to obtain

(7.1a) \[ zp(z) = p(z) G^v G^\wedge \]

and

(7.1b) \[ zp'(z) = p'(z) G^\wedge G^v. \]

However, these are—in matrix notation—just the recurrences for the FOP1s $P_n$ and those for the FOP1s $P_n'$. Therefore, they must be identical to (3.11), which describes this recurrence for $P_n$, and with the corresponding relation for the set $\{P_n'\}$, respectively. This leads to the following result.

Theorem 7.1. Under the assumptions of Theorem 6.1 let $H$ and $H'$ be the block tridiagonal matrices from (3.6) that describe, according to (3.11), the recurrences for the FOP1s $\{P_n\}$ and $\{P_n'\}$, respectively, that correspond to the linear functionals $\Phi$ and $\Phi'$. Then $H$ has the block LU factorization

(7.2a) \[ H = G^v G^\wedge, \]

and $H'$ has the block UL factorization

(7.2b) \[ H' = G^\wedge G^v, \]

where $G^v$ and $G^\wedge$ are the block bidiagonal matrices defined by (6.16)–(6.17) and (6.9). The resulting relation between the entries of $G^v$, $G^\wedge$, and $H$ or $H'$ are listed in Tables.
Computation of $H'$ from $G^\vee$ and $G^\wedge$ by equating the elements of $H'$ with those of $G^\wedge G^\vee$.

<table>
<thead>
<tr>
<th>$h_j^\wedge$</th>
<th>$h_j^\vee$</th>
<th>$h_{j+1}^\wedge$</th>
<th>$h_{j+1}^\vee$</th>
<th>$n_{i-1}^i$</th>
<th>$n_i^i$</th>
<th>$n_{i+1}^i$</th>
<th>$n_{i+2}^i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$&gt; 1$</td>
<td>$&gt; 1$</td>
<td>$&gt; 1$</td>
<td>$n_j^\wedge$</td>
<td>$n_j^\wedge$</td>
<td>$n_{j+1}^\wedge$</td>
<td>$n_{j+1}^\wedge$</td>
<td>$n_{j+1}^\wedge$</td>
</tr>
<tr>
<td>$= 1$</td>
<td>$&gt; 1$</td>
<td>$&gt; 1$</td>
<td>$n_j^\vee$</td>
<td>$n_j^\vee$</td>
<td>$n_j^\vee$</td>
<td>$n_{j+1}^\vee$</td>
<td>$n_{j+1}^\vee$</td>
</tr>
<tr>
<td>$= 1$</td>
<td>$= 1$</td>
<td>$&gt; 1$</td>
<td>$n_j^\wedge$</td>
<td>$n_j^\vee$</td>
<td>$n_j^\wedge$</td>
<td>$n_j^\vee$</td>
<td>$n_{j+1}^\wedge$</td>
</tr>
<tr>
<td>$&gt; 1$</td>
<td>$&gt; 1$</td>
<td>$= 1$</td>
<td>$n_j^\vee$</td>
<td>$n_j^\vee$</td>
<td>$n_j^\wedge$</td>
<td>$n_j^\vee$</td>
<td>$n_{j+1}^\vee$</td>
</tr>
<tr>
<td>$= 1$</td>
<td>$= 1$</td>
<td>$= 1$</td>
<td>$n_j^\vee$</td>
<td>$n_j^\wedge$</td>
<td>$n_j^\vee$</td>
<td>$n_j^\wedge$</td>
<td>$n_{j+1}^\vee$</td>
</tr>
<tr>
<td>$= 1$</td>
<td>$= 1$</td>
<td>$= 1$</td>
<td>$n_j^\wedge$</td>
<td>$n_j^\vee$</td>
<td>$n_j^\vee$</td>
<td>$n_j^\vee$</td>
<td>$n_{j+1}^\wedge$</td>
</tr>
</tbody>
</table>

1 and 2. There $n_i$ and $n_i'$ denote the sequences of the indices of the regular FOP1s out of $\{P_n\}$ and $\{P_n'\}$, respectively; the entries of $H'$ are also distinguished by a prime from those of $H$. The functions $\Omega_j$ and $\Omega'_j$ are defined by

\[
\Omega_j(e_j^\wedge, \varphi) := T_{h_j^\wedge -1}^W \begin{bmatrix} e_j^\wedge \\ -1 \end{bmatrix} + \begin{bmatrix} \varphi \\ 0 \\ \vdots \\ 0 \end{bmatrix},
\]

(7.3a)

\[
\Omega'_j(e_j^\vee, \varphi) := T_{h_j^\vee -1}^W \begin{bmatrix} e_j^\vee \\ -1 \end{bmatrix} + \begin{bmatrix} \varphi \\ 0 \\ \vdots \\ 0 \end{bmatrix}.
\]

(7.3b)

Remark. In the case $W_m(z) = z^m$, the definitions (7.3) reduce to

\[
\Omega_j(e_j^\wedge, \varphi) := \begin{bmatrix} \varphi \\ e_j^\wedge \end{bmatrix}, \quad \Omega'_j(e_j^\wedge, \varphi) := \begin{bmatrix} \varphi \\ e_j^\vee \end{bmatrix}.
\]

(7.4)

Proof. We have already derived (7.2), so it remains to relate the entries of $H$ and $H'$ to those of $G^\vee$ and $G^\wedge$. For this we need to compute explicitly the elements of the products $G^\vee G^\wedge$ and $G^\wedge G^\vee$. The task is complicated by the fact that the structure of the blocks depends on the quantities $h_j^\wedge$ and $h_j^\vee$ being larger than or equal to 1.

Let us start with the off-diagonal blocks of the products. First, the $(j+1,j)$-blocks of $G^\vee G^\wedge$ and $G^\wedge G^\vee$ are $L_j^\vee G_j^\wedge = L_j^\vee$ and $G_j^\wedge L_j^\vee = L_j^\vee$, respectively. The
(j − 1, 1)-blocks are $G_{j−1}^\gamma F_j^\wedge$ and $F_j^\wedge G_j^\gamma$, and their structure depends on the value of $h_j^\gamma$ and $h_{j−1}^\gamma$, respectively. Since $F_j^\wedge$ has only one nonzero entry, namely, the $(n_j^\gamma−1, n_j^\gamma)$ entry $\varphi_j^\gamma$ of $G^\wedge$, only the $n_j^\gamma−1$th column of $G_j^\gamma$ and the $n_j^\gamma$th row of $G_j^\wedge$ matter. (Columns and rows of the blocks of $G^\gamma$ and $G^\wedge$ are here again numbered according to their indices in the whole matrix.) The $n_j^\gamma−1$th column of $G_j^\gamma$ contains 1 as its $(n_j^\gamma−1, n_j^\gamma)$ entry if $h_j^\gamma > 1$, and $\varphi_j^\gamma$ as its $(n_j^\gamma−1, n_j^\gamma)$ entry if $h_j^\gamma = 1$. Consequently,

\[
G_j^\gamma F_j^\wedge = \begin{cases} 
\varphi_j^\gamma & \text{if } (m, n) = (n_j^\gamma−1, n_j^\gamma) \\
0 & \text{otherwise}
\end{cases} \quad \text{if } h_j^\gamma > 1,
\]

and

\[
F_j^\wedge G_j^\gamma = \begin{cases} 
\varphi_j^\gamma & \text{if } (m, n) = (n_j^\gamma−1, n_j^\gamma) \\
0 & \text{otherwise}
\end{cases} \quad \text{if } h_j^\gamma = 1.
\]

The $n_j^\gamma$th row of $G_j^\gamma$ contains 1 as its $(n_j^\gamma, n_j^\gamma−1)$ entry if $h_j^\gamma > 1$, and $\varphi_j^\gamma$ as its $(n_j^\gamma−1, n_j^\gamma−1)$ entry if $h_j^\gamma = 1$, so that

\[
F_j^\wedge F_j^\gamma = \begin{cases} 
\varphi_j^\gamma & \text{if } (m, n) = (n_j^\gamma−1, n_j^\gamma−1) \\
0 & \text{otherwise}
\end{cases} \quad \text{if } h_j^\gamma > 1,
\]

and

\[
F_j^\wedge G_j^\gamma = \begin{cases} 
\varphi_j^\gamma & \text{if } (m, n) = (n_j^\gamma−1, n_j^\gamma) \\
0 & \text{otherwise}
\end{cases} \quad \text{if } h_j^\gamma = 1.
\]

The diagonal blocks of $G^\gamma G^\wedge$ and $G^\wedge G^\gamma$ are $G_j^\gamma G_j^\wedge + L_j^\gamma−1 F_j^\wedge$ and $G_j^\wedge G_j^\gamma + F_j^\wedge+1 L_j^\wedge$. Here we obtain

\[
L_j^\gamma−1 F_j^\wedge = \mathbf{0} \quad \text{if } h_j^\gamma−1 > 1,
\]

\[
L_j^\gamma−1 F_j^\wedge = \begin{cases} 
\varphi_j^\gamma & \text{if } (m, n) = (n_j^\gamma−1, n_j^\gamma−1) \\
0 & \text{otherwise}
\end{cases} \quad \text{if } h_j^\gamma−1 = 1,
\]

\[
F_j^\wedge+1 L_j^\wedge = \mathbf{0} \quad \text{if } h_j^\wedge+1 > 1,
\]

\[
F_j^\wedge+1 L_j^\wedge = \begin{cases} 
\varphi_j^\wedge+1 & \text{if } (m, n) = (n_j^\gamma, n_j^\gamma−1) \\
0 & \text{otherwise}
\end{cases} \quad \text{if } h_j^\wedge+1 = 1.
\]

Furthermore, if $h_j^\gamma > 1$ and $h_j^\wedge > 1$,}

\[
G_j^\gamma G_j^\wedge = \begin{bmatrix}
T_j^\wedge & T_j^\wedge e_j^\wedge & f_j^\wedge y_j^\gamma T_j^\wedge \\
\varepsilon_j^\wedge−j & 0 & 0 \\
0 & c_j^\gamma & T_j^\gamma + e_j^\gamma y_j^\gamma T_j^\wedge
\end{bmatrix},
\]

\[
G_j^\wedge G_j^\gamma = \begin{bmatrix}
T_j^\wedge & e_j^\wedge y_j^\gamma T_j^\wedge & f_j^\gamma y_j^\gamma T_j \\\
c_j^\gamma y_j^\gamma T_j^\wedge & T_j^\gamma & T_j^\gamma e_j^\gamma y_j^\gamma \\
0 & y_j^\gamma T_j & \varepsilon_j^\wedge−j
\end{bmatrix};
\]

(7.9a)
if \( h^\gamma_j = 1 \) and \( h^\gamma_j > 1 \),

\[
G_j^\gamma G_j^\gamma = \begin{bmatrix}
0 & \varphi_j^\gamma I_j^T \\
\varepsilon_j^\gamma & T_j^T + e_j^\gamma I_j^T
\end{bmatrix}, \quad G_j^\gamma G_j^\gamma = \begin{bmatrix}
T_j^\gamma & c_j^\gamma \varphi_j^\gamma + T_j^\gamma e_j^\gamma \\
\varphi_j^\gamma I_j^T & \varepsilon_j^\gamma - 2, j
\end{bmatrix};
\]

if \( h^\gamma_j > 1 \) and \( h^\gamma_j = 1 \),

\[
G_j^\gamma G_j^\gamma = \begin{bmatrix}
T_j^\gamma & T_j^\gamma e_j^\gamma + f_j^\gamma \\
\varphi_j^\gamma I_j^T & \varepsilon_j^\gamma - 2, j
\end{bmatrix}, \quad G_j^\gamma G_j^\gamma = \begin{bmatrix}
T_j^\gamma + \varphi_j^\gamma I_j^T & f_j^\gamma \\
\varphi_j^\gamma I_j^T & 0
\end{bmatrix};
\]

and, if \( h^\gamma_j = h^\gamma_j = 1 \),

\[
G_j^\gamma G_j^\gamma = [\varphi_j^\gamma], \quad G_j^\gamma G_j^\gamma = [\varphi_j^\gamma].
\]

To relate the entries of \( H \) and \( H' \) with those of \( G^\gamma G^\gamma \) and \( G^\gamma G^\gamma \), we need to associate the blocks \( A_i, B_i, C_i \) of \( H \) and the blocks \( A_i', B_i', C_i' \) of \( H' \) with those of \( G^\gamma G^\gamma \) and \( G^\gamma G^\gamma \), respectively. However, depending on \( h^\gamma_j \) and \( h^\gamma_j \) being equal to or greater than 1, a diagonal block

\[ H_j^\gamma := G_j^\gamma G_j^\gamma + L_j^\gamma F_j; \]

of order \( h^\gamma_j + h^\gamma_j - 1 \) of \( G^\gamma G^\gamma = H \) and such a block

\[ H_j^\gamma := G_j^\gamma G_j^\gamma + F_{j+1}^\gamma L_j^\gamma \]

of \( G^\gamma G^\gamma = H' \) corresponds either to a single block \( A_i \) or \( A_i' \) or to a 2 \times 2 block matrix

\[
\begin{bmatrix}
A_i & B_{i+1} \\
C_i & A_{i+1}
\end{bmatrix} \quad \text{or} \quad \begin{bmatrix}
A_{i-1} & B'_i \\
C'_{i-1} & A'_i
\end{bmatrix},
\]

respectively. Likewise, the off-diagonal blocks may correspond to a \( 1 \times 1, 1 \times 2, 2 \times 1, \) or \( 2 \times 2 \) block matrix.

The indices of the regular elements of the sequences \( \{P_n\} \) and \( \{P_n'\} \) are now denoted by \( n_i \) and \( n_i' \), respectively, while \( \{n_i^\gamma\} \subseteq \{n_i\} \) and \( \{n_i^\gamma\} \subseteq \{n_i'\} \) still denote the index subsequences of the regular elements in the generalized staircase. Since the upper left corners of \( A_i \) and \( G_j^\gamma G_j^\gamma \) are at \( (n_i, n_i) \) and \( (n_j^\gamma, n_j^\gamma) \), the association of the blocks is based on the identification \( n_i = n_j^\gamma \). Likewise, the upper left corner of \( A_i' \) at \( (n_i', n_i') \) corresponds to the \( (n_j^\gamma, n_j^\gamma) \) element of \( G_j^\gamma G_j^\gamma \), hence we have \( n_i' = n_j^\gamma \).

From (7.3)-(7.9) we can then read off the associations listed in Tables 1 and 2.

Tables 1 and 2 describe the structure and the entries of the products \( H = G^\gamma G^\gamma \) and \( H' = G^\gamma G^\gamma \) in terms of the entries of \( G^\gamma \) and \( G^\gamma \). Next, we are interested in inverting these two operations, i.e., in computing the lower block bidiagonal matrix \( G^\gamma \) and the upper block bidiagonal matrix \( G^\gamma \), either from \( H \) or from \( H' \). Of course, \( G^\gamma \) and \( G^\gamma \) are required to have the structure specified by (6.16) and (6.17). While \( H = G^\gamma G^\gamma \) is a block LU decomposition, \( H' = G^\gamma G^\gamma \) is a (not so often encountered) block UL decomposition. Theoretically, in view of

\[
(H')^{-1} = (G^\gamma)^{-1} (G^\gamma)^{-1},
\]

the latter could be obtained via a block LU decomposition of \( (H')^{-1} \) followed by the inversion of the factors, but we can directly determine the block UL decomposition.
with ease. From Theorem 7.1 we know that these two block decompositions exist, but it is not yet clear how the sizes of the blocks of $G^\vee$ and $G^\wedge$ can be determined from those in $H$ or from those in $H'$.

In the $j$th step of the block LU decomposition, $H = G^\vee G^\wedge$, the block pivot element is obtained by subtracting $L_{j-1}^\vee F_j^\wedge$ from a diagonal block $H_j^\wedge$ of $H$ of the appropriate size. It must then be split into $G_j^\vee G_j^\wedge$:

$$H_j^\wedge - L_{j-1}^\vee F_j^\wedge = G_j^\vee G_j^\wedge.$$  

(7.13)

However, given $H$ structured according to (3.6)-(3.10), we do not know a priori the sizes of these diagonal blocks, since they may correspond to a single block $A_i$ or to a $2 \times 2$ block matrix given on the left-hand side of (7.11). (This is also the reason for calling these blocks $H_j^\wedge$ and not $H_j^\wedge$; the block sizes are the same as in $G^\wedge$ and $G^\vee$.) But any block pivot has to be nonsingular, and from this requirement we can determine the correct size of the block $H_j^\wedge$. Any tentative $1 \times 1$ block pivot is either a $1 \times 1$ matrix obtained by updating $[\alpha_{0,0}]$ or a unit upper Hessenberg matrix. The latter is a companion matrix if we assume for the moment that $W_m(z) = z^m$. Hence it is nonsingular if and only if the element in its upper right corner is nonzero. Thus, identifying $n_i = n_i^\vee$, we then have $H_j^\wedge := A_i$ if and only if $\alpha_{0,0} \neq 0$ or the $(n_i, n_{i+1} - 1)$-element of $L_{j-1}^\vee F_j^\wedge$ does not vanish, i.e., in view of (7.7) and $n_i^\vee = n_i + 1 - 1$ (cf. Table 1), and that this quantity can be computed at this moment. Moreover, in the case of a $1 \times 1$ block pivot, i.e., when (7.14) holds, we conclude from Table 1 and (7.4) that $\varphi_j^\vee = \varphi_j^\vee$. Hence, (7.14) means that we test whether the tentative value $\varphi_j^\vee$ of $\varphi_j^\vee$ does not vanish. This value does not depend on the basis $\{W_m\}$ chosen, as long as this basis consists of monic polynomials, since $\varphi_j^\vee$ is the coefficient of a regular FOP1 in one of our mixed three-term recurrence formulas (5.25). Therefore, in the general case, we can replace (7.14) by a test for the nonvanishing of the tentative value of $\varphi_j^\vee$, which can be found by inversion of the function $\Omega_j$, cf. Table 1.

If the test fails, we end up with a $2 \times 2$ block pivot,

$$H_j^\wedge := \begin{bmatrix} A_i & B_{i+1} \\ C_i & A_{i+1} \end{bmatrix},$$

(7.15)

cf. (7.11), for which the unit Hessenberg matrix $H_j^\wedge - L_{j-1}^\vee F_j^\wedge$ is always nonsingular, since $\beta_{i+1} \neq 0$, while the $(n_i, n_{i+2} - 2)$-element of $L_{j-1}^\vee F_j^\wedge$ is always zero because $n_j^\vee < n_{i+2} - 2$.

The matrices $G^\vee$ and $G^\wedge$ can thus be built up by successive determination of the sizes of the block pivots and simultaneous computation of the entries by using the formulas of Table 1. The result is summarized in the following theorem and in Table 3.

**Theorem 7.2.** Given $H$, one can compute the block LU factorization $H = G^\vee G^\wedge$ by a Gauss block elimination process. In step $j$, where the upper left corner of the block pivot $H_j^\wedge - L_{j-1}^\vee F_j^\wedge$ is at $(n_i^\vee, n_j^\vee) = (n_i, n_i)$, the block $H_j^\wedge$ of $H$ is defined by identifying it with either the $1 \times 1$ block $A_i$ or the $2 \times 2$ block (7.15) (hence the size of the block pivot is either $h_i$ or $h_i + h_{i+1}$), depending on whether the tentative value
\( \hat{\varphi}_j^\gamma \) of \( \varphi_j^\gamma \) given in Table 3 is nonzero or zero. The relevant entries \( e_j^\gamma, \varphi_j^\gamma, e_j^\gamma, \) and \( \varphi_j^\gamma \) of \( G_j^\gamma \) and \( G^\gamma \) are then obtained according to Table 3. There the functions \( \Omega_j^\gamma \) and \( \omega_j^\gamma \) are together the inverse of \( \Omega_j \) defined in (7.3a); they are computed as follows: Partition the matrix \( T_{h_j}^{w_{j-1}} \) and the vector \( \alpha_i \) according to

\[
T_{h_j}^{w_{j-1}} = \begin{bmatrix} s^T & \tau \\ S & t \end{bmatrix}, \quad \alpha_i = \begin{bmatrix} \alpha \\ a \end{bmatrix};
\]

then set

\[
\Omega_j^\gamma(a_i) := S^{-1}(a + t),
\]

\[
\omega_j^\gamma(a_i) := \alpha + \tau - s^T \Omega_j^\gamma(a_i).
\]

The process starts at \( j = i = 0 \) with \( n_0^\gamma := n_0, \varphi_0^\gamma := 0, L_{-1}^\gamma F_0^\gamma := 0; \) (\( h_{-1}^\gamma \) does not matter.)

Remark. In the case where \( W_m(z) = z^m \), the definitions (7.17) reduce to

\[
\Omega_j^\gamma(a_i) := a = J_{h_j^\gamma-1} a_i,
\]

\[
\omega_j^\gamma(a_i) := \alpha = \alpha_{0,i}.
\]

### Table 3

**Formulas for the block LU factorization** \( H = G^\gamma G^\land \).

<table>
<thead>
<tr>
<th>( h_j^\gamma - 1 )</th>
<th>( h_j )</th>
<th>( h_j^\gamma )</th>
<th>( n_j^\gamma )</th>
<th>( n_j^\gamma + 1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( &gt; 1 )</td>
<td>( &gt; 1 )</td>
<td>( \omega_j^\gamma(a_i) = 0 )</td>
<td>( n_i )</td>
<td>( n_i+1 - 1 )</td>
</tr>
<tr>
<td>( = 1 )</td>
<td>( &gt; 1 )</td>
<td>( \omega_j^\gamma(a_i) - \varphi_j^\gamma = 0 )</td>
<td>( n_i )</td>
<td>( n_i+1 - 1 )</td>
</tr>
<tr>
<td>( &gt; 1 )</td>
<td>( = 1 )</td>
<td>( \alpha_{0,i} = 0 )</td>
<td>( n_i )</td>
<td>( n_i = n_i+1 - 1 )</td>
</tr>
<tr>
<td>( = 1 )</td>
<td>( = 1 )</td>
<td>( \alpha_{0,i} - \varphi_j^\gamma = 0 )</td>
<td>( n_i )</td>
<td>( n_i = n_i+1 - 1 )</td>
</tr>
<tr>
<td>( &gt; 1 )</td>
<td>( &gt; 1 )</td>
<td>( \omega_j^\gamma(a_i) \neq 0 )</td>
<td>( n_i )</td>
<td>( n_i+1 - 1 )</td>
</tr>
<tr>
<td>( = 1 )</td>
<td>( &gt; 1 )</td>
<td>( \omega_j^\gamma(a_i) - \varphi_j^\gamma \neq 0 )</td>
<td>( n_i )</td>
<td>( n_i+1 - 1 )</td>
</tr>
<tr>
<td>( = 1 )</td>
<td>( = 1 )</td>
<td>( \alpha_{0,i} \neq 0 )</td>
<td>( n_i )</td>
<td>( n_i = n_i+1 - 1 )</td>
</tr>
<tr>
<td>( &gt; 1 )</td>
<td>( &gt; 1 )</td>
<td>( \alpha_{0,i} - \varphi_j^\gamma \neq 0 )</td>
<td>( n_i )</td>
<td>( n_i = n_i+1 - 1 )</td>
</tr>
</tbody>
</table>

For the block UL decomposition \( H' = G^\land G^\gamma \) we must likewise find a diagonal block \( H_j^\gamma \) of \( H' \) of appropriate size, so that \( H_j^\gamma - F_j^{\land+1} L_j^\gamma \) is nonsingular and can be split into \( G_j^\gamma G_j^\gamma \):

\[
H_j^\gamma - F_j^{\land+1} L_j^\gamma = G_j^\gamma G_j^\gamma.
\]
In the case \( W_m(z) = z^m \) any tentative \( 1 \times 1 \) block pivot is again either a \( 1 \times 1 \) matrix or a companion matrix. The condition for it to be nonsingular is now that \( \alpha'_{0,i} \) minus the \((n'_i, n'_{i+1} - 1)\)-element of \( F^\wedge_{j+1} L_j \) does not vanish; i.e., in view of (7.8), \( n'_i = n'_j \) and \( n'_{i+1} = n'_{j+1} \) (cf. Table 2) if and only if

\[
\begin{align*}
(7.20a) & \quad \alpha'_{0,i} \neq 0 \quad \text{in case } h_{j+1}^\wedge > 1, \\
(7.20b) & \quad \alpha'_{0,i} - \varphi_{j+1}^\wedge \neq 0 \quad \text{in case } h_{j+1}^\wedge = 1.
\end{align*}
\]

However, \( \varphi_{j+1}^\wedge \) is not known at this moment. But we can apply a different argument.

Let \( n'_j := n'_i \) and assume that the elements of \( G^\wedge \) and \( G^\wedge \) have already been determined up to the \((n'_j - 1)\)th column. Then \( h_j^\wedge, e_j^\wedge, \) and \( \varphi_j^\wedge \) are known, and \( n'_j \) is equal to the dimension \( h'_j \) of \( a'_i \). From Table 2 we see that \( \varphi_j^\wedge \) is determined by \( \beta'_j \) and \( \varphi_j^\wedge \). Moreover, by analogy to (7.17) and (7.18) there are functions \( \Omega_j^\wedge \) and \( \omega_j^\wedge \) inverting \( \Omega_j^\wedge \); the first one yields \( e_j^\wedge \) if \( h_j^\wedge > 1 \).

Next, let us first assume that \( h_j^\wedge > 1 \) and consider

\[
\varphi_{j+1}^\wedge := \begin{cases} 
\omega_j^\wedge(a'_i) & \text{in case } h_j^\wedge > 1, \\
\alpha'_{0,i} & \text{in case } h_j^\wedge = 1
\end{cases}
\]

as a tentative value for \( \varphi_{j+1}^\wedge \). If nonvanishing, we let it be the true value of \( \varphi_{j+1}^\wedge \) and set \( h_{j+1}^\wedge := 1 \) (\( e_{j+1}^\wedge \) is then void). Otherwise, \( h_{j+1}^\wedge > 1 \), hence \( n'_{j+1} := n'_{i+2} \), \( e_{j+1}^\wedge := a_{i+1}^\wedge, \varphi_{j+1}^\wedge := \beta_{i+1}^\wedge \).

If \( h_j^\wedge = 1 \), we instead let

\[
\varphi_{j+1}^\wedge := \begin{cases} 
\omega_j^\wedge(a'_i) - \varphi_j^\wedge & \text{in case } h_j^\wedge > 1, \\
\alpha'_{0,i} - \varphi_j^\wedge & \text{in case } h_j^\wedge = 1
\end{cases}
\]

be the tentative value for \( \varphi_{j+1}^\wedge \). The rest of the step is the same as before.

To complete the definition of the procedure we have to describe its start. At this point we must note that although the functional \( \Phi' \) is uniquely determined by \( \Phi \), the converse is not true, since \( \Phi \) depends on \( \Phi(1) = \phi_l \), while \( \Phi' \) is independent of \( \phi_l \). Hence, the set \( \{P_n\} \) of POP1s determined by \( \Phi \) cannot be uniquely determined by the set \( \{P'_n\} \) corresponding to \( \Phi' \). Moreover, the value of \( \phi_l \) determines whether the \((l - 1, 0)\) and the \((l, 0)\) Padé approximants of \( f(z) = \sum \phi_k z^k \) belong to the same block of the table or not. In fact, they do if and only if \( \phi_l = 0 \). (Recall that these Padé approximants are polynomial interpolants.)

Therefore, given the recurrence formulas for \( \{P'_n\} \) (i.e., given the matrix \( H' \)), those for \( \{P_n\} \) and those for the mixed recurrence (i.e., the matrices \( H, G^\wedge, \) and \( G^\wedge \)) are only determined after \( \phi_l \) has been specified.

According to (5.17a), \( \varphi_0^\wedge \) satisfies

\[
(7.23) \quad \varphi_0^\wedge \Phi(z n_0^\wedge) = \Phi(z n_0^\wedge + 1 W_{n_0^\wedge - 1} P_{n_0^\wedge}').
\]

In the case \( h_0^\wedge = 1 \), where \( n_0^\wedge := n_0^\wedge := n_0^\wedge = 0 \) (i.e., \( i = j = 0 \), \( h_0^\wedge := h_0^\wedge, n_1^\wedge := n_1^\wedge \) and where both \( e_0^\wedge \) and \( \varphi_0^\wedge \) are void, we obtain

\[
(7.24) \quad \varphi_0^\wedge := \Phi(z W_{n_1^\wedge - 1})/\phi_l,
\]

which for the monomial basis \( W_m(z) = z^m \) reduces to

\[
(7.25) \quad \varphi_0^\wedge := \phi_{l+n'_l}/\phi_l.
\]
In the case $h > 1$, where $n^\wedge_0 := n_0 = 0, n^\wedge_1 := n_1$ (i.e., $i = 1, j = 0$), $h^\wedge_0 := h_0 + 1, h^\wedge_1 := h_1$, we have $e^\wedge_0 := a_0$ (cf. (7.9a) and (7.9c) with (3.7)) and, from (7.23),

\begin{equation}
\varphi^\wedge_0 := \Phi(z^{n_1+1}W_{h'_1-1}P_{n'_1})/\phi_{l+n'_1},
\end{equation}

which for $W_m(z) = z^m$ reduces to

\begin{equation}
\varphi^\wedge_0 := \phi_{l+n'_2}/\phi_{l+n'_1}.
\end{equation}

In both cases $e^\wedge_0, e^\wedge_1$, and $\varphi^\wedge_1$ are obtained according to the general formulas in Table 4. Altogether we get the following analog of Theorem 7.2.

**Theorem 7.3.** Given $H'$ and either $\phi_1 \neq 0$ and $\phi_{l+n'_i}$ or $\phi_1 = 0, \phi_{l+n'_i}$, and $\phi_{l+n'_i}$, we can compute the block UL factorization $H' = G^\wedge G^\wedge$ by the following process. In step $j$, where we compute columns $n^\wedge_j = n'_i$ through $n^\wedge_{j+1} - 1$ of $G^\wedge$ and $G^\wedge$, the diagonal block $H^\wedge_j$ of $H'$ (containing rows and columns $n^\wedge_j$ through $n^\wedge_{j+1} - 1$) is defined as either the $1 \times 1$ block $H^\wedge_j := A^\wedge_j$ or the $2 \times 2$ block

\begin{equation}
H^\wedge_j := \begin{bmatrix} A^\wedge_{i-1} & B^\wedge_i \\ C^\wedge_{i-1} & A^\wedge_i \end{bmatrix},
\end{equation}

depending on whether the tentative value $\varphi^\wedge_{j+1}$ of $\varphi^\wedge_{j+1}$ given in Table 4 is nonzero or zero. The relevant entries $e^\wedge_{j+1}, e^\wedge_j, e^\wedge_{j+1}$, and $\varphi^\wedge_{j+1}$ of $G^\wedge$ and $G^\wedge$ are then obtained according to Table 4. There the functions $\Omega^\wedge_j$ and $\omega^\wedge_j$ are together the inverse of $\Omega^\wedge_j$ defined in (7.3b); they are computed as follows: Partition the matrix $T^W_{h^\wedge_{j-1}}$ and the vector $a^\wedge_i$ according to

\begin{equation}
T^W_{h^\wedge_{j-1}} := \begin{bmatrix} s^T & \tau \\ S & t \end{bmatrix}, \quad a^\wedge_i := \begin{bmatrix} \alpha \\ \mathbf{a} \end{bmatrix};
\end{equation}

then

\begin{equation}
\Omega^\wedge_j(a^\wedge_i) := S^{-1}(a + t),
\end{equation}

\begin{equation}
\omega^\wedge_j(a^\wedge_i) := \alpha + \tau - s^T \Omega^\wedge_j(a^\wedge_i).
\end{equation}

The first step depends on the value $\phi_1 = \Phi(1)$ (on which $H'$ does not depend, while $H$ does):

If $\phi_1 \neq 0$, then

\begin{align}
(7.31a) \quad & h^\wedge_0 := 1, \quad h^\wedge_0 := h_0, \\
(7.31b) \quad & n^\wedge_0 := n_0 = 0, \quad n^\wedge_1 := n_1,
\end{align}

e^\wedge_0 is void, and $\varphi^\wedge_0$ is given by (7.24); $e^\wedge_0, e^\wedge_1, \varphi^\wedge_0$ are then obtained from the general formulas in Table 4, with $i = j = 0$.

If $\phi_1 = 0$, then

\begin{align}
(7.32a) \quad & h^\wedge_0 := h_0 + 1, \quad h^\wedge_0 := h_1, \\
(7.32b) \quad & n^\wedge_0 := n_0 = 0, \quad n^\wedge_1 := n_1, \quad n^\wedge_2 := n_2, \\
(7.32c) \quad & e^\wedge_0 := a_0,
\end{align}

and $\varphi^\wedge_0$ is given by (7.26); $e^\wedge_0, e^\wedge_1, \varphi^\wedge_0$ are obtained from Table 4 by setting $i = 1, j = 0$. 

---

**40 MARTIN H. GUTKNECHT**
Starting from the recurrence coefficients for some sequence \( \{P_n\}_{n=0}^\infty \) of FOP1s, say from \( \{P_{n0}\}_{n=0}^\infty \), inductive application of Theorem 7.2 and of the \( H' G G^\vee \) part of Theorem 7.1 allows us to compute the recurrence coefficients of any sequence \( \{P_{nl}\}_{n=0}^\infty \), \( l = 1, 2, \ldots \). These two theorems therefore define the \textit{progressive} qd algorithm, even for nongeneric situations. In the generic case, for which the algorithm is due to Rutishauser [23], only the formulas for \( h_j^\wedge = h_j = h_i = h_i' = 1 \) (\( \forall i, j \)) are used, i.e., only those in the last rows of Tables 2 and 3. In this generic case the qd table contains the recurrence coefficients for every diagonal sequence of FOP1s, i.e., our coefficients \( \alpha_{0,i} (i = 0, 1, \ldots) \) and \( \beta_i (i = 1, 2, \ldots) \) for every diagonal \( (l = 0, 1, \ldots) \). The progressive qd algorithm allows us to build up the qd table from its main diagonal, where \( l = 0 \). (More generally, one can proceed downwards from any diagonal or row.) Rutishauser had some heuristic rules for dealing with nongeneric situations, namely, rules for filling the then appearing gaps in the qd table with zeros and \( \infty \) symbols. Draux [6] also formulated and established such rules. However, according to the above result, we can define a qd table that is valid in every nongeneric situation and contains as entries on its \( l \)-th diagonal the nontrivial entries \( a_i \) (\( i = 0, 1, \ldots \)) and \( \beta_i \) (\( i = 1, 2, \ldots \)) of the matrix \( H \) for this \( l \).

Likewise, starting from the recurrence coefficients of \( \{P_{n0}\}_{n=0}^\infty \) inductive application of Theorem 7.3 and of the \( H = G^\vee G^\wedge \) part of Theorem 7.1 allows us to compute those of any sequence \( \{P_{ni}\}_{n=0}^\infty \), \( l = -1, -2, \ldots \). In each step a new “moment” \( \phi_i \), \( l = -1, -2, \ldots \), has to be provided. Hence, we can proceed from the main (or any other) diagonal upwards and to the right. For the generic case this process is well known. We call this the \textit{backward} qd algorithm.

The progressive qd algorithm enables us in particular to compute the moments

---

**Table 4**

Formulas for the block UL factorization \( H' = G^\wedge G^\vee \).

<table>
<thead>
<tr>
<th>( h_j^\wedge )</th>
<th>( h_i' )</th>
<th>( \phi_j^\wedge+1 )</th>
<th>( h_j^\wedge+1 )</th>
<th>( n_j^\vee )</th>
<th>( n_i^\wedge+1 )</th>
<th>( n_j^\wedge+1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt; 1</td>
<td>&gt; 1</td>
<td>( \omega_j^\vee (a_i') = 0 )</td>
<td>( n_j^\vee )</td>
<td>( n_i^\wedge+1 )</td>
<td>( n_j^\wedge+1 )</td>
<td></td>
</tr>
<tr>
<td>= 1</td>
<td>&gt; 1</td>
<td>( \omega_j^\vee (a_i') - \phi_j^\vee = 0 )</td>
<td>( n_i^\wedge )</td>
<td>( n_i^\wedge+1 )</td>
<td>( n_i^\wedge+1 )</td>
<td></td>
</tr>
<tr>
<td>&gt; 1</td>
<td>= 1</td>
<td>( \alpha_{0,i} = 0 )</td>
<td>( n_i^\wedge )</td>
<td>( n_i^\wedge+1 = n_i^\wedge+1 )</td>
<td>( n_i^\wedge+1 )</td>
<td></td>
</tr>
<tr>
<td>= 1</td>
<td>= 1</td>
<td>( \alpha_{0,i} - \phi_j^\vee = 0 )</td>
<td>( n_i^\wedge )</td>
<td>( n_i^\wedge+1 = n_i^\wedge+1 )</td>
<td>( n_i^\wedge+1 )</td>
<td></td>
</tr>
<tr>
<td>&gt; 1</td>
<td>&gt; 1</td>
<td>( \omega_j^\vee (a_i') \neq 0 )</td>
<td>( n_i^\wedge )</td>
<td>( n_i^\wedge+1 )</td>
<td>( n_i^\wedge+1 )</td>
<td></td>
</tr>
<tr>
<td>= 1</td>
<td>&gt; 1</td>
<td>( \omega_j^\vee (a_i') - \phi_j^\vee \neq 0 )</td>
<td>( n_i^\wedge )</td>
<td>( n_i^\wedge+1 = n_i^\wedge+1 )</td>
<td>( n_i^\wedge+1 )</td>
<td></td>
</tr>
<tr>
<td>&gt; 1</td>
<td>= 1</td>
<td>( \alpha_{0,i} \neq 0 )</td>
<td>( n_i^\wedge )</td>
<td>( n_i^\wedge+1 = n_i^\wedge+1 )</td>
<td>( n_i^\wedge+1 )</td>
<td></td>
</tr>
<tr>
<td>= 1</td>
<td>= 1</td>
<td>( \alpha_{0,i} - \phi_j^\vee \neq 0 )</td>
<td>( n_i^\wedge )</td>
<td>( n_i^\wedge+1 = n_i^\wedge+1 )</td>
<td>( n_i^\wedge+1 )</td>
<td></td>
</tr>
</tbody>
</table>
\( \phi_l \) \((l = 0, 1, \ldots)\) from the coefficients \( a_l \) \((i = 0, 1, \ldots)\) and \( \beta_l \) \((i = 1, 2, \ldots)\) of the main diagonal. Conversely, given the moments, the ordinary qd algorithm yields in the generic case the recurrence coefficients on the main diagonal. This process is known to be highly unstable. The same task can be done with the Chebyshev algorithm, which is less likely to break down or to be unstable, since it requires only that all FOP1s \( P_{0,n} \) on the main diagonal are regular. However, often the problem itself is ill conditioned, and there is no chance for numerically stable computations. It has, therefore, been proposed to replace the moments by modified moments if possible. Golub and Gutknecht [9] have extended the corresponding modified Chebyshev algorithm to the nongeneric case. The nongeneric Chebyshev algorithm is included there as a special case.

The progressive and the backward qd algorithms are well known to have interesting convergence properties. Basically, by proceeding downwards in the qd table we obtain the poles of \( f \), and by moving to the right we find its zeros, see, e.g., [3].

8. The nongeneric biconjugate gradient algorithm (BCG or BIOMIN) and nongeneric BIODIR. The biconjugate gradient (BCG) algorithm is closely related to the Lanczos biorthogonalization (BO) method. It can be traced back to Lanczos [20], where it was introduced as "the complete algorithm for minimized iterations." More than 20 years later, Fletcher [7] revived and popularized it. It generates the same biorthogonal vector sequences \( \{x_n\} \), \( \{y_n\} \) characterized by (4.6)–(4.7), and the same iterates \( \{z_n\} \) satisfying (4.61) as the normalized BIORES algorithm [11], but additionally it generates two biconjugate vector sequences \( \{u_n\} \) and \( \{v_n\} \) taken from the same nested sequences of Krylov spaces as \( \{x_n\} \) and \( \{y_n\} \):

\[
\begin{align*}
(8.1a) \quad u_n & \in K_{n+1} := \text{span} \left( x_0, A x_0, A^2 x_0, \ldots, A^n x_0 \right), \\
(8.1b) \quad v_n & \in L_{n+1} := \text{span} \left( y_0, A^H y_0, (A^H)^2 y_0, \ldots, (A^H)^n y_0 \right)
\end{align*}
\]

with

\[
\langle v_m, A u_n \rangle_B \begin{cases} = 0 & \text{if } m \neq n, \\ \neq 0 & \text{if } m = n. \end{cases}
\]

This process can break down for various reasons, cf. [11]. From \S 4 we know already that \( \{x_n\} \) and \( \{y_n\} \) satisfying (4.6) and (4.7) may not exist and that a suitable modification of the process can be based on the theory of formal orthogonal polynomials. An argument analogous to the one given in \S 4 shows that if (8.1) and (8.2) can be fulfilled for \( n = 0, 1, \ldots, \nu - 1 \), then they are fulfilled by

\[
(8.3) \quad u_n = P_n'(A) x_0 \Gamma'_n, \quad v_n = \overline{P_n'}(A^H) y_0 \bar{\Gamma}'_n,
\]

where the scale factors \( \Gamma'_n \) and \( \bar{\Gamma}'_n \) are not necessarily the same as the factors \( \Gamma_n \) and \( \bar{\Gamma}_n \) in (4.14), and where \( P_n' \) is the \( n \)th monic FOP1 with respect to the linear functional \( \Phi' = \Phi \) defined by

\[
(8.4) \quad \Phi'(z^k) := \phi_{k+1} := \langle y_0, A^{k+1} x_0 \rangle_B.
\]

(Recall that if (8.2) holds for all \( m < \nu \) and \( n < \nu \), all the polynomials \( P_n' \) \((n < \nu)\) in (8.3) are regular FOP1s.)

In case of a breakdown, the formulas (8.3) point again to the correct generalization of the process: \( u_n \) and \( v_n \) must still have the same form, with \( P_n' \) being an \( n \)th FOP1 for \( \Phi' \), even if it is not a regular FOP1; hence (8.2) does not hold for this \( n \). The
recurrences of §5 allow us to find a recursive algorithm for computing $u_n$ and $v_n$ along with $x_n$ and $y_n$, and the relations of §§6 and 7 yield corresponding matrix results.

First we formulate the matrix algorithm that is based on a reinterpretation of Theorems 5.3 and 5.4. Since it yields both a pair of biorthogonal and a pair of biconjugate sequences, we call it, as in [11], the BOBC algorithm.

As an extension of (4.20a), for $i, n \in \mathbb{N}$ we let

\begin{align}
(8.5a) & \quad \gamma_{n,i} := \Gamma_n / \Gamma_{n-i}, \quad \bar{\gamma}_{n,i} := \bar{\Gamma}_n / \bar{\Gamma}_{n-i}, \\
(8.5b) & \quad \gamma'_{n,i} := \Gamma'_n / \Gamma'_{n-i}, \quad \bar{\gamma'}_{n,i} := \bar{\Gamma}'_n / \bar{\Gamma}'_{n-i},
\end{align}

and

\begin{align}
(8.6a) & \quad \gamma_{n,i}^\wedge := \Gamma_n / \Gamma_{n-i}, \quad \bar{\gamma}_{n,i}^\wedge := \bar{\Gamma}_n / \bar{\Gamma}_{n-i}, \\
(8.6b) & \quad \gamma'_{n,i}^\wedge := \Gamma'_n / \Gamma'_{n-i}, \quad \bar{\gamma'}_{n,i}^\wedge := \bar{\Gamma}'_n / \bar{\Gamma}'_{n-i}.
\end{align}

Recall that the mixed three-term recurrence relations (5.25) allow us to generate the two sequences $\{P_n^\gamma\}_{j=0}^\infty$ and $\{P_n'^\gamma\}_{j=0}^\infty$ consisting of regular FOP1s for $\Phi$ and $\Phi'$, respectively, and that two full sequences of FOP1s for these two functionals are then defined by (5.27). If $h_j := n_j - n_{j-1} + 1 > 1$ or $h'_j := n'_j + 1 - n'_j > 1$, some polynomials on the two diagonals coincide or differ only by a factor of $z$, cf. (5.28); consequently,

\begin{align}
(8.7a) & \quad x_n = u_n \gamma_{n,0}^\wedge, \quad y_n = v_n \gamma_{n,0}^\wedge, \text{ if } n_j \leq n < n_{j+1}, \\
(8.7b) & \quad x_n = A u_{n-1} \gamma_{n,1}^\wedge, \quad y_n = A^H v_{n-1} \bar{\gamma}_{n,1}^\wedge, \text{ if } n_{j-1} < n < n_{j+1}.
\end{align}

The polynomials $W_m$ in (5.27) are for practicality again assumed to satisfy the three-term recurrence (2.10), so that the recurrences (5.29) hold, which translate into

\begin{align}
(8.8a) & \quad x_{n+1} = [A x_n - x_n a_n^{-1} \gamma_{n+1,1} - x_n - 1 \beta_n \gamma_{n+1,2}, \\
(8.8b) & \quad x_{n+1} = [A x_n - x_n a_n^{-1} \gamma_{n+1,1} - x_n - 1 \beta_n \gamma_{n+1,2}, \\
(8.8c) & \quad u_{n+1} = [A u_n - u_n a_n^{-1} \gamma_{n+1,1} - u_n - 1 \beta_n \gamma_{n+1,2}, \\
(8.8d) & \quad u_{n+1} = [A u_n - u_n a_n^{-1} \gamma_{n+1,1} - u_n - 1 \beta_n \gamma_{n+1,2},
\end{align}

Of course, analogous formulas with the complex conjugate coefficients $\bar{\alpha}_m$ and $\bar{\beta}_m$, with the scale factors $\bar{\gamma}_{n,i}$ and $\bar{\gamma'}_{n,i}$ of (8.6) and with $A$ replaced by $A^H$ hold for $\{y_n\}$ and $\{v_n\}$, but from now on we only give those for $\{x_n\}$ and $\{u_n\}$. For simplicity we refer to these analogous formulas as the conjugate recurrences, although $\bar{\gamma}_{n,i}$ and $\bar{\gamma'}_{n,i}$ need not be complex conjugate to $\gamma_{n,i}$ and $\gamma_{n,i}$.

If $\varepsilon_{i,j}^\vee$ and $\varepsilon_{i,j}^\wedge$ denote the coefficients of the polynomials $e_j^\vee$ and $e_j^\wedge$, respectively, (as is the case in (6.5)), the mixed three-term recurrence formulas (5.25) yield

\begin{align}
(8.9a) & \quad u_{n_j}^\vee = x_n^\vee \gamma_{n_j,0}^\vee - x_n^\vee - 1 e_{n_j-2,j}^\wedge \gamma_{n_j,1}^\wedge - x_n^\vee - 2 e_{n_j-3,j}^\wedge \gamma_{n_j,2}^\wedge - \cdots \\
& \quad - x_n^\vee \varepsilon_{0,j}^\wedge \gamma_{n_j,h_j-1}^\wedge - u_{n_{j-1}} \phi_{j}^\vee \gamma_{n_j,h'_j}^\vee + h_{j-1}^\vee - 1,
\end{align}
when \( h_j^+ > 1 \), cf. (5.27a) and (5.31a). In view of (5.27b), (5.29b), and (5.28b), when \( h_j^+ > 1 \),

\[
(8.9b) \quad x_{n_{j+1}}^+ = A[u_{n_{j+1}}^+ - 1\gamma_{n_{j+1}}^+ - u_{n_{j+1}}^+ - 2\varepsilon_{h_j^+ - 2,j}^+ y_{n_{j+1}}^+ - 2 - u_{n_{j+1}}^+ - 3\varepsilon_{h_j^+ - 3,j}^+ y_{n_{j+1}}^+ - 3 - \cdots - u_{n_{j+1}}^+ - \varepsilon_{0,j}^+ y_{n_{j+1}}^+ + h_j^+ - h_{j-1}^+]
\]

\[
(8.9c) \quad A \text{ Un}_j^+ - 1\gamma_{n_{j+1}}^+ - x_{n_{j+1}}^+ - 1\varepsilon_{h_j^+ - 2,j}^+ y_{n_{j+1}}^+ - 1 - x_{n_{j+1}}^+ - 2\varepsilon_{h_j^+ - 3,j}^+ y_{n_{j+1}}^+ - 2 - \cdots - x_{n_{j+1}}^+ - \varepsilon_{0,j}^+ y_{n_{j+1}}^+ + h_j^+ - h_{j-1}^+ - 1
\]

\[
(8.9d) \quad = [A x_{n_{j+1}}^+ - 1 - x_{n_{j+1}}^+ - 1(\varepsilon_{h_j^+ - 2,j}^+ + \alpha_{h_j^+ - 3}^+) y_{n_{j+1}}^+ - 1 - x_{n_{j+1}}^+ - \varepsilon_{0,j}^+ y_{n_{j+1}}^+ + h_j^+ - h_{j-1}^+ - 1
\]

(cf. (5.31b)). However, when \( h_j^+ = 1 \), i.e., \( n_j^+ = n_j^+ \), then \( W_{h_j^+ - 1}(z) \equiv 1 \) and \( e_j^+(z) \equiv 0 \), so that instead of (8.9a) we simply obtain

\[
(8.9e) \quad u_{n_j^+}^+ = x_{n_j^+}^+ \gamma_{n_j^+}^+ - 0 - u_{n_j^+ - 1}^+ \varphi_j^+ y_{n_j}^+ + h_{j-1}^+.
\]

Likewise, if \( h_j^+ = 1 \), i.e., \( n_j^+ = n_j^+ + 1 \), (8.8b)–(8.8d) are replaced by

\[
(8.9f) \quad x_{n_{j+1}}^+ = A u_{n_j^+}^+ \gamma_{n_{j+1}}^+ - x_{n_j^+}^+ \varphi_j^+ y_{n_{j+1}}^+ + h_j^+.
\]

It remains to give formulas for the index sequences \( \{n_j^+\} \) and \( \{n_j^+\} \) and for the coefficients \( \varepsilon_{n,j}^+, \varepsilon_{n,j}^+ \), \( \varphi_j^+ \), and \( \varphi_j^+ \) that appear in (8.9). First, according to (5.26) and (5.27), \( h_j^+ := n_j^+ - n_j^+ + 1 \) and \( h_j^+ := n_{j+1}^+ - n_j^+ \) are given by

\[
(8.10a) \quad h_j^+ := \min \{k \in \mathbb{N}^+; \langle y_{n_j^+}, x_{n_j^+ - 1} \rangle_B \neq 0\},
\]

\[
(8.10b) \quad h_j^+ := \min \{k \in \mathbb{N}^+; \langle y_{n_j^+}, A u_{n_j^+ - 1} \rangle_B \neq 0\}.
\]

Second, equations for the mentioned coefficients follow from Theorem 5.4; we choose relations (5.32a)–(5.32c) and (5.32f) in order to work without \( v_{n_j^+}, \ldots, v_{n_j^+ - 1} \):

\[
(8.11a) \quad \varphi_{j}^+ y_{n_j^+} - h_j^1 \gamma_{n_j^+}^+ - h_j^1 h_{j-1}^1 \gamma_{y_{n_j^+}}^+, A u_{n_j^+ - 1} = \langle y_{n_j^+}, x_{n_j^+} \rangle_B
\]

\[
(8.11b) \quad \sum_{s=1}^{k} \varepsilon_{s,j}^+ y_{n_j^+ - 1, s}^+ x_{n_j^+ - s, k} = \langle y_{n_j^+ \rightarrow k}, x_{n_j^+ \rightarrow s} \rangle_B
\]

\[
(8.11c) \quad \varphi_{j}^+ y_{n_j^+} - h_j^1 \gamma_{n_j^+}^+ - h_j^1 h_{j-1}^1 \gamma_{y_{n_j^+}}^+, A u_{n_j^+ - 1} = \langle y_{n_{j+1}^+ - 1}, A u_{n_j^+} \rangle_B
\]

\[
(8.11d) \quad \sum_{s=1}^{k} \varepsilon_{s,j}^+ y_{n_j^+ - 1, s}^+ x_{n_j^+ - s, 1} = \langle A^H y_{n_j^+ - 1} - y_{n_j^+ - 1, 1} \alpha_{h_j^+ - 2} \gamma_{w_{n_j^+ - 1}, 1}, u_{n_j^+ + k} \rangle_B
\]

Again, (8.11a) and (8.11c) are single linear equations, hence explicit formulas, for \( \varphi_{j}^+ \) and \( \varphi_{j}^+ \), respectively. Equations (8.11b) and (8.11d) are triangular systems for the
coefficients $\varepsilon_{s,j}^x$ and $\varepsilon_{s,j}^y$ in the representation (6.5) of the polynomials $e_j^x$ and $e_j^y$. If $h_j^x = 1$, the system (8.11b) is void, and, if $h_j^x = 1$, the system (8.11d) is void.

We now obtain the following algorithm.

**Algorithm 4 (Nongeneric BOBC Algorithm).** Given a bounded linear operator $A : \mathcal{H} \to \mathcal{H}$ and two initial vectors $x_0, y_0 \in \mathcal{H}$ satisfying $(y_0, x_0)_B \neq 0$, set $u_0 := x_0$, $v_0 := y_0$, $h_0^x := 1$. Then construct sequences $\{x_n\}_{n=0}^{\infty}$, $\{y_n\}_{n=0}^{\infty}$, $\{u_n\}_{n=0}^{\infty}$, and $\{v_n\}_{n=0}^{\infty}$ according to the inductive process, which, for $j = 0, 1, \ldots$, consists of:

(i) If $h_j^x > 1$, then $\{x_n_{j+k}\}_{k=1}^{h_j^x-1}$, $\{y_n_{j+k}\}_{k=1}^{h_j^x-1}$, and $h_j^x$ are defined by executing concurrently (8.8a), the corresponding conjugate recurrence for $y_n_{j+k}$, and (8.10a); if $h_j^x = \infty$, then $J^x := j$, $J^y := j - 1$; in particular, if $x_{n_j^x} = 0$ or $y_{n_j^x} = 0$, then $h_j^x = \infty$ and $x_{n_j^x+k} = 0 \, (\forall k \geq 0)$ or $y_{n_j^x+k} = 0 \, (\forall k \geq 0)$, respectively, and the algorithm terminates (in practice, $x_{n_j^x+k}$ and $y_{n_j^x+k}$ are then not needed);

(ii) once $h_j^x$ has been determined, the nonzero constant $\varphi_j^x$ is given by (8.11a) and, if $h_j^x > 1$, the coefficients $\{\varepsilon_{s,j}^x\}_{s=1}^{h_j^x-2}$ are obtained by solving the triangular linear system (8.11b);

(iii) depending on whether or not $h_j^x > 1$, $u_{n_j}^x$ and $v_{n_j}^x$ are then given by (8.9a) or (8.9c) and the conjugate recurrence; if $(v_{n_j}^x, Au_{n_j}^x)_B \neq 0$, set $h_j^x := 1$; otherwise $h_j^x > 1$;

(iv) if $h_j^y > 1$, then $\{u_{n_j+k}\}_{k=1}^{h_j^y-1}$, $\{y_{n_j+k}\}_{k=1}^{h_j^y-1}$, and $h_j^y$ are defined by (8.8d), (8.7b), the recurrence conjugate to (8.8b), and by (8.10b); if $h_j^y = \infty$, then $J^y := J^x := j$; in particular, if $u_{n_j}^y = 0$ or $v_{n_j}^y = 0$, then $h_j^y = \infty$ and $x_{n_j+k} = 0 \, (\forall k \geq 1)$ or $y_{n_j+k} = 0 \, (\forall k \geq 1)$, respectively, and the algorithm terminates (in practice, $x_{n_j+k}$ and $y_{n_j+k}$ are then not needed);

(v) once $h_j^y$ has been determined, the nonzero constant $\varphi_j^y$ is given by (8.11c) and, if $h_j^y > 1$, the coefficients $\{\varepsilon_{s,j}^y\}_{s=1}^{h_j^y-2}$ are obtained by solving the triangular linear system (8.11d);

(vi) depending on whether $h_j^y > 1$ or not, $x_{n_j+1}$ and $y_{n_j+1}$ are either given by (8.9b) and the conjugate recurrence to (8.9d) or by (8.9f) and its conjugate recurrence; if $(v_{n_j+1}^x, u_{n_j+1}^x)_B \neq 0$, set $h_j^{x+1} := 1$, otherwise $h_j^{x+1} > 1$.

The recurrence coefficients $\alpha^x_m$ and $\beta^x_m$ in (8.8) and the nonvanishing scale factors $\Gamma_n$, $\Gamma_n^x$, $\Gamma_n^y$, and $\Gamma_n^w$ ($n \in \mathbb{N}$), which determine $\gamma_{n,i}^x$, $\gamma_{n,i}^y$, $\gamma_{n,i}^w$, $\gamma_{n,i}^\nu$, $\gamma_{n,i}^\nu$, $\gamma_{n,i}^\nu$, and $\gamma_{n,i}^\nu$ according to (8.5) and (8.6), can be chosen freely. (For the sake of simplicity, we assume that $\Gamma_0 := \Gamma_0^x := \Gamma_0^y := \Gamma_0^w := 1$.)

As in Algorithm 1 (§4) we could set $\Gamma_n := \Gamma_n^x := \Gamma_n^y := \Gamma_n^w := 1 \, (n \in \mathbb{N})$, which would imply that $\gamma_{n,i}^x = \gamma_{n,i}^y = \gamma_{n,i}^w = \gamma_{n,i}^\nu = \gamma_{n,i}^\nu = \gamma_{n,i}^\nu = \gamma_{n,i}^\nu = 1 \, (\forall n, \forall i)$, but might lead to overflow or underflow.

Of course, Algorithm 4 also has a matrix interpretation, which is analogous to the one for Algorithm 1 that was formulated in Theorem 4.2 of Part I.

**Theorem 8.1.** Gather the vectors generated by Algorithm 4 into

$$
(8.12a) \quad X := [x_0, x_1, x_2, \ldots], \quad Y := [y_0, y_1, y_2, \ldots],
$$

$$
(8.12b) \quad U := [u_0, u_1, u_2, \ldots], \quad V := [v_0, v_1, v_2, \ldots],
$$

and the scale factors used into

$$
(8.13a) \quad \Gamma := \text{diag} \, [\Gamma_0, \Gamma_1, \Gamma_2, \ldots], \quad \tilde{\Gamma} := \text{diag} \, [\tilde{\Gamma}_0, \tilde{\Gamma}_1, \tilde{\Gamma}_2, \ldots],
$$
Let $D^\wedge$ and $D^\vee$ be the block diagonal matrices

\[(8.14)\quad D^\wedge := \Phi((p)^T p), \quad D^\vee := \Phi((p')^T p')\]

expressing the formal orthogonality of the two sequences of FOP1s, and let

\[(8.15)\quad D^\wedge_r := \tilde{\Gamma}D^\wedge \tilde{\Gamma}, \quad D^\vee_r := \tilde{\Gamma}D^\vee \tilde{\Gamma}\]

be corresponding diagonally scaled matrices. Furthermore, using the block matrices $E^\wedge, E^\vee, F^\wedge, F^\vee, G^\wedge,$ and $G^\vee$ from §6, introduce the scaled matrices

\begin{align*}
(8.16a) & \quad E^\wedge_r := (\tilde{\Gamma}')^{-1}E^\wedge \tilde{\Gamma}, \quad \overline{E^\wedge_r} := (\tilde{\Gamma}')^{-1}\overline{E^\wedge} \tilde{\Gamma}, \\
(8.16b) & \quad E^\vee_r := (\tilde{\Gamma})^{-1}E^\vee \tilde{\Gamma}, \quad \overline{E^\vee_r} := (\tilde{\Gamma})^{-1}\overline{E^\vee} \tilde{\Gamma}, \\
(8.16c) & \quad F^\wedge_r := (\tilde{\Gamma})^{-1}F^\wedge \tilde{\Gamma}, \quad \overline{F^\wedge_r} := (\tilde{\Gamma})^{-1}\overline{F^\wedge} \tilde{\Gamma}, \\
(8.16d) & \quad F^\vee_r := (\tilde{\Gamma}')^{-1}F^\vee \tilde{\Gamma}, \quad \overline{F^\vee_r} := (\tilde{\Gamma}')^{-1}\overline{F^\vee} \tilde{\Gamma}, \\
(8.16e) & \quad G^\wedge_r := (\tilde{\Gamma})^{-1}G^\wedge \tilde{\Gamma}, \quad \overline{G^\wedge_r} := (\tilde{\Gamma})^{-1}\overline{G^\wedge} \tilde{\Gamma}, \\
(8.16f) & \quad G^\vee_r := (\tilde{\Gamma}')^{-1}G^\vee \tilde{\Gamma}, \quad \overline{G^\vee_r} := (\tilde{\Gamma}')^{-1}\overline{G^\vee} \tilde{\Gamma}.
\end{align*}

Then Algorithm 4 induces the relations

\begin{align*}
(8.17a) & \quad AUE^\wedge_r = X F^\wedge_r, \quad A^H V E^\wedge_r = Y F^\wedge_r, \\
(8.17b) & \quad X E^\vee_r = U F^\vee_r, \quad Y E^\vee_r = V F^\vee_r,
\end{align*}

which imply

\begin{align*}
(8.18a) & \quad AU = X G^\wedge_r, \quad A^H V = Y G^\wedge_r, \\
(8.18b) & \quad X = U G^\vee_r, \quad Y = V G^\vee_r.
\end{align*}

Moreover, if we write the infinite matrix with $(m, n)$-element $(y_m, x_n)_B$ formally as $Y^H X$, and the one with $(m, n)$-element $(v_m, A u_n)_B$ as $V^H A U$, then we have

\[(8.19)\quad Y^H X = D^\wedge_r, \quad V^H A U = D^\vee_r.\]

Note that $(8.9b)$ corresponds to $(8.17a)$, while $(8.9c)$ translates into the relation $(8.18a)$, which is equivalent to $(8.17a)$.

From Algorithm 4 it is a small step to a nongeneric version of the BCG method, which also goes under the names Lanczos/ORTHOMIN [16] and BIOMIN [11]. This normalized nongeneric BIOMIN algorithm is a nearly straightforward application of the above BOBC algorithm to the problem of solving a linear system of equations $A z = b$. As in the generic case [11], the basic strategy is to define a sequence $\{z_n\}$ of approximants in such a way that the vectors $x_n$ generated by Algorithm 4 are the residuals, which means that $\Gamma_n P_n$ is the $n$th residual polynomial. Consequently, for the normalized algorithm we have to choose $\Gamma_n := 1/P_n(0)$, thus producing a breakdown whenever $P_n(0) = 0$. However, the latter equality holds whenever $h^\vee > 1$ and $n^\vee < n < n^\wedge_{l+1}$, cf. (5.28b). This mirrors the fact that the restriction of $A$ that

---

4 The solid overbar denotes complex conjugation.
is implicitly constructed at this stage is singular, and, hence, the projected system cannot be solved in general. Recall that the same difficulty occurred in the nongeneric normalized BiORES algorithm (Algorithm 2 of Part I). But here, the difficulty is easier to recognize since it corresponds exactly to the case $h_j^\gamma > 1$. It is also easier to understand how to circumnavigate it.

By translating (8.9b) into a polynomial recurrence (i.e., by inserting (8.3) and (4.14)) we see that

$$P_{n_{j+1}}(0) \Gamma_{n_{j+1}} = -P_{n_j}(0) \Gamma_{n_j} \varphi_j^\gamma \gamma_{n_{j+1},h_j^\gamma + h_j^\gamma - 1},$$

so that normalization is inherited from $P_{n_j}(0) \Gamma_{n_j}$ to $P_{n_{j+1}}(0) \Gamma_{n_{j+1}}$, i.e., from $x_{n_j}$ to $x_{n_{j+1}}$, simply by choosing

$$\gamma_{n_{j+1},h_j^\gamma + h_j^\gamma - 1} = \frac{-1}{\varphi_j^\gamma}.$$  

In contrast to Algorithm 2, only the iterates $z_{n_j}$ are considered as approximants, and only the corresponding vectors $x_{n_j}$ are true residuals. (It is possible to modify Algorithm 2 accordingly, thus avoiding the breakdown due to normalization. The resulting version is, in fact, just the unnormalized Algorithm 3 with scale factors $\Gamma_{n_j}$, which yield normalized iterates when $n = n_j$.)

**Algorithm 5 (Normalized Nongeneric BOBC Algorithm for Linear Systems: Normalized Nongeneric BiMIN).** For solving $Az = b$, choose an initial approximation $z_0$, set $u_0 := x_0 := b - Az_0$, choose $v_0 := y_0$ with $(y_0, x_0)_B \neq 0$, and apply Algorithm 4 with the special choice (8.21) for the relative scale factors $\gamma_{n_{j+1},h_j^\gamma + h_j^\gamma - 1}$ (which determine $\Gamma_{n_{j+1}}$, while the other scale factors $\Gamma_n (n \neq n_{j+1})$, $\Gamma'$, $\Gamma_n$, and $\Gamma'$ may be chosen arbitrarily nonzero).

Additionally, compute for $j = 0, 1, \ldots$ the approximant $z_{n_{j+1}}$ according to

$$z_{n_{j+1}} = -[u_{n_{j+1}} \gamma_{n_{j+1},1} - u_{n_{j+1}} - 2e_j^\gamma - 2,j \gamma_{n_{j+1},2} - u_{n_{j+1}} - 3e_j^\gamma - 3,j \gamma_{n_{j+1},3} - \cdots - u_{n_{j+1}} \varphi_j^{\gamma} \gamma_{n_{j+1},h_j^\gamma + h_j^\gamma - 1} + z_{n_{j+1}} \varphi_j^{\gamma} \gamma_{n_{j+1},h_j^\gamma + h_j^\gamma - 1}].$$

The algorithm terminates when $n = n_{j+1}$ and $x_n = 0$. Then $n_{j+1} = n$ and $x_n$ is the solution of $Az = b$. However, if $n = n_{j+1} = n_{j+1}$, but $x_n \neq 0$, the solution cannot be found using those initial vectors (a case of incurable breakdown).

As an analogy to the generic case [11] and to Algorithm 3 of Part I, we also suggest an unnormalized version of the nongeneric BiMIN algorithm. It not only avoids the danger of breakdown due to normalization (as our nongeneric normalized BiMIN algorithm does too), but allows to monitor independently the damping effect of the Lanczos polynomials $P_n$ and the often adverse effect of normalization at 0. In this unnormalized version of BiMIN we can choose all the scale factors $\Gamma_n$ arbitrarily. We keep track of them by evaluating a recurrence for $\rho_{n_j} := \Gamma_{n_j} P_n(0)$, which follows from (8.20). Note that $\rho_{n_j} \neq 0 \quad (\forall j)$ in view of $\varphi_j^{\gamma} \neq 0 \quad (\forall j)$. In contrast to our unnormalized nongeneric BiORES algorithm of Part I (Algorithm 3), we restrict ourselves here to this subsequence; thus there is now only a small difference between the normalized and the unnormalized version.

**Algorithm 6 (Unnormalized Nongeneric BOBC Algorithm for Linear Systems: Unnormalized Nongeneric BiMIN).** For solving $Az = b$ choose an
initial approximation \( z_0 \), set \( u_0 := x_0 := b - Az_0 \), choose \( v_0 := y_0 \) with \( (y_0, x_0)_B \neq 0 \), and apply Algorithm 4 (with arbitrary nonzero scale factors \( \Gamma_n, \Gamma'_n, \Gamma_n, \) and \( \Gamma'_n \)). Additionally, compute recursively the vector sequence \( \{z_n^\wedge\} \) according to (8.22) and the scalar sequence \( \{\rho_n^\wedge\} \) according to

\[
\rho_{n+1}^\wedge := -\varphi_j^\wedge \gamma_{n+1}^\wedge, h_j^\wedge + h_j^\wedge - 1 \rho_n^\wedge.
\]

The algorithm terminates when \( n = n_j^\wedge + 1 \) and \( x_n = 0 \). Then, \( n_j^\wedge = n \) and \( z_n/\rho_n \) is the solution of \( Az = b \). However, if \( n = n_j^\wedge + 1 = n_j^\wedge \), but \( x_n \neq 0 \), the solution cannot be found using those initial vectors (a case of incurable breakdown).

By an induction argument we obtain the following theorem, which is analogous to Theorem 4.6.

**Theorem 8.2.** (i) In Algorithm 5 (normalized nongeneric BIOMIN) holds

\[
x_n^\wedge = b - Az_n^\wedge, \quad j = 0, 1, 2, \ldots
\]

(ii) In Algorithm 6 (unnormalized nongeneric BIOMIN) holds

\[
x_n^\wedge = b \rho_n^\wedge - Az_n^\wedge, \quad j = 0, 1, 2, \ldots
\]

**Proof.** Assume that (8.25) holds up to a certain \( j \). Using the formulas (8.23), (8.22), and (8.9b) of Algorithm 6 we get

\[
b \rho_{n+1}^\wedge - Az_{n+1}^\wedge = -b \varphi_j^\wedge \gamma_{n+1}^\wedge, h_j^\wedge + h_j^\wedge - 1 \rho_n^\wedge + A[u_{n+1}^\wedge - \gamma_{n+1}^\wedge, 1]
\]

Hence, (8.25) follows by induction. In Algorithm 5, (8.21) guarantees that \( \rho_n^\wedge \), defined by (8.23), is 1 for all \( j \), so that (8.24) holds. \( \square \)

Finally, we want to sketch the nongeneric generalization of yet another important algorithm, namely, of BIODIR [11] (or Lanczos/ORTHODIR [16]). As in the generic case [11], we first define a “biconjugation algorithm,” which is nothing more than the BO algorithm with the inner product \( (\cdot, \cdot)_B \) instead of \( (\cdot, \cdot)_B \). This algorithm can be used to generate the sequences \( \{u_n\} \) and \( \{v_n\} \) alone, without concurrently building up \( \{x_n\} \) and \( \{y_n\} \).

**Algorithm 7 (Nongeneric “Biconjugation (BC) Algorithm”).** Given a bounded linear operator \( A : H \to H \) and two initial vectors \( u_0, v_0 \in H \) satisfying \( (u_0, Av_0)_B \neq 0 \), apply the nongeneric BO algorithm (Algorithm 1) with the inner product \( (\cdot, \cdot)_B \) (instead of \( (\cdot, \cdot)_B \)) to produce the two vector sequences \( \{u_n\} \) and \( \{v_n\} \) with scale factors \( \Gamma_n \) and \( \Gamma_n \) and the matrix \( H' \) containing the recurrence coefficients of the corresponding FOP1s. Denote the indices of the regular FOP1s by \( n_i \), and let \( h_i := n_i + 1 - n_i \).

In view of their orthogonality properties, the resulting vector sequences are the same as the sequences \( \{u_n\} \), \( \{v_n\} \) generated by the BOBC algorithm (if the initial vectors and the scale factors are the same). By applying half a step of the nongeneric backward qd algorithm (specified by Theorem 7.3), we can find the factors \( G^\wedge \) and \( G^\vee \) of the relevant block UL decomposition of \( H' \). Finally, we can apply formulas
(8.9b) and (8.22) to compute the subsequences $x_{n_j}$ and $z_{n_j}$, and (8.23) to find the appropriate scale factors for the normalization given below.

**Algorithm 8 (Normalized Nongeneric BIODIR).** For solving $Ax = b$, choose an initial approximation $z_0$, set $u_0 := x_0 := b - Ax_0$, choose $v_0 := y_0$ with $\langle y_0, x_0 \rangle_B \neq 0$, and apply Algorithm 7 (with arbitrary nonzero scale factors $\Gamma'_n$ and $\Gamma''_n$) in order to produce the two vector sequences $\{u_n\}$ and $\{v_n\}$ and the matrix $H'$ of recurrence coefficients of the corresponding FOP1s. Concurrently, compute block by block the relevant block UL decomposition $H' = G^\dagger G^\vee$ according to Theorem 7.3. The initial moments required for that are

\begin{equation}
\phi_0 := \frac{\langle y_0, x_0 \rangle_B}{\Gamma_0 \Gamma_0}, \quad \phi_{n_1'} := \frac{\langle v_0, Au_{n_1'-1} \rangle_B}{\Gamma_0 \Gamma_0' n_1' - 1}.
\end{equation}

Additionally, for $j = 0, 1, \ldots$ compute the subsequences $\{x_{n_j}\}$ and $\{z_{n_j}\}$ according to (8.9b) and (8.22), the value of $\gamma_{n_j+1} h_{n_j}^+ + h_{n_j}^- - 1$ being given by (8.21). (This determines the scale factors $\Gamma_{n_j}$, while $\Gamma_n (n \neq n_j)$ can be chosen arbitrarily nonzero.)

The algorithm terminates when $n = n_1$ and $x_n = 0$. Then $n_1 = n$ and $z_n$ is the solution of $Ax = b$. If $n = n_1 + 1 = n_1'$, but $x_n \neq 0$, the solution cannot be found using those initial vectors (a case of incurable breakdown).

This algorithm is normalized in the same sense as Algorithm 5; the residual polynomials $\Gamma_{n_j} P_{n_j}$ of the approximants $z_{n_j}$ are normalized to 1 at 0. Of course, we could try to replace the backward qd step by the solution of an extra triangular system of equations similar to (8.11c) and (8.11d). But this would require computation of both extra vectors and inner products.

Note that the breakdown conditions for the nongeneric versions of unnormalized BIORES, normalized and unnormalized BIOMIN, and normalized BIODIR are all the same. This is in contrast to the generic versions of these algorithms [11].

**9. The treatment of near-breakdown for diagonal sequences.** So far we have assumed that we work with exact arithmetic and that, therefore, the regular formal orthogonal polynomials (FOP1s) are well defined, and the corresponding elements of the various vector sequences generated by Lanczos-type algorithms can be computed accurately. Index steps $h_{j'}$ of size greater than 1 between regular FOP1s occur as a consequence of serious, but curable, breakdown. However, in practice exact curable breakdown is very unlikely, but near-breakdown may occur as a consequence of either an exact breakdown contaminated by roundoff or a very small $|\delta_j|$. Any such near-breakdown means that the recurrence coefficients $\alpha_{j+1}$ and $\beta_{j+1}$ are probably large and numerically not well determined; then the subsequent FOP1s and the corresponding Krylov space vectors must be expected to be inaccurate.

Therefore, one must find a way to treat near-breakdowns. The simplest approach would be to use exactly the same formulas as for the exact curable breakdown. This would mean that we proceed implicitly with slightly modified data, but process them in a stable way, instead of treating the original data in an instable way. However, in this section we show that we can do even better. It is possible to treat near-breakdown exactly and still fairly efficiently. If our previously defined algorithms are modified accordingly, then, in exact arithmetic, those regular FOP1s that are well conditioned are obtained independently of the threshold used to define near-breakdown. The same is true for the corresponding Krylov space vectors. (Of course, the number of “well-conditioned” regular FOP1s depends on the threshold.) Here we present only
the polynomial formulation of these algorithms. Details of implementation for the corresponding Lanzos-type algorithms are described in joint work with Freund and Nachtigal [8].

In this section we treat ordinary, diagonal sequences of FOP1s. Recall that \( \{n_j\}_{j=0}^J \) denotes the index sequence of the regular FOP1s \( P_{n_j} \) for some functional \( \Phi \), and that this sequence is characterized by

\[
\begin{align*}
(9.1a) & \quad \Phi(p_{P_{n_j}}^2) = 0 \quad (\forall p \in P_{n_{j+1}-2}), \\
(9.1b) & \quad \Phi(z^{h_j-1}p_{n_j}^2) =: \delta_j \neq 0,
\end{align*}
\]

where \( h_j := n_{j+1} - n_j, \) \( j = 0, \ldots, J - 1 \) \( (J \leq \infty) \). This recursive definition of the sequence \( \{n_j\} \) implies that for \( j < J \) the diagonal blocks

\[
D_j = \begin{bmatrix}
\delta_j & \ast \\
& \ddots & \ddots & \ast \\
& \ast & \ast & \ddots & \ast \\
\ast & \ast & \ast & \ast & \ast
\end{bmatrix}
\]

of the formal Gramian \( D \) in (3.23) and (8.14) are nonsingular. For the intermediate values of \( n \) (i.e., for those satisfying \( n_j < n < n_{j+1} \) for some \( j \)) the FOP1s are not uniquely determined, and we made the particular choice (1.28) for these inner FOP1s.

Now, we want to extract an index subsequence \( \{\tilde{n}_k\}_{k=0}^K \subset \{n_j\}_{j=0}^J \) that marks the well-conditioned regular FOP1s \( P_{\tilde{n}_k} \). Before we come to its recursive definition we choose first, by analogy to (1.28), tentative inner FOP1s for this subsequence:

\[
(9.2) \quad \hat{p}_n(z) := W_{n-\tilde{n}_k}(z)P_{\tilde{n}_k}(z) \quad \text{if} \quad \tilde{n}_k \leq n < \tilde{n}_{k+1}.
\]

Here \( W_m \) is still a prescribed monic sequence, for example, one satisfying a three-term recurrence (2.10). In the latter case the tentative inner polynomials \( \hat{p}_n \) themselves are obtained by a three-term recurrence. For simplicity, we could choose \( W_m(z) = z^m \), although in practice this is often a rather inappropriate basis. To simplify formulas we include in (9.2) \( n = \tilde{n}_k \), where \( \hat{p}_n = P_{\tilde{n}_k} \) is regular and thus not inner. Actually, when computing the next well-conditioned regular FOP1 \( P_{\tilde{n}_{k+1}} \), we also start from a polynomial of the form (9.2), with \( n = \tilde{n}_{k+1} \), which is then orthogonalized with respect to the previous blocks. We assume here that for those FOP1s this orthogonalization process has already been carried out, so that (9.2) holds for \( \tilde{n}_k \leq n < \tilde{n}_{k+1} \) instead of for \( \tilde{n}_k < n \leq \tilde{n}_{k+1} \).

One is tempted to define the index steps \( \tilde{h}_k := \tilde{n}_{k+1} - \tilde{n}_k \) by analogy to (9.1) by

\[
\begin{align*}
(9.3a) & \quad \Phi(p_{P_{\tilde{n}_k}}^2) = 0 \quad (\forall p \in P_{\tilde{n}_{k-1}}), \\
(9.3b) & \quad |\Phi(z^{i}p_{\tilde{n}_k}^2)| \leq \varepsilon \quad (0 \leq i \leq \tilde{h}_k - 2), \\
(9.3c) & \quad \Phi(z^{\tilde{h}_k-1}p_{\tilde{n}_k}^2) =: \tilde{\delta}_k, \quad |\tilde{\delta}_k| > \varepsilon,
\end{align*}
\]

where \( \varepsilon > 0 \) is some prescribed small constant. This would imply that

\[
\Phi(\hat{p}_m\hat{p}_n) = \begin{cases} 
0 & \text{if} \quad \tilde{n}_k \leq n \leq \tilde{n}_{k+1} - 2 \quad \text{and} \quad m + n < 2\tilde{n}_k, \\
O(\varepsilon) & \text{if} \quad \tilde{n}_k \leq n \leq \tilde{n}_{k+1} - 2 \quad \text{and} \quad 2\tilde{n}_k \leq m + n \leq \tilde{n}_k + \tilde{n}_{k+1} - 2, \\
\tilde{\delta}_k + O(\varepsilon) & \text{if} \quad \tilde{n}_k \leq n \leq \tilde{n}_{k+1} - 1 \quad \text{and} \quad m + n = \tilde{n}_k + \tilde{n}_{k+1} - 1.
\end{cases}
\]
For example, if $K = 5$ and $\tilde{n}_0 = 0$, $\tilde{n}_1 = 1$, $\tilde{n}_2 = 2$, $\tilde{n}_3 = 5$, $\tilde{n}_5 = 7$, and if $\varepsilon$, $\delta_k$, and $\ast$ are abbreviations for $O(\varepsilon)$, $\delta_k + O(\varepsilon)$, and $O(1)$, respectively, the formal Gramian of $\{\tilde{p}_n\}$ is then of the form

\[
(9.5) \quad [\Phi(\tilde{p}_m \tilde{p}_n)]_{m,n=0}^\infty = \begin{bmatrix}
\delta_0 \\
\delta_1 \\
\delta_2 \\
\delta_3 \\
\delta_4 \\
\delta_5 \\
\vdots 
\end{bmatrix}
\]

There are two serious problems with this approach: First, while the formal Gramian $D$ of the FOP1s $\{P_n\}$ is block diagonal, this is obviously no longer true for the one of $\{\tilde{p}_n\}$. Here, small $O(\varepsilon)$ elements can penetrate into several off-diagonal blocks. As a consequence, it can be seen that the Hessenberg matrix of recurrence coefficients for $\{\tilde{p}_n\}$ is no longer block tridiagonal; the formula for the regular FOP1 $\tilde{p}_{n_{k+1}}$ may contain not only terms from the two previous blocks but even terms from older ones. Hence, the major advantage of the generic and the nongeneric Lanczos algorithms is lost here. Second, $\delta_k$ and thus the sequence $\{\tilde{p}_n\}$ is generally not appropriately defined. Since $|\delta_k|$ need not be much larger than $\varepsilon$, $|\delta_k + O(\varepsilon)|$ may be small or may even vanish, and the diagonal blocks of the formal Gramian may be singular.

The second difficulty can be overcome if we transform the sequence, $\{\tilde{p}_n\}$, by a block orthogonalization process into a new formally block orthogonal sequence, $\{\tilde{p}_n\}$, whose formal Gramian is block diagonal. The index subsequence $\{\tilde{n}_k\}$ is then defined by requiring that the diagonal blocks

\[
(9.6) \quad \tilde{D}_k = [\Phi(\tilde{p}_{n_k+i} \tilde{p}_{n_k+j})]_{i,j=0}^{n_k-1}
\]

of this formal Gramian $\tilde{D}$ are not only nonsingular, but neither near-singular nor ill conditioned:

\[
(9.7) \quad \sigma_{\text{min}}(\tilde{D}_k) > \varepsilon, \quad \kappa(\tilde{D}_k) \leq \kappa'.
\]

The appropriate order of magnitude of $\varepsilon$ and of the upper bound $\kappa'$ of the condition number $\kappa$ depend on the functional $\Phi$. If the latter or the inner polynomials are scaled such that $\sigma_{\text{max}}(\tilde{D}_k)$ is a priori bounded, it suffices—at least in theory—to make the first check. This is in accordance with the recommendation of Parlett [21]. Note that—also just in theory—the near-breakdown threshold $\varepsilon$ need not be very small and the block size chosen need not be the minimal one satisfying (9.7). Unfortunately, in practice, roundoff spoils the orthogonality and the determination of the block size according to (9.7) may not work for large-scale systems. Therefore, a different strategy has to be chosen in the implementation (see [8] for details).

---

5 Confer footnote 2 at the end of the introduction.
However, as is seen from (9.5), the block orthogonalization process suggested above also has the disadvantage that it may involve sums that contain elements from several previous blocks, i.e., the upper triangular matrix describing this process is not block bidiagonal, as one might have hoped. In fact, "long" recurrences may occur here not only for some \( \tilde{p}_n \), but also for some inner \( \tilde{p}_n \).

Things change if we build up the block orthogonal sequence \( \{\tilde{p}_n\} \) directly by orthogonalizing \( \tilde{z}\tilde{p}_n \) with respect to previous blocks. In fact, by a standard argument we are going to conclude that the Hessenberg matrix \( \tilde{H} \) of recurrence coefficients is then block tridiagonal. The freedom of choice present in (9.2) is gained from the possibility of adding any linear combination of polynomials that have already been computed in the block under construction. (The coefficients in this linear combination are the same as those in the recurrence (2.10).) Since this linear combination is already orthogonal to all previous blocks, it has no effect on the off-diagonal blocks of \( \tilde{H} \).

Let \( \tilde{p} := [\tilde{p}_0, \tilde{p}_1, \ldots] \) be the infinite row vector whose elements are the resulting monic block orthogonalized polynomials \( \tilde{p}_n \). Then the block diagonal formal Gramian of these polynomials can be written as

\[
\tilde{D} := \Phi(\tilde{p}^T\tilde{p}) := [\Phi(\tilde{p}_m\tilde{p}_n)]_{m,n=0}^\infty.
\]

(Recall that its diagonal blocks satisfy (9.7) by definition of \( \{\tilde{n}_k\} \).)

Since the polynomials are monic of ascending degree, they can certainly be generated by a recurrence that has the matrix form

\[
z\tilde{p}(z) = \tilde{p}(z)\tilde{H},
\]

with \( \tilde{H} \) being unit upper Hessenberg. Consequently,

\[
\begin{align*}
\tilde{D}\tilde{H} &= \Phi(\tilde{p}^T\tilde{p})\tilde{H} = \Phi(\tilde{p}^T\tilde{p}\tilde{H}) = \Phi(\tilde{z}\tilde{p}) \\
&= \Phi(\tilde{H}^T\tilde{p}^T\tilde{p}) = \tilde{H}^T\Phi(\tilde{p}^T\tilde{p}) = \tilde{H}^T\tilde{D}.
\end{align*}
\]

Since \( \tilde{H} \) is upper Hessenberg and \( \tilde{D} \) is block diagonal, the left-hand side is upper block Hessenberg and the right-hand side is lower block Hessenberg. Consequently, both sides are block tridiagonal, and multiplication by \( \tilde{D}^{-1} \) shows that the same is true for \( \tilde{H} \) itself, i.e., by analogy to (3.6),

\[
\tilde{H} := \begin{bmatrix}
\tilde{A}_0 & \tilde{B}_1 & & \\
\tilde{C}_0 & \tilde{A}_1 & \tilde{B}_2 & \\
& \ddots & \ddots & \\
& & \tilde{C}_{K-1} & \tilde{B}_K \\
& & & \tilde{A}_K
\end{bmatrix}.
\]

In general the block structure is coarser here than in (3.6), but all block boundaries in this matrix are also present in the matrix \( H \) of (3.6). In view of \( \tilde{H} \) being unit upper Hessenberg, the blocks \( \tilde{C}_k \) have again just a 1 in the upper right corner, and the blocks \( \tilde{A}_k \) are either 1 \( \times \) 1 or unit upper Hessenberg. It remains to investigate the upper triangular part of \( \tilde{A}_k \) and the superdiagonal blocks \( \tilde{B}_k \). Moreover, we have to discuss how to determine the elements of these blocks.

The only condition we are imposing is that \( \{\tilde{p}_n\} \) be formally block orthogonal, i.e., \( \Phi(\tilde{p}^T\tilde{p}) = \tilde{D} \) is block diagonal. Splitting the row vector \( \tilde{p} \) into blocks of size \( h_k \), we set \( \tilde{p} = [\tilde{p}_0, \tilde{p}_1, \ldots, (\tilde{p}_K)] \) to get

\[
\tilde{D}_k = \Phi(\tilde{p}_k^T\tilde{p}_k) \tag{9.12a}
\]
From the $k$th block column of (9.10a) we extract the three conditions

\begin{align}
(9.12b) \quad \tilde{D}_{k-1} \tilde{B}_k &= \Phi(\tilde{p}_{k-1}^T \tilde{p}_k), \\
(9.12c) \quad \tilde{D}_k \tilde{A}_k &= \Phi(\tilde{p}_k^T \tilde{z} \tilde{p}_k), \\
(9.12d) \quad \tilde{D}_{k+1} \tilde{C}_k &= \Phi(\tilde{p}_{k+1}^T \tilde{z} \tilde{p}_k).
\end{align}

If $\tilde{p}_{k-1}$, $\tilde{p}_{n_k}$, and $\tilde{D}_{k-1}$ are known, (9.12a)-(9.12c) allow us, if they are used column by column and in parallel, to build up $\tilde{p}_k$, $\tilde{B}_k$, $\tilde{A}_k$, and $\tilde{D}_k$, and to determine $\tilde{h}_{k+1}$ and $\tilde{p}_{n_{k+1}}$. Actually, except for its last column, $\tilde{A}_k$ can be chosen as an arbitrary unit upper Hessenberg matrix, hence we assume it given, except for the last column. For example, each of the other columns may be zero except for a 1 on the subdiagonal.

First, since $\Phi(\tilde{p}_p) = 0^T$ for all $p \in \mathcal{P}_{n_k-1}$, only the last line of (9.12b) is nonzero:

\begin{equation}
(9.13a) \quad \tilde{D}_{k-1} \tilde{B}_k = \tilde{I}_k \phi_k^T,
\end{equation}

where

\begin{equation}
(9.13b) \quad \phi_k^T := \Phi(\tilde{z} \tilde{p}_{n_k-1} \tilde{p}_k), \quad \tilde{I}_k := [0, \ldots, 0, 1]^T \in \mathbb{C}^{\tilde{n}_k}.
\end{equation}

Hence, $\tilde{B}_k$ has rank 1. Once $\tilde{p}_n$ is known for some $n$ with $(\tilde{n}_k \leq n \leq \tilde{n}_{k+1} - 1)$, column $n$ of $\tilde{B}_k$ is obtained by solving a linear system with the coefficient matrix $\tilde{D}_{k-1}$, which is no longer triangular, but has constant nonzero anti-diagonal elements and a small upper left triangular part.\(^6\) If $n < \tilde{n}_{k+1} - 1$, the corresponding column of $\tilde{A}_k$ is prescribed, and thus $\tilde{p}_{n+1}$ can be computed according to (9.9). Moreover, the element $n + 1$ in the first row (i.e., row $\tilde{n}_k + 1$) of $\tilde{D}_k$ can be evaluated explicitly using the definition (9.12a) of $\tilde{D}_k$. Once $n + 1 = \tilde{n}_{k+1} - 1$, the whole first column of $\tilde{D}_k$ is known, as is its first row, thanks to symmetry.

After splitting off the first row and column of $\tilde{D}_k$ and the first row and the last column of $\tilde{A}_k$, (9.12c) is seen to yield a set of $\tilde{n}_k - 1$ triangular systems for computing the yet unknown elements of $\tilde{D}_k$. Then, by (9.12b), too, the last column $\tilde{a}_k$ of $\tilde{A}_k$ is also found by solving a linear system with coefficient matrix $\tilde{D}_k$:

\begin{equation}
(9.14) \quad \tilde{D}_k \tilde{a}_k = \Phi(\tilde{p}_k^T \tilde{z} \tilde{p}_{n_{k+1}-1}).
\end{equation}

Finally, now that the last columns of $\tilde{A}_k$ and $\tilde{B}_k$ are known, the recurrence for the next well-conditioned regular FOP1, $\tilde{p}_{n_{k+1}}$, is ready.

Due to the special structure of $\tilde{C}_k$, the only part of $\tilde{D}_k$ that matters in (9.12d) is the first column. It is easy to see that for the elements in this first column (9.12d) provides formulas that are mathematically equivalent to those from the definition (9.12a) of $\tilde{D}_k$.

Summarizing, we see that on the basis of the relations (9.12) we can build up $\tilde{H}$ and $\tilde{D}$ column by column.

One may wonder what can be said about the order of magnitude of the elements of $\tilde{D}_k$ and $\tilde{B}_k$ if (9.3) holds for some $\varepsilon << 1$ and with $|\delta_k'| >> 1$ ($k' = k - 1, k$). In

---

\(^6\) Column $n$ of $\tilde{B}_k$ refers to the part of column $n$ of $\tilde{H}$ that lies in $\tilde{B}_k$; rows and elements are referred to in an analogous fashion.
the notation of (9.5) we clearly have, for $0 \leq k < K$,

$$
\tilde{D}_k = \begin{bmatrix}
\varepsilon & \varepsilon & \cdots & \varepsilon & \tilde{\delta}_k \\
\varepsilon & & \ddots & \ddots & \ddots \\
\varepsilon & & \ddots & \ddots & \ddots \\
\tilde{\delta}_k & \ddots & \ddots & \ddots & \ddots \\
\end{bmatrix}.
$$

(9.15)

(When $K < \infty$, the last diagonal block $\tilde{D}_K$ is infinite and consists of elements of order $O(\varepsilon)$.) From (9.15) it is easy to conclude that

$$
\tilde{D}_k^{-1} = \begin{bmatrix}
* & * & \cdots & * \\
* & \varepsilon & \cdots & \varepsilon \\
\ddots & \ddots & \ddots & \ddots \\
* & \tilde{\delta}_k & \cdots & \varepsilon \\
\end{bmatrix},
$$

(9.16)

and, therefore, that

$$
\tilde{B}_k = \tilde{D}_k^{-1} \tilde{I}_k \Phi_k^T = \begin{bmatrix}
\varepsilon & \cdots & \tilde{\delta}_k & * \\
\varepsilon^2 & \cdots & \tilde{\delta}_k & \varepsilon \\
\ddots & \ddots & \ddots & \ddots \\
\varepsilon^2 & \cdots & \tilde{\delta}_k & \varepsilon \\
\end{bmatrix},
$$

(9.17)

where $\varepsilon^2 := O(\varepsilon^2)$.

Actually, the order-of-magnitude statement in (9.17) can be seen to be a consequence of the block diagonality of $\tilde{D}$, the block tridiagonality of $\tilde{H}$, and the special structure of the blocks $D_k$ and $C_k$. Comparing superdiagonal blocks in the relation (9.10), we get $\tilde{D}_{k-1} \tilde{B}_k = C_{k-1}^T \tilde{D}_k$; hence

$$
\tilde{B}_k = \tilde{D}_{k-1}^{-1} C_{k-1}^T \tilde{D}_k = \begin{bmatrix}
* & * & \cdots & * \\
* & \varepsilon & \cdots & \varepsilon \\
\ddots & \ddots & \ddots & \ddots \\
* & \varepsilon & \cdots & \varepsilon \\
\end{bmatrix} \begin{bmatrix}
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
\end{bmatrix} = \begin{bmatrix}
\varepsilon & \cdots & \tilde{\delta}_k & * \\
\varepsilon^2 & \cdots & \tilde{\delta}_k & \varepsilon \\
\ddots & \ddots & \ddots & \ddots \\
\varepsilon^2 & \cdots & \tilde{\delta}_k & \varepsilon \\
\end{bmatrix}.
$$

Comparing (9.13), (9.17), and (9.18), we get the additional relation

$$
\phi_k^T := \Phi(z \tilde{p}_{\gamma_k} \tilde{p}_k) = \tilde{I}_k \phi_k^T = C_{k-1}^T \tilde{D}_k = \Phi(\tilde{p}_{\gamma_k} \tilde{p}_k),
$$

(9.19)

which is, indeed, a consequence of $\tilde{p}_n$ ($n \geq \gamma_k$) being orthogonal to $\mathcal{P}_{\gamma_{k-1}}$.

10. The treatment of near-breakdown for staircase sequences. In this section we modify the ideas of the previous section in order to compute a block staircase sequence of well-conditioned FOP1s instead of a block diagonal sequence.

The aim is to determine two full sequences $\tilde{p} := \{\tilde{p}_n\}_{n=0}^\infty$, $\tilde{p}' := \{\tilde{p}'_n\}_{n=0}^\infty$ of monic polynomials ($\tilde{p}_n$ and $\tilde{p}'_n$ being of degree $n$) that are block orthogonal, well conditioned, and block compatible in the following sense. There are two (finite or infinite) index sequences $\{\tilde{n}_k^\wedge\}$, $\{\tilde{n}_k^\vee\}$ ($k = 0, \ldots, K^\wedge + 1 \leq K^\vee + 1 \leq \infty$) satisfying

$$
\tilde{n}_k^\wedge \leq \tilde{n}_k^\vee (k = 0, \ldots, K^\vee), \quad \tilde{n}_k^\wedge < \tilde{n}_{k+1}^\wedge (k = 0, \ldots, K^\wedge - 1)
$$

(10.1)
such that the following two conditions hold: First

\( (10.2a) \quad \Phi(p_m p_n) = 0 \) if \( m < \bar{n}_k^\vee \leq n \) for some \( k \)

or \( m \leq \bar{n}_k^\wedge < n \) for some \( k \),

\( (10.2b) \quad \Phi'(p'_m p'_n) = 0 \) if \( m < \bar{n}_k^\wedge \leq n \) for some \( k \)

or \( m < \bar{n}_{k+1}^\wedge \leq n \) for some \( k \),

which means that the two formal Gramians

\( (10.3) \quad \tilde{D}^\wedge := \Phi(p^T p), \quad \tilde{D}^\vee := \Phi'((p')^T p') \)

are block diagonal, with blocks starting at the columns \( \bar{n}_k^\wedge, \bar{n}_k^\vee + 1 \) \( (\leq \bar{n}_{k+1}^\wedge) \) and \( \bar{n}_k^\wedge, \bar{n}_{k+1}^\wedge \) \( (\leq \bar{n}_{k+1}^\vee) \), respectively. Second, for their diagonal blocks

\( (10.4a) \quad \tilde{D}_k^\wedge := [\Phi(p_m p_n)]_{m,n=\bar{n}_k^\wedge} \), \( \tilde{D}_{k+\frac{1}{2}}^\wedge := [\Phi(p_m p_n)]_{m,n=\bar{n}_{k+1}^\wedge} \),

\( (10.4b) \quad \tilde{D}_k^\vee := [\Phi'(p'_m p'_n)]_{m,n=\bar{n}_k^\vee} \)

holds, by analogy to (9.7), (for some \( \varepsilon > 0, \kappa' > 1 \)):

\( (10.5a) \quad \sigma_{\min}(\tilde{D}_k^\wedge) > \varepsilon, \quad \kappa(\tilde{D}_k^\wedge) \leq \kappa' \),

\( (10.5b) \quad \sigma_{\min}(\tilde{D}_{k+\frac{1}{2}}^\wedge) > \varepsilon, \quad \kappa(\tilde{D}_{k+\frac{1}{2}}^\wedge) \leq \kappa' \),

\( (10.5c) \quad \sigma_{\min}(\tilde{D}_k^\vee) > \varepsilon, \quad \kappa(\tilde{D}_k^\vee) \leq \kappa' \),

\( (10.5d) \quad \sigma_{\min}(\tilde{D}_{k-\frac{1}{2}}^\vee) > \varepsilon, \quad \kappa(\tilde{D}_{k-\frac{1}{2}}^\vee) \leq \kappa' \).

Note that some of the blocks \( \tilde{D}_{k+\frac{1}{2}}^\wedge \) and \( \tilde{D}_{k-\frac{1}{2}}^\vee \) may be void. They exist only if the respective index step

\( (10.6) \quad \bar{h}_k^\vee := \bar{n}_{k+1}^\vee - \bar{n}_k^\vee, \quad \bar{h}_k^\wedge := \bar{n}_k^\wedge - \bar{n}_{k-1}^\wedge \)

is larger than 1. Then \( \tilde{D}_{k+\frac{1}{2}}^\wedge \) lies in \( \tilde{D}^\wedge \) between \( \tilde{D}_k^\wedge \) and \( \tilde{D}_{k+1}^\wedge \), and \( \tilde{D}_{k-\frac{1}{2}}^\vee \) lies in \( \tilde{D}^\vee \) between \( \tilde{D}_{k-1}^\vee \) and \( \tilde{D}_k^\vee \). The blocks in (10.4a) have order \( \bar{h}_k^\wedge \) and \( \bar{h}_k^\wedge - 1 \); those in (10.4b) have order \( \bar{h}_k^\vee \) and \( \bar{h}_k^\vee - 1 \), respectively.

The recursive process for constructing these two sequences, like the one in §5, alternates between the two sequences. The block structure generated in this way in the corresponding analog of the Padé table (which is a “near-FOP1 table”) still resembles the one of Fig. 3, but at this point it is restricted to the two diagonals specified by \( l \) and \( l + 1 \). A conflict arises because we require the four conditions (10.5a)–(10.5d), although (10.5a) and (10.5c) would be enough to determine the two index sequences \( \{\bar{n}_k^\wedge\} \) and \( \{\bar{n}_k^\vee\} \). But it is important that \( \tilde{p}_{\bar{n}_k^\wedge} \) lies in the first row of a block, and \( \tilde{p}_{\bar{n}_{k+1}^\wedge} \) in the first column, as happens automatically in §5, cf. Fig. 3. This must now be enforced by a condition guaranteeing that the blocks \( \tilde{D}_{k+\frac{1}{2}}^\wedge \) and \( \tilde{D}_{k-\frac{1}{2}}^\vee \) are nonsingular. In order that \( \tilde{p}_{\bar{n}_k^\wedge} \) and \( \tilde{p}_{\bar{n}_{k+1}^\wedge} \) be well-conditioned regular FOP1s, these blocks can be neither near-singular nor ill-conditioned. Consequently, on each diagonal the sizes of some blocks may turn out to be larger than when we apply the algorithm of the previous section to the respective diagonal. But this is the price we have to pay for a “compatible” block structure on the two diagonals, i.e., index sequences \( \{\bar{n}_k^\wedge\}, \{\bar{n}_k^\vee\} \), which together define the blocks on both diagonals.

Now we come to the details of the algorithm. Its basic pattern is as follows:
(I) When \( n = \tilde{n}_k^\wedge \), \( \Phi \)-orthogonalize \( z\tilde{p}_{n-1}^\wedge \) with respect to \( \mathcal{P}_{\tilde{n}_k^\wedge -1} \) to get \( \tilde{p}_n^\wedge \); check (10.5a) to determine whether \( \tilde{D}_k^\wedge \) is just \( 1 \times 1 \). If yes, set \( \tilde{n}_k^\wedge := n \) and proceed with (iii).

(ii) If not, i.e., if \( \tilde{n}_k^\wedge > \tilde{n}_k^\wedge \), then, for \( n = \tilde{n}_k^\wedge + 1, \tilde{n}_k^\wedge + 2, \ldots \), \( \Phi' \)-orthogonalize \( z\tilde{p}_{n-1}^\wedge \) with respect to \( \mathcal{P}_{\tilde{n}_k^\wedge -1} \) to get \( \tilde{p}_n^\wedge \), and set \( \tilde{p}_n := z\tilde{p}_{n-1}^\wedge \); check (10.5a) and (10.5d) to determine whether \( \tilde{D}_k^\wedge \) and \( \tilde{D}_{k-1}^\wedge \) are completed, in which case \( \tilde{n}_k^\wedge := n \).

(iii) When \( n = \tilde{n}_k^\wedge \), \( \Phi' \)-orthogonalize \( \tilde{p}_n \) with respect to \( \mathcal{P}_{\tilde{n}_k^\wedge -1} \) to get \( \tilde{p}_n^\wedge \); check (10.5c) to determine whether \( \tilde{n}_k^\wedge \) is just \( 1 \times 1 \). If yes, set \( \tilde{n}_k^\wedge := n \) and proceed with (i).

(iv) If not, i.e., if \( \tilde{n}_k^\wedge + 1 > \tilde{n}_k^\wedge + 1 \), then, for \( n = \tilde{n}_k^\wedge + 1, \tilde{n}_k^\wedge + 2, \ldots \), \( \Phi \)-orthogonalize \( z\tilde{p}_{n-1}^\wedge \) with respect to \( \mathcal{P}_{\tilde{n}_k^\wedge -1} \) to get \( \tilde{p}_n \) and set \( \tilde{p}_n^\wedge := \tilde{p}_n \); check (10.5c) and (10.5b) to determine whether \( \tilde{D}_k^\wedge \) and \( \tilde{D}_{k-1}^\wedge \) are completed, in which case \( \tilde{n}_k^\wedge + 1 := n + 1 \).

Note that the choice \( \tilde{p}_n := z\tilde{p}_{n-1}^\wedge \) in (ii) implies that \( \tilde{p}_n \) is \( \Phi \)-orthogonal to \( \mathcal{P}_{\tilde{n}_k^\wedge -1} \); analogously, the choice \( \tilde{p}_n^\wedge := \tilde{p}_n \) in (iv) implies that \( \tilde{p}_n^\wedge \) is \( \Phi' \)-orthogonal to \( \mathcal{P}_{\tilde{n}_k^\wedge -1} \).

By analogy to the generalities introduced in §9, one can modify \( \tilde{p}_{n-1} \) and \( z\tilde{p}_{n-1}^\wedge \) in (ii), and \( \tilde{p}_n \) and \( \tilde{p}_n^\wedge \) in (iv), by adding a linear combination of polynomials (of the other type) that have already been found in the respective substep.

Since the polynomials \( \tilde{p}_n \) (and, likewise, \( \tilde{p}_n^\wedge \)) are not orthogonal to each other within the blocks, the orthogonalization procedures called in the algorithm have to make use of the inverses \( (\tilde{D}_k^\wedge)^{-1} \), \( (\tilde{D}_{k+1}^\wedge)^{-1} \), \( (\tilde{D}_{k-1}^\wedge)^{-1} \), and \( (\tilde{D}_{k-1}^\wedge)^{-1} \), respectively.

As is seen from this recipe, another complication that arises is that the analog of (5.28) does not hold. We still have

\[
\begin{align*}
(10.7a) & \quad \Phi(p\tilde{p}_n) = 0 \quad (\forall p \in \mathcal{P}_{\tilde{n}_k^\wedge -1}) \Rightarrow \Phi'(p\tilde{p}_n) = 0 \quad (\forall p \in \mathcal{P}_{\tilde{n}_k^\wedge -2}), \\
(10.7b) & \quad \Phi'(p\tilde{p}_n) = 0 \quad (\forall p \in \mathcal{P}_{\tilde{n}_k^\wedge -1}) \Rightarrow \Phi(zp\tilde{p}_{n-1}^\wedge) = 0 \quad (\forall p \in \mathcal{P}_{\tilde{n}_k^\wedge -1}), \\
(10.7c) & \quad \Phi(p\tilde{p}_n) = 0 \quad (\forall p \in \mathcal{P}_{\tilde{n}_k^\wedge -1}) \Rightarrow \Phi'(p\tilde{p}_n) = 0 \quad (\forall p \in \mathcal{P}_{\tilde{n}_k^\wedge -1}), \\
(10.7d) & \quad \Phi'(p\tilde{p}_n) = 0 \quad (\forall p \in \mathcal{P}_{\tilde{n}_k^\wedge -1}) \Rightarrow \Phi(zp\tilde{p}_{n-1}^\wedge) = 0 \quad (\forall p \in \mathcal{P}_{\tilde{n}_k^\wedge -1}),
\end{align*}
\]

and we know that the left-hand sides are true when \( \tilde{n}_k^\wedge \leq n \) in (a), \( \tilde{n}_k^\wedge < n \) in (b) and (c), and \( \tilde{n}_k^\wedge < n \) in (d), respectively; but, in general,

\[
(10.8) \quad \Phi'(z\tilde{n}_k^\wedge -1\tilde{p}_n) \neq 0 \quad \text{and} \quad \Phi(z\tilde{n}_k^\wedge z\tilde{p}_{n-1}^\wedge) \neq 0.
\]

Therefore, \( \tilde{p}_n \) needs to be \( \Phi' \)-orthogonalized with respect to \( z\tilde{n}_k^\wedge -1 \) or \( z\tilde{n}_k^\wedge -1 \) to get \( \tilde{p}_n^\wedge \); likewise \( z\tilde{p}_{n-1}^\wedge \) needs to be \( \Phi \)-orthogonalized with respect to \( z\tilde{n}_k^\wedge \) or \( \tilde{n}_k^\wedge \) to get \( \tilde{p}_n \). In the recurrence formulas for \( \tilde{p}_n \) (or \( \tilde{p}_n \)) the polynomials \( \tilde{p}_m \) (or \( \tilde{p}_m \)), respectively, of the previous block appear, but not those from older blocks.

There are other, theoretically equivalent formulations for the algorithm. Our version, which is truly sequential, suggests an implementation by recurrence formulas that have the matrix form

\[
(10.9) \quad \tilde{p}(z) = \tilde{p}'(z)\tilde{G}^\wedge, \quad z\tilde{p}'(z) = \tilde{p}(z)\tilde{G}^\wedge,
\]

as in (6.15), with \( \tilde{G}^\wedge \) unit upper triangular and \( \tilde{G}^\wedge \) unit upper Hessenberg. As in (7.1), elimination of \( \tilde{p}' \) or \( \tilde{p} \), respectively, leads to

\[
(10.10a) \quad z\tilde{p}(z) = \tilde{p}(z)\tilde{H}^\wedge, \quad \text{where} \quad \tilde{H}^\wedge := \tilde{G}^\wedge\tilde{G}^\wedge,
\]
and

\[(10.10b) \quad z \tilde{p}'(z) = \tilde{p}'(z) \tilde{H}^v, \quad \text{where} \quad \tilde{H}^v := \tilde{G}^v \tilde{G}^v.\]

Each of these equalities is of the same type as (9.9). But as we mentioned above, if we apply the algorithm of §9 to \(\Phi\), the block structure of \(\tilde{H}\) in (9.9) may turn out to be finer than the one of \(\tilde{H}^v\) in (10.10a), because here we have the four conditions (10.5) instead of the condition (9.7) determining the block sizes in (9.9). But we have also pointed out that there is no need to choose in §9 the minimum block sizes satisfying (9.7), and since each of the pairs (10.5a)–(10.5b) and (10.5c)–(10.5d) is equivalent to (9.7) (applied to the two different functionals \(\Phi\) and \(\Phi'\)), each of (10.10a) and (10.10b) is, indeed, identical to a case of (9.9). From §9 it follows in particular that \(\tilde{H}^v\) and \(\tilde{H}^v\) are block tridiagonal and block lower bidiagonal, respectively:

\[(10.11) \quad \tilde{G}^v := \begin{bmatrix}
\tilde{G}_0^v \\
\tilde{L}_0^v \\
\tilde{G}_1^v \\
\tilde{L}_1^v \\
\vdots \\
\tilde{L}_{j-1}^v \\
\tilde{G}_j^v
\end{bmatrix}, \quad \tilde{G}^v := \begin{bmatrix}
\tilde{G}_0^v \\
\tilde{F}_1^v \\
\tilde{G}_1^v \\
\tilde{F}_2^v \\
\vdots \\
\tilde{F}_j^v
\end{bmatrix}, \]

analogous to (6.16). From (10.9) and our description of the algorithm, it is clear that the blocks \(\tilde{L}_k^v\) have again just a 1 in their upper right corner, that the blocks \(\tilde{G}_k^v\) are itself unit upper Hessenberg, and that the blocks \(\tilde{G}_k^v\) are unit upper triangular. Further results on the structure of the blocks can be obtained by mixing the approach of §6 with the one of §9, but the details become somewhat tedious.

**Conclusions.** In the two parts of this paper we have reviewed the theory of formal orthogonal polynomials (FOPs) of the first and second kind (FOP1s and FOP2s, respectively), and we have derived old and new recurrences for recursively constructing certain sequences of such FOPs. From the beginning we dealt with the so-called non-normal or nongeneric case, where some of the FOPs are not regular and the associated Padé table has singular blocks. This led to "nongeneric" algorithms for constructing such FOPs. By translating these nongeneric algorithms into algorithms for sequences of vectors in Krylov space, we found nongeneric versions of the Lanczos-type algorithms BIORES, BIOMIN, and BIODIR. They can overcome most exact breakdowns of the previously known standard versions of these algorithms. (The exception is the so-called incurable breakdown.)

Finally, in §§9 and 10, we addressed the near-breakdown and presented algorithms for generating sequences of FOP1s in a stable way. By translating these "look-ahead" algorithms for constructing FOP1s into algorithms for sequences of vectors in Krylov space, we readily find stable look-ahead versions of BIORES, BIOMIN, and BIODIR. Since the FOP1s are denominators of Padé approximants, and since the numerators of Padé approximants satisfy the same recurrences, with different initial conditions, the algorithms given in this work can also be considered as algorithms for generating sequences of Padé approximants on either a diagonal or a generalized staircase of the Padé table. The look-ahead versions can be expected to be stable, in contrast to other algorithms [2], [4]–[6], [10] that have been proposed for nonnormal Padé tables and the related partial realization problem of systems theory. Applications to the fast solution of Hankel systems are also foreseeable.
Acknowledgment. The author would like to thank Susan Göldi for typing the first version of this manuscript and to Noël Nachtigal for making a number of linguistic corrections.

REFERENCES