Contemporary Mathematics

A Matrix Interpretation of the Extended Euclidean Algorithm

Martin H. Gutknecht

ABSTRACT. The extended Euclidean algorithm for polynomials and formal power series that is used for the recursive computation of Padé approximants can be viewed in various ways as a sequence of successive matrix multiplications that are applied to a Sylvester matrix with the original data. Here we present this result in a general version that includes the treatment of the Cabay–Meleshko look-ahead algorithm, which generalizes the extended Euclidean algorithm and yields a weakly stable (forward stable) method for computing Padé fractions if it is combined with an appropriate rule for choosing the look-ahead step length. Moreover, we choose for the matrix interpretation a particularly appealing form where also the product of all the matrices that are applied has a meaning: this product yields at the end four Toeplitz blocks with the coefficients of the polynomials (which belong to Padé forms) that are generated by the extended Euclidean algorithm in addition to those resulting from the ordinary Euclidean algorithm.

1. Introduction and notation

The Euclidean algorithm for polynomials or power series is closely related to Padé approximation. It is, in fact, up to minor modifications, one of the standard algorithms to compute Padé approximants along a diagonal of the Padé table and to generate the corresponding continued fraction (called a P-fraction), although the name Euclid comes only up in part of the Padé literature. An excellent, detailed account of these connections is given by Bultheel and Van Barel [**BV97**]; some of the main aspects are also treated by Gragg and Gutknecht [**GG94**] and others.

Since the Euclidean algorithm (in the form discussed here) concerns linear combinations of polynomials and power series, it is no surprise that one can interpret it in terms of matrices. Recall that the ring of formal power series is isomorphic to the ring of infinite lower triangular Toeplitz matrices. Such matrix interpretations of the Euclidean algorithm for polynomials are not very common, but some sketches can be found in the literature, at least for the (generic) case where in each step the degrees of the polynomials reduce only by 1; see, for example, Householder and Stewart [**HS69**] — we do not attempt to give a complete set of references here. A

1991 *Mathematics Subject Classification.* Primary 41A21; Secondary 30E05.

This paper was first presented at the conference "Orthogonal Polynomials in Numerical Analysis" in Oberwolfach, Germany, March 23–27, 1998.

single step of the algorithm is described as a matrix multiplication, and the whole algorithm takes the form of a matrix product consisting of a Sylvester matrix with the original data multiplied by a sequence of matrices that contain the quotients from the Euclidean algorithm. However, there is some freedom in arranging the data and in defining the factors of the product. We have succeeded to do this in a way so that the product of the matrix factors becomes a 4×4 block matrix with Toeplitz blocks that contain the coefficients of the Padé forms that are implicitly constructed in the process. In the terminology of Padé approximation, the Euclidean algorithm produces residuals of Padé approximants, while the Padé approximants itself, or rather the Padé forms (which are pairs of polynomials that are the numerators and denominators of the Padé approximants, respectively) are constructed additionally in the so-called extended Euclidean algorithm [**BV97**]. The matrix identity that we derive links the Sylvester matrix containing the given data not only to the recurrence coefficients (quotients) of the Euclidean algorithm but also with the six different sets of polynomials that are generated by the extended version. Moreover, the well-known symmetric LDU and inverse symmetric LDU decompositions of the Hankel moment matrix associated with the Padé problem can be read off the matrix identity.

Our motivation for this work came from section 9 of [**GH95a**], where, building up on earlier work by Delosme and Ipsen [**DI89**], Gutknecht and Hochbruck derived an analogous matrix representation of the look-ahead Levinson and Schur(-Bareiss) algorithms introduced in that paper. There too the resulting matrix identities can be interpreted in many ways. In particular, they are linked to the LDU and inverse LDU decomposition of the Toeplitz matrix of the given trigonometric moments.

Before we can start with the derivation and description of our matrix form of the Euclidean algorithm we present in sections 2–6 the prerequisites. This partly well-known material is extracted from [**GG94**] and suitably adapted to the situation and needs of this paper.

Notation. We let $\mathcal{L}$ be the set of formal Laurent series

$$h_0(\zeta) = \sum_{k=-\infty}^{\infty} \mu_k \zeta^k$$

with complex coefficients. The following subsets of $\mathcal{L}$ will play a role:

$$\begin{aligned}
\mathcal{L}_{l:m} &:\equiv \{h_0 \in \mathcal{L}\,;\ \mu_k = 0 \text{ if } k < l \text{ or } k > m\},\\
\mathcal{L}_m^* &:\equiv \mathcal{L}_{-\infty:m} = \{h_0 \in \mathcal{L}\,;\ \mu_k = 0 \text{ if } k > m\},\\
\mathcal{L}_0^* &:\equiv \mathcal{L}_{-\infty:0} = \text{formal power series in } \zeta^{-1},\\
\mathcal{P}_m &:\equiv \mathcal{L}_{0:m} = \text{polynomials in } \zeta \text{ of degree } m,\\
\mathcal{P} &:\equiv \bigcup \mathcal{P}_m = \text{polynomials in } \zeta.
\end{aligned}$$

For $h_0 \in \mathcal{L}_m^*$ or $h_0 \in \mathcal{P}$ we define the *exact degree* ∂h_0 by

$$\partial h_0 :\equiv \{n \in \mathbb{Z}\,;\ \mu_n \neq 0, \mu_k = 0 \text{ for } k > n\}.$$

We also use the notation $h_0(\zeta) = \mathcal{O}_-(\zeta^m)$ to indicate that $h_0 \in \mathcal{L}_m^*$, while $h_0(\zeta) \equiv \mathcal{O}_-(\zeta^m)$ means that $h_0(\zeta) \in \mathcal{L}_m^*$, but $h_0(\zeta) \notin \mathcal{L}_{m-1}^*$; in other words, it means the same as $\partial h_0 = m$. Moreover, if $h_0 \in \mathcal{L}_0^*$ we define $h_0(\infty) :\equiv \mu_0$.

Finally, $\mathcal{R}_{m,n}$ denotes the set of rational functions of type (m,n). Such a function r that can be represented as the quotient of two not necessarily relatively prime polynomials, $r = p/q$, where the numerator p has degree at most m, and the denominator q has degree at most n and is not the zero polynomial.

While $:\equiv$ is used to denote definitions, $:=$ indicates assignments that may be used in an algorithm. (Sometimes either one could be applied.)

2. Padé approximants of a pair of formal power series

Given a pair of formal power series in ζ^{-1},

$$\begin{aligned} f_0 &\in \mathcal{L}^*_{-1}, & f_0 &\neq 0, & f_0(\zeta) &\equiv: \textstyle\sum_{j=1}^{\infty} \phi_{j,0}\zeta^{-k}, \\ g_0 &\in \mathcal{L}^*_{0}, & g_0(\infty) &= 1, & g_0(\zeta) &\equiv: \textstyle\sum_{j=0}^{\infty} \gamma_{j,0}\zeta^{-k}, \end{aligned} \tag{2.1}$$

we consider, for fixed $n > 1$, pairs of pairs of polynomials,

$$(\grave{p}_n, \grave{q}_n) \in \mathcal{P}_{n-2} \times \mathcal{P}_{n-1}, \qquad (p_n, q_n) \in \mathcal{P}_{n-1} \times \mathcal{P}_n \quad \text{with } q_n \text{ monic}, \tag{2.2}$$

satisfying the following conditions: there are two new formal power series in ζ^{-1} of the form

$$f_n \in \mathcal{L}^*_{-n-1}, \qquad g_n \in \mathcal{L}^*_{-n} \quad \text{with } g_n(\zeta) = \zeta^{-n} + \mathcal{O}_-(\zeta^{-n-1}) \tag{2.3}$$

such that

$$\underbrace{g_0(\zeta)\grave{p}_n(\zeta)}_{\in \mathcal{L}^*_{n-2}} + \underbrace{f_0(\zeta)\grave{q}_n(\zeta)}_{\in \mathcal{L}^*_{n-2}} = \underbrace{g_n(\zeta)}_{\in \mathcal{L}^*_{-n}}, \tag{2.4a}$$

$$\underbrace{g_0(\zeta)p_n(\zeta)}_{\in \mathcal{L}^*_{n-1}} + \underbrace{f_0(\zeta)q_n(\zeta)}_{\in \mathcal{L}^*_{n-1}} = \underbrace{f_n(\zeta)}_{\in \mathcal{L}^*_{-n-1}}. \tag{2.4b}$$

This means that (p_n, q_n) and $(\grave{p}_n, \grave{q}_n)$ are *Padé forms* of the pair (g_0, f_0) at $\zeta = \infty$. The Padé form (p_n, q_n) consists of the numerator p_n and the denominator q_n of the (n, n) *Padé approximant* $r_{n,n} :\equiv p_n/q_n$, and, likewise, $(\grave{p}_n, \grave{q}_n)$ contains the (differently normalized, not necessarily mutually prime) numerator and denominator of a $(n-1, n-1)$ Padé approximant, if we consider the fractions as functions of ζ^{-1}. Therefore, these Padé approximants lie on the main diagonal of the *Padé table* of (g_0, f_0). We call the formal power series g_n and f_n in ζ^{-1} *residuals* of (p_n, q_n) and $(\grave{p}_n, \grave{q}_n)$, respectively. (In the literature, often the shifted series $\zeta^n g_n(\zeta)$ and $\zeta^n f_n(\zeta)$ are referred to as the residuals; but here the formulas will be simpler when we let the leading n or $n+1$ coefficients be zero.)

Note that (2.4a) implies that the $2n-2$ coefficients of the terms $\zeta^{-n+1}, \ldots, \zeta^{n-2}$ on the left-hand side must be set to 0, and by (2.3) the one of ζ^{-n} is 1. These $2n-1$ conditions are matched by the same number of degrees of freedom in the coefficients of $\grave{p}_n$ and $\grave{q}_n$. Likewise, in (2.4b) the $2n$ coefficients of the terms $\zeta^{-n}, \ldots, \zeta^{n-1}$ are zero, and the polynomials p_n and q_n have exactly $2n$ free coefficients since q_n must be monic.

Normally, Padé approximation is defined for a single formal power series. We can return to this situation by noting that due to $g_0(\infty) = 1$ the quotient $h_0(\zeta) :\equiv -f_0(\zeta)/g_0(\zeta) \in \mathcal{L}^*_{-1}$ is well defined. If we then redefine in the above formulas

$$g_0(\zeta) :\equiv 1, \qquad f_0(\zeta) :\equiv -h_0(\zeta), \tag{2.5}$$

it is easy to verify that (2.4a)–(2.4b) match with the usual Padé conditions. The series

$$h_0(\zeta) :\equiv \sum_{k=1}^{\infty} \frac{\mu_k}{\zeta^k} \quad \in \mathcal{L}^*_{-1}, \tag{2.6}$$

is the generating function of the *Markov parameters* or *moments* $\{\mu_k\}_{k=1}^{\infty}$ and can be viewed as the *symbol* of an infinite *Hankel matrix* M defined by $m_{k,l} :\equiv \mu_{k+l-1}$.

It is well known and easy to verify, see for example [**GG94**], that unique solutions $(\grave{p}_n, \grave{q}_n)$ and (p_n, q_n) to the conditions (2.2)–(2.4b) exist if and only if the nth principal leading submatrix

$$\mathsf{M}_n := \begin{bmatrix} \mu_1 & \mu_2 & \cdots & \mu_n \\ \mu_2 & & \iddots & \vdots \\ \vdots & \iddots & & \vdots \\ \mu_n & \cdots & \cdots & \mu_{2n-1} \end{bmatrix} \tag{2.7}$$

of M is nonsingular. In this case, we call q_n a *regular formal orthogonal polynomial* or *regular FOP*, the pair of denominators $(\grave{q}_n, q_n)$ a normalized nth *regular pair* and n a *regular index.* Heinig and Rost [**HR84**] refer to the pair as *fundamental solutions*, since it allows them to give an explicit formula for M^{-1}. Note that in the situation (2.5) it is trivial to determine the numerators $\grave{p}_n$ and p_n once the denominators $\grave{q}_n$ and q_n are known.

3. The recursive computation of Padé approximants by a generalized extended Euclidean algorithm

By using a 2×2 matrix with polynomial entries and row vectors with formal power series as entries we can write (2.4a)–(2.4b) as

$$\begin{bmatrix} g_0 & f_0 \end{bmatrix} \begin{bmatrix} \grave{p}_n & p_n \\ \grave{q}_n & q_n \end{bmatrix} = \begin{bmatrix} g_n & f_n \end{bmatrix}. \tag{3.1}$$

Note that if we extracted from the row vector on the right-hand side the common factor ζ^{-n}, the remainder would be of the same form as the row vector $[\ g_0 \quad f_0\ ]$. This indicates, but does not prove yet, that we can set up a recursive procedure. One still has to show that the corresponding 2×2 matrices fit together appropriately.

Generalizing earlier work of Gragg, Gustavson, Warner, and Yun [**GGWY82**] that was restricted to the generic case (where *all* leading principal submatrices of M are nonsingular) Cabay and Meleshko [**CM93**] established stable recurrences of this sort for computing well-conditioned Padé approximants along a diagonal of the Padé table. Here, well-conditioned means that the submatrix M_n, which could alternatively be used to determine these Padé approximants, is in a certain sense (left vague here) well conditioned. These Cabay-Meleshko recurrences are a special case of the following theorem, which could be used to compute *any* ordered sequence of regular pairs along the main diagonal of the table. The generalization to other diagonals is trivial. In the current notation this theorem is given and proved in [**GG94**].

THEOREM 3.1. *Let* $[\ \grave{q}_n \quad q_n\ ]$ *be a normalized* n*th regular pair for* $[\ g_0 \quad f_0\ ]$, *so that* (3.1) *holds with* (2.2)–(2.3)*. Likewise, let* $[\ \grave{a} \quad a\ ]$ *be a normalized* k*th* *regular pair for* $\zeta^n[\, g_n \quad f_n \,]$, *so that*

$$(\grave{b}, \grave{a}) \in \mathcal{P}_{k-2} \times \mathcal{P}_{k-1}, \qquad (b, a) \in \mathcal{P}_{k-1} \times \mathcal{P}_k \quad \textit{with } a \textit{ monic,} \tag{3.2}$$

and the residuals

$$[\, g_{n+k} \quad f_{n+k} \,] := [\, g_n \quad f_n \,] \begin{bmatrix} \grave{b} & b \\ \grave{a} & a \end{bmatrix} \tag{3.3}$$

satisfy

$$f_{n+k} \in \mathcal{L}^*_{-n-k-1}, \qquad g_{n+k} \in \mathcal{L}^*_{-n-k} \tag{3.4}$$

with

$$g_{n+k}(\zeta) = \zeta^{-n-k} + \mathcal{O}_-(\zeta^{-n-k-1}). \tag{3.5}$$

Then

$$\begin{bmatrix} \grave{p}_{n+k} & p_{n+k} \\ \grave{q}_{n+k} & q_{n+k} \end{bmatrix} := \begin{bmatrix} \grave{p}_n & p_n \\ \grave{q}_n & q_n \end{bmatrix} \begin{bmatrix} \grave{b} & b \\ \grave{a} & a \end{bmatrix} \tag{3.6}$$

yields a normalized $(n+k)$*th regular pair* $[\, \grave{q}_{n+k} \quad q_{n+k} \,]$ *for* $[\, g_0 \quad f_0 \,]$ *as well as the corresponding numerators* $[\, \grave{p}_{n+k} \quad p_{n+k} \,]$.

The corresponding pair of residuals is $[\, g_{n+k} \quad f_{n+k} \,]$. *Hence,* (3.3) *is the recurrence for the residuals.*

To obtain stable recurrences one needs a rule for identifying well-conditioned normalized regular pairs; see Cabay and Meleshko [**CM93**] and, for further results and comments on this and related matter, Beckermann [**Bec96**] and Gragg and Gutknecht [**GG94**]. The basis of this rule is Heinig's inversion formula for M_n, an improved variation of the well-known Gohberg-Semencul formula; see Heinig and Rost [**HR84**] and Gutknecht and Hochbruck [**GH95b**].

The condition (3.3) for the "recurrence coefficients", that is, the Padé forms $(\grave{b}, \grave{a})$ and (b, a), translates easily into a pair of linear systems for the coefficients of these Padé forms. Let

$$g_n(\zeta) \equiv: \sum_{j=0}^{\infty} \gamma_{j,n} \zeta^{-n-j}, \qquad f_n(\zeta) \equiv: \sum_{j=1}^{\infty} \phi_{j,n} \zeta^{-n-j}, \tag{3.7}$$

with $\gamma_{0,n} = 1$, and consider the $2k \times 2k$ *Sylvester matrix*

$$\mathsf{S}_k(g_n, f_n) :\equiv \left[\begin{array}{cccc|cccc} 0 & \cdots & 0 & 1 & 0 & \cdots & \cdots & 0 \\ \vdots & \iddots & 1 & \gamma_{1,n} & \vdots & & 0 & \phi_{1,n} \\ 0 & \iddots & \iddots & \vdots & \vdots & \iddots & \iddots & \vdots \\ 1 & \iddots & & \gamma_{k-1,n} & 0 & \iddots & & \phi_{k-1,n} \\ \gamma_{1,n} & & \iddots & \gamma_{k,n} & \phi_{1,n} & & \iddots & \phi_{k,n} \\ \vdots & \iddots & \iddots & \vdots & \vdots & \iddots & \iddots & \vdots \\ \gamma_{k-1,n} & \iddots & & \vdots & \phi_{k-1,n} & \iddots & & \vdots \\ \gamma_{k,n} & \cdots & \cdots & \gamma_{2k-1,n} & \phi_{k,n} & \cdots & \cdots & \phi_{2k-1,n} \end{array}\right]. \tag{3.8}$$

Moreover, set

$$\mathsf{e}_{2k,2k} :\equiv [\, 0 \quad \cdots \quad 0 \quad 1 \,]^\top \in \mathbb{C}^{2k}, \qquad \mathsf{f}_{2k,n} :\equiv -[\, \phi_{1,n} \cdots \phi_{2k,n} \,]^\top, \tag{3.9}$$

and let

$$\grave{\mathsf{b}}, \quad \grave{\mathsf{a}}, \quad \mathsf{b} \in \mathbb{C}^k, \qquad \mathsf{a} = \begin{bmatrix} \dot{\mathsf{a}} \\ 1 \end{bmatrix} \in \mathbb{C}^{k+1} \tag{3.10}$$

be the coefficient vectors of the polynomials $\grave{b}$, $\grave{a}$, b, and a, the first one being augmented by a zero component. Then (3.3) becomes, when the last component 1 of a is moved to the right-hand side,

$$\mathsf{S}_k(g_n, f_n) \begin{bmatrix} \grave{\mathsf{b}} & \mathsf{b} \\ \grave{\mathsf{a}} & \dot{\mathsf{a}} \end{bmatrix} = [\ \mathsf{e}_{2k,2k} \quad \mathsf{f}_{2k,n}\ ]. \tag{3.11}$$

When $k = 1$, which is the generic case, (3.11) reduces to

$$\begin{bmatrix} 1 & 0 \\ \gamma_{1,n} & \phi_{1,n} \end{bmatrix} \begin{bmatrix} \grave{\beta}_0 & \beta_0 \\ \grave{\alpha}_0 & \alpha_0 \end{bmatrix} = \begin{bmatrix} 0 & -\phi_{1,n} \\ 1 & -\phi_{2,n} \end{bmatrix}, \tag{3.12}$$

from which it follows that

$$\grave{\beta}_0 = 0, \qquad \grave{\alpha}_0 = \frac{1}{\phi_{1,n}}, \qquad \beta_0 = -\phi_{1,n}, \qquad \alpha_0 = \gamma_{1,n} - \frac{\phi_{2,n}}{\phi_{1,n}}. \tag{3.13}$$

4. J-fractions, P-fractions, and the extended Euclidean algorithm for power series

The recurrences of Theorem 3.1 are very general, and it is not so obvious what they have in common with the Euclidean algorithm. To explain this, we consider here two special cases, namely the one where $k = 1$ for all n, and thus all n are regular indices, and the one where $k > 1$ is allowed, but n and $n + k$ are two successive regular indices.

Therefore, for a moment assume that for all n, q_n is a regular FOP. Then the recursion (3.6) can be applied for all n with $k = 1$, where the coefficients are given by (3.13). Recall that this holds, if and only if $\phi_{1,n} \neq 0$ for all n. Then

$$(\grave{p}_{n+1}, \grave{q}_{n+1}) = (p_n, q_n)/\phi_{1,n}, \tag{4.1}$$

and hence

$$\begin{bmatrix} p_n & p_{n+1} \\ q_n & q_{n+1} \end{bmatrix} := \begin{bmatrix} p_{n-1} & p_n \\ q_{n-1} & q_n \end{bmatrix} \begin{bmatrix} 0 & -\beta_{n+1} \\ 1 & \zeta - \alpha_{n+1} \end{bmatrix}, \tag{4.2}$$

where

$$\beta_{n+1} := \frac{\phi_{1,n}}{\phi_{1,n-1}}, \qquad \alpha_{n+1} := \frac{\phi_{2,n}}{\phi_{1,n}} - \gamma_{1,n}, \tag{4.3}$$

with $\phi_{1,-1} := -1$, $\phi_{k,0} := \phi_k$, $\gamma_{k,0} := \gamma_k$.

Eliminating the trivial part of (4.2) we find the well-known generic diagonal recurrences for Padé forms and FOPs,

$$\begin{aligned} p_n(\zeta) &:= (\zeta - \alpha_n)p_{n-1}(\zeta) - \beta_n p_{n-2}(\zeta), \\ q_n(\zeta) &:= (\zeta - \alpha_n)q_{n-1}(\zeta) - \beta_n q_{n-2}(\zeta), \end{aligned} \tag{4.4}$$

with initial values

$$(p_{-1}(\zeta), q_{-1}(\zeta)) :\equiv (1, 0), \qquad (p_0(\zeta), q_0(\zeta)) :\equiv (0, 1). \tag{4.5}$$

They reveal that p_n and q_n are the nth numerator and denominator of the *Jacobi fraction (J-fraction)* of $h_0 = -f_0/g_0$,

$$h_0(\zeta) = -\frac{f_0(\zeta)}{g_0(\zeta)} = \frac{\beta_1|}{|\zeta - \alpha_1} - \frac{\beta_2|}{|\zeta - \alpha_2} - \cdots, \tag{4.6}$$

that is,

$$\frac{p_n(\zeta)}{q_n(\zeta)} = \frac{\beta_1|}{|\zeta - \alpha_1} - \cdots - \frac{\beta_n|}{|\zeta - \alpha_n}. \tag{4.7}$$

This J-fraction does not exist if $\phi_{1,n} = 0$ for some n. It may have very large $|\alpha_{n+1}|$ and $|\beta_{n+2}|$ if $|\phi_{1,n}|$ is small.

In exact arithmetic the case $\phi_{1,n} = 0$ can be treated as follows: if

$$\phi_{0,n} = \cdots = \phi_{k-1,n} = 0, \qquad \phi_{k,n} \neq 0, \tag{4.8}$$

then $\mathsf{S}_1, \ldots, \mathsf{S}_{k-1}$ are singular, but S_k is nonsingular. The left system in (3.11) has then the solution

$$\grave{b}(\zeta) \equiv 0, \qquad \grave{a}(\zeta) \equiv \frac{1}{\phi_{k,n}}. \tag{4.9}$$

Consequently,

$$(\grave{p}_{n+k}, \grave{q}_{n+k}) = \frac{1}{\phi_{k,n}}(p_n, q_n), \qquad g_{n+k} = \frac{\zeta^k f_n}{\phi_{k,n}}. \tag{4.10}$$

In this way, we can proceed from any regular Padé form (p_n, q_n) and its upper-left neighbor $(\grave{p}_n, \grave{q}_n)$ to the next one, (p_{n+k}, q_{n+k}), and its upper-left neighbor $(\grave{p}_{n+k}, \grave{q}_{n+k})$, where k is determined by (4.8).

Let $\{n_j\}_{j=0}^{J}$ $(J \leq \infty)$ be the sequence of *all* regular indices, starting with $n_0 = 0$, and let

$$k_j :\equiv n_{j+1} - n_j \qquad (j \geq 1). \tag{4.11}$$

Additionally, we set $n_{-1} := -1$, $k_0 := 1$. Then, $\phi_{k_j, n_j} \neq 0$ $(\forall j)$. In analogy to (4.2), recurrence (3.6) becomes

$$\begin{bmatrix} p_{n_j} & p_{n_{j+1}} \\ q_{n_j} & q_{n_{j+1}} \end{bmatrix} := \begin{bmatrix} p_{n_{j-1}} & p_{n_j} \\ q_{n_{j-1}} & q_{n_j} \end{bmatrix} \begin{bmatrix} 0 & -\beta_{j+1} \\ 1 & a_{j+1}(\zeta) \end{bmatrix} \qquad (j \geq 0), \tag{4.12}$$

with

$$\beta_{j+1} := \frac{\phi_{k_j, n_j}}{\phi_{k_{j-1}, n_{j-1}}}, \qquad a_{j+1}(\zeta) := q_{k_j}^{(n_j)}(\zeta) \tag{4.13}$$

and initial conditions (4.5).

(4.12) is equivalent to the three-term recurrence of Magnus [**Mag62**] and Struble [**Str63**]:

$$\begin{aligned} p_{n_j}(\zeta) &:= a_j(\zeta) p_{n_{j-1}}(\zeta) - \beta_j\, p_{n_{j-2}}(\zeta), \\ q_{n_j}(\zeta) &:= a_j(\zeta) q_{n_{j-1}}(\zeta) - \beta_j\, q_{n_{j-2}}(\zeta), \end{aligned} \qquad (j \geq 1) \tag{4.14}$$

with the initial conditions (4.5); see also [**Gut92**] and further references given there.

They correspond to the Magnus *P-fraction*

$$h_0(\zeta) = -\frac{f_0(\zeta)}{g_0(\zeta)} = \frac{\beta_1|}{|a_1(\zeta)} - \frac{\beta_2|}{|a_2(\zeta)} - \cdots. \tag{4.15}$$

whose partial numerators and denominators are p_{n_j} and q_{n_j}, respectively.

In view of (4.10), the recurrence (3.3) for the residuals becomes

$$f_{n_j}(\zeta) = a_j(\zeta) f_{n_{j-1}}(\zeta) - \beta_j f_{n_{j-2}}(\zeta) \quad (j \geq 1) \tag{4.16}$$

with initial conditions

$$f_{n_{-1}}(\zeta) := f_{-1}(\zeta) := -g_0(\zeta), \qquad f_{n_0}(\zeta) := f_0(\zeta). \tag{4.17}$$

Recall that

$$\partial f_{n_{j-1}} = -n_j > \partial f_{n_j} = -n_{j+1} > \partial f_{n_{j+1}} \tag{4.18}$$

and

$$k_{j+1} = \partial a_{j+1} = \partial f_{n_{j-1}} - \partial f_{n_j}. \tag{4.19}$$

β_j is just used to make a_{j+1} monic.

This reveals that the P-fraction can be constructed by a simple generalization of the *Euclidean algorithm* for polynomials. In fact, the Euclidean algorithm for two polynomials f_{-1} and f_0 with $\partial f_{-1} > \partial f_0$, generates a finite sequence of polynomials f_j of decreasing degree according to

$$f_j(\zeta) := a_j(\zeta) f_{j-1}(\zeta) - \beta_j f_{j-2}(\zeta) \qquad (j \geq 1), \tag{4.20}$$

with $0 \neq \beta_j \in \mathbb{C}$ arbitrary and $a_j \in \mathcal{P}$. Note that this is the same formula as (4.16) except for a slightly different indexing scheme.

(4.14) and (4.16) define the *extended Euclidean algorithm for power series in* ζ^{-1}. The word "extended" refers to the fact that not only the residuals f_{n_j} are updated, which are the tails of the P-fraction (4.15), but also the corresponding partial numerators and denominators p_{n_j} and q_{n_j}; see Bultheel and Van Barel [**BV97**] for a detailed discussion of the extended Euclidean algorithm.

In the generic case, where $a_j(\zeta) = \zeta - \alpha_j$, the Euclidean algorithm for power series is also referred to as *Chebyshev algorithm*. An alternative generalization of the latter to the case where $k > 1$ for some n was proposed by Golub and Gutknecht [**GG90**].

Dividing (4.16) by $f_{n_{j-1}}(\zeta)$ shows that a_j is the *polynomial* or *principal part* of the formal Laurent series at ∞ of $\beta_j f_{n_{j-2}}/f_{n_{j-1}}$. This fact gave rise to the name P-fraction.

5. Inner polynomials

Let us return to the Padé approximation problems for (g_0, f_0) or, equivalently, for $h_0 :\equiv -f_0/g_0$ that we considered in section 2. If n is not regular, that is, if M_n of (2.7) is singular, there is no or not a unique polynomial $q_n \in \mathcal{P}_n$ satisfying the corresponding Padé condition (2.4b) with some $p_n \in \mathcal{P}_{n-1}$ and some $f_n \in \mathcal{L}^*_{-n-1}$.

However, there are still solutions to a *relaxed Padé condition*: if n is regular and $k > 0$, we can always find a Padé form

$$(p_{n+k}, q_{n+k}) \in \mathcal{P}_{n+k-1} \times \mathcal{P}_{n+k} \quad \text{with } q_{n+k} \text{ monic} \tag{5.1}$$

such that the corresponding residual f_{n+k} satisfies

$$f_{n+k} :\equiv g_0 p_{n+k} + f_0 q_{n+k} \in \mathcal{L}^*_{-n-1}. \tag{5.2}$$

The quotient $r_{n+k,n+k} :\equiv p_{n+k}/q_{n+k}$ has been called an *underdetermined Padé approximant* in [**Gut93**]. In conjunction with the algorithm of Theorem 3.1, which allows us to compute a sequence of regular indices and pairs, we will refer to those

q_{n+k} that are not part of this sequence but still satisfy (5.1)–(5.2) as *inner polynomials.*

The following theorem, which has been proved in [**GG94**], establishes not only the well-known existence of such inner polynomials and underdetermined Padé approximants, but also how to compute them from an earlier regular pair.

THEOREM 5.1. *Let* $[\ \grave{q}_n \quad q_n\ ]$ *be a normalized* n*th regular pair for* $[\ g_0 \quad f_0\ ]$, *and let* k *be a positive integer. Consider any pair* (p_{n+k}, q_{n+k}) *constructed according to*

$$\begin{bmatrix} p_{n+k} \\ q_{n+k} \end{bmatrix} := \begin{bmatrix} \grave{p}_n & p_n \\ \grave{q}_n & q_n \end{bmatrix} \begin{bmatrix} b \\ a \end{bmatrix} \tag{5.3}$$

from a pair

$$(b, a) \in \mathcal{P}_{k-1} \times \mathcal{P}_k \quad \textit{with a monic} \tag{5.4}$$

that satisfies

$$f_{n+k} :\equiv g_n b + f_n a \in \mathcal{L}^*_{-n-1}. \tag{5.5}$$

Then (5.1) *and* (5.2) *hold.*

Condition (5.5) *is fulfilled when*

$$b(\zeta) := -\frac{\zeta^k f_n(\zeta)}{g_n(\zeta)} + \mathcal{O}_-(\zeta^{-1}), \qquad a(\zeta) := \zeta^k\,; \tag{5.6}$$

that is,

$$b(\zeta) := \mu_1^{(n)} \zeta^{k-1} + \mu_2^{(n)} \zeta^{k-2} + \cdots + \mu_k^{(n)} \tag{5.7}$$

if

$$h_n(\zeta) :\equiv -\frac{f_n(\zeta)}{g_n(\zeta)} \equiv: \mu_1^{(n)} \zeta^{-1} + \mu_2^{(n)} \zeta^{-2} + \cdots. \tag{5.8}$$

The computation of $b(\zeta)$ in (5.6) requires only the solution of a $k \times k$ triangular Toeplitz system, and as k is growing, these systems are nested, so that the solution of the last one contains the solution for all.

6. Matrix representations of Padé conditions and orthogonality

The most relevant aspect of the Padé condition (2.4b) for the Padé form (p_n, q_n) is that the terms in $\zeta^{-n}, \ldots, \zeta^{n-1}$ on the left-hand side cancel. If $g_0(\zeta) \equiv 1$ as in (2.5), $p_n \in \mathcal{P}_{n-1}$ is chosen so that the positive powers cancel, while q_n must be chosen so that the $n+1$ terms in $\zeta^{-n}, \ldots, \zeta^{-1}, \zeta^0$ cancel. If we choose some fixed $N > n$ and assume that all $n < N$ are regular indices, this condition can be formulated in terms of the $N \times N$ Hankel moment matrix M_N for $h_0 :\equiv -f_0/g_0$ as

$$\mathsf{M}_N\,\mathsf{R}_N = \mathsf{F}_N, \tag{6.1}$$

where

$$\mathsf{R}_N :\equiv \begin{bmatrix} 1 & \rho_{0,1} & \cdots & \cdots & \rho_{0,N-1} \\ & 1 & \cdots & \cdots & \rho_{1,N-1} \\ & & \ddots & & \vdots \\ & & & \ddots & \rho_{N-2,N-1} \\ & & & & 1 \end{bmatrix}, \tag{6.2}$$

is a unit upper triangular matrix containing in its columns the coefficients of the polynomials $q_0, q_1, \ldots, q_{N-1}$, while F_N is a lower triangular matrix

$$\mathsf{F}_N :\equiv \begin{bmatrix} \phi_{1,0} & & & \\ \phi_{2,0} & \phi_{1,1} & & \\ \vdots & \vdots & \ddots & \\ \phi_{N,0} & \phi_{N-1,1} & \cdots & \phi_{1,N-1} \end{bmatrix} \tag{6.3}$$

whose columns contain the coefficients, down to those of the term in ζ^{-N}, of $f_0, f_1, \ldots, f_{N-1}$.

If we multiply (6.1) from the left by $\mathsf{R}_N^\top$, we obtain

$$\mathsf{R}_N^\top \mathsf{M}_N \mathsf{R}_N = \mathsf{R}_N^\top \mathsf{F}_N,$$

where the left-hand side is clearly symmetric and the right-hand side is clearly lower triangular, hence both sides must equal a diagonal matrix D_N: on the left,

$$\mathsf{R}_N^\top \mathsf{M}_N \mathsf{R}_N = \mathsf{D}_N :\equiv \mathrm{diag}\,[\delta_0 \cdots \delta_{N-1}], \tag{6.4}$$

where

$$\delta_n :\equiv \mathsf{r}_n^\top \mathsf{M}_{n+1} \mathsf{r}_n, \qquad \mathsf{r}_n :\equiv [\ \rho_{0,n} \quad \ldots \rho_{n,n}\ ]^\top. \tag{6.5}$$

Eq. (6.4) is often referred to as an *inverse symmetric LDU decomposition* of M_N. On the right, $\mathsf{R}_N^\top \mathsf{F}_N = \mathsf{D}_N$, or, $(\mathsf{R}_N^\top)^{-1} = \mathsf{F}_N \mathsf{D}_N^{-1}$, and, therefore, substituting R_N in (6.4) yields a *symmetric LDU decomposition* of M_N:

$$\mathsf{M}_N = \mathsf{F}_N \mathsf{D}_N^{-1} \mathsf{F}_N^\top. \tag{6.6}$$

The decompositions (6.4) and (6.6) are well known and are fully analogous to decomposition related to the Levinson and Schur algorithms for Toeplitz matrices, except that the latter are not necessarily symmetric; see, for example, [**GH95a**] for the decompositions of Toeplitz matrices. In particular, (6.4) is just a matrix representation of the mutual formal orthogonality of the FOPs.

If q_n is not regular for some n, the decompositions (6.4) and (6.6) do not exist in this form. But by applying Theorem 5.1 we can redefine all those q_n and f_n so that (6.4) and (6.6) hold as a *inverse symmetric block LDU decomposition* and a *symmetric block LDU decomposition*, respectively. Then, D_N is *block diagonal*, R_N is still *unit upper triangular*, and F_n is *lower block triangular*.

We may make the blocks bigger than theoretically needed and only keep those regular q_n that are well-conditioned. Then the columns of R_n and F_n with well-conditioned regular index n still contain the coefficients of the FOPs and the corresponding residuals. They are the first columns of a block, while the other columns contain the coefficients of the inner polynomials and the corresponding residuals, respectively.

If we want to aim at an analogous treatment of the general Padé conditions (2.4a)–(2.4b) for pairs of Padé forms $(\grave{p}_n, \grave{q}_n)$, (p_n, q_n) that interpolate pairs of power series (g_0, f_0), then we need to include the powers $\zeta^0, \ldots, \zeta^{n-1}$. Using our notation (3.8) for Sylvester matrices we can write (with e_1 and e_N the first and the last standard unit vector of $\mathbb{C}^N$)

$$\mathsf{S}_N(g_0, f_0) \left[\begin{array}{c|c} \grave{\mathsf{P}}_N & \mathsf{P}_N \\ \hline \grave{\mathsf{R}}_N & \mathsf{R}_N \end{array}\right] = \left[\begin{array}{c|c} \mathsf{e}_N^\top \mathsf{e}_1 & \mathsf{O} \\ \hline \mathsf{G}_N & \mathsf{F}_N \end{array}\right], \tag{6.7}$$

where R_N is defined in (6.2) and $\grave{\mathsf{R}}_N$, P_N, and $\grave{\mathsf{P}}_N$ are analogous upper triangular matrices with the coefficients of the polynomials $\grave{q}_n$, p_n, and $\grave{p}_n$ $(0 \le n < N)$, respectively, while G_N is a $N \times N$ lower Hessenberg matrix that contains in its columns the coefficients (down to those of ζ_{-N}) of the residuals g_n (except that $\gamma_{0,0} = 1$ is missing at the top of the first column), and F_N is a $N \times N$ lower triangular matrix with the coefficients of the residuals f_n as before.

However, compared to these residuals, those whose coefficients are in F_N of (6.3) are devided by $-g_0$, because this is how the general case of (2.4a)–(2.4b) is related to the special case of (2.5). Therefore, to return from (6.7) to the special case (6.1), we multiply (6.7) from the left by a $2N \times 2N$ lower triangular Toeplitz matrix that contains the first $2N$ coefficients of the formal power series of $-1/g_0$. Then, after redefining F_N, the $(2,2)$ block of the resulting block matrix identity yields exactly (6.1).

7. Matrix form of the extended Euclidean algorithm

Now, we can finally aim at the intended reformulation of the extended Euclidean algorithm as a matrix product applied to the Sylvester matrix with the given data. First, we discuss a single step and return to (3.3),

$$\begin{bmatrix} g_n & f_n \end{bmatrix} \begin{bmatrix} \grave{b} & b \\ \grave{a} & a \end{bmatrix} = \begin{bmatrix} g_{n+k} & f_{n+k} \end{bmatrix},$$

and its matrix form (3.11), rewritten with two additional columns and two additional rows as

$$\mathsf{S}_{k+1}(g_n, f_n) \left[\begin{array}{c|c} \grave{\mathsf{b}} & \mathsf{b} \\ 0 & 0 \\ \hline \grave{\mathsf{a}} & \dot{\mathsf{a}} \\ 0 & 1 \end{array}\right] = \begin{bmatrix} \mathsf{e}_{2k+1,2k+1} & \mathsf{o}_{2k+1} \\ \gamma_{1,n+k} & \phi_{1,n+k} \end{bmatrix} \tag{7.1}$$

(where o_{2k+1} is the zero vector in $\mathbb{C}^{2k+1}$). This corresponds to the terms from ζ^{-n+k} downto ζ^{-n-k-1} of (3.3). We can take into account as many leading terms as we want by adding additional rows to the Sylvester matrix and to the right-hand side. In particular, if $g_n \in \mathcal{L}^*_{-n}$ and $f_n \in \mathcal{L}^*_{-n-1}$ are polynomials in ζ^{-1} of degree at most $2N$, we can process all their coefficients if the Sylvester matrix is of size $(2N - n + 2) \times (2k + 2)$.

Starting from a pair

$$f_0 \in \mathcal{L}_{-2N:-1}, \qquad g_0 \in \mathcal{L}_{-2N:0}$$

with $f_0 \ne 0$ and $g_0(\infty) :\equiv \gamma_{0,0} = 1$, we will apply (3.3) recursively until we reach some $f_n \in \mathcal{L}^*_{-N-1}$. This will happen for some regular $n \equiv: n_J \le N$. We let $\{n_j\}_{j=0}^J$ be the corresponding sequence of regular indices, starting with $n_0 = 0$, and set

$$k_j := n_{j+1} - n_j \qquad (j \ge 1).$$

Then, $\phi_{k_j,n_j} \ne 0$ $(\forall j)$. Additionally, we set $n_{-1} := -1$, $k_0 := 1$. The sequence $\{n_j\}_{j=0}^J$ need not contain *all* regular indices, but only the chosen well-conditioned ones.

Adding further terms of order $\mathcal{O}_-(-2N-1)$ to f_0 and g_0 will not change the resulting pair (p_{n_J}, q_{n_J}).

If, in the step starting at $n = n_j$, we write $\grave{a} \equiv: \grave{a}_j$, $\grave{b} \equiv: \grave{b}_j$, $\grave{\mathsf{a}} \equiv: \grave{\mathsf{a}}_j$, $\grave{\mathsf{b}} \equiv: \grave{\mathsf{b}}_j$, etc., then (3.3) and (3.6) yield

$$\begin{bmatrix} g_{n_{j+1}} & f_{n_{j+1}} \end{bmatrix} := \begin{bmatrix} g_{n_j} & f_{n_j} \end{bmatrix} \begin{bmatrix} \grave{b}_j & b_j \\ \grave{a}_j & a_j \end{bmatrix} \qquad (j \geq 0), \tag{7.2}$$

and

$$\begin{bmatrix} \grave{p}_{n_{j+1}} & p_{n_{j+1}} \\ \grave{q}_{n_{j+1}} & q_{n_{j+1}} \end{bmatrix} := \begin{bmatrix} \grave{p}_{n_j} & p_{n_j} \\ \grave{q}_{n_j} & q_{n_j} \end{bmatrix} \begin{bmatrix} \grave{\mathsf{b}}_j & \mathsf{b}_j \\ \grave{\mathsf{a}}_j & \mathsf{a}_j \end{bmatrix} \qquad (j \geq 0). \tag{7.3}$$

Consequently,

$$\begin{bmatrix} g_{n_J} & f_{n_J} \end{bmatrix} := \begin{bmatrix} g_{n_0} & f_{n_0} \end{bmatrix} \prod_{j=0}^{J-1} \begin{bmatrix} \grave{b}_j & b_j \\ \grave{a}_j & a_j \end{bmatrix} \tag{7.4}$$

and

$$\begin{bmatrix} \grave{p}_{n_J} & p_{n_J} \\ \grave{q}_{n_J} & q_{n_J} \end{bmatrix} := \begin{bmatrix} \grave{p}_{n_0} & p_{n_0} \\ \grave{q}_{n_0} & q_{n_0} \end{bmatrix} \prod_{j=1}^{J-1} \begin{bmatrix} \grave{\mathsf{b}}_j & \mathsf{b}_j \\ \grave{\mathsf{a}}_j & \mathsf{a}_j \end{bmatrix} \tag{7.5}$$

with $f_{n_J} \in \mathcal{L}^*_{-N-1}$. Obviously, we could combine the two equations (7.4) and (7.5) into one.

Let us again look first at the *generic case*, where $k = 1$ for all n, and thus $n_j = n$. Here, (7.1) reduces to

$$\left[\begin{array}{cc|cc} 0 & 1 & 0 & 0 \\ 1 & \gamma_{1,n} & 0 & \phi_{1,n} \\ \gamma_{1,n} & \gamma_{2,n} & \phi_{1,n} & \phi_{2,n} \\ \gamma_{2,n} & \gamma_{3,n} & \phi_{2,n} & \phi_{3,n} \end{array}\right] \left[\begin{array}{c|c} 0 & \beta_0^{(j)} \\ 0 & 0 \\ \hline \grave{\alpha}_0^{(j)} & \alpha_0^{(j)} \\ 0 & 1 \end{array}\right] = \left[\begin{array}{c|c} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ \gamma_{1,n+1} & \phi_{1,n+1} \end{array}\right]. \tag{7.6}$$

This implies trivially that

$$\left[\begin{array}{cc|cc} 0 & 1 & 0 & 0 \\ 1 & \gamma_{1,n} & 0 & \phi_{1,n} \\ \gamma_{1,n} & \gamma_{2,n} & \phi_{1,n} & \phi_{2,n} \\ \gamma_{2,n} & \gamma_{3,n} & \phi_{2,n} & \phi_{3,n} \end{array}\right] \left[\begin{array}{cc|cc} 1 & 0 & 0 & \beta_0^{(j)} \\ 0 & 0 & 0 & 0 \\ \hline 0 & \grave{\alpha}_0^{(j)} & 1 & \alpha_0^{(j)} \\ 0 & 0 & 0 & 1 \end{array}\right] = \left[\begin{array}{cc|cc} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ \gamma_{1,n} & 1 & \phi_{1,n} & 0 \\ \gamma_{2,n} & \gamma_{1,n+1} & \phi_{2,n} & \phi_{1,n+1} \end{array}\right].$$

Note that the two additional columns in the second matrix allow us to "store" the coefficients of the residuals in the resulting matrix on the right-hand side.

To obtain a representation that can accommodate all N steps we need to "imbed" this relationship into $2N \times 2N$ matrices. For example, in the first step, where $n = 0$, we get

$$\mathsf{S}_N(g_0, f_0)\mathsf{C}_1 = \mathsf{V}_1$$

with

$$\mathsf{S}_N(g_0,f_0) :\equiv \left[\begin{array}{cccc|cccc} 0 & \cdots & 0 & 1 & 0 & \cdots & \cdots & 0 \\ \vdots & \iddots & 1 & \gamma_{1,0} & \vdots & & 0 & \phi_{1,0} \\ 0 & \iddots & \iddots & \vdots & \vdots & \iddots & \iddots & \vdots \\ 1 & \iddots & & \vdots & 0 & \iddots & & \vdots \\ \gamma_{1,0} & & & \gamma_{N,0} & \phi_{1,0} & & & \phi_{N,0} \\ \vdots & & \iddots & \vdots & \vdots & & \iddots & \vdots \\ \vdots & \iddots & & \vdots & \vdots & \iddots & & \vdots \\ \gamma_{N,0} & \cdots & \cdots & \gamma_{2N-1,0} & \phi_{N,0} & \cdots & \cdots & \phi_{2N-1,0} \end{array}\right],$$

$$\mathsf{C}_1 :\equiv \left[\begin{array}{cccc|cccc} 1 & 0 & & & 0 & \beta_0^{(0)} & & \\ & 0 & \ddots & & & 0 & \ddots & \\ & & \ddots & 0 & & & \ddots & \beta_0^{(0)} \\ & & & 0 & & & & 0 \\ \hline 0 & \grave{\alpha}_0^{(0)} & & & 1 & \alpha_0^{(0)} & & \\ & 0 & \ddots & & & 1 & \ddots & \\ & & \ddots & \grave{\alpha}_0^{(0)} & & & \ddots & \alpha_0^{(0)} \\ & & & 0 & & & & 1 \end{array}\right],$$

$$\mathsf{V}_1 :\equiv \left[\begin{array}{c|ccc|c|ccc} 0 & 0 & \ldots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \iddots & 0 & \vdots & \vdots & \iddots & 0 \\ 0 & 0 & \iddots & 1 & \vdots & 0 & \iddots & 0 \\ 1 & 0 & \iddots & \gamma_{1,1} & 0 & 0 & \iddots & \phi_{1,1} \\ \gamma_{1,0} & 1 & \iddots & \vdots & \phi_{1,0} & 0 & \iddots & \vdots \\ \vdots & \gamma_{1,1} & & \vdots & \vdots & \phi_{1,1} & & \vdots \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \gamma_{N,0} & \gamma_{N-1,1} & \cdots & \gamma_{2N-3,1} & \phi_{N,0} & \phi_{N-1,1} & \cdots & \phi_{2N-3,1} \end{array}\right].$$

In the second step,

$$\mathsf{V}_1\mathsf{C}_2 = \mathsf{V}_2,$$

C_2 will have two 2×2 unit matrices where C_1 had two 1's, and therefore the columns with indices $1, 2, N+1, N+2$ of V_1 will remain untouched in the transition to V_2:

$$\mathsf{V}_2 = \left[\begin{array}{cc|ccc|cc|ccc}
0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \iddots & \vdots & \vdots & \vdots & \vdots & \iddots & \vdots \\
0 & 0 & 0 & & 0 & \vdots & 0 & 0 & & \vdots \\
1 & 0 & \vdots & \iddots & 1 & 0 & \vdots & \vdots & & 0 \\
\gamma_{1,0} & 1 & 0 & \iddots & \gamma_{1,2} & \phi_{1,0} & 0 & \vdots & \iddots & \phi_{1,2} \\
\vdots & \gamma_{1,1} & 1 & \iddots & \vdots & \vdots & \phi_{1,1} & 0 & \iddots & \vdots \\
\vdots & \vdots & \gamma_{1,2} & & \vdots & \vdots & \vdots & \phi_{1,2} & & \vdots \\
\vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots \\
\gamma_{N,0} & \gamma_{N-1,1} & \gamma_{N-2,2} & \cdots & \gamma_{2N-5,2} & \phi_{N,0} & \phi_{N-1,1} & \phi_{N-2,2} & \cdots & \phi_{2N-5,2}
\end{array}\right],$$

while all the Toeplitz blocks in C_2 and V_2 are by one column and one or two rows smaller than the corresponding ones in C_1 and V_1.

By now it should be clear, how, in the generic case, the matrices V_n change from step to step until at the end,

$$\mathsf{V}_N = \left[\begin{array}{c|c} \mathsf{e}_N^\top \mathsf{e}_1 & \mathsf{O} \\ \hline \mathsf{G}_N & \mathsf{F}_N \end{array}\right]$$

is the matrix on the right-hand side of (6.7).

Now, we turn to the *look-ahead case* where $n = n_j$ and $k = k_j > 1$, either due to the situation (4.8) or because of a general look-ahead step that also leaves out ill-conditioned regular pairs. We append $N - n_j - k_j - 1 = N - n_{j+1} - 1$ rows at the bottom of the $(2k_j + 2) \times (2k_j + 2)$ matrix $\mathsf{S}_{k_j+1}(g_{n_j}, f_{n_j})$ to get the $(N - n_j + k_j + 1) \times (2k_j + 2)$ matrix

$$\underline{\mathsf{S}}_{k_j+1}(g_{n_j}, f_{n_j}) :\equiv$$
$$\left[\begin{array}{cccc|cccc}
0 & \cdots & 0 & 1 & 0 & \cdots & \cdots & 0 \\
\vdots & \iddots & 1 & \gamma_{1,n_j} & \vdots & & 0 & \phi_{1,n_j} \\
0 & \iddots & \iddots & \vdots & \vdots & \iddots & \iddots & \vdots \\
1 & \iddots & & \vdots & 0 & \iddots & & \vdots \\
\gamma_{1,n_j} & & & \gamma_{k+1,n_j} & \phi_{1,n_j} & & & \phi_{k+1,n_j} \\
\vdots & & \iddots & \gamma_{k+2,n_j} & \vdots & & \iddots & \phi_{k+2,n_j} \\
\vdots & \iddots & \iddots & \vdots & \vdots & \iddots & \iddots & \vdots \\
\gamma_{k+1,n_j} & \iddots & & \gamma_{2k+1,n_j} & \phi_{k+1,n_j} & \iddots & & \phi_{2k+1,n_j} \\
\hline
\gamma_{k+2,n_j} & \cdots & \cdots & \gamma_{2k+2,n_j} & \phi_{k+2,n_j} & \cdots & \cdots & \phi_{2k+2,n_j} \\
\vdots & & & \vdots & \vdots & & & \vdots \\
\gamma_{N-n_j,n_j} & \cdots & \cdots & \gamma_{N-n_{j+1},n_j} & \phi_{N-n_j,n_j} & \cdots & \cdots & \phi_{N-n_{j+1},n_j}
\end{array}\right]$$

(for simplicity, we write k instead of k_j in the indices). Now, (7.1) reads

$$\underline{\mathsf{S}}_{k_j+1}(g_{n_j},f_{n_j})\left[\begin{array}{c|c} 0 & \beta_0^{(j)} \\ \grave{\beta}_1^{(j)} & \beta_1^{(j)} \\ \vdots & \vdots \\ \grave{\beta}_{k-1}^{(j)} & \beta_{k-1}^{(j)} \\ 0 & 0 \\ \hline \grave{\alpha}_0^{(j)} & \alpha_0^{(j)} \\ \grave{\alpha}_1^{(j)} & \alpha_1^{(j)} \\ \vdots & \vdots \\ \grave{\alpha}_{k-1}^{(j)} & \alpha_{k-1}^{(j)} \\ 0 & 1 \end{array}\right] = \left[\begin{array}{c|c} 0 & 0 \\ 0 & 0 \\ \vdots & \vdots \\ \vdots & \vdots \\ 0 & 0 \\ 1 & 0 \\ \gamma_{1,n_{j+1}} & \phi_{1,n_{j+1}} \\ \hline \gamma_{2,n_{j+1}} & \phi_{2,n_{j+1}} \\ \vdots & \vdots \\ \gamma_{N-n_j-k,n_{j+1}} & \phi_{N-n_j-k,n_{j+1}} \end{array}\right]. \tag{7.7}$$

This system contains $N - n_{j+1} + 1$ equations and reflects the terms from $\zeta^{-n_{j+1}-1}$ downto ζ^{-N+1} of (3.3). We could add $N + n_j - k_j - 1$ zero rows at the top to get a total of $2N$ rows and to cover the terms upto ζ^N.

Next we want to include the inner polynomials. To this end, we multiply the $(N - n_j + k_j + 1) \times (2k_j + 2)$ matrix $\underline{\mathsf{S}}_{k_j+1}(g_{n_j},f_{n_j})$ by the $(2k_j + 2) \times (2k_j + 2)$ matrix

$$\mathsf{C}_j^\circ :\equiv \left[\begin{array}{c|c} \grave{\mathsf{B}}_j^\circ & \mathsf{B}_j^\circ \\ \hline \grave{\mathsf{A}}_j^\circ & \mathsf{A}_j^\circ \end{array}\right] :\equiv \left[\begin{array}{ccccc|ccccc} 1 & 1 & \cdots & 1 & 0 & 0 & \mu_1^{(j)} & \cdots & \mu_k^{(j)} & \beta_0^{(j)} \\ 0 & 0 & \cdots & 0 & \grave{\beta}_1^{(j)} & & 0 & \ddots & \vdots & \beta_1^{(j)} \\ \vdots & & & \vdots & \vdots & & & \ddots & \mu_1^{(j)} & \vdots \\ \vdots & & & \vdots & \grave{\beta}_{k-1}^{(j)} & & & & 0 & \beta_{k-1}^{(j)} \\ 0 & 0 & \cdots & 0 & 0 & & & & & 0 \\ \hline 0 & 0 & \cdots & 0 & \grave{\alpha}_0^{(j)} & 1 & 0 & \cdots & 0 & \alpha_0^{(j)} \\ \vdots & & & \vdots & \grave{\alpha}_1^{(j)} & & 1 & \ddots & \vdots & \alpha_1^{(j)} \\ \vdots & & & \vdots & \vdots & & & \ddots & 0 & \vdots \\ \vdots & & & \vdots & \grave{\alpha}_{k-1}^{(j)} & & & & 1 & \alpha_{k-1}^{(j)} \\ 0 & 0 & \cdots & 0 & 0 & & & & & 1 \end{array}\right].$$

This yields

$$
\underline{\mathsf{S}}_{k_j+1}(g_{n_j}, f_{n_j})\,\mathsf{C}_j^\circ =
$$

$$
\left[\begin{array}{cccc|cccc}
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\
\vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\
0 & \cdots & 0 & \vdots & \vdots & & \vdots & \vdots \\
1 & \cdots & 1 & \vdots & 0 & \cdots & 0 & \vdots \\
\gamma_{1,n_j} & \cdots & \gamma_{1,n_j} & \vdots & \phi_{1,n_j} & \cdots & \phi_{1,n_{j+1}-1} & \vdots \\
\vdots & & \vdots & 0 & \vdots & & \vdots & 0 \\
\vdots & & \vdots & 1 & \vdots & & \vdots & 0 \\
\gamma_{k+1,n_j} & \cdots & \gamma_{k+1,n_j} & \gamma_{1,n_{j+1}} & \phi_{k+1,n_j} & \cdots & \phi_{k+1,n_{j+1}-1} & \phi_{1,n_{j+1}} \\
\hline
\gamma_{k+2,n_j} & \cdots & \gamma_{k+2,n_j} & \gamma_{2,n_{j+1}} & \phi_{k+2,n_j} & \cdots & \phi_{k+2,n_{j+1}-1} & \phi_{2,n_{j+1}} \\
\vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots \\
\gamma_{N-n_j,n_j} & \cdots & \gamma_{N-n_j,n_j} & \gamma_{N-n_j-k,n_{j+1}} & \phi_{N-n_j,n_j} & \cdots & \phi_{N-n_j,n_{j+1}-1} & \phi_{N-n_j-k,n_{j+1}}
\end{array}\right].
$$

We still need to imbed these matrices into even bigger ones of size $2N \times 2N$. First, as mentioned, $\underline{\mathsf{S}}_{k_j+1}(g_{n_j}, f_{n_j})$ is extended on top by $N+n_j-k_j-1$ zero rows to make it a $2N \times (2k_j+2)$ matrix. Then each half of it is extended on the right by $N-n_j-k_j-1$ columns, so that each half has size $2N \times (N-n_j)$. Finally, at the left of each half, n_j columns with the previously computed coefficients of the residuals g_n, $n = 0, \ldots n_j - 1$, and f_n, $n = 0, \ldots n_j - 1$, are appended. This yields a $2N \times 2N$ matrix that we call again V_{j-1}. At the beginning and at the end, we have, respectively,

$$
\mathsf{V}_0 = \mathsf{S}_N(g_0, f_0), \qquad \mathsf{V}_J = \left[\begin{array}{c|c} \mathsf{e}_N^\top \mathsf{e}_1 & \mathsf{O} \\ \hline \mathsf{G}_N & \mathsf{F}_N \end{array}\right], \tag{7.8}
$$

where the last matrix is the same one that appears in (6.7).

To imbed C_j°, each of the four square blocks of size k_j+1 of C_j° is first extended to a square block of size $N-n_j+1$ by appending $N-n_j-k_j$ rows and columns, so that the last $N-n_j-k_j+1$ columns have Toeplitz structure. This yields the blocks $\grave{\mathsf{B}}_j$, B_j, $\grave{\mathsf{A}}_j$, and A_j, which are imbedded into a $2N \times 2N$ matrix C_j as follows:

$$
\mathsf{C}_j :\equiv \left[\begin{array}{cc|cc}
\mathsf{I}_{n_j-1} & \mathsf{O} & \mathsf{O}_{n_j-1} & \mathsf{O} \\
\mathsf{O} & \grave{\mathsf{B}}_j & \mathsf{O} & \mathsf{B}_j \\
\hline
\mathsf{O}_{n_j-1} & \mathsf{O} & \mathsf{I}_{n_j-1} & \mathsf{O} \\
\mathsf{O} & \grave{\mathsf{A}}_j & \mathsf{O} & \mathsf{A}_j
\end{array}\right]. \tag{7.9}
$$

Note that in view of the 1's in the upper left corners of $\grave{\mathsf{B}}_j$ and A_j the blocks I_{n_j-1} are part of unit matrices of size n_j, whose boundaries do not match those of the blocks $\grave{\mathsf{B}}_j$, $\grave{\mathsf{A}}_j$, B_j, and A_j, however. Now we have

$$
\mathsf{V}_{j-1}\mathsf{C}_j = \mathsf{V}_j \tag{7.10}
$$

and, consequently,

$$
\mathsf{S}_N(g_0, f_0) \prod_{j=1}^{J} \mathsf{C}_j = \left[\begin{array}{c|c} \mathsf{e}_N^\top \mathsf{e}_1 & \mathsf{O} \\ \hline \mathsf{G}_N & \mathsf{F}_N \end{array}\right], \tag{7.11}
$$

where G_N is now lower block Hessenberg, and F_N is lower block triangular. So, finally, the comparison with (6.7) reveals that

$$\prod_{j=1}^{J} \mathsf{C}_j = \left[\begin{array}{c|c} \grave{\mathsf{P}}_N & \mathsf{P}_N \\ \hline \grave{\mathsf{R}}_N & \mathsf{R}_N \end{array}\right]. \tag{7.12}$$

Eqs. (7.8)–(7.12) are the matrix relations we have been looking for. Eq. (7.10) describes a single step of the Euclidean algorithm as a matrix multiplication, and (7.11) summarizes the whole process as a matrix product with the factors C_j. Moreover, (7.11) shows how the coefficients of the Padé forms constructed are also obtained by multiplying all these matrices C_j. Finally, as we have seen at the end of section 6, it is an easy matter to link (7.8) to the block LDU decomposition and to the inverse block LDU decomposition of the Hankel matrix M_n.

Acknowledgement. The author would like to thank Paul van Dooren for pointing out some references to earlier related work such as [**HS69**]. He also mentioned that his diploma student Paul Bailey wrote a related thesis, which, however, was not available to the author.

References

[Bec96] B. Beckermann, *The stable computation of formal orthogonal polynomials*, Numerical Algorithms **11** (1996), 1–23.

[BV97] A. Bultheel and M. Van Barel, *Linear algebra, rational approximation and orthogonal polynomials*, Elsevier, Amsterdam, 1997.

[CM93] S. Cabay and R. Meleshko, *A weakly stable algorithm for Padé approximants and the inversion of Hankel matrices*, SIAM J. Matrix Anal. Appl. **14** (1993), 735–765.

[DI89] J.-M. Delosme and I. C. F. Ipsen, *From Bareiss' algorithm to the stable computation of partial correlations*, J. Comput. Appl. Math. **27** (1989), 53–91.

[GG90] G. H. Golub and M. H. Gutknecht, *Modified moments for indefinite weight functions*, Numer. Math. **57** (1990), 607–624.

[GG94] W. B. Gragg and M. H. Gutknecht, *Stable look-ahead versions of the Euclidean and Chebyshev algorithms*, Approximation and Computation (R. V. M. Zahar, ed.), ISNM, vol. 119, Birkhäuser Verlag, Basel-Boston-Berlin, 1994, pp. 231–260.

[GGWY82] W. B. Gragg, F. G. Gustavson, D. D. Warner, and D. Y. Y. Yun, *On fast computation of superdiagonal Padé fractions*, Math. Programming Stud. **18** (1982), 39–42.

[GH95a] M. H. Gutknecht and M. Hochbruck, *Look-ahead Levinson and Schur algorithms for non-Hermitian Toeplitz systems*, Numer. Math. **70** (1995), 181–227.

[GH95b] M. H. Gutknecht and M. Hochbruck, *The stability of inversion formulas for Toeplitz matrices*, Linear Algebra Appl. **223/224** (1995), 307–324.

[Gut92] M. H. Gutknecht, *A completed theory of the unsymmetric Lanczos process and related algorithms, Part I*, SIAM J. Matrix Anal. Appl. **13** (1992), 594–639.

[Gut93] M. H. Gutknecht, *Stable row recurrences in the Padé table and generically superfast lookahead solvers for non-Hermitian Toeplitz systems*, Linear Algebra Appl. **188/189** (1993), 351–421.

[HR84] G. Heinig and K. Rost, *Algebraic methods for Toeplitz-like matrices and operators*, Akademie-Verlag, Berlin, and Birkhäuser, Basel/Stuttgart, 1984.

[HS69] A. S. Householder and G. W. Stewart, *Bigradients, Hankel determinants, and the Padé table*, Constructive aspects of the fundamental theorem of algebra (B. Dejon and P. Henrici, eds.), Wiley-Interscience, 1969, pp. 131–150.

[Mag62] A. Magnus, *Certain continued fractions associated with the Padé table*, Math. Z. **78** (1962), 361–374.

[Str63] G. W. Struble, *Orthogonal polynomials: variable-signed weight functions*, Numer. Math **5** (1963), 88–94.

Seminar for Applied Mathematics, ETH-Zurich, ETH-Zentrum HG, CH-8092 Zurich
E-mail address: `mhg@sam.math.ethz.ch`