Other Manifestations of the Schur Complement

Claude Brezinski
Laboratoire d'Analyse Numérique et d'Optimisation
UFR IEEA — M3
Université de Lille Flandres-Artois
59655 Villeneuve d'Ascq-Cedex, France

Submitted by Richard A. Brualdi

ABSTRACT

Many series and sequence transformations used in numerical analysis are defined as ratios of determinants. They are implemented by recursive algorithms based on determinantal identities. The aim of this paper is to study the connections of these transformations and algorithms with the Schur complement of a matrix. The Schur's formula is extended to the vector case, thus providing the same treatment for vector sequence transformations and the corresponding recursive algorithms. Thanks to these connections, particular rules for avoiding division by zero or numerical instability are obtained for these algorithms. Some fixed-point methods and continued fractions also fit in this framework.

1. INTRODUCTION

Let M be a matrix partitioned into four blocks

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix},$$

where the submatrix A is assumed to be square and nonsingular. The Schur complement of A in M, denoted by (M/A), is defined by

$$(M/A) = D - CA^{-1}B.$$

When M is square, matrices of this form are connected with block Gaussian decomposition, since we have

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} I & 0 \\ CA^{-1} & I \end{pmatrix} \begin{pmatrix} A & B \\ 0 & (M/A) \end{pmatrix}.$$

655 Avenue of the Americas, New York, NY 10010 0024-3795/88/$3.50

Moreover

$$|M| = |A||(M/A)|$$

where $|\cdot|$ denotes determinant, an identity first proved by Schur [41].

Let us remark that having the nonsingular square submatrix A in the upper left-hand corner of M is a matter of convention. We can similarly define the Schur complements

$$(M/B) = C - DB^{-1}A,$$

$$(M/C) = B - AC^{-1}D,$$

$$(M/D) = A - BD^{-1}C,$$

which are related to Gaussian elimination by

$$M = \begin{pmatrix} I & 0 \\ DB^{-1} & I \end{pmatrix} \begin{pmatrix} A & B \\ (M/B) & 0 \end{pmatrix}, \qquad |M| = -|B||(M/B)|,$$

$$M = \begin{pmatrix} I & AC^{-1} \\ 0 & I \end{pmatrix} \begin{pmatrix} 0 & (M/C) \\ C & D \end{pmatrix}, \qquad |M| = -|C||(M/C)|,$$

$$M = \begin{pmatrix} I & BD^{-1} \\ 0 & I \end{pmatrix} \begin{pmatrix} (M/D) & 0 \\ C & D \end{pmatrix}, \qquad |M| = |D||(M/D)|.$$

These past few years Schur complements have received much attention. A survey paper of 1981 by Ouellette [37] contains 76 references, and a recent book by Fiedler [24] gives the latest developments. As examplified in [20], Schur complements have many applications in computing inertias of matrices, tests for positive definiteness, and the Minkowsky property; they also arise in statistics and mathematical programming among other fields. They have been recently proved to be the basis for a common derivation of the determinantal identities named after Cauchy, Sylvester, Kronecker, Jacobi, Binet, Laplace, Muir, and Cayley [17]. Schweins' identity can also be included in the list, since it follows from Sylvester's. A generalization of the Schur complement introduced in [25] was proved to be useful in deriving an extension of Sylvester's identity [35] and a test for the total positivity of a matrix.

The aim of this paper is to show the strong connection between Schur complements and methods for transforming sequences and series such as extrapolation methods used to accelerate the convergence of sequences of numbers or vectors, Padé approximants, and continued fractions. Due to their relations with fixed-point methods, some results in this direction will also be given. Such a connection is not surprising, since many sequence transformations are defined as a ratio of two determinants, as in Schur's formula given above, and many recursive algorithms used for their implementation are obtained via Sylvester's determinantal identity or by the bordering method for solving a system of linear equations, whose connection with Schur complements and Gaussian elimination is well known [23, p. 167; 24, p. 237; 27, p. 58].

As we shall see below, the use of Schur's complement provides us a systematic way for obtaining singular rules for avoiding division by zero or numerical instability in the algorithms studied. Such rules, which are very important in practice, are given here for the first time. Nuttall's compact type formulas were also obtained for these algorithms.

2. PADÉ APPROXIMANTS

Let $f(t) = c_0 + c_1 t + c_2 t^2 + \cdots$ be a formal power series. The rational fraction $[p/q]_f(t) = (a_0 + a_1 t + \cdots + a_p t^p)/(1 + b_1 t + \cdots + b_q t^q)$ is called a Padé approximant of f if $a_p b_q \neq 0$ and if its power-series expansion agrees with f up to the degree $p+q$ inclusive. $[p/q]$ can be written as a ratio of two determinants, a formula due to Jacobi [1]:

$$[p/q] = \frac{\begin{vmatrix} \sum_{i=0}^{p-q} c_i t^i & c_{p-q+1} & \cdots & c_p \\ -t^{p-q+1} c_{p-q+1} & c_{p-q+1} - t c_{p-q+2} & \cdots & c_p - t c_{p+1} \\ \cdots & \cdots & \cdots & \cdots \\ -t^{p-q+1} c_p & c_p - t c_{p+1} & \cdots & c_{p+q-1} - t c_{p+q} \end{vmatrix}}{|W|}$$

where

$$W = \begin{pmatrix} c_{p-q+1} - t c_{p-q+2} & \cdots & c_p - t c_{p+1} \\ \cdots & \cdots & \cdots \\ c_p - t c_{p+1} & \cdots & c_{p+q-1} - t c_{p+q} \end{pmatrix}.$$

Taking the Schur complement of W in the numerator of $[p/q]$ and applying Schur's formula, we obtain

$$\begin{aligned}[p/q]_f(t) &= \sum_{i=0}^{p-q} c_i t^i - \left(-t^{p-q+1}c, W^{-1}c\right) \\ &= \sum_{i=0}^{p-q} c_i t^i + t^{p-q+1}\left(c, W^{-1}c\right),\end{aligned}$$

where $c = (c_{p-q+1}, \ldots, c_p)^T$. This formula is known as Nuttall's compact formula for Padé approximants [36; 7, pp. 17–18, 35–36].

A similar formula can be obtained for the vector Padé approximants introduced by Van Iseghem [47] or for the multivariate Padé approximants studied in [19, p. 106] or for the partial Padé approximants defined in [13].

3. CONTINUED FRACTIONS

We consider the continued fraction

$$C = b_0 + \frac{a_1|}{|b_1} + \frac{a_2|}{|b_2} + \cdots.$$

Let $C_n = A_n/B_n$ be its nth convergent. It is well known [38, p. 8] that

$$A_n = \begin{vmatrix} b_0 & -1 & & & & \\ a_1 & b_1 & -1 & & 0 & \\ & a_2 & b_2 & -1 & & \\ & & \ddots & \ddots & \ddots & \\ & 0 & & a_{n-1} & b_{n-1} & -1 \\ & & & & a_n & b_n \end{vmatrix},$$

$$B_n = \begin{vmatrix} b_1 & -1 & & & & \\ a_2 & b_2 & -1 & & 0 & \\ & a_3 & b_3 & -1 & & \\ & & \ddots & \ddots & \ddots & \\ & 0 & & a_{n-1} & b_{n-1} & -1 \\ & & & & a_n & b_n \end{vmatrix}.$$

By Schur's formula we have

$$A_n = A_{n-1}\left|b_n - (0,\ldots,0,a_n)A_{n-1}^{-1}(0,\ldots,0,\ -1)\right|^T$$
$$= A_{n-1}\left(b_n + a_n\frac{A_{n-2}}{A_{n-1}}\right)$$

and a similar formula for B_n. This is the usual recurrence formula for the successive convergents of a continued fraction. Reciprocally, starting from this recurrence relation and using Schur's formula leads to the above determinantal formula for the convergent of a continued fraction. The complete algebraic theory of continued fractions in a noncommutative algebra due to Draux [22] makes an extensive implicit use of Schur complements as well as the generalized continued fractions introduced by Rutishauser [39] and developed by Magnus [31].

4. THE E-ALGORITHM

The E-algorithm is a very general sequence transformation including as particular cases most of the sequence transformations actually known [5, 28] and, in particular, Shanks's transformation [42], that is, the ε-algorithm of Wynn [48].

In this transformation the sequence (S_n) is transformed into a set of sequences whose members $E_k^{(n)}$ are defined by ratios of determinants. It has been proved that [14]

$$E_{k+m}^{(n)} = \frac{\begin{vmatrix} E_k^{(n)} & \Delta E_k^{(n)} & \cdots & \Delta E_k^{(n+m-1)} \\ g_{k,k+1}^{(n)} & \Delta g_{k,k+1}^{(n)} & \cdots & \Delta g_{k,k+1}^{(n+m-1)} \\ \cdots & \cdots & \cdots & \cdots \\ g_{k,k+m}^{(n)} & \Delta g_{k,k+m}^{(n)} & \cdots & \Delta g_{k,k+m}^{(n+m-1)} \end{vmatrix}}{\begin{vmatrix} \Delta g_{k,k+1}^{(n)} & \cdots & \Delta g_{k,k+1}^{(n+m-1)} \\ \cdots & \cdots & \cdots \\ \Delta g_{k,k+m}^{(n)} & \cdots & \Delta g_{k,k+m}^{(n+m-1)} \end{vmatrix}},$$

where Δ operates on the upper indexes. The quantities $g_{k+m,i}^{(n)}$ can be expressed by similar ratios by replacing the first row of the numerator by $g_{k,i}^{(n)}, \Delta g_{k,i}^{(n)}, \dots, \Delta g_{k,i}^{(n+m-1)}$. In these formulas we have $E_0^{(n)} = S_n$ and $g_{0,i}^{(n)} = g_i(n)$, where the g_i's are given (known) auxiliary sequences.

Applying Schur's formula, we immediately obtain

$$E_{k+m}^{(n)} = E_k^{(n)} - \left(\Delta E_k^{(n)}, \dots, \Delta E_k^{(n+m-1)}\right)$$

$$\times \begin{pmatrix} \Delta g_{k,k+1}^{(n)} & \cdots & \Delta g_{k,k+1}^{(n+m-1)} \\ \cdots & \cdots & \cdots \\ \Delta g_{k,k+m}^{(n)} & \cdots & \Delta g_{k,k+m}^{(n+m-1)} \end{pmatrix}^{-1} \begin{pmatrix} g_{k,k+1}^{(n)} \\ \vdots \\ g_{k,k+m}^{(n)} \end{pmatrix}$$

and a similar relation for $g_{k+m,i}^{(n)}$. For $m = 1$ these relations reduce to the usual recursive formulas for the E-algorithm,

$$E_{k+1}^{(n)} = E_k^{(n)} - g_{k,k+1}^{(n)} \frac{\Delta E_k^{(n)}}{\Delta g_{k,k+1}^{(n)}},$$

$$g_{k+1,i}^{(n)} = g_{k,i}^{(n)} - g_{k,k+1}^{(n)} \frac{\Delta g_{k,i}^{(n)}}{\Delta g_{k,k+1}^{(n)}}.$$

These relations can only be used if $\Delta g_{k,k+1}^{(n)} \neq 0$. If such a singularity arises, the preceding formulas for $m > 1$ enable us to jump over the singularity, thus providing us a singular rule for the E-algorithm. If $|\Delta g_{k,k+1}^{(n)}|$ is small, the computation can be affected by a cancellation error and the usual rule of the E-algorithm becomes numerically unstable. This drawback can sometimes be avoided by using the singular rule.

Let us give a numerical example to illustrate this point. We consider the case where S_n is the partial sum of a power series f:

$$S_n = c_0 + c_1 x + \cdots + c_n x^n, \qquad n = 0, 1, \dots.$$

with the choice $g_i(n) = c_{n+i} x^{n+i}$, $E_k^{(n)} = [n + k/k]_f(x)$. Thus if f is the power-series expansion of a rational function with a numerator of degree at most k and a denominator of degree k, we shall have $E_k^{(n)} = f(x)$ for all n and x.

Let us take $f(x) = (1 + \varepsilon x)/(1 - x^2) = 1 + \varepsilon x + x^2 + \varepsilon x^3 + x^4 + \cdots$. If $\varepsilon = 0$, the normal rule of the E-algorithm cannot be used, since a division by zero occurs. The singular rule provides the exact answer $E_2^{(n)} = f(x)$.

For small values of $|\varepsilon|$, numerical instability affects the normal rule of the E-algorithm, but not the singular rule.

For $\varepsilon = 10^{-7}$ and $x = 0.7$, $f(x) = 1.96078$. The values of $E_2^{(n)}$ obtained are the following:

n	Normal rule	Singular rule
0	1.96078	1.96078
1	1.49000	1.96078
2	1.96078	1.96078
3	1.73010	1.96078
4	1.96078	1.96078
5	1.84775	1.96078

For $x = 0.8$, $f(x) = 2.77778$ and we have

n	Normal rule	Singular rule
0	2.77778	2.77778
1	3.27840	2.77778
2	2.77778	2.77778
3	2.04960	2.77778

For $x = 0.9$ a division by zero occurs in the normal rule, while the singular rule gives us an error of one in the last digit. The computations were performed on an Epson HX-20 working with six decimal digits in single precision.

For $k = 0$ our relation gives a Nuttall-type formula for the E-algorithm:

$$E_m^{(n)} = S_n - (\Delta S_n, \dots, \Delta S_{n+m-1})$$

$$\times \begin{pmatrix} \Delta g_1(n) & \cdots & \Delta g_1(n+m-1) \\ \cdots & \cdots & \cdots \\ \Delta g_m(n) & \cdots & \Delta g_m(n+m-1) \end{pmatrix}^{-1} \begin{pmatrix} g_1(n) \\ \vdots \\ g_m(n) \end{pmatrix}.$$

The E-algorithm can be also derived from a generalization of the Neville-Aitken scheme for recursive interpolation which is due to Mühlbach [32]; it is the so-called MNA algorithm, whose proof can be based on Sylvester's determinantal identity [6]. Obviously, similar results also hold for the MNA algorithm, thus providing a singular rule and a Nuttall-type formula for it. In [30] the MNA algorithm was used to construct generalized rational interpolants, that is, interpolants where the numerator and the denominator are linear combinations of given functions forming a complete or quasicomplete Chebyshev system on the set of interpolating points; a singular rule, in the case of an isolated singularity, is also given. Using the above technique, a more general singular rule could be obtained.

The E-algorithm is the most complete generalization of the Richardson extrapolation process actually existing. Another generalization, dealing with a special case, was given by Sidi [43] together with a recursive algorithm for its implementation [44]. This algorithm is based on the special features of the system of linear equations defining the transformation, which can be written as a ratio of two determinants. Schur's formula can thus be also applied to this generalization.

The E and MNA algorithms have recently received applications for implementing composite sequence transformations [11] and are used, in conjunction with extrapolation methods, for predicting the unknown terms of a sequence whose first ones are given [12]. Thus, again, Schur's complements have applications to these questions.

Interpolation and extrapolation methods are based on the solution of systems of linear equations. Many investigations have been recently conducted on this topic in order to obtain efficient (that is, recursive) algorithms for the implementation of these methods. As stated by Mühlbach [33], there are three possibilities for obtaining such algorithms: determinantal identities (as in [5]), use of some special features of the linear system (as in [44]), and elimination strategies (as in [34]). In this section we showed that a fourth possibility is offered by Schur's complements and formula and that this approach is useful for obtaining singular rules for the algorithms in order to jump over the singularities and/or to avoid numerical instability (see [46]). Let us mention that many other methods for computing ratios of determinants [4] can receive a similar treatment.

5. THE VECTOR CASE

Let E be a vector space on K ($\mathbb{R}$ or $\mathbb{C}$), and $x_0, x_1, \dots, x_k$ elements of E. We shall set $x^T = (x_1, \dots, x_k)$. Let $u \in K^k$ and $A \in K^{k \times k}$. The determinant

$$\begin{vmatrix} x_0 & x^T \\ u & A \end{vmatrix}$$

denotes the linear combination of $x_0, x_1, \dots, x_k$ obtained by expanding it with respect to its first row by using the classical rule for this purpose.

It was proved in [10] that

$$\frac{\begin{vmatrix} x_0 & x^T \\ u & A \end{vmatrix}}{|A|} = x_0 - x * A^{-1}u,$$

where $x * a = a * x$ stands for the linear combination $a_1 x_1 + \cdots + a_k x_k$ if $a \in K^k$ has components $a_1, \dots, a_k$. This identity, which was called the generalized Magnus identity, clearly appears now as an extension of that of Schur. It can also be derived directly from Cauchy's development [17]. It was used, together with the bordering method, to derive an extension of Sylvester's identity for determinants whose first row consists of elements of E, and an extension of Schweins's as well. If $E = \mathbb{C}^p$, which is the most frequent case in numerical analysis, it provides an extension of Schur's classical identity to the case of a vector left upper block, thus leading to a determinantal interpretation of some classical methods of numerical analysis. For example, let us consider a quasi-Newton method for solving a system of nonlinear equations $f(x) = 0$ [21]. It gives rise to iterations of the form

$$x_{n+1} = x_n - J_n^{-1} f(x_n),$$

where J_n is an approximation of the Jacobian matrix of f at the point x_n (or its exact value for Newton's method). Using our generalized Schur's identity, we immediately obtain the following determinantal formula

$$x_{n+1} = \frac{\begin{vmatrix} x_n & I \\ f(x_n) & J_n \end{vmatrix}}{|J_n|}.$$

In particular, for $J_n = f'(x_n)$ Newton's method is recovered.

Let us write $f(x) = x - F(x)$. Henrici's method [29] is a quadratic method (under some assumptions) for finding the fixed point of F. At the nth iteration we compute

$$u_0 = x_n$$

$$u_{k+1} = F(u_k), \qquad k = 0, \dots, p,$$

where p is the dimension of the system. Let ΔU_i be the matrix whose columns are $\Delta u_i, \Delta u_{i+1}, \dots, \Delta u_{i+p-1}$ and $\Delta^2 U_0 = \Delta U_1 - \Delta U_0$. Then, in the above framework, Henrici's method corresponds to the choice $J_n = (\Delta^2 U_0)(\Delta U_0)^{-1}$, thus providing a determinantal formula for it.

We now consider a system of linear equations $Bx = b$. It was proved in [7] that the sequence of vectors $x_0 = 0, x_1, x_2, \dots$ generated by the conjugate-gradient method (or the biconjugate-gradient method in the non-symmetric case) was given by

$$x_k = \frac{\begin{vmatrix} 0 & -b & \cdots & -B^{k-1}b \\ c_0 & c_1 & \cdots & c_k \\ \cdots & \cdots & \cdots & \cdots \\ c_{k-1} & c_k & \cdots & c_{2k-1} \end{vmatrix}}{\begin{vmatrix} c_1 & \cdots & c_k \\ \cdots & \cdots & \cdots \\ c_k & \cdots & c_{2k-1} \end{vmatrix}}$$

with $c_i = (b, B^i b)$. Using the above extension of Schur's formula, we get

$$x_k = \left(b, Bb, \dots, B^{k-1}b\right) * \begin{pmatrix} c_1 & \cdots & c_k \\ \cdots & \cdots & \cdots \\ c_k & \cdots & c_{2k-1} \end{pmatrix}^{-1} \begin{pmatrix} c_0 \\ \vdots \\ c_{k-1} \end{pmatrix}.$$

Similar ratios of determinants arise also in orthogonal projection. Let $\mathcal{N}_k$ be the subspace spanned by $v_1, \dots, v_k$, and let w be the orthogonal projection of v on $\mathcal{N}_k$. We have

$$w = -\frac{\begin{vmatrix} 0 & v_1 & \cdots & v_k \\ (v, v_1) & (v_1, v_1) & \cdots & (v_k, v_1) \\ \cdots & \cdots & \cdots & \cdots \\ (v, v_k) & (v_1, v_k) & \cdots & (v_k, v_k) \end{vmatrix}}{\begin{vmatrix} (v_1, v_1) & \cdots & (v_k, v_1) \\ \cdots & \cdots & \cdots \\ (v_1, v_k) & \cdots & (v_k, v_k) \end{vmatrix}},$$

and by Schur's extended formula

$$w = (v_1, \dots, v_k) * \begin{pmatrix} (v_1, v_1) & \cdots & (v_k, v_1) \\ \cdots & \cdots & \cdots \\ (v_1, v_k) & \cdots & (v_k, v_k) \end{pmatrix}^{-1} \begin{pmatrix} (v, v_1) \\ \vdots \\ (v, v_k) \end{pmatrix}.$$

More generally, let E be a vector space on K, and E^* its dual. We shall denote by $\langle \cdot, \cdot \rangle$ the bilinear form of the duality between E and E^*. Let $x_0, x_1, \dots \in E$ and $z_1, z_2, \dots \in E^*$. We consider the determinants

$$D_k^{(i)} = \begin{vmatrix} \langle z_1, x_{i+1} \rangle & \cdots & \langle z_1, x_{i+k} \rangle \\ \cdots & \cdots & \cdots \\ \langle z_k, x_{i+1} \rangle & \cdots & \langle z_k, x_{i+k} \rangle \end{vmatrix}$$

and

$$N_k^{(i)} = \begin{vmatrix} x_i & x_{i+1} & \cdots & x_{i+k} \\ \langle z_1, x_i \rangle & \langle z_1, x_{i+1} \rangle & \cdots & \langle z_1, x_{i+k} \rangle \\ \cdots & \cdots & \cdots & \cdots \\ \langle z_k, x_i \rangle & \langle z_k, x_{i+1} \rangle & \cdots & \langle z_k, x_{i+k} \rangle \end{vmatrix}.$$

In [9] it was proved that the ratios of determinants

$$E_k = N_k^{(0)} / D_k^{(0)}$$

can be recursively computed by the so-called recursive projection algorithm (RPA):

$$E_0 = x_0, \quad g_{0,i} = x_i \qquad \text{for} \quad i \geqslant 1$$

$$E_k = E_{k-1} - \frac{\langle z_k, E_{k-1} \rangle}{\langle z_k, g_{k-1,k} \rangle} g_{k-1,k}, \qquad k > 0,$$

$$g_{k,i} = g_{k-1,i} - \frac{\langle z_k, g_{k-1,i} \rangle}{\langle z_k, g_{k-1,k} \rangle} g_{k-1,k}, \qquad i > k > 0.$$

The $g_{k,i}$'s are given by a ratio of determinants similar to the one of E_k by replacing the first column of $N_k^{(0)}$ by $x_i, \langle z_1, x_i\rangle, \ldots, \langle z_k, x_i\rangle$.

We now consider the ratios given by

$$e_k^{(i)} = N_k^{(i)}/D_k^{(i)}.$$

Of course $e_k^{(0)} = E_k$. The $e_k^{(i)}$'s can be recursively computed by the compact recursive projection algorithm (CRPA):

$$e_0^{(i)} = x_i, \qquad i \geqslant 0,$$

$$e_k^{(i)} = e_{k-1}^{(i)} - \frac{\langle z_k, e_{k-1}^{(i)}\rangle}{\langle z_k, e_{k-1}^{(i+1)}\rangle} e_{k-1}^{(i+1)}, \qquad i \geqslant 0 \quad \text{and} \quad k \geqslant 1.$$

The proofs for these two algorithms are obtained via the extended Sylvester's identity following Schur's. These algorithms can be used to compute recursively the vectors arising in the abovementioned methods: Henrici's, conjugate- and biconjugate-gradient, and projection on a subspace. As shown in [9], it also has connections with the general interpolation and extrapolation problems, Fourier expansion and the Gram-Schmidt orthonormalization process, and Rosen's method used in optimization for projection on the orthogonal complementary space of a subspace.

Of course, by an argument similar to the one used for the scalar E-algorithm, we have for the CRPA

$$e_{k+m}^{(i)} = \frac{\begin{vmatrix} e_k^{(i)} & \cdots & e_k^{(i+m)} \\ \langle z_{k+1}, e_k^{(i)}\rangle & \cdots & \langle z_{k+1}, e_k^{(i+m)}\rangle \\ \cdots\cdots & & \cdots\cdots \\ \langle z_{k+m}, e_k^{(i)}\rangle & \cdots & \langle z_{k+m}, e_k^{(i+m)}\rangle \end{vmatrix}}{\begin{vmatrix} \langle z_{k+1}, e_k^{(i+1)}\rangle & \cdots & \langle z_{k+1}, e_k^{(i+m)}\rangle \\ \cdots\cdots & & \cdots\cdots \\ \langle z_{k+m}, e_k^{(i+1)}\rangle & \cdots & \langle z_{k+m}, e_k^{(i+m)}\rangle \end{vmatrix}}.$$

If $m = 1$ this relation reduces to the above rule of the CRPA. For $m \geqslant 1$, we have, by Schur's extended formula,

$$e_{k+m}^{(i)} = e_k^{(i)} - \left(e_k^{(i+1)}, \dots, e_k^{(i+m)}\right)$$

$$* \begin{pmatrix} \langle z_{k+1}, e_k^{(i+1)} \rangle & \cdots & \langle z_{k+1}, e_k^{(i+m)} \rangle \\ \cdots & \cdots & \cdots \\ \langle z_{k+m}, e_k^{(i+1)} \rangle & \cdots & \langle z_{k+m}, e_k^{(i+m)} \rangle \end{pmatrix}^{-1} \begin{pmatrix} \langle z_{k+1}, e_k^{(i)} \rangle \\ \vdots \\ \langle z_{k+m}, e_k^{(i)} \rangle \end{pmatrix},$$

thus providing an extension of the normal rule of the algorithm to obtain the $e_{k+m}^{(i)}$'s directly from the $e_k^{(i)}$'s without computing the intermediate vectors. This is the singular rule of the CRPA, which is to be used instead of the normal one when the denominator vanishes or is too small, in order to avoid numerical instability. For $k = 0$ the above formula is a Nuttall-type one for the CRPA.

In Section 4 we saw a scalar sequence transformation, namely the E-algorithm. A vector version of it exists in which the following ratios of determinants are considered:

$$E_k^{(n)} = \frac{\begin{vmatrix} S_n & g_1(n) & \cdots & g_k(n) \\ \langle y, \Delta S_n \rangle & \langle y, \Delta g_1(n) \rangle & \cdots & \langle y, \Delta g_k(n) \rangle \\ \cdots & \cdots & \cdots & \cdots \\ \langle y, \Delta S_{n+k-1} \rangle & \langle y, \Delta g_1(n+k-1) \rangle & \cdots & \langle y, \Delta g_k(n+k-1) \rangle \end{vmatrix}}{\begin{vmatrix} \langle y, \Delta g_1(n) \rangle & \cdots & \langle y, \Delta g_k(n) \rangle \\ \cdots & \cdots & \cdots \\ \langle y, \Delta g_1(n+k-1) \rangle & \cdots & \langle y, \Delta g_k(n+k-1) \rangle \end{vmatrix}},$$

where the g_i's are auxiliary sequences of vectors and y an arbitrary vector such that the denominator does not vanish. Thus Schur's extended formula also applies to such ratios. By using the extended Sylvester's identity, derived from Schur's, it can be proved that the $E_k^{(n)}$ can be recursively computed by an algorithm quite similar to the scalar E-algorithm given in Section 4.

We also can prove that

$$E_{k+m}^{(n)} = \frac{\begin{vmatrix} E_k^{(n)} & g_{k,k+1}^{(n)} & \cdots & g_{k,k+m}^{(n)} \\ \langle y, \Delta E_k^{(n)} \rangle & \langle y, \Delta g_{k,k+1}^{(n)} \rangle & \cdots & \langle y, \Delta g_{k,k+m}^{(n)} \rangle \\ \cdots & \cdots & \cdots & \cdots \\ \langle y, \Delta E_k^{(n+k-1)} \rangle & \langle y, \Delta g_{k,k+1}^{(n+k-1)} \rangle & \cdots & \langle y, \Delta g_{k,k+m}^{(n+k-1)} \rangle \end{vmatrix}}{\begin{vmatrix} \langle y, \Delta g_{k,k+1}^{(n)} \rangle & \cdots & \langle y, \Delta g_{k,k+m}^{(n)} \rangle \\ \cdots & \cdots & \cdots \\ \langle y, \Delta g_{k,k+1}^{(n+k-1)} \rangle & \cdots & \langle y, \Delta g_{k,k+m}^{(n+k-1)} \rangle \end{vmatrix}},$$

where $g_{k,i}^{(n)}$ has an expression similar to that of $E_k^{(n)}$ by replacing the S_n's in its first column by $g_i(n)$'s. The $g_{k+m,i}^{(n)}$'s can also be expressed like the $g_{k,i}^{(n)}$'s by replacing the $E_k^{(n)}$'s in the first column of the above expression by $g_{k,i}^{(n)}$'s.

Again Schur's extended formula leads to a singular rule for the vector E-algorithm, whose writing is left to the reader. A particular case of the vector E-algorithm is the topological ε-algorithm, whose ratio of determinants is recovered by the choice $g_i(n) = \Delta S_{n+i-1}$ [2]. Thus a particular rule for it has also been obtained; it has to be compared with that given in [18]. This topological ε-algorithm provides a quadratically convergent method (under some assumptions) for fixed-point problems; see [3] for example. It also has applications to vector Padé approximation [2, 47].

Another interesting vector sequence transformation is due to Germain-Bonne [26, p. 71, Proposition 12]. It is also given by a ratio of determinants similar to the preceding ones. Thus Schur's extended formula also applies to it, leading to a Nuttall-type one. This transformation can be implemented by the RPA.

In [8] the following ratio was considered in connection with the already mentioned Henrici's method for solving fixed-point problems:

$$H_k^{(n)} = \frac{\begin{vmatrix} S_n & \Delta S_n & \cdots & \Delta S_{n+k-1} \\ g_1(n) & \Delta g_1(n) & \cdots & \Delta g_1(n+k-1) \\ \cdots & \cdots & \cdots & \cdots \\ g_k(n) & \Delta g_k(n) & \cdots & \Delta g_k(n+k-1) \end{vmatrix}}{\begin{vmatrix} \Delta g_1(n) & \cdots & \Delta g_1(n+k-1) \\ \cdots & \cdots & \cdots \\ \Delta g_k(n) & \cdots & \Delta g_k(n+k-1) \end{vmatrix}},$$

where the S_n's are vectors and the $g_i(n)$ scalars. The vectors $H_k^{(n)}$ can be recursively computed by the so-called H-algorithm, whose proof is based on the extended Sylvester's formula. Its rules are exactly those of the scalar algorithm given in Section 4 (after replacing the $E_k^{(n)}$ by the $H_k^{(n)}$), except that now the $H_k^{(n)}$ are no longer scalars, but vectors. Thus, similarly to the scalar E-algorithm, the $H_{k+m}^{(n)}$'s and the $g_{k+m,i}^{(n)}$'s can be expressed in a determinantal form in terms of the $H_k^{(n)}$'s and the $g_{k,i}^{(n)}$'s, and Schur's extended formula provides a particular rule for it.

Most of the vector algorithms studied in this section have strong connections with fixed-point methods. Such connections were recently studied in [16, 40, 45].

The bordering method, which is closely connected with Schur's formula, provides the basis for computing the initial values needed for implementing the progressive forms of the above algorithms [15].

REFERENCES

1 G. A. Baker Jr. and P. R. Graves-Morris, *Padé Approximants*, Addison-Wesley, Reading, Mass., 1981.
2 C. Brezinski, Généralisations de la transformation de Shanks, de la table de Padé et de l'ε-algorithme, *Calcolo* 12:317–360 (1975).
3 C. Brezinski, *Accélération de la Convergence en Analyse Numérique*, LNM 584, Springer, Heidelberg, 1977.
4 C. Brezinski, Sur le calcul de certains rapports de déterminants, in *Padé Approximation and Its Applications* (L. Wuytack, Ed.), LNM 765, Springer, Heidelbgerg, 1979.
5 C. Brezinski, A general extrapolation algorithm, *Numer. Math.* 35:175–187 (1980).
6 C. Brezinski, The Mühlbach-Neville-Aitken algorithm and some extensions, *BIT* 20:444–451 (1980).
7 C. Brezinski, *Padé-type Approximation and General Orthogonal Polynomials*, ISNM Vol. 50, Birkhäuser, Basel, 1980.
8 C. Brezinski, About Henrici's method for nonlinear equations, presented at Symposium on Numerical Analysis and Computational Complex Analysis, Zürich, 15–17 Aug. 1983, unpublished.
9 C. Brezinski, Recursive interpolation, extrapolation and projection, *J. Comput. Appl. Math.* 9:369–376 (1983).
10 C. Brezinski, Some determinantal identities in a vector space, with applications, in *Padé Approximation and Its Applications* (H. Werner and H. J. Bünger, Eds.), LNM 1071, Springer, Heidelberg, 1984.
11 C. Brezinski, Composite sequence transformations, *Numer. Math.* 46:311–321 (1985).
12 C. Brezinski, Prediction properties of some extrapolation methods, *Appl. Numer. Math.* 1:457–462 (1985).
13 C. Brezinski, Partial Padé approximants, *J. Approx. Theory*, 54:210–233 (1988).

14 C. Brezinski, Algebraic properties of the *E*-transformation, in *Numerical Methods*, Banach Center Publications, to appear.
15 C. Brezinski, Bordering methods and progressive forms for sequence transformations, *Zastos. Math.*, to appear.
16 C. Brezinski and H. Sadok, Vector sequence transformations and fixed point methods, in *Numerical Methods in Laminar and Turbulent Flows* (C. Taylor et al., Eds.), Prineridge, Swansea, 1987.
17 R. A. Brualdi and H. Schneider, Determinantal identities: Gauss, Schur, Cauchy, Sylvester, Kronecker, Jacobi, Binet, Laplace, Muir, and Cayley, *Linear Algebra Appl.*, 52–53:769–791 (1983).
18 F. Cordellier, Particular rules for the vector ε-algorithm, *Numer. Math.* 27:203–207 (1977).
19 A. Cuyt and L. Wuytack, *Nonlinear methods in numerical analysis*, North-Holland, Amsterdam, 1987.
20 R. W. Cottle, Manifestations of the Schur complement, *Linear Algebra Appl.* 8:189–211 (1974).
21 J. E. Dennis, A brief introduction to quasi-Newton methods, *Proc. Symp. Appl. Math.* 22:19–52 (1978).
22 A. Draux, The Padé approximants in a non-commutative algebra and their applications, in *Padé Approximation and Its Applications* (H. Werner and H. J. Bünger, Eds.), LNM 1071, Springer, Heidelberg, 1984.
23 D. K. Faddeev and V. N. Faddeeva, *Computational Methods of Linear Algebra*, Freeman, San Francisco, 1963.
24 M. Fiedler, *Special Matrices and Their Applications in Numerical Mathematics*, Nijhoff, Dordrecht, 1986.
25 M. Gasca and G. Mühlbach, Generalized Schur-complements and a test for total positivity, *Appl. Numer. Math.* 3:215–232 (1987).
26 B. Germain-Bonne, Estimation de la Limite de Suites et Formalisation de Procédés d'Accélération de Convergence, Thèse, Univ. de Lille I, 1978.
27 G. H. Golub and C. F. Van Loan, *Matrix Computations*, Johns Hopkins U.P., Baltimore, 1983.
28 T. Håvie, Generalized Neville type extrapolation schemes, *BIT* 19:204–213 (1979).
29 P. Henrici, *Elements of Numerical Analysis*, Wiley, New-York, 1964.
30 S. L. Loi and A. W. McInnes, An algorithm for generalized rational interpolation, *BIT* 23:105–117 (1983).
31 A. P. Magnus, Fractions Continues Généralisées: Théorie et Applications, Thèse, Univ. Catholique de Louvain, 1976.
32 G. Mühlbach, The general Neville-Aitken algorithm and some applications, *Numer. Math.* 31:97–110 (1978).
33 G. Mühlbach, Algorithmes d'Extrapolation, Pub. ANO 118, Univ. de Lille I, 1984.
34 G. Mühlbach, Linear and quasilinear extrapolation algorithms, in *Numerical Mathematics and Applications* (R. Vichnevetsky and J. Vignes, Eds.), North-Holland, Amsterdam, 1986.

35 G. Mühlbach and M. Gasca, A generalization of Sylvester's identity on determinants and some applications, *Linear Algebra Appl.* 66:221–234 (1985).
36 J. Nuttall, Convergence of Padé approximants for the Bethe-Salpeter amplitude, *Phys. Rev.* 157:1312–1316 (1967).
37 D. V. Ouellette, Schur complements and statistics, *Linear Algebra Appl.* 36:187–295 (1981).
38 O. Perron, *Die Lehre von dem Kettenbrüchen*, Chelsea, New York, 1950.
39 H. Rutishauser, Über eine Verallgemeinerung der Kettenbrüche, *Z. Angew. Math. Mech.* 39:278–279 (1958).
40 H. Sadok, Accélération de la Convergence de Suites Vectorielles et Méthodes de Point Fixe. Thèse, Univ. de Lille I, 1988.
41 I. Schur, Über Potenzreihen, die im Innern des Einheitskreises beschränkt sind, *J. Reine Angew. Math.* 147:205–232 (1917).
42 D. Shanks, Nonlinear transformations of divergent and slowly convergent sequences, *J. Math. Phys.* 34:1–42 (1955).
43 A. Sidi, Some properties of a generalization of the Richardson extrapolation process, *J. Inst. Math. Appl.* 24:327–346 (1979).
44 A. Sidi, An algorithm for a special case of a generalization of the Richardson extrapolation process, *Numer. Math.* 38:299–307 (1982).
45 D. A. Smith, W. F. Ford, and A. Sidi, Extrapolation methods for vector sequences, *SIAM Rev.* 29:199–233 (1987).
46 U. Tempelmeier, Schurkomplemente in Rekursiven Extrapolationsalgorithmen, Dissertation, Univ. Hannover, 1986.
47 J. Van Iseghem, Vector Padé approximants, in *Numerical Mathematics and Applications* (R. Vichnevetsky and J. Vignes, Eds.), North-Holland, Amsterdam, 1986.
48 P. Wynn, On a device for computing the $e_m(S_n)$ transformation, *MTAC* 10:91–96 (1956).

Received 9 October 1987; final manuscript accepted 21 April 1988