Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Linear Algebra and its Applications 368 (2003) 279–301
LINEAR ALGEBRA AND ITS APPLICATIONS
www.elsevier.com/locate/laa

A Schur complement approach to a general extrapolation algorithm

C. Brezinski [a,*], M. Redivo Zaglia [b]

[a] *Laboratoire de Mathématiques Appliquées, FRE CNRS 2222, UFR de Mathématiques Pures et Appliquées, Université des Sciences et Technologies de Lille, 59655 Villeneuve d'Ascq cedex, France*
[b] *Università degli Studi di Padova, Dipartimento di Matematica Pura ed Applicata, Via G.B. Belzoni 7, 35131 Padova, Italy*

Received 24 April 2002; accepted 8 November 2002

Submitted by R.A. Brualdi

Abstract

This paper is devoted to a Schur complement approach to the E-transformation which is the most general scalar sequence transformation known so far for accelerating the convergence. A new derivation of known results on Schur complements is given in the first part of the paper. Then, Schur complements and their properties are used to obtain various interpretations of the E-transformation. The recursive rules of the E-algorithm for its implementation are also recovered. New results on its kernel are derived and issues on its convergence are discussed. This approach can be extended to the vector case, thus leading to new vector sequence transformations.

Keywords: Schur complement; Extrapolation method; Sequence transformation; E-algorithm

1. Introduction

Iterative methods are aimed at producing a sequence converging to the solution of the problem to be solved. However, when such a sequence converges too slowly, its practical interest is quite limited. Thus, in this case, a *sequence transformation* can be used to transform the slowly converging sequence into a new sequence converging

* Corresponding author. Tel.: +33-3-2043-4296; fax: +33-3-2043-6869.
E-mail addresses: claude.brezinski@univ-lille1.fr (C. Brezinski), michela.redivozaglia@unipd.it (M. Redivo Zaglia).

0024-3795/03/$ - see front matter © 2003 Elsevier Science Inc. All rights reserved.
doi:10.1016/S0024-3795(02)00686-9

faster, under some assumptions, to the same limit. The most well-known example of such a sequence transformation is Aitken Δ^2 process. Sequence transformations are based on the idea of extrapolation [18].

The E-transformation is the most general scalar sequence transformation known so far. The elements of the new sequence it produces are defined by the solution of a system of linear equations and are expressed as ratios of determinants. These ratios can be computed by a triangular recursive algorithm called the E-algorithm. Its rules can be obtained either by an elimination strategy [34], or by means of Sylvester determinantal identity [6], or by the bordering method [12], or by an annihilation operator approach connected to elimination [19]. Ratios of determinants, Sylvester identity, Schur complements, and the bordering method are, in fact, all related to Gaussian elimination for the solution of systems of linear equations. In particular, Sylvester determinantal formula can be obtained from Gaussian elimination [30, Chapter 2, Section 3], or by the bordering method and Schur complements [8], or even directly [14, Appendix 3]. Thus, it is not surprising that the rules of the E-algorithm could be derived by using any of these tools. However, although the ratios of determinants involved in the E-transformation were related to Schur complements in [10], the rules of the E-algorithm have never been established employing only Schur complements. The main motivation of this paper is to derive the E-transformation and the E-algorithm by different approaches and to give various representations that could be extended to the vector and matrix cases [20]. New algebraic and convergence results are also presented. They can be obtained directly without using Schur complements.

Schur complements have many applications in numerical mathematics [10,25–27, 47,48,52]. So, any new derivation concerning this topic could be of some interest. This is, in particular, the case in the domain of sequence transformations [10]. So, the first part of this paper will be devoted to Schur complements. Our goal is not to obtain new results on Schur complements, but only to derive known ones by a different procedure (that could be extended to the vector case). In the second part of the paper, we will exhibit the tight bonds between Schur complements and the E-transformation. Many results scattered in the literature will be recovered, related, and interpreted in an alternative way even those which are not directly connected to Schur complements. This Schur complement approach will also allow us to obtain additional representations and new results for the E-transformation. Moreover, in our derivation, we will only make use of operations that can be carried over to the vector case, thus opening the way to new vector sequence transformations [20].

Section 2 is devoted to Schur complements and their relations with Gaussian factorization and the solution of a system of linear equations. The quotient property of Schur complements will be the subject of Section 3 where it will also be related to Gaussian elimination. These ingredients will be used, in Section 4, to establish the recursive rules of the E-algorithms employing only Schur complements and their quotient property. Then, the quantities computed by the E-transformation will be represented as Schur complements, ratios of determinants, linear combinations, and by recurrence relations. They will also be associated with the solution of different systems of linear equations. We will relate some of these results to important properties of sequence transformations. A new expression, which sheds some light on the structure of the kernel of the E-transformation, will be derived. Various necessary and sufficient conditions for a sequence to belong to the kernel of the E-transformation will be proved. Convergence results will be stated. Examples will be provided as illustrations of the results. Extensions of the techniques of this paper to vector sequences and other topics will be evoked in Section 5.

2. Schur complements

We consider the following partitioned matrices

$$A' = \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \quad B' = \begin{pmatrix} B & E \\ D & F \end{pmatrix},$$
$$C' = \begin{pmatrix} C & D \\ G & H \end{pmatrix}, \quad D' = \begin{pmatrix} D & F \\ H & L \end{pmatrix}.$$

Assuming that D is square and nonsingular, we define the following Schur complements of D in the corresponding partitioned matrix

$$(A'/D) = A - BD^{-1}C, \quad (B'/D) = E - BD^{-1}F,$$
$$(C'/D) = G - HD^{-1}C, \quad (D'/D) = L - HD^{-1}F.$$

Schur complements are related to the Gaussian factorization of a matrix. Indeed, we have

$$A' = \begin{pmatrix} I & BD^{-1} \\ 0 & I \end{pmatrix} \begin{pmatrix} (A'/D) & 0 \\ C & D \end{pmatrix},$$

$$B' = \begin{pmatrix} I & BD^{-1} \\ 0 & I \end{pmatrix} \begin{pmatrix} 0 & (B'/D) \\ D & F \end{pmatrix},$$

$$C' = \begin{pmatrix} I & 0 \\ HD^{-1} & I \end{pmatrix} \begin{pmatrix} C & D \\ (C'/D) & 0 \end{pmatrix},$$

$$D' = \begin{pmatrix} I & 0 \\ HD^{-1} & I \end{pmatrix} \begin{pmatrix} D & F \\ 0 & (D'/D) \end{pmatrix}.$$

So, if the matrices are square, we have the Schur determinantal formulae [50]

$$\det A' = \det(A'/D) \det D,$$
$$\det B' = (-1)^n \det(B'/D) \det D,$$
$$\det C' = (-1)^n \det(C'/D) \det D,$$
$$\det D' = \det(D'/D) \det D,$$

where n is the dimension of the block opposite to D on the diagonal.

If A' is square and nonsingular, we consider the system of linear equations

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} u \\ v \end{pmatrix}.$$

Since D is square and nonsingular, we see from the Gaussian factorization of A' (or from the corresponding determinantal formula), that (A'/D) is nonsingular and it follows that

$$x = (A'/D)^{-1}(u - BD^{-1}v). \tag{1}$$

Once x has been obtained, y can be easily deduced but it has no interest for our purposes.

Similar expressions hold for the systems with matrices B', C' and D'.

Schur complements can be defined also for matrices partitioned into an arbitrary number of blocks. We consider the $n \times m$ block matrix

$$M = \begin{pmatrix} A_{11} & \cdots & A_{1j} & \cdots & A_{1m} \\ \vdots & & \vdots & & \vdots \\ A_{i1} & \cdots & A_{ij} & \cdots & A_{im} \\ \vdots & & \vdots & & \vdots \\ A_{n1} & \cdots & A_{nj} & \cdots & A_{nm} \end{pmatrix}.$$

We denote by $A^{(i,j)}$ the $(n-1) \times (m-1)$ block matrix obtained by deleting the ith row of blocks and the jth column of blocks of M, and we set

$$B_j = \begin{pmatrix} A_{1j} \\ \vdots \\ A_{i-1,j} \\ A_{i+1,j} \\ \vdots \\ A_{nj} \end{pmatrix}, \quad C_i = (A_{i1}, \ldots, A_{i,j-1}, A_{i,j+1}, \ldots, A_{im}).$$

Assuming that A_{ij} is square and nonsingular, the Schur complement of A_{ij} in M is defined as

$$(M/A_{ij}) = A^{(i,j)} - B_j A_{ij}^{-1} C_i. \tag{2}$$

For Schur complements, consult, for example, [29, pp. 19–23] or [56, pp. 36–39].

3. The quotient property

Let us consider the matrix

$$M = \begin{pmatrix} A & B & E \\ C & D & F \\ G & H & L \end{pmatrix}.$$

The *quotient property* of Schur complements was established in [28] using the Schur determinantal formula (for another proof, see [46]). We will give a proof based on the inverse of a bordered matrix.

Property 1. *If* (D'/D) *is nonsingular, then*

$$(M/D') = (A'/D) - (B'/D)(D'/D)^{-1}(C'/D) \tag{3}$$
$$= ((M/D)/(D'/D)). \tag{4}$$

Proof. By definition

$$(M/D') = A - (BE)(D')^{-1}\begin{pmatrix} C \\ G \end{pmatrix}.$$

Setting for simplicity $S = (D'/D)$, we have, by the block bordering method [37],

$$(D')^{-1} = \begin{pmatrix} D & F \\ H & L \end{pmatrix}^{-1} = \begin{pmatrix} D^{-1} + D^{-1}FS^{-1}HD^{-1} & -D^{-1}FS^{-1} \\ -S^{-1}HD^{-1} & S^{-1} \end{pmatrix},$$

and (3) follows easily from the expression of (M/D').

We set

$$M' = \begin{pmatrix} (A'/D) & (B'/D) \\ (C'/D) & (D'/D) \end{pmatrix}.$$

So,

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

But, by (2),

$$\begin{aligned}(M/D) &= \begin{pmatrix} A & E \\ G & L \end{pmatrix} - \begin{pmatrix} B \\ H \end{pmatrix} D^{-1}(C\ F) \\ &= \begin{pmatrix} A - BD^{-1}C & E - BD^{-1}F \\ G - HD^{-1}C & L - HD^{-1}F \end{pmatrix} \\ &= M'\end{aligned}$$

which proves (4). $\square$

Remark 1. In the case where the matrices involves are square, (4) gives

$$\det(M/D) = \det M/\det D = \det(M/D')\det(D'/D).$$

If A, E, G and L are numbers, then the four Schur complements involved in the right hand side of (3) are also numbers and Schur determinantal identity for (M/D') gives

$$\det M \det D = \det A' \det D' - \det B' \det C'$$

which is the Sylvester identity. We also note (see [14, p. 143]) that the Schweins determinantal identity can be obtained by applying the Sylvester identity to the matrix

$$M = \left(\begin{array}{c|ccc|c} 0 & b_1 & \cdots & b_{n-1} & b_n \\ \hline 0 & a_{1,1} & \cdots & a_{1,n-1} & a_{1,n} \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & a_{n-2,1} & \cdots & a_{n-2,n-1} & a_{n-2,n} \\ 1 & a_{n-1,1} & \cdots & a_{n-1,n-1} & a_{n-1,n} \\ \hline 0 & c_1 & \cdots & c_{n-1} & c_n \end{array}\right).$$

So, the Schweins identity can be seen as arising from Schur complements and the quotient property. It can be extended to the case where A, E, G and L are matrices instead of numbers. For Sylvester and Schweins identities, see [1, pp. 45ff, 107ff].

For the connections between Schur complements, Sylvester's and other determinantal identities, see [24].

Many other algebraic properties of Schur complements were given in [42].

Now, we consider the system

$$\begin{pmatrix} A & B & E \\ C & D & F \\ G & H & L \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} u \\ v \\ w \end{pmatrix}. \tag{5}$$

Let us apply the block Gaussian elimination to this system with D as the pivot. The first step consists in eliminating B and H

$$\begin{pmatrix} (A'/D) & 0 & (B'/D) \\ C & D & F \\ (C'/D) & 0 & (D'/D) \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} u - BD^{-1}v \\ v \\ w - HD^{-1}v \end{pmatrix}.$$

Let us notice that this system leads to

$$\begin{pmatrix} (A'/D) & (B'/D) \\ (C'/D) & (D'/D) \end{pmatrix} \begin{pmatrix} x \\ z \end{pmatrix} = \begin{pmatrix} u - BD^{-1}v \\ w - HD^{-1}v \end{pmatrix}. \tag{6}$$

The second step of Gaussian elimination removes (B'/D) and gives

$$\begin{pmatrix} (M/D') & 0 & 0 \\ C & D & F \\ (C'/D) & 0 & (D'/D) \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} (u - BD^{-1}v) - (B'/D)(D'/D)^{-1}(w - HD^{-1}v) \\ v \\ w - HD^{-1}v \end{pmatrix}$$

and it follows

$$x = (M/D')^{-1}[(u - BD^{-1}v) - (B'/D)(D'/D)^{-1}(w - HD^{-1}v)]. \tag{7}$$

This formula can also be obtained directly by applying (1) to the system (6). The expressions for y and z can easily be deduced but they have no interest for our purposes.

We also have

$$M=\begin{pmatrix} I & BD^{-1} & (B'/D)(D'/D)^{-1}\\ 0 & I & 0\\ 0 & HD^{-1} & I\end{pmatrix}\begin{pmatrix}(M/D') & 0 & 0\\ C & D & F\\ (C'/D) & 0 & (D'/D)\end{pmatrix}.$$

4. The E-algorithm

Let (S_n) be a scalar sequence. The E-transformation consists of converting (S_n) into a set of new sequences denoted by $\left(E_k^{(n)}\right)$. The quantities $E_k^{(n)}$, for $k=1,2,\dots$ and $n=0,1,\dots,$ are defined as the first component of the solution of the system (partitioned as the system (5), and using the same denominations for the blocks)

$$\left(\begin{array}{c|ccc|c} 1 & g_1(n) & \cdots & g_{k-1}(n) & g_k(n)\\ \hline 1 & g_1(n+1) & \cdots & g_{k-1}(n+1) & g_k(n+1)\\ \vdots & \vdots & & \vdots & \vdots\\ 1 & g_1(n+k-1) & \cdots & g_{k-1}(n+k-1) & g_k(n+k-1)\\ \hline 1 & g_1(n+k) & \cdots & g_{k-1}(n+k) & g_k(n+k)\end{array}\right)\left(\begin{array}{c} E_k^{(n)}\\ \hline \gamma_1^{(k,n)}\\ \vdots\\ \gamma_{k-1}^{(k,n)}\\ \hline \gamma_k^{(k,n)}\end{array}\right)$$

$$=\left(\begin{array}{c} S_n\\ \hline S_{n+1}\\ \vdots\\ S_{n+k-1}\\ \hline S_{n+k}\end{array}\right). \qquad (8)$$

The $(g_i(n))$ are known given sequences (which can depend on the sequence (S_n)), and $\gamma^{(k,n)}=\left(\gamma_1^{(k,n)},\dots,\gamma_k^{(k,n)}\right)^{\mathrm{T}}$ is a vector of parameters depending on k and n. We set $E_0^{(n)}=S_n$.

The quantities $E_k^{(n)}$ can also be expressed as a ratio of determinants and computed by a recursive algorithm called the E-algorithm; see, for example, [18, pp. 55–72] for more detailed explanations. Both will be recovered from our Schur complement approach.

The system (8) shows that the idea behind the E-transformation is that of extrapolation. According to the choice of the auxiliary sequences $(g_i(n))$, many well-known sequence transformations are recovered as particular cases.

Writing the system (8) for $E_{k-1}^{(n)}$ and $E_{k-1}^{(n+1)}$ gives

$$\begin{pmatrix} A & B\\ C & D\end{pmatrix}\begin{pmatrix} E_{k-1}^{(n)}\\ \alpha^{(k-1,n)}\end{pmatrix}=\begin{pmatrix} u\\ v\end{pmatrix},\qquad \begin{pmatrix} C & D\\ G & H\end{pmatrix}\begin{pmatrix} E_{k-1}^{(n+1)}\\ \beta^{(k-1,n+1)}\end{pmatrix}=\begin{pmatrix} v\\ w\end{pmatrix}, \qquad (9)$$

where $u=S_n$, $v=(S_{n+1},\dots,S_{n+k-1})^{\mathrm{T}}$, $w=S_{n+k}$, and where $\alpha^{(k-1,n)}$ and $\beta^{(k-1,n+1)}$ are vectors of parameters similar to $\gamma^{(k,n)}$.

If, in the system (8) giving $E_k^{(n)}$, the right hand side is replaced by $(g_i(n), \ldots, g_i(n+k))^{\mathrm{T}}$ then the first component of its solution will be denoted by $g_{k,i}^{(n)}$ for $k = 1, 2, \ldots$ and $n = 0, 1, \ldots$ Similar changes in the systems for $E_{k-1}^{(n)}$ and $E_{k-1}^{(n+1)}$ lead to $g_{k-1,i}^{(n)}$ and $g_{k-1,i}^{(n+1)}$ respectively. So, in these systems, u is identical to the block E in (8), v to F, and w to L. We also see that $g_{k,i}^{(n)} = 0$ for $i \leqslant k$. We set $g_{0,i}^{(n)} = g_i(n)$. As will be seen below, the $g_{k,i}^{(n)}$ are auxiliary quantities needed in the recursive rule of the E-algorithm.

In this section, we will make use of Schur complements and their properties for obtaining various representations of the quantities $E_k^{(n)}$. First, the recursive rules of the E-algorithm will be derived by applying the quotient property given in Section 3 to the definition (8) of the transformation. Then, the quantities $E_k^{(n)}$ will be exhibited in different ways as Schur complements and written as a ratio of determinants. They will also be given as a linear combination of successive terms of the sequence to be transformed with coefficients as the solution of a system of linear equations whose matrix could depend on the initial sequence. The kernel of the E-transformation will be characterized from this expression. Other interpretations connected to the solution of linear systems will also be discussed. Necessary and sufficient conditions related to the kernel will be proved. Then, convergence results will be deduced from the expression of $E_k^{(n)}$. Examples conclude the section.

In the sequel, the quantities to be computed will always be assumed to be well defined, and the matrices whose inverses are needed to be nonsingular. In particular, the matrices of the systems (8) are assumed to be nonsingular for all $k \geqslant 0$ and all $n \geqslant 0$. Thus $\forall\, k, n \geqslant 0$, $E_k^{(n)}$ and $g_{k,i}^{(n)}$ exist. We also assume that the sequences $(g_i(n))$, $i = 1, \ldots, k$ are linearly independent (they are said to form a *fundamental set*) which means that, $\forall\, n$,

$$\sum_{i=1}^{k} a_i g_i(n) = 0,$$

implies $a_i = 0$ for $i = 1, \ldots, k$. Let $H_k(u_n)$ be the Casorati's determinant

$$H_k(n) = \begin{vmatrix} g_1(n) & \cdots & g_k(n) \\ \vdots & & \vdots \\ g_1(n+k-1) & \cdots & g_k(n+k-1) \end{vmatrix}. \tag{10}$$

A sufficient condition for the linear independence of the g_i's is that there exists an index $\bar{n}$ such that $H_k(\bar{n}) \neq 0$. In the sequel, we will assume that, $\forall\, k, n \geqslant 0$, $H_k(n) \neq 0$. It is a necessary condition for the matrices of the systems (8) to be nonsingular. This condition also insures that, $\forall\, k, n$, $g_{k-1,k}^{(n)} \neq 0$. The meaning of these conditions will again be discussed in Section 4.2.

4.1. A variety of representations

We will now exhibit several representations for the quantities $E_k^{(n)}$.

4.1.1. Representation 1

Let us first derive the recursive rules of the E-algorithm with the help of Schur complements. In [34], these rules were found by elimination, while they were obtained from Sylvester determinantal identity in [6]. Other derivations are given in [19].

Applying formula (1) (and the similar one with (C'/D)) to the systems (9) gives

$$\begin{aligned}
(B'/D) &= (A'/D) g_{k-1,k}^{(n)}, \\
(D'/D) &= (C'/D) g_{k-1,k}^{(n+1)}, \\
u - BD^{-1}v &= (A'/D) E_{k-1}^{(n)}, \\
w - HD^{-1}v &= (C'/D) E_{k-1}^{(n+1)}.
\end{aligned}$$

Let us now exploit the quotient property and its consequences for the system (8) which defines $E_k^{(n)}$, and make use of (7). We have

$$\begin{aligned}
(M/D') &= (A'/D)[1 - (A'/D)^{-1}(B'/D)(D'/D)^{-1}(C'/D)] \\
&= (A'/D)\left[1 - g_{k-1,k}^{(n)}\left(g_{k-1,k}^{(n+1)}\right)^{-1}\right].
\end{aligned}$$

Similarly

$$\begin{aligned}
&(u - BD^{-1}v) - (B'/D)(D'/D)^{-1}(w - HD^{-1}v) \\
&\quad = (A'/D)E_{k-1}^{(n)} - (B'/D)(D'/D)^{-1}(C'/D)E_{k-1}^{(n+1)} \\
&\quad = (A'/D)\left[E_{k-1}^{(n)} - g_{k-1,k}^{(n)}\left(g_{k-1,k}^{(n+1)}\right)^{-1} E_{k-1}^{(n+1)}\right].
\end{aligned}$$

Substituting into (7) gives

$$E_k^{(n)} = \left[1 - g_{k-1,k}^{(n)}\left(g_{k-1,k}^{(n+1)}\right)^{-1}\right]^{-1}\left[E_{k-1}^{(n)} - g_{k-1,k}^{(n)}\left(g_{k-1,k}^{(n+1)}\right)^{-1} E_{k-1}^{(n+1)}\right].$$

This formula can also be written in a more symmetrical form as

$$\begin{aligned}
E_k^{(n)} &= \left[1 - g_{k-1,k}^{(n)}\left(g_{k-1,k}^{(n+1)}\right)^{-1}\right]^{-1}\left(g_{k-1,k}^{(n)}\right)^{-1} g_{k-1,k}^{(n)} \\
&\quad \times \left[E_{k-1}^{(n)} - g_{k-1,k}^{(n)}\left(g_{k-1,k}^{(n+1)}\right)^{-1} E_{k-1}^{(n+1)}\right] \\
&= \left[\left(g_{k-1,k}^{(n+1)}\right)^{-1} - \left(g_{k-1,k}^{(n)}\right)^{-1}\right]^{-1} \\
&\quad \times \left[\left(g_{k-1,k}^{(n+1)}\right)^{-1} E_{k-1}^{(n+1)} - \left(g_{k-1,k}^{(n)}\right)^{-1} E_{k-1}^{(n)}\right].
\end{aligned}$$

From the preceding relation, we also have

$$
\begin{aligned}
E_k^{(n)} &= \left[\left(g_{k-1,k}^{(n+1)}\right)^{-1} - \left(g_{k-1,k}^{(n)}\right)^{-1}\right]^{-1} \\
&\quad \times \left[\left(g_{k-1,k}^{(n+1)}\right)^{-1} E_{k-1}^{(n)} - \left(g_{k-1,k}^{(n+1)}\right)^{-1} E_{k-1}^{(n)} \right. \\
&\qquad \left. + \left(g_{k-1,k}^{(n+1)}\right)^{-1} E_{k-1}^{(n+1)} - \left(g_{k-1,k}^{(n)}\right)^{-1} E_{k-1}^{(n)}\right] \\
&= \left[(g_{k-1,k}^{(n+1)})^{-1} - (g_{k-1,k}^{(n)})^{-1}\right]^{-1} \\
&\quad \times \left[\left(\left(g_{k-1,k}^{(n+1)}\right)^{-1} - \left(g_{k-1,k}^{(n)}\right)^{-1}\right) E_{k-1}^{(n)} + \left(g_{k-1,k}^{(n+1)}\right)^{-1} \Delta E_{k-1}^{(n)}\right] \\
&= E_{k-1}^{(n)} + \left[g_{k-1,k}^{(n+1)}\left(\left(g_{k-1,k}^{(n+1)}\right)^{-1} - \left(g_{k-1,k}^{(n)}\right)^{-1}\right)\right]^{-1} \Delta E_{k-1}^{(n)} \\
&= E_{k-1}^{(n)} + \left[1 - g_{k-1,k}^{(n+1)}\left(g_{k-1,k}^{(n)}\right)^{-1}\right]^{-1} \Delta E_{k-1}^{(n)} \\
&= E_{k-1}^{(n)} - g_{k-1,k}^{(n)}\left(\Delta g_{k-1,k}^{(n)}\right)^{-1} \Delta E_{k-1}^{(n)}
\end{aligned}
$$

where the forward difference operator Δ acts on the upper indexes.

Similar rules hold for the quantities $g_{k,i}^{(n)}$, $i = k+1, k+2, \ldots,$

$$
\begin{aligned}
g_{k,i}^{(n)} &= \left[1 - g_{k-1,k}^{(n)}\left(g_{k-1,k}^{(n+1)}\right)^{-1}\right]^{-1}\left[g_{k-1,i}^{(n)} - g_{k-1,k}^{(n)}\left(g_{k-1,k}^{(n+1)}\right)^{-1} g_{k-1,i}^{(n+1)}\right] \\
&= \left[\left(g_{k-1,k}^{(n+1)}\right)^{-1} - \left(g_{k-1,k}^{(n)}\right)^{-1}\right]^{-1}\left[\left(g_{k-1,k}^{(n+1)}\right)^{-1} g_{k-1,i}^{(n+1)} - \left(g_{k-1,k}^{(n)}\right)^{-1} g_{k-1,i}^{(n)}\right] \\
&= g_{k-1,i}^{(n)} - g_{k-1,k}^{(n)}\left(\Delta g_{k-1,k}^{(n)}\right)^{-1} \Delta g_{k-1,i}^{(n)}.
\end{aligned}
$$

These rules are exactly those of the E-algorithm as given in [6,34] (see also [40,49]). Obviously, since we are in the scalar case, they can still be written in various other forms. However, we restrict ourselves to manipulations that could be carried over to the vector or the matrix case. In [43], the rules of the vector and the matrix E-algorithm were established from Schur complements. Of course, the same methodology is also valid for the scalar E-algorithm. However, the derivation given above is simpler. A similar approach to other vector and matrix sequence transformations is also possible as shown in [41].

4.1.2. Representation 2

The quantities $E_k^{(n)}$ can be expressed as a ratio of determinants. From (8), we immediately recover the formula

$$
E_k^{(n)} = \begin{vmatrix} S_n & g_1(n) & \cdots & g_k(n) \\ S_{n+1} & g_1(n+1) & \cdots & g_k(n+1) \\ \vdots & \vdots & & \vdots \\ S_{n+k} & g_1(n+k) & \cdots & g_k(n+k) \end{vmatrix} \Bigg/ \begin{vmatrix} 1 & g_1(n) & \cdots & g_k(n) \\ 1 & g_1(n+1) & \cdots & g_k(n+1) \\ \vdots & \vdots & & \vdots \\ 1 & g_1(n+k) & \cdots & g_k(n+k) \end{vmatrix}. \tag{11}
$$

A quite similar determinantal formula allows us to express $E_{k+m}^{(n)}$ in terms of $E_m^{(n+i)}$ and $g_{m,m+j}^{(n+i)}$ for $i, j = 0, \ldots, k$ (the roles of k and m can also be interchanged)

[13]. A formula analogous to (11) holds for the quantities $g_{k,i}^{(n)}$, $i > k$, by replacing $S_n, \dots, S_{n+k}$ in the first column by $g_i(n), \dots, g_i(n+k)$.

4.1.3. Representation 3

Let us show that $E_k^{(n)}$ can be expressed as a Schur complement. The system (8) can be written as

$$\left(\begin{array}{c|ccc} 1 & g_1(n) & \cdots & g_k(n) \\ \hline 0 & \Delta g_1(n) & \cdots & \Delta g_k(n) \\ \vdots & \vdots & & \vdots \\ 0 & \Delta g_1(n+k-1) & \cdots & \Delta g_k(n+k-1) \end{array}\right) \left(\begin{array}{c} E_k^{(n)} \\ \hline \gamma_1^{(k,n)} \\ \vdots \\ \gamma_k^{(k,n)} \end{array}\right) = \left(\begin{array}{c} S_n \\ \hline \Delta S_n \\ \vdots \\ \Delta S_{n+k-1} \end{array}\right).$$

So, the vector $\gamma^{(k,n)}$ is solution of

$$\begin{pmatrix} \Delta g_1(n) & \cdots & \Delta g_k(n) \\ \vdots & & \vdots \\ \Delta g_1(n+k-1) & \cdots & \Delta g_k(n+k-1) \end{pmatrix} \begin{pmatrix} \gamma_1^{(k,n)} \\ \vdots \\ \gamma_k^{(k,n)} \end{pmatrix} = \begin{pmatrix} \Delta S_n \\ \vdots \\ \Delta S_{n+k-1} \end{pmatrix}. \quad (12)$$

Thus

$$\begin{aligned} E_k^{(n)} = S_n &- (g_1(n), \dots, g_k(n)) \\ &\times \begin{pmatrix} \Delta g_1(n) & \cdots & \Delta g_k(n) \\ \vdots & & \vdots \\ \Delta g_1(n+k-1) & \cdots & \Delta g_k(n+k-1) \end{pmatrix}^{-1} \begin{pmatrix} \Delta S_n \\ \vdots \\ \Delta S_{n+k-1} \end{pmatrix}. \end{aligned} \quad (13)$$

4.1.4. Representation 4

Applying the Schur determinantal formula to (13) leads to

$$E_k^{(n)} = \frac{\begin{vmatrix} S_n & g_1(n) & \cdots & g_k(n) \\ \Delta S_n & \Delta g_1(n) & \cdots & \Delta g_k(n) \\ \vdots & \vdots & & \vdots \\ \Delta S_{n+k-1} & \Delta g_1(n+k-1) & \cdots & \Delta g_k(n+k-1) \end{vmatrix}}{\begin{vmatrix} \Delta g_1(n) & \cdots & \Delta g_k(n) \\ \vdots & & \vdots \\ \Delta g_1(n+k-1) & \cdots & \Delta g_k(n+k-1) \end{vmatrix}}. \quad (14)$$

This ratio of determinants is the same as (11). Formulae similar to (13) and (14) hold for the quantities $g_{k,i}^{(n)}$, $i > k$, after replacing $S_n, \dots, S_{n+k}$ by $g_i(n), \dots, g_i(n+k)$, see [10].

4.1.5. Representation 5

Let us now express $E_k^{(n)}$ as a combination of terms of the sequence (S_n). This expression will be useful in the sequel for deriving the results on the kernel of the E-transformation (Section 4.2) and those concerning its convergence (Section 4.3). It is not derived from the Schur complement but directly from the system (8) which may be written

$$E_k^{(n)} = S_{n+i} - \sum_{j=1}^{k} \gamma_j^{(k,n)} g_j(n+i), \quad i = 0, \ldots, k, \tag{15}$$

where the $\gamma_j^{(k,n)}$ are solution of the system (12).

Let $a_0^{(k,n)}, \ldots, a_k^{(k,n)}$ be parameters, depending on k and n, such that $a_0^{(k,n)} a_k^{(k,n)} \neq 0$, and $a_0^{(k,n)} + \cdots + a_k^{(k,n)} = 1$ (which does not restrict the generality). We multiply the first equation of (15) by $a_0^{(k,n)}$, the second one by $a_1^{(k,n)}$, and so on until the last one which is multiplied by $a_k^{(k,n)}$. Adding them together gives

$$\begin{aligned} E_k^{(n)} &= \sum_{i=0}^{k} a_i^{(k,n)} S_{n+i} - \sum_{i=0}^{k} a_i^{(k,n)} \sum_{j=1}^{k} \gamma_j^{(k,n)} g_j(n+i) \\ &= \sum_{i=0}^{k} a_i^{(k,n)} S_{n+i} - \sum_{j=1}^{k} \gamma_j^{(k,n)} \sum_{i=0}^{k} a_i^{(k,n)} g_j(n+i). \end{aligned}$$

Since, for $l > k$,

$$g_{k,l}^{(n)} = g_l(n+i) - \sum_{j=1}^{k} \gamma_j^{(k,n)} g_j(n+i), \quad i = 0, \ldots, k,$$

a similar formula holds for $g_{k,l}^{(n)}$.

Thus, as shown in [49] in a slightly less general setting, we have the

Lemma 1. *For all k and n,*

$$\begin{aligned} E_k^{(n)} &= \sum_{i=0}^{k} a_i^{(k,n)} S_{n+i}, \\ g_{k,l}^{(n)} &= \sum_{i=0}^{k} a_i^{(k,n)} g_l(n+i), \quad l > k, \end{aligned} \tag{16}$$

where the coefficients $a_i^{(k,n)}$ satisfy the system

$$\begin{aligned} &\sum_{i=0}^{k} a_i^{(k,n)} = 1, \\ &\sum_{i=0}^{k} a_i^{(k,n)} g_j(n+i) = 0, \quad j = 1, \ldots, k. \end{aligned} \tag{17}$$

It must be noticed that the condition $H_k(n) \neq 0$ insures that $a_0^{(k,n)} a_k^{(k,n)} \neq 0$ in (17). The coefficients $a_i^{(k,n)}$ can also be recursively computed by the E-algorithm [18, pp. 59–60].

4.1.6. Representation 6

Another expression by Schur complements can be obtained. It follows from Lemma 1 that

$$E_k^{(n)} = (S_n, \ldots, S_{n+k}) \begin{pmatrix} 1 & \cdots & 1 \\ g_1(n) & \cdots & g_1(n+k) \\ \vdots & & \vdots \\ g_k(n) & \cdots & g_k(n+k) \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{18}$$

which shows that $E_k^{(n)}$ is also the Schur complement of the matrix of the system (17) in the matrix

$$\left(\begin{array}{ccc|c} 1 & \cdots & 1 & -1 \\ g_1(n) & \cdots & g_1(n+k) & 0 \\ \vdots & & \vdots & \vdots \\ g_k(n) & \cdots & g_k(n+k) & 0 \\ \hline S_n & \cdots & S_{n+k} & 0 \end{array}\right). \tag{19}$$

Applying the Schur determinantal formula yields (14) again. Obviously, the $g_{k,l}^{(n)}$'s can also be expressed as Schur complements.

It must be noticed that the determinants of the systems (12) and (17) are the same, and that $\left(a_0^{(k,n)}, \ldots, a_k^{(k,n)}\right)^{\mathrm{T}}$ is the first column of the inverse of the matrix of the system (17).

4.1.7. Representation 7

Let us give another interpretation of the E-transformation and relate it to the previous one. We consider the system (whose solution depends on k and n)

$$\begin{pmatrix} S_n & \cdots & S_{n+k} \\ g_1(n) & \cdots & g_1(n+k) \\ \vdots & & \vdots \\ g_k(n) & \cdots & g_k(n+k) \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_k \end{pmatrix} = \begin{pmatrix} \alpha \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{20}$$

with $\alpha \neq 0$. Replacing each column by its difference with the previous one, we observe that this system is equivalent to

$$\begin{pmatrix} S_n & \Delta S_n & \cdots & \Delta S_{n+k-1} \\ g_1(n) & \Delta g_1(n) & \cdots & \Delta g_1(n+k-1) \\ \vdots & \vdots & & \vdots \\ g_k(n) & \Delta g_k(n) & \cdots & \Delta g_k(n+k-1) \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{pmatrix} = \begin{pmatrix} \alpha \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

with $\beta_i = \sum_{j=i}^{k} \alpha_j$ for $i = 0, \ldots, k$. So, it is easy to see from (14) (or directly from (11)) that

$$E_k^{(n)} = \frac{\alpha}{\sum_{i=0}^{k} \alpha_i}.$$

This relation generalizes the interpretation of the ε-algorithm given in [3] (see also [5, pp. 51–52]).

We have $\beta_0 = \sum_{i=0}^{k} \alpha_i$. Adding this condition to (20), we get

$$\begin{pmatrix} 1 & \cdots & 1 & -1 \\ g_1(n) & \cdots & g_1(n+k) & 0 \\ \vdots & & \vdots & \vdots \\ g_k(n) & \cdots & g_k(n+k) & 0 \\ S_n & \cdots & S_{n+k} & 0 \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_k \\ \beta_0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \alpha \end{pmatrix}. \tag{21}$$

The matrix of this system is (19). Solving it for the unknown β_0 provides the determinantal expression (11). Comparing with the solution of (17) shows that $\alpha_i = \beta_0 a_i^{(k,n)}$ for $i = 0, \ldots, k$, and it follows $\beta_0 E_k^{(n)} = \sum_{i=0}^{k} \alpha_i S_{n+i} = \alpha$. Since α is arbitrary, it can be adjusted so that $\beta_0 = 1$ and, thus, $\alpha = E_k^{(n)}$. So (21) becomes

$$\begin{aligned} &\alpha_0 + \cdots + \alpha_k - 1 = 0, \\ &\alpha_0 g_1(n) + \cdots + \alpha_k g_1(n+k) = 0, \\ &\vdots \\ &\alpha_0 g_k(n) + \cdots + \alpha_k g_k(n+k) = 0, \\ &\alpha_0 S_n + \cdots + \alpha_k S_{n+k} - E_k^{(n)} = 0. \end{aligned}$$

As this system has a nonzero solution, its determinant is zero and (11) is recovered. This result generalizes the heuristic interpretation of the ε-algorithm given in [32] (see also [5, pp. 52–53]).

4.1.8. Representation 8

We will now examine some of the preceding results in the light of certain properties of sequence transformations which are quite useful for a deeper understanding. Let F be a function of $x_0, \ldots, x_k$. It is assumed to be *translative*, that is, $\forall\, b$, $F(x_0 + b, \ldots, x_k + b) = F(x_0, \ldots, x_k) + b$. As proved in [2], a necessary and sufficient condition for F to be translative is that it can be written as $F(x_0, \ldots, x_k) = f(x_0, \ldots, x_k)/\mathrm{D}f(x_0, \ldots, x_k)$, where $\mathrm{D}f$ denotes the sum of the partial derivatives of f, and where $\mathrm{D}^2 f$ is identically zero. We also assume that F is *homogeneous*, that is, $\forall\, a$, $F(ax_0, \ldots, ax_k) = aF(x_0, \ldots, x_k)$. A function which is translative and homogeneous is called *quasi-linear*. In this case, it holds [11] (see also [18, pp. 11–18])

$$F = \sum_{i=0}^{k} x_i f_i'/\mathrm{D}f \quad \text{with } f_i' = \partial f/\partial x_i, \tag{22}$$

$$= \sum_{i=0}^{k} x_i F_i' \quad \text{with } F_i' = \partial F/\partial x_i. \tag{23}$$

From (11), we see that the E-transformation (as many other sequence transformations) has the form

$$E_k^{(n)} = F(S_n, \ldots, S_{n+k}), \quad k, n = 0, 1, \ldots \tag{24}$$

with F quasi-linear and depending on k and n. Indeed, it corresponds to

$$f(x_0, \ldots, x_k) = \begin{vmatrix} x_0 & \cdots & x_k \\ g_1(n) & \cdots & g_1(n+k) \\ \vdots & & \vdots \\ g_k(n) & \cdots & g_k(n+k) \end{vmatrix}.$$

It can easily be verified that (24) is identical to (18). Indeed, the inverse of the matrix in (18) is the ratio of its adjunct matrix divided by its determinant. This determinant is precisely $\mathrm{D}f(S_n, \ldots, S_{n+k})$. The adjunct of a matrix is the matrix whose element (j, i) is the determinant of the submatrix obtained by deleting the ith row and the jth column, multiplied by $(-1)^{i+j}$. When this adjunct matrix is multiplied by the vector $(1, 0, \ldots, 0)^{\mathrm{T}}$, and then scalarly by the row vector $(S_n, \ldots, S_{n+k})$, we get $f(S_n, \ldots, S_{n+k})$ and (24) is recovered.

It follows from (22)–(24) that $F_i' = a_i^{(k,n)}$, $i = 0, \ldots, k$, the solution of the system (17). So, the coefficients $a_i^{(k,n)}$ have been identified as the partial derivatives of F and, conversely, these partial derivatives have been expressed as the solution of a system of linear equations.

4.2. The kernel

We will now discuss the *kernel* of the E-transformation. By definition, it is the set of sequences such that, $\forall\, n$, $E_k^{(n)} = S$.

We have the following result which was, apparently, first reported in [9] but without a formal proof. Of course, the proof given in [43] for the matrix case is valid in the scalar case, but, as we will see, a simpler one could be obtained in our situation. It is based on the Representations 2, 4, and 5 above.

Theorem 1. *Let k be fixed. A necessary and sufficient condition that, $\forall\, n$, $E_k^{(n)} = S$ is that, $\forall\, n$,*

$$S_n = S + \gamma_1 g_1(n) + \cdots + \gamma_k g_k(n), \tag{25}$$

where the γ_i are constants independent of k and n.

Proof. By construction of the transformation, we see from (8), or (11), or (14), or (15) that, if the sequence (S_n) satisfies (25), then, $\forall n$, $E_k^{(n)} = S$.

Let us prove the reciprocal. From (8), we see that, $\forall n$, $E_k^{(n)} = S_n - \gamma_1^{(k,n)} g_1(n) - \cdots - \gamma_k^{(k,n)} g_k(n)$. If, $\forall n$, $E_k^{(n)} = S$, we will show that the coefficients $\gamma_i^{(k,n)}$ are independent of n.

The systems (12) for the indexes n and $n+1$ have $k-1$ equations in common. Subtracting them gives

$$\Delta g_1(n+i)\Delta\gamma_1^{(k,n)} + \cdots + \Delta g_k(n+i)\, \Delta\gamma_k^{(k,n)} = 0, \quad i = 1, \ldots, k-1,$$

where the forward difference operator Δ acts on n. One more equation is needed. Writing (15) for n and $i = 1$, and for $n+1$ and $i = 0$ provides

$$E_k^{(n)} = S_{n+1} - \sum_{j=1}^{k} \gamma_j^{(k,n)} g_j(n+1),$$

$$E_k^{(n+1)} = S_{n+1} - \sum_{j=1}^{k} \gamma_j^{(k,n+1)} g_j(n+1).$$

But, since, $\forall n$, $E_k^{(n)} = S$, subtracting the first equation from the second leads to

$$\sum_{j=1}^{k} g_j(n+1)\, \Delta\gamma_j^{(k,n)} = 0,$$

where the operator Δ acts on the superscript n. So, we have obtained an homogeneous system of k equations in k unknowns $\Delta\gamma_i^{(k,n)}$, $i = 1, \ldots, k$. It is easy to see that the determinant of this system is the numerator of $g_{k-1,k}^{(n+1)}$ which, by our assumptions, is different from zero. So it follows $\Delta\gamma_i^{(k,n)} = 0$ for $i = 1, \ldots, k$, which ends the proof. □

We will now express this kernel in a different form. From (16) and the first equation in (17), it follows that

$$E_k^{(n)} - S = \sum_{i=0}^{k} a_i^{(k,n)} (S_{n+i} - S).$$

So, we have the

Theorem 2. *Let k be fixed. A necessary and sufficient condition that*, $\forall n$, $E_k^{(n)} = S$ *is that*, $\forall n$,

$$\sum_{i=0}^{k} a_i^{(k,n)} (S_{n+i} - S) = 0, \tag{26}$$

with the coefficients $a_i^{(k,n)}$ *satisfying* (17).

Let us comment on Theorems 1 and 2. From the theory of linear difference equations (see, for example, [33, pp. 292–294], [38, pp. 27–31], [44, pp. 351–360], [45, pp. 1–14] or [54, pp. 265–269]), we know that (26) has a unique solution. The vector space of its solutions has dimension k, and its general solution is a linear combination of k linearly independent particular solutions. From (17), we see that $(g_i(n))$ is a particular solution of (26) for $i = 1, \ldots, k$. Since we assumed the linear independence of the g_i's, it follows that the general solution of (26) is given by the expression (25), and the result of Theorem 1 is recovered. The Casorati's determinant $H_k(n)$ defined by (10) satisfies

$$a_k^{(k,n)} H_k(n+1) = (-1)^k a_0^{(k,n)} H_k(n).$$

So, if we assume that $\exists\, \bar{n}$ such that $H_k(\bar{n}) \neq 0$, and since the $(g_i(n))$ are solutions of (26), then $\forall\, n \geqslant \bar{n},\ H_k(n) \neq 0$ [38, Theorem 2.1.3, p. 30]. Conversely, if $\forall\, n$, $H_k(n) \neq 0$, then $\forall\, n,\ a_0^{(k,n)} a_k^{(k,n)} \neq 0$. The first condition in (17) simply means that the sum of the $a_i^{(k,n)}$'s must be different from zero (if their sum equals $a \neq 0$, each coefficient can be divided by a, thus making the sum equal to 1). If the sum is zero, then S in (26) is not uniquely defined, and, moreover, the system (17) has only the zero solution.

Although the result of Theorem 2 is evident from (16) and (17), it seems that it has not been noticed so far.

The reason why the coefficients γ_i in (25) are independent of k and n, while the coefficients $a_i^{(k,n)}$ in (26) depend on k and n, is that there are $k-1$ equations in common between the systems (12) giving the coefficients γ_i for the indexes n and $n+1$, and no common equation between the systems (17) for the coefficients $a_i^{(k,n)}$ and $a_i^{(k,n+1)}$.

Let us now relate the function F which characterizes a quasi-linear transformation (see Section 4.1.8) to its kernel. It follows from (22) and (23) (see [2] or [11]) that a sequence (S_n) belongs to the kernel of a quasi-linear transformation if and only if, $\forall\, n$,

$$\sum_{i=0}^{k} (S_{n+i} - S) f_i'(S_n - S, \ldots, S_{n+k} - S) = 0,$$

or if and only if, $\forall\, n$,

$$\sum_{i=0}^{k} (S_{n+i} - S) F_i'(S_n - S, \ldots, S_{n+k} - S) = 0. \tag{27}$$

So, the expression (26) of Theorem 2 for the kernel of the E-transformation and its expression as (27) are identical.

4.3. Convergence results

We consider the sequence transformation $T : (S_i) \longmapsto (T_p)$ where

$$T_p = \sum_{i=0}^{\infty} \alpha_{pi} S_i, \quad p = 0, 1, \ldots$$

The transformation T is called *regular* if, for *any* convergent sequence (S_i), the sequence (T_p) converges and its limit is equal to the limit of (S_i). The regularity of T is governed by the *Toeplitz Theorem* (see, for example, [5, pp. 23–24] or [53, pp. 24–27]).

Theorem 3. *The transformation T is regular if and only if*

(1) $\exists M, \ \forall p, \quad \sum_{i=0}^{\infty} |\alpha_{pi}| < M,$

(2) $\forall i, \quad \lim_{p\to\infty} \alpha_{pi} = 0,$

(3) $\lim_{p\to\infty} \sum_{i=0}^{\infty} \alpha_{pi} = 1.$

An important point to notice is that, if the coefficients α_{pi} depend on some constraints (which can involve the sequence (S_i) itself), then the regularity is only valid under the restriction that these constraints are satisfied. In particular, it means that the regularity could hold only for *some* classes of sequences (those satisfying the constraints) and not for *all* convergent sequences.

From (16), we see that the E-transformation enters into this framework either when k is fixed and n tends to infinity or the contrary. Indeed, when k is fixed, we consider the sequence transformation defined by $T_p = E_k^{(p)}$. It corresponds to

$$\alpha_{pi} = \begin{cases} 0, & i = 0, \ldots, p-1, \\ a_{i-p}^{(k,p)}, & i = p, \ldots, p+k, \\ 0, & i > p+k. \end{cases}$$

When n is fixed, we consider the transformation given by $T_p = E_p^{(n)}$, which corresponds to

$$\alpha_{pi} = \begin{cases} 0, & i = 0, \ldots, n-1, \\ a_{i-n}^{(p,n)}, & i = n, \ldots, p+n, \\ 0, & i > p+n. \end{cases}$$

So, in both cases, the Toeplitz Theorem can be applied to the E-transformation. The third condition is always satisfied and we have

Theorem 4. *Let (S_n) be a sequence converging to S.*

When k is fixed, $(E_k^{(n)})$ converges to S when n goes to infinity if and only if

$$\exists M, \ \forall n, \quad \sum_{i=0}^{k} \left|a_i^{(k,n)}\right| < M.$$

When n is fixed, $(E_k^{(n)})$ *converges to S when k goes to infinity if and only if*

$$\lim_{k\to\infty} a_i^{(k,n)} = 0, \quad i = 0, 1, \ldots,$$

$$\exists M, \ \forall k, \quad \sum_{i=0}^{k} \left|a_i^{(k,n)}\right| < M.$$

We see that it is harder to obtain convergence when n is fixed and k goes to infinity than the contrary.

As a consequence of Theorem 4, we have the

Corollary 1. *Let* (S_n) *be a sequence converging to S and let k be fixed. We assume that* $\exists\, a_0^{(k)}, \ldots, a_k^{(k)}$ *such that* $\lim_{n\to\infty} a_i^{(k,n)} = a_i^{(k)}$ *for* $i = 0, \ldots, k$. *Then* $(E_k^{(n)})$ *converges to S when n tends to infinity.*

Proof. Under the condition of existence of the $a_i^{(k)}$'s, there exists M such that, $\forall\, n$ (maybe greater than some $\bar{n}$), $\sum_{i=0}^{k} \left|a_i^{(k,n)}\right| < M$, and the result follows from the first statement of Theorem 4. □

Applying the Toeplitz Theorem to the recursive rule of the E-algorithm

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

gives

Theorem 5. *Let k be fixed. We assume that the sequence* $\left(E_{k-1}^{(n)}\right)$ *converges to S when n goes to infinity. Then* $(E_k^{(n)})$ *tends to S when n tends to infinity if and only if*

$$\exists M, \ \forall n, \quad \frac{\left|g_{k-1,k}^{(n+1)}\right| + \left|g_{k-1,k}^{(n)}\right|}{\left|g_{k-1,k}^{(n+1)} - g_{k-1,k}^{(n)}\right|} < M.$$

Obviously, the difficulty lies in finding cases where the conditions of Theorems 4 and 5, and Corollary 1 are satisfied.

Thanks to the results given at the end of Section 4.2, convergence and acceleration conditions expressed in terms of the partial derivatives of f or F [11] can be recast in terms of the $a_i^{(k,n)}$'s.

Let us now give some examples which illustrate the preceding convergence results.

The Richardson extrapolation process is a particular case of the E-algorithm. It corresponds to $g_i(n) = x_n$, where (x_n) is an auxiliary sequence. It can be seen that

$g_{k-1,k}^{(n)} = x_n \cdots x_{n+k-1}$. So, from the Toeplitz Theorem, we have the following result [39].

Theorem 6. *We assume that (S_n) converges to S. Let (x_n) be a strictly decreasing sequence of positive terms. Then, a necessary and sufficient condition that $\forall\, n$, the sequence $\left(E_k^{(n)}\right)$ converges to S when k goes to infinity is that $\exists\, M < 1$, such that $\forall\, n,\ x_{n+1}/x_n \leqslant M$.*

Under this condition, convergence also occurs for k fixed and n tending to infinity.

The Shanks transformation [51] (see also [18, pp. 78ff]) corresponds to the choice $g_i(n) = \Delta S_{n+i-1}$ in the E-algorithm. It can be implemented by the ε-algorithm [55]. Its kernel is the set of sequences such that, $\forall\, n$,

$$\sum_{i=0}^{k} a_i (S_{n+i} - S) = 0, \tag{28}$$

with $a_0 a_k \neq 0$ and $a_0 + \cdots + a_k \neq 0$. The conditions (17) are satisfied since they are, in fact, a consequence of (28). The relation (28) is similar to (26), but with the coefficients a_i independent of k and n. So, the kernel of Shanks transformation is clearly a particular case of the kernel of the E-algorithm. The difference equation (28) can be solved, thus leading to an explicit expression for the sequences belonging to the kernel of Shanks transformation (see [17]).

Aitken's Δ^2 process is the Shanks transformation for $k = 1$. Since $g_1(n) = \Delta S_n$, the system (17) gives

$$a_0^{(1,n)} = \frac{\Delta S_{n+1}}{\Delta^2 S_n}, \quad a_1^{(1,n)} = -\frac{\Delta S_n}{\Delta^2 S_n}$$

and we have

$$E_1^{(n)} = \varepsilon_2^{(n)} = \frac{S_n\, \Delta S_{n+1} - S_{n+1}\, \Delta S_n}{\Delta S_{n+1} - \Delta S_n}.$$

If $\exists\ -1 \leqslant \lambda < 1$ such that $\lim_{n\to\infty} (S_{n+1} - S)/(S_n - S) = \lambda$, then $\lim_{n\to\infty} \Delta S_{n+1}/\Delta S_n = \lambda$ [31]. Thus, when n goes to infinity, $a_0^{(1,n)}$ converges to $a_0^{(1)} = \lambda/(\lambda - 1)$ and $a_1^{(1,n)}$ tends to $a_1^{(1)} = -1/(\lambda - 1)$. Thus, the conditions of Corollary 1 are satisfied and a well-known sufficient condition for the convergence of Aitken Δ^2 process is recovered.

Since many other sequence transformations are particular cases of the E-transformation, the preceding convergence results apply.

5. Perspectives

An extension of the E-transformation for treating vector sequences was given in [6]. It is based on the solution of a system of linear equations and an extension of

Sylvester determinantal identity established in [8]. A matrix version of the transformation also exists. The elimination and annihilation operator approaches can be carried out to these systems, thus leading to the vector and matrix E-algorithm [22]. The interest of the Schur complement approach presented in this paper lies in the fact that it can be applied to the vector case as in [41], where the RPA [7] is treated, and in [43], which is devoted to the E-algorithm. The topological ε-algorithm and its variant [4], the β-algorithm [35] (see also [18, pp. 225–227]), and the H-algorithm [21], which were investigated by the extension of Sylvester identity given in [8], could easily be covered by the Schur complement approach. It also allows us to treat the new vector sequence transformations introduced in [20].

It was proved in [23] that ratios of determinants similar to those appearing in the E-transformation can be recursively computed by a triangular recursive scheme and that, reciprocally, quantities computed by such a scheme can be expressed as a ratio of determinants. Such a theory applies, in particular, to B-splines, Bernstein polynomials, orthogonal polynomials, Padé approximants, generalized divided differences, and projection methods. Thus, following the results of this paper, these items can also be interpreted as Schur complements, another application of the results given above.

Sequence transformation are also related to fixed point iterations for systems of linear and nonlinear equations [16,36], and to numerical methods for differential equations [15]. So, Schur complement techniques could also be beneficial in these domains.

Acknowledgements

We would like to thank the referees whose remarks and corrections helped us to improve the presentation of the paper.

References

[1] A.C. Aitken, Determinants and Matrices, 6th ed., Oliver and Boyd, Edinburgh and London, 1949.
[2] M.D. Benchiboun, Étude de Certaines Généralisations du Δ^2 d'Aitken et Comparaison de Procédés d'Accélération de la Convergence, Thèse de 3ième Cycle, Université des Sciences et Techniques de Lille, 1987.
[3] C. Brezinski, Résultats sur les procédés de sommation et l'ε-algorithme, R.I.R.O., 4ième année R-3 (1970) 147–153.
[4] C. Brezinski, Généralisation de la transformation de Shanks, de la table de Padé et de l'ε-algorithme, Calcolo 12 (1975) 317–360.
[5] C. Brezinski, Lecture Notes in Mathematics, Springer-Verlag, Berlin, 1977.
[6] C. Brezinski, A general extrapolation algorithm, Numer. Math. 35 (1980) 175–187.
[7] C. Brezinski, Recursive interpolation, extrapolation and projection, J. Comput. Appl. Math. 9 (1983) 369–376.

[8] C. Brezinski, Some determinantal identities in a vector space, with applications, in: H. Werner, H.J. Bünger (Eds.), Padé Approximation and its Applications, Bad Honnef 1983, Lecture Notes in Mathematics, vol. 1071, Springer-Verlag, Berlin, 1984, pp. 1–11.
[9] C. Brezinski, Convergence acceleration methods: the past decade, J. Comput. Appl. Math. 12–13 (1985) 19–36.
[10] C. Brezinski, Other manifestations of the Schur complement, Linear Algebra Appl. 111 (1988) 231–247.
[11] C. Brezinski, Quasi-linear extrapolation processes, in: R.P. Agarwal et al. (Eds.), Numerical Mathematics, Singapore 1988, ISNM vol. 86, Birkhäuser, Basel, 1988, pp. 61–78.
[12] C. Brezinski, Bordering methods and progressive forms for sequence transformations, Zastos. Mat. 20 (1990) 435–443.
[13] C. Brezinski, Algebraic properties of the E-transformation, in: A. Wakulicz (Ed.), Numerical Analysis and Mathematical Modelling, Banach Center Publications, vol. 24, PWN, Warsaw, 1990, pp. 85–90.
[14] C. Brezinski, Biorthogonality and its Applications of Numerical Analysis, Marcel Dekker, New York, 1992.
[15] C. Brezinski, Dynamical systems and sequence transformations, J. Phys. A: Math. Gen. 34 (2001) 10659–10669.
[16] C. Brezinski, A classification of quasi-Newton methods, Numer. Algorithms, in press.
[17] C. Brezinski, M. Crouzeix, Remarques sur le procédé Δ^2 d'Aitken, C. R. Acad. Sci. Paris A 270 (1970) 896–898.
[18] C. Brezinski, M. Redivo Zaglia, Extrapolation Methods: Theory and Practice, North-Holland, Amsterdam, 1991.
[19] C. Brezinski, M. Redivo Zaglia, A general extrapolation algorithm revisited, Adv. Comput. Math. 2 (1994) 461–477.
[20] C. Brezinski, M. Redivo Zaglia, New vector sequence transformations, in preparation.
[21] C. Brezinski, H. Sadok, Vector sequence transformations and fixed point methods, in: C. Taylor et al. (Eds.), Numerical Methods in Laminar and Turbulent Flows, Pineridge Press, Swansea, 1987, pp. 3–11.
[22] C. Brezinski, A. Salam, Matrix and vector sequence transformations revisited, Proc. Edinb. Math. Soc. 38 (1995) 495–510.
[23] C. Brezinski, G. Walz, Sequences of transformations and triangular recursion schemes with applications in numerical analysis, J. Comput. Appl. Math. 34 (1991) 361–383.
[24] R.A. Brualdi, H. Schneider, Determinantal identities: Gauss, Schur, Cauchy, Sylvester, Kronecker, Jacobi, Binet, Laplace, Muir, and Cayley, Linear Algebra Appl. 52/53 (1983) 769–791.
[25] F. Burns, D. Carlson, E. Haynsworth, T. Markham, Generalized inverse formulas using the Schur complement, SIAM J. Appl. Math. 26 (1974) 254–259.
[26] R.W. Cottle, Manifestations of the Schur complement, Linear Algebra Appl. 8 (1974) 189–211.
[27] R.W. Cottle, J.S. Pang, R.E. Stone, The Linear Complementarity Problem, Academic Press, New York, 1992.
[28] D.E. Crabtree, E.V. Haynsworth, An identity for the Schur complement of a matrix, Proc. Am. Math. Soc. 22 (1969) 364–366.
[29] M. Fiedler, Special Matrices and their Applications in Numerical Mathematics, Martinus Nijhoff Publishers, Dordrecht, 1986.
[30] F.R. Gantmacher, The Theory of Matrices, Chelsea Publishing Copany, New York, 1959.
[31] B. Germain-Bonne, Estimation de la Limite de Suites et Formalisation de Procédés d'Accélération de Convergence, Thèse de Doctorat ès Sciences, Université des Sciences et Techniques de Lille, 1978.
[32] T.N.E. Greville, On some conjectures of P. Wynn concerning the ε-algorithm, MRC Technical Summary Report no. 877, Mathematics Research Center, The University of Wisconsin, Madison, May 1968.

[33] A.O. Guelfond, Calcul des Différences Finies, Dunod, Paris, 1963.
[34] T. Håvie, Generalized Neville type extrapolation schemes, BIT 19 (1979) 204–213.
[35] K. Jbilou, Méthodes d'Extrapolation et de Projection. Applications aux Suites de Vecteurs, Thèse de 3ième Cycle, Université des Sciences et Technologies de Lille, 1988.
[36] K. Jbilou, H. Sadok, Some results about vector extrapolation methods and related fixed-point iterations, J. Comput. Appl. Math. 36 (1991) 385–398.
[37] H.B. Keller, The bordering algorithm and path following near singular points of higher nullity, SIAM J. Sci. Stat. Comput. 4 (1983) 573–582.
[38] V. Lakshmikantham, D. Trigiante, Theory of Difference Equations: Numerical Methods and Applications, Academic Press, San Diego, 1988.
[39] P.-J. Laurent, Une théorème de convergence pour le procédé d'extrapolation de Richardson, C. R. Acad. Sci. Paris 256 (1963) 1435–1437.
[40] G. Meinardus, G.D. Taylor, Lower estimates for the error of the best uniform approximation, J. Approx. Theory 16 (1976) 150–161.
[41] A. Messaoudi, Matrix recursive projection and interpolation algorithms, Linear Algebra Appl. 202 (1994) 71–89.
[42] A. Messaoudi, Résolution des Systèmes Linéaires par des Méthodes d'Interpolation et d'Extrapolation, Thèse d'État ès Sciences, Université Cadi Ayyad, Marrakesh, Morocco, 1996.
[43] A. Messaoudi, Matrix extrapolation algorithms, Linear Algebra Appl. 256 (1997) 49–73.
[44] L.M. Milne-Thomson, The Calculus of Finite Differences, Macmillan Press, London, 1933 (reprint by Dover, New York, 1981).
[45] N.-E. Nörlund, Leçons sur les Équations Linéaires aux Différences Finies, Gauthier-Villars, Paris, 1929.
[46] A. Ostrowski, A new proof of Haynsworth's quotient formula for Schur complements, Linear Algebra Appl. 4 (1971) 389–392.
[47] D.V. Ouellette, Schur complements and statistics, Linear Algebra Appl. 36 (1981) 187–295.
[48] A. Quarteroni, A. Valli, Domain Decomposition Methods for Partial Differential Equations, Oxford University Press, Oxford, 1999.
[49] C. Schneider, Vereinfachte Rekursionen zur Richardson—Extrapolation in Spezialfällen, Numer. Math. 24 (1975) 177–184.
[50] I. Schur, Potenzreihen im Innern des Einheitskreises, J. Reine Angew. Math. 147 (1917) 205–232.
[51] D. Shanks, Non linear transformations of divergent and slowly convergent sequences, J. Math. Phys. 34 (1955) 1–42.
[52] A.W. Tucker, A combinatorial equivalence of matrices, in: R.E. Bellman, M. Hall (Eds.), Proceedings of Symposia in Applied Mathematics, vol. 10, American Mathematical Society, Providence, 1960, pp. 129–140.
[53] J. Wimp, Sequence Transformations and their Applications, Academic Press, New York, 1981.
[54] J. Wimp, Computation with Recurrence Relations, Pitman, Boston, 1984.
[55] P. Wynn, On a device for computing the $e_m(S_n)$ transformation, MTAC 10 (1956) 91–96.
[56] F.-Z. Zhang, Matrix Theory: Basic Results and Techniques, Springer-Verlag, New York, 1999.