

# MATRIX AND VECTOR SEQUENCE TRANSFORMATIONS REVISITED

by C. BREZINSKI and A. SALAM

(Received 18th March 1994)

Sequence transformations are extrapolation methods. They are used for the purpose of convergence acceleration. In the scalar case, such algorithms can be obtained by two different approaches which are equivalent. The first one is an elimination approach based on the solution of a system of linear equations and it makes use of determinants. The second approach is based on the notion of annihilation difference operators. In this paper, these two approaches are generalized to the matrix and the vector cases.

1991 *Mathematics subject classification*: 65B05.

## 0. Introduction

There exist many algorithms for transforming a sequence of numbers, or a sequence of vectors, or a sequence of matrices into a new sequence of objects of the same type. Such sequence transformations are used for accelerating the convergence of the initial sequence. They are often useful, and even essential, since many sequences and many iterative processes used in numerical analysis and in applied mathematics converge so slowly that their practical use is very limited. Sequence transformations are based on the notion of extrapolation as explained, for example, in [5, 9, 21] and their theory is now fully understood. Quite often, the elements of the new sequence obtained by an extrapolation method are expressed as a ratio of two determinants since the elements of the new sequence are, in fact, expressed as the solutions of systems of linear equations. Of course, a numerical analyst would avoid computing the value of a determinant (because there are too many arithmetical operations and then too many rounding errors), and prefer to have recursive algorithms at his disposal for computing these ratios of determinants. The equivalence between ratios of determinants and triangular recursive schemes was studied in [8]. As shown in [4] and in [7], acceleration algorithms can also be derived by means of difference operators, an approach first introduced by Weniger [20]. In this paper, we shall extend these two procedures to the vector and the matrix cases. In the first section, we shall begin by recalling the main points of these two approaches.

## 1. The scalar case

Let  $(S_n)$  be a sequence of complex numbers such that,  $\forall n$

$$S_n - S = aD_n \tag{1}$$

where  $S$  and  $a$  are unknown numbers and  $(D_n)$  a known sequence. We shall try to compute the value of  $S$ . Writing (1) for the index  $n$  and for the index  $n+1$  and subtracting the first equation from the second one, we obtain the system

$$\left. \begin{aligned} S_n &= S + aD_n \\ \Delta S_n &= 0 + a\Delta S_n \end{aligned} \right\}$$

Assuming that  $\Delta D_n = D_{n+1} - D_n \neq 0$ , we have

$$S = \frac{\begin{vmatrix} S_n & D_n \\ \Delta S_n & \Delta D_n \end{vmatrix}}{\begin{vmatrix} 1 & D_n \\ 0 & \Delta D_n \end{vmatrix}} = S_n - \frac{\Delta S_n}{\Delta D_n} D_n.$$

This approach will be called the *elimination approach*.

The second approach makes use of difference operators. Although several authors used this idea before, it really originated in [20]. It was fully developed and exploited in [4]. It allows to recover known results, and to obtain new ones, without using the notion of determinant. Such an approach was followed in [7] in the case of the  $E$ -algorithm which is the most general extrapolation algorithm actually known. As above, we assume that (1) holds and that  $\forall n, D_n \neq 0$ . Thus, (1) can be written as

$$\frac{S_n}{D_n} - \frac{S}{D_n} = a.$$

Since  $a$  is a constant, the difference operator  $\Delta$  is such that  $\Delta a = 0$ . A linear operator  $P$  such that  $\forall n, Pa_n = 0$  is called an annihilation difference operator for the sequence  $(a_n)$ . Thus  $\Delta$  is an annihilation operator for the constant sequence  $(a_n = a)$ . Applying  $\Delta$  to the preceding relation yields

$$\Delta(S_n/D_n) - S\Delta(1/D_n) = \Delta a = 0.$$

It follows that

$$S = \frac{\Delta(S_n/D_n)}{\Delta(1/D_n)}.$$

This approach will be called the *annihilation operator approach*.

Let us now show that the elimination and the annihilation operator approaches are equivalent in the scalar case. Using the Leibniz rule for the operator  $\Delta$

$$\Delta(a_n b_n) = b_{n+1} \Delta a_n + a_n \Delta b_n,$$

we have, starting from the second approach

$$\Delta(S_n/D_n) = \Delta S_n/D_{n+1} + S_n\Delta(1/D_n).$$

Then

$$\frac{\Delta(S_n/D_n)}{\Delta(1/D_n)} = S_n + \frac{\Delta S_n}{D_{n+1}\Delta(1/D_n)} = S_n - \frac{\Delta S_n}{\Delta D_n} D_n$$

since  $D_{n+1}(1/D_n) = D_{n+1}(1/D_{n+1} - 1/D_n) = -\Delta D_n/D_n$ . Thus, the annihilation operator approach has been recovered. Conversely, in this approach, we can write

$$S = \frac{\begin{vmatrix} S_n/D_n & 1 \\ S_{n+1}/D_{n+1} & 1 \end{vmatrix}}{\begin{vmatrix} 1/D_n & 1 \\ 1/D_{n+1} & 1 \end{vmatrix}} = \frac{\begin{vmatrix} S_n & D_n \\ S_{n+1} & D_{n+1} \end{vmatrix}}{\begin{vmatrix} 1 & D_n \\ 1 & D_{n+1} \end{vmatrix}} = \frac{\begin{vmatrix} S_n & D_n \\ \Delta S_n & \Delta D_n \end{vmatrix}}{\begin{vmatrix} 1 & D_n \\ 0 & \Delta D_n \end{vmatrix}}$$

which is the expression obtained by the elimination approach.

Obviously, in both approaches, if  $(S_n)$  does not satisfy (1), we can define the sequence transformation  $T:(S_n) \mapsto (T_n)$  by

$$T_n = S_n - \frac{\Delta S_n}{\Delta D_n} D_n \quad \text{for } n=0, 1, \dots$$

This is the so-called  $\Theta$ -procedure introduced in [2] (see also [6]) whose convergence and acceleration properties have been studied. By construction,  $T_n = S_n, \forall n$  if  $\forall n$ , (1) holds.

In Section 2, the two approaches defined above will be generalized to the matrix case, while the vector case will be considered in Section 3.

Of course, the form considered in equation (1) is the simplest one. It can be generalized by assuming that the sequence  $(S_n)$  satisfies,  $\forall n$

$$S_n = S + a_1 g_1(n) + \dots + a_k g_k(n) \tag{2}$$

where the  $(g_i(n))$ 's are given known auxiliary sequences which can depend on the sequence  $(S_n)$  itself. This case, which is the most general one to have been treated so far, leads to the  $E$ -algorithm of Håvie [12] and Brezinski [1]. Using the elimination and the annihilation difference operator approaches, it will be generalized to matrices in Section 4 and to vectors in Section 5.

**2. The matrix case**

Let now  $(S_n)$  be a sequence of  $p \times q$  matrices,  $(D_n)$  a sequence of  $q \times q$  matrices, and  $A$  and  $S$  two unknown  $p \times q$  matrices. We assume that,  $\forall n$

$$S_n = S + AD_n. \tag{3}$$

As before, the problem consists in computing  $S$ . Then, a sequence transformation  $T:(S_n) \mapsto (T_n)$  could be defined as above and its convergence and acceleration properties could be studied. By construction, we shall have  $T_n = S, \forall n$  if (3) holds,  $\forall n$ .

**2.1. The elimination approach**

In the scalar case, the elimination approach is based on the solution of a system of linear equations with scalar coefficients and it makes use of the notion of determinant. For systems of linear equations with elements in a non-commutative algebra, this notion no longer exists [10, 11, 15] and it has to be replaced by a more general one. In this case, the corresponding object is that of *designant*. This notion was introduced by Heyting [14] in 1927 and it was recently extensively used by Salam [17, 18, 19] in connection with vector sequence acceleration methods and, in particular, with the vector  $\epsilon$ -algorithm which was proved to correspond to a ratio of designants instead of a ratio of determinants as for the scalar case.

Thus, let us first remind the definition of designants and give some of their properties that will be useful later.

Let  $\mathcal{M}_q$  be the algebra of  $q \times q$  matrices with complex entries and  $\mathcal{M}_{p,q}$  the set of  $p \times q$  matrices with complex entries. We consider the system of linear equations

$$\left. \begin{aligned} x_1 a_{11} + x_2 a_{12} &= b_1 \\ x_1 a_{21} + x_2 a_{22} &= b_2 \end{aligned} \right\} \tag{4}$$

where  $a_{ij} \in \mathcal{M}_q, b_i \in \mathcal{M}_{p,q}$  and the unknown matrices  $x_i$  belong to  $\mathcal{M}_{p,q}$ . Let us remark that, in this system, the unknowns are multiplied on the right by the coefficients.

Let us assume that the matrix  $a_{11}$  is regular. For eliminating the first unknown  $x_1$  from the first equation of (4), let us proceed as in Gaussian elimination, that is, let us multiply it by  $a_{11}^{-1}$  on the right, then by  $a_{21}$  and finally subtract it from the second equation. We obtain

$$x_2(a_{22} - a_{12}a_{11}^{-1}a_{21}) = b_2 - b_1a_{11}^{-1}a_{21}.$$

The square matrix  $a_{22} - a_{12}a_{11}^{-1}a_{21}$  is called the *designant* (more precisely, the right-designant) of the system (4) and it will be denoted by

$$\left| \begin{array}{cc} a_{11} & a_{12} \\ a_{21} & a_{22} \end{array} \right|_r$$

where the subscript  $r$  stands for ‘‘right’’.

If the designant is regular, then

$$x_2 = (b_2 - b_1a_{11}^{-1}a_{21})(a_{22} - a_{12}a_{11}^{-1}a_{21})^{-1}$$

that is, using the notation introduced above for designants

$$x_2 = \left| \begin{array}{cc|cc} a_{11} & b_1 & a_{11} & a_{12} \\ a_{21} & b_2 & a_{21} & a_{22} \end{array} \right|_r^{-1}.$$

Similarly, the second unknown matrix is given by

$$x_1 = \left| \begin{array}{cc|cc} b_1 & a_{12} & a_{11} & a_{12} \\ b_2 & a_{22} & a_{21} & a_{22} \end{array} \right|_r^{-1}.$$

**-Remark.** The notion of designant is related to that of Schur complement. Indeed, the Schur complement  $(M/a_{11})$  of  $a_{11}$  in the matrix

$$M = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

is the designant whose inverse appears in the expressions of  $x_1$  and  $x_2$ . The Schur complement was already proved to be very much related to scalar and vector sequence transformations [3].

The notion of designant can be generalized to systems in a non-commutative algebra with an arbitrary number of equations and unknowns. Let us consider the system

$$\left. \begin{array}{l} x_1 a_{11} + \dots + x_n a_{1n} = b_1 \\ \vdots \\ x_1 a_{n1} + \dots + x_n a_{nn} = b_n \end{array} \right\} \tag{5}$$

as a generalization of (4). We have

$$x_n = \left| \begin{array}{ccc|cc} a_{11} & \dots & a_{1,n-1} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{n1} & \dots & a_{n,n-1} & b_n \end{array} \right|_r \left| \begin{array}{ccc} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{array} \right|_r^{-1}.$$

Let us set

$$\Delta_n = \left| \begin{array}{ccc} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{array} \right|_r$$

and let  $A_{r,s}^{(n-2)}$  be the designant of order  $n-1$  obtained by keeping the rows  $1, 2, \dots, n-2, r$  and the columns  $1, 2, \dots, n-2, s$  of  $\Delta_n$ .  $\Delta_n$  can be written as a designant and computed from  $A_{n-1,n-1}^{(n-2)}, A_{n-1,n}^{(n-2)}, A_{n,n-1}^{(n-2)}$  and  $A_{n,n}^{(n-2)}$  by using Sylvester's identity [14]

$$\Delta_n = \begin{vmatrix} A_{n-1,n-1}^{(n-2)} & A_{n-1,n}^{(n-2)} \\ A_{n,n-1}^{(n-2)} & A_{n,n}^{(n-2)} \end{vmatrix}_r = A_{n,n}^{(n-2)} - A_{n-1,n}^{(n-2)} [A_{n-1,n-1}^{(n-2)}]^{-1} A_{n,n-1}^{(n-2)}.$$

More information about designants can be found in [14, 17, 18].

We shall now use these notions for computing  $S$  when the sequence  $(S_n)$  satisfies (3). Writing this relation for the indices  $n$  and  $n + 1$  leads to the system

$$\left. \begin{aligned} S + AD_n &= S_n \\ 0 + A\Delta D_n &= \Delta S_n \end{aligned} \right\} \tag{6}$$

If the designant of the system is regular, we obtain

$$S = \begin{vmatrix} D_n & S_n \\ \Delta D_n & \Delta S_n \end{vmatrix}_r \begin{vmatrix} D_n & I \\ \Delta D_n & 0 \end{vmatrix}_r^{-1} \tag{7}$$

Another procedure for computing  $S$  is first to obtain  $A$  from the second equation of the system (6), assuming that  $\Delta D_n$  is regular, and then to replace it in the first equation, thus leading to

$$S = S_n - \Delta S_n (\Delta D_n)^{-1} D_n. \tag{8}$$

The relations (7) and (8) are equivalent since, from (7)

$$\begin{aligned} S &= (\Delta S_n - S_n D_n^{-1} \Delta D_n) (0 - I \cdot D_n^{-1} \Delta D_n)^{-1} \\ &= (-\Delta S_n + S_n D_n^{-1} \Delta D_n) (\Delta D_n)^{-1} D_n \\ &= -\Delta S_n (\Delta D_n)^{-1} D_n + S_n \end{aligned}$$

which is the relation (8).

Before ending this subsection, let us mention that designants corresponding to a system of linear equations where the unknowns in (4) and (5) are multiplied on the left by the coefficients can also be defined and that the results obtained are very similar to those given above.

### 2.2. The annihilation operator approach

Let us now treat the matrix case along the lines proposed in [4]. Assuming that the square  $q \times q$  matrix  $D_n$  is regular, we have  $S_n D_n^{-1} - S D_n^{-1} = A$ . Applying the difference operator  $\Delta$  to both sides leads to

$$\Delta(S_n D_n^{-1}) - S \Delta(D_n^{-1}) = 0$$

that is

$$S = \Delta(S_n D_n^{-1})(\Delta(D_n^{-1}))^{-1}. \tag{9}$$

This approach is equivalent to the elimination approach explained in the preceding subsection. Indeed, we have  $\Delta(S_n D_n^{-1}) = (\Delta S_n) D_{n+1}^{-1} + S_n \Delta(D_n^{-1})$ . It follows that

$$\begin{aligned} S &= S_n + \Delta S_n D_{n+1}^{-1} (\Delta D_n^{-1})^{-1} \\ &= S_n + \Delta S_n (\Delta(D_n^{-1}) D_{n+1})^{-1} \\ &= S_n + \Delta S_n (D_n^{-1} (D_n - D_{n+1}))^{-1} \\ &= S_n - \Delta S_n (\Delta D_n)^{-1} D_n \end{aligned}$$

which is the relation (8).

It is possible to obtain the same results using the notion of designant. (9) can be written as

$$S = \left| \begin{array}{cc|c} I & S_n D_n^{-1} & I \\ I & S_{n+1} D_{n+1}^{-1} & I \end{array} \right|_r \left| \begin{array}{c|c} I & D_n^{-1} \\ I & D_{n+1}^{-1} \end{array} \right|_r^{-1}.$$

Right multiplying the first row by  $D_n$  and the second row by  $D_{n+1}$  does not change the designants and we obtain

$$S = \left| \begin{array}{cc|c} D_n & S_n & D_n \\ D_{n+1} & S_{n+1} & D_{n+1} \end{array} \right|_r \left| \begin{array}{c|c} D_n & I \\ D_{n+1} & I \end{array} \right|_r^{-1} = \left| \begin{array}{cc|c} D_n & S_n & D_n \\ \Delta D_n & \Delta S_n & \Delta D_n \end{array} \right|_r \left| \begin{array}{c|c} D_n & I \\ \Delta D_n & 0 \end{array} \right|_r^{-1}.$$

which is exactly (7).

### 3. The vector case

We shall now apply the techniques used in the matrix case to the vector case. Let  $(S_n)$  be a sequence of vectors of  $\mathbb{C}^p$ , let  $A \in \mathcal{M}_{p,q}$  and let  $(D_n)$  be a sequence of vectors of  $\mathbb{C}^q$ . We assume that,  $\forall n$

$$S_n = S + A D_n. \tag{10}$$

As above, the problem consists in computing the unknown vector  $S \in \mathbb{C}^p$ . The idea will be to transform this vector problem into a matrix one and then to use the approaches of the previous section.

#### 3.1. The elimination approach

We shall make use of the following notation: let  $a_n$  be any sequence of vectors of

dimension  $m$ , we shall denote by  $[a_n]$  the  $m \times q$  matrix whose columns are  $a_n, \dots, a_{n+q-1}$ . In the sequel, we shall take  $m=p$  or  $m=q$ . Thus, such matrices will be square for  $m=q$ .

Writing (10) for the indices  $n, \dots, n+q-1$ , we obtain

$$[S] = [S_n] - [\Delta S_n][\Delta D_n]^{-1}[D_n].$$

The vector case is then recovered by considering the first column of each side of this relation, that is

$$S = S_n - [\Delta S_n][\Delta D_n]^{-1}D_n. \tag{11}$$

The transformation  $H : (S_n) \mapsto (H_n)$  where  $H_n$  is the right hand side of (11) generalizes Henrici's transformation [13, formula 5.35, p. 116] which was studied in details by Sadok [16]. Obviously, it is also related to the Schur complement. Indeed, if we set

$$M_n = \begin{pmatrix} S_n & [\Delta S_n] \\ D_n & [\Delta D_n] \end{pmatrix}$$

we have  $(M_n / [\Delta D_n]) = S_n - [\Delta S_n][\Delta D_n]^{-1}D_n = H_n$ .

If  $H_n$  is a square matrix, which in our case can only arise if  $p=1$ , then  $\det M_n = H_n \det [\Delta D_n]$  that is

$$H_n = \left| \begin{array}{cc} S_n & [\Delta S_n] \\ D_n & [\Delta D_n] \end{array} \right|_r / \left| [\Delta D_n] \right|_r$$

which is nothing else than the  $E$ -algorithm [1, 12] since

$$H_n = \left| \begin{array}{cccc} S_n & \Delta S_n & \dots & \Delta S_{n+q-1} \\ g_1(n) & \Delta g_1(n) & \dots & \Delta g_1(n+q-1) \\ \vdots & \vdots & \dots & \vdots \\ g_q(n) & \Delta g_q(n) & \dots & \Delta g_q(n+q-1) \end{array} \right| \left/ \begin{array}{ccc} \Delta g_1(n) & \dots & \Delta g_1(n+q-1) \\ \vdots & & \vdots \\ \Delta g_q(n) & \dots & \Delta g_q(n+q-1) \end{array} \right|$$

### 3.2. The annihilation operator approach

We shall now treat the vector case by the annihilation operator approach. Using the same notations as in the elimination approach we have, from (10),  $[S_n] - [S] = A[D_n]$ . It follows

$$\begin{aligned} [S] &= \Delta([S_n][D_n]^{-1})(\Delta[D_n]^{-1})^{-1} \\ &= [S_n] - [\Delta S_n][\Delta D_n]^{-1}[D_n]. \end{aligned}$$



Keeping only the first columns of the matrices in both sides, we obtain

$$S = S_n - [\Delta S_n][\Delta D_n]^{-1} D_n$$

which shows that the two approaches are equivalent in the vector case.

**4. The general matrix case**

We shall now extend the general scalar relation (2) to the matrix case. More precisely we assume that,  $\forall n$

$$S_n = S + A_1 g_1(n) + A_2 g_2(n) + \dots + A_k g_k(n) \tag{12}$$

where  $S_n, S, A_i \in \mathcal{M}_{p,q}$  and where  $S_n$  and the  $k$  matrices  $g_i(n) \in \mathcal{M}_q$  are assumed to be known.

**4.1 The elimination approach**

$S$  is obtained by solving the system

$$\begin{array}{rcl} S_n & = & S + A_1 g_1(n) + \dots + A_k g_k(n) \\ \dots & \dots & \dots \\ S_{n+k} & = & S + A_1 g_1(n+k) + \dots + A_k g_k(n+k). \end{array}$$

If the designant of this system is non-singular, then

$$S = \left| \begin{array}{cccc} g_1(n) & \dots & g_k(n) & S_n \\ \vdots & & \vdots & \vdots \\ g_1(n+k) & \dots & g_k(n+k) & S_{n+k} \end{array} \right|_r \left| \begin{array}{cccc} g_1(n) & \dots & g_k(n) & I \\ \vdots & & \vdots & \vdots \\ g_1(n+k) & \dots & g_k(n+k) & I \end{array} \right|_r^{-1} \tag{13}$$

If the sequence  $(S_n)$  does not satisfy (12), the expression in the right hand side of (13) will depend on the indices  $k$  and  $n$ . In that case, it will be denoted by  $E_k^{(n)}$ . It is a matrix belonging to  $\mathcal{M}_{p,q}$ . We shall denote by  $g_{k,i}^{(n)}$  the  $q \times q$  matrix obtained by replacing the last column in the numerator of  $E_k^{(n)}$  by  $g_i(n), \dots, g_i(n+k)$ . Obviously  $g_{k,i}^{(n)} = 0$  for  $i \leq k$ . These matrices can be recursively computed by a matrix  $E$ -algorithm generalizing the scalar one [1, 12]. It is as follows (the operator  $\Delta$  acting on the upper index  $n$ )

**Theorem 1.**

$$\begin{aligned} E_{k+1}^{(n)} &= E_k^{(n)} - \Delta E_k^{(n)} (\Delta g_{k,k+1}^{(n)})^{-1} g_{k,k+1}^{(n)} \\ g_{k+1,i}^{(n)} &= g_{k,i}^{(n)} - \Delta g_{k,i}^{(n)} (\Delta g_{k,k+1}^{(n)})^{-1} g_{k,k+1}^{(n)} \end{aligned}$$

with  $E_0^{(n)} = S_n$  and  $g_{0,i}^{(n)} = g_i(n)$ .

**Proof.** We have

$$E_k^{(n)} = \left| \begin{array}{cccc} g_1(n) & \dots & g_k(n) & S_n \\ \vdots & & \vdots & \vdots \\ g_1(n+k) & \dots & g_k(n+k) & S_{n+k} \end{array} \right|_r \left| \begin{array}{cccc} g_1(n) & \dots & g_k(n) & I \\ \vdots & & \vdots & \vdots \\ g_1(n+k) & \dots & g_k(n+k) & I \end{array} \right|_r^{-1}$$

Let us set

$$g_{k,i}^{(n)} = \left| \begin{array}{cccc} g_1(n) & \dots & g_k(n) & g_i(n) \\ \vdots & & \vdots & \vdots \\ g_1(n+k) & \dots & g_k(n+k) & g_i(n+k) \end{array} \right|_r \left| \begin{array}{cccc} g_1(n) & \dots & g_k(n) & I \\ \vdots & & \vdots & \vdots \\ g_1(n+k) & \dots & g_k(n+k) & I \end{array} \right|_r^{-1}$$

We shall prove that these quantities can be recursively computed by the relations given in the theorem. Let us denote, for simplicity, by  $A$  the numerator of  $E_k^{(n)}$  and by  $D$  its denominator. In other words, we write  $E_k^{(n)} = AD^{-1}$ . The value of the designant  $A$  does not change if its first row is permuted with one of the first  $k-1$  ones [14]. Thus

$$A = \left| \begin{array}{cccc} g_1(n+1) & \dots & g_k(n+1) & S_{n+1} \\ \vdots & & \vdots & \vdots \\ g_1(n+k-1) & \dots & g_k(n+k-1) & S_{n+k-1} \\ g_1(n) & \dots & g_k(n) & S_n \\ g_1(n+k) & \dots & g_k(n+k) & S_{n+k} \end{array} \right|_r$$

Applying Sylvester's identity, extended to designants [17], we obtain

$$A = \left| \begin{array}{cccc} g_1(n+1) & \dots & g_{k-1}(n+1) & S_{n+1} \\ \vdots & & \vdots & \vdots \\ g_1(n+k) & \dots & g_{k-1}(n+k) & S_{n+k} \end{array} \right|_r - CJ$$

with

$$J = \left| \begin{array}{ccc} g_1(n+1) & \dots & g_k(n+1) \\ \vdots & & \vdots \\ g_1(n+k) & \dots & g_k(n+k) \end{array} \right|_r$$

and

$$C = \left| \begin{array}{cccc} g_1(n+1) & \dots & g_{k-1}(n+1) & S_{n+1} \\ \vdots & & \vdots & \vdots \\ g_1(n+k-1) & \dots & g_{k-1}(n+k-1) & S_{n+k-1} \\ g_1(n) & \dots & g_{k-1}(n) & S_n \end{array} \right|_r \left| \begin{array}{cccc} g_1(n+1) & \dots & g_k(n+1) & I \\ \vdots & & \vdots & \vdots \\ g_1(n+k-1) & \dots & g_k(n+k-1) & I \\ g_1(n) & \dots & g_k(n) & I \end{array} \right|_r^{-1}$$

From a property of designants [14, 17] we have

$$C = \left| \begin{array}{cccc} g_1(n) & \dots & g_{k-1}(n) & S_n \\ g_1(n+1) & \dots & g_{k-1}(n+1) & S_{n+1} \\ \vdots & & \vdots & \vdots \\ g_1(n+k-1) & \dots & g_{k-1}(n+k-1) & S_{n+k-1} \end{array} \right|_r \left| \begin{array}{ccc} g_1(n) & \dots & g_k(n) \\ g_1(n+1) & \dots & g_k(n+1) \\ \vdots & & \vdots \\ g_1(n+k-1) & \dots & g_k(n+k-1) \end{array} \right|_r^{-1}$$

Thus

$$B = E_{k-1}^{(n+1)}(g_{k-1}^{(n+1)})^{-1}, \quad C = E_{k-1}^{(n)}(g_{k-1}^{(n)})^{-1},$$

$$AJ^{-1} = B - C = E_{k-1}^{(n+1)}(g_{k-1}^{(n+1)})^{-1} - E_{k-1}^{(n)}(g_{k-1}^{(n)})^{-1}.$$

We have

$$D = \left| \begin{array}{ccc|c} g_1(n+1) & \dots & g_k(n+1) & I \\ \vdots & & \vdots & \vdots \\ g_1(n+k-1) & \dots & g_k(n+k-1) & I \\ g_1(n) & \dots & g_k(n) & I \\ g_1(n+k) & \dots & g_k(n+k) & I \end{array} \right|_r$$

Applying Sylvester's identity to  $D$  we obtain

$$DJ^{-1} = E - F$$

with

$$E = \left| \begin{array}{ccc|c} g_1(n+1) & \dots & g_{k-1}(n+1) & I \\ \vdots & & \vdots & \vdots \\ g_1(n+k) & \dots & g_{k-1}(n+k) & I \end{array} \right|_r \left| \begin{array}{ccc} g_1(n+1) & \dots & g_k(n+1) \\ \vdots & & \vdots \\ g_1(n+k) & \dots & g_k(n+k) \end{array} \right|_r^{-1} = (g_{k-1}^{(n+1)})^{-1}$$

and

$$F = \left| \begin{array}{ccc|c} g_1(n+1) & \dots & g_{k-1}(n+1) & I \\ \vdots & & \vdots & \vdots \\ g_1(n+k-1) & \dots & g_{k-1}(n+k-1) & I \\ g_1(n) & \dots & g_{k-1}(n) & I \end{array} \right|_r \left| \begin{array}{ccc} g_1(n+1) & \dots & g_k(n+1) \\ \vdots & & \vdots \\ g_1(n+k-1) & \dots & g_k(n+k-1) \\ g_1(n) & \dots & g_k(n) \end{array} \right|_r^{-1}$$

It is known [14] that a designant changes if its last row is permuted with another one but that the product of the two designants appearing in the definition of  $F$  does not change when permuting the last rows of both of them with the same other row. Thus, it follows that

$$F = \begin{pmatrix} g_1(n) & \dots & g_{k-1}(n) & I \\ g_1(n+1) & \dots & g_{k-1}(n+1) & I \\ \vdots & & \vdots & \vdots \\ g_1(n+k-1) & \dots & g_{k-1}(n+k-1) & I \end{pmatrix}_r \begin{pmatrix} g_1(n) & \dots & g_k(n) \\ g_1(n+1) & \dots & g_k(n+1) \\ \vdots & & \vdots \\ g_1(n+k-1) & \dots & g_k(n+k-1) \end{pmatrix}_r^{-1}$$

$$= (g_{k-1,k}^{(n)})^{-1}.$$

Finally we have

$$E_k^{(n)} = AD^{-1} = AJ^{-1}JD^{-1} = AJ^{-1}(DJ^{-1})^{-1}$$

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

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

which is the first rule of the algorithm given in the theorem. The proof of the rule for the  $g_{k,i}^{(n)}$ 's is the same after replacing the  $S_n$ 's by the  $g_i(n)$ 's. The algorithm is initialized with

$$E_0^{(n)} = S_n \quad \text{and} \quad g_{0,i}^{(n)} = g_i(n). \quad \square$$

**4.2. The annihilation operator approach**

We shall now generalize to the matrix case the operator approach introduced in [7] for the scalar  $E$ -algorithm. For any matrix sequence  $U = (U_n)$  we shall define recursively the operators  $N_k: U = (U_n)_n \mapsto N_k(U) = (N_k^{(n)}(U))_n$  by

$$N_{k+1}^{(n)}(U) = \Delta(N_k^{(n)}(U)(N_k^{(n)}(g_{k+1}))^{-1})$$

with  $N_0^{(n)}(U) = U_n$ .

We set

$$E_k^{(n)}(U) = N_k^{(n)}(U)(N_k^{(n)}(I))^{-1}.$$

For simplicity we shall also make use of the notation  $N_k(U_n)$  for denoting  $N_k^{(n)}(U)$  and  $E_k(U_n)$  for denoting  $E_k^{(n)}(U)$ . Sometimes both notations will be used simultaneously.

If  $A \in \mathcal{M}_{p,q}$  and  $B \in \mathcal{M}_q$  with  $B$  regular, the product  $AB^{-1}$  will be denoted by  $A/B$ . Thus, (12) can be written as

$$N_0(S_n) - SN_0^{(n)}(I) = A_1N_0(g_1(n)) + \dots + A_kN_0(g_k(n)). \tag{14}$$

Assuming that  $N_0(g_1(n))$  is regular, we can multiply (14) on the right by its inverse and then apply the operator  $\Delta$  to both sides. Since  $\Delta A_1 = 0$ , we obtain, using the definition of  $N_1$

$$N_1(S_n) - SN_1^{(n)}(I) = A_2N_1(g_2(n)) + \dots + A_kN_1(g_k(n)) \tag{15}$$

and

$$E_1(S_n) = N_1(S_n)/N_1^{(n)}(I).$$

Assuming that  $N_1(g_2(n))$  is regular, multiplying (15) on the right by its inverse, applying  $\Delta$ , and using the definition of  $N_2$ , we get

$$N_2(S_n) - SN_2^{(n)}(I) = A_3N_2(g_3(n)) + \cdots + A_kN_2(g_k(n))$$

and

$$E_2(S_n) = N_2(S_n)/N_2^{(n)}(I).$$

And so on, we have

$$\begin{aligned} N_{k+1}(S_n) &= \Delta(N_k(S_n)(N_k(g_{k+1}(n)))^{-1}) \\ &= \Delta\left(\frac{N_k(S_n)}{N_k^{(n)}(I)}\left(\frac{N_k(g_{k+1}(n))}{N_k^{(n)}(I)}\right)^{-1}\right) \\ &= \Delta(E_k(S_n)(E_k(g_{k+1}(n)))^{-1}) \\ &= \Delta\left(\frac{E_k(S_n)}{E_k(g_{k+1}(n))}\right). \end{aligned}$$

Similarly

$$\begin{aligned} N_{k+1}^{(n)}(I) &= \Delta(N_k^{(n)}(I)(N_k(g_{k+1}(n)))^{-1}) \\ &= \Delta\left(\frac{E_k^{(n)}(I)}{E_k(g_{k+1}(n))}\right). \end{aligned}$$

Since  $E_k^{(n)}(I) = I$ , we finally obtain the

**Theorem 2.**

$$\begin{aligned} E_{k+1}(S_n) &= \Delta\left(\frac{E_k(S_n)}{E_k(g_{k+1}(n))}\right) / \Delta\left(\frac{I}{E_k(g_{k+1}(n))}\right) \\ E_{k+1}(g_i(n)) &= \Delta\left(\frac{E_k(g_i(n))}{E_k(g_{k+1}(n))}\right) / \Delta\left(\frac{I}{E_k(g_{k+1}(n))}\right). \end{aligned}$$

These expressions are exactly those obtained in the elimination approach. Thus, the  $E_k(S_n)$  and the  $E_k(g_i(n))$  coincide with the  $E_k^{(n)}$  and the  $g_{k,i}^{(n)}$  given in the previous subsection.

### 5. The general vector case

We shall assume that the sequence of vectors  $S_n \in \mathbb{C}^p$  satisfies,  $\forall n$

$$S_n = S + A_1 g_1(n) + A_2 g_2(n) + \cdots + A_k g_k(n) \quad (16)$$

where now  $A_i \in \mathcal{M}_{p,q}$ ,  $S \in \mathbb{C}^p$  and  $g_i(n) \in \mathbb{C}^q$ .

The problem is again to compute  $S$ . We shall follow successively the two approaches used in the general matrix case.

#### 5.1. The elimination approach

Transforming the vector relation (16) into a matrix one, we obtain

$$[S_n] = [S] + A_1 [g_1(n)] + \cdots + A_k [g_k(n)].$$

Using the matrix  $E$ -algorithm given in the theorem above and keeping only the first columns in each side of the relation, we obtain the following algorithm

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

with  $E_0^{(n)} = S_n$  and  $g_{0,i}^{(n)} = g_i(n)$ . This algorithm could be called the *matrix-vector E-algorithm* since it is a vector algorithm coming out from the matrix  $E$ -algorithm given in the previous section.

In this algorithm, the recursive computation of  $E_k^{(n)}$  needs the knowledge of  $S_n, \dots, S_{n+qk}$  (and similarly for  $g_{k,i}^{(n)}$ ). In the topological  $E$ -algorithm proposed in [1], this computation only needs  $S_n, \dots, S_{n+2k}$ . However, for the topological  $E$ -algorithm, the  $A_i$ 's must be scalar matrices  $A_i = a_i I$  while, now, they can be arbitrary  $p \times q$  matrices.

#### 5.2. The annihilation operator approach

Let  $E_k[S_n]$  be the  $p \times q$  matrix whose columns are  $E_k(S_n), \dots, E_k(S_{n+q-1})$  and similarly for  $E_k[g_i(n)]$ . After transforming, as above, the vector case into a matrix relation, we obtain

$$E_{k+1}[S_n] = E_k[S_n] - \Delta E_k[S_n](\Delta E_k[g_{k+1}(n)])^{-1} E_k[g_{k+1}(n)].$$

Keeping only the first columns of the matrices in both sides, we get

$$E_{k+1}(S_n) = E_k(S_n) - \Delta E_k[S_n](\Delta E_k[g_{k+1}(n)])^{-1} E_k(g_{k+1}(n))$$

and

$$E_{k+1}(g_i(n)) = E_k(g_i(n)) - \Delta E_k[g_i(n)](\Delta E_k[g_{k+1}(n)])^{-1} E_k(g_{k+1}(n))$$

with  $E_0(S_n) = S_n$  and  $E_0(g_i(n)) = g_i(n)$ .

Thus, the elimination and annihilation operator approaches are equivalent.

**Acknowledgements.** We would like to thank the referees for several improvements of the paper.

#### REFERENCES

1. C. BREZINSKI, A general extrapolation algorithm, *Numer. Math.* **35** (1980), 175–187.
2. C. BREZINSKI, Some new convergence acceleration methods, *Math. Comput.* **39** (1982), 133–145.
3. C. BREZINSKI, Other manifestations of the Schur complement, *Linear Algebra Appl.* **111** (1988), 231–247.
4. C. BREZINSKI and A. C. MATOS, A derivation of extrapolation algorithms based on error estimates, *J. Comput. Appl. Math.*, to appear.
5. C. BREZINSKI and M. REDIVO ZAGLIA, *Extrapolation Methods. Theory and Practice* (North-Holland, Amsterdam, 1991).
6. C. BREZINSKI and M. REDIVO ZAGLIA, Construction of extrapolation processes, *Appl. Numer. Math.* **8** (1991), 11–23.
7. C. BREZINSKI and M. REDIVO ZAGLIA, A general extrapolation procedure revisited, *Adv. Comput. Math.* **2** (1994), 461–477.
8. C. BREZINSKI and G. WALZ, Sequences of transformations and triangular recursion schemes with applications in numerical analysis, *J. Comput. Appl. Math.* **34** (1991), 361–383.
9. J. P. DELAHAYE, *Sequence Transformations* (Springer-Verlag, Berlin, 1988).
10. J. DIEUDONNÉ, Les déterminants sur un corps non commutatif, *Bull. Soc. Math. France* **7** (1943), 27–45.
11. F. J. DYSON, Quaternion determinants, *Helv. Phys. Acta* **45** (1972), 289–302.
12. T. HAVIE, Generalized Neville type extrapolation schemes, *BIT* **19** (1979), 204–213.
13. P. HENRICI, *Elements of Numerical Analysis* (Wiley, New York, 1964).
14. A. HEYTING, Die Theorie der linearen Gleichungen in einer Zahlenspezies mit nicht-kommutativer Multiplikation, *Math. Ann.* **98** (1927), 465–490.
15. M. L. MEHTA, *Matrix Theory, Selected Topics and Useful Results* (Les Editions de Physique, Les Ullis, 1989).
16. H. SADOK, About Henrici's transformation for accelerating vector sequences, *J. Comput. Appl. Math.* **29** (1990), 101–110.
17. A. SALAM, *Extrapolation: Extension et Nouveaux Résultats* (Thèse, Université des Sciences et Technologies de Lille, 1993).
18. A. SALAM, Non-commutative extrapolation algorithms, *Numerical Algorithms* **7** (1994), 225–251.
19. A. SALAM, On the vector-valued Padé approximants and the vector  $\varepsilon$ -algorithm, in *Nonlinear Numerical Methods and Rational Approximation, II* (A. Cuyt ed., Kluwer, Dordrecht, 1994), 291–301.

20. E. J. WENIGER, Nonlinear sequence transformations for the acceleration of convergence and the summation of divergent series, *Comput. Phys. Reports* **10** (1989), 189–371.

21. J. WIMP, *Sequence Transformations and Their Applications* (Academic Press, New York, 1981).

LABORATOIRE D'ANALYSE NUMERIQUE ET D'OPTIMISATION

UFR IEEA-M3

UNIVERSITÉ DES SCIENCES ET TECHNOLOGIES DE LILLE

59655 VILLENEUVE D'ASCQ CEDEX

FRANCE

*E-mail address:* [brezinsk@omega.univ-lille1.fr](mailto:brezinsk@omega.univ-lille1.fr)