Comp. & Maths. with Appls. Vol. 11, No. 6, pp. 541–564, 1985
Printed in Great Britain

0097-4943/85 $3.00 + .00

ROUNDING ERROR ANALYSIS FOR FORWARD CONTINUED FRACTION ALGORITHMS

ANNIE CUYT and PAUL VAN DER CRUYSSEN
University of Antwerp, UIA, Universiteitsplein 1, B-2610 Antwerp, Belgium

(Received February 1983; *in revised form January* 1984)

Communicated by R. B. Kelman

Abstract—There exist several algorithms for the calculation of convergents of a continued fraction. We will investigate the effect of data perturbations and rounding errors for some algorithms, using the ideas of Stummel's perturbation theory [3] which is a forward error analysis.

In Section 1 we briefly repeat the forward *a priori* error analysis which we shall use. In Section 2 we present three forward recurrence algorithms (including a method which we believe to be new) and the well-known backward recurrence algorithm for the calculation of a convergent of a given continued fraction. The next four sections are devoted to the *a priori* error analysis of the four algorithms. The theoretical results are applied to numerical examples in Section 7. As far as the rounding errors were concerned no algorithm was better or worse than the others for all the examples and no error bounds were especially more accurate than the other ones.

1. FORWARD A PRIORI ERROR ANALYSIS OF NUMERICAL ALGORITHMS.

The algorithms we will consider are defined by a finite sequence of input operations and arithmetic operations $(F_1, \ldots, F_N)$ for the determination of $(u_1, \ldots, u_N)$ which are data, intermediate or final results:

$$u_l = F_l(u_1, \ldots, u_{l-1}) \qquad l = 1, \ldots, N.$$

Under perturbations, an algorithm yields approximations v_1 of u_1 such that

$$v_l = (1 + e_l)F_l(v_1, \ldots, v_{l-1}).$$

We shall assume that the local errors e_l are bounded by

$$|e_l| \leq \gamma_l \eta \qquad l = 1, \ldots, n \tag{1.1}$$

where γ_l are suitable non-negative weights and η is an accuracy constant. For the relative *a priori* errors ([3], p. 439) $r_l = (v_l - u_l)/u_l$, $l = 1, \ldots, N$ one can write down the following linearized error equations ([3], p. 449)

$$r - \mathcal{B}r = e \tag{1.2}$$

where $r = (r_1, \ldots, r_N)$, $e = (e_1, \ldots, e_N)$ and $\mathcal{B} = (\mathcal{B}_{lt})$ is an $N \times N$ lower triangular matrix defined as follows: if F_l is an input operation then $\mathcal{B}_{lt} = 0$, $t = 1, \ldots, l$, and if F_l is an arithmetic operation $u_l = u_i \; o_l \; u_j$ with $i, j < l$, $i \neq j$ and $o_1 \in \{+, -, /, .\}$ then $\mathcal{B}_{lt}$ is given by Table 1.1.

Table 1.1.

o_l	$t = i$	$t = j$	$t \neq i, j$
$+$	u_i/u_l	u_j/u_l	0
$-$	u_i/u_l	$-u_j/u_l$	0
$/$	1		0
$\cdot$	1	1	0

For the solution of the linearized error equations (1.2) one has to calculate $(\mathcal{I} - \mathcal{B})^{-1} = \mathcal{A} = (\mathcal{A}_{lt})$ where $\mathcal{I}$ is the $N \times N$ unit matrix. The matrix $\mathcal{A}$ is also a lower triangular matrix.

The weighted relative *a priori* condition number ρ_l is then defined by

$$\rho_l = \sum_{t=1}^{l} |\mathcal{A}_{lt}| \gamma_t \qquad l = 1, \ldots, N$$

and it permits the estimates ([3], p. 455)

$$|r_l| \leq \rho_l \, \eta + 0(\eta^2) \qquad l = 1, \ldots, N. \tag{1.3}$$

One can also define weighted relative *a priori* data condition numbers ρ_l^D and weighted relative *a priori* rounding condition numbers ρ_l^R: if $D = \{l | F_l$ is an input operation$\}$ and $R = \{l / F_l$ is an arithmetic operation$\}$, we write

$$\rho_l^D = \sum_{t \in D} |\mathcal{A}_{lt}| \gamma_t$$
$$\rho_l^R = \sum_{t \in R} |\mathcal{A}_{lt}| \gamma_t \tag{1.4}$$

Clearly $\rho_1 = \rho_l^D + \rho_l^R$.

2. CALCULATION OF CONVERGENTS OF A CONTINUED FRACTION

A continued fraction is an infinite expression of the form

$$\cfrac{a_0}{b_0 + \cfrac{a_1}{b_1 + \cfrac{a_2}{b_2 + \ddots}}}$$

or, in a more suitable notation

$$\frac{a_0}{b_0} + \frac{a_1}{b_1} + \frac{a_2}{b_2} + \cdots$$

with a_k and b_k real numbers or functions.

The sub-expression

$$c_n = \frac{a_0}{b_0} + \frac{a_1}{b_1} + \frac{a_2}{b_2} + \cdots + \frac{a_n}{b_n}$$

is called the nth convergent of the continued fraction.

If

$$\lim_{n \to \infty} c_n = c$$

exists and is finite, then the continued fraction is said to converge to c. In order to approximate this limit, a sequence of successive convergents can be computed until sufficient convergence is obtained.

To this end, forward continued fraction algorithms (i.e. algorithms which permit the computation of c_n out of c_{n-1}) are very well suited. Mikloško[2] has shown that c_n is the first unknown of the tridiagonal system of $n+1$ linear equations

$$H_n(a_1, b_0)\mathbf{x}_n = \mathbf{a}_0 \tag{2.1}$$

where $H_n(a_i, b_j)$ denotes the tridiagonal matrix

$$\begin{pmatrix} b_j & \ddots & -1 & & 0 \\ a_i & \ddots & \ddots & \ddots & -1 \\ 0 & & a_{i+n-1} & \ddots & b_{j+n} \end{pmatrix}$$

and the vectors $\mathbf{x}_n$ and $\mathbf{a}_k$ are given by

$$\mathbf{x}_n = (x_{0,n}\ x_{1,n}\ \ldots\ x_{n,n})' \qquad n \geqslant 0$$

$$\mathbf{a}_k = (a_k\ 0\ \ldots\ 0)' \qquad k \geqslant 0.$$

Consequently, algorithms for the solution of a linear tridiagonal system, and especially for the solution of the first unknown, are also algorithms for the computation of a convergent. In this way Mikloško derives several well-known algorithms.

If the shooting method is used for the solution of (2.1), we get the following popular algorithm:
let

$$A_{-2} = 1,\ A_{-1} = 0\ \ A_k = b_k A_{k-1} + a_k A_{k-2} \quad (k \geqslant 0) \tag{2.2a}$$

$$B_{-2} = 0,\ B_{-1} = 1\ \ B_k = b_k B_{k-1} + a_k B_{k-2} \quad (k \geqslant 0) \tag{2.2b}$$

then

$$c_n = \frac{A_n}{B_n} \quad (n \geqslant 0) \tag{2.3}$$

The method based on the formulas (2.2-3) will be referred to as the *AB*-algorithm.

Gaussian elimination for (2.1) leads to another method:
let

$$f_0 = b_0,\ f_k = b_k + a_k/f_{k-1}\ (k \geqslant 1) \tag{2.4}$$

then

$$\begin{aligned} c_n &= \sum_{k=0}^{n} (-1)^k \frac{a_0 \ldots a_k}{f_0^2 \ldots f_{k-1}^2 f_k} \\ &= c_{n-1} - \left((-1)^{n-1} \frac{a_0 \ldots a_{n-1}}{f_0^2 \ldots f_{n-2}^2 f_{n-1}}\right)\left(\frac{a_n}{f_{n-1} f_n}\right). \end{aligned} \tag{2.5}$$

Since c_n is represented by a sum, we call the technique using (2.4-5) the SM-algorithm.

The third algorithm, which we believe to be new, is based on the following theorem.

THEOREM 2.1

Let $x_{0,n}$ and $x_{n,n}$ respectively be the first and last unknown of (2.1.) and let $y_{n-1,n-1}$ be the last unknown of the system $H_{n-1}(a_2, b_1)\mathbf{y}_{n-1} = \mathbf{a}_1$ then $c_n = x_{0,n} = -x_{n,n}/y_{n-1,n-1}$.

Proof. Using Cramer's rule we get:

$$x_{0,n}+\frac{x_{n,n}}{y_{n-1,n-1}}=\frac{\begin{vmatrix}\mathbf{a}_0 & -1 \quad 0 \ldots 0\\ & H_{n-1}(a_2,b_1)\end{vmatrix}}{|H_n(a_1,b_0)|}+\frac{\begin{vmatrix}H_{n-1}(a_1,b_0) & \mathbf{a}_0\\ 0 \ldots 0 \; a_n & \end{vmatrix}\;|H_{n-1}(a_2,b_1)|}{|H_n(a_1,b_0)|\;\begin{vmatrix}H_{n-2}(a_2,b_1) & \mathbf{a}_1\\ 0 \ldots 0 \; a_n & \end{vmatrix}}$$

$$=\frac{a_0}{|H_n(a_1,b_0)|}\left\{|H_{n-1}(a_2,b_1)| - \frac{|H_{n-1}(a_2,b_1)|\cdot|D|}{|D|}\right\}$$

$$=0$$

where D is the upper-triangular matrix

$$\begin{pmatrix} a_1 & b_1 & -1 & & 0\\ & \ddots & \ddots & \ddots & -1\\ & & \ddots & \ddots & b_{n-1}\\ 0 & & & & a_n \end{pmatrix}.$$

The unknowns $x_{n,n}$ and $y_{n-1,n-1}$ can be computed by Gaussian elimination: if

$$f_0=b_0,\; f_k=b_k+a_k/f_{k-1} \quad (k\geqslant 1) \tag{2.6a}$$

$$p_0=a_0,\; p_k=-p_{k-1}a_k/f_{k-1} \quad (k\geqslant 1)$$

and

$$g_1=b_1,\; g_k=b_k+a_k/g_{k-1} \quad (k\geqslant 2) \tag{2.6b}$$

$$q_1=a_1,\; q_k=-q_{k-1}a_k/g_{k-1} \quad (k\geqslant 2)$$

then

$$x_{n,n}=\frac{p_n}{f_n},\; y_{n-1,n-1}=\frac{q_n}{g_n}$$

so that

$$c_n=-\frac{p_n}{q_n}\cdot\frac{g_n}{f_n}.$$

Because

$$\frac{p_n}{q_n}=-\frac{a_0g_1\ldots g_{n-1}}{f_0f_1\ldots f_{n-1}}$$

we have

$$c_n=\frac{a_0g_1\ldots g_n}{f_0f_1\ldots f_n}=c_{n-1}\frac{g_n}{f_n}. \tag{2.7}$$

Now c_n is written as a product and thus we call (2.6–2.7) the PR-algorithm. These three forward algorithms will be compared with the backward algorithm (abbreviated by BW-algorithm) which is a result of solving (2.1) by means of backward Gaussian elimination:
if

$$r_{n+1,n} = 0,\ r_{k,n} = a_k/(b_k + r_{k+1,n}) \quad k = n, \ldots, 0 \tag{2.8}$$

then

$$c_n = r_{o,n}. \tag{2.9}$$

Since this algorithm is not forward, it must fully be repeated for each convergent we want to compute.

3. ERROR ANALYSIS OF THE PR-ALGORITHM

First we will compose the matrix $\mathfrak{B}$ for the computation of the convergent c_n via the PR-algorithm. This computation consists of a sequence of operations where the lth operation determines the lth row of $\mathfrak{B}$ since $\mathfrak{B}$ is composed according to the rules given in Table 1.1. The operation F_l is written down in front of the lth row of $\mathfrak{B}$. We only give the elements $\mathfrak{B}_{lt}$, $t = 1, \ldots, l$ because the others are zero.

Let us introduce the notations:

$k(0)$ = row-vector consisting of k zeroes

$$F_k = \frac{a_k}{f_{k-1}f_k} \qquad (k \geq 1)$$

$$G_k = \frac{a_k}{g_{k-1}g_k} \qquad (k \geq 2)$$

$$S_k = \frac{b_k}{f_k} \qquad (k \geq 1)$$

$$T_k = \frac{b_k}{g_k} \qquad (k \geq 2)$$

$$\tau_{i,k}(F) = 1 - F_i + F_iF_{i+1} - \ldots + (-1)^{k-i+1}F_i \ldots F_k \quad (1 \leq i \leq k+1)$$

$$\tau_{i,k}(G) = 1 - G_i + G_iG_{i+1} - \ldots + (-1)^{k-i+1}G_i \ldots G_k \quad (2 \geq i \geq k+1)$$

with the convention that

$$\tau_{k+1,k}(F) = 1 = T_{k+1,k}(G)$$

$$\tau_{i,k}(F) = 0 = T_{i,k}(G) \text{ for } i \geq k+2$$

The matrix $\mathfrak{B}$ which will have $N = 8n + 1$ rows and columns, can be subdivided into n submatrices

$$\mathfrak{B} = \begin{pmatrix} \bar{\mathfrak{B}}^{(1)} \\ \vdots \\ \bar{\mathfrak{B}}^{(n)} \end{pmatrix}$$

where $\bar{\mathfrak{B}}^{(1)}$ is a $9 \times N$ matrix and $\bar{\mathfrak{B}}^{(k)}$ is an $8 \times N$ matrix ($k = 2, \ldots, n$).

We will only give the elements $\mathfrak{B}_{lt}$, $t = 1, \ldots, l$ because the others are zero and we will also write the performed operation in front of the lth row of $\mathfrak{B}$.

Then $\bar{\mathcal{B}}^{(1)}$ is given by

input a_0	0								
input $b_0=f_0$	0	0							
$a_0/b_0=c_0$	1	-1	0						
input a_1	0	0	0	0					
a_1/f_0	0	-1	0	1	0				
input $b_1=g_1$	0	0	0	0	0	0			
$b_1+\dfrac{a_1}{f_0}=f_1$	0	0	0	0	F_1	S_1	0		
g_1/f_1	0	0	0	0	0	1	-1	0	
$c_0\cdot(g_1/f_1)=c_1$	0	0	1	0	0	0	0	1	0

and for $k\geqslant 2$, $\bar{\mathcal{B}}^{(k)}=(\mathcal{B}_{lt})$ $(l=8k-6,\ldots,8k+1$ and $t=1,\ldots,l)$ is equal to

input a_k	$8k$-11	(0)	0	0	0	0	0							
a_k/g_{k-1}	$8k$-11	(0)	-1	0	0	0	1	0						
a_k/f_{k-1}	$8k$-11	(0)	0	-1	0	0	1	0	0					
input b_k	$8k$-11	(0)	0	0	0	0	0	0	0	0				
$b_k+\dfrac{a_k}{g_{k-1}}=g_k$	$8k$-11	(0)	0	0	0	0	0	G_k	0	T_k	0			
$b_k+\dfrac{a_k}{f_{k-1}}=f_k$	$8k$-11	(0)	0	0	0	0	0	0	F_k	S_k	0	0		
g_k/f_k	$8k$-11	(0)	0	0	0	0	0	0	0	0	1	-1	0	
$c_{k-1}\cdot(g_k/f_k)=c_k$	$8k$-11	(0)	0	0	0	1	0	0	0	0	0	0	1	0

For the computation of $\mathcal{A}=(\mathcal{I}-\mathcal{B})^{-1}$ we start from the relation

$$(\mathcal{I}-\mathcal{B})\,\mathcal{A}=\mathcal{I}$$

or

$$\sum_{t=1}^{l}(\delta_{lt}-\mathcal{B}_{lt})\,\mathcal{A}_{tj}=\delta_{lj}\qquad 1\leqslant l,j\leqslant N$$

$(\delta_{lt}-\mathcal{B}_{lt}=0$ for $t>l)$.

If we use the fact that $\delta_{lt}=0$ for $t<l$ and $\mathcal{B}_{ll}=0$ then we get

$$\mathcal{A}_{lj}=\sum_{t=1}^{l-1}\mathcal{B}_{lt}\,\mathcal{A}_{tj}+\delta_{lj}\qquad 1\leqslant l,j\leqslant N \tag{3.1}$$

which means that the lth row of $\mathcal{A}$ is a linear combination of the preceding rows and with coefficients coming from the lth row of $\mathcal{B}$.

The matrix $\mathcal{A}$ can, just like $\mathcal{B}$, be subdivided into submatrices, as will be illustrated in the following theorem.

THEOREM 3.1.

$\mathscr{A}$ can be subdivided into

$$\mathscr{A}=\begin{pmatrix}\bar{\mathscr{A}}^{(1)}\\ \bar{\mathscr{A}}^{(2)}\\ \vdots\\ \bar{\mathscr{A}}^{(n)}\end{pmatrix}$$

where $\bar{\mathscr{A}}^{(1)}$ is a $9\times N$ matrix and $\bar{\mathscr{A}}^{(k)}=(\mathscr{A}_{lt})_{\substack{l=8k-6,\dots,8k+1\\ t=1,\dots,l}}$ is an $8\times N$ matrix ($k\geq 2$) given by:

$$\bar{\mathscr{A}}^{(1)}=\begin{pmatrix}
1 \\
0 & 1 \\
1 & -1 & 1 \\
0 & 0 & 0 & 1 \\
0 & -1 & 0 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 1 \\
0 & -F_1 & 0 & F_1 & F_1 & S_1 & 1 \\
0 & F_1 & 0 & -F_1 & -F_1 & 1-S_1 & -1 & 1 \\
1 & -1+F_1 & 1 & -F_1 & -F_1 & 1-S_1 & -1 & 1 & 1
\end{pmatrix}$$

$$\bar{\mathscr{A}}^{(k)}=\begin{pmatrix}
6(0) & 8(0)\ \dots\ 8(0) & 0\ 0\ 0\ 1\ 0\ 0\ 0\ 0 \\
\theta_2^{(k)} & \varphi_2^{(k)}(2)\ \dots\ \varphi_2^{(k)}(k-1) & \varphi_2^{(k)}(k) \\
\theta_3^{(k)} & \varphi_3^{(k)}(2)\ \dots\ \varphi_3^{(k)}(k-1) & \varphi_3^{(k)}(k) \\
6(0) & 8(0)\ \dots\ 8(0) & 0\ 0\ 0\ 0\ 0\ 0\ 1\ 0 \\
\theta_5^{(k)} & \varphi_5^{(k)}(2)\ \dots\ \varphi_5^{(k)}(k-1) & \varphi_5^{(k)}(k) \\
\theta_6^{(k)} & \varphi_6^{(k)}(2)\ \dots\ \varphi_6^{(k)}(k-1) & \varphi_6^{(k)}(k) & 1 \\
\theta_7^{(k)} & \varphi_7^{(k)}(2)\ \dots\ \varphi_7^{(k)}(k-1) & \varphi_7^{(k)}(k) & -1 & 1 \\
\theta_8^{(k)} & \varphi_8^{(k)}(2)\ \dots\ \varphi_8^{(k)}(k-1) & \varphi_8^{(k)}(k) & -1 & 1 & 1
\end{pmatrix}$$

(remember that $\mathscr{A}$ is a lower triangular matrix), with

$\theta_2^{(k)}=(-1)^{k-1}\ (5(0),\ T_1G_2\ \dots\ G_{k-1})$
$\theta_3^{(k)}=(-1)^{k-1}\ (0,\ -F_1\ \dots\ F_{k-1},\ 0,\ F_1\ \dots\ F_{k-1},\ F_1\ \dots\ F_{k-1},\ S_1F_2\ \dots\ F_{k-1})$
$\theta_5^{(k)}=(-1)^{k-1}\ (5(0),\ T_1G_2\ \dots\ G_k)$
$\theta_6^{(k)}=(-1)^{k-1}\ (0,\ -F_1\ \dots\ F_k,\ 0,\ F_1\ \dots\ F_k,\ F_1\ \dots\ F_k,\ S_1F_2\ \dots\ F_k)$
$\theta_7^{(k)}=(-1)^{k-1}\ (0,\ F_1\ \dots\ F_k,\ 0,\ -F_1\ \dots\ F_k,\ -F_1\ \dots\ F_k,\ T_1G_2\ \dots\ G_k-S_1F_2\ \dots\ F_k)$
$\theta_8^{(k)}=(1,\ -\tau_{1,k}(F),\ 1,\ \tau_{1,k}(F)-1,\ \tau_{1,k}(F)-1,\ T_1\tau_{2,k}(G)-S_1\tau_{2,k}(F)$

and for $i=2,\ \dots,k$

$\varphi_2^{(k)}(i)=(-1)^{k-i}\ (0,\ 0,\ 0,\ G_i\ \dots\ G_{k-1},\ G_i\ \dots\ G_{k-1},\ 0,\ T_iG_{i+1}\ \dots\ G_{k-1},\ G_{i+1}\ \dots\ G_{k-1})$
$\varphi_3^{(k)}(i)=(-1)^{k-i}\ (-F_i\ \dots\ F_{k-1},\ 0,\ 0,\ F_i\ \dots\ F_{k-1},\ 0,\ F_i\ \dots\ F_{k-1},\ S_iF_{i+1}\ \dots\ F_{k-1},\ 0)$
$\varphi_5^{(k)}(i)=(-1)^{k-i}\ (0,\ 0,\ 0,\ G_i\ \dots\ G_k,\ G_i\ \dots\ G_k,\ 0,\ T_iG_{i+1}\ \dots\ G_k,\ G_{i+1}\ \dots\ G_k)$
$\varphi_6^{(k)}(i)=(-1)^{k-i}\ (-F_i\ \dots\ F_k,\ 0,\ 0,\ F_i\ \dots\ F_k,\ 0,\ F_i\ \dots\ F_k,\ S_iF_{i+1}\ \dots\ F_k,\ 0)$
$\varphi_7^{(k)}(i)=(-1)^{k-i}\ (F_i\ \dots\ F_k,\ 0,\ 0,\ G_i\ \dots\ G_k-F_i\ \dots\ F_k,\ G_i\ \dots\ G_k,\ -F_i\ \dots\ F_k,$
$\qquad T_iG_{i+1}\ \dots\ G_k-S_iF_{i+1}\ \dots\ F_k,\ G_{i+1}\ \dots\ G_k)$
$\varphi_8^{(k)}(i)=(-\tau_{i,k}(F),\ 1,\ 1,\ \tau_{i,k}(F)-\tau_{i,k}(G),\ 1-\tau_{i,k}(G),\ -1+\tau_{i,k}(F),\ T_i\tau_{i+1,k}(G)-S_i\tau_{i+1,k}(F),\ \tau_{i+1,k}(G))$

and with the convention that $F_i \ldots F_k$ and $G_i \ldots G_k$ are equal to 1 if $i=k+1$ and equal to 0 if $i \geq k+2$.

Proof. First we show that the first 9 rows of $\mathscr{A}$ are formed by $\bar{\mathscr{A}}^{(1)}$. To this end we take (3.1) for $l = 1,\ldots,9$ and fill in the proper coefficients $\mathscr{B}_{lt}$:

$l=1 \quad \mathscr{A}_{1j} = \delta_{1j}$

$l=2 \quad \mathscr{A}_{2j} = \delta_{2j}$

$l=3 \quad \mathscr{A}_{3j} = \mathscr{A}_{1j} - \mathscr{A}_{2j} + \delta_{3j} = \delta_{1j} - \delta_{2j} + \delta_{3j}$

$l=4 \quad \mathscr{A}_{4j} = \delta_{4j}$

$l=5 \quad \mathscr{A}_{5j} = -\mathscr{A}_{2j} + \mathscr{A}_{4j} + \delta_{5j} = -\delta_{2j} + \delta_{4j} + \delta_{5j}$

$l=6 \quad \mathscr{A}_{6j} = \delta_{6j}$

$l=7 \quad \mathscr{A}_{7j} = F_1 \mathscr{A}_{5j} + S_1 \mathscr{A}_{6j} + \delta_{7j} = -F_1\delta_{2j} + F_1\delta_{4j} + F_1\delta_{5j} + S_1\delta_{6j} + \delta_{7j}$

$l=8 \quad \mathscr{A}_{8j} = \mathscr{A}_{6j} - \mathscr{A}_{7j} + \delta_{8j} = F_1\delta_{2j} - F_1\delta_{4j} - F_1\delta_{5j} + (1-S_1)\delta_{6j} - \delta_{7j} + \delta_{8j}$

$l=9 \quad \mathscr{A}_{9j} = \mathscr{A}_{3j} + \mathscr{A}_{8j} + \delta_{9j}$
$\qquad = \delta_{1j} + (-1+F_1)\delta_{2j} + \delta_{3j} - F_1\delta_{4j} - F_1\delta_{5j} + (1-S_1)\delta_{6j} - \delta_{7j} + \delta_{8j} + \delta_{9j}.$

These results agree with the elements of $\bar{\mathscr{A}}^{(1)}$.

To prove the formula for $\bar{\mathscr{A}}^{(k)}$ we use induction: we show that $\bar{\mathscr{A}}^{(2)}$ is of the right form and then we show it for $\bar{\mathscr{A}}^{(k)}$ starting from the expression for $\bar{\mathscr{A}}^{(k-1)}$.

The equations for $\bar{\mathscr{A}}^{(2)}$ are:

$l=10 \quad \mathscr{A}_{10j} = \delta_{10j}$

$l=11 \quad \mathscr{A}_{11j} = -\mathscr{A}_{6j} + \mathscr{A}_{10j} + \delta_{11j}$

$l=12 \quad \mathscr{A}_{12j} = -\mathscr{A}_{7j} + \mathscr{A}_{10j} + \delta_{12j}$

$l=13 \quad \mathscr{A}_{13j} = \delta_{13j}$

$l=14 \quad \mathscr{A}_{14j} = G_k \mathscr{A}_{11j} + T_k \mathscr{A}_{13j} + \delta_{14j}$

$l=15 \quad \mathscr{A}_{15j} = F_k \mathscr{A}_{12j} + S_k \mathscr{A}_{13j} + \delta_{15j}$

$l=16 \quad \mathscr{A}_{16j} = \mathscr{A}_{14j} - \mathscr{A}_{15j} + \delta_{16j}$

$l=17 \quad \mathscr{A}_{17j} = \mathscr{A}_{9j} + \mathscr{A}_{16j} + \delta_{17j}.$

By filling in the values for $\mathscr{A}_{6j}$, $\mathscr{A}_{7j}$ and $\mathscr{A}_{9j}$ and by substitution we get expressions for $\mathscr{A}_{lj}$ ($l=10, \ldots, 17$) which agree with the elements of $\bar{\mathscr{A}}^{(2)}$, regarding the given formulas for $\theta_r^{(2)}$ and $\varphi_r^{(2)}(2)$ ($r=2,3,5,6,7,8$).

For example for $l=12$ we have

$$\mathscr{A}_{12j} = F_1\delta_{2j} - F_1\delta_{4j} - F_1\delta_{5j} - S_1\delta_{6j} - \delta_{7j} + \delta_{10j} + \delta_{12j}$$

so that the 12th row of $\mathscr{A}$ is

$$(0, F_1, 0, -F_1, -F_1, -S_1, -1, 0, 0, 1, 0, 1)$$

On the other hand

$$\theta_3^{(2)} = (0, F_1, 0, -F_1, -F_1, -S_1)$$

and

$$\varphi_3^{(2)}(2) = (-1, 0, 0, 1, 0, 1, 0, 0)$$

Hence $(\theta_3^{(2)}, \varphi_3^{(2)}(2))$ forms the 12th row of $\mathcal{A}$, i.e. the third row of $\bar{\mathcal{A}}^{(2)}$. We assume now that the formulas are exact for $\bar{\mathcal{A}}^{(k-1)}$. To prove the formulas for $\bar{\mathcal{A}}^{(k)}$ we need the equations for the rows $8k-6, \ldots, 8k+1$. The coefficients $\mathcal{B}_{ll}$ for the equations can be found in the block $\bar{\mathcal{B}}^{(k)}$ of $\mathcal{B}$:

$l = 8k-6 \quad \mathcal{A}_{8k-6,j} = \delta_{8k-6,j}$

$l = 8k-5 \quad \mathcal{A}_{8k-5,j} = -\mathcal{A}_{8k-10,j} + \mathcal{A}_{8k-6,j} + \delta_{8k-5,j}$

$l = 8k-4 \quad \mathcal{A}_{8k-4,j} = -\mathcal{A}_{8k-9,j} + \mathcal{A}_{8k-6,j} + \delta_{8k-4,j}$

$l = 8k-3 \quad \mathcal{A}_{8k-3,j} = \delta_{8k-3,j}$

$l = 8k-2 \quad \mathcal{A}_{8k-2,j} = G_k \mathcal{A}_{8k-5,j} + T_k \mathcal{A}_{8k-3,j} + \delta_{8k-2,j}$

$l = 8k-1 \quad \mathcal{A}_{8k-1,j} = F_k \mathcal{A}_{8k-4,j} + S_k \mathcal{A}_{8k-3,j} + \delta_{8k-1,j}$

$l = 8k \quad \mathcal{A}_{8k,j} = \mathcal{A}_{8k-2,j} - \mathcal{A}_{8k-1,j} + \delta_{8k,j}$

$l = 8k+1 \quad \mathcal{A}_{8k+1,j} = \mathcal{A}_{8k-7,j} + \mathcal{A}_{8k,j} + \delta_{8k+1,j}.$

The correspondence between the elements of $\mathcal{A}$ and those of the blocks $\bar{\mathcal{A}}^{(k)}$ $(k \geqslant 2)$ is given by

$$\mathcal{A}_{lj} = \bar{\mathcal{A}}^{((l+6)+8)}_{l-8[(l+6)/8]+7,j} \qquad (l \geqslant 10)$$

where $[x]$ is the integer part of x.

With these transformations the equations become:

$$\bar{\mathcal{A}}_{1j}^{(k)} = \delta_{8k-6,j}$$

$$\bar{\mathcal{A}}_{2j}^{(k)} = -\bar{\mathcal{A}}_{5j}^{(k-1)} + \bar{\mathcal{A}}_{1j}^{(k)} + \delta_{8k-5,j}$$

$$\bar{\mathcal{A}}_{3j}^{(k)} = -\bar{\mathcal{A}}_{6j}^{(k-1)} + \bar{\mathcal{A}}_{1j}^{(k)} + \delta_{8k-4,j}$$

$$\bar{\mathcal{A}}_{4j}^{(k)} = \delta_{8k-3,j}$$

$$\bar{\mathcal{A}}_{5j}^{(k)} = G_k \bar{\mathcal{A}}_{2j}^{(k)} + T_k \bar{\mathcal{A}}_{4j}^{(k)} + \delta_{8k-2,j}$$

$$\bar{\mathcal{A}}_{6j}^{(k)} = F_k \bar{\mathcal{A}}_{3j}^{(k)} + S_k \bar{\mathcal{A}}_{4j}^{(k)} + \delta_{8k-1,j}$$

$$\bar{\mathcal{A}}_{7j}^{(k)} = \bar{\mathcal{A}}_{5j}^{(k)} - \bar{\mathcal{A}}_{6j}^{(k)} + \delta_{8k,j}$$

$$\bar{\mathcal{A}}_{8j}^{(k)} = \bar{\mathcal{A}}_{8j}^{(k-1)} + \bar{\mathcal{A}}_{7j}^{(k-1)} + \delta_{8k+1,j}.$$

We observe already that these formulas confirm the fact that the first and the fourth row of $\bar{\mathscr{A}}_k$ have a 1 on the diagonal. If we combine the equations above for $j=1,\ldots,6$ then we get equations for the vectors $\theta_r^{(k)}$:

$$\theta_2^{(k)} = -\theta_5^{(k-1)}$$

$$\theta_3^{(k)} = -\theta_6^{(k-1)}$$

$$\theta_5^{(k)} = G_k\theta_2^{(k)}$$

$$\theta_6^{(k)} = F_k\theta_3^{(k)}$$

$$\theta_7^{(k)} = \theta_5^{(k)} - \theta_6^{(k)}$$

$$\theta_8^{(k)} = \theta_8^{(k-1)} - \theta_7^{(k)}$$

which can be rewritten as:

$$\theta_2^{(k)} = -\theta_5^{(k-1)}$$

$$\theta_3^{(k)} = -\theta_6^{(k-1)}$$

$$\theta_5^{(k)} = -G_k\theta_5^{(k-1)}$$

$$\theta_6^{(k)} = -F_k\theta_6^{(k-1)}$$

$$\theta_7^{(k)} = -G_k\theta_5^{(k-1)} + F_k\theta_6^{(k-1)}$$

$$\theta_8^{(k)} = \theta_8^{(k-1)} - G_k\theta_5^{(k-1)} + F_k\theta_6^{(k-1)}$$

By induction we have expressions for the $\theta_r^{(k-1)}$ which appear in the right hand side, so that it is now easy to check the formulas for the $\theta_r^{(k)}$.
For example for $r=8$ we have:

$$\begin{aligned}\theta_8^{(k)} = {} & (1,\ -\tau_{1,k-1}(F),\ 1,\ \tau_{1,k-1}(F)-1,\ \tau_{1,k-1}(F)-1,\ T_1\tau_{2,k-1}(G)-S_1\tau_{2,k-1}(F)) \\ & -G_k(-1)^{k-2}(0,\ 0,\ 0,\ 0,\ 0,\ T_1G_2\ldots G_{k-1}) \\ & +F_k(-1)^{k-2}(0,\ -F_1\ldots F_{k-1},\ 0,\ F_1\ldots F_{k-1},\ F_1\ldots F_{k-1},\ S_1F_2\ldots F_{k-1})\end{aligned}$$

which becomes, regarding the meaning of the τ-notation:

$$\theta_8^{(k)} = (1,\ -\tau_{1,k}(F),\ 1,\ \tau_{1,k}(F)-1,\ \tau_{1,k}(F)-1,\ T_1\,\tau_{2,k}(G)-S_1\,\tau_{2,k}(F))$$

and this coincides with the formula given in the formulation of the theorem.

The same can be done for the $\varphi_r^{(k)}(i)$ $(i=2,\ldots,k)$: if the equations for $j=8(i-2)-1,\ldots,8(i-2)+6$ are combined then we get equations for $\varphi_r^{(k)}(i)$ $(r=2,3,5,6,7,8)$.

Some attention must be paid to the case $i=k$ since, unlike in the case $2\leqslant i\leqslant k-1$, the vectors

$$(\bar{\mathscr{A}}_{1,8k-9}^{(k)},\ \ldots,\ \bar{\mathscr{A}}_{1,8k-2}^{(k)}) \text{ and } (\bar{\mathscr{A}}_{4,8k-9}^{(k)},\ \ldots,\ \bar{\mathscr{A}}_{4,8k-2}^{(k)})$$

are not zero-vectors but (0,0,0,1,0,0,0,0) and (0,0,0,0,0,0,1,0) respectively. Finally the equations for $j=8k-1, 8k, 8k+1$ must be solved.

We have $\bar{\mathscr{A}}_{rj}^{(k)}=0$ for $r=1,\ldots,5$ and $\bar{\mathscr{A}}_{6j}^{(k)}=\delta_{8k-1,j}, \bar{\mathscr{A}}_{7j}^{(k)}=-\delta_{8k-1,j}+\delta_{8k,j}, \bar{\mathscr{A}}_{8j}^{(k)}=-\delta_{8k-1,j}+\delta_{8k,j}+\delta_{8k+1,j}$ as it should be. ■

According to the definitions in Section 1 we have for the PR-algorithm

$$D=\{1,2,4,6\}\cup\{8k-6,8k-3|k=2,\ldots,n\} \text{ and } R=\{1,\ldots,8n+1\}\backslash D.$$

For the weights γ_l of formula (1.1) we take $\gamma_l=1$ for l in R because the local errors are indeed bounded by η when F_l is an arithmetic operation (addition, subtraction, division, multiplication). For l in D the $\gamma_l\eta$ are upper bounds for the local error caused by the input of a_0, b_0, a_1, b_1,.... and thus we will write from now on:

$$\gamma_1=\gamma_0^a$$

$$\gamma_2=\gamma_0^b$$

$$\gamma_4=\gamma_1^a$$

$$\gamma_6=\gamma_1^b$$

$$\gamma_{8k-6}=\gamma_k^a \quad k=2,\ldots,n$$

$$\gamma_{8k-3}=\gamma_k^b \quad k=2,\ldots,n$$

The weighted relative *a priori* data and rounding condition numbers for the calculation of the nth convergent $c_n(n\geqslant 2)$ are

$$\begin{aligned}\rho_N^D&=\sum_{l\in D}|\bar{\mathscr{A}}_{8l}^{(n)}|\gamma_l\\
&=|\theta_{8.1}^{(n)}|\gamma_0^a+|\theta_{8.2}^{(n)}|\gamma_0^b+|\theta_{8.4}^{(n)}|\gamma_1^a+|\theta_{8.6}^{(n)}|\gamma_1^b+\sum_{k=2}^{n}(|\varphi_{8.4}^{(n)}(k)|\gamma_k^a+|\varphi_{8.7}^{(n)}(k)|\gamma_k^b)\\
&=\gamma_0^a+|\tau_{1,n}(F)|\gamma_0^b+|1-\tau_{1,n}(F)|\gamma_1^a+|T_1\tau_{2,n}(G)-S_1\tau_{2,n}(F)|\gamma_1^b\\
&\quad+\sum_{k=2}^{n}(|\tau_{k,n}(F)-\tau_{k,n}(G)|\gamma_k^a+|T_k\tau_{k+1,n}(G)-S_k\tau_{k+1,n}(F)|\gamma_k^b)\end{aligned}$$

and

$$\begin{aligned}\rho_N^R&=\sum_{l\in R}|\bar{\mathscr{A}}_{8l}^{(n)}|\\
&=|\theta_{8.3}^{(n)}|+|\theta_{8.5}^{(n)}|+\sum_{k=2}^{n}\left(\sum_{\substack{s=1\\ s\neq 4,7}}^{8}|\varphi_{8.s}^{(n)}(k)|\right)+3\\
&=1+|1-\tau_{1,n}(F)|+\sum_{k=2}^{n}(2+|\tau_{k,n}(F)|+|1-\tau_{k,n}(G)|+|1-\tau_{k,n}(F)|+|\tau_{k+1,n}(G)|)+3\end{aligned}$$

where $\theta_{8.s}^{(n)}$ and $\varphi_{8.s}^{(n)}(k)$ denote the sth element of $\theta_8^{(n)}$ and $\varphi_8^{(n)}(k)$ respectively. We will calculate upper bounds for ρ_N^D and ρ_N^R for the case $a_k\geqslant 0$ and $b_k\geqslant 0$ for $k=0,1,2,\ldots$ by means of the following lemma.

LEMMA 3.1

If $a_k \geq 0$ and $b_k \geq 0$ for $k=0,1,\ldots,n$ then the following numbers belong to the real interval $[0,1]$:

F_k $(k \geq 1)$

G_k $(k \geq 2)$

S_k $(k \geq 1)$

T_k $(k \geq 2)$

$\tau_{k,n}(F)$ $(k \geq 1)$

$\tau_{k,n}(G)$ $(k \geq 2)$

$|\tau_{k,n}(F) - \tau_{k,n}(G)|$ $(k \geq 2)$

$|T_k \tau_{k+1,n}(G) - S_k \tau_{k+1,n}(F)|$ $(k \geq 1)$.

Using Lemma 3.1 we obtain

$$\rho_N^D \leqslant \gamma_0{}^a + \tau_{1,n}(F)\gamma_0{}^b + (1 - \tau_{1,n}(F))\gamma_1{}^a + \gamma_1{}^b + \sum_{k=2}^{n} (\gamma_k{}^a + \gamma_k{}^b) \leqslant \sum_{k=0}^{n} (\gamma_k{}^a + \gamma_k{}^b)$$

(if $\gamma_0{}^b = \gamma_1{}^a$ then one can get a sharper upper bound for ρ_N^D because then

$$\tau_{1,n}(F)\gamma_0{}^b + (1 - \tau_{1,n}(F))\gamma_1{}^a = \gamma_1{}^a)$$

and

$$\rho_N^R \leqslant 2 - \tau_{1,n}(F) + \sum_{k=2}^{n} (4 + \tau_{k+1,n}(G) - \tau_{k,n}(G)) + 3$$

$$= 5 - \tau_{1,n}(F) + 4(n-1) + \tau_{n+1,n}(G) - \tau_{2,n}(G) \leqslant 4n + 2.$$

For numerical examples we refer to Section 7. Since γ_k^a and $\gamma_k^b (k \geq 0)$ depend on the chosen continued fraction, an upper bound for ρ_N^D in function of n will be calculated for each example separately.

4. ERROR ANALYSIS OF THE AB-ALGORITHM

We introduce the following notations $(k \geq 0)$:

$$\alpha_k = \frac{a_k A_{k-2}}{A_k}$$

$$\beta_k = \frac{b_k B_{k-1}}{B_k}$$

Then by formulas (2.2a–b):

$$1-\alpha_k=\frac{b_k A_{k-1}}{A_k}$$

$$1-\beta_k=\frac{a_k B_{k-2}}{B_k}$$

If we take a look at the $N\times N$ matrix $\mathcal{B}$ for the computation of the convergent c_n via the AB-algorithm, we see that we can subdivide $\mathcal{B}$ in n submatrices and one row

$$\mathcal{B}=\begin{pmatrix}\bar{\mathcal{B}}^{(1)}\\ \vdots \\ \bar{\mathcal{B}}^{(n)}\\ \mathcal{B}_N\end{pmatrix}.$$

Where $\bar{\mathcal{B}}^{(1)}$ is a $7\times N$ matrix, $\bar{\mathcal{B}}^{(k)}$ is an $8\times N$ matrix $(k=2,\ .\ .\ .\ ,n)$ and $\mathcal{B}_N$ is the last row of $\mathcal{B}(N=8n)$. Again we write the lth operation in front of the lth row of $\mathcal{B}$.

Then $\bar{\mathcal{B}}^{(1)}$ is given by

input $b_0=B_0$	0						
input $a_0=A_0$	0	0					
input b_1	0	0	0				
b_1B_0	1	0	1	0			
input a_1	0	0	0	0	0		
$(b_1B_0)+a_1=B_1$	0	0	0	β_1	$1-\beta_1$	0	
$a_0\cdot b_1=A_1$	0	1	1	0	0	0	0

and $\bar{\mathcal{B}}^{(k)}$ is equal to

input b_k	$8k$-19(0)	0	0	z_k(0)	0	0	0							
$b_k B_{k-1}$	$8k$-19(0)	0	0	z_k(0)	1	0	1	0						
$b_k A_{k-1}$	$8k$-19(0)	0	0	z_k(0)	0	1	1	0	0					
input a_k	$8k$-19(0)	0	0	z_k(0)	0	0	0	0	0	0				
$a_k B_{k-2}$	$8k$-19(0)	1	0	z_k(0)	0	0	0	0	0	1	0			
$a_k A_{k-2}$	$8k$-19(0)	0	1	z_k(0)	0	0	0	0	0	1	0	0		
$(b_k B_{k-1})+(a_kB_{k-2})=B_k$	$8k$-19(0)	0	0	z_k(0)	0	0	0	β_k	0	0	$1-\beta_k$	0	0	
$(b_k A_{k-1})+(a_kB_{k-1})=A_k$	$8k$-19(0)	0	0	z_k(0)	0	0	0	0	$1-\alpha_k$	0	0	α_k	0	0

where $z_k=3$ for $k=2$ and $z_k=6$ for $k=3,\ .\ .\ .\ ,n$

The last row $\mathcal{B}_N$ is

$A_n/B_n=c_n$	$8n$-3(0)	-1	1	0

For the inverse matrix $\mathcal{A}=(\mathcal{I}-\mathcal{B})^{-1}$ we have an analogous theorem as in Section 3.

THEOREM 4.1.

$\mathscr{A}$ consists of blocks $\bar{\mathscr{A}}^{(1)}, \ldots, \bar{\mathscr{A}}^{(n)}$ plus one row $\mathscr{A}_N$ where $\bar{\mathscr{A}}^{(1)}$ is a $7\times N$ matrix and $\bar{\mathscr{A}}^{(k)}$ is an $8\times N$ matrix for $k=2, \ldots, n$.

For the last two rows of each block we can write down the following recursion for $k=1, \ldots, n$.

$$\mathscr{A}_{8k-2}=(\theta^{(k)}\ \varphi^{(k)}(2)\ \ldots\ \varphi^{(k)}(n)\ 0)$$
$$\mathscr{A}_{8k-1}=(\chi^{(k)}\ \pi^{(k)}(2)\ \ldots\ \pi^{(k)}(n)\ 0)$$

where $\theta^{(k)}$, $\chi^{(k)}$ are vectors of length 7 and $\varphi^{(k)}(i)$, $\pi^{(k)}(i)$ $(i=2, \ldots, n)$ are vectors of length 8 defined by

$$\begin{cases}\theta^{(1)}=(\beta_1 & 0 & \beta_1 & \beta_1 & 1-\beta_1 & 1 & 0)\\ \chi^{(1)}=(0 & 1 & 1 & 0 & 0 & 0 & 1)\end{cases}$$

$$\begin{cases}\theta^{(2)}=(\beta_1\beta_2+(1-\beta_2) & 0 & \beta_1\beta_2 & \beta_1\beta_2 & (1-\beta_1)\beta_2 & \beta_2 & 0\ )\\ \chi^{(2)}=(\quad 0 & 1 & 1-\alpha_2 & 0 & 0 & 0 & 1-\alpha_2)\end{cases}$$

$$\begin{cases}\theta^{(k)}=\beta_k\theta^{(k-1)}+(1-\beta_k)\theta^{(k-2)} & (4.1a)\\ \chi^{(k)}=(1-\alpha_k)\chi^{(k-1)}+\alpha_k\chi^{(k-2)} & (4.1b)\end{cases}$$

$$\begin{cases}\varphi^{(k)}(i)=8(0)\\ \pi^{(k)}(i)=8(0)\end{cases}\quad \text{for } i=k+1, \ldots, n$$

$$\begin{cases}\varphi^{(k)}(k)=(\ \beta_k & \beta_k & 0 & 1-\beta_k & 1-\beta_k & 0 & 1 & 0) & (4.2a)\\ \pi^{(k)}(k)=(1-\alpha_k & 0 & 1-\alpha_k & \alpha_k & 0 & \alpha_k & 0 & 1) & (4.2b)\end{cases}$$

$$\begin{cases}\varphi^{(k)}(i)=\beta_k\varphi^{(k-1)}(i)+(1-\beta_k)\varphi^{(k-2)}(i) & (4.3a)\\ \pi^{(k)}(i)=(1-\alpha_k)\pi^{(k-1)}(i)+\alpha_k\pi^{(k-2)}(i) & (4.3b)\end{cases}\quad \text{for } i=2, \ldots, k-1$$

The last row of $\mathscr{A}$ is

$$\mathscr{A}_N=(\chi^{(n)}-\theta^{(n)},\ \pi^{(n)}(2)-\varphi^{(n)}(2), \ldots, \pi^{(n)}(n)-\varphi^{(n)}(n),\ 1)$$

The vectors $\theta^{(k)}$, $\chi^{(k)}$, $\varphi^{(k)}(i)$ and $\pi^{(k)}(i)$ have a fixed pattern of zero-elements:

$$\begin{cases}\theta_2^{(k)}=\theta_7^{(k)}=0\\ \chi_1^{(k)}=\chi_4^{(k)}=\chi_5^{(k)}=\chi_6^{(k)}=0\end{cases}\quad k\geq 1$$

$$\begin{cases}\varphi_3^{(k)}(i)=\varphi_6^{(k)}(i)=\varphi_8^{(k)}(i)=0\\ \pi_2^{(k)}(i)=\pi_5^{(k)}(i)=\pi_7^{(k)}(i)=0\end{cases}\quad i=2, \ldots, n \text{ and } k\geq 1$$

where $\theta_s^{(k)}$, $\chi_s^{(k)}$, $\varphi_s^{(k)}(i)$ and $\pi_s^{(k)}(i)$ denote the sth element of $\theta^{(k)}$, $\chi^{(k)}$, $\varphi^{(k)}(i)$ and $\pi^{(k)}(i)$ respectively.

Proof. First we will calculate an expression for the last two rows of $\bar{\mathscr{A}}^{(1)}$, which can be done by taking the equations (3.1) for $l = 1,\ldots,7$:

$$\mathscr{A}_{1j}=\delta_{1j}$$
$$\mathscr{A}_{2j}=\delta_{2j}$$
$$\mathscr{A}_{3j}=\delta_{3j}$$
$$\mathscr{A}_{4j}=\mathscr{A}_{1j}+\mathscr{A}_{3j}+\delta_{4j}$$
$$\mathscr{A}_{5j}=\delta_{5j}$$
$$\mathscr{A}_{6j}=\beta_1\mathscr{A}_{4j}+(1-\beta_1)\mathscr{A}_{5j}+\delta_{6j}$$
$$\mathscr{A}_{7j}=\mathscr{A}_{2j}+\mathscr{A}_{3j}+\delta_{7j}.$$

Hence

$$\mathscr{A}_{6j}=\beta_1\delta_{1j}+\beta_1\delta_{3j}+\beta_1\delta_{4j}+(1-\beta_1)\delta_{5j}+\delta_{6j}$$

$$\mathscr{A}_{7j}=\delta_{2j}+\delta_{3j}+\delta_{7j}$$

In an analogous way we can obtain expressions for $\mathscr{A}_{14j}$ and $\mathscr{A}_{15j}$, i.e. the last two rows of $\bar{\mathscr{A}}^{(2)}$.

For the block $\bar{\mathscr{A}}^{(k)}$ $(k\geqslant 3)$ we have to take the equations (3.1) with $l=8k-8,\ldots,8k-1$.

If we take in these equations $1\leqslant j\leqslant 8k-9$, then all the δ-terms are zero and we get:

$$\mathscr{A}_{8k-2,j}=\beta_k\mathscr{A}_{8k-7,j}+(1-\beta_k)\mathscr{A}_{8k-4,j}=\beta_k\mathscr{A}_{8k-10,j}+(1-\beta_k)\mathscr{A}_{8k-18,j} \qquad (4.4a)$$

$$\mathscr{A}_{8k-1,j}=(1-\alpha_k)\mathscr{A}_{8k-6,j}+\alpha_k\mathscr{A}_{8k-3,j}=(1-\alpha_k)\mathscr{A}_{8k-9,j}+\alpha_k\mathscr{A}_{8k-17,j}. \qquad (4.4b)$$

If $j\geqslant 8k-8$, then the term $\mathscr{A}_{lj}$ with $l<8k-8$ in the r.h.s. of the equations vanishes so that

$$\begin{aligned}\mathscr{A}_{8k-2,j}&=\beta_k\mathscr{A}_{8k-7,j}+(1-\beta_k)\mathscr{A}_{8k-4,j}+\delta_{8k-2,j}\\&=\beta_k(\delta_{8k-8,j}+\delta_{8k-7,j})+(1-\beta_k)(\delta_{8k-5,j}+\delta_{8k-4,j})+\delta_{8k-2,j}\\ \mathscr{A}_{8k-1,j}&=(1-\alpha_k)\mathscr{A}_{8k-6,j}+\alpha_k\mathscr{A}_{8k-3,j}+\delta_{8k-1,j}\\&=(1-\alpha_k)(\delta_{8k-8,j}+\delta_{8k-6,j})+\alpha_k(\delta_{8k-5,j}+\delta_{8k-3,j})+\delta_{8k-1,j}.\end{aligned}$$

Observe that (4.4a) defines the first $8k-9$ elements of the last row but one of $\bar{\mathscr{A}}^{(k)}$ as a linear combination of the first $8k-9$ elements of the last rows but one of $\bar{\mathscr{A}}^{(k-1)}$ and $\bar{\mathscr{A}}_{(k-2)}$. Since $\theta^{(k)}$, $\varphi^{(k)}(2),\ldots,\varphi^{(k)}(k-1)$ contain exactly the first $8k-9$ elements of the last row but one of $\mathscr{A}^{(k)}$, the same recursion holds for $\theta^{(k)}$, $\varphi^{(k)}(2),\ldots,\varphi^{(k)}(k-1)$. Analogously (4.4b) can be rewritten as a recursion for $\chi^{(k)}$, $\pi^{(k)}(2),\ldots,\pi^{(k)}(k-1)$.
If we define

$$\varphi^{(k)}(k)=(\mathscr{A}_{8k-2,8k-8},\ldots,\mathscr{A}_{8k-2,8k-1})$$

$$\pi^{(k)}(k)=(\mathscr{A}_{8k-1,8k-8},\ldots,\mathscr{A}_{8k-1,8k-1}).$$

then this agrees with the formulas (4.2 a–b) for $\varphi^{(k)}(k)$ and $\pi^{(k)}(k)$. Since $\mathscr{A}$ is a lower triangular matrix, indeed $\varphi^{(k)}(i)$ must be 8(0) for $k<i\leqslant n$. The reason why we have subdivided the rows in vectors $\theta^{(k)}$, $\varphi^{(k)}(i)$, $\chi^{(k)}$, $\pi^{(k)}(i)$ is that these vectors contain zeroes forming a pattern (e.g. the third, sixth and eighth element of $\varphi^{(k)}(i)$ is always zero for $i=2,\ldots,n$), which is important for the computation of the weighed relative *a priori* condition numbers later on. The proof of this pattern is by induction. We give it for $\varphi^{(k)}(i)$; for $\pi^{(k)}(i)$, $\theta^{(k)}$ and $\chi^{(k)}$ everything is similar.

For $k=1$ we have $\varphi^{(1)}(i)=8(0)$ $(2\leq i\leq n)$. If $k=2$ then $\varphi^{(2)}(2)$ is given by (4.3a) and $\varphi^{(2)}(i)=8(0)$ $(3\leqslant i\leqslant n)$, which shows that the pattern holds for $\varphi^{(1)}(i)$ and $\varphi^{(2)}(i)$. Since for $k\geqslant 3$, $\varphi^{(k)}(i)$ is a linear combination of $\varphi^{(k-1)}(i)$ and $\varphi^{(k-2)}(i)$, the zero-pattern holds for $\varphi^{(k)}(i)$ too. ■

For the AB-algorithm we have

$$D=\{1,2,3,5\}\cup\{8k-8,8k-5\,|\,k=2,\ldots,n\}\text{ and }R=\{1,\ldots,8n\}\backslash D.$$

Again $\gamma_l = 1$ for l in R. For the input of $a_0, b_0, a_1, b_1, \ldots$, we write

$$\gamma_1 = \gamma_0{}^b$$

$$\gamma_2 = \gamma_0{}^a$$

$$\gamma_3 = \gamma_1{}^b$$

$$\gamma_5 = \gamma_1{}^a$$

$$\gamma_{8k-8} = \gamma_k{}^b \quad k = 2, \ldots, n$$

$$\gamma_{8k-5} = \gamma_k{}^a \quad k = 2, \ldots, n.$$

The weighted relative *a priori* data and rounding condition numbers for the calculation of the nth convergent c_n $(n \geq 2)$ are

$$\rho_N^D = \sum_{t \in D} |\mathscr{A}_{Nt}| \gamma_t$$

$$= |\chi_1^{(n)} - \theta_1^{(n)}|\gamma_0{}^b + |\chi_2^{(n)} - \theta_2^{(n)}| \gamma_0{}^a + |\chi_3^{(n)} - \theta_3^{(n)}|\gamma_1{}^b + |\chi_5^{(n)} - \theta_5^{(n)}|\gamma_1{}^a$$

$$+ \sum_{k=2}^{n} (|\pi_1^{(n)}(k) - \varphi_1^{(n)}(k)|\gamma_k{}^b + |\pi_4^{(n)}(k) - \varphi_4^{(n)}(k)|\gamma_k{}^a)$$

$$= |\theta_1^{(n)}|\gamma_0{}^b + |\chi_2^{(n)}|\gamma_0{}^a + |\chi_3^{(n)} - \theta_3^{(n)}|\gamma_1{}^b + |\theta_5^{(n)}|\gamma_1{}^a$$

$$+ \sum_{k=2}^{n} (|\pi_1^{(n)}(k) - \varphi_1^{(n)}(k)|\gamma_k{}^b + |\pi_4^{(n)}(k) - \varphi_4^{(n)}(k)|\gamma_k{}^a)$$

and

$$\rho_N^R = \sum_{t \in R} |\mathscr{A}_{N,t}| = |\chi_4^{(n)} - \theta_4^{(n)}| + |\chi_6^{(n)} - \theta_6^{(n)}| + |\chi_7^{(n)} - \theta_7^{(n)}|$$

$$+ \sum_{k=2}^{n} \left(\sum_{\substack{s=1 \\ s \neq 1,4}}^{8} |\pi_s^{(n)}(k) - \varphi_s^{(n)}(k)| \right) + 1$$

$$= |\theta_4^{(n)}| + |\theta_6^{(n)}| + |\chi_7^{(n)}| + \sum_{k=2}^{n} (|\varphi_2^{(n)}(k)| + |\pi_3^{(n)}(k)|$$

$$+ |\varphi_5^{(n)}(k)| + |\pi_6^{(n)}(k)| + |\varphi_7^{(n)}(k)| + |\pi_8^{(n)}(k)| \,) + 1.$$

We need the following lemma to calculate upper bounds for ρ_N^D and ρ_N^R in case $a_k \geq 0$ and $b_k \geq o$ for $k = 0, 1, 2, \ldots$

LEMMA 4.1.

If $a_k \geqslant 0$ and $b_k \geqslant 0$ for $k=0,1,2,\ldots,n$, then the following numbers belong to the real interval [0,1]:

$$\begin{cases} \theta_s^{(n)} \\ \varphi_s^{(n)} & s=1,\ldots,7 \\ |\varphi_s^{(n)}-\theta_s^{(n)}| \end{cases}$$

$$\begin{cases} \varphi_s^{(n)}(i) \\ \pi_s^{(n)}(i) & s=1,\ldots,8 \text{ and } i=2,\ldots,n \\ |\pi_s^{(n)}(i)-\varphi_s^{(n)}(i)| \end{cases}$$

and the following numbers belong to the interval [0,2]:

$$\begin{cases} \varphi_2^{(n)}(i)+\varphi_5^{(n)}(i)+\varphi_7^{(n)}(i) \\ \pi_3^{(n)}(i)+\pi_6^{(n)}(i)+\pi_8^{(n)}(i) \end{cases} \quad i=2,\ldots,n$$

By means of lemma 4.1. we obtain for positive coefficients

$$\rho_N^D \leqslant \sum_{k=0}^{n} (\gamma_k{}^a+\gamma_k{}^b)$$

and

$$\rho_N^R \leqslant 4+4(n-1)=4n.$$

Numerical examples can again be found in Section 7.

5. ERROR ANALYSIS OF THE SM-ALGORITHM

Before we give the matrix $\mathcal{B}$, we rewrite (2.5) in the following way.
Let

$$H_{i,k}=F_i\ldots F_k \quad i\leqslant k$$

$$H_{k+1,k}=1$$

$$H_{i,k}=0 \quad i\geqslant k+2$$

(F_k is defined at the beginning of Section 3), then

$$c_0=H_{0,0}=F_0=\frac{a_0}{b_0}$$

$$c_k=c_{k-1}+(-1)^k H_{0,k} \quad (k=1,\ldots,n)$$

Notice also that

$$\tau_{i,k}(F)=\sum_{j=i-1}^{k}(-1)^{j-i+1}H_{i,j}$$

Now

$$\mathfrak{B}=\begin{pmatrix}\bar{\mathfrak{B}}^{(1)}\\ \vdots \\ \bar{\mathfrak{B}}^{(n)}\end{pmatrix}$$

with $\bar{\mathfrak{B}}^{(1)}$ the $10\times N$ matrix $(N=7n+3)$

input a_0	0									
input $b_0=f_0$	0	0								
$a_0/b_0=c_0=H_{0,0}=F_0$	1	-1	0							
input a_1	0	0	0	0						
input b_1	0	0	0	0	0					
a_1/f_0	0	-1	0	1	0	0				
$b_1+(a_1/f_0)=f_1$	0	0	0	0	S_1	F_1	0			
$(a_1/f_0)/f_1=F_1$	0	0	0	0	0	1	-1	0		
$F_0F_1=H_{0,1}$	0	0	1	0	0	0	0	1	0	
$c_0+(-1)^1H_{0,1}=c_1$	0	0	$\frac{c_0}{c_1}$	0	0	0	0	0	$(-1)^1\frac{H_{0,1}}{c_1}$	0

and $\bar{\mathfrak{B}}^{(k)}$ for $k\geq 2$ the $7\times N$ matrix

input a_k	$7k$-8(0)	0	0	0	0	0					
input b_k	$7k$-8(0)	0	0	0	0	0	0				
a_k/f_{k-1}	$7k$-8(0)	-1	0	0	0	1	0	0			
$b_k+(a_k/f_{k-1})=f_k$	$7k$-8(0)	0	0	0	0	0	S_k	F_k	0		
$(a_k/f_{k-1})/f_k=F_k$	$7k$-8(0)	0	0	0	0	0	0	1	-1	0	
$H_{0,k-1}F_k=H_{0,k}$	$7k$-8(0)	0	0	1	0	0	0	0	0	1	0
$c_{k-1}+(-1)^kH_{0,k}=c_k$	$7k$-8(0)	0	0	0	$\frac{c_{k-1}}{c_k}$	0	0	0	0	0	$(-1)^k\frac{H_{0,k}}{c_k}$ 0.

$\mathcal{A}$ consists of a $10\times N$ submatrix $\bar{\mathcal{A}}^{(1)}$ and $(n-1)$, each $7\times N$ matrices $\bar{\mathcal{A}}^{(k)}=(\mathcal{A}_{lt})_{l=7k-3,7k+3}$ $(k=2,\ldots,n)$ such that

$$\mathcal{A}=\begin{pmatrix}\bar{\mathcal{A}}^{(1)}\\ \vdots \\ \bar{\mathcal{A}}^{(n)}\end{pmatrix}$$

with

$$\bar{\mathcal{A}}^{(1)}=$$

$$\begin{pmatrix}
0 & 1 & & & & & & & & \\
1 & -1 & 1 & & & & & & & \\
0 & 0 & 0 & 1 & & & & & & \\
0 & 0 & 0 & 0 & 1 & & & & & \\
0 & -1 & 0 & 1 & 0 & 1 & & & & \\
0 & -F_1 & 0 & F_1 & S_1 & F_1 & 1 & & & \\
0 & F_1-1 & 0 & 1-F_1 & -S_1 & 1-F_1 & -1 & 1 & & \\
1 & F_1-2 & 1 & 1-F_1 & -S_1 & 1-F_1 & -1 & 1 & 1 & \\
\frac{c_0}{c_1}(1-F_1) & \frac{-c_0}{c_1}(1-F_1)^2 & \frac{c_0}{c_1}(1-F_1) & \frac{-c_0}{c_1}F_1(1-F_1) & \frac{c_0}{c_1}F_1S_1 & \frac{-c_0}{c_1}F_1(1-F_1) & \frac{c_0}{c_1}F_1 & \frac{-c_0}{c_1}F_1 & \frac{-c_0}{c_1}F_1 & 1
\end{pmatrix}$$

and for $k \geqslant 2$

$$\bar{\mathcal{A}}^{(k)} =$$

$$\begin{pmatrix}
6(0) & 7(0)\ \dots\ 7(0) & 0\ \ 0\ \ 0\ \ 0\ \ 1\ \ 0\ \ 0 & & & & \\
6(0) & 7(0)\ \dots\ 7(0) & 0\ \ 0\ \ 0\ \ 0\ \ 0\ \ 1\ \ 0 & & & & \\
\theta_3^{(k)} & \varphi_3^{(k)}(2)\ \dots\ \varphi_3^{(k)}(k-1) & \varphi_3^{(k)}(k) & & & & \\
\theta_4^{(k)} & \varphi_4^{(k)}(2)\ \dots\ \varphi_4^{(k)}(k-1) & \varphi_4^{(k)}(k) & 1 & & & \\
\theta_5^{(k)} & \varphi_5^{(k)}(2)\ \dots\ \varphi_5^{(k)}(k-1) & \varphi_5^{(k)}(k) & -1 & 1 & & \\
\theta_6^{(k)} & \varphi_6^{(k)}(2)\ \dots\ \varphi_6^{(k)}(k-1) & \varphi_6^{(k)}(k) & -1 & 1 & 1 & \\
\theta_7^{(k)} & \varphi_7^{(k)}(2)\ \dots\ \varphi_7^{(k)}(k-1) & \varphi_7^{(k)}(k) & (-1)^{k+1}\dfrac{H_{0,k}}{c_k} & (-1)^k\dfrac{H_{0,k}}{c_k} & (-1)^k\dfrac{H_{0,k}}{c_k} & 1
\end{pmatrix}$$

with

$$\theta_3^{(k)} = (-1)^k\,(0,\ H_{1,k-1},\ 0,\ -H_{1,k-1},\ -S_1H_{2,k-1},\ -H_{1,k-1})$$
$$\theta_4^{(k)} = (-1)^k\,(0,\ H_{1,k},\ 0,\ -H_{1,k},\ -S_1H_{2,k},\ -H_{1,k})$$
$$\theta_5^{(k)} = (-1)^k\,(0,\ H_{1,k-1}-H_{1,k},\ 0,\ H_{1,k}-H_{1,k-1},\ S_1(H_{2,k}-H_{2,k-1}),\ H_{2,k}-H_{2,k-1})$$
$$\theta_6^{(k)} = (1,\ -\tau_{1,k}(F)-\tau_{1,k-1}(F),\ 1,\ \tau_{1,k}(F)-F_1\tau_{2,k-1}(F),\ -S_1(\tau_{2,k-1}(F)+\tau_{2,k}(F)),\ \tau_{1,k}(F)-F_1\tau_{2,k-1}(F))$$
$$\theta_7^{(k)} = \left(1,\ -\tau_{1,k}(F),\ 1,\ -F_1\tau_{2,k}(F),\ \frac{S_1H_{0,1}}{c_k}\,(\tau_{2,k}(F))^2,\ -F_1\tau_{2,k}(F)\right)$$

and for $i = 2, \dots, k$

$$\varphi_3^{(k)}(i) = (-1)^{k-i}\,(-H_{i,k-1},\ 0,\ 0,\ 0,\ H_{i,k-1},\ S_iH_{i+1,k-1},\ H_{i,k-1})$$
$$\varphi_4^{(k)}(i) = (-1)^{k-i}\,(-H_{i,k},\ 0,\ 0,\ 0,\ H_{i,k},\ S_iH_{i+1,k},\ H_{i,k})$$
$$\varphi_5^{(k)}(i) = (-1)^{k-i}\,(H_{i,k}-H_{i,k-1},\ 0,\ 0,\ 0,\ H_{i,k-1}-H_{i,k},\ S_i(H_{i+1,k-1}-H_{i+1,k}),\ H_{i,k-1}-H_{i,k})$$
$$\varphi_6^{(k)}(i) = (-\tau_{i,k-1}(F)-\tau_{i,k}(F),\quad 1,\quad 1,\quad 0,\quad \tau_{i,k}(F)-F_i\tau_{i+1,k-1}(F),\quad -S_2(\tau_{i+1,k-1}(F)+\tau_{i+1,k}F)),$$
$$\tau_{i,k}(F)-F_i\tau_{i+1,k-1}(F))$$
$$\varphi_7^{(k)}(i) = (-1)^i\quad \frac{H_{0,i-1}}{c_k}\left[(\tau_{i,k}(F))^2,\quad -\tau_{i,k}(F),\quad -\tau_{i,k}(F),\quad \frac{c_{i-1}}{(-1)^iH_{0,i-1}},\quad \tau_{i,k}(F)\tau_{i+1,k}(F)F_i,\right.$$
$$\left.-S_iF_i(\tau_{i+1,k}(F))^2,\ \tau_{i,k}(F)\tau_{i+1,k}(F)F_i)\right]$$

THEOREM 5.7.

The proof is similar to that of Theorem 3.1.

For the SM-algorithm we have

$$D = \{1,2,4,5\} \cup \{7k-3, 7k-2 \mid k = 2,\dots,n\} \text{ and } R = \{1,\dots,7n+3\} \backslash D.$$

For l in R we put $\gamma_l = 1$ and for the input of $a_0, b_0, a_1, b_1, \dots,$ we put

$$\gamma_1 = {\gamma_0}^a$$

$$\gamma_2 = {\gamma_0}^b$$

$$\gamma_4 = {\gamma_1}^a$$

$$\gamma_5 = {\gamma_1}^b$$

$$\gamma_{7k-3} = {\gamma_k}^a \quad k = 2,\dots,n$$

$$\gamma_{7k-2} = {\gamma_k}^b \quad k = 2,\dots,n.$$

The weighted relative *a priori* condition numbers ρ_N^D and ρ_N^R for the calculation of $c_n (n \geq 2)$ with positive data a_k and b_k $(k = 0, \dots, n)$ are:

$$\begin{aligned}\rho_N^D &= \sum_{t \in D} |\mathscr{A}_{Nt}| \gamma_t \\ &= |\theta_{7,1}^{(n)}| \gamma_0{}^a + |\theta_{7,2}^{(n)}| \gamma_0{}^b + |\theta_{7,4}^{(n)}| \gamma_1{}^a + |\theta_{7,5}^{(n)}| \gamma_1{}^b + \sum_{k=2}^{n} (|\varphi_{7,5}^{(n)}(k)| \gamma_k{}^a + |\varphi_{7,6}^{(n)}(k)| \gamma_k{}^b) \\ &= \gamma_0{}^a + \tau_{1,n}(F)\gamma_0{}^b + F_1 \tau_{2,n}(F) \gamma_1{}^a + \frac{S_1 H_{0,1}}{c_n} (\tau_{2,n}(F))^2 \gamma_1{}^b \\ &\quad + \sum_{k=2}^{n} \frac{H_{0,k-1}}{c_n} [\tau_{k,n}(F)\tau_{k+1,n}(F) F_k \gamma_k{}^a + S_k F_k (\tau_{k+1,n}(F))^2 \gamma_k{}^b]\end{aligned}$$

and

$$\begin{aligned}\rho_N^R &= \sum_{t \in R} |\mathscr{A}_{Nt}| = |\theta_{7,3}^{(n)}| + |\theta_{7,6}^{(n)}| + \sum_{k=2}^{n} \left(\sum_{s=1}^{7} |\varphi_{7,s}^{(n)}(k)| \right) \\ &= 1 + F_1 \tau_{2,n}(F) + \sum_{k=2}^{n} \frac{H_{0,k-1}}{c_n} \left[(\tau_{k,n}(F))^2 + 2\tau_{k,n}(F) + \frac{c_{k-1}}{H_{0,k-1}} + F_k \tau_{k,n}(F) \tau_{k+1,n}(F) \right]\end{aligned}$$

where we have already used Lemma 3.1.

Because $H_{0,k-1} \leq H_{0,0} = c_0$ we obtain the following upper bounds for ρ_N^D and ρ_N^R.

$$\begin{aligned}\rho_N^D &\leq \gamma_0{}^a + \tau_{1,n}(F)\gamma_0{}^b + F_1 \tau_{2,n}(F) \gamma_1{}^a + \gamma_1{}^b + \frac{c_0}{c_n} \sum_{k=2}^{n} (\gamma_k{}^a + \gamma_k{}^b) \\ &\leq \sum_{k=0}^{1} (\gamma_k{}^a + \gamma_k{}^b) + \frac{c_0}{c_n} \sum_{k=2}^{n} (\gamma_k{}^a + \gamma_k{}^b)\end{aligned}$$

(if $\gamma_0{}^b = \gamma_1{}^a$ one can use the fact that $\tau_{1,n}(F) + F_1 \tau_{2,n}(F) = 1$ to get a sharper upper bound)

$$\rho_N{}^R \leq 2 + \sum_{k=2}^{n} \left(3 \frac{H_{0,k-1}}{c_n} \tau_{k,n}(F) + \frac{c_{k-1}}{c_n} \right)$$

(which we obtain because $\tau_{k,n}(F) + F_k \tau_{k+1,n}(F) = 1$ for $k = 2, \dots, n$)

$$\leq 2 + 3\frac{c_0}{c_n}(n-1) + \frac{1}{c_n} \sum_{k=1}^{n-1} c_k$$

For a continued fraction with positive coefficients one can say that

$$\frac{c_0}{c_n} > 1$$

Numerical examples are discussed in Section 7.

6. ERROR ANALYSIS OF THE BW-ALGORITHM

We introduce the notations

$$g_k^{(n)} = \frac{r_{k+1,n}}{b_k + r_{k+1,n}} \quad k = n, \ldots, 0$$

$$G_0^{(n)} = 1$$

$$G_k^{(n)} = g_0^{(n)} \ldots g_{k-1}^{(n)} \quad k = n, \ldots, 1$$

where the $r_{k,n}$ are given by formula (2.8)

Via an analogous method as in the previous sections ($N = 4n + 3$) or using the ideas of Stummel in [4], one obtains:

$$\rho_N^D = \sum_{k=0}^{n} |G_k^{(n)}| \gamma_k{}^a + \sum_{k=0}^{n} |(1 - g_k^{(n)}) G_k^{(n)}| \gamma_k{}^b \tag{6.1}$$

$$\rho_N^R = 2\left(\sum_{k=0}^{n-1} |G_k^{(n)}|\right) + |G_n^{(n)})|.$$

For a continued fraction with positive coefficients a_k and b_k ($k = 0, \ldots, n$), the $g_k^{(n)}$ and $G_k^{(n)}$ belong to [0, 1] and the condition numbers are bounded by

$$\rho_N^D \leq \sum_{k=0}^{n} (\gamma_k^a + \gamma_k^b)$$

$$\rho_N^D \leq 2n + 1.$$

If we rewrite (6.1) as

$$\rho_N^D = \gamma_0{}^a + \sum_{k=0}^{n-1} G_k^{(n)} (g_k^{(n)} \gamma_{k+1}^a + (1 - g_k^{(n)}) \gamma_k^b) + G_n^{(n)} \gamma_n{}^b$$

then it is easy to see that for positive a_k and b_k and in case $\gamma_{k+1}^a = \gamma_k{}^b$ for $k = 0, \ldots, n - 1$ we have the estimate

$$\rho_N^D \leq \gamma_0{}^a + \sum_{k=0}^{n} \gamma_k{}^b.$$

Sharper bounds for $g_k^{(n)}$ can be obtained by methods given in [1]. These methods can also be used to estimate $|g_k^{(n)}|$ for some classes of continued fractions which need not have positive coefficients.

7. NUMERICAL EXAMPLES

We examined several continued fractions and did not remark any regular pattern in the numerical results as far as the rounding errors were concerned: no algorithm was better or worse than the others in all the examples, no error bounds were especially more accurate than the other ones. Therefore we discuss now only the following continued fraction with positive coefficients, chosen among those we tested:

$$a_k = x \quad k \geq 0$$

$$b_k = 1 - x \quad k \geq 0$$

$$0 < x < 1.$$

Table 7.1. $x = 2/3$

	AB-algorithm			PR-algorithm			SM-algorithm			BW-algorithm		
n	c_n	$\lvert r_N \rvert$	$\frac{\bar\rho_N^D + \bar\rho_N^R}{\lvert r_N \rvert}\eta$	c_n	$\lvert r_N \rvert$	$\frac{\bar\rho_N^D + \bar\rho_N^R}{\lvert r_N \rvert}\eta$	c_n	$\lvert r_N \rvert$	$\frac{\bar\rho_N^D + \bar\rho_N^R}{\lvert r_N \rvert}\eta$	c_n	$\lvert r_N \rvert$	$\frac{\rho_N^D + \rho_N^R}{\lvert r_N \rvert}\eta$
5	5.745141E-01	9.075E-08	2.1E+01	5.745140E-01	1.300E-08	1.6E+02	5.745142E-01	1.945E-07	2.5E+01	5.745141E-01	9.075E-08	1.5E+01
6	7.343379E-01	5.336E-08	4.2E+01	7.343377E-01	1.901E-07	1.3E+01	7.343379E-01	1.345E-07	3.5E+01	7.343377E-01	1.090E-07	1.5E+01
7	6.244121E-01	1.227E-07	2.1E+01	6.244120E-01	2.725E-08	1.0E+02	6.244121E-01	2.182E-07	3.0E+01	6.244121E-01	1.227E-07	1.5E+01
8	6.960794E-01	1.396E-07	2.1E+01	6.960791E-01	2.029E-07	1.5E+01	6.960794E-01	1.396E-07	4.8E+01	6.960792E-01	1.173E-07	1.8E+01
9	6.476186E-01	1.168E-07	2.9E+01	6.476185E-01	2.477E-08	1.4E+02	6.476187E-01	2.088E-07	3.9E+01	6.476186E-01	1.168E-07	2.0E+01
10	6.796120E-01	3.211E-08	1.2E+02	6.796119E-01	1.433E-07	2.7E+01	6.796122E-01	2.075E-07	4.2E+01	6.796120E-01	5.559E-08	4.6E+01
11	6.581467E-01	2.035E-08	2.0E+02	6.581467E-01	2.035E-08	2.1E+02	6.581469E-01	2.015E-07	5.0E+01	6.581468E-01	1.109E-07	2.5E+01
12	6.723955E-01	4.702E-08	9.4E+01	6.723953E-01	1.303E-07	3.5E+02	6.723956E-01	2.243E-07	4.8E+01	6.723954E-01	4.163E-08	7.3E+01
13	6.628692E-01	4.009E-08	1.2E+02	6.628692E-01	4.009E-08	1.2E+02	6.628694E-01	2.297E-07	5.2E+01	6.628693E-01	4.983E-08	6.6E+01
14	6.692079E-01	9.369E-08	5.5E+01	6.692079E-01	9.369E-08	5.6E+01	6.692081E-01	2.626E-07	4.8E+01	6.692079E-01	4.620E-09	7.6E+02
15	6.649767E-01	1.009E-07	5.4E+01	6.649767E-01	1.009E-07	5.6E+01	6.649770E-01	2.577E-07	5.3E+01	6.649768E-01	1.124E-08	3.3E+02
16	6.677951E-01	1.507E-07	3.9E+01	6.677951E-01	6.148E-08	9.7E+01	6.677954E-01	2.955E-07	5.0E+01	6.677952E-01	2.778E-08	1.4E+02
17	6.659151E-01	7.874E-08	7.9E+01	6.659151E-01	7.874E-08	8.0E+01	6.659154E-01	2.793E-07	5.6E+01	6.659152E-01	1.076E-08	3.9E+02
18	6.671680E-01	1.053E-07	6.2E+01	6.671680E-01	1.053E-07	6.3E+01	6.671682E-01	2.520E-07	6.6E+01	6.671680E-01	1.600E-08	2.8E+02
19	6.663325E-01	1.612E-07	4.3E+01	6.663325E-01	7.172E-08	9.8E+01	6.663328E-01	2.861E-07	6.1E+01	6.663326E-01	1.773E-08	2.7E+02
20	6.668894E-01	1.333E-07	5.5E+01	6.668894E-01	1.333E-07	5.5E+01	6.668897E-01	3.135E-07	5.9E+01	6.668894E-01	4.397E-08	1.1E+02
21	6.665181E-01	1.197E-07	6.4E+01	6.665182E-01	3.024E-08	2.6E+02	6.665184E-01	3.275E-07	5.9E+01	6.665182E-01	5.918E-08	8.8E+01
22	6.667656E-01	7.142E-08	1.1E+02	6.667656E-01	1.608E-07	5.0E+01	6.667659E-01	2.862E-07	7.1E+01	6.667656E-01	7.142E-08	7.6E+01
23	6.666006E-01	3.172E-08	2.6E+02	6.666006E-01	3.172E-08	2.7E+02	6.666009E-01	3.259E-07	6.5E+01	6.666007E-01	5.770E-08	9.8E+01
24	6.667106E-01	8.255E-08	1.1E+02	6.667106E-01	1.719E-07	5.1E+01	6.667109E-01	3.645E-07	6.1E+01	6.667106E-01	8.255E-08	7.1E+01
25	6.666373E-01	4.624E-08	2.0E+02	6.666373E-01	4.624E-08	2.0E+02	6.666376E-01	4.008E-07	5.8E+01	6.666374E-01	4.317E-08	1.4E+02
26	6.666862E-01	2.003E-08	4.7E+02	6.666861E-01	1.588E-07	6.0E+01	6.666865E-01	3.777E-07	6.4E+01	6.666862E-01	6.937E-08	9.2E+01
27	6.666536E-01	6.131E-09	1.6E+03	6.666536E-01	8.328E-08	1.2E+02	6.666539E-01	3.638E-07	6.9E+01	6.666536E-01	6.131E-09	1.1E+03
28	6.666754E-01	4.541E-08	2.2E+02	6.666753E-01	1.334E-07	7.7E+01	6.666756E-01	4.030E-07	6.5E+01	6.666753E-01	4.399E-08	1.6E+02
29	6.666608E-01	4.028E-08	2.6E+02	6.666608E-01	4.028E-08	2.6E+02	6.666611E-01	4.068E-07	6.7E+01	6.666609E-01	4.912E-08	1.4E+02
30	6.666705E-01	4.271E-08	2.5E+02	6.666704E-01	1.321E-07	8.3E+01	6.666708E-01	4.043E-07	6.9E+01	6.666705E-01	4.271E-08	1.7E+02
31	6.666641E-01	4.833E-08	2.3E+02	6.666641E-01	4.108E-08	2.8E+02	6.666644E-01	4.060E-07	7.1E+01	6.666641E-01	4.833E-08	1.6E+02
32	6.666684E-01	4.216E-08	2.7E+02	6.666683E-01	1.316E-07	8.9E+01	6.666687E-01	4.049E-07	7.4E+01	6.666684E-01	4.216E-08	1.9E+02
33	6.666656E-01	4.797E-08	2.5E+02	6.666655E-01	4.143E-08	2.9E+02	6.666658E-01	4.056E-07	7.6E+01	6.666656E-01	4.797E-08	1.7E+02
34	6.666674E-01	4.192E-08	2.9E+02	6.666673E-01	1.313E-07	9.4E+01	6.666677E-01	4.051E-07	7.9E+01	6.666674E-01	4.192E-08	2.0E+02
35	6.666662E-01	1.179E-08	1.1E+03	6.666662E-01	1.179E-08	1.1E+03	6.666664E-01	4.352E-07	7.5E+01	6.666662E-01	7.762E-08	1.1E+02
36	6.666670E-01	3.187E-08	4.1E+02	6.666669E-01	1.213E-07	1.1E+02	6.666673E-01	4.152E-07	8.1E+01	6.666670E-01	3.187E-08	2.7E+02
37	6.666664E-01	1.132E-08	1.2E+03	6.666664E-01	7.809E-08	1.7E+02	6.666667E-01	4.584E-07	7.6E+01	6.666664E-01	1.132E-08	8.0E+02
38	6.666668E-01	1.748E-08	7.8E+02	6.666667E-01	1.069E-07	1.3E+02	6.666671E-01	4.296E-07	8.3E+01	6.666668E-01	1.748E-08	5.3E+02
39	6.666666E-01	1.719E-09	8.2E+03	6.666666E-01	1.719E-09	8.3E+03	6.666669E-01	4.488E-07	8.2E+01	6.666666E-01	9.113E-08	1.0E+02
40	6.666667E-01	7.068E-08	2.0E+02	6.666666E-01	1.601E-07	9.1E+01	6.666670E-01	4.658E-07	8.1E+01	6.666667E-01	7.068E-08	1.4E+02

All computations were performed on a VAX 11 of the University of Antwerp

Table 7.2. $x = 0.96875$

	AB-algorithm			PR-algorithm			SM-algorithm			BW-algorithm		
n	c_n	$\lvert r_N\rvert$	$\frac{\bar\rho_N^D+\bar\rho_N^R}{\lvert r_N\rvert}\eta$	c_n	$\lvert r_N\rvert$	$\frac{\bar\rho_N^D+\bar\rho_N^R}{\lvert r_N\rvert}\eta$	c_n	$\lvert r_N\rvert$	$\frac{\bar\rho_N^D+\bar\rho_N^R}{\lvert r_N\rvert}\eta$	c_n	$\lvert r_N\rvert$	$\frac{\bar\rho_N^D+\bar\rho_N^R}{\lvert r_N\rvert}\eta$
200	9.719887E-01	1.996E-07	2.4E+02	9.719845E-01	1.058E-06	4.5E+01	9.719869E-01	3.572E-06	3.2E+02	9.719840E-01	5.062E-07	4.7E+01
201	9.656281E-01	4.736E-07	1.1E+02	9.656277E-01	5.536E-09	8.7E+03	9.656311E-01	3.586E-06	3.2E+02	9.656271E-01	5.500E-07	4.4E+01
202	9.717844E-01	1.207E-07	4.0E+02	9.717853E-01	1.102E-06	4.4E+01	9.717877E-01	3.555E-06	3.3E+02	9.717848E-01	5.500E-07	4.4E+01
203	9.658200E-01	4.960E-07	9.8E+01	9.658194E-01	5.939E-08	8.2E+02	9.658229E-01	3.582E-06	3.3E+02	9.658189E-01	6.148E-07	3.9E+01
204	9.715974E-01	8.656E-08	5.6E+02	9.715984E-01	1.129E-06	4.3E+01	9.716008E-01	3.563E-06	3.3E+02	9.715979E-01	5.773E-07	4.2E+01
205	9.659999E-01	3.949E-07	1.2E+02	9.659995E-01	3.702E-08	1.3E+03	9.660030E-01	3.603E-06	3.3E+02	9.659989E-01	5.923E-07	4.1E+01
250	9.694105E-01	3.290E-07	1.8E+02	9.694111E-01	9.439E-07	6.3E+01	9.694138E-01	3.711E-06	3.9E+02	9.694105E-01	3.905E-07	7.6E+01
251	9.681111E-01	2.724E-07	2.2E+02	9.681109E-01	8.767E-08	6.8E+02	9.681144E-01	3.720E-06	3.9E+02	9.681104E-01	4.664E-07	6.4E+01
252	9.6 3699E-01	3.098E-07	1.9E+02	9.693705E-01	9.862E-07	6.1E+01	9.693732E-01	3.692E-06	3.9E+02	9.693700E-01	4.328E-07	7.0E+01
253	9.681504E-01	2.391E-07	2.5E+02	9.681503E-01	1.159E-07	5.2E+02	9.681537E-01	3.687E-06	4.0E+02	9.681497E-01	4.382E-07	6.9E+01
254	9.693317E-01	2.928E-07	2.1E+02	9.693324E-01	9.692E-07	6.3E+01	9.693350E-01	3.675E-06	4.0E+02	9.693319E-01	4.158E-07	7.3E+01
255	9.681873E-01	2.398E-07	2.5E+02	9.681872E-01	1.782E-07	3.4E+02	9.681906E-01	3.687E-06	4.0E+02	9.681867E-01	3.758E-07	8.1E+01
300	9.688854E-01	4.915E-07	1.5E+02	9.688860E-01	1.107E-06	6.5E+01	9.688886E-01	3.814E-06	4.5E+02	9.688855E-01	5.531E-07	6.5E+01
301	9.686196E-01	3.054E-07	2.3E+02	9.686192E-01	6.377E-08	1.1E+03	9.686230E-01	3.813E-06	4.6E+02	9.686187E-01	6.176E-07	5.8E+01
302	9.688771E-01	4.484E-07	1.6E+02	9.688777E-01	1.125E-06	6.4E+01	9.688804E-01	3.832E-06	4.5E+02	9.688772E-01	5.714E-07	6.3E+01
303	9.686277E-01	3.725E-07	1.9E+02	9.686273E-01	3.286E-09	2.2E+04	9.686310E-01	3.818E-06	4.6E+02	9.686268E-01	5.505E-07	6.6E+01
304	9.688693E-01	4.316E-07	1.7E+02	9.688699E-01	1.108E-06	6.6E+01	9.688725E-01	3.815E-06	4.6E+02	9.688694E-01	5.547E-07	6.5E+01
305	9.686353E-01	3.964E-07	1.8E+02	9.686348E-01	3.432E-08	2.1E+03	9.686386E-01	3.842E-06	4.6E+02	9.686343E-01	5.881E-07	6.2E+01
350	9.687780E-01	4.399E-07	1.9E+02	9.687787E-01	1.178E-06	7.1E+01	9.687812E-01	3.762E-06	5.4E+02	9.687782E-01	6.245E-07	6.7E+01
351	9.687237E-01	3.922E-07	2.1E+02	9.687232E-01	1.001E-07	8.4E+02	9.687269E-01	3.776E-06	5.4E+02	9.687226E-01	6.538E-07	6.4E+01
352	9.687763E-01	4.692E-07	1.8E+02	9.687771E-01	1.208E-06	7.0E+01	9.687796E-01	3.792E-06	5.4E+02	9.687765E-01	6.538E-07	6.4E+01
353	9.687253E-01	3.562E-07	2.4E+02	9.687248E-01	1.361E-07	6.2E+02	9.687286E-01	3.802E-06	5.4E+02	9.687243E-01	6.898E-07	6.1E+01
354	9.687747E-01	4.522E-07	1.9E+02	9.687755E-01	1.252E-06	6.8E+02	9.687780E-01	3.775E-06	5.4E+02	9.687750E-01	6.984E-07	6.1E+01
355	9.687269E-01	4.861E-07	1.7E+02	9.687263E-01	1.907E-07	4.4E+02	9.687301E-01	3.747E-06	5.5E+02	9.687257E-01	7.445E-07	5.7E+01
400	9.687562E-01	5.766E-07	1.7E+02	9.687565E-01	8.842E-07	1.1E+02	9.687592E-01	3.653E-06	6.3E+02	9.687560E-01	3.305E-07	1.4E+02
401	9.687453E-01	7.796E-07	1.2E+02	9.687448E-01	2.259E-07	4.2E+02	9.687481E-01	3.671E-06	6.3E+02	9.687442E-01	3.279E-07	1.5E+02
402	9.687559E-01	6.272E-07	1.5E+02	9.687561E-01	8.733E-07	1.1E+02	9.687588E-01	3.642E-06	6.4E+02	9.687556E-01	3.196E-07	1.5E+02
403	9.687456E-01	7.402E-07	1.3E+02	9.687451E-01	2.480E-07	3.9E+02	9.687484E-01	3.632E-06	6.4E+02	9.687446E-01	3.057E-07	1.6E+02
404	9.687556E-01	6.557E-07	1.5E+02	9.687558E-01	8.403E-07	1.1E+02	9.687585E-01	3.609E-06	6.5E+02	9.687552E-01	2.865E-07	1.7E+02
405	9.687458E-01	6.607E-07	1.5E+02	9.687455E-01	2.915E-07	3.3E+02	9.687487E-01	3.614E-06	6.5E+02	9.697449E-01	2.622E-07	1.8E+02
450	9.687521E-01	1.631E-07	1.1E+02	9.687518E-01	7.170E-07	1.5E+02	9.687547E-01	3.670E-06	7.1E+02	9.687518E-01	1.633E-07	3.3E+02
451	9.687492E-01	3.533E-07	3.0E+02	9.687493E-01	4.148E-07	2.6E+02	9.697524E-01	3.676E-06	7.1E+02	9.687487E-01	1.389E-07	3.9E+02
452	9.687520E-01	9.748E-07	1.1E+02	9.687517E-01	6.672E-07	1.6E+02	9.687546E-01	3.682E-06	7.1E+02	9.687512E-01	1.134E-07	4.8E+02
453	9.687493E-01	3.439E-07	3.1E+02	9.687494E-01	4.669E-07	2.3E+02	9.687525E-01	3.666E-06	7.1E+02	9.687489E-01	8.682E-01	6.2E+02
454	9.687520E-01	9.820E-07	1.1E+02	9.687516E-01	6.129E-07	1.8E+02	9.687546E-01	3.689E-06	7.1E+02	9.687511E-01	5.911E-09	9.2E+02
455	9.687493E-01	3.388E-07	3.2E+02	9.687495E-01	5.234E-07	2.1E+02	9.687526E-01	3.661E-06	7.2E+02	9.687490E-01	3.034E-09	1.8E+03

All computations were performed on a VAX 11 of the University of Antwerp

This continued fraction converges to the value x itself and for x close to 1 the convergence is very slow.

In the following tables of numerical results one can find for each of the four algorithms:

c_n	nth convergent calculated in 24 digit binary arithmetic
$\lvert r_N\rvert = \left\lvert \frac{v_N - u_N}{u_N} \right\rvert$	relative *a priori* error for $c_n = v_N$ where u_N is the value of the nth convergent calculated in 56 binary arithmetic (v_N, u_N are defined in Section 1).
$\frac{(\bar{\rho}_N{}^D + \bar{\rho}_N{}^R)\eta}{\lvert r_N\rvert}$	upper bound for $\lvert r_N\rvert(\eta = 2^{-24})$, divided by $\lvert r_N\rvert$ (see (1.3) and (1.4)), where $\bar{\rho}_N{}^D$ and $\bar{\rho}_N{}^R$ are upper bounds for $\rho_N{}^D$ and $\rho_N{}^R$ respectively.

If we take $x = 2/3$ then $\gamma_k{}^a = 1 = \gamma_k{}^b$ for $k = 0,\ldots,n$ (see Table 7.1) and if we take $x = 0.96875 = 31 \cdot 2^{-5}$ then $\gamma_k{}^a = 0 = \gamma_k{}^b$ for $k = 0,\ldots,n$ (see Table 7.2).

We can make two important remarks:

(a) If we want a reasonable approximation c_n for the limit of a slowly converging continued fraction, n must be large. Since the upper bounds $\bar{\rho}_N{}^D$ and $\bar{\rho}_N{}^R$ grow with n, while $\lvert r_N\rvert$ can oscillate and stay very small the ratio $(\bar{\rho}_N{}^D + \bar{\rho}_N{}^R)\eta/\lvert r_N\rvert$ can become quite large for a slowly converging continued fraction. In such cases $(\bar{\rho}_N{}^D + \bar{\rho}_N{}^R)\eta$ is an irrealistic estimate for $\lvert r_N\rvert$.

(b) Let us again take a look at the values of $\lvert r_N\rvert$. If the $\lvert r_N\rvert$ are quite constant then clearly $(\bar{\rho}_N{}^D + \bar{\rho}_N{}^R)\eta/\lvert r_N\rvert$ grows linearly (see the SM-algorithm in Table 7.2). If these values oscillate very much then of course also $(\rho_N^D + \rho_N^R)\eta/\lvert r_N\rvert$ oscillates.

REFERENCES

1. W. B. Jones and W. J. Thron, Numerical stability in evaluating continued fractions. *Math. Comp.* **28**(127), 795–810 (1974).
2. Miklŏsko, Investigation of algorithms for numerical computation of continued fractions. *USSR Comp. Math. Math. Phys.* **16**(4), 1–12 (1976).
3. F. Stummel, Perturbation theory for evaluation algorithms of arithmetic expressions. *Math. Comp.* **37**(156), 435–473 (1981).
4. F. Stummel, Rounding error analysis of elementary numerical algorithms. *Computing Suppl.* **2**, 169–195 (1980).