A Worpitzky theorem for vector valued continued fractions

Huanxi Zhao\textsuperscript{a,b,}\textcopyright, Gongqin Zhu\textsuperscript{b}

\textsuperscript{a}Department of Applied mathematics, Central-South University, Changsha 410083, China
\textsuperscript{b}Department of mathematics, University of Science and Technology of China, Hefei 230026, China

Received 3 January 2002; received in revised form 19 September 2002

Abstract

Generalizations of Worpitzky’s convergence theorem for ordinary continued fractions are proved for vector valued continued fractions defined by the Samelson inverse for vectors. Our convergence proof is constructive and thus yields more refined truncation-error estimates. This work shows that some of our results are exact generalizations or improvements of the scalar ones.

\textcopyright 2003 Elsevier Science B.V. All rights reserved.

Keywords: Vector valued continued fractions; Worpitzky; Convergence theorem; Truncation error

1. Introduction

We concern ourselves with the formal expression

\[
\tilde{b}_0 + \frac{a_1}{\tilde{b}_1 + \frac{a_2}{\tilde{b}_2 + \ldots}} \tag{1.1}
\]

in which the \(a_i\) are complex numbers and the \(\tilde{b}_i\) are complex vectors in \(\mathbb{C}^d, d \in \mathbb{N}\). The above evaluation process is based on the use of the Samelson inverse for vectors [see \(2,6\)]:

\[
\tilde{b}^{-1} = \frac{\tilde{b}^*}{\|\tilde{b}\|^2}, \tag{1.2}
\]

where \(\tilde{b}^*\) is denotes the complex conjugate of the proper vector \(\tilde{b}\) and \(\|\tilde{b}\|\) is the norm of \(\tilde{b}\).

(1.1) is called the vector valued continued fractions. Graves-Morris [2,3] and Gu [4] showed explicitly that vector valued continued fractions could be used to define vector valued Thiele-type
rational interpolants and vector valued rational approximants. Zhao et al. [7] gave a few simple convergence theorems for vector valued continued fractions by means of a so-called backward three-term recurrence algorithm. But many of the best known and most often used convergence criteria for ordinary continued fractions (such as see [5]) are not still generalized to the vector case.

The first object of this paper is generalization of the well-known Worpitzky convergence criterion for ordinary continued fractions.

Let \( \{t_n(\tilde{w})\} \) be a sequence of vector linear fractional transformations (l.f.ts) of the form
\[
t_0(\tilde{w}) = \tilde{b}_0 + \tilde{w}, \quad t_n(\tilde{w}) = \frac{a_n}{\tilde{b}_n + \tilde{w}}, \quad a_n \neq 0, \quad n \geq 1, \quad \tilde{b}_n, \tilde{w} \in \mathbb{C}^d.
\]

Define
\[
T_n(\tilde{w}) = t_0 \circ \cdots \circ t_n(\tilde{w}), \quad n \geq 1,
\]
then \( T_n \) is a vector valued l.f.t. Clearly, \( T_n(0) \) is a vector continued fraction itself, i.e.,
\[
T_n(0) = \tilde{b}_0 + \frac{a_1}{\tilde{b}_1} + \frac{a_2}{\tilde{b}_1 + \tilde{b}_2} + \cdots + \frac{a_n}{\tilde{b}_1 + \cdots + \tilde{b}_n},
\]
so \( T_n(0) \) is called its \( n \)th convergent. Without lose of generality, we let \( \tilde{b}_0 = 0 \) in this paper. Continued fractions \( K(a_n/\tilde{b}_n) \) with \( |a_n| \leq 1 \) and \( \|\tilde{b}_n\| = 2 \) for all \( n \) will be called Worpitzky type vector valued continued fractions (in short W-fractions) in this paper. If \( \|\tilde{z}\| \leq 1 \), then \( \|(\tilde{z})\| \leq |a_n| \leq 1 \) so that \( t_n(\tilde{z}) \subseteq \mathbb{D} \), where \( \mathbb{D} \) and \( \tilde{\mathbb{D}} \) are the open and closed unit discs, respectively, in \( \mathbb{C}^d \). It follows that
\[
T_n(\tilde{D}) \subseteq T_{n-1}(\tilde{D}) \subseteq \cdots \subseteq T_1(\tilde{D}) \subseteq \tilde{\mathbb{D}},
\]
so that the \( T_n(\tilde{D}) \) form a nested decreasing vector sequence of compact discs. We denote the radius of \( T_n(\mathbb{D}) \) by \( r_n \), i.e.,
\[
r_n = 1/2 \text{diam}(T_n(\mathbb{D})) = 1/2 \sup \{d(T_n(\tilde{w}_1), T_n(\tilde{w}_2)); \tilde{w}_1, \tilde{w}_2 \in \mathbb{D}\}. \tag{1.3}
\]

In the scalar case, it is known that the radius \( r_n \) of \( T_n(\mathbb{D}) \) (at this moment, \( d = 1 \)) satisfies
\[
\begin{align*}
(1) \quad & r_n \leq \frac{1}{2n + 1}, \tag{1.4} \\
(2) \quad & r_n \leq \rho^n \text{ if } |a_n| \leq \rho < 1 \text{ for all } n. \tag{1.5}
\end{align*}
\]
In [1], by using a geometric argument, Beardon gave the following result:
\[
\begin{align*}
(3) \quad & r_n \leq \frac{1}{2/|a_1| + 2/|a_1a_2| + \cdots + 2/|a_1\cdots a_{n-1}| + 3/|a_1\cdots a_n|}. \tag{1.6}
\end{align*}
\]
Clearly, as \( |a_n| \leq 1 \), inequality (1.6) is stronger than (1.4) so that bound (1.6) does indeed contain the best possible bound for \( r_n \). Also (1.6) implies that if \( |a_n| \leq \rho < 1 \) for all \( n \), then (1.5) is valid. Moreover, with the assumption \( \sum_k |a_1\cdots a_k| = A < \infty \), he gave a faster rate of convergence:
\[
(4) \quad r_n \leq \frac{A}{2n^2}. \tag{1.7}
\]

The second objective of this paper is to give estimates of the rate of convergence for vector valued W-fractions, which are more refined than (1.6) or (1.7) (see (3.1) and Corollary 1 in Section 3), and so it also improves the corresponding scalar results.
2. Definition, notation and difference formula

Throughout this paper, we denote by $\tilde{R}_n$ the $n$th convergent of the vector valued continued fractions, namely,

$$
\tilde{R}_n = \tilde{b}_0 + \frac{a_1}{\tilde{b}_1 + \ldots + \tilde{b}_n}
$$

and we denote by $\tilde{R}_n^i$ the $i$th tail of $\tilde{R}_n$, namely

$$
\tilde{R}_n^i = \tilde{b}_i + \frac{a_{i+1}}{\tilde{b}_{i+1} + \ldots + \tilde{b}_n}
$$

For $n = 0$ one obtains the continued fractions itself and hence one usually omits the superscripts in this case, i.e.

$$
\tilde{R}_n = \tilde{R}_n^0.
$$

The proofs of our main results are based on the use of the following Theorem 1. Before proving Theorem 1, we give two lemmas.

Lemma 1 directly comes from the definition (1.2).

Lemma 1. Let $\tilde{b}_1, \tilde{b}_2 \in \mathbb{C}^d, a \neq 0 \in \mathbb{C}$, then we find

1. $\left\| (a\tilde{b}_1)^{-1} \right\| = \frac{1}{|a| \cdot \|\tilde{b}_1\|}$,

2. $\left\| \frac{1}{\tilde{b}_1} - \frac{1}{\tilde{b}_2} \right\| = \frac{1}{\|\tilde{b}_1\| \cdot \|\tilde{b}_2\|} \|\tilde{b}_1 - \tilde{b}_2\|.$

Proof. (1) It is clearly from the definition (1.2).

(2) The proof of (2) can be achieved by noting that

$$
\left( \frac{2}{\|\tilde{b}_1\|} \right) \left( \frac{2}{\|\tilde{b}_2\|} \right) = \|\tilde{b}_1 - \tilde{b}_2\|^2.
$$

We leave the details to the reader.

The following backward algorithm for vector valued continued fractions is stated in [7].

Lemma 2. For any $n \in \mathbb{N}$, let

$$
\tilde{A}_n = \tilde{b}_n, \quad B_n = 1, \quad B_n^{-1} = \|\tilde{b}_n\|^2 \quad (n = 1, 2, \ldots)
$$

$$
\tilde{A}_n = \tilde{b}_n B_n + a_{i+1} \tilde{A}_{i+1} \ast (i = n - 1, \ldots, 0)
$$

$$
B_n = \|\tilde{b}_{i+1}\|^2 B_{i+1} + 2 \text{Re}(a_{i+2} \tilde{A}_{i+2} \ast \tilde{b}_{i+1}) + |a_{i+2}|^2 B_{i+2}^2, \quad (i = n - 2, \ldots, 0)
$$

then we have

1. $B_i \geq 0 \quad (n = 0, 1, \ldots, 0 \leq i \leq n),$

(2.8)
\( \| \vec{A}_i \|^2 = B_i B_{i-1}^{-1} \quad (1 \leq i \leq n), \) \hfill (2.9)

\( \vec{R}_n = \frac{\vec{A}_n}{B_n} = \vec{b}_1 + \frac{a_{i+1}}{|b_{i+1}|} + \frac{a_{i+2}}{|b_{i+2}|} \cdots + \frac{a_n}{|b_n|} \quad (0 \leq i \leq n). \) \hfill (2.10)

Similar to (2.2), for \( i = 0 \) one usually omits the superscripts, i.e.
\[ \vec{A}_n = \vec{A}_0 \quad \text{and} \quad B_n = B_0. \]

The algorithm (2.5)–(2.7) is not of technical nature, but it is of fundamental interest as it leads to some interesting results. Now we can give a difference formula for two convergents of vector valued continued fractions, i.e.

**Theorem 1.** For any \( n \in \mathbb{N}, \vec{w}_1, \vec{w}_2 \in \mathbb{C}^d \), the formula

\[ \| T_n(\vec{w}_1) - T_n(\vec{w}_2) \| = \left\| \frac{\vec{A}_n(\vec{w}_1)}{B_n(\vec{w}_1)} - \frac{\vec{A}_n(\vec{w}_2)}{B_n(\vec{w}_2)} \right\| = \frac{|a_1| \cdots |a_n| \cdot \| \vec{w}_1 - \vec{w}_2 \|}{\sqrt{B_n(\vec{w}_1)} \sqrt{B_n(\vec{w}_2)}} \] \hfill (2.11)

holds true for two convergents \( \vec{A}_n(\vec{w}_1)/B_n(\vec{w}_1) \) and \( \vec{A}_n(\vec{w}_2)/B_n(\vec{w}_2) \) of the vector valued continued fraction (1.1).

**Proof.** Let \( D_i = \left\| \frac{\vec{A}_i(\vec{w}_1)}{B_i(\vec{w}_1)} - \frac{\vec{A}_i(\vec{w}_2)}{B_i(\vec{w}_2)} \right\| \), \( 0 \leq i \leq n \), for \( i = 0 \), from (2.6) and (2.9), one finds

\[ D_0 = |a_1| \cdot \left| \frac{\vec{A}_1(\vec{w}_1)}{B_1(\vec{w}_1)} - \frac{\vec{A}_1(\vec{w}_2)}{B_1(\vec{w}_2)} \right| = |a_1| \cdot \left| \frac{\| \vec{A}_1(\vec{w}_1) \|^2}{B_1(\vec{w}_1)} - \frac{\| \vec{A}_1(\vec{w}_2) \|^2}{B_1(\vec{w}_2)} \right| \]

\[ = |a_1| \cdot \left| \frac{B_1(\vec{w}_1)}{\vec{A}_1(\vec{w}_1)} - \frac{B_1(\vec{w}_2)}{\vec{A}_1(\vec{w}_2)} \right|. \]

Using Lemma 1, one gets
\[ D_0 = |a_1| \cdot \frac{B_1(\vec{w}_1) \cdot B_1(\vec{w}_2)}{\| \vec{A}_1(\vec{w}_1) \| \cdot \| \vec{A}_1(\vec{w}_2) \|} \cdot D_1. \]

By continuing the above process, one sees that
\[ D_0 = \prod_{i=1}^{n} |a_i| \cdot \frac{B_i(\vec{w}_1) \cdot B_i(\vec{w}_2)}{\| \vec{A}_i(\vec{w}_1) \| \cdot \| \vec{A}_i(\vec{w}_2) \|} \cdot D_n, \] \hfill (2.12)

where
\[ D_n = \left| \frac{\vec{A}_n(\vec{w}_1)}{B_n(\vec{w}_1)} - \frac{\vec{A}_n(\vec{w}_2)}{B_n(\vec{w}_2)} \right| = \| \vec{w}_1 - \vec{w}_2 \|, \]

since \( \vec{b}_n \) has to be replaced by \( \vec{b}_n + \vec{w} \).
Paying attention to (2.9), one can write (2.12) as

\[ D_0 = \left| a_1 \right| \cdots \left| a_n \right| \cdot \| \tilde{w}_1 - \tilde{w}_2 \| \]

\[ \sqrt{B_0^n(\tilde{w}_1)} \sqrt{B_0^n(\tilde{w}_2)} \]

which completes the proof of Theorem 1. \( \square \)

3. A convergence theorem and truncation error

Now, by Theorem 1, we can give an exact generalization of Worpitzky’s Theorem.

**Theorem 2** (Worpitzky’s Theorem). If \( |a_n| \leq 1 \) for all \( n \), \( \| \tilde{b}_n \| = 2 \), then vector continued fractions \( K(a_n/\tilde{b}_n) \) converge, the radius \( r_n \) of \( T_n(\tilde{D}) \) satisfies

\[ r_n \leq \min\{ \varepsilon_1, \varepsilon_2 \}, \]  

(3.1)

where

\[ \varepsilon_1 = \frac{1}{\sum_{i=1}^{n-1} 1/|a_1 \cdots a_i| + 2/|a_1 \cdots a_n|} \frac{1}{2^{n-1-\sum_{i=2}^{n-1} |a_i|}}, \]  

(3.2)

\[ \varepsilon_2 = \frac{|a_1| \cdots |a_n|}{2^{n-1-\sum_{i=2}^{n-1} |a_i|} \cdot 2^{n-\sum_{i=2}^{n-1} |a_i|}}. \]

**Proof.** We restrict our considerations to the \( n \)th convergent \( T_n(\tilde{w}) \) \( (T_n(\tilde{w}) = \tilde{A}_n(\tilde{w})/B_n(\tilde{w})) \) where \( \tilde{w} \in \tilde{D} \). In view of (2.7) and \( \| \tilde{b}_i \| = 2 \), one can write

\[ \sqrt{B_i^n(\tilde{w})} \geq 2 \sqrt{B_i^{n+1}(\tilde{w})} - |a_{i+2}| \sqrt{B_i^{n+2}(\tilde{w})} \]

for any \( \tilde{w} \in \tilde{D} \) and \( 0 \leq i \leq n - 2 \). (3.3)

Furthermore, from \( |a_i| \leq 1 \) for all \( i \in \mathbb{N} \), one gets

\[ \sqrt{B_i^n(\tilde{w})} - \sqrt{B_i^{n+1}(\tilde{w})} \geq |a_{i+2}| \left( \sqrt{B_i^{n+1}(\tilde{w})} - \sqrt{B_i^{n+2}(\tilde{w})} \right) \]

for \( 0 \leq i \leq n - 2 \). (3.4)

Continuing the above process, one finds that

\[ \sqrt{B_i^n(\tilde{w})} \geq \sqrt{B_i^{n+1}(\tilde{w})} + \left| a_{i+2} \right| \cdots \left| a_n \right| \left( \sqrt{B_i^{n-1}(\tilde{w})} - 1 \right) \]

for \( 0 \leq i \leq n - 2 \). (3.4)

By letting \( i = 0 \) in (3.4), and repeated application of (3.4), one has

\[ \sqrt{B_0^n(\tilde{w})} \geq \sqrt{B_0^n(\tilde{w})} + \left| a_2 \right| \cdots \left| a_n \right| \left( \sqrt{B_0^{n-1}(\tilde{w})} - 1 \right) \]

\[ \geq \sqrt{B_0^{n-1}(\tilde{w})} + \sum_{k=2}^{n} \left| a_k \right| \cdots \left| a_n \right| \left( \sqrt{B_0^{n-1}(\tilde{w})} - 1 \right). \]

As \( \sqrt{B_0^{n-1}(\tilde{w})} = \| \tilde{b}_n + \tilde{w} \| \), one can write

\[ \sqrt{B_0^n(\tilde{w})} \geq \| \tilde{b}_n + \tilde{w} \| + \sum_{k=2}^{n} \left| a_k \right| \cdots \left| a_n \right| (\| \tilde{b}_n + \tilde{w} \| - 1). \]  

(3.5)
On the other hand, letting \( i = 0 \) in (3.3), and using (3.3) and (3.4), one obtains
\[
\sqrt{B_n^0(\tilde{w})} \geq (2 - |a_{i+2}|) \sqrt{B_{n-1}^n(\tilde{w})} \quad \text{for } 0 \leq i \leq n - 2. \tag{3.6}
\]
Applying (3.6) repeatedly yields
\[
\sqrt{B_n^0(\tilde{w})} \geq (2 - |a_2|) \cdots (2 - |a_n|) \sqrt{B_0^n(\tilde{w})}. \tag{3.7}
\]
From \( \sqrt{B_n^0(\tilde{w})} = \| \tilde{b}_n + \tilde{w} \| \), we can write (3.7) as
\[
\sqrt{B_n^0(\tilde{w})} \geq (2 - |a_2|) \cdots (2 - |a_n|) \| \tilde{b}_n + \tilde{w} \|. \tag{3.8}
\]
Now, from (3.5), (3.8) and Theorem 1, one finds
\[
\| T_n(\tilde{w}_1) - T_n(\tilde{w}_2) \| \leq \| T_n(\tilde{w}_1) - T_n(0) \| + \| T_n(\tilde{w}_2) - T_n(0) \|
\leq \frac{2^m}{2} \left( 2 + \sum_{k=2}^{n} |a_k \cdots a_n| \cdot \left[ \prod_{i=2}^{n} (2 - |a_i|) \| \tilde{b}_n + \tilde{w} \| \right] \right)
\tag{3.9}
\]
and
\[
\| T_n(\tilde{w}_1) - T_n(\tilde{w}_2) \| \leq \frac{2^m}{2} \left( \prod_{i=2}^{n} (2 - |a_i|) \right) \cdot \left[ \prod_{i=2}^{n} (2 - |a_i|) \| \tilde{b}_n + \tilde{w} \| \right]. \tag{3.10}
\]
By the definition of the radius \( r_n \) of \( T_n(\tilde{D}) \), it follows from (3.9) that
\[
r_n = 1/2 \sup_{\tilde{w}_1, \tilde{w}_2 \in \tilde{D}} \{ \| T_n(\tilde{w}_1) - T_n(\tilde{w}_2) \| \} \leq \frac{|a_1| \cdots |a_n|}{(2 + \sum_{k=2}^{n} |a_k \cdots a_n|) \cdot \prod_{i=2}^{n} (2 - |a_i|)}.
\tag{3.11}
\]
Similarly, from 3.10, one has
\[
r_n \leq \frac{|a_1| \cdots |a_n|}{2 \prod_{i=2}^{n} (2 - |a_i|)^2}. \tag{3.12}
\]
Furthermore, by \((1 + x) \geq 2^x \) for \( 0 \leq x \leq 1 \), and from \( |a_i| \leq 1 \) for all \( i \in \mathbb{N} \), one has
\[
(2 - |a_2|) \cdots (2 - |a_n|) \geq 2^{n-1} - \sum_{i=2}^{n} |a_i|. \tag{3.13}
\]
So from (3.11)–(3.13), one obtains
\[
r_n \leq \min\{\varepsilon_1, \varepsilon_2\}, \tag{3.14}
\]
where \( \varepsilon_1, \varepsilon_2 \) are defined by (3.2).
In particular, for any \( n, m \in \mathbb{N} \), one obtains
\[
\left| \frac{\tilde{A}_{n+m} - \tilde{A}_n}{B_{n+m} - B_n} \right| \leq 2r_n \leq 2 \min\{\varepsilon_1, \varepsilon_2\} \leq \frac{2}{n+1}. \tag{3.15}
\]
Hence, from (3.15), the vector sequence \( \{\tilde{A}_n/B_n\} \) is a Cauchy sequence, and thus according to the Cauchy criterion for convergence, it converges to a vector value \( \tilde{f} \). Let \( m \to \infty \) in (3.15),
clearly, the truncation error bounds \( \|f - (\tilde{A}\tilde{n}/B\tilde{n})\| \leq 2r_n \leq 2\min\{\varepsilon_1, \varepsilon_2\} \), which completes the proof of Theorem 2. \( \square \)

As \(|a_i| \leq 1\) for all \( i \in \mathbb{N} \), inequality (3.1) is stronger than (1.6). By using the fact that the arithmetic mean of the positive numbers \( x_1, \ldots, x_n \) is not smaller than their harmonic mean. Indeed, for all \( i \) and \( j \), one has \( (x_i/x_j + x_j/x_i) \geq 2 \) so that

\[
(x_1 + x_2 + \cdots + x_n)(1/x_1 + 1/x_2 + \cdots 1/x_n) \geq n + 2 \left( \frac{n}{2} \right) = n^2.
\]

Using this with (3.1) (after we have replaced the 2 in (3.2) by 1) and with assumption \( \sum_k |a_1 \cdots a_k| < \infty \), we obtain the following sufficient condition for a faster rate of convergence.

**Corollary 1.** Suppose that for all \( n \), \( 0 < a_n \leq 1 \) and \( \|	ilde{b}_n\| = 2 \). If the series \( \sum_k |a_1 \cdots a_k| < \infty \), then for \( K(a_n/\tilde{b}_n) \) the truncation error bounds is

\[
r_n = \min \left( O\left( \frac{1}{n^22^n} \right), O\left( \frac{1}{4^n} \right) \right) = O\left( \frac{1}{4^n} \right).
\]

In particular, in the scalar case, we have \( r_n = O(1/n^2) \), clearly, this result is better than (1.7).

Now, in the scalar case, we illustrate the effect of the improved truncation error bounds by a numerical example.

**Example 1.** Let us estimate the truncation error bounds of the continued fraction \( K(a_n/2) \) given by

\[
a_n = (-1)^n \frac{n}{n + 1}, \quad n = 1, 2, \ldots.
\]

**Solution:** Clearly, the series \( \sum_k |a_1 \cdots a_k| \) diverges. Using (1.6), one gets

\[
r_n \leq \frac{1}{2[2 + 3 + \cdots + n + 1.5(n + 1)]} = \frac{1}{n^2 + 4n + 1} = O\left( \frac{1}{n^2} \right)
\]

(3.16)

As \( \sum_{k=1}^{n} 1/k = \ln n + r + o(1) \), where \( r \) (\( r = 0.57721566 \)) is Euler constant,

\[
|\varepsilon_1| \leq \frac{2^{1.5 - r}}{(n^2 + 5n + 2)(n + 1)^{\ln 2}} \quad \text{and} \quad |\varepsilon_2| \leq \frac{2^{2 - 2r}}{(n + 1)^{1 + \ln 2} \cdot (n + 1)^{\ln 2}}.
\]

Hence, using (3.1), one has

\[
r_n \leq \min(|\varepsilon_1|, |\varepsilon_2|) = \frac{2^{2.5 - r}}{(n^2 + 5n + 2)(n + 1)^{\ln 2}} = O\left( \frac{1}{n^2 + \ln n^2} \right).
\]

(3.17)

So (3.17) is stronger than (3.16).

It would be desirable to extend Śleszyński–Pringsheim theorem to the vector case, but it appears that it will require a proof of a different type than that above.

**Acknowledgements**

The authors would like to thank the referees for their corrections and valuable comments.
References