JOURNAL OF APPROXIMATION THEORY **86**, 72–80 (1996)
ARTICLE NO. 0055

Convergence Theorems for Continued Fractions in Banach Spaces

ANDREAS SCHELLING

Abteilung Mathematik I, Universität Ulm, 89069 Ulm, Germany

Communicated by Alphonse P. Magnus

Received September 8, 1993; accepted in revised form July 25, 1995

Generalizations of Śleszyński–Pringheim's convergence criteria for ordinary continued fractions are proved for noncommutative continued fractions in Banach spaces. Some of them are exact generalizations of the scalar results. © 1996 Academic Press, Inc.

1. INTRODUCTION

We consider expressions of the form

$$A_1(B_1 + A_2(\cdots)^{-1} C_2)^{-1} C_1 \tag{1}$$

where A_n, B_n, C_n are elements of a complex Banach algebra M, called noncommutative continued fractions. They occur in computations of various mathematical and physical problems, for example in control theory for expansion of the transfer function of multivariate control systems, as solution of quadratic equations in Banach spaces, also in numerical mathematics for calculation of square roots of matrices, in theoretical physics for investigations of the Brownian motion and of the anharmonic oscillator eigenvalues as well as in perturbation theory. An extensive bibliography of their applications we find in [1].

The object of this paper are generalizations of the well-known Śleszyński–Pringheim convergence criteria for ordinary continued fractions cf. [5].

A first result, which is related to one of Śleszyński–Pringheim's Theorems is given by Denk and Riederle [1]. However they need an additional condition, so it is inconvenient to apply and for creation of other criteria like in the scalar case.

The new convergence criteria for noncommutative continued fractions, contained in this paper, are closely connected with the main results of Śleszyński–Pringheim, in some cases they are even exact generalizations.

0021-9045/96 $18.00

2. Definitions and Notations

Throughout this paper M denotes a complex non-commutative Banach algebra with norm $\|\cdot\|$, identity E and $\|E\|=1$. M^* will be the set of invertible elements of M.

Definition 1. For $k\in\mathbb{N}$ let $S_k\colon N_k\to M$, $N_k\subseteq M$,

$S_k(X):=A_k(B_k+X)^{-1}C_k$, where $A_k, B_k, C_k\in M$.

If $R_n:=S_1\circ\cdots\circ S_n(0)$ exists we call R_n the nth approximant of (1). If R_n exists for all $n\in\mathbb{N}$ and $\lim_{n\to\infty}R_n=R\in M$ then we say (1) converges and has the value R. We write

$$A_1(B_1+A_2(\cdots)^{-1}C_2)^{-1}C_1=R.$$

If R_n is well-defined we have

$$R_n=A_1(B_1+A_2(\cdots(B_{n-2}+A_{n-1}(B_{n-1}+A_nB_n^{-1}C_n)^{-1}C_{n-1})^{-1}\cdots)^{-1}C_2)^{-1}C_1.$$

In order to discuss the behaviour of the functions S_k, we define a new distance on the ball $K_\delta:=\{X\in M\mid \|X\|<\delta\}$ with a suitable $\delta<1$, which has originally been introduced by Earl and Hamilton [2] and Hayden and Suffridge [4]:

A function $f\colon K_\delta\to M$ is called differentiable in K_δ, if its Frechét derivative $Df(X)$ exists for all $X\in K_\delta$.

For differentiable f let $Df(X)[Y]$ be the Frechét-Differential of f in $X\in K_\delta$. f is called continuously differentiable in K_δ, if the mapping $Df\colon X\to Df(X)$ is continuous for all $X\in K_\delta$. Now let F_δ be the set of all continuously differentiable functions $f\colon K_\delta\to K_\delta$. For $X\in K_\delta$, $Y\in M$ we define

$$\alpha_\delta(X,Y):=\sup_{f\in F_\delta}\|Df(X)[Y]\|.$$

Let γ be a continuously differentiable curve in K_δ with $\gamma\colon[0,1]\to K_\delta$ and $\gamma(0)=X$ $\gamma(1)=Y$. For the set of all such curves we write $C_1(X,Y)$. Then the δ-length of γ is defined as

$$L_\delta(\gamma):=\int_0^1\alpha_\delta(\gamma(t),\gamma'(t))\,dt$$

and the δ-distance between X and $Y\in K_\delta$ as

$$\rho_\delta(X,Y):=\inf_{\gamma\in C_1(X,Y)}L_\delta(\gamma).$$

LEMMA 1 (Properties of $\rho_\delta(X, Y)$ [4]).

(i) *For* $X, Y \in K_{\delta'}$, $\delta' < \delta$, *we have*

$$\|X - Y\| \leqslant \rho_\delta(X, Y) \leqslant \frac{2\delta}{\delta - \delta'} \|X - Y\|$$

(ii) *Let* $h \in F_\delta$ *be a function, such that for some* $c > 0$

$g: K_\delta \to K_\delta,\ Y \mapsto h(Y) + c(h(Y) - h(X))$ *is* $\in F_\delta$ *for all* $X \in K_\delta$,

then $\rho_\delta(h(X), h(Y)) \leqslant \dfrac{1}{1+c} \rho_\delta(X, Y)$ *for all* $X, Y \in K_\delta$.

3. CONVERGENCE THEOREMS

We consider the function

$$S(X) = A(B + X)^{-1} C, \qquad A, C \in M, \quad B \in M^*$$

and require that

$$\|AB^{-1}\| \ \|C\| \leqslant (1 - \varepsilon)(1 - \|B^{-1}\|) \qquad \text{holds for some} \quad \varepsilon > 0.$$

Then for all $X \in K_{1-\varepsilon/2}$ the Neumann series implies: $S(X)$ is well defined and $\|S(X)\| \leqslant 1 - \varepsilon$.

It is easily seen that its Frechét-Differential

$DS(X)[Y] = -A(B + X)^{-1} \cdot Y(B + X)^{-1} C$ is continuous for all $X \in K_{1-\varepsilon/2}$ and $Y \in M$, so that $S \in F_{1-\varepsilon/2}$.

On the other hand the function

$g(Y) := S(Y) + (\varepsilon/4)(S(Y) - S(X)) \in F_{1-\varepsilon/2}$ for all $X \in K_{1-\varepsilon/2}$ and so by Lemma 1 we have

$$\rho_{1-\varepsilon/2}(S(X), S(Y)) \leqslant \frac{1}{1 + \varepsilon/4} \rho_{1-\varepsilon/2}(X, Y) \qquad \text{for all} \quad X, Y \in K_{1-\varepsilon/2}.$$

This property is fundamental to the following

THEOREM 1. *Let* (A_n), (B_n), (C_n) *be sequences with* $A_n, C_n \in M$, $B_n \in M^*$ $\forall n \in \mathbb{N}$, *suppose that for some* $\varepsilon > 0$ *and for each* $n \in \mathbb{N}$ *either*

$$\|A_n B_n^{-1}\| \ \|C_n\| \leqslant (1 - \varepsilon)(1 - \|B_n^{-1}\|) \tag{2}$$

or

$$\|A_n\| \ \|B_n^{-1} C_n\| \leqslant (1 - \varepsilon)(1 - \|B_n^{-1}\|) \tag{3}$$

holds. (*Not necessary* ($\forall n(2)$) *or* ($\forall n(3)$).)

Then (1) *converges and for its value* R *and* n*th approximants* R_n *we have*

$$\|R\| \leqslant 1-\varepsilon \qquad \text{and} \qquad \|R_n - R\| \leqslant \frac{4(1-\varepsilon)(1-\varepsilon/2)}{\varepsilon}\left(\frac{1}{1+\varepsilon/4}\right)^n.$$

Proof. Let $\|A_n B_n^{-1}\| \, \|C_n\| \leqslant (1-\varepsilon)(1-\|B_n^{-1}\|)$. The alternative inequality (3) is treated in the analogous way. For all $X \in K_{1-\varepsilon/2}$ we obtain

$$\|S_k(X)\| \leqslant 1-\varepsilon \quad \forall k \in \mathbb{N}, \qquad \text{thus} \quad S_k(X) \in K_{1-\varepsilon},$$

hence $S_1 \circ \cdots \circ S_n(X)$ is well defined $\forall X \in K_{1-\varepsilon/2}$, especially for $X=0$ we have

$$R_n = S_1 \circ \cdots \circ S_n(0) \text{ exists} \qquad \text{and} \qquad \|R_n\| \leqslant 1-\varepsilon \;\; \forall n \in \mathbb{N}.$$

The preliminary note shows, that we get a common contraction factor for all functions S_n on the ball $K_{1-\varepsilon/2}$, so by Lemma 1

$$\begin{aligned}
\|R_{n+k} - R_n\| &= \|S_1 \circ \cdots \circ S_{n+k}(0) - S_1 \circ \cdots \circ S_n(0)\| \\
&\leqslant \rho_{1-\varepsilon/2}(S_1(S_2 \circ \cdots \circ S_{n+k}(O)), S_1(S_2 \circ \cdots \circ S_n(0))) \\
&\leqslant \frac{1}{1+\varepsilon/4}\, \rho_{1-\varepsilon/2}(S_2 \circ \cdots \circ S_{n+k}(0), S_2 \circ \cdots \circ S_n(0)) \\
&\;\;\vdots \\
&\leqslant \left(\frac{1}{1+\varepsilon/4}\right)^n \rho_{1-\varepsilon/2}(S_{n+1} \circ \cdots \circ S_{n+k}(0), 0) \\
&\leqslant \left(\frac{1}{1+\varepsilon/4}\right)^n \frac{4(1-\varepsilon/2)(1-\varepsilon)}{\varepsilon} \qquad \forall n, k \in \mathbb{N}.
\end{aligned}$$

Hence R_n is a Cauchy sequence and therefore converges. Further the inequalities hold for $k \to \infty$ and we obtain the estimation for $\|R - R_n\|$. This completes the proof of Theorem 1.

Remark. If we put $M = \mathbb{C}$, then we have the ordinary continued fractions, where A_n and C_n together play the role of the numerators. In this case Theorem 1 is comparable with the Śleszyński–Pringheim convergence criterion cf. [5].

We know from the scalar case, that equivalence transformations (i.e. transformation of a continued fraction into another with the same approximants for all $n \in \mathbb{N}$) are suitable to create new convergence criteria. A simple but very useful transformation for our continued fractions is the following

LEMMA 2. *Let* $(D_n) \subset M^*$ *be any sequence with* $D_0 = E$. *Then the continued fractions*

$$A_1(B_1 + A_2(\cdots)^{-1} C_2)^{-1} C_1$$

and

$$\tilde{A}_1(\tilde{B}_1 + \tilde{A}_2(\cdots)^{-1} \tilde{C}_2)^{-1} \tilde{C}_1,$$

with

$$\tilde{A}_n = A_n D_n, \quad \tilde{B}_n = B_n D_n \quad \text{and} \quad \tilde{C}_n = C_n D_{n-1} \qquad \text{for} \quad n \in \mathbb{N},$$

are equivalent.

Proof. For $\tilde{S}_n(X) = \tilde{A}_n(\tilde{B}_n + X)^{-1} \tilde{C}_n$ it follows by induction:

$$\tilde{S}_1 \circ \cdots \circ \tilde{S}_n(X) = S_1 \circ \cdots \circ S_n(X D_n^{-1}).$$

Since $D_n \in M^*$, $\tilde{S}(X)$ exists exactly if $S_n(X D_n^{-1})$ exists, hence we have

$$\tilde{R}_n = \tilde{S}_1 \circ \cdots \circ \tilde{S}_n(0) = S_1 \circ \cdots \circ S_n(0) = R_n.$$

Let (p_n) be a sequence of numbers, $p_n \geqslant 1$, $B_n \in M^*$. If we choose $D_n = p_n \, \|B_n^{-1}\| \, E$ for $n \in \mathbb{N}$, then condition (2) for $\tilde{A}_n$, $\tilde{B}_n$, $\tilde{C}_n$ as in Lemma 2, implies

$$\|A_n D_n D_n^{-1} B_n^{-1}\| \, \|C_n D_{n-1}\| \leqslant (1-\varepsilon)(1 - \|D_n^{-1} B_n^{-1}\|)$$

$$\Leftrightarrow \|A_n B_n^{-1}\| \, \|C_n\| \, \|B_{n-1}^{-1}\| \, p_{n-1} \leqslant (1-\varepsilon)\left(1 - \frac{1}{p_n}\right)$$

$$\Leftrightarrow \|A_n B_n^{-1}\| \, \|C_n\| \, \|B_{n-1}^{-1}\| \leqslant (1-\varepsilon) \frac{p_n - 1}{p_n p_{n-1}}.$$

This leads to

THEOREM 2. *The continued fraction* (1) *converges if there exists a sequence of numbers* (p_n), $p_n \geqslant 1$ $\forall n \geqslant 1$ *and some* $\varepsilon > 0$ *with*

$$\|A_n B_n^{-1}\| \, \|C_n\| \, \|B_{n-1}^{-1}\| \leqslant (1-\varepsilon) \frac{p_n - 1}{p_{n-1} p_n} \tag{4}$$

or

$$\|A_n\| \, \|B_n^{-1} C_n\| \, \|B_{n-1}^{-1}\| \leqslant (1-\varepsilon) \frac{p_n - 1}{p_{n-1} p_n} \qquad \forall n \geqslant 2. \tag{5}$$

Proof. If (4) holds for $n \geqslant 2$ then $\tilde{A}_2(\tilde{B}_2 + \tilde{A}_3(\cdots)^{-1}\tilde{C}_3)^{-1}\tilde{C}_2$ converges by Theorem 1 and its value $\tilde{T}$ and approximants $\tilde{T}_n$ satisfy

$$\|\tilde{T}\| \leqslant 1-\varepsilon \text{ and } \|\tilde{T}_n\| \leqslant 1-\varepsilon, \text{ furthermore } \|B_1^{-1}\| = \frac{1}{p_1} \leqslant 1.$$

hence by the Neumann series: $\tilde{S}_1(\tilde{T}_n) = \tilde{R}_n$ and $\tilde{S}_1(\tilde{T}) = \tilde{R}$ exist. It follows by Lemma 2 that $R_n \to R$ $(n \to \infty)$, for we have two equivalent continued fractions. Condition (5) implies convergence in the analogous way.

Remark. Theorem 2 is a generalization of a further Theorem of Śleszyński–Pringheim cf. [5].

A first generalization of this Theorem is due to Denk and Riederle [1]. They need a strong additional condition for the numbers p_n, which has unpleasant consequence for applications, especially for creation of other convergence criterions by proper choice of p_n. The new results are suitable to generalize many of the known criteria for ordinary continued fractions to continued fractions in Banach Algebras. We renounce to specify this.

Let us have a look upon a special form of continued fractions. We choose

$$C_n = E \text{ for all } n \in \mathbb{N}, \text{ then (1) becomes } A_1(B_1 + A_2(\cdots)^{-1})^{-1}. \tag{6}$$

In this case we are able to discuss convergence also by the following recurrence relations:

$$\begin{aligned} &\text{Let} \quad P_{-1} = E, \quad P_0 = 0, \quad P_n = P_{n-1}B_n + P_{n-2}A_n \quad \text{for} \quad n \geqslant 1 \\ &\text{and} \quad Q_{-1} = 0, \quad Q_0 = E, \quad Q_n = Q_{n-1}B_n + Q_{n-2}A_n \quad \text{for} \quad n \geqslant 1. \end{aligned}$$

If R_n exists we have $R_n = P_n Q_n^{-1}$ (especially $Q_n \in M^*$), therefore we call P_n the nth numerator and Q_n the nth denominator of (6).

On the other side let

$$Q_{-1}^{(k)} = 0, \quad Q_0^{(k)} = E \qquad \text{and} \qquad Q_n^{(k)} = Q_{n-1}^{(k)} B_{n+k} + Q_{n-2}^{(k)} A_{n+k},$$

then R_n of (6) exists if

$$Q_{n-k}^{(k)} \in M^* \qquad \text{for all} \quad 0 \leqslant k < n \quad \text{see [6].}$$

For P_n and Q_n we have the identity

$$P_n Q_n^{-1} = \sum_{k=1}^{n-1} (-1)^k A_1 B_1^{-1} Q_0 A_2 Q_2^{-1} Q_1 A_3 Q_3^{-1} \cdots Q_{k-1} A_{k+1} Q_{k+1}^{-1} \tag{7}$$

(if $Q_k \in M^*$ for all k) [3].

Now we state

THEOREM 3. (6) *converges if* $A_n \neq 0$, $B_n \in M^*$,

$$\|A_n B_n^{-1}\| \leqslant 1 - \|B_n^{-1}\| \tag{8}$$

and

$$\|A_n Q_n^{-1}\| \, \|Q_n\| \, \|B_n^{-1}\| \leqslant 1 - \|B_n^{-1}\| \qquad \forall n \in \mathbb{N}. \tag{9}$$

The value R *of* (6) *satisfies* $\|R\| \leqslant 1$.

Remark. Inequality (8) ensures that Q_n^{-1} exists, so the left side of (9) makes sense, subsequently (8) follows immediately from (9).

Proof of Theorem 3. Let $G_n := \{X \in M \mid \|XB_n^{-1}\| \leqslant \|B_n^{-1}\|\}$, then $S_n(X) = A_n(B_n + X)^{-1}$ exists $\forall X \in G_n$, because $A_n \neq 0$ and (8) implies $\|B_n^{-1}\| < 1$.

Further (8) guarantees

$$\|S_n(X)\, B_{n-1}^{-1}\| \leqslant \|A_n(B_n + X)^{-1}\| \, \|B_{n-1}^{-1}\| \leqslant \frac{\|A_n B_n^{-1}\|}{1 - \|B_n^{-1}\|} \|B_{n-1}^{-1}\| \leqslant \|B_{n-1}^{-1}\|,$$

hence $S_n(X) \in G_{n-1}$ $\forall X \in G_n$.

Because $0 \in G_n$ $\forall n \in \mathbb{N}$, we have

$$R_n = S_1 \circ \cdots \circ S_n(0) \text{ exists and furthermore } Q_n \in M^*.$$

By the recurrence formula in connection with (8) it follows:

$$\|Q_n\| = \|(Q_{n-1} + Q_{n-2} A_n B_n^{-1})\, B_n\| \geqslant \frac{\|Q_{n-1}\| - \|Q_{n-2}\| \, \|A_n B_n^{-1}\|}{\|B_n^{-1}\|}$$

$$\geqslant \frac{1}{\|B_n^{-1}\|} \|Q_{n-1}\| - \left(\frac{1}{\|B_n^{-1}\|} - 1\right) \|Q_{n-2}\|$$

hence

$$\|Q_n\| - \|Q_{n-1}\| \geqslant \left(\frac{1}{\|B_n^{-1}\|} - 1\right) (\|Q_{n-1}\| - \|Q_{n-2}\|)$$

and therefore

$$\|Q_n\| - \|Q_{n-1}\| \geqslant \left(\frac{1}{\|B_n^{-1}\|} - 1\right) \cdots \left(\frac{1}{\|B_2^{-1}\|} - 1\right) (\|Q_1\| - \|Q_0\|)$$
$$\geqslant \left(\frac{1}{\|B_n^{-1}\|} - 1\right) \cdots \left(\frac{1}{\|B_1^{-1}\|} - 1\right)$$
$$\geqslant \frac{\|A_n B_n^{-1}\|}{\|B_n^{-1}\|} \cdots \frac{\|A_1 B_1^{-1}\|}{\|B_1^{-1}\|} > 0,$$

so $\|Q_n\|$ strictly increases.

Thus

$$\sum_{k=1}^{\infty} \frac{1}{\|Q_{k-1}\|} - \frac{1}{\|Q_k\|}$$

converges. (9) and the above computation imply

$$\frac{1}{\|Q_{k-1}\|} - \frac{1}{\|Q_k\|}$$
$$= \frac{\|Q_k\| - \|Q_{k-1}\|}{\|Q_{k-1}\| \, \|Q_k\|}$$
$$\geqslant \frac{1 - \|B_k^{-1}\|}{\|B_k^{-1}\|} \cdots \frac{1 - \|B_1^{-1}\|}{\|B_1^{-1}\|} \frac{1}{\|Q_k\| \, \|Q_{k-1}\|}$$
$$= \frac{(1 - \|B_k^{-1}\|) \, \|Q_{k-2}\|}{\|B_k^{-1}\| \, \|Q_k\|} \cdots \frac{(1 - \|B_2^{-1}\|) \, \|Q_0\|}{\|B_2^{-1}\| \, \|Q_2\|} \frac{(1 - \|B_1^{-1}\|)}{\|B_1^{-1}\| \, \|Q_1\|}$$
$$\geqslant \|A_k Q_k^{-1}\| \, \|Q_{k-2}\| \cdots \|A_2 Q_2^{-1}\| \, \|Q_0\| \, \|A_1 Q_1^{-1}\|.$$

Thus

$$\sum_{k=1}^{\infty} \|A_1 Q_1^{-1}\| \, \|Q_0\| \, \|A_2 Q_2^{-1}\| \cdots \|Q_{k-2}\| \, \|A_k Q_k^{-1}\|$$

converges and by (7) we obtain $\lim_{n \to \infty} R_n = \lim_{n \to \infty} P_n Q_n^{-1}$ exists.

At last $\|Q_0\| = \|E\|$ implies

$$\sum_{k=1}^{\infty} \frac{1}{\|Q_{k-1}\|} - \frac{1}{\|Q_k\|} \leqslant 1$$

and so $\|R\| \leqslant 1$.

Using the transformation method in a similar way like above we obtain:

THEOREM 4. *Let* $A_n \neq 0$, $B_n \in M^*$ *then* (6) *converges, if there exist real numbers* $p_n \geq 1$ *such that*

$$\|A_n B_n^{-1}\| \; \|B_{n-1}^{-1}\| \leq \frac{p_n - 1}{p_{n-1} p_n}$$

and

$$\|A_n Q_n^{-1}\| \; \|Q_n\| \; \|B_n^{-1}\| \; \|B_{n-1}^{-1}\| \leq \frac{p_n - 1}{p_{n-1} p_n}, \qquad \forall n \geq 2$$

and strict inequality holds for at least one n.

Proof. We have to use another equivalence transformation:
For $D_n \in M^*$ $\forall n \geq 1$ we put

$\tilde{A}_1 = A_1 D_1$, $\tilde{A}_2 = A_2 D_2$, $\tilde{A}_n = D_{n-2}^{-1} A_n D_n$, $n \geq 3$ and

$\tilde{B}_1 = B_1 D_1$, $\tilde{B}_n = D_{n-1}^{-1} B_n D_n$, $n \geq 2$.

$\tilde{A}_1(\tilde{B}_1 + \tilde{A}_2(\cdots)^{-1})^{-1}$ and (6) are equivalent. For the nth numerators and denominators we have $\tilde{P}_n = P_n D_n$ and $\tilde{Q}_n = Q_n D_n$. This follows by induction (compare [6]).

Now we choose $D_n = p_1 \cdots p_n \|B_1^{-1}\| \cdots \|B_n^{-1}\| E$, then

$$\tilde{A}_2(\tilde{B}_2 + \tilde{A}_3(\cdots)^{-1})^{-1} \tag{10}$$

complies with the prerequisites of Theorem 3. Thus (10) tends to $\tilde{T} \in M$.

The strict inequality for at least one $n \geq 2$ ensures that $\tilde{A}_1(\tilde{B}_1 + \tilde{T})^{-1}$ and all its approximants exist. Altogether (6) converges.

Remarks. 1. Both Theorem 3 and Theorem 4 are generalizations of the well-known scalar convergence criteria, similar to Theorem 1 and Theorem 2. The outstanding property of these results is, that if we put $M = \mathbb{C}$ we have exactly the Śleszyński–Pringheim results.

2. Theorem 3 and Theorem 4 of course hold in the analogous way if we consider continued fractions $(B_1 + (B_2 + (\cdots)^{-1} C_3)^{-1} C_2)^{-1} C_1$.

REFERENCES

1. H. DENK AND M. RIEDERLE, A generalization of a theorem of Pringheim, *J. Approx. Theory* **35** (1982), 355–363.
2. C. J. EARL AND R. S. HAMILTON, A fixed point theorem for holomorphic mappings, *Proc. Sympos. Pure Math.* **16** (1970), 61–65.
3. W. FAIR, Noncommutative continued fractions, *SIAM J. Math. Anal.* **2** (1971), 226–232.
4. T. L. HAYDEN AND T. J. SUFFRIDGE, Fixed points of holomorphic maps in Banach spaces, *Proc. Amer. Math. Soc.* **60** (1976), 95–105.
5. O. PERRON, "Die Lehre von den Kettenbrüchen, II," 3rd ed., Teubner Verlag/Stuttgart, 1957.
6. P. PFLUGER, "Matrizenkettenbrüche," Dissertation ETH. Zürich, 1966.