Reprinted from JOURNAL OF NUMBER THEORY
Vol. 5, No. 2, April 1973
Printed in Belgium
All Rights Reserved by Academic Press, New York and London

Some Remarks on Kuzmin's Theorem for F-Expansions

F. SCHWEIGER

Mathematisches Institut der Universität Salzburg, Porschestr. 1/1, A5020 Salzburg, Austria

AND

M. WATERMAN*

Department of Mathematics, Idaho State University, Pocatello, Idaho 83201, U.S.A.

Communicated by E. Hlawka

Received December 14, 1970; revised October 13, 1971

In this paper a general Kuzmin theorem for a class of multidimensional F-expansions is given. This result gives the uniform rate at which a certain sequence of approximates converges to the density of the invariant measure associated with an F-expansion. Some metric theorems are also given. This work extends and corrects some earlier results. The Jacobi algorithm is included as an example.

1. INTRODUCTION

In a recent paper [6], one of the authors gave a general Kuzmin theorem for a class of multidimensional F-expansions. The theorem applies to the Jacobi algorithm only when $n = 2$. This is due to the fact that $TB_{\nu+1} = B_\nu$, $\nu \geqslant 1$, is not satisfied for $n \geqslant 3$. This difficulty can easily be shown by an example based on the restriction of digits [2, 3]. In the proof of the theorem the assumption that $B_\nu \cap A_i \neq \varnothing$ implies $B_\nu \subseteq A_i$, $\nu \geqslant 1$, is implicitly used. Also, in the proof of Theorem 2 of [6] an inequality is obtained incorrectly. The notation of [6] is retained and we utilize much of the work done there.

* This research was partially supported by National Science Foundation Grant GP-28313.

Copyright © 1973 by Academic Press, Inc.
All rights of reproduction in any form reserved.

2. Kuzmin's Theorem

For $F \in \mathscr{F}$, we partition $(0, 1)_F^n$ by the sets $T^\nu B_\nu(x)$, $\nu \geqslant 1$, $x \in (0, 1)_F^n$. This partition is assumed countable and is denoted by $\{A_i\}_{i \geqslant 1}$. Next we define

$$E_{i,\nu} = \{\langle k \rangle = (k^{(1)}, k^{(2)}, \dots, k^{(\nu)}) \mid T^\nu B(k) \supset A_i\}.$$

We remark that, if $B_\nu(\langle k \rangle)$ is a proper cylinder of order ν, then $\langle k \rangle \in E_{i,\nu}$.

Our first lemma generalizes Lemma 1 of [6].

LEMMA 1. *Suppose $F \in \mathscr{F}$. Let Ψ_0 be given and define Ψ_ν, $\nu \geqslant 1$, by*

$$\Psi_\nu(x) = \sum_{k \in E_{i,1}} \Psi_{\nu-1}(f_k(x)) \mid J_{f_k}(x)|, \qquad x \in A_i. \tag{1}$$

Then

$$\Psi_\nu(x) = \sum_{k \in E_{i,\nu}} \Psi_0(f_\nu(x)) \mid J_\nu(x)|, \qquad x \in A_i, \tag{2}$$

where $f_\nu(x) = \prod_{i=1}^{\nu} \circ f_{k^{(i)}}(x)$ and $J_\nu = J_{f_\nu}$.

Proof. Our proof is by induction. Formula (2) is clearly true for $\nu = 1$. Assume that (2) holds for ν. Then, by definition,

$$\Psi_{\nu+1}(x) = \sum_{k \in E_{i,1}} \sum_{E_{j,\nu}} \Psi_0(f_\nu(f_k(x))) \mid J_\nu(f_k(x))| \mid J_{f_k}(x)|,$$

$x \in A_i$, where j in the second summation satisfies $f_k(x) \in A_j$.

Since $f_k(x) \in A_j$ and $E_{j,\nu}$ is the set of all $(k^{(1)}, k^{(2)}, \dots, k^{(\nu)})$ such that $T^\nu B_\nu \supset A_j$, we have that $(k^{(1)}, k^{(2)}, \dots, k^{(\nu)}, k)$ runs over all sequences ending in k such that $f_{k^{(1)}} \circ f_{k^{(2)}} \circ \cdots \circ f_{k^{(\nu)}} \circ f_k(x)$ is defined. Therefore, we have $x \in T^{\nu+1}B(k^{(1)}, \dots, k^{(\nu)}, k)$ which implies $T^{\nu+1}B(k^{(1)}, \dots, k^{(\nu)}, k) \supset A_i$. Conversely, if $T^{\nu+1}B(k^{(1)}, \dots, k^{(\nu+1)}) \supset A_i$ we have

$$TB(k^{(\nu+1)}) \supset T^{\nu+1}B(k^{(1)}, \dots, k^{(\nu+1)}) \supset A_i.$$

Therefore $k^{(\nu+1)} \in E_{i,1}$. We have thus shown that $(k^{(1)}, \dots, k^{(\nu)}, k)$ runs over all sequences of length $\nu + 1$ in $E_{i,\nu+1}$.

LEMMA 2. *Suppose $F \in \mathscr{F}$ and define $\{\Psi_\nu\}_{\nu \geqslant 0}$ as in Lemma 1. Then, if $\{A_i\}$ is countable,*

$$\int_{(0,1)^n} \Psi_\nu(x)\, dx = \int_{(0,1)^n} \Psi_0(x)\, dx, \qquad \nu \geqslant 1.$$

Proof. The proof proceeds as in Corollary 6.1 of [5].

We next state our main theorem. The proof uses the technique of the proof of Theorem 2 in [6] but the difficulties mentioned in the introduction are corrected here. In addition, we make clear the uniformity of convergence to ρ.

THEOREM 1. *Let $F \in \mathscr{F}$ satisfy conditions (c), (q), and (L), and $\sigma(\nu) = \sup_x \operatorname{diam} B_\nu(x) \to 0$ as $\nu \to \infty$. In addition, assume the partition $\{A_i\}$ countable and that there is an $N \geqslant 1$ such that $B_\nu \cap T^\mu B_\mu \neq \varnothing$ implies $B_\nu \subseteq T^\mu B_\mu$ for all $\mu \geqslant 1$ and $\nu \geqslant N$. We suppose there exist constants A and D such that*

$$|(\partial f_\nu)_k / \partial x_j| \leqslant A \qquad (x \in A_i)$$

uniformly in ν, k, and j; and

$$\big|\, |J_\nu(x)| - |J_\nu(y)| \,\big| \leqslant Dm(B_\nu) \| x - y \| \qquad (x, y \in B_\nu),$$

uniformly in ν. Let $\{\Psi_\nu\}_{\nu > 1}$ be the sequence of functions recursively defined in Lemma 1 with Ψ_0 satisfying

$$0 < m \leqslant \Psi_0 \leqslant M \qquad \textit{and} \qquad |\Psi_0(x) - \Psi_0(y)| \leqslant N \| x - y \|.$$

Then

$$|\Psi_\nu(x) - a\rho(x)| < b\sigma(\nu) \tag{3}$$

where ρ is the density of the invariant measure for F, $a = \int \Psi_0 \, dm$ and b are constants.

Proof. As in the proof of Theorem 2 of [6] we have

$$|\Psi_\nu(x) - \Psi_\nu(y)| \leqslant C_1 \| x - y \| \tag{4}$$

for $x, y \in A_i$, and

$$g_0 \Psi_\nu(x) < \Psi_{\nu+\mu}(x) < G_0 \Psi_\nu(x)$$

uniformly in x, μ, and ν.

We define

$$\phi_\nu(x) = \Psi_{\mu+\nu}(x) - g_0 \Psi_\nu(x)$$

and

$$\zeta_\nu(x) = G_0 \Psi_\nu(x) - \Psi_{\mu+\nu}(x).$$

By Lemma 1 and condition (C)

$$\phi_\nu(x) \geqslant C^{-1} \sum^{(i)} \phi_0(f_\nu(x)) \, m(B_\nu).$$

We let $\mathscr{P}_\nu = \bigcup B_\nu$ where the union is over all proper cylinders of order ν. There exists $y_\nu' \in B_\nu \subseteq (0, 1)_F^n$ such that

$$\int_{B_\nu} \phi_0 \leqslant \phi_0(y_\nu')\, m(B_\nu).$$

Hence

$$-\int_{\mathscr{P}_\nu} \phi_0 \geqslant -\sum_{\langle k\rangle} \phi_0(y_\nu')\, m(B_\nu).$$

Then for $\nu \geqslant N$

$$\begin{aligned}\phi_\nu(x) - C^{-1}\int_{\mathscr{P}_\nu} \phi_0(y)\, dy &\geqslant C^{-1}\sum_{\langle k\rangle} \{\phi_0(f_\nu(x)) - \phi_0(y_\nu')\}\, m(B_\nu)\\ &\geqslant -C^{-1}C_1(1+g_0)\,\sigma(\nu)\sum_{\langle k\rangle} m(B_\nu)\\ &\geqslant -C_2\sigma(\nu).\end{aligned}$$

Here the sums are taken over all proper cylinders of order ν and the second inequality is by (4) applied to $\phi_0 = \Psi_\mu - g_0\phi_0$. For the second inequality we use the fact that $f_\nu(x)$ and y_ν' are contained in the same element A_i of the partition. This follows from the assumption that $B_\nu \cap T^\mu B_\mu \neq \varnothing$ implies $B_\nu \subset T^\mu B_\mu$ for $\mu \geqslant 1$ and $\nu \geqslant N$. This was tacitly used in [5] and [6]. That is,

$$\Psi_{\mu+\nu}(x) - g_0\Psi_\nu(x) \geqslant C^{-1}\int_{\mathscr{P}_\nu} (\Psi_\mu - g_0\Psi_0)\, dm - C_2\sigma(\nu).$$

In the same manner we obtain

$$G_0\Psi_\nu(x) - \Psi_{\mu+\nu}(x) \geqslant C^{-1}\int_{\mathscr{P}_\nu} (G_0\Psi_0 - \Psi_\mu)\, dm + C_3\sigma(\nu).$$

The expressions for l_i and l_i^1 in [6] are functions of g_0 which makes ν_0 a function of g_0. Therefore, when we iterate, it is possible that $\nu_0(g_1) > \nu_0(g_0)$ and so on. From this observation it is clear that we cannot let $r \to \infty$ without $\nu \to \infty$, and we see that inequality (9) of [6] was obtained incorrectly. We now give another derivation of a similar inequality.

We define

$$\alpha(\nu) = 1 - (M_1C)^{-1}\int_{\mathscr{P}_\nu} \Psi_0\, dm,$$

$$\beta(\nu) = (M_1C)^{-1}\int_{\mathscr{P}_\nu} \Psi_\mu\, dm - \frac{C_3\sigma(\nu)}{m_1},$$

and

$$\delta(\nu) = (M_1C)^{-1}\int_{\mathscr{P}_\nu} \Psi_\mu \, dm + \frac{C_3\sigma(\nu)}{m_1}.$$

From the equations

$$g_{r+1} = g_r + (M_1C)^{-1}\int_{\mathscr{P}_\nu} (\Psi_\mu - g_r\Psi_0)\, dm - \frac{C_2\sigma(\nu)}{m_1}$$

and

$$G_{r+1} = G_r - (M_1C)^{-1}\int_{\mathscr{P}_\nu} (G_0\Psi_0 - \Psi_\mu)\, dm + \frac{C_3\sigma(\nu)}{m_1},$$

we obtain the recursion relations

$$g_{r+1} = g_r\alpha(\nu) + \beta(\nu)$$

and

$$G_{r+1} = G_r\alpha(\nu) + \delta(\nu).$$

We have the following inequalities

$$\alpha(\nu) \leqslant 1 - m_1q(M_1C)^{-1} < 1,$$
$$\beta(\nu) \geqslant \frac{m_1q}{M_1C} - \frac{C_2\sigma(\nu)}{m_1} > 0,$$
$$\delta(\nu) \geqslant m_1q(M_1C)^{-1} > 0$$

for $\nu \geqslant \nu_0$. We have assumed, without loss of generality, that $C > 1$ and have used the inequality $m(\mathscr{P}_\nu) \geqslant q$. It now follows that

$$g(\nu) = \lim_{r\to\infty} g_r = Q(\nu, \mu) + O(\sigma(\nu))$$

and

$$G(\nu) = \lim_{r\to\infty} G_r = Q(\nu, \mu) + O(\sigma(\nu))$$

where

$$Q(\nu, \mu) = \left(\int_{\mathscr{P}_\nu} \Psi_\mu \, dm\right)\Big/\int_{\mathscr{P}_\nu} \Psi_0 \, dm.$$

From

$$g_r\Psi_\nu(x) \leqslant \Psi_{\mu+\nu}(x) \leqslant G_r\Psi_\nu(x)$$

we have

$$g(\nu)\,\Psi_\nu(x) \leqslant \Psi_{\mu+\nu}(x) \leqslant G(\nu)\,\Psi_\nu(x)$$

and hence

$$|\, Q(\nu, \mu)\,\Psi_\nu(x) - \Psi_{\mu+\nu}(x)| = O(\sigma(\nu)).$$

Now we integrate over $(0, 1)^n$ and, by Lemma 2, obtain

$$| Q(\nu, \mu) - 1 | \leqslant O(\sigma(\nu)).$$

Thus

$$| \Psi_\nu(x) - \Psi_{\mu+\nu}(x)| = O(\sigma(\nu))$$

and the proof is completed as in Theorem 2 of [6].

The assumption of Theorem 1 that there is an $N \geqslant 1$ such that $B_\nu \cap T^\mu B_\mu \neq \varnothing$ implies $B_\nu \subseteq T^\mu B_\mu$ for all $\mu \geqslant 1$ and $\nu \geqslant N$ is implied by the following condition: There is an $N \geqslant 1$ such that $TB_{\nu+1} = B_\nu$ for $\nu \geqslant N$. To see this, take $B_\mu = \bigcup B_{\mu+\nu}$. Then $T^\mu B_\mu = \bigcup T^\mu B_{\mu+\nu} = \bigcup B_\nu$. Since distinct cylinders are disjoint we are done. This observation will be used in Section 4 in dealing with the Jacobi algorithm.

3. Applications

As a first application of Theorem 1 we generalize Theorem 6 of [4] and Theorem 6.2 of [5]. The rate of convergence is the same as given in [4].

THEOREM 2 (Gauss). *Let F satisfy the assumptions of Theorem* 1. *If μ denotes the invariant measure, then*

$$| m(T^{-\nu}(E)) - \mu(E)| < bm(E)\,\sigma(\nu) \tag{5}$$

for all Borel sets E.

Proof. Define $\Psi_0(x) = 1$. By Theorem 1, we obtain

$$| \Psi_\nu(x) - \rho(x)| < b\sigma(\nu), \tag{6}$$

since $a = \int \Psi_0 = 1$. We multiply (6) by I_E and integrate to obtain

$$\left| \int I_E \Psi_\nu \, dm - \mu(E) \right| < bm(E)\,\sigma(\nu).$$

Calculation similar to that in the proof of Theorem 6.2 in [4] shows

$$\int I_E \Psi_\nu \, dm = mT^{-\nu}E$$

and (5) follows.

The next application is a generalization of Theorem 7 of [4] and

Theorem 6.3 of [5]. We again obtain a rate of convergence identical with that in [4]. This result gives us a rate on the mixing.

THEOREM 3. *Let F satisfy the assumptions of Theorem* 1. *Then for all Borel sets F and cylinders E*

$$| \mu(E \cap T^{-\nu}F) - \mu(E)\,\mu(F)| \leqslant \frac{\langle Cb \rangle}{q}\, \mu(E)\,\mu(F)\,\sigma([\nu/2]). \tag{7}$$

Proof. First let B_{ν_0} be any proper cylinder and define

$$\Psi_0(x) = \frac{I_{B_{\nu_0}}(x)}{\mu(B_{\nu_0})}\,\rho(x).$$

The sequence $\{\Psi_\nu\}_{\nu \geqslant \nu_0}$ satisfies the hypotheses of Theorem 1. Our proof is completed as in Theorem 6.3 of [5], and we obtain

$$| \mu(B_{\nu_0} \cap T^{-\nu-\nu_0}(F)) - \mu(B_{\nu_0})\,\mu(F)| \leqslant \frac{\langle Cb \rangle}{q}\, \mu(B_{\nu_0})\,\mu(F)\,\sigma(\nu). \tag{8}$$

A similar inequality holds for improper cylinders. The proof of inequality (7) from inequality (8) follows as in Folgerung 1 in [4].

4. THE JACOBI ALGORITHM

The main purpose of this paper is to prove a Kuzmin theorem which applies to the Jacobi algorithm. In earlier statements of Theorem 1 [5], [6], it has been assumed that $\lim_{\nu\to\infty} \operatorname{diam} B_\nu(x) = 0$ a.e. However, it is necessary to replace that assumption by $\sigma(\nu) \to 0$ as $\nu \to \infty$. In a paper of Paley and Ursell [1] it is shown, for $n = 2$, that $\sigma(\nu) = O(\theta^{-\nu})$, where θ is the unique root of $x^3 - x^2 - 1 = 0$ with $1 < \theta$. Unfortunately no proof is present for their results on the Jacobi algorithm with $n > 2$. It is possible, however, to give a proof that $\sigma(\nu) \to 0$ as $\nu \to \infty$ which is valid for all $n \geqslant 1$. This proof, which will not be included here, is somewhat lengthly and depends on additional facts concerning the Jacobi algorithm.

An earlier paper [6] handled all other assumptions of Theorem 1 with the exception of proving $\{A_i\}$ countable and showing $TB_{\nu+1} = B_\nu$ for all $\nu \geqslant n - 1$. Thus we must consider the partition generated by all sets $T^\nu B_\nu$.

We consider the transformation

$$T(x_1, x_2, \ldots, x_n) = \left(\frac{x_2}{x_1} - \left[\frac{x_2}{x_1}\right], \ldots, \frac{x_n}{x_1} - \left[\frac{x_n}{x_1}\right], \frac{1}{x_1} - \left[\frac{1}{x_1}\right]\right) \tag{9}$$

which is defined on the unit cube after the deletion of an appropriate set of measure zero. This transformation corresponds to an F-expansion with

$$F(x_1, \ldots, x_n) = \left(\frac{1}{x_n}, \frac{x_1}{x_n}, \ldots, \frac{x_{n-1}}{x_n}\right).$$

We denote

$$k^{(1)}(x) = (k_1^{(1)}, \ldots, k_n^{(1)}) = ([x_2/x_1], \ldots, [1/x_1])$$

and

$$k^{(s)}(x) = k^{(1)}(T^{s-1}(x)), \qquad s \geqslant 1.$$

In this fashion we obtain a sequence of n-dimensional digits

$$k^{(1)}, k^{(2)}, \ldots$$

which is in one to one correspondence with the point x. The restrictions (see [1, 2]) on this sequence are as follows:

$$k_n^{(s)} \geqslant 1, \qquad k_n^{(s)} \geqslant k_i^{(s)} \geqslant 0, \qquad i = 1, 2, \ldots, n-1, \tag{10}$$

and

$$\text{if} \quad k_i^{(s)} = k_n^{(s)}, \qquad \text{then} \qquad k_{i-1}^{(s+1)} \leqslant k_{n-1}^{(s+1)}. \tag{11}$$

If we then obtain $k_{i-1}^{(s+1)} = k_{n-1}^{(s+1)}$, we have $k_{i-2}^{(s+2)} \leqslant k_{n-2}^{(s+2)}$, and so on. We formally set $k_0^{(t)} = 1$ and $k_r^{(t)} = 0$ for $r < 0$.

The dependence of the digits is of length $n - 1$. That is, the occurrence of a digit satisfying (10) depends only on the $n - 1$ preceeding digits. Thus, we have $TB_{\nu+1} = B_\nu$ for $\nu \geqslant n - 1$. The structure of T^{-1}, namely

$$x_1 = \frac{1}{k_n + y_n}, \qquad x_2 = \frac{k_1 + y_1}{k_n + y_n}, \ldots, \qquad x_n = \frac{k_{n-1} + y_{n-1}}{k_n + y_n},$$

shows that the restrictions (11) correspond to ones of the type $y_j < y_{j+h}$. Therefore, the sets $T^\nu B_\nu$ partition the unit cube into $n!$ simplices A_i defined by the intersections of the sets $\{(x_1, \ldots, x_n) \mid x_j < x_{j+h}\}$ and their complements. This completes our proof that Theorem 1 (and, hence, Theorems 2 and 3) applies to the Jacobi algorithm.

Acknowledgment

The authors wish to thank Dr. Fischer of the University of Salzburg for pointing out certain errors in our earlier proof of Theorem 1

References

1. R. Paley and H. D. Ursell, Continued fractions in several dimensions, *Proc. Cambridge Philos. Soc.* **26** (1930), 127–144.
2. O. Perron, Grundlagen für eine Theorie des Jacobischen Kettenbruch-Algorithmus, *Math. Ann.* **64** (1907), 1–76.
3. F. Schweiger, Metrische Sätze über der Jacobischen Algorithmus, *Monatsh. Math.* **69** (1965), 243–255.
4. F. Schweiger, Metrische Theorie einer Klasse Zahlentheoretischer Transformationen, *Acta Arith.* **15** (1968), 1–18.
5. M. Waterman, Some ergodic properties of multi-dimensional *F*-expansions, *Z. Wahrscheinlichkeitstheorie Verw. Gebiete* **16** (1970), 77–103.
6. M. Waterman, A Kuzmin theorem for a class of number theoretic endomorphisms, *Acta Arith.* **19** (1971), 59–69.

Printed by the St Catherine Press Ltd., Tempelhof 37, Bruges, Belgium.