

A Kuzmin theorem for a class of number theoretic endomorphisms

by

MICHAEL S. WATERMAN (Pocatello, Idaho)

Recently several papers ([3], [4], [5], [6], [7]) have been concerned with generalizations of a 1928 theorem of Kuzmin. His result gives a rate of $e^{-N^{\sqrt{n}}}$ for the convergence of the iteration of an arbitrary function to the invariant measure for the continued fraction. The present paper gives a generalized Kuzmin theorem for a class of multi-dimensional F -expansions which includes the n -dimensional continued fraction. An earlier paper ([6]) presented such a theorem with a rate of $(e^{-N^{\sqrt{n}}} + \sigma(\sqrt{n}))$. Our present theorem improves the rate to $\sigma(\sqrt{n})$.

Our F -expansions were first considered in [6], and we include a short summary of notation and assumptions here. Let A be a fixed convex subset of R^n . Suppose F is a one-to-one continuous map of A onto $(0, 1)^n$. We assume $J_F(\cdot)$, the Jacobian of F , exists, the components of F have continuous first order partial derivatives, and $J_F(x) \neq 0$ for almost all $x \in A$. Let $D = F^{-1}$, $T(x) = D(x) - [D(x)]$, and $a_\nu(x) = [D(T^{\nu-1}x)]$ (where $[z] = ([z_1], [z_2], \dots, [z_n])$). We call $a_\nu(x)$ the ν -th coordinate of the F -expansion of x . Letting

$$(0, 1)_F^n = \{x \in (0, 1)^n : T^\nu(x) \in (0, 1)^n \text{ for all } \nu \geq 1\},$$

we impose the assumption $m((0, 1)_F^n) = 1$, where m denotes n -dimensional Lebesgue measure. We will write $F \in \mathcal{F}$ to indicate the satisfaction of these assumptions.

We define the cylinder of order ν generated by a realizable set of coordinates k_1, k_2, \dots, k_ν as

$$B_\nu = B_\nu(k_1, k_2, \dots, k_\nu) = \{x \in (0, 1)_F^n : a_i(x) = k_i, i = 1, \dots, \nu\},$$

and the cylinder of order ν generated by $x \in (0, 1)_F^n$ as

$$B_\nu = B_\nu(x) = \{y \in (0, 1)_F^n : a_i(y) = a_i(x), i = 1, \dots, \nu\}.$$

Of course $T(B_\nu(k_1, k_2, \dots, k_\nu)) \subseteq B_{\nu-1}(k_2, \dots, k_\nu)$ so that T is the shift on the coordinates of the expansion. If B_ν is generated by k_1, k_2, \dots, k_ν and we let $f_{k_i}(t) = F(a_i + t)$, then we define

$$f_\nu(t) = f_{k_1} \circ f_{k_2} \circ \dots \circ f_{k_\nu}(t) = \prod_{i=1}^{\nu} \circ f_{k_i}(t), \quad t \in T^\nu B_\nu.$$

Below are three additional assumptions on F . The first generalizes condition (C) of Renyi [2].

$$(C) \quad \frac{\sup_{t \in T^\nu B_\nu} |J_{f_\nu}(t)|}{\inf_{t \in T^\nu B_\nu} |J_{f_\nu}(t)|} \leq C < +\infty$$

uniformly where f_ν runs over all $\nu \geq 1$ and all realizable cylinders $B_\nu(k_1, k_2, \dots, k_\nu)$.

If $m(T^\nu B_\nu) = 1$ we say that B_ν is *proper*; otherwise B_ν is said to be *improper*. Difficulties with improper cylinders necessitate the next two conditions.

$$(L) \quad 0 < L \leq m(T^\nu B_\nu(x)) \quad \text{for all } x \in (0, 1)_F^\nu, \quad \nu \geq 1.$$

For each $B_\nu(x)$, there exists $\hat{B}_{\nu+1}$, a collection of proper cylinders of order $\nu+1$ contained in $B_\nu(x)$, such that

$$(q) \quad 0 < q \leq \frac{m(\hat{B}_{\nu+1})}{m(B_\nu(x))} \quad \text{for all } x \in (0, 1)_F^\nu, \quad \nu \geq 1.$$

The following theorem appears in [6] and is basic to the problem considered here.

THEOREM 1. *Suppose $F \in \mathcal{F}$ satisfies condition (C), condition (L), and condition (q). Then there exists a unique probability measure μ on $(0, 1)_F^\infty$ such that $\mu \ll m$ and T is a measure preserving transformation for μ . If we let $\rho(x) = \frac{d\mu}{dm}(x)$, we have*

$$\frac{q}{C} \leq \rho(x) \leq \frac{C}{L}.$$

Of course we could conclude $\mu \sim m$ but we will only need $\mu \ll m$ in our proof. Also, adding the assumption $m\{\alpha: \text{diam } B_\nu(x) \rightarrow 0\} = 1$ allows us to conclude T ergodic. This assumption is included in Theorem 2 below.

To formulate a Kuzmin theorem for \mathcal{F} we need to partition $(0, 1)_F^\infty$. For each cylinder of order 1, $B(k)$, we have $TB(k) \subset (0, 1)_F^\infty$. We use the collection $TB(k)$ to partition $(0, 1)_F^\infty$ and assume the partition is essen-

tially countable. Denote this partition by $\{A_i\}_{i \geq 1}$. With each A_i we associate

$$\mathcal{E}_i = \{k: TB(k) \supset A_i\}.$$

This allows us to calculate

$$(1) \quad \varrho(x) = \sum_{k \in \mathcal{E}_i} \varrho(f_k(x)) |J_{f_k}(x)|, \quad x \in A_i.$$

The two lemmas below are taken from [6] and depend only on the properties of f_k . Both are related to the form of equation (1).

LEMMA 1. Suppose $F \in \mathcal{F}$ and assume $TB_{v+1} = B_v$, $v \geq 1$. Let Ψ_0 be given and Ψ_v be defined by

$$\Psi_v(x) = \sum_{k \in \mathcal{E}_i} \Psi_{v-1}(f_k(x)) |J_{f_k}(x)|, \quad x \in A_i \quad (i = 1, 2, \dots).$$

Then

$$\Psi_v(x) = \sum^{(i)} \Psi_0(f_v(x)) |J_{f_v}(x)|, \quad x \in A_i \quad (i = 1, 2, \dots),$$

where the last summation is over all realizable cylinders (k_1, \dots, k_v) where $k_v \in \mathcal{E}_i$.

LEMMA 2. Let $F, \{\Psi_v\}_{v \geq 0}$ be as in Lemma 1. Then

$$\int_{(0,1)^n} \Psi_v(x) dx = \int_{(0,1)^n} \Psi_0(x) dx \quad \text{for } v \geq 1.$$

The theorem below was motivated by a paper of Schweiger ([5]) in which he proves a Kuzmin theorem for a class of F -expansions which has the restriction that all cylinders be proper. Since the n -dimensional Jacobi algorithm has improper cylinders, it was not included. Difficulties are encountered in our proof which do not exist if all cylinders are proper. The assumption $\lim_{v \rightarrow \infty} \text{diam } B_v(x) = 0$ almost everywhere is to insure our F -expansions converge and $\sigma(v) \rightarrow 0$ as $v \rightarrow \infty$. To circumvent notational difficulty, we will tacitly assume $x \in (0, 1)_{\mathbb{F}}^n$ implies $\lim_{v \rightarrow \infty} \text{diam } B_v(x) = 0$, which involves the deletion of a set of measure zero from the conclusion of our theorem.

THEOREM 2. Let $F \in \mathcal{F}$ satisfy conditions (C), (q), (L) and $m\{x: \lim_{v \rightarrow \infty} \text{diam } B_v(x) = 0\} = 1$. In addition, suppose $TB_{v+1}(x) = B_v(x)$, $v \geq 1$, $x \in (0, 1)_{\mathbb{F}}^n$. Assume there is a constant A such that

$$\left| \frac{\partial(f_v)_k}{\partial x_j} \right| \leq A \quad \text{uniformly in } v, k, \text{ and } j.$$

Also suppose there exists a constant D such that

$$|J_v(x) - J_v(y)| \leq Dm(B_v)\|x - y\| \quad (x, y \in T^v B_v)$$

uniformly in v . Let $\{\Psi_v\}_{v \geq 0}$ be a sequence of functions recursively defined by

$$\Psi_v(x) = \sum_{k \in \mathcal{E}_i} \Psi_{v-1}(f_k(x)) |J_{f_k}(x)|, \quad x \in A_i, \quad i \geq 1,$$

where Ψ_0 is an arbitrary measurable function satisfying

$$0 < m \leq \Psi_0(x) \leq M$$

and

$$|\Psi_0(x) - \Psi_0(y)| \leq N\|x - y\|.$$

Then

$$|\Psi_v(x) - a_\varrho(x)| < b\sigma(v)$$

where ϱ is the density of the invariant measure for F ,

$$a = \int_{(0,1)^n} \Psi_0(x) dx \text{ and } b \text{ are constants,}$$

and

$$\sigma(v) = \sup \{ \text{diam } B_v(y) : y \in (0, 1)_F^n \}.$$

Proof. By Lemma 1, we have

$$\Psi_v(x) = \sum^{(i)} \Psi_0(f_v(x)) |J_v(x)|, \quad x \in A_i.$$

Using this formula and the bounds assumed above, we can show, for $x, y \in A_i$,

$$|\Psi_v(x) - \Psi_v(y)| \leq N \sum^{(i)} \|f_v(x) - f_v(y)\| \cdot |J_v(x)| + MD\|x - y\| \sum^{(i)} m(B_v).$$

Now, applying the mean value theorem to the components of f_v , we obtain

$$\|f_v(x) - f_v(y)\| \leq nA\|x - y\|,$$

and use of condition (C) and condition (L) yields

$$\sum^{(i)} |J_v(x)| \leq \frac{C}{L}.$$

Therefore

$$|\Psi_v(x) - \Psi_v(y)| \leq (NnACL^{-1} + MD)\|x - y\| = C_1\|x - y\| \quad \text{for } x, y \in A_i.$$

Application of Lemma 1, equation (1), and condition (O) yields

$$(2) \quad 0 < m_1 = mLqC^{-2} < \Psi_\nu(x) < M_1 = C^2(MLq)^{-1}, \quad \text{uniformly in } \nu, x.$$

This allows us to obtain $0 < g_0 < G_0$ such that

$$(3) \quad g_0\Psi_\nu(x) < \Psi_{\mu+\nu}(x) < G_0\Psi_\nu(x) \quad \text{uniformly in } x, \mu, \text{ and } \nu.$$

For $\mu \geq 0$, we define

$$(4) \quad \Phi_\nu(x) = \Psi_{\mu+\nu}(x) - g_0\Psi_\nu(x)$$

and

$$(5) \quad \zeta_\nu(x) = G_0\Psi_\nu(x) - \Psi_{\mu+\nu}(x).$$

By application of Lemma 1, we have

$$\Phi_\nu(x) = \sum^{(i)} \Phi_0(f_\nu(x)) |J_\nu(x)|$$

and

$$\zeta_\nu(x) = \sum^{(i)} \zeta_0(f_\nu(x)) |J_\nu(x)|, \quad x \in A_i.$$

We obtain

$$\Phi_\nu(x) \geq C^{-1} \sum^{(i)} \Phi_0(f_\nu(x)) m(B_\nu)$$

from condition (C). We let

$$\mathcal{C}'_i = \bigcup_{k_r \in \mathcal{C}'_i} B_r(k_1, \dots, k_r).$$

By the mean value theorem for integrals

$$\int_{\mathcal{C}'_i} \Phi_0(y) dy = \sum^{(i)} \Phi_0(y'_i) m(B_\nu).$$

Therefore

$$\begin{aligned} \Phi_\nu(x) - C^{-1} \int_{\mathcal{C}'_i} \Phi_0(y) dy &\geq C^{-1} \sum^{(i)} \{\Phi_0(f_\nu(x)) - \Phi_0(y'_i)\} m(B_\nu) \\ &\geq -C^{-1} C_1 (1 + g_0) \sigma(\nu) \sum^{(i)} m(B_\nu) \geq -C_2 \sigma(\nu). \end{aligned}$$

That is,

$$\Psi_{\nu+\mu}(x) - g_0\Psi_\nu(x) \geq C^{-1} \int_{\mathcal{C}'_i} (\Psi_\mu(x) - g_0\Psi_0(x)) dx - C_2 \sigma(\nu).$$

In the same manner we obtain ($x \in A_i$)

$$\zeta_\nu \geq C^{-1} \sum^{(i)} \zeta_0(f_\nu(x)) m(B_\nu),$$

$$\int_{\mathcal{C}_i^\nu} \zeta_0(x) dx = \sum^{(i)} \zeta_0(y'_\nu) m(B_\nu),$$

and

$$G_0 \Psi_\nu(x) - \Psi_{\nu+\mu}(x) \geq C^{-1} \int_{\mathcal{C}_i^\nu} (G_0 \Psi_0(x) - \Psi_\mu(x)) dx - C_3 \sigma(\nu).$$

Letting

$$l_i = C^{-1} \int_{\mathcal{C}_i^\nu} (\Psi_\mu(x) - g_0 \Psi_0(x)) dx,$$

and

$$l'_i = C^{-1} \int_{\mathcal{C}_i^\nu} (G_0 \Psi_0(x) - \Psi_\mu(x)) dx,$$

we can show

$$(6) \quad \Psi_{\nu+\mu}(x) \geq \Psi_\nu(x) \left(g_0 + \frac{l_i}{M_1} - (m_1)^{-1} C_2 \sigma(\nu) \right) = g_1 \Psi_\nu(x)$$

and

$$(7) \quad \Psi_{\nu+\mu}(x) \leq \Psi_\nu(x) \left(G_0 - \frac{l'_i}{M_1} + (m_1)^{-1} C_3 \sigma(\nu) \right) = G_1 \Psi_\nu(x).$$

There exists ν_0 such that for $\nu \geq \nu_0$,

$$g_0 < g_1 < G_1 < G_0.$$

Now

$$l_i + l'_i = C^{-1} \int_{\mathcal{C}_i^\nu} (G_0 - g_0) \Psi_0(x) dx \geq C^{-1} (G_0 - g_0) m_1 \sum^{(i)} m(B_\nu)$$

$$\geq C^{-1} (G_0 - g_0) m_1 \sum' m(B_\nu) = C^{-1} m q (G_0 - g_0),$$

where \sum' denotes summation over proper cylinders of order ν and the last inequality is by condition (q). The importance of this bound is its independence of both μ and ν .

From these results we obtain

$$(8) \quad G_1 - g_1 = G_0 - g_0 - \frac{1}{M_1} (l'_i + l_i) + (m_1)^{-1} (C_2 + C_3) \sigma(\nu)$$

$$\leq (G_0 - g_0) (1 - m_1 q (C M_1)^{-1}) + C_4 \sigma(\nu).$$

We note that

$$0 < \lambda = 1 - m_1 q (CM_1)^{-1} < 1,$$

since without loss of generality $C > 1$.

Now we summarize the result just obtained. From

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

we have proved (for $\nu \geq \nu_0$)

$$g_1 \Psi_\nu(x) < \Psi_{\nu+\mu}(x) < G_1 \Psi_\nu(x)$$

where

$$g_0 < g_1 < G_1 < G_0$$

and

$$G_1 - g_1 \leq (G_0 - g_0)\lambda + C_4 \sigma(\nu).$$

The argument for g_1 and G_1 can be repeated to obtain

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

where

$$g_0 < g_1 < \dots < g_r < G_r < \dots < G_1 < G_0$$

and

$$\begin{aligned} G_r - g_r &\leq (G_{r-1} - g_{r-1})\lambda + C_4 \sigma(\nu) \\ &\leq \lambda^r (G_0 - g_0) + C_4 \sigma(\nu) (1 + \lambda + \dots + \lambda^r) \\ &\leq \lambda^r (G_0 - g_0) + \frac{C_4}{1 - \lambda} \sigma(\nu). \end{aligned}$$

It should be emphasized that G_r and g_r are functionally dependent on r, i, ν, μ , and Ψ_0 .

Now

$$\lim_{\nu, r \rightarrow \infty} (G_r - g_r) = 0$$

implies

$$\lim_{\nu, r \rightarrow \infty} G_r = \lim_{\nu, r \rightarrow \infty} g_r = Q(\mu).$$

Thus we can write

$$|\Psi_{\nu+\mu}(x) - Q(\mu)\Psi_\nu(x)| < (G_r - g_r)\Psi_\nu(x) \leq M_1 C \left(\lambda^r (G_0 - g_0) + \frac{C_4}{1 - \lambda} \sigma(\nu) \right)$$

Finally, since $Q(\mu) \equiv 1$, we have by (9)

$$(11) \quad |\Psi_{\nu+\mu}(x) - \Psi_{\nu}(x)| < b\sigma(\nu), \quad \nu \geq \nu_0.$$

Therefore $\{\Psi_{\nu}(x)\}_{\nu \geq 1}$ is a Cauchy sequence. Letting $\Psi(x) = \lim_{\nu \rightarrow \infty} \Psi_{\nu}(x)$, $a = \int \Psi(x) = \int \Psi_0(x)$, and $\varrho(x) = a^{-1}\Psi(x)$, we have

$$|\Psi_{\nu}(x) - a\varrho(x)| < b\varrho(\nu).$$

Since $\varrho(x)$ satisfies (1), $\int \varrho(x) dx = 1$, ϱ is the unique invariant measure $\ll m$. This completes the proof.

The following corollary corresponds to F. Schweiger's result ([5]) of $\varrho \in \text{Lip}^1(0, 1)^n$.

COROLLARY 1. *The density function, $\varrho(\cdot)$, of Theorem 2 satisfies a Lipschitz condition of order 1 on each of the sets A_i . That is,*

$$|\varrho(x) - \varrho(y)| \leq K \|x - y\|, \quad x, y \in A_i.$$

Proof. The result follows directly from $\Psi_{\nu}(\cdot) \in \text{Lip}^1(A_i)$, and the conclusion of Theorem 2. Note that K has the same value for each of the A_i .

COROLLARY 2. *Let $F \in \mathcal{F}$ and Ψ_0 be as in Theorem 2. Then for all $\mu \geq 0$ and $i \geq 1$,*

$$\lim_{\nu \rightarrow \infty} \frac{\int_{\mathcal{C}_i^{\nu}} \Psi_{\mu}(x) dx}{\int_{\mathcal{C}_i^{\nu}} \Psi_0(x) dx} = 1.$$

Proof. We remark that if $\mathcal{C}_i^{\nu} \equiv (0, 1)^n_F$ (for fixed i), then the result is obvious from Lemma 2. In general, however, it seems necessary to return to an explicit determination of g_r .

$$g_0 = \frac{m_1}{2M_1},$$

$$\begin{aligned} g_r &= g_{r-1} (1 - (CM_1)^{-1}) \int_{\mathcal{C}_i^{\nu}} \Psi_0(x) dx + (CM_1)^{-1} \int_{\mathcal{C}_i^{\nu}} \Psi_{\mu}(x) dx - C_5 \sigma(\nu) \\ &= ag_{r-1} + b = a^r g_0 + b(1 + a + \dots + a^{r-1}). \end{aligned}$$

By choosing $\nu \geq \nu_3$ and making M_1 sufficiently large, we have $0 < a, b < 1$. Therefore,

$$(12) \quad \lim_{r \rightarrow \infty} g_r = \frac{b}{1-a} = \frac{(CM_1)^{-1} \int_{\mathcal{C}_i^{\nu}} \Psi_{\mu}(x) dx - C_5 \sigma(\nu)}{(CM_1)^{-1} \int_{\mathcal{C}_i^{\nu}} \Psi_0(x) dx}$$

and

$$1 = Q(\mu) = \lim_{r \rightarrow \infty} (\lim_{r \rightarrow \infty} g_r) = \lim_{r \rightarrow \infty} \frac{\int_{\mathcal{E}_i^r} \Psi_\mu(x) dx}{\int_{\mathcal{E}_i^r} \Psi_0(x) dx}.$$

A formula very similar to (12) exists for $\lim_{r \rightarrow \infty} G_r$, so that an alternate method of proving our theorem would be to conclude Corollary 2 without benefit of $Q(\mu) \equiv 1$. However, this is essentially asking for an explicit calculation of the nature of $\lim_{r \rightarrow \infty} \mathcal{E}_i^r$ which does not seem to be easy. If, for example, $m(\lim_{r \rightarrow \infty} \mathcal{E}_i^r) = 1$, the result would follow.

To apply our theorem to the Jacobi algorithm we refer to [6]. There we showed that

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

belongs to \mathcal{F} and the assumptions are satisfied with

$$C = (1 + 2n)^{n+1},$$

$$L = \frac{1}{n!},$$

and

$$q = \frac{1}{n!(1+n)^{n+1}(1+2n)^{n+1}}.$$

Also, following Schweiger [3], we can verify the assumptions on f , and J . Thus our Kuzmin theorem holds for the Jacobi algorithm.

The author would like to express his appreciation to F. Schweiger for making available a manuscript containing a corrected version of his Kuzmin theorem ([5]). The work referred to in [6] will appear elsewhere.

References

- [1] R. O. Kuzmin, *Sur un probleme de Gauss*, Atti del Congresso Internazionale del Matematici Bologna 6 (1928), pp. 83-89.
- [2] A. Renyi, *Representations for real numbers and their ergodic properties*, Acta Math. Acad. Sci. Hungaricae 8 (1957), pp. 477-493.
- [3] F. Schweiger, *Ein Kuzmin'scher Satz über den Jacobischen Algorithmus*, J. Reine Angew. Math. 232 (1968), pp. 35-40.
- [4] — *Metrische Theorie einer Klasse zahlentheoretischer Transformationen*, Acta Arith. 15 (1968), pp. 1-18.

- [5] F. Schweiger, *Metrische Theorie einer Klasse zahlentheoretischer Transformationen (Corrigendum)*, Acta Arith. 16 (1969), pp. 217-219.
- [6] M. Waterman, *Some ergodic properties of multi-dimensional F-expansions*, Z. Wahrscheinlichkeitstheorie verw. Geb. 16 (1970), pp. 77-103.
- [7] T. Vinh-Hien, *The central limit theorem for stationary processes generated by number theoretic endomorphisms* (in Russian), Vestnik Moskov Univ. Ser. I. Mat. Meh. 5, 1 (1963), pp. 28-34.

IDAHO STATE UNIVERSITY

Received on 2. 2. 1970

(30)