Z. Wahrscheinlichkeitstheorie verw. Geb. 16, 77–103 (1970)

Some Ergodic Properties of Multi-Dimensional F-Expansions

MICHAEL S. WATERMAN

I. Introduction

This paper is concerned with probabilistic aspects of the expansion of points in n-dimensional Euclidean space. The expansions we consider need not converge although previous work has required convergence.

The classical continued fraction was first examined from a measure theoretic point of view by Gauss [11]. Defining $T(x)=\frac{1}{x}-\left[\frac{1}{x}\right]$, where $x\in(0,1)$ and $[\cdot]$ denotes the greatest integer function, he found the Lebesgue measure of $\{\alpha: T^n(\alpha)<x\}$ to have the limiting value $\frac{\log(1+x)}{\log 2}$. Gauss posed the problem of estimating the difference between the approximate and the limiting values. In 1928 Kuzmin [14] solved this problem by considering

$$\psi_{n+1}(x)=\sum_{k=1}^{\infty}\frac{1}{(k+x)^2}\,\psi_n\left(\frac{1}{k+x}\right),\qquad n\geqq 0$$

where ψ_0 is an arbitrary function satisfying regularity conditions. He shows ψ_n approaches $((1+x)\log 2)^{-1}$, Gauss' measure.

Ryll-Nardzewski [21] put this work in modern terminology by noting that $T(\cdot)$ was the shift on the digits of the continued fraction expansion. With his proof that $T(\cdot)$ was ergodic with respect to Lebesgue measure, the ergodic theory was completed by noting that Gauss' measure is an invariant measure for $T(\cdot)$.

Renyi [18] generalized Ryll-Nardzewski's results to the f-expansions of Everett [6] and Bissinger [3]. Here the shift is $T(x)=f^{-1}(x)-[f^{-1}(x)]$, $x\in(0,1)$. To show T ergodic with respect to Lebesgue measure, Renyi imposed condition (C), a regularity condition. Rohlin [20] obtained some information theoretic results which were applied to Renyi's f-expansions. Recently Vinh-Hien [31] has extended Kuzmin's theorem to f-expansions and obtained a central limit theorem. Reznik [19] has used Vinh-Hien's work to obtain a law of the iterated logarithm.

In 1869 Jacobi [10] presented an extension of the continued fraction to two dimensions. Perron [17] extended Jacobi's work to n-dimensions. In 1964 Schweiger [23] began an examination of the measure theoretic properties of Jacobi's algorithm (see [24 to 29]). It was this work which motivated our paper. However Schweiger [30] has recently published some results which also concern general F-expansions for n-dimensions. The class of algorithms he considers does not include the Jacobi algorithm and is a natural generalization of Renyi [18]. Our results generalize most of Schweiger's work and have the Jacobi algorithm as an example. We also include a central limit theorem and a law of the iterated logarithm.

In obtaining a Kuzmin theorem and Rohlin's results the chain rule is of primary importance. The identity is $(g\circ h)'=g'(h)h'$ in one-dimension and $J_{g\circ h}(\cdot)=J_g(h(\cdot))\,J_h(\cdot)$ in n-dimensions ([1], p. 140). Our method of obtaining the chain rule is related to some work of Parry ([15, 16]).

II. Notation and Assumptions

In this section we define the symbols that we use throughout the paper. The underlying assumptions are made and some special assumptions are stated and labeled. To avoid repetition, we later use the symbols defined here without redefining them. The reason for making some of the assumptions will be pointed out when they come into use.

Suppose F is a 1-1, continuous map of A onto $(0,1)^n$ where A is a convex set contained in R^n. We assume $J_F(x)$, the Jabobian of F evaluated at x, exists, the components of F have continuous first order partial derivatives, and $J_F(x)\neq 0$ for almost all $x\in A$. Let $D=F^{-1}$. We define the sequence of coordinates (n tuples of integers) associated with $x\in(0,1)^n$ as follows:

$$a^1(x)=[D(x)],\qquad \delta^1(x)=D(x)-[D(x)],$$
$$\dots$$
$$a^k(x)=[D(\delta^{k-1}(x))],\qquad \delta^k(x)=D(\delta^{k-1}x)-[D(\delta^{k-1}(x))],$$
$$\dots$$

where $[z]=([z_1],\dots,[z_n])$ for $z\in R^n$. ($[z_i]$ denotes the integral part of the real number z_i.)

Note that $\delta^k(x)$ need not belong to $(0,1)^n$, that is, $\delta^k_j(x)=0$ may occur for some $j\in\{1,\dots,n\}$. To avoid these difficulties we define the algorithm on the restricted set

$$(0,1)^n_F=\{x\in(0,1)^n:\delta^k(x)\in(0,1)^n \text{ for } k=1,2,\dots\}.$$

We impose the assumption that

$$m(0,1)^n_F=1$$

where m denotes n-dimensional Lebesgue measure. Our underlying measure space, then, is to be m on the Lebesgue measurable subsets of $(0,1)^n_F$.

We define the cylinder of order v generated by a specified set of coordinates $k^1,\dots,k^v$ as

$$B^v=B^v(k^1,\dots,k^v)=\{x\in(0,1)^n_F:a^j(x)=k^j,\ j=1,\dots,v\}.$$

We remark that

$$\delta B^v(k^1,\dots,k^v)\subseteq B^{v-1}(k^2,\dots,k^v).$$

That is, δ is a shift operator. Here $B^0=(0,1)^n_F$ by definition. Of course we only consider $k^1,\dots,k^v$ which are admissible, that is which could arise by application of the algorithm to some $x\in B^0$. This is to assume $B^v\neq\emptyset$. Also

$$B^v(x)=B^v(a^1(x),\dots,a^v(x))$$

is assumed to satisfy $m(B^v(x))>0$ for all $x\in(0,1)^n_F$.

From the definition of a^i and δ, it is easily verified that

$$x = F\left(a^1(x) + F\left(a^2(x) + \cdots + F\left(a^\nu(x) + \delta^\nu(x)\right)\cdots\right)\right).$$

We define

$$f_{a^i}(t) = F(a^i + t).$$

The above identity can now be rewritten as $x = f_{a^1} \circ f_{a^2} \circ \cdots \circ f_{a^\nu}(t)$. Since there is a chain rule for Jacobians [1] just as there is for functions of one variable, and since the integral of a Jacobian is area [1], we make fundamental use of this composition of functions. The entire theory we construct is based on the following observation:

Lemma 2.1. *If B^ν is generated by $k^1, \ldots, k^\nu$, then*

$$B^\nu = f_{k^1} \circ f_{k^2} \circ \cdots \circ f_{k^\nu}(\delta^\nu B^\nu) = \prod_{i=1}^{\nu} \circ f_{k^i}(\delta^\nu B^\nu).$$

Proof.

$$f_{k^1} \circ f_{k^2} \circ \cdots \circ f_{k^\nu}(\delta^\nu B^\nu) = \bigcup_{x \in B^\nu} \left\{F\left(k^1 + F\left(k^2 + \cdots + F\left(k^\nu + \delta^\nu(x)\right)\ldots\right)\right)\right\} = \bigcup_{x \in B^\nu} \{x\} = B^\nu.$$

The assumptions on F made above will be used throughout and when they hold we will write $F \in \mathfrak{F}$. Below we make three additional assumptions on F. Every use of these conditions will be clearly indicated.

The following condition generalizes condition (C) of Renyi [18]:

$$\text{(C)} \qquad \frac{\sup_{t \in \delta^\nu B^\nu} |J_{f_\nu}(t)|}{\inf_{t \in \delta^\nu B^\nu} |J_{f_\nu}(t)|} \leqq C < +\infty$$

where $f_\nu = \prod_{i=1}^{\nu} \circ f_{k^i}$ for each $\nu \geqq 1$ and all admissible cylinders $B^\nu(k^1, \ldots, k^\nu)$. This condition is to assure us that no path of the process is too flat or too steep.

It should be remarked that $\delta^\nu B^\nu \subseteq (0,1)^n_F$ but strict containment can occur. In the n-dimensional continued fraction (to be discussed below) for example, $m(\delta^\nu B^\nu) = (n!)^{-1}$ occurs a countable number of times (whenever $k^\nu = (k, \ldots, k), k \geqq 1$). If $\delta^\nu B^\nu = (0,1)^n_F$, we say B^ν is proper; otherwise B^ν is said to be improper. Difficulties associated with improper cylinders lead us to the next two conditions.

$$\text{(L)} \qquad m\left(\delta^\nu B^\nu(x)\right) \geqq L > 0 \quad \text{for all } x \in B^0, \quad \nu = 1, 2, \ldots.$$

Condition (L) assures us that $\delta^\nu B^\nu$ is never too small.

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

$$\text{(q)} \qquad \frac{m(\hat{B}^{\nu+1})}{m(B^\nu(x))} \geqq q > 0 \quad \text{for all } x \in B^0, \quad \nu = 1, 2, \ldots.$$

Condition (q) implies that the sum of the measures of the proper cylinders of any given order is at least $q > 0$.

Renyi's usage of notation in [18] is followed throughout. For example, if we assume condition (L) holds for $f \in \mathfrak{F}$, we are assuming that $m(\delta^v B^v(x)) \geqq L > 0$ for all $x \in B^0$, $v \geqq 1$. The ambiguity between condition (L) and the numerical value of L is retained for notational convenience.

III. Convergence of *F*-Expansions

A very useful property for $F \in \mathfrak{F}$ to possess is that F expand almost all $x \in (0, 1)^n$, that is diam $B^v(x) \to 0$ as $v \to \infty$. We write this convergence as $B^v(x) \to x$. Previous work of Renyi [18] on f-expansions, Schweiger ([25 to 27]) on the Jacobi algorithm, and Schweiger [30] on certain n-dimensional expansions has considered classes of such algorithms. Convergence of the algorithm is useful in proving ergodicity of the shift. The following theorem characterizes this convergence.

Theorem 3.1. *Suppose* $F \in \mathfrak{F}$ *and condition* (q) *holds. Then*

$$m\{x\colon B^v(x) \to x\} = 1$$

if and only if for all B^v *and all* $\varepsilon > 0$, *there exist countably many disjoint* $B^\mu \subset B^v$ *such that* diam $B^\mu < \varepsilon$ *and* $m(B^v \sim \bigcup_\mu B^\mu) = 0$.

Proof. Assume $m\{x\colon B^v(x) \to x\} = 1$. For a.a. x there exists $v(x)$ such that $\operatorname{diam}(B^{v(x)}(x)) < \varepsilon$. For each μ we remove the set of B^μ such that $x \in B^v$ and $B^\mu(x) = B^{v(x)}(x)$. Collecting these sets, $m(B^v \sim \bigcup_\mu B^\mu) = 0$. The converse is obvious.

Recent work of Schweiger [30] deals with a class of n-dimensional F-expansions. We will refer to his conditions as condition (S). For $f_v = \prod_{i=1}^{v} \circ f_{k^i}$, $f_{k^i}(t) = F(k^i + t)$, define

$$\|J_v\| = \sup\left\{\left|\frac{\partial (f_v)_i}{\partial x_j}\right| \colon 1 \leqq i, j \leqq n, x \in \delta^v B^v\right\}.$$

Then F satisfies condition (S) if

(S 1) $$F \in \mathfrak{F}$$

(S 2) $$\|J_v\| \leqq \tau(v)$$

where $\tau(\cdot)$ is a non-negative, decreasing function on $[1, \infty)$ such that $\lim_{x \to \infty} \tau(x) = 0$. This inequality must hold uniformly for admissible $k^1, \ldots, k^v$.

(S 3) All cylinders of all orders are proper.

Theorem 3.2 (Schweiger). *If* F *satisfies condition* (S), *then* diam $B^v(x) \to x$ *as* $v \to \infty$ *for all* x.

The theorem is proved by considering the components of f_v and applying the mean value theorem for functions of n variables. Then (S 2) yields the result. This result gives a relationship between the Jacobian and the diameter of the cylinders. However it is an open question whether or not the Jacobi algorithm satisfies (S 2). In Section VII we show that Theorem 3.3 below applies to the Jacobi algorithm. The hypothesis clearly contains a version of condition (q).

Theorem 3.3. *Let $F \in \mathfrak{F}$. Also suppose there exists $1 \leqq \ell$, $\lambda \in (0, 1)$, $0 < q'$ such that for each B^v there exists $\hat{B}^{v+\ell} \subseteq B^v$, a cylinder of order $v+\ell$, satisfying*

a) $\operatorname{diam}(\hat{B}^{v+\ell}) \leqq \lambda \operatorname{diam}(B^v)$

and

b) $0 < q' \leqq \dfrac{m(\hat{B}^{v+\ell})}{m(B^v)}$.

Then

$$m\{x\colon B^v(x) \to x\} = 1.$$

Proof. Choose B^v and $\varepsilon > 0$. Repeatedly applying our hypothesis we obtain $\hat{B}^{v+k\ell}$ such that

$$\operatorname{diam} \hat{B}^{v+k\ell} \leqq \lambda^k \operatorname{diam} B^v \leqq \lambda^k < \varepsilon$$

and

$$0 < q'' = (q')^k \leqq \frac{m(\hat{B}^{v+k\ell})}{m(\hat{B}^{v+(k-1)\ell})} \cdot \ldots \cdot \frac{m(\hat{B}^{v+\ell})}{m(B^v)} = \frac{m(\hat{B}^{v+k\ell})}{m(B^v)}$$

where $k = \min\{r\colon \lambda^r < \varepsilon\}$. This implies (taking the union over all $v+k\ell$ cylinders with less diameter less than ε)

$$m(B^v \sim \bigcup B^{v+k\ell}) \leqq m(B^v) - m(\hat{B}^{v+k\ell}) \leqq (1-q'')\, m(B^v).$$

For any $B^{v+k\ell} \subset B^v$ with diameter greater than or equal to ε, we have $\hat{B}^{v+2k\ell}$ such that

$$q'' \leqq \frac{m(\hat{B}^{v+2k\ell})}{m(B^{v+k\ell})}$$

with diam $(\hat{B}^{v+2k\ell}) < \varepsilon$. This implies

$$m\left(B^v \sim \bigcup_{\mu=1}^{2} B^{v+\mu k\ell}\right) \leqq (1-q'')\, m(B^v \sim \bigcup B^{v+k\ell}) \leqq (1-q'')^2$$

where the union is over cylinders satisfying $\operatorname{diam}(B^{v+\mu k\ell}) < \varepsilon$. By induction we obtain

$$m\left(B^v \sim \bigcup_{\mu=1}^{p} B^{v+\mu k\ell}\right) \leqq (1-q'')^p$$

and the theorem is proved with an application of Theorem 3.1.

As a simple illustration of Theorem 3.3 we consider $F(x, y) = (x/2, y/2)$ with domain $(0, 2)^2$. Then $D(x, y) = (2x, 2y)$ and the cylinders are squares. If we consider $B^v(k^1, \ldots, k^v)$, then $B^{v+1}(k^1, \ldots, k^v, (1,1)) = \hat{B}^{v+1}$ is the square in the upper right hand corner of B^v. We have $m(\hat{B}^{v+1}) = \frac{1}{4} m(B^v)$ and $\operatorname{diam}(\hat{B}^{v+1}) = \frac{1}{2} \operatorname{diam}(B^v)$. Thus Theorem 3.3 applies with $\ell = 1$, $\lambda = \frac{1}{2}$, and $q = q' = \frac{1}{4}$.

Another application of Theorem 3.3 will be made in Section VII to a generalization of the Jacobi algorithm.

IV. Ergodic Theory of the Shift

For any $F \in \mathfrak{F}$ with conditions (C) and (L) satisfied, we show there exists a probability measure μ invariant with respect to δ. This is the initial step in establishing an ergodic theory for $\mathfrak{F}$. The proofs follow Renyi's [18] technique for the one-dimensional case but must be modified to handle $\delta^\nu B^\nu \neq (0,1)^n_F$. It is this technicality that complicates Schweiger's [26] calculation of the results of Ryll-Nardzewski's theorem for the n-dimensional Jacobi algorithm.

Theorem 4.1. *Suppose that $F \in \mathfrak{F}$ satisfies condition* (C) *and condition* (L). *Then there exists a probability measure μ on $(0,1)^n$ such that $\mu << m$ and δ is a measure preserving transformation for μ. Moreover $\mu(E) \leqq \frac{C}{L} m(E)$ holds for measurable $E \subset (0,1)^n$. Therefore, if g is any Lebesgue integrable function on $(0,1)^n$, we have (by an Ergodic theorem)*

$$\lim_{\nu\to\infty} \frac{1}{\nu} \sum_{j=0}^{\nu-1} g(\delta^j(x)) = \hat{g}(x)$$

exists for a.a. $[m]$ $x \in (0,1)^n$. $\hat{g} \in L_1$ and $\hat{g} = E(g \mid \mathscr{I})$ where $\mathscr{I} = \{E: \delta^{-1}E = E\}$ and $E(\cdot)$ is with respect to μ above.

Proof. Let $J_\nu(t) = J_{f_\nu}(t)$, $f_\nu = \prod_{i=1}^{\nu} \circ f_{k^i}$, and $B^\nu = B^\nu(k^1, \ldots, k^\nu)$.

By an area theorem for Jacobians [1] and Lemma 2.1

$$\int_{\delta^\nu B^\nu} |J_\nu(t)|\, dt = m(B^\nu).$$

Let $\mathscr{E}_\nu$ denote an admissible sequence $k^1, \ldots, k^\nu$. If $\mathscr{E}_\nu$ runs over all admissible sequences

$$\sum_{\mathscr{E}_\nu} \int_{\delta^\nu B^\nu} |J_\nu(t)|\, dt = \sum_{\mathscr{E}_\nu} m(B^\nu) = 1.$$

From this we obtain

$$\sum_{\mathscr{E}_\nu} \inf_{\delta^\nu B^\nu} |J_\nu(t)|\, m(\delta^\nu B^\nu) \leqq 1 \leqq \sum_{\mathscr{E}_\nu} \sup_{\delta^\nu B^\nu} |J_\nu|\, m(\delta^\nu B^\nu)$$

and

$$L \sum_{\mathscr{E}_\nu} \inf_{\delta^\nu B^\nu} |J_\nu(t)| \leqq 1 \leqq \sum_{\mathscr{E}_\nu} \sup_{\delta^\nu B^\nu} |J_\nu|. \tag{4.1}$$

Now we set $E = \mathop{\mathsf{X}}_{i=1}^{n} [a_i, b_i)$ where $0 \leqq a_i \leqq b_i \leqq 1$.

$$\delta^{-\nu}(E) = \bigcup_{\mathscr{E}_\nu} \{F(k^1 + F(k^2 + \cdots + F(k^\nu + t)\ldots)): t \in E \cap \delta^\nu B^\nu\}.$$

Due to the 1-1 property of F, the above union is disjoint. By a Jacobian argument similar to the one above, we obtain

$$m(\delta^{-\nu}E) = \sum_{\mathscr{E}_\nu} \int_{\delta^\nu B^\nu \cap E} |J_\nu(t)|\, dt \leqq \sum_{\mathscr{E}_\nu} \sup_{\delta^\nu B^\nu \cap E} |J_\nu|\, m(\delta^\nu B^\nu \cap E)$$

$$\leqq \sum_{\mathscr{E}_\nu} \sup_{\delta^\nu B^\nu} |J_\nu|\, m(E) = m(E) \sum_{\mathscr{E}_\nu} \sup_{\delta^\nu B^\nu} |J_\nu|.$$

Using the earlier inequality (4.1),

$$m(\delta^{-\nu}E) \leqq \frac{m(E)}{L} \frac{\sum_{\mathscr{E}_\nu} \sup |J_\nu|}{\sum_{\mathscr{E}_\nu} \inf |J_\nu|} \leqq \frac{C}{L} m(E).$$

This yields

$$\frac{1}{\nu} \sum_{j=0}^{\nu-1} m(\delta^{-j}E) \leqq \frac{C}{L} m(E) \tag{4.2}$$

where E is the product of rectangles, and hence (4.2) holds all measurable subsets of $(0, 1)^n$.

By the direct application a theorem of Dunford and Miller [5], (4.2) implies

$$\lim_{\nu\to\infty} \frac{1}{\nu} \sum_{j=0}^{\nu-1} g(\delta^k x) = \hat{g}(x) \tag{4.3}$$

exists for all $g \in L_1(0, 1)^n$ and a.a. $x \in (0, 1)^n$. We may also define (applying (4.3) above)

$$\mu(E) = \lim_{\nu\to\infty} \frac{1}{\nu} \sum_{j=0}^{\nu-1} m(\delta^{-j}E)$$

which easily has the properties

$$\mu(E) \leqq \frac{C}{L} m(E)$$

and

$$\mu(E) = \lim_{\nu\to\infty} \frac{1}{m} \sum_{j=0}^{m-1} m(\delta^{-j}E) = \lim_{\nu\to\infty} \frac{1}{m} \sum_{j=0}^{m-1} m(\delta^{-j}(\delta^{-1}E)) = \mu(\delta^{-1}E).$$

If condition (q) also holds, the proper subcylinders of any cylinder B^ν have measure greater than or equal to $q\,m(B^\nu)$. This allows us to obtain a lower bound on $\mu(E)$ and have μ equivalent to m. Renyi has no difficulty here because all of his cylinders are proper.

Theorem 4.2. *Suppose we have an $F \in \mathfrak{F}$ with conditions* (C), (L), *and* (q) *holding. Then there exists a probability measure μ on $(0, 1)^n$ such that $\mu \sim m$ and δ is a measure preserving transformation for μ. Moreover*

$$\frac{q}{C} m(E) \leqq \mu(E) \leqq \frac{C}{L} m(E)$$

holds for measurable $E \subset (0, 1)^n$. Therefore, if g is any Lebesgue integrable function on $(0, 1)^n$, we have (by the Ergodic theorem)

$$\lim_{\nu\to\infty} \frac{1}{\nu} \sum_{j=0}^{\nu-1} g(\delta^j(x)) = \hat{g}(x)$$

exists a.e. $[m] \cdot \hat{g} \in L_1(m)$ and $\hat{g} = E(g \mid \mathscr{I})$ where $\mathscr{I} = \{E\colon \delta^{-1}E = E\}$.

Proof. The only additional work beyond Theorem 4.1 is to establish the lower bound on $\mu(E)$. Below $\sum_{\mathscr{E}_\nu}$ will denote a summation over all admissible cylinders of order ν;

$\sum'_{\mathscr{E}_\nu}$ will denote a summation over all proper admissible cylinders of order ν.

For each $\nu-1$ cylinder $B^{\nu-1}$, condition (q) states that there exists $\hat{B}^\nu$, a collection of proper cylinders of order ν such that

$$m\,\hat{B}^\nu \geqq q\,m(B^{\nu-1}).$$

Thus

$$\sum_{\mathscr{E}_\nu}' m(B^\nu) \geqq \sum_{\mathscr{E}_{\nu-1}} m(\hat{B}^\nu) \geqq \sum_{\mathscr{E}_{\nu-1}} q\,m(B^{\nu-1}) = q,$$

and

$$\sum_{\mathscr{E}_\nu}' \sup |J_\nu| \geqq \sum_{\mathscr{E}_\nu}' \int_{\delta^\nu B^\nu} |J_\nu(t)|\,dt = \sum_{\mathscr{E}_\nu}' m\,B^\nu \geqq q.$$

Now

$$\begin{aligned} m(\delta^{-\nu} E) &= \sum_{\mathscr{E}_\nu} \int_{\delta^\nu B^\nu \cap E} |J_\nu(t)|\,dt \\ &\geqq \sum_{\mathscr{E}_\nu}' \int_{\delta^\nu B^\nu \cap E} |J_\nu(t)|\,dt \\ &\geqq \sum_{\mathscr{E}_\nu}' \inf |J_\nu|\, m(E) \\ &\geqq q \frac{\sum'_{\mathscr{E}_\nu} \inf |J_\nu|}{\sum'_{\mathscr{E}_\nu} \sup |J_\nu|}\, m(E) \geqq \frac{q}{C}\, m(E). \end{aligned}$$

To complete the ergodic theory for $\mathfrak{F}$ we show our measure preserving transformation δ ergodic with respect to m (and hence with respect to μ) for all F such that $m\{x\colon B^\nu(x) \to x\} = 1$. This allows us to replace the $\hat{g}(x)$ in Theorems 4.1 and 4.2 by the constant $\int g(x)\,d\mu(x)$. Our technique is used by Billingsley ([2], p. 44) in connection with continued fractions and by Schweiger [30] in a similar context. Renyi [18] used a theorem of Knopp [13] to show ergodicity while Schweiger [25] has used the Lebesgue density theorem [22].

The following lemma shows the invariant sets have positive density in proper cylinders.

Lemma 4.1. *Suppose* $F \in \mathfrak{F}$ *and condition* (C) *holds. Then, if* $E \in \mathscr{I}_d = \{D\colon D$ *a rectangle,* $m(D) = d$, $\delta^{-1} D = D\}$ $(0 < d < 1)$,

$$\frac{m(E \cap B^\nu)}{m(B^\nu)} \geqq \frac{d}{C}$$

uniformly in $E \in \mathscr{I}_d$ *and proper* B^ν, $\nu \geqq 0$.

Proof. As usual we write $J_\nu(x) = J_{f_\nu}(x)$, $f_\nu = \prod_{i=1}^{\nu} \circ f_{k^i}$, where $B^\nu = B^\nu(k_1, \ldots, k^\nu)$.

Since $\delta^{-1} E = E$,

$$I_E(x) = I_{\delta^{-\nu} E}(x) = I_E(\delta^\nu x) = I_E\big(f_\nu^{-1}(x)\big), \qquad \text{for } x \in B^\nu.$$

Then, using the transformation theorem for Jacobians,

$$\int_{B^\nu} I_E(x)\,dx = \int_{B^\nu} I_E\big(f_\nu^{-1}(x)\big)\,dx = \int_{\delta^\nu B^\nu} I_E(t)\,|J_\nu(t)|\,dt.$$

Using this relationship (and $\delta^\nu B^\nu = (0,1)_F^n$),

$$m(B^\nu \cap E) \geqq \inf |J_\nu(\cdot)| \int_{(0,1)_F^n} I_E(t)\,dt = m(E) \inf |J_\nu(\cdot)| \geqq m(E)\frac{\sup |J_\nu(\cdot)|}{C}.$$

But

$$\sup_{(0,1)_F^n} |J_\nu(\cdot)| \geqq \int_{(0,1)_F^n} |J_\nu(t)|\,dt = m(B^\nu),$$

so that

$$m(E \cap B^\nu) \geqq \frac{m(E)}{C}\, m(B^\nu)$$

or

$$\frac{m(E \cap B^\nu)}{m(B^\nu)} \geqq \frac{d}{C} > 0.$$

Next we note that condition (q) allows us to replace an improper cylinder by the proper cylinders contained in it and lose an arbitrarily small portion of the cylinder. Lemma 4.2 corresponds directly to a result of Schweiger [25] concerning the Jacobi algorithm.

Lemma 4.2. *Suppose $F \in \mathfrak{F}$ and satisfies condition* (q). *Then for an arbitrary cylinder B^ν and $\varepsilon > 0$, there exist countably many disjoint proper cylinders B^μ and an integer $\nu_0 = \nu_0(\varepsilon)$ such that*

(i) $B^\nu \supset \bigcup_{\mu=\nu+1}^{\nu_0} B^\mu$

and

(ii) $m\left(B^\nu \sim \bigcup_{\mu=\nu+1}^{\nu_0} B^\mu\right) < \varepsilon$.

It is clear that (ii) *can only be improved by taking the union over all proper cylinders of order μ contained in B^ν.*

Proof. Let $\hat{B}^{\nu+1}$ be the union of a collection of proper cylinders of order $\nu+1$ contained in B^ν such that

$$q \leqq \frac{m(\hat{B}^{\nu+1})}{m(B^\nu)}.$$

Below $\bigcup_\mu{}'$ denotes the union over proper cylinders of order μ.

$$m(B^\nu \sim \textstyle\bigcup' B^{\nu+1}) = m\,B^\nu - m(\bigcup' B^{\nu+1}) \leqq m(B^\nu) - m(\hat{B}^{\nu+1}) \leqq m(B^\nu) - q\,m(B^\nu) = (1-q)\,m(B^\nu).$$

Take any $B^{\nu+1} \subset B^{\nu}$ which is improper. Let $\hat{B}^{\nu+2}$ be a proper cylinder such that

$$q \leqq \frac{m(\hat{B}^{\nu+2})}{m(B^{\nu+1})}.$$

Then

$$m(B^{\nu+1} \sim \bigcup'_{\hat{B}^{\nu+2} \subset B^{\nu+1}} \hat{B}^{\nu+2}) \leqq (1-q)\, m(B^{\nu+1})$$

which implies

$$m(B^{\nu} \sim ((\bigcup' B^{\nu+1}) \cup (\bigcup' B^{\nu+2}))) \leqq (1-q)\, m(B^{\nu} \sim \bigcup' B^{\nu+1}) \leqq (1-q)^2\, m(B^{\nu}).$$

We proceed by induction and let

$$\nu_0(\varepsilon) = \min\{p: (1-q)^p < \varepsilon\}.$$

Lemma 4.2 allows us to conclude that $E \in \mathscr{I}_d$ has positive density with respect to *any* cylinder.

Corollary 4.1. *Suppose* $F \in \mathfrak{F}$ *and satisfies conditions* (C) *and* (q). *Then*

$$\frac{m(E \cap B^{\nu})}{m(B^{\nu})} \geqq \frac{d}{C}$$

uniformly in $E \in \mathscr{I}_d$ *and* B^{ν} $(\nu \geqq 0)$.

Proof. We apply Lemmas 4.1 and 4.2. $\sum_{\mu}'$ denotes a summation over proper cylinders as above. For all $\varepsilon > 0$,

$$\frac{m(E \cap B^{\nu})}{m(B^{\nu})} = \frac{\sum' m(E \cap B^{\nu+1}) + \cdots + \sum' m(E \cap B^{\nu_0}) + m\left(E \cap B^{\nu} \sim \bigcup_{\nu+1}^{\nu_0} B^{\mu}\right)}{\sum' m(B^{\nu+1}) + \cdots + \sum' m(B^{\nu_0}) + m\left(B^{\nu} \sim \bigcup_{\nu+1}^{\nu_0} B^{\mu}\right)}$$

$$\geqq \frac{\sum' m(E \cap B^{\nu+1}) + \cdots + \sum' m(E \cap B^{\nu_0})}{\sum' m(B^{\nu+1}) + \cdots + \sum' m(B^{\nu_0}) + \varepsilon}.$$

If we let $\varepsilon \to 0$, we can obtain

$$\frac{m(E \cap B^{\nu})}{m(B^{\nu})} \geqq d'.$$

Finally we can prove δ ergodic if $m\{x: B^{\nu}(x) \to x\} = 1$.

Theorem 4.3. *Suppose* $F \in \mathfrak{F}$ *satisfies conditions* (C), (q), *and* $m\{x: B^{\nu}(x) \to x\} = 1$. *Then* δ *is ergodic under* m.

Proof. By definition, δ is ergodic under m if $\delta^{-1}E = E$ implies $m(E) = 0$ or 1. Corollary 4.1 states $m(E \cap B^{\nu}) \geqq C^{-1} m(E)\, m(B^{\nu})$ for all cylinders B^{ν}. If A is any rectangle it is easy to see that A can be approximated in measure by a countable union of cylinders (since diam $B^{\nu}(x) \to 0$ a.e.). Thus $m(E \cap A) \geqq C^{-1} m(E)\, m(A)$. Letting $A = \tilde{E}$, $m(E \cap \tilde{E}) \geqq C^{-1} m(E)\, m(\tilde{E})$ and $m(E) = 0$ or 1.

V. Rohlin's Formula

In this section we obtain bounds on the μ-measure and the m-measure of $B^\nu(x)$ for large ν. Our formula will be similar to that which Rohlin [20] used to compute the entropy of one-dimensional f-expansions. Kinney and Pitcher [12] have used Rohlin's formula for bounds on one-dimensional f-expansions when the formula did not represent entropy. The calculations of this section follow Rohlin [20] as Schweiger [27] did in his computation of entropy for the Jacobi algorithm. $J_\nu(t)$ will denote $J_{f_\nu}(t)$, $f_\nu = \prod_{i=1}^{\nu} \circ f_{k^i}$. The next theorem corresponds to Theorem 4.1 and only gives an upper bound for $\mu(B^\nu(x))$.

Theorem 5.1. *Suppose that* $F \in \mathfrak{F}$ *satisfies conditions* (C) *and* (L). *Assume* $\log |J_D(\cdot)|$ *is Lebesgue integrable on* $(0,1)^n$ *and that* δ *is ergodic (indecomposable) with respect to m. Then, if we take* μ *to be the measure of Theorem* 4.1,

$$\lim_{\nu\to\infty} \frac{1}{\nu} \log \frac{1}{m(B^\nu(x))} = +\int \log |J_D(t)|\, d\mu(t)$$

for a.a. x. Also

$$\varliminf_{\nu\to\infty} \frac{1}{\nu} \log \frac{1}{\mu(B^\nu(x))} \geqq +\int \log |J_D(t)|\, d\mu(t).$$

Proof. By Theorem 4.1, $\mu(E) \leqq \frac{C}{L} m(E)$. Then

$$\frac{1}{\nu} \log \frac{1}{m(B^\nu(x))} - \frac{1}{\nu} \log \frac{1}{\mu(B^\nu(x))} = \frac{1}{\nu} \log \frac{\mu(B^\nu)}{m(B^\nu)} \leqq \frac{1}{\nu} \log \frac{C}{L}.$$

Thus, if $\lim_{\nu\to\infty} \frac{1}{\nu} \log \frac{1}{m(B^\nu x)}$ exists,

$$\lim_{\nu\to\infty} \frac{1}{\nu} \log \frac{1}{m(B^\nu(x))} \leqq \varliminf_{\nu\to\infty} \frac{1}{\nu} \log \frac{1}{\mu(B^\nu(x))}.$$

Conditions (L) and (C) imply

$$\left| \frac{1}{\nu} \log \frac{1}{m\, B^\nu(x)} - \frac{1}{\nu} \log \frac{1}{|J_\nu(\delta^\nu x)|} \right| \leqq \frac{1}{\nu} \log \frac{1}{L \inf_{\delta^\nu B^\nu(x)} |J_\nu(\cdot)|}$$
$$- \frac{1}{\nu} \log \frac{1}{\sup_{\delta^\nu B^\nu(x)} |J_\nu(\cdot)|} \leqq \frac{1}{\nu} \log \frac{C}{L}.$$

We next obtain a limit for $\frac{1}{\nu} \log \frac{1}{|J_\nu(\delta^\nu x)|}$ which will then be equal to

$$\lim \frac{1}{\nu} \log \frac{1}{m(B^\nu(x))}.$$

For $x \in B^\nu$, $f_\nu \circ \delta^\nu(x) = x$ so that

$$J_{f_\nu \circ \delta^\nu}(x) = 1,$$

and by the chain rule [1]

$$J_{f_\nu \circ \delta^\nu}(x) = J_{f_\nu}(\delta^\nu(x)) \cdot J_{\delta^\nu}(x) = J_{f_\nu}(\delta^\nu(x)) \prod_{i=0}^{\nu-1} J_\delta(\delta^i(x)).$$

Therefore

$$\frac{1}{J_{f_\nu}(\delta^\nu(x))} = \prod_{i=0}^{\nu-1} J_\delta(\delta^i(x)).$$

Noting $J_\delta = J_D$ we have

$$\frac{1}{\nu} \log \frac{1}{|J_{f_\nu}(\delta^\nu(x))|} = \frac{1}{\nu} \sum_{i=0}^{\nu-1} \log |J_D(\delta^i(x))|.$$

By the Ergodic Theorem

$$\lim_{\nu\to\infty} \log \frac{1}{|J_{f_\nu}(\delta^\nu(x))|} = + \int_{(0,1)^n} \log |J_D(t)| \, d\mu(t)$$

exists for a. a. $t \in (0, 1)^n$.

If we add condition (q), then Theorem 4.2 gives us a $\mu \sim m$ and the new condition allows us to conclude that $m(B^\nu(x))$ and $\mu(B^\nu(x))$ have the same bounds. The conclusions of this theorem fully correspond to Rohlin [20].

Theorem 5.2. *Suppose we have an* $F \in \mathfrak{F}$ *with additional conditions* (C), (L), *and* (q). *Assume* $\log |J_D(\cdot)|$ *is Lebesgue integrable on* $(0, 1)^n$ *and that* δ *is ergodic (indecomposable) with respect to m. Then if we take* μ *to be the measure of Theorem* 4.2,

$$\lim_{\nu\to\infty} \frac{1}{\nu} \log \frac{1}{m(B^\nu(x))} = \lim_{\nu\to\infty} \frac{1}{\nu} \log \frac{1}{\mu(B^\nu(x))} = + \int \log |J_D(t)| \, d\mu(t),$$

for a.a. $x \in (0, 1)^n$.

Proof. Since Theorem 5.1 can be applies, we need only establish

$$\lim_{\nu\to\infty} \frac{1}{\nu} \log \frac{1}{m(B^\nu(x))} = \lim_{\nu\to\infty} \frac{1}{\nu} \log \frac{1}{\mu(B^\nu(x))}.$$

By Theorem 4.2

$$\frac{q}{C} m(E) \leqq \mu(E) \leqq \frac{C}{L} m(E).$$

This implies (letting $C' = \max\{C/L, C/q\}$)

$$\left| \frac{1}{\nu} \log \frac{1}{m B^\nu(x)} - \frac{1}{\nu} \log \frac{1}{\mu B^\nu(x)} \right| \leqq \frac{1}{\nu} \log C'.$$

Rohlin [20] has shown that in the case of one-dimensional f-expansions the entropy of δ, $h(\delta)$, is given by

$$h(\delta)=\int \log|\varphi'(t)|\,d\mu(t)$$

where $\varphi=f^{-1}$. Since Theorem 5.2 above has an integral of the same form, it is natural to conjecture that, if $m\{x: B^\nu(x)\to x\}=1$, then

$$h(\delta)=\int \log|J_D(t)|\,d\mu(t).$$

Utilizing some results of Rohlin [20] we will establish this result below.

In Rohlin's terminology a measure preserving transformation on (Ω, B, P) is called an endomorphism. A transformation T is said to be exact if $\bigcap_{\nu=0}^{\infty} T^{-\nu}B$ is composed of sets of P-measure 0 or 1.

Theorem 5.3. *Let* $F\in\mathfrak{F}$ *satisfy condition* (C), (L), (q) *and* $m\{x: B^\nu(x)\to x\}=1$. *Then* δ *is an exact endomorphism.*

Proof. We apply a theorem of Rohlin ([20], p. 30). Let $\mathfrak{A}$ be the collection of all proper cylinders of all orders. To index $\mathfrak{A}$ we define $B_\nu(x)$, $x\in(0,1)_F^n$, to be the ν-th proper cylinder containing x. For a.a. x $B_\nu(x)$ exists and $B_\nu(x)\to x$. It is easy to see that $\mathfrak{A}$ generates the Lebesgue measurable sets.

To complete the proof we need to establish

$$\mu(\delta^\nu X)\leqq M\,\frac{\mu(X)}{\mu(B^\nu)}$$

for proper cylinders B^ν and measurable $X\subset B^\nu$ with $\delta^\nu B^\nu$ measurable. It is easy to show that

$$m(\delta^\nu X)\leqq\frac{m(X)}{\inf|J_\nu(\cdot)|}\leqq\frac{C\,m(X)}{\sup|J_\nu(\cdot)|}\leqq C\,\frac{m(X)}{m(B^\nu)}.$$

Using the inequalities of Theorem 4.2

$$\frac{L}{C}\,\mu(\delta^\nu X)\leqq m(\delta^\nu X)\leqq C\,\frac{m(X)}{m(B^\nu)}\leqq\frac{C^3}{Lq}\,\frac{\mu(X)}{\mu(B^\nu)}.$$

Rohlin ([20], p. 17) also proves an exact endomorphism is mixing of all degrees.

Corollary 5.1. *Let* $F\in\mathfrak{F}$ *satisfy conditions* (C), (L), (q), *and* $m\{x: B^\nu(x)\to x\}=1$. *Then* δ *is mixing of all degrees.*

Our main result is established as a result of δ being an exact endomorphism. The following lemma will be necessary. We define $\xi=\{B^1(x): x\in B^0\}$.

Lemma 5.1. *Suppose* $F\in\mathfrak{F}$ *satisfies conditions* (C), (L), *and* (q). *Then*

$$H(\xi)=-\sum\mu(B^1)\log\mu(B^1)<+\infty$$

if and only if

$$\log|J_D(\cdot)|\in L_1\big(m,(0,1)^n\big).$$

Proof. Using condition (L) and summing over ξ, we obtain

$$\sum m(B)\log\frac{1}{\sup\limits_{\delta B}|J_1(\cdot)|}\leqq\sum m(B)\log\frac{1}{m(B)}\leqq\sum m(B)\log\frac{1}{L\inf\limits_{\delta B}|J_1(\cdot)|}.$$

As in the proof of Theorem 6.1,

$$J_D(x)=J_\delta(x)=\frac{1}{J_{f_1}(\delta(x))}.$$

Taking supremum and infimum over B followed by integration over B and finally symmation over ξ, we obtain

$$\sum m(B)\log\frac{1}{\sup\limits_{\delta B}|J_1(\cdot)|}\leqq\int\log|J_D(x)|\,dx\leqq\sum m(B)\log\frac{1}{\inf\limits_{\delta B}|J_1(\cdot)|}.$$

But

$$0\leqq m(B)\log\frac{1}{\inf\limits_{\delta B}|J_1(\cdot)|}-m(B)\log\frac{1}{\sup\limits_{\delta B}|J_1(\cdot)|}\leqq m(B)\log C$$

so that the sums bounding $-\sum m(B)\log m(B)$ and $\int\log|J_D(x)|\,dx$ converge and diverge together. By virtue of the inequalities of Theorem 4.2 $-\sum m(B)\log m(B)$ and $H(\xi)$ converge or diverge together.

Theorem 5.4. *Suppose* $F\in\mathfrak{F}$ *satisfies condition* (C), (L), *and* (q) *with*

$$\log|J_D(\cdot)|\in L_1(m,(0,1)^n).$$

Then, if $m\{x: B^\nu(x)\to x\}=1$, *we have*

$$h(\delta)=\int\log|J_D(t)|\,d\mu(t)$$

and

$$h(\delta)=\lim_{\nu\to\infty}\frac{1}{\nu}\log\frac{1}{m(B^\nu(x))}=\lim_{\nu\to\infty}\frac{1}{\nu}\log\frac{1}{\mu(B^\nu(x))}\quad a.e.\quad[m].$$

Proof. By Rohlin's generalization ([20], p. 21) of the Kolmogorov-Sinai Theorem [2],

$$h(\delta)=h(\delta,\xi).$$

The Shannon-Mcmillan-Breiman-Chung Theorem ([2, 4]) then allows us to conclude

$$\lim_{\nu\to\infty}\frac{1}{\nu}\log\frac{1}{\mu(B^\nu(x))}=h(\delta)\qquad\text{for a.a. } x$$

if we have

$$H(\xi)=\sum\mu(B^1)\log\frac{1}{\mu(B^1)}<+\infty.$$

Applications of Lemma 5.1 and Theorem 5.2 then complete the proof.

VI. A Kuzmin Theorem for F-Expansions

Much of the value of our work is dependent on calculating $h(\delta)$, the entropy of the shift associated with a given F-expansion. Unter the conditions of Section V, $h(\delta)$ is given by

$$h(\delta)=\int_{(0,1)^n} \log |J_D(t)|\, d\mu(t)$$

where $\mu \sim m$ has been discussed in Section III. But even in the one-dimensional case μ or, equivalently, $d\mu/dm$ is difficult to find for a given F. However, in the one-dimensional case, some recent work of S. Guthery finds an F for which a given μ is the invariant measure (personal communication). Unfortunately his results do not extend to the n-dimensional case.

In the introduction we noted results of Kuzmin [14] and Vinh-Hien [30] which give a rate at which

$$\psi_{n+1}(x)=\sum_k \psi_n(f(x+k))\,|f'(x+k)|$$

approaches $\rho(x)$, the density for the invariant measure associated with f (here f refers to a one-dimensional f-expansion). Recently Schweiger [29] has generalized the Kuzmin theorem to the Jacobi algorithm when $n=2$. In this section we generalize the Kuzmin theorem to arbitrary F-expansions satisfying conditions (C), (L), (q), and $m\{x\colon B^v(x)\to x\}=1$. We also require that

$$\delta B^v(k^1, k^2, \ldots, k^v)=B^{v-1}(k^2, \ldots, k^v), \qquad v \geqq 2.$$

Our results include those of Schweiger on the Jacobi algorithm [29], and our theorem holds for Schweiger's general F-expansions [30]. However, Schweiger's general theorem has a bound of $\sigma(v)$ instead of $\sigma(\sqrt{v})+e^{-\lambda\sqrt{v}}$ (see Theorem 6.1 below), which we have been unable to obtain in our situation.

For each admissible cylinder of order 1, $B(k)$, we have $\delta\, B(k) \subseteq 0, 1)^n_F$. The sets $\delta\, B(k)$ can be used to partition the set B^0. (The induced partition is the finest partition measurable with respect to the σ-algebra generated by $\delta\, B(k)$.) We assume this partition is countable and denote our partition by $\{A_i\}_{i=1}^\infty$. With each A_i we associate

$$\mathscr{E}_i=\{k\colon \delta\, B(k)\supset A_i\}.$$

Since our interest is in statements true almost everywhere we may assume $m(A_i)>0$ for all i without loss of generality. It is clear that

$$\mathscr{E}_i \supset \{k\colon B(k) \text{ is proper}\}.$$

Next we obtain a Kuzmin relation for $\rho(\cdot)$. Suppose $E\subset A_i$. $\mu(E)=\mu(\delta^{-1}E)$ and

$$\delta^{-1}(E)=\{x\colon \delta(x)\in E\}=\bigcup_{k\in\mathscr{E}_i}\{f_k(t)\colon t\in E\}=\bigcup_{k\in\mathscr{E}_i} f_k(E).$$

Due to the 1-1 nature of F the unions above are disjoint. Thus

$$\mu(E)=\int_E \rho(x)\,dx=\int_{\delta^{-1}E}\rho(x)\,dx=\sum_{k\in\mathscr{E}_i}\int_{f_kE}\rho(x)\,dx$$
$$=\sum_{k\in\mathscr{E}_i}\int_E \rho(f_k(x))\,|J_{f_k}(x)|\,dx.$$

This implies $\rho(\cdot)$ satisfies

$$\rho(x)=\sum_{k\in\mathscr{E}_i}\rho(f_k(x))\,|J_{f_k}(x)|,\qquad x\in A_i. \tag{6.1}$$

The relation (6.1) suggests that we formulate a Kuzmin theorem where the recursion is done in separate components. The following lemma points out that such an approach is feasible:

Lemma 6.1. *Let $\psi_0(x)$ be given and $\psi_v(x)$ be defined by*

$$\psi_v(x)=\sum_{k\in\mathscr{E}_i}\psi_{v-1}(f_k(x))\,|J_{f_k}(x)|,\qquad x\in A_i,\ i=1,2,\ldots.$$

Then

$$\psi_v(x)=\sum^{(v)_i}\psi_0(f_v(x))\,|J_{f_v}(x)|,\qquad x\in A_i,$$

where $f_v=\prod_{i=1}^{v}\circ f_{k^i}$ and the last summation is over all admissible cylinders $(k^1,\ldots,k^v)$ where $k^v\in\mathscr{E}_i$.

Proof. The proof is by induction. Assume

$$\psi_v(x)=\sum^{(v)}\psi_0(f_v(x))\,|J_{f_v}(x)|,\qquad x\in A_i,\ i=1,2,\ldots,$$

where the sum is over all admissible $(k^1,\ldots,k^v)$ with $k^v\in\mathscr{E}_i$. For $v=1$ the induction hypothesis holds by definition. Now, for $x\in A_i$,

$$\begin{aligned}\psi_{v+1}(x)&=\sum_{k\in\mathscr{E}_i}\psi_v(f_k(x))\,|J_{f_k}(x)|\\&=\sum_{k\in\mathscr{E}i}\sum^{(v)_j}\psi_0(f_v(f_k(x)))\,|J_{f_v}(f_k(x))|\,|J_{f_k}(x)|,\end{aligned}$$

where the second summation is taken over all admissible $(k^1,\ldots,k^v)$ with $k^v\in\mathscr{E}_j$ and $f_k(x)\in A_j$.

Since $f_k(x)\in A_j$ and $(v)_j$ are all admissible sequences such that $\delta^v B^v\supset A_j$, then $(k^1,\ldots,k^v,k)$ with $(k^1,\ldots,k^v)\in(v)_j$ runs over all admissible sequences ending in k. This allows us to write (using the chain rule)

$$\psi_{v+1}(x)=\sum^{(v+1)_i}\psi_0(f_{v+1}(x))\,|J_{f_{v+1}}(x)|,\qquad x\in\mathscr{E}_i.$$

Lemma 6.1 motivates Theorem 6.1 below.

Theorem 6.1. *Let $\rho(\cdot)$ be any function $(q/C\leqq\rho(x)\leqq C/L)$ on B^0 satisfying*

$$\rho(x)=\sum_{k\in\mathscr{E}_i}\rho(f_k(x))\,|J_{f_k}(x)|,\qquad x\in A_i,\ i=1,2,\ldots.$$

Let $\{\psi_v\}_{v\geqq 0}$ be a sequence of recursively defined real-valued functions satisfying

$$\psi_{v+1}(x)=\sum_{k\in\mathscr{E}_i}\psi_v(f_k(x))\,|J_{f_k}(x)|,\qquad x\in A_i,\ i=1,2,\ldots,$$

where $0<m\leqq\psi_0\leqq M$; ψ_0 *and* ρ *have continuous first order partials satisfying*

$$\left|\frac{\partial\psi_0}{\partial x_j}\right|<N; \qquad \left|\frac{\partial\rho}{\partial x_j}\right|<A, \qquad j=1,\dots,n.$$

Further we assume $|J_{f_v}(x)|\leqq D_1<+\infty$ *uniformly in* v *and* x, *and*

$$\sum^{(v)_i}\left|\frac{\partial J_{f_v}(x)}{\partial x_j}\right|\leqq D_2, \qquad \sum^{(v)_i}\left|\frac{\partial(f_v(x))_k}{\partial x_j}\right|\leqq D_3,$$

uniformly in v *and* x. *Then*

$$|\psi_v(x)-a\,\rho(x)|<b\left(e^{-\lambda\sqrt{v}}+\sigma(\sqrt{v})\right)$$

where a, b, λ *are constants and*

$$\sigma(x)=\sup_{y\in B^0}(\operatorname{diam} B^v(y)), \qquad v\leqq x<v+1.$$

Remark. (i) $\sigma(x)$ is monotone decreasing.

(ii) $\sigma(x)\to 0$ as $x\to\infty$.

(iii) diam $B^v(x)\leqq\sigma(v)$.

Proof. We fix $x\in A_i$. By Lemma 6.1

$$\psi_v(x)=\sum^{(v)_i}\psi_0(f_v(x))\,|J_{f_v}(x)|, \qquad x\in A_i.$$

$$\left|\frac{\partial\psi_v(x)}{\partial x_j}\right|\leqq\sum^{(v)_i}\left\{\sum_{k=1}^{n}\left|\frac{\partial\psi_0}{\partial x_k}(f_v(x))\right|\left|\frac{\partial(f_v(x))_k}{\partial x_j}\right||J_{f_v}(x)|+\psi_0(f_v(x))\left|\frac{\partial J_{f_v}(x)}{\partial x_j}\right|\right\}$$

and

$$\left|\frac{\partial\psi_v(x)}{\partial x_j}\right|\leqq N D_1\sum^{(v)_i}\sum_{k=1}^{n}\left|\frac{\partial(f_v(x))_k}{\partial x_j}\right|+M\sum^{(v)_i}\left|\frac{\partial J_{f_v}(x)}{\partial x_j}\right|$$

$$\leqq n N D_1 D_3+M D_2=D_4.$$

We now state another Lemma to be used in the remainder of the proof:

Lemma 6.2. *If constants* t, T *exist such that*

$$t\,\rho(x)<\psi_v(x)<T\rho(x), \qquad x\in B^0,$$

then

$$t\,\rho(x)<\psi_{v+1}(x)<T\rho(x), \qquad x\in B^0.$$

We make the first application of Lemma 6.2 by noting

$$g\,\rho(x)<\frac{mL}{C}\,\rho(x)\leqq\psi_0(x)\leqq\frac{CM}{q}\,\rho(x)<G\,\rho(x),$$

where $g=\dfrac{mL}{2C}$ and $G=\dfrac{2CM}{2q}$. Then

$$g\,\rho(x)<\psi_v(x)<G\,\rho(x).$$

We define

$$\varphi_\nu(x)=\psi_\nu(x)-g\,\rho(x)$$

and

$$\zeta_\nu(x)=G\,\rho(x)-\psi_\nu(x).$$

By application of Lemma 6.1, we have

$$\varphi_\nu(x)=\sum^{(\nu)_i}\varphi_0(f_\nu(x))\,|J_{f_\nu}(x)|$$

and

$$\zeta_\nu(x)=\sum^{(\nu)_i}\zeta_0(f_\nu(x))\,|J_{f_\nu}(x)|.$$

Using condition (C)

$$\varphi_\nu(x)\geqq\frac{1}{C}\sum^{(\nu)_i}\varphi_0(f_\nu(x))\,m(B^\nu).$$

We let $\mathfrak{C}_i^\nu=\bigcup_{k^\nu\in\mathscr{E}_i}B(k^1,\ldots,k^\nu)$. By the mean value theorem for integrals ([1], p. 269),

$$\int_{\mathfrak{C}_i^\nu}\varphi_0(y)\,dy=\sum^{(\nu)_i}\varphi_0(y'_\nu)\,m(B^\nu).$$

By the mean value theorem for functions of several variables ([1], p. 117),

$$\begin{aligned}|\varphi_0(y)-\varphi_0(y')|&\leqq|\psi_0(y)-\psi_0(y')|+g\,|\rho(y)-\rho(y')|\\&\leqq\sqrt{n}(N+g\,A)\,\|y-y'\|\leqq\sqrt{n}(N+g\,A)\,\sigma(\nu)\end{aligned}$$

where y and y' belong to B^ν, and $\|\cdot\|$ is Euclidean distance. Combining the above results

$$\varphi_\nu(x)>\frac{1}{C}\int_{\mathfrak{C}_i^\nu}\varphi_0(x)\,dx-\frac{\sqrt{n}}{C}(N+g\,A)\,\sigma(\nu).$$

Now define

$$\mathfrak{l}_i=\frac{1}{C}\int_{\mathfrak{C}_i^\nu}\varphi_0(x)\,dx>0.$$

We now write

$$\varphi_\nu(x)>\mathfrak{l}_i-\frac{\sqrt{n}}{C}(N+g\,A)\,\sigma(\nu)=\mathfrak{l}_i-C_1\,\sigma(\nu).$$

In the same manner we obtain

$$\zeta_\nu(x)>\frac{1}{C}\int_{\mathfrak{C}_i^\nu}\zeta_0(x)\,dx-\frac{\sqrt{n}}{C}(N+GA)\,\sigma(\nu)$$

or, defining

$$\mathfrak{l}_i^1=\frac{1}{C}\int_{\mathfrak{C}_i^\nu}\zeta_0(x)\,dx>0,$$

$$\zeta_\nu(x)>\mathfrak{l}_i^1-C_2\,\sigma(\nu).$$

It now follows that

$$\psi_\nu(x) > g\,\rho(x) + \mathfrak{l}_i - C_1\,\sigma(\nu) = \rho(x)\left(g + \frac{\mathfrak{l}_i}{\rho(x)} - \frac{C_1\,\sigma(x)}{\rho(x)}\right)$$

$$\geqq \rho(x)\left(g + \frac{L}{C}\,\mathfrak{l}_i - \frac{C\,c_1\,\sigma(\nu)}{q}\right),$$

and

$$\psi_\nu(x) < G\,\rho(x) - \mathfrak{l}_i^1 + C_2\,\sigma(\nu) = \rho(x)\left(G - \frac{\mathfrak{l}_i^1}{\rho(x)} + \frac{C_2\,\sigma(\nu)}{\rho(x)}\right)$$

$$\leqq \rho(x)\left(G - \frac{L\,\mathfrak{l}_i^1}{C} + \frac{C_2\,C}{q}\,\sigma(\nu)\right).$$

There exists a ν_0 such that $\nu \geqq \nu_0$ implies

$$g_1 = g + \frac{L}{C}\,\mathfrak{l}_i - \frac{C\,c_1\,\sigma(\nu)}{q} > g$$

and

$$G_1 = G - \frac{L}{C}\,\mathfrak{l}_i^1 + \frac{C_2\,C}{q}\,\sigma(\nu) < G.$$

$$G_1 - g_1 = G - g - \frac{L}{C}(\mathfrak{l}_i + \mathfrak{l}_i^1) + \left(\frac{C}{q}(C_1 + C_2)\right)\sigma(\nu).$$

Now

$$\mathfrak{l}_i + \mathfrak{l}_i^1 = \frac{1}{C}\int_{\mathfrak{C}_i^\nu} (\zeta_0(x) + \varphi_0(x))\,dx = \frac{G-g}{C}\int_{\mathfrak{C}_i^\nu} \rho(x)\,dx = \frac{G-g}{C}\sum^{(\nu)_i} m(B^\nu)$$

$$\geqq \frac{G-g}{C}\sum{}' m(B^\nu) \geqq \frac{G-g}{C}(q),$$

where the last summation ($\sum'$) is over proper cylinders of order ν and the last inequality is due to condition (q). The importance of this result is to establish lower bound on $\mathfrak{l}_i + \mathfrak{l}_i^1$ which is *uniform* in i and ν. It is this bound which allows us to iterate below and obtain the geometric bound. Now

$$G_1 - g_1 \leqq (G-g)\left(1 - \frac{q\,L}{C^2}\right) + C_3\,\sigma(\nu).$$

Now $0 < q \leqq 1$, $0 < L \leqq 1$, and without loss of generality $C > 1$. Thus $0 < \hat{q} = 1 - \frac{q\,L}{C^2} < 1$, and

$$G_1 - g_1 \leqq (G-g)\,\hat{q} + C_3\,\sigma(\nu).$$

We now summarize the result just obtained. From the conditions that

$$g\,\rho(x) < \psi_0(x) < G\,\rho(x),$$

$$\left|\frac{\partial \psi_0}{\partial x_j}\right| < N, \qquad j = 1, \dots, n,$$

we have shown that for sufficiently large v

$$g_1\rho(x)<\psi_v(x)<G_1\rho(x)$$

where $g<g_1<G_1<G$; $G_1-g_1\leqq(G-g)\hat{q}+C_3\sigma(v)$.

By noting that we have earlier shown

$$\left|\frac{\partial\psi_v}{\partial x_j}\right|<D_4, \quad j=1,\ldots,n,$$

we can repeat the argument to obtain (for a fixed $v\geqq v_0$)

$$g_2\rho(x)<\psi_{2v}(x)<G_2\rho(x)$$

where $g_1<g_2<G_2<G_1$, and

$$G_2-g_2\leqq(G_1-g_1)\hat{q}+C_4\sigma(v)$$

where C_4 is a function of D_4.

The argument for g_2 and G_2 can be iterated to obtain

$$g_r\rho(x)<\psi_{rv}(x)<G_r\rho(x)$$

where

$$\begin{aligned}G_r-g_r&\leqq(G_{r-1}-g_{r-1})\hat{q}+C_4\sigma(v)\leqq\cdots\\&\leqq\hat{q}^r(G-g)+C_4\sigma(v)(\hat{q}^{r-1}+\hat{q}^{r-2}+\cdots+1)\\&\leqq\hat{q}^r(G-g)+\frac{C_4}{1-\hat{q}}\sigma(v).\end{aligned}$$

Therefore

$$G_r-g_r\leqq C_5e^{-\lambda r}+C_6\sigma(v)$$

where $-\lambda=\log\hat{q}$, $C_5=G-g$, $C_6=\dfrac{C_4}{1-\hat{q}}$, and $\lim\limits_{v=r\to\infty}G_r=\lim\limits_{v=r\to\infty}g_r=a$ follows. If we set $r=v$,

$$|\psi_{v^2}(x)-a\rho(x)|<\rho(x)(G_v-g_v)\leqq\frac{C}{L}(C_5e^{-\lambda v}+C_6\sigma(v)).$$

This gives us

$$a=\lim_{v\to\infty}\int_{B^0}\psi_{v^2}(x)\,dx.$$

Finally choose $N\in[v^2,(v+1)^2)$. Applying Lemma 6.2 we obtain

$$|\psi_N(x)-a\rho(x)|<\frac{C}{L}(C_5e^{-\lambda v}+C_6\sigma(v))\leqq\frac{C}{L}(C_5e^{+\lambda}e^{-\lambda(v+1)}+C_6\sigma(\sqrt{N})).$$

Finally

$$|\psi_N(x)-a\rho(x)|<b(e^{-\lambda\sqrt{N}}+\sigma(\sqrt{N})),$$

setting $b=\dfrac{C}{L}C_5e^{\lambda}+\dfrac{C}{L}C_6$.

Corollary 6.1. *The constant a of the previous theorem has the value*

$$a = \int_{B^0} \psi_0(x)\,dx.$$

Proof.

$$\int_{B^0} \psi_\nu(x)\,dx = \sum_i \int_{A_i} \psi_\nu(x)\,dx = \sum_i \sum^{(\nu)_i} \int_{A_i} \psi_0(f_\nu(x))\,|J_\nu(x)|\,dx.$$

But f_ν takes A_i 1-1 onto $f_\nu(A_i)$. Thus

$$\begin{aligned}\int_{B^0} \psi_\nu(x)\,dx &= \sum_i \sum^{(\nu)_i} \int_{f_\nu^{-1} f_\nu(A_i)} \psi_0(f_\nu(x))\,|J_\nu(x)|\,dx\\ &= \sum_i \sum^{(\nu)_i} \int_{f_\nu A_i} \psi_0(x)\,dx = \int_{B^0} \psi_0(x)\,dx.\end{aligned}$$

The Kuzmin theorem appears in Schweiger's paper [30] with an error bound of the form $b\,\tau(\nu)$. He then uses this result to generalize a result of Gauss that states

$$m(\delta^{-\nu}(E)) \to \mu(E).$$

We have no difficulty generalizing Schweiger's theorem to the F-expansions considered above.

Theorem 6.2. *Let* $F \in \mathfrak{F}$ *satisfy conditions* (C), (L), (q), *and* $m\{X: B^\nu(x) \to x\} = 1$. *Also assume that the hypothesis of Kuzmin's Theorem* 9.1 *is satisfied for* ρ *and* f_ν. *Then*

$$|m(\delta^{-\nu}(E)) - \mu(E)| < b\, m(E)\big(e^{-\lambda\sqrt{\nu}} + \sigma(\sqrt{\nu})\big),$$

for all Borel sets E.

Proof. Define $\psi_0(x) = 1$. Then Theorem 6.1 states

$$|\psi_\nu(x) - a\,\rho(x)| < b\big(e^{-\lambda\sqrt{\nu}} + \sigma(\sqrt{\nu})\big). \tag{6.2}$$

But Corollary 6.1 shows $a = \int_{B^0} \psi_0(x)\,dx = 1$. Multiplying (6.2) by $I_E(x)$ and integrating, we have

$$\Big|\int_{B^0} I_E(x)\,\psi_\nu(x)\,dx - \mu(E)\Big| < b\, m(E)\big(e^{-\lambda\sqrt{\nu}} + \sigma(\sqrt{\nu})\big).$$

We complete the proof by

$$\begin{aligned}\int_{B^0} I_E(x)\,\psi_\nu(x)\,dx &= \sum_i \int_{A_i} \psi_\nu(x)\, I_E(x)\,dx\\ &= \sum_i \sum^{(\nu)_i} \int_{A_i} \psi_0(f_\nu(x))\, I_E(x)\,|J_\nu(x)|\,dx\\ &= \sum_i \sum^{(\nu)_i} \int_{f_\nu A_i} I_E(\delta^\nu x)\,dx\\ &= \int_{B^0} I_{\delta^{-\nu}E}(x)\,dx = m(\delta^{-\nu}E).\end{aligned}$$

The next application of the Kuzmin Theorem gives us a rate on the strong mixing condition. It is this result which allows us to obtain the limit theory.

Theorem 6.3. *Let $F \in \mathfrak{F}$ satisfy conditions* (C), (L), (q), *and* $m\{x: B^v(x) \to x\} = 1$. *Also assume the hypothesis of Theorem* 6.1 *is satisfied for* ρ *and* f_v. *Then*

$$|\mu(B(k^1, \ldots, k^s) \cap \delta^{-v-s} E) - \mu(B(k^1, \ldots, k^s))\,\mu(E)|$$
$$\leqq \frac{Cb}{q}\,(e^{-\lambda\sqrt{v}} + \sigma(\sqrt{v}))\,\mu(B(k^1, \ldots, k^s))\,\mu(E),$$

for all admissible $k^1, \ldots, k^s$ *and Borel sets* E.

Proof. Let $B(k^1, \ldots, k^s) = B^s$ be a proper cylinder. Define

$$\psi_0(x) = \frac{I_{B^s}(x)}{\mu(B^s)}\,\rho(x),$$

with ψ_v defined recursively as usual.

Since $\psi_0(x) = 0$ for $x \notin B^s$, we cannot apply Kuzmin's Theorem to $\{\psi_v\}_{v \geqq 0}$. However, since $\mathscr{E}_i \supset \{k: B(k) \text{ is proper}\}$,

$$\psi_s(x) = \frac{\rho(f_s(x))}{\mu(B^s)}\,|J_s(x)|, \qquad x \in B^0.$$

The hypothesis of Theorem 6.1 then applies to the sequence $\{\psi_v\}_{v \geqq s}$, and

$$|\psi_v(x) - a\,\rho(x)| < b(e^{-\lambda\sqrt{v}} + \sigma(\sqrt{v})). \tag{6.3}$$

Applying Corollary 6.1 to $\{\psi_v\}_{v \geqq 0}$ gives $a = 1$. We multiply (6.3) by $I_E(\cdot)$ and integrate to obtain

$$\Big|\int_{B^0} I_E(x)\,\psi_v(x)\,dx - \mu(E)\Big| < b(e^{-\lambda\sqrt{v}} + \sigma(\sqrt{v}))\,m(E) \leqq \frac{Cb}{q}\,(e^{-\lambda\sqrt{v}} + \sigma(\sqrt{v}))\,\mu(E).$$

Proceeding as in Corollary 6.1,

$$\begin{aligned}
\int_{B^0} I_E(x)\,\psi_v(x)\,dx &= \sum_i \int_{A_i} I_E(x)\,\psi_v(x)\,dx \\
&= \sum_i \sum^{(s+v)_i} \int_{A_i} I_E(x)\,\psi_0(f_{v+s}(x))\,|J_v(x)|\,dx \\
&= \sum_i \sum^{(s+v)_i} \int_{f_{v+s}(A_i)} \psi_0(x)\,I_E(\delta^{s+v}(x))\,dx = \int_{B^0} \psi_0(x)\,I_{\delta^{-s-v}E}(x)\,dx \\
&= \frac{\mu(B^s \cap \delta^{-s-v} E)}{\mu(B^s)}.
\end{aligned}$$

Thus the conclusion of our theorem holds for proper cylinders. Lemma 4.2 states that the measure of any cylinder can be approximated by the measure of countably many disjoint proper cylinders whenever condition (q) holds. This completes the proof.

To apply the limit theory developed by Ibragimov ([7 to 9] and Reznik [19] we need the following assumption

$$\text{(M)} \qquad \sum_{v=1}^{\infty} \left(e^{-\lambda\sqrt{v}} + \sigma(\sqrt{v})\right)^{\frac{1}{2}} < +\infty.$$

For the cylinder $B^v(x)$ and $g \in L_2(0,1)^n$ we write

$$[g]_v^{(x)} = \frac{1}{\mu(B^v(x))} \int_{B^v(x)} f\, d\mu$$

and

$$\|g\|_\mu = \{\int g^2\, d\mu\}^{\frac{1}{2}}.$$

Then we have

Theorem 6.4. *Suppose* $F \in \mathfrak{F}$ *satisfies conditions* (C), (L), (q), (M) *and* $m\{x: \operatorname{diam} B^v(x) \to 0\} = 1$. *Also suppose the hypothesis of Theorem* 6.1 *is satisfied for* ρ *and* f_v. *Assume* $g \in L_2(0,1)^n$ *and* $\int g\, d\mu = 0$. *Then if*

$$\sum_{v=1}^{\infty} \|g - [g]_v\|_\mu < +\infty,$$

it follows that

$$\|g\|_\mu^2 + 2 \sum_{v=0}^{\infty} \int_{(0,1)^n} g(t)\, g(\delta^v(t))\, dt = \sigma^2 < +\infty$$

and, if $\sigma \neq 0$,

(i) (Ibragimov)

$$\mu\left\{t: \frac{\sum_{k=0}^{v-1} g(\delta^k(t))}{\sigma\sqrt{v}} < z\right\} \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{u^2}{2}}\, du \qquad \text{as } v \to \infty,$$

and

(ii) (Reznik)

$$\varlimsup_{v\to\infty} \frac{\left|\sum_{k=0}^{v-1} g(\delta^k(t))\right|}{(2\sigma^2 \ln\ln v)^{\frac{1}{2}}} = 1 \qquad \text{a.e. } [m].$$

One-dimensional examples of Theorem 6.4 (i) and (ii) can be found in Vinh Hien [31], Ibragimov [8], and Reznik [19].

VII. Examples

In this section we present two main examples to which our theory has application. Besides the independence example developed first, we develop a generalization of the Jacobi algorithm.

The simplest and most natural algorithm is

$$F(x_1, \ldots, x_n) = (f_1(x_1), f_2(x_2), \ldots, f_n(x_n)),$$

$$F: \mathop{\times}_{i=1}^{n} (A_i, B_i) \to (0,1)^n,$$

where f_i is either in the class of increasing f's considered by Everett [6] or decreasing f's considered by Bissinger [3]. It is clear that $\delta(t)=(\delta_1(t_1),\ldots,\delta_n(t_n))$ where $\delta_i(\cdot)$ is the shift for the process generated by f_i. If each f_i (with $n=1$) satisfies condition (C) with C_i, then we have invariance (μ_i) and ergodicity. Thus the F being considered has $C=\prod_{i=1}^{n} C_i$, $q=L=1$, and $\mu=\mathop{\times}_{i=1}^{n}\mu_i$. The results available to us could of course have been obtained as direct application of Renyi [20].

Next we define a modified Jacobi algorithm and indicate its development. The proofs follow Perron [17] and Schweiger ([23, 25, 26]) and are not included here due to their excessive length. Our algorithm generalizes the Jacobi algorithm and an example of Kinney and Pitcher [12] who consider $f(x)=M/x$.

Let $M_1, M_2, \ldots, M_n$ be a fixed set of positive reals such that $M_j \geqq 1$ for $j=1,\ldots,n$. We define $F(\cdot)$ by

$$F(x_1,\ldots,x_n)=\left(\frac{M_1}{x_n},\frac{M_2 x_1}{x_n},\ldots,\frac{M_n x_{n-1}}{x_n}\right)$$

where $x\in\left\{(z_1,\ldots,z_n): z_n\geqq M_1;\ \frac{z_n}{z_{j-1}}\geqq M_j, j=2,\ldots,n\right\}$. It is easy to see that $D=F^{-1}$ has the form

$$D(y_1,\ldots,y_n)=\left(\frac{M_1}{M_2}\frac{y_2}{y_1},\frac{M_1}{M_3}\frac{y_3}{y_1},\ldots,\frac{M_1}{M_n}\frac{y_n}{y_1},\frac{M_1}{y_1}\right).$$

Some simple computations show $\{x\in(0,1)^n:\ \delta_1^k(x)=0$ for some $k\geqq 1\}$ to be a countable set of hyperplanes intersected with $(0,1)^n$ and hence to have measure 0.

The coefficient matrices defined below correspond to those which Schweiger used for the Jacobi algorithm. Relationships of our matrices to F are given below.

$$A^0=\begin{pmatrix} 0 & 0 & 0\cdots 0 & 0 & \frac{1}{F_n} \\ 1 & 0 & 0\cdots 0 & 0 & 0 \\ 0 & \frac{M_2}{F_1} & 0\cdots 0 & 0 & 0 \\ & \cdots & & \cdots & \\ 0 & 0 & 0\cdots 0 & \frac{M_n}{F_{n-1}} & 0 \end{pmatrix},$$

$$A^1=\begin{pmatrix} 0 & 0\cdots 0 & M_1 \\ 1 & 0\cdots 0 & F_1 a_1^k \\ 0 & 1\cdots 0 & F_2 a_2^k \\ \cdots & & \cdots \\ 0 & 0\cdots 1 & F_n a_n^k \end{pmatrix},$$

and

$$A^k=\begin{pmatrix} 0 & 0\cdots 0 & 1 \\ 1 & 0\cdots 0 & F_1 a_1^k \\ 0 & 1\cdots 0 & F_2 a_2^k \\ \cdots & & \cdots \\ 0 & 0\cdots 1 & F_n a_n^k \end{pmatrix}, \quad k\geqq 2,$$

where

$$F_j = \frac{\left(\prod_{i=1}^{n} M_i\right)^{j/n+1}}{\prod_{i=1}^{j} M_i}, \qquad j=1,2,\ldots,n.$$

Next we define

$$\begin{aligned}\Omega^k(x) &= \prod_{v=0}^{k-1} \Lambda^v(x)\\ &= (\omega_{ij}^k), \quad i=0,\ldots,n,\ j=0,\ldots,n.\end{aligned}$$

It is not hard to see that det $\Omega^k(x) = (-1)^{nk} \prod_{i=1}^{n} \frac{M_i}{F_i}$, $k \geqq 2$.

The following lemma allows us to recover x and corresponds to a result of Perron [17].

Lemma 7.1. *If* $x \in (0,1)_F^n$, *then*

$$x_i = \frac{\omega_{i,n}^{k+1} + \sum_{j=1}^{n} \omega_{i,j}^k \delta_j^k F_j}{\omega_{0,n}^{k+1} + \sum_{j=1}^{n} \omega_{0,j}^k \delta_j^k F_j},$$

where $k=1,2,\ldots,$ $i=1,2,\ldots,n$, *and* $\omega_{i,j}^k \in \Omega^k(x)$.

Using Lemma 7.1 and following Perron [17], the following convergence theorem can be obtained.

Theorem 7.1. *If* $x \in (0,1)_F^n$, *then*

$$\lim_{k\to\infty} \frac{\omega_{i,n}^k(x)}{\omega_{0,n}^k(x)} = x_i, \qquad i=1,2,\ldots,n.$$

Our most difficult task comes in evaluating $J_{f_v}(t)$ where $f_v(t) = \prod_{i=1}^{v} \circ f_{k^i}(t)$. A determinant evaluation similar to one which Schweiger ([23], p. 72) makes must be carried out. We find

$$|J_v(t)| = \frac{\prod_{i=1}^{n} M_i}{\left(\omega_{0n}^{v+1} + \sum_{l=1}^{n} \omega_{0l}^v t_l F_l\right)^{n+1}}.$$

This yields

$$\frac{\sup |J_v(\cdot)|}{\inf |J_v(\cdot)|} \leqq \left(\frac{1+2\sum_{l=1}^{n} F_l}{F_n}\right)^{n+1} = C < +\infty$$

so that condition (C) is verified.

It is difficult to verify condition (L) for arbitrary M_i. If $M_2 = M_3 = \cdots = M_n = 1$ with M_1 an integer, then $\delta^v B^v \supset \{X : x_j \leqq x_n,\ x_n \in (0,1)$ and $x \in (0,1)_F\}$, which occurs if and only if $k^v = (k,k,k,\ldots,k)$, $k \geqq 1$. Thus $m(\delta^v B^v) \geqq 1/n! = L > 0$ and condition (L) is verified.

Given any cylinder B^ν we choose a proper cylinder $\hat{B}^{\nu+1} \subset B^\nu$ with $a^{\nu+1} = (0, 0, \ldots, \hat{M})$. This gives condition (q) with

$$q = L\left(\frac{F_n}{\left(1+\sum_{l=1}^{n} F_l\right)\left(1+2\sum_{l=1}^{n} F_l\right)\hat{M}}\right)^{n+1} > 0.$$

Thus Theorems 4.2, 4.3 and 5.4 hold. Since the form of $\rho(x)$ is not known we cannot conclude Kuzmin's theorem holds.

As an alternative to the classical approach to Theorem 7.1, we can conclude a.e. convergence as a result of Theorem 3.3, since it can be shown that

$$\operatorname{diam}(B^{\nu+n}) \leqq \lambda \operatorname{diam}(B^\nu).$$

where $\lambda = (2nF)(\max\{M_1, 2([2nF]+1)\})^{-1}$ and

$$q' = \left(\frac{LF_n^{n+1}}{\left(1+\sum_{l=1}^{n} F_l\right)^{n+1}\left(1+\sum_{l=1}^{n} F_l\right)^{n+1}}\right)^n \left(\frac{1}{M_1^{n-1}(\max\{M_1, 2([2nF]+1)\})}\right)^{n+1}.$$

It is quite natural to ask whether all $F \in \mathfrak{F}$ satisfying conditions (C), (q), and δ ergodic have the property that $m\{x: B^\nu(x) \to x\} = 1$. This is not the case as our consideration of $F(x, y) = (1/x, y + \xi \bmod 1)$ (where ξ is irrational) will show. Now $|J_F(x, y)| = 1/x^2 \neq 0$ for all $x \in [1, \infty)$ and the cylinders are rectangles. The invariant measure is simply the product of Lebesgue measure and Gauss' measure:

$$d_\mu(x, y) = \frac{dx\, dy}{(\log 2)(1+x)}.$$

Since δ_1 and δ_2 are both ergodic on (0, 1) (see [2]) we may conclude $\delta = \delta_1 \times \delta_2$ ergodic on $(0, 1)^2$. Also $L = q = 1$ and, following Renyi [20], we may choose $C = 4$. Thus all conditions are satisfied but $B^\nu(x_1, x_2)$ is a rectangle of height 1 erected above the one-dimensional continued fraction cylinder for x_1. Therefore $\operatorname{diam}(B^\nu(x)) \geqq 1$.

Acknowledgments. I would like to express my gratitude to Professor J. R. Kinney for introducing me to information theory and for his guidance in the preparation of my thesis from which this paper is taken. Also I would like to express my appreciation to the Department of Statistics and Probability at Michigan State University and to the National Science Foundation for providing financial support.

References

1. Apostol, Tom M.: Mathematical analysis. A modern approach to advanced calculus. Reading, Mass. and London: Addison-Wesley Publishing Company, Inc. 1960.
2. Billingsley, Patrick: Ergodic theory and information. New York-London-Sidney: John Wiley & Sons, Inc. 1965.
3. Bissinger, B. H.: A generalization of continued fractions. Bull. Amer. math. Soc. **50**, 868–876 (1944).
4. Chung, K. L.: A note on the ergodic theorem of information theory. Ann. math. Statistics **32**, 612–614 (1961).
5. Dunford, N., Miller, D. S.: On the ergodic theorem. Trans. Amer. math. Soc. **60**, 538–549 (1946).
6. Everett, G. I.: Representations for real numbers. Bull. Amer. math. Soc. **52**, 861–869 (1946).

7. Ibragimov, I.A.: Some limit theorems for stochastic processes stationary in the strict sense. Doklady Akad. Nauk. SSSR u. Sev. **125**, 711-714 (1959): English transl. Select. Transl. math. Statist. and Probab. **4**, 99-103 (1963).
8. – Asymptotic distribution of the values of certain sums. Vestnik Leningrad. Univ. **15**, 7, 81-115 (1960): English transl. Select. Transl. math. Statist. and Probab. **5**, 33-48 (1965).
9. – Some limit theorems for stationary processes. Theor. Probab. Appl. **7**, **4**, 349-382 (1962).
10. Jacobi, C.G.J.: Allgemeine Theorie der kettenbruchähnlichen Algorithmen, in welchen jede Zahl aus drei vorhergehenden gebildet wird. J. reine angew. Math. **69**, 29-64 (1868).
11. Khinchin, A. Ya: Continued fractions. Chicago and London: The University of Chicago Press 1964.
12. Kinney, J., Pitcher, T.: The dimension of some sets defined in terms of f-expansions. Z. Wahrscheinlichkeitstheorie verw. Geb. **4**, 293-315 (1966).
13. Knopp, K.: Mengen theoretische Behandlung einiger Probleme der diophantischen Approximationen und der transfiniten Wahrscheinlichkeiten. Math. Ann. **95**, 409-426 (1926).
14. Kuzmin, R.O.: Sur un probleme de Gauss. Atti del Congresso Internazionale del Matematici Bologna **6**, 83-89 (1928).
15. Parry, W.: On Rohlin's formula for entropy. Acta math. Acad. Sci. Hungar. **15**, 107-113 (1964).
16. Parry, William: Entropy and generators in ergodic theory. New York and Amsterdam: Benjamin, Inc. 1969.
17. Perron, Oskar: Grundlagen für eine Theorie des Jacobischen Kettenbruchalgorithmus. Math. Ann. **64**, 1-76 (1907).
18. Renyi, A.: Representations for real numbers and their ergodic properties. Acta math. Acad. Sci. Hungar. **8**, 477-493 (1957).
19. Reznik, M.H.: The law of the iterated logarithm for some classes of stationary processes (in Russian). Teor. Verojatn. Primen. **8**, 4, 642-656 (1968).
20. Rohlin, V.A.: Exact endomorphisms of a Lebesgue space. Izvestya Akad. Nauk. SSSR, Ser. mat. **24**, 499-530 (1960): English Transl. Amer. math. Soc. Translat. II Ser. **39**, 1-36 (1964).
21. Ryll-Nardzewski, C.: On the ergodic theorems (II). Ergodic theory of continued fractions. Studia math. **12**, 74-79 (1951).
22. Saks, S.: Theory of the integral. New York: Hafner Publishing Company 1937.
23. Schweiger, F.: Geometrische und elementar metrische Sätze über den Jacobischen Algorithmus. Österr. Akad. Wiss. Wien, math.-naturw. Kl., S.-Ber. Abt. II, **173**, 59-92 (1964).
24. – Metrische Sätze über den Jacobischen Algorithmus. Monatsh. Math. **69**, 243-255 (1965).
25. – Ergodische Theorie des Jacobischen Algorithmus. Acta arithmetica **11**, 451-460 (1966).
26. – Existenz eines invarianten Masses beim Jacobischen Algorithmus. Acta arithmetica **12**, 263-268 (1967).
27. – Mischungseigenschaften und Entropie beim Jacobischen Algorithmus. J. reine angew. Math. **229**, 50-56 (1968).
28. – Induzierte Masse und Jacobischen Algorithmus. Acta arithmetica **13**, 419-422 (1968).
29. – Ein Kuzminscher Satz über den Jacobischen Algorithmus. J. reine angew. Math. **232**, 35-40 (1968).
30. – Metrische Theorie einer Klasse zahlentheoretischer Transformationen. Acta arithmetica **15**, 1-18 (1968).
31. Vinh-Hien, Tran.: The central limit theorem for stationary processes generated by number theoretic endomorphisms (in Russian). Vestnik Moskov Univ. Ser. I. Mat. Meh. **5**, 1 28-34 (1963).

Dr. Michael S. Waterman
Department of Mathematics
Idaho State University
Pocatello, Idaho 83201

(Received June 16, 1969)

Universitätsdruckerei H. Stürtz AG Würzburg
Printed in Germany