Z. Wahrscheinlichkeitstheorie verw. Gebiete 31, 163–173 (1975)

Hausdorff Dimension of Some Continued-Fraction Sets

Patrick Billingsley* and Inge Henningsen

Section 1

Introduction. Each irrational ω in the unit interval Ω has a unique continued-fraction expansion

$$\omega = \frac{1\,|}{|a_1(\omega)} + \frac{1\,|}{|a_2(\omega)} + \cdots$$

with integral partial quotients $a_j(\omega)$. We shall compute the Hausdorff dimension of certain sets defined in terms of the frequencies with which the $a_j(\omega)$ assume various values.

Let $N_n(i, \omega)$ be the number of j, $1 \leqq j \leqq n$, for which $a_j(\omega) = i$. For a probability vector $p = (p_1, p_2, \ldots)$, let $L(p)$ be the set of ω for which

$$\lim_{n\to\infty} \frac{1}{n} N_n(i, \omega) = p_i, \qquad i = 1, 2, \ldots, \tag{1.1}$$

and for a set A of probability vectors, let $L(A) = \bigcup_{p\in A} L(p)$. Thus $L(A)$ is the set of ω for which the frequency vector $(n^{-1} N_n(1, \omega), n^{-1} N_n(2, \omega), \ldots)$ approaches (in each component) some element of A.

We first give a lower bound for $\dim L(A)$ in terms of entropy. Call a probability measure v on the Borel subsets of Ω stationary if under v the stochastic process $[a_1(\omega), a_2(\omega), \ldots]$ is stationary; in this case, let $h(v)$ be the entropy of the process. Call v ergodic if under v the process is ergodic. Put

$$\ell(v) = -2 \int \log \omega \, v(d\omega). \tag{1.2}$$

On the set of k-tuples of positive integers, define a probability measure v_k by

$$v_k(i_1, \ldots, i_k) = v[\omega\colon a_j(\omega) = i_j, j \leqq k]. \tag{1.3}$$

If for each k, P_k is a probability measure on the k-tuples, and if $\{P_k\}$ is consistent, then there exists a v such that $v_k = P_k$ for all k.

The entropy $h(v)$ is finite if and only if $-\sum_i v_1(i) \log v_1(i)$ converges. The quantity $\ell(v)$ is always positive, and since $a_1(\omega) \leqq \omega^{-1} \leqq a_1(\omega) + 1$, it is finite if and only if $\sum_i v_1(i) \log i$ converges. Let $\mathcal{N}$ be the set of stationary and ergodic v for which $h(v)$ and $\ell(v)$ are both finite.

Theorem 1. *For each set A of probability vectors,*

$$\dim L(A) \geqq \sup\left[\frac{h(v)}{\ell(v)}\colon\ v \in \mathcal{N},\ v_1 \in A\right]. \tag{1.4}$$

---

* Supported by NSF GP 32037X.

This result, related to theorems of Kinney and Pitcher [6], is proved in Section 2. For inequalities going the other way, we must restrict attention to the set Ω_r of ω for which $a_j(\omega) \leqq r$, $j=1, 2, \ldots$. If A is a set of r-dimensional probability vectors $p=(p_1, \ldots, p_r)$, let $U_r(A)$ be the set of ω in Ω_r such that for every positive ε and n_0 there exist a p in A and an integer $n \geqq n_0$ that satisfy

$$\left|\frac{1}{n} N_n(i, \omega) - p_i\right| < \varepsilon, \qquad i=1, \ldots, r. \tag{1.5}$$

Thus $U_r(A)$ is the set of ω in Ω_r for which some limit point of the frequency vectors $(n^{-1} N_n(1, \omega), \ldots, n^{-1} N_n(r, \omega))$ lies in the closure of the set A.

Let $\mathcal{N}_r$ be the set of v in $\mathcal{N}$ supported by Ω_r. A v in $\mathcal{N}$ lies in $\mathcal{N}_r$ if and only if $v_1(i)=0$ for $i>r$; in this case we may regard v_1 as an r-dimensional probability vector. Finally, let $\mathcal{M}_r$ consist of the v in $\mathcal{N}_r$ under which $[a_1(\omega), a_2(\omega), \ldots]$ is a Markov chain of some order.

Theorem 2. *For each set A of r-dimensional probability vectors,*

$$\dim U_r(A) \leqq \sup\left[\frac{h(v)}{\ell(v)}:\ v \in \mathcal{M}_r,\ v_1 \in A\right]. \tag{1.6}$$

This theorem is proved in Section 4. If p is an r-dimensional probability vector, let $L_r(p)$ be the set of ω in Ω_r that satisfy (1.1) for $i \leqq r$, and put $L_r(A) = \bigcup_{p \in A} L_r(p)$. Since $L_r(A) \subset U_r(A)$ and $\mathcal{M}_r \subset \mathcal{N}_r$, it is a simple matter (see Section 3) to combine Theorems 1 and 2:

Theorem 3. *For each set A of r-dimensional probability vectors,*

$$\begin{aligned}\dim L_r(A) = \dim U_r(A) &= \sup\left[\frac{h(v)}{\ell(v)}:\ v \in \mathcal{N}_r,\ v_1 \in A\right] \\ &= \sup\left[\frac{h(v)}{\ell(v)}:\ v \in \mathcal{M}_r,\ v_1 \in A\right].\end{aligned} \tag{1.7}$$

Let C_k be the set of r^k-dimensional vectors p with nonnegative components $p(i_1, \ldots, i_k)$ satisfying the constraint

$$\sum p(i_1, \ldots, i_k) = 1 \tag{1.8}$$

(in this sum, as in (1.11), the indices range from 1 to r) and, for each $i_1, \ldots, i_{k-1}$, the constraint

$$\sum_i p(i_1, \ldots, i_{k-1}, i) = \sum_i p(i, i_1, \ldots, i_{k-1}). \tag{1.9}$$

Define τ_k: $C_k \to C_1$ by

$$(\tau_k(p))_i = \sum_{i_1, \ldots, i_{k-1}} p(i, i_1, \ldots, i_{k-1}). \tag{1.10}$$

Let $q_k(i_1, \ldots, i_k)$ be the denominator of the fraction

$$\frac{1|}{|i_1} + \cdots + \frac{1|}{|i_k}$$

(in lowest terms). Finally, for $A \subset C_1$, define

$$\alpha_k(A) = \sup_{\substack{p \in C_k \\ \tau_k(p) \in A}} \frac{-\sum p(i_1, \ldots, i_k) \log p(i_1, \ldots, i_k)}{2 \sum p(i_1, \ldots, i_k) \log q_k(i_1, \ldots, i_k)} \tag{1.11}$$

(with the convention $0 \log 0 = 0$). In all this, r is fixed and hence does not appear in the notation.

Theorem 4. *For each $A \subset C_1$,*

$$\dim L_r(A) = \dim U_r(A) = \lim_{k \to \infty} \alpha_k(A). \tag{1.12}$$

This theorem is proved in Section 5. Since $U_r(C_1) = \Omega_r$, (1.7) and (1.12) determine $\dim \Omega_r$. Good [5] (see also [7, 8]) has given a different expression for $\dim \Omega_r$; Section 6 deals with the connection between the two results and with a conjecture the connection suggests.

The sets $L_r(A)$ and $U_r(A)$ have analogues defined in terms of the frequencies with which the v-tuples $(a_j(\omega), \ldots, a_{j+v-1}(\omega))$ assume various values. It will be clear from the proofs that the corresponding analogues of the above theorems hold, with the appropriate restrictions on v_v in the suprema. It will also be clear that the set $[1, \ldots, r]$ to which the partial quotients have been restricted can be replaced by any finite set.

Section 2

Proof of Theorem 1. For each n, Ω splits into the fundamental intervals of order n, the various sets

$$[\omega\colon a_j(\omega) = i_j, j \leqq n];$$

let $u_n(\omega)$ be that fundamental interval of order n containing ω. If v is stationary and ergodic and the entropy $h(v)$ is finite, then according to [4],

$$\lim_{n \to \infty} \left[-\frac{1}{n} \log v(u_n(\omega)) \right] = h(v) \tag{2.1}$$

except on a set of v-measure 0.

In all that follows, λ will represent Lebesgue measure. If $q_n(\omega)$ is the denominator of the nth convergent, then $\lambda^{-1}(u_n(\omega)) = q_n(\omega)(q_n(\omega) + q_{n-1}(\omega))$ (see [3; (4.10)], for example), and hence

$$2 \log q_n(\omega) \leqq -\log \lambda(u_n(\omega)) \leqq 2 \log q_n(\omega) + \log 2. \tag{2.2}$$

Moreover [3; (4.21)],

$$\left| \log q_n(\omega) + \sum_{j=1}^{n} \log \left(\frac{1|}{|a_j(\omega)} + \frac{1|}{|a_{j+1}(\omega)} + \cdots \right) \right| \leqq 4. \tag{2.3}$$

If v is stationary and ergodic and $t(v)$ is finite, then

$$-\frac{2}{n} \sum_{j=1}^{n} \log \left(\frac{1|}{|a_j(\omega)} + \frac{1|}{|a_{j+1}(\omega)} + \cdots \right) \to t(v)$$

almost everywhere by the ergodic theorem, and hence, by (2.2) and (2.3),

$$\lim_{n\to\infty}\left[-\frac{1}{n}\log\lambda(u_n(\omega))\right]=\ell(\nu). \tag{2.4}$$

If $\nu\in\mathcal{N}$, then (2.1) and (2.4) both hold except for ω in a set of ν-measure 0.

Let $\dim_\nu M$ denote dimension relative to ν and to coverings by fundamental intervals. That is, let $\nu_\alpha(M,\rho)=\inf\sum_i \nu(v_i)^\alpha$, the infimum extending over countable coverings $[v_i]$ of M by fundamental intervals with $\nu(v_i)<\rho$, let

$$\nu_\alpha(M)=\lim_{\rho\to 0}\nu_\alpha(M,\rho),$$

and take $\dim_\nu M$ as that α_0 such that $\nu_\alpha(M)=\infty$ for $\alpha<\alpha_0$ and $\nu_\alpha(M)=0$ for $\alpha>\alpha_0$. See [1].

By Theorem 2.4 of [2] (or see [3; Section 14]), if $\log\nu(u_n(\omega))/\log\lambda(u_n(\omega))\to\delta$ on M, then $\dim_\lambda M=\delta\dim_\nu M$, so that $\dim_\lambda M=\delta$ if $\nu(M)>0$. By Lemma 3.1 of [6], if $-n^{-1}\log\lambda(u_n(\omega))$ goes to a positive constant on M, then $\dim_\lambda M$ coincides with ordinary Hausdorff dimension $\dim M$. Thus we have the following result.

Lemma 1. *If* $\alpha\geqq 0$ *and* $\beta>0$, *then the three relations*

$$M\subset\left[\omega\colon \lim_{n\to\infty}\left[-\frac{1}{n}\log\nu(u_n(\omega))\right]=\alpha\right], \tag{2.5}$$

$$M\subset\left[\omega\colon \lim_{n\to\infty}\left[-\frac{1}{n}\log\lambda(u_n(\omega))\right]=\beta\right], \tag{2.6}$$

$$\nu(M)>0, \tag{2.7}$$

together imply

$$\dim M=\frac{\alpha}{\beta}. \tag{2.8}$$

Let M_ν be the intersection of the two sets on the right in (2.5) and (2.6) with $\alpha=h(\nu)$ and $\beta=\ell(\nu)$. By (2.1) and (2.4), $\nu\in\mathcal{N}$ implies $\nu(M_\nu)=1$. Applying Lemma 1 to $M\cap M_\nu$, we see that $\nu(M)>0$ implies $\dim M\geqq\dim M\cap M_\nu=h(\nu)/\ell(\nu)$, which gives this result:

Lemma 2. *For arbitrary* M,

$$\dim M\geqq\sup\left[\frac{h(\nu)}{\ell(\nu)}\colon\ \nu\in\mathcal{N},\ \nu(M)>0\right]. \tag{2.9}$$

If $\nu\in\mathcal{N}$ and $\nu_1=p=(p_1,p_2,\ldots)$, then by the ergodic theorem, $\nu(L(p))=1$. Therefore, by Lemma 2,

$$\dim L(p)\geqq\sup\left[\frac{h(\nu)}{\ell(\nu)}\colon\ \nu\in\mathcal{N},\ \nu_1=p\right], \tag{2.10}$$

and Theorem 1 follows from this and the definition of $L(A)$.

Section 3

Proof of Theorem 3. If $v \in \mathcal{N}_r$, then $v(\Omega_r)=1$, so that, in addition to (2.10), Lemma 2 gives

$$\dim L_r(p) \geqq \sup \left[\frac{h(v)}{\ell(v)} : v \in \mathcal{N}_r,\ v_1 = p\right]$$

for r-dimensional probability vectors p, and hence

$$\dim L_r(A) \geqq \sup \left[\frac{h(v)}{\ell(v)} : v \in \mathcal{N}_r,\ v_1 \in A\right] \tag{3.1}$$

for sets A of such vectors. Since (1.7) is a consequence of (1.6), (3.1), and the relations $L_r(A) \subset U_r(A)$ and $\mathcal{M}_r \subset \mathcal{N}_r$, Theorem 3 will follow if we prove Theorem 2.

Section 4

Proof of Theorem 2. Since $\dim M \leqq \dim_\lambda M$, the following result is a consequence of Theorem 2.1 of [2].

Lemma 3. *For* $\delta \geqq 0$,

$$\dim \left[\omega : \liminf_{n\to\infty} \frac{-n^{-1}\log v(u_n(\omega))}{-n^{-1}\log \lambda(u_n(\omega))} \leqq \delta\right] \leqq \delta. \tag{4.1}$$

The proof of (1.6) is based on this lemma. To find an upper bound for the limit inferior in (4.1), we require an upper bound for the numerator of the ratio and a lower bound for the denominator; these we shall find separately.

We first show that

$$\ell(v) > \tfrac{1}{6} \tag{4.2}$$

and

$$-2\int \log\left(\frac{1|}{|a_1(\omega)} + \cdots + \frac{1|}{|a_k(\omega)}\right) v(d\omega) > \tfrac{1}{6}, \qquad k \geqq 3, \tag{4.3}$$

for stationary v. If $v_1(1)=p$, then $v_3(1,1,1) \geqq 1-3(1-p)$. The quantities in question are at least

$$-2\big(1-v_1(1)\big)\log\frac{1|}{|2} = 2(1-p)\log 2,$$

which exceeds $\frac{1}{6}$ for $0 \leqq p \leqq \frac{3}{4}$; they are also at least

$$-2v_3(1,1,1)\log\left(\frac{1|}{|1} + \frac{1|}{|1} + \frac{1|}{|1}\right) \geqq 2(3p-2)\log\tfrac{3}{2},$$

which exceeds $\frac{1}{6}$ for $\frac{3}{4} \leqq p \leqq 1$.

From now on, r is a fixed positive integer, each ω considered lies in Ω_r, and the indices $i_1, i_2, \ldots$ range from 1 to r. We shall bound the numerator in (4.1) for probability measures v corresponding to higher-order Markov chains. As in Section 1, let C_k be the set of r^k-dimensional vectors p with nonnegative components satisfying (1.8) and (1.9). Let C_k^0 be the set of p in C_k with strictly positive components.

For p in C_k put

$$s_p(i_1, \ldots, i_{k-1}) = \sum_i p(i_1, \ldots, i_{k-1}, i) \tag{4.4}$$

and define

$$t_p(i_1, \ldots, i_{k-1}; i_k) = \begin{cases} \dfrac{p(i_1, \ldots, i_k)}{s_p(i_1, \ldots, i_{k-1})} & \text{if } s_p(i_1, \ldots, i_{k-1}) > 0 \\ \delta_{i_{k-1}, i_k} & \text{otherwise.} \end{cases} \tag{4.5}$$

Now (4.5) is a set of transition probabilites for a Markov chain of order $k-1$ with states $1, \ldots, r$. Because of (1.8) and (1.9), (4.4) are stationary probabilites for the chain. There is a probability measure on Ω_r under which $[a_1(\omega), a_2(\omega), \ldots]$ is a stationary chain with these stationary and transition probabilities; call this measure ν_p.

For p in C_k, define (here and later $\sum$ extends over all k-tuples from $[1, \ldots, r]$)

$$h_k(p) = -\sum p(i_1, \ldots, i_k) \log t_p(i_1, \ldots, i_{k-1}; i_k) \tag{4.6}$$

(with $0 \log 0 = 0$). The entropy $h(\nu_p)$ of ν_p is just $h_k(p)$. Clearly $h_k(p)$ is continuous on C_k^0; that it is continuous on all of C_k follows from the fact that

$$-p(i_1, \ldots, i_k) \log t_p(i_1, \ldots, i_{k-1}; i_k)$$

vanishes for $p(i_1, \ldots, i_k) = 0$, and is at most $-p(i_1, \ldots, i_k) \log p(i_1, \ldots, i_k)$ for positive $p(i_1, \ldots, i_k)$ and hence tends to 0 with $p(i_1, \ldots, i_k)$. Also define

$$\text{\l}_k(p) = -2 \sum p(i_1, \ldots, i_k) \log \left(\frac{1|}{|i_1} + \cdots + \frac{1|}{|i_k} \right). \tag{4.7}$$

Clearly $\text{\l}_k(p)$ is continuous on C_k.

Let $N_n(i_1, \ldots, i_k; \omega)$ be the number of j, $1 \leqq j \leqq n$, for which

$$\big(a_j(\omega), \ldots, a_{j+k-1}(\omega)\big) = (i_1, \ldots, i_k). \tag{4.8}$$

If $p \in C_k^0$, then

$$\begin{aligned} -\log \nu_p\big(u_n(\omega)\big) = & -\sum N_{n-k+1}(i_1, \ldots, i_k; \omega) \log t_p(i_1, \ldots, i_{k-1}; i_k) \\ & -\log s_p\big(a_1(\omega), \ldots, a_{k-1}(\omega)\big), \end{aligned} \tag{4.9}$$

a formula useful for bounding the numerator in (4.1). Let $N_n'(i_1, \ldots, i_k; \omega)$ be the number of j, $1 \leqq j \leqq n-k+1$, for which (4.8) holds, plus the number of j, $n-k+2 \leqq j \leqq n$, for which

$$\big(a_j(\omega), \ldots, a_n(\omega), a_1(\omega), \ldots, a_{j-n-1+k}(\omega)\big) = (i_1, \ldots, i_k) \tag{4.10}$$

holds. Let $\pi_n(k, \omega)$ have components $\pi_n(i_1, \ldots, i_k; \omega) = n^{-1} N_n(i_1, \ldots, i_k; \omega)$, and let $\pi_n'(k, \omega)$ have components $\pi_n'(i_1, \ldots, i_k; \omega) = n^{-1} N_n'(i_1, \ldots, i_k; \omega)$. Then $\pi_n'(k, \omega)$ satisfies (1.9) as well as (1.8) and hence lies in C_k (which is not generally true of $\pi_n(k, \omega)$). Clearly $N_{n-k+1} \leqq N_n'$, so (4.9) implies

$$\begin{aligned} -\log \nu_p\big(u_n(\omega)\big) \leqq & -\sum \pi_n'(i_1, \ldots, i_k; \omega) \log t_p(i_1, \ldots, i_{k-1}; i_k) \\ & -\frac{1}{n} \log s_p\big(a_1(\omega), \ldots, a_{k-1}(\omega)\big). \end{aligned} \tag{4.11}$$

For p in C_k^0 and $\varepsilon>0$, let V_p^ε be the set of π in C_k for which

$$-\sum|\pi(i_1,\ldots,i_k)-p(i_1,\ldots,i_k)|\log t_p(i_1,\ldots,i_{k-1};i_k)<\varepsilon \tag{4.12}$$

and

$$-2\sum|\pi(i_1,\ldots,i_k)-p(i_1,\ldots,i_k)|\log\left(\frac{1|}{|i_1}+\cdots+\frac{1|}{|i_k}\right)<\varepsilon. \tag{4.13}$$

Note that V_p^ε is open relative to C_k. If $p\in C_k^0$, if $\pi_n'(k,\omega)\in V_p^\varepsilon$, and if n exceeds some $n_1(\varepsilon,p,\omega)$, then by (4.11),

$$-\frac{1}{n}\log v_p(u_n(\omega))\leqq h_k(p)+2\varepsilon. \tag{4.14}$$

To estimate the corresponding denominator in (4.1), note first that

$$-\log\lambda(u_n(\omega))\geqq-2\sum_{j=1}^{n}\log\left(\frac{1|}{|a_j(\omega)}+\frac{1|}{|a_{j+1}(\omega)}+\cdots\right)-8$$

by (2.2) and (2.3). Since

$$\left|\log\left(\frac{1|}{|b_1}+\frac{1|}{|b_2}+\cdots\right)-\log\left(\frac{1|}{|b_1}+\cdots+\frac{1|}{|b_k}\right)\right|\leqq\frac{4}{2^k} \tag{4.15}$$

(see [3; (4.7)], for example), we have

$$-\log\lambda(u_n(\omega))\geqq-2\sum N_n(i_1,\ldots,i_k;\omega)\log\left(\frac{1|}{|i_1}+\cdots+\frac{1|}{|i_k}\right)-8-\frac{8n}{2^k}.$$

Since $i_1\leqq r$, and since N_n and N_n' differ by less than k, this gives

$$\begin{aligned}-\frac{1}{n}\log\lambda(u_n(\omega))\geqq&-2\sum\pi_n'(i_1,\ldots,i_k;\omega)\log\left(\frac{1|}{|i_1}+\cdots\frac{1|}{|i_k}\right)\\&-\frac{2k}{n}r^k\log(r+1)-\frac{8}{n}-\frac{8}{2^k}.\end{aligned}$$

Therefore, if $p\in C_k^0$, if $\pi_n'(k,\omega)\in V_p^\varepsilon$, and if n exceeds some $n_2(\varepsilon,p)$, then

$$-\frac{1}{n}\log\lambda(u_n(\omega))\geqq t_k(p)-2\varepsilon-8\cdot2^{-k}. \tag{4.16}$$

From now on consider only $k\geqq7$ and $\varepsilon<1/24$; for such pairs $2\varepsilon+8\cdot2^{-k}<\frac{1}{6}$, and so the right side of (4.16) is positive by (4.3). Suppose p lies in C_k^0. If $\pi_n(k,\omega)$ lies in V_p^ε i.o. (that is, for infinitely many values of n), then $\pi_n'(k,\omega)$ lies in $V_p^{2\varepsilon}$ i.o., so that (4.14) and (4.16) with 2ε in place of ε are simultaneously true i.o. It follows by Lemma 3 that

$$\dim[\omega\in\Omega_r\colon\pi_n(k,\omega)\in V_p^\varepsilon\text{ i.o.}]\leqq\frac{h_k(p)+4\varepsilon}{t_k(p)-4\varepsilon-8\cdot2^{-k}}. \tag{4.17}$$

For a nonempty subset B of C_k, define

$$G(B)=\left[\omega\in\Omega_r\colon\liminf_{n\to\infty}d(\pi_n(k,\omega),B)=0\right], \tag{4.18}$$

where d denotes ordinary distance in r^k-space. Let $B^\varepsilon=[\pi\in C_k\colon d(\pi, B)<\varepsilon]$. If $\pi\in C_k$ and $\varepsilon>0$, then there exists a p in C_k^0 such that $d(\pi, p)<\varepsilon$ and $\pi\in V_p^\varepsilon$ (because the sums in (4.12) and (4.13) go to 0 as p approaches π and because C_k is convex and C_k^0 is its interior relative to the hyperplane determined by (1.8) and (1.9)). Thus the V_p^ε with p in $B^\varepsilon\cap C_k^0$ cover the closure of B, and therefore B can be covered by finitely many $V_{p_1}^\varepsilon, \ldots, V_{p_a}^\varepsilon$ with $p_a\in B^\varepsilon\cap C_k^0$. Clearly

$$G(B)\subset \bigcup_{\alpha=1}^{a} [\omega\in\Omega_r\colon \pi_n(k,\omega)\in V_{p_\alpha}^\varepsilon \text{ i.o.}].$$

Therefore $\dim G(B)$ is at most the supremum of the right side of (4.17) with p ranging over $B^\varepsilon\cap C_k^0$. Since ε is arbitrary, and since (4.6) and (4.7) are continuous in p,

$$\dim G(B)\leqq \sup_{p\in B} \frac{h_k(p)}{\ell_k(p)-8\cdot 2^{-k}}. \tag{4.19}$$

Now suppose that A is a nonempty set of r-dimensional probability vectors—a subset of C_1. With τ_k defined by (1.10), $G(\tau_k^{-1}A)$ coincides with $U_r(A)$ as defined in Section 1 (see (1.5)). Now (4.19) applied to $\tau_k^{-1}A$ gives

$$\dim U_r(A)\leqq \sup_{p\in\tau_k^{-1}A} \frac{h_k(p)}{\ell_k(p)-8\cdot 2^{-k}}. \tag{4.20}$$

For p in C_k, it follows by (1.9) that

$$\sum_i s_p(i, i_1, \ldots, i_{k-2})=\sum_i s_p(i_1, \ldots, i_{k-2}, i).$$

Call this common value $c_p(i_1, \ldots, i_{k-2})$ and define

$$p^*(i_1, \ldots, i_k)=\frac{s_p(i_1, \ldots, i_{k-1})\, s_p(i_2, \ldots, i_k)}{c_p(i_2, \ldots, i_{k-1})} \tag{4.21}$$

if the denominator is positive, and define $p^*(i_1, \ldots, i_k)=0$ otherwise. Then p^* lies in C_k and has the same stationary probabilities as p; moreover $t_{p^*}(i_1, \ldots, i_{k-1}; i_k)>0$ if $s_p(i_1, \ldots, i_{k-1})$ and $s_p(i_2, \ldots, i_k)$ are both positive. Any convex combination $xp+(1-x)p^*$, $0\leqq x<1$, has the same properties and the corresponding Markov chain is therefore ergodic. Thus any p in C_k can be approximated by elements π of C_k for which v_π is ergodic and for which the stationary probabilities are as for p. Since (4.6) and (4.7) are continuous in p, the supremum in (4.20) is unaltered if in addition to $p\in\tau_k^{-1}A$ we require that v_p be ergodic.

Now $\ell_k(p)$ differs from $\ell(v_p)$ by at most $8\cdot 2^{-k}$ because of (4.15), and $h_k(p)$ coincides with $h(v_p)$. Therefore (if $k\geqq 7$, so that $16\cdot 2^{-k}<\frac{1}{6}$)

$$\dim U_r(A)\leqq \sup \frac{h(v)}{\ell(v)-16\cdot 2^{-k}},$$

the supremum extending over all ergodic Markov chains of order $k-1$ with $v_1\in A$. Since a chain of order $k-1$ is also a chain of order k, (1.6) follows.

Section 5

Proof of Theorem 4. Let $H_\nu(a_1, \ldots, a_k)$ be the entropy of the random variables $a_1(\omega), \ldots, a_k(\omega)$ under ν. Note first that

$$\alpha_k(A) = \sup_{\substack{\nu \in \mathscr{M}_r \\ \nu_1 \in A}} \frac{H_\nu(a_1, \ldots, a_k)}{2 \int \log q_k(\omega)\, \nu(d\omega)}. \tag{5.1}$$

Indeed, if ν lies in $\mathscr{M}_r$ (or even $\mathscr{N}_r$) and $p = \nu_k$, then $p \in C_k$ and the ratios in (5.1) and (1.11) coincide. On the other hand, as the argument involving (4.21) shows, the supremum in (1.11) is unaltered if further restricted to p for which the corresponding Markov chain is ergodic; for such a p, if $\nu = \nu_p$, then $\nu \in \mathscr{M}_r$ and the ratios in (5.1) and (1.11) coincide.

By (1.2) and (2.3),

$$\left| \frac{2}{k} \int \log q_k(\omega)\, \nu(d\omega) - \ell(\nu) \right| \leqq \frac{8}{k}. \tag{5.2}$$

And (see [3; p. 82], for example)

$$\frac{1}{k} H_\nu(a_1, \ldots, a_k) \geqq h(\nu). \tag{5.3}$$

For positive ε, there is by (1.7) a ν in $\mathscr{M}_r$ with $h(\nu)/\ell(\nu) > \dim L_r(A) - \varepsilon$. Choose k_0 so that $h(\nu)/(\ell(\nu) + 8/k_0) > \dim L_r(A) - \varepsilon$. By (5.2) and (5.3), $k \geqq k_0$ implies

$$\alpha_k(A) \geqq \frac{H_\nu(a_1, \ldots, a_k)}{2 \int \log q_k(\omega)\, \nu(d\omega)} \geqq \frac{h(\nu)}{\ell(\nu) + 8/k} > \dim L_r(A) - \varepsilon.$$

Thus $\liminf_k \alpha_k(A) \geqq \dim L_r(A)$.

If $j \leqq k$ and $\nu \in \mathscr{N}_r$, then by standard properties of conditional entropy,

$$\begin{aligned} H_\nu(a_1, \ldots, a_k) &= H_\nu(a_1, \ldots, a_j) + \sum_{v=j}^{k-1} H_\nu(a_{v+1} | a_1, \ldots, a_v) \\ &\leqq \log r^j + k H_\nu(a_j | a_1, \ldots, a_{j-1}). \end{aligned}$$

If ν^j is the measure corresponding to that Markov chain of order $j-1$ whose j-dimensional distribution coincides with that of ν (that is, if $\nu_j^j = \nu_j$), then

$$H_\nu(a_1, \ldots, a_k) \leqq j \log r + k h(\nu^j), \quad j \leqq k. \tag{5.4}$$

By (4.15), $\ell(\nu)$ is within $8/2^j$ of

$$-2 \int \log \left(\frac{1}{\lfloor a_1(\omega) \rfloor} + \cdots + \frac{1}{\lfloor a_j(\omega) \rfloor} \right) \nu(d\omega), \tag{5.5}$$

and similarly for $\ell(\nu^j)$; but (5.5) is the same for ν^j as for ν, and hence $|\ell(\nu) - \ell(\nu^j)| \leqq 16/2^j$.

By this, (5.4), (5.2), and (4.2),

$$\frac{H_\nu(a_1,\ldots,a_k)}{2\int\log q_k(\omega)\,\nu(d\omega)} \leqq \frac{\frac{j}{k}\log r + h(\nu^j)}{\ell(\nu^j)-\frac{8}{k}-\frac{16}{2^j}} \leqq \frac{\frac{j}{k}\log r}{\frac{1}{6}-\frac{8}{k}-\frac{16}{2^j}} + \frac{1}{1-6\left(\frac{8}{k}+\frac{16}{2^j}\right)}\cdot\frac{h(\nu^j)}{\ell(\nu^j)}$$

for $j \leqq k$.

We next show that $\nu\in\mathcal{N}_r$ implies $\nu^j\in\mathcal{M}_r$. Consider cylinders

$$M_n = [\omega\colon a_{n+v}(\omega)=i_v,\ 0\leqq v\leqq j-2]$$

and

$$M_n' = [\omega\colon a_{n+v}(\omega)=i_v',\ 0\leqq v\leqq j-2].$$

Since ν is ergodic,

$$n^{-1}\sum_{k=1}^{n}\nu(M_1\cap M_k')\to\nu(M_1)\,\nu(M_1')=\nu^j(M_1)\,\nu^j(M_1').$$

Thus $\nu(M_1\cap M_n')$ is positive for some n if $\nu^j(M_1)$ and $\nu^j(M_1')$ are both positive; but then $\nu^j(M_1\cap M_n')$ is also positive for some n, because $\nu(M)>0$ implies $\nu^j(M)>0$ for cylinders M. Therefore, if two sequences (of length $j-1$) of states are possible under ν^j, it is also possible to pass from one to the other; hence ν^j is ergodic.

It therefore follows from (1.7) that we may replace $h(\nu^j)/\ell(\nu^j)$ by $\dim L_r(A)$ in the preceding inequality if $\nu\in\mathcal{N}_r$. Applying (5.1) to the left side of the inequality, letting $k\to\infty$, and then letting $j\to\infty$, we obtain $\limsup_k \alpha_k(A)\leqq\dim L_r(A)$, which proves Theorem 4.

Section 6

A Conjecture. Let $\beta_k(A)$ be the supremum of the ratio in (1.11) over those p with nonnegative components satisfying (1.8) and $\tau_k(p)\in A$. In other words, drop the constraints (1.9). Then $\alpha_k(A)\leqq\beta_k(A)$, and it may be conjectured that

$$\lim_{k\to\infty}\left(\beta_k(A)-\alpha_k(A)\right)=0, \tag{6.1}$$

so that $\alpha_k(A)$ can be replaced by $\beta_k(A)$ in (1.12).

The conjecture is true for $A=C_1$: If s satisfies

$$\sum q_k(i_1,\ldots,i_k)^{-2s}=1 \tag{6.2}$$

($i_1,\ldots,i_k$ ranging from i to r as usual), and if $p(i_1,\ldots,i_k)=q_k(i_1,\ldots,i_k)^{-2s}$, then p satisfies (1.8) and a computation shows that the ratio in (1.11) has the value s. On the other hand, if s satisfies (6.2) and the $p(i_1,\ldots,i_k)$ are nonnegative and add to 1, then

$$\sum p(i_1, \ldots, i_k) \log p(i_1, \ldots, i_k)^{-1} q_k(i_1, \ldots, i_k)^{-2s}$$
$$\leqq \log \sum q_k(i_1, \ldots, i_k)^{-2s} = 0$$

by convexity, so the ratio in (1.11) is at most s. Thus $\beta_k(C_1)$ is the root of the Eq. (6.2). Good [5] has shown that this root converges to $\dim \Omega_r = \dim U_r(C_1)$, so (6.1) does hold for $A = C_1$. It would be interesting to have a simple, direct proof of (6.1), at least for this special case.

References

1. Billingsley, P.: Hausdorff dimension in probability theory. Illinois. J. Math. **4**, 187–209 (1960)
2. Billingsley, P.: Hausdorff dimension in probability theory II. Illinois. J. Math. **5**, 291–298 (1961)
3. Billingsley, P.: Ergodic theory and information. New York: Wiley 1965
4. Chung, K. L.: A note on the ergodic theorem of information theory. Ann. Math. Statist. **32**, 612–614 (1961)
5. Good, I. J.: The fractional dimensional theory of continued fractions. Proc. Cambridge Philos. Soc. **37**, 199–228 (1941)
6. Kinney, John R., Pitcher, Tom S.: The dimension of some sets defined in terms of f-expansions. Z. Wahrscheinlichkeitstheorie verw. Gebiete **4**, 293–315 (1966)
7. Schweiger, F., Stradner, W.: Die Billingsleydimension von Mengen mit vorgeschriebenen Ziffern I. Österreich. Akad. Wiss. Math.-Natur. Kl. S.-B. II. **180**, 95–109 (1972)
8. Schweiger, F.: Die Billingsleydimension von Mengen mit vorgeschriebenen Ziffern II. Österreich. Akad. Wiss. Math.-Natur. Kl. S.-B. II. **181**, 151–155 (1973)

Patrick Billingsley
Department of Statistics
University of Chicago
Chicago, Illinois 60637
U.S.A.

Inge Henningsen
Institute of Mathematical Statistics
University of Copenhagen
Copenhagen
Denmark

(Received April 4, 1974)