The Annals of Mathematical Statistics
1970, Vol. 41, No. 5, 1472–1490

AN INVERSION ALGORITHM FOR ONE-DIMENSIONAL *F*-EXPANSIONS[1]

By Scott Bates Guthery

Michigan State University[2]

If f is a monotone function subject to certain restrictions and φ its inverse, then one can associate with any x, a real number between zero and one, a sequence $\{a_n\}$ of integers such that

$$x = f(a_1 + f(a_2 + f(a_3 + f(a_4 + \cdots.$$

If T is the transformation $\langle \varphi(x) \rangle$ where $\langle\ \rangle$ stands for the fractional part, it has been shown that there is a unique measure μ invariant under T which is absolutely continuous with respect to Lebesgue measure. Examples are $f(x) = x/10$ which gives rise to the decimal expansion with invariant measure Lebesgue measure, or $f(x) = 1/x$ which gives rise to the continued fraction, with measure $dx/\ln 2(1+x)$. This induces a measure P on the sequences $\{a_n\}$ which is stationary ergodic and has other interesting properties. However, a large class of pairs $\{f, \mu\}$ gives rise to the pair $\{\{a_n\}, P\}$. The paper is concerned with the problem of how, given a measure μ to find, when possible, an f, which corresponds to a pair $\{\{a_n\}, P\}$, or given an $\{f, \mu\}$ pair, to reduce it to a canonical form. Interesting observations about the "memory" of the process arise from the "canonical form".

1. Introduction. This paper examines a variety of one-dimensional f-expansions along with their invariant measures and associated stochastic processes. To introduce the material, we present the following brief summary of relevant work in the field.

1.1. Background. The classical f-expansion is the continued fraction. Beginning with $x \in (0, 1)$ and $f(x) = 1/x$ and letting [ ] denote the greatest integer function and $\langle\ \rangle$ the fractional part, we use the expansion algorithm:

$$a_1(x) = [f^{-1}(x)], \qquad r_1(x) = \langle f^{-1}(x) \rangle$$

and for $i \geqq 1$, if $r_i(x) \neq 0$, then

$$a_{i+1}(x) = [f^{-1}(r_i(x))] \quad \text{and} \quad r_{i+1}(x) = \langle f^{-1}(r_i(x)) \rangle.$$

Setting $\rho_n(x) = f(a_1(x) + f(a_2(x) + \cdots + f(a_n(x))$

$$= \cfrac{1}{a_1(x) + \cfrac{1}{a_2(x) + \cfrac{\ddots}{+ \cfrac{1}{a_n(x)}}}}$$

Received November 14, 1969.

[1] This paper is a revised version of the author's doctoral dissertation which was written at Michigan State University in 1969 under the direction of Professor John R. Kinney.

[2] The author is currently with Bell Telephone Laboratories, Holmdel, New Jersey.

we have $x = \rho_n(x)$ if $r_n(x) = 0$ and otherwise $x = \lim_{n\to\infty} \rho_n(x)$. Properties of this expansion have been studied extensively and an excellent survey is provided by Khinchin [6].

Let us set

$$(0,1)_f = \{x \mid r_n(x) \neq 0 \text{ for all } n\}$$

and note that since we have excluded only a countable number of elements of (0. 1), we have

$$\lambda((0,1)_f) = 1$$

where λ is Lebesgue measure. Any non-atomic probability measure on (0, 1) induces, in the obvious way, a probability measure on $(0, 1)_f$. The underlying σ-field is assumed to be the Borel field $\mathscr{B}$ and almost everywhere (a.e.) statements are made relative to λ.

In 1951, Ryll–Nardzewski [10] considered the transformation $T(x) = \langle 1/x \rangle$ on (0, 1) and found that the measure ω on (0, 1) defined by

$$\frac{d\omega}{d\lambda} = \frac{1}{\log 2}\frac{1}{(x+1)}$$

was preserved by T and that T was ergodic with respect to ω. By noting that for a.e. $x \in (0, 1)$

$$T^n(x) = r_n(x),$$

where we let $r_0(x) = x$, and hence

$$a_{n+1}(x) = [1/T^n(x)],$$

he was able to deduce many of the measure theoretic properties of the continued fraction expansion through applications of the individual ergodic theorem.

Then, in 1957, Rényi [8] extended this result to the work of Everett [3] and Bissinger [1] who had investigated the use of an arbitrary monotone function in the expansion algorithm and conditions under which $x = \lim_{n\to\infty} \rho_n(x)$. Such functions were said to be valid for f-expansions.

Citing the following conditions on f:

(A1) $f(1) = 1$;

(A2) $f(t)$ is nonnegative, continuous, and strictly decreasing for $1 \leqq t \leqq N$ and $f(t) = 0$ for $t \geqq N$ where $N > 2$ is an integer or $+\infty$;

(A3) $$|f(t_2)-f(t_1)| \leqq |t_2-t_1| \quad \text{for} \quad 1 \leqq t_1 < t_2 \qquad \text{and}$$
$$|f(t_2)-f(t_1)| < |t_2-t_1| \quad \text{if} \quad \tau-\varepsilon < t_1 < t_2$$

where τ is the solution of the equation $1+f(\tau) = \tau$ and $0 < \varepsilon < \tau$ is arbitrary;

(B1) $f(0) = 0$;

(B2) $f(t)$ is nonnegative, continuous, and strictly increasing for $0 \leqq t \leqq N$ and $f(t) = 1$ for $t \geqq N$ where $N > 1$ is an integer or $+\infty$;

(B3) $|f(t_2)-f(t_1)| < |t_2-t_1|$ for $0 \leqq t_1 < t_2$; and

(C) if $H_n(x, t) = d/dt\, f(a_1(x)+f(a_2(x)+ \cdots +f(a_n(x)+t)))$ then

$$\frac{\sup_{0<t<1} H_n(x, t)}{\inf_{0<t<1} H_n(x, t)} \leqq C < +\infty$$

where the constant $C \geqq 1$ depends neither on x nor on n; Rényi proved

THEOREM 1.1. *If f satisfies conditions* A *or* B *then f is valid for f-expansions. If f further satisfies condition* C, *then there exists a unique probability measure ω on* $(0, 1)$ *such that*:

(i) *ω is equivalent to λ*;
(ii) *ω is preserved by $T(x) = \langle f^{-1}(x)\rangle$*;
(iii) *T is ergodic with respect to ω; and*
(iv) $C^{-1} \leqq d\omega/d\lambda \leqq C$.

This theorem defines an entire class of functions whose measure theoretic f-expansion properties can be investigated using the individual ergodic theorem. However, since it utilizes a non-constructive proof, it leaves open the problem of finding the measure ω for each function in the class. This problem has been solved in only a very few cases and is the primary impetus behind the present work.

Next, in 1960, Rokhlin [9] obtained an approximate rate of convergence for the f-expansion of numbers using the functions and measure described by Rényi. Defining $\varphi = f^{-1}$ and

$$B_n(x) = \{y \mid a_i(y) = a_i(x), i = 1(1)n\}$$

he proved

THEOREM 1.2. *If f satisfies* A *and* C *or* B *and* C *and* $\log|\varphi'|$ *is Lebesgue integrable on* $(0, 1)$, *then*

$$h(T) = -\lim_{n\to\infty} \frac{\log \omega(B_n(x))}{n} = -\lim_{n\to\infty} \frac{\log \lambda(B_n(x))}{n} = \int_0^1 \log|\varphi'(t)|\, d\omega(t) \quad \text{a.e.}$$

The number $h(T)$ is called the entropy of the transformation T.

Finally, in 1966, Kinney and Pitcher [7] considered the discrete stochastic process $[a_i, \nu, (0, 1)_f]$ associated with an f-expansion formed by the coefficients (a_i) of an f-expansion and a measure ν on $(0, 1)$. Using this construct, they were able to calculate the dimension of some sets defined in terms of f-expansions and connect certain properties of the processes with properties of the f-expansions.

1.2. *Terminology.* Suppose we consider the following conditions on a function f:

(A′) $f(1) = 1$; $f(t)$ is nonnegative, continuous and non-increasing for $1 \leqq t \leqq N$; and $f(t) = 0$ for $t \geqq N$ where $N > 2$ is an integer or $+\infty$;

(B′) $f(0) = 0$; $f(t)$ is nonnegative, continuous, and non-decreasing for $0 \leqq t \leqq N$; and $f(t) = 1$ for $t \geqq N$ where $N > 1$ is an integer or $+\infty$.

If f satisfies A′ let us define

$$f^{-1}(x) = \text{g.l.b.}\, \{t \mid f(t) \leqq x\}$$

and if f satisfies B′ let us define

$$f^{-1}(x) = \text{g.l.b.}\,\{t \mid f(t) \geqq x\}$$

for all $x \in (0, 1)$.

Now, if f satisfies A′ or B′, ω is a λ-equivalent measure on (0, 1) and the transformation $T = \langle f^{-1} \rangle$ is an endomorphism on $((0, 1), \mathscr{B}, \omega)$; i.e., T is measurable and $\omega(T^{-1}B) = \omega(B)$ for all $B \in \mathscr{B}$; then we shall call the pair $(f, d\omega/d\lambda)$ an expansion pair. The measure ω will be said to be invariant with respect to or preserved by f. If an expansion pair (f, h) is such that f is valid for f-expansions, the pair is called a valid expansion pair. Similarly, if T is an ergodic endomorphism the pair is called an ergodic expansion pair. Using this terminology Rényi's theorem states that if f satisfies A and C or B and C then there exists a unique λ-equivalent probability measure ω such that $C^{-1} \leqq d\omega/d\lambda \leqq C$ and $(f, d\omega/d\lambda)$ is a valid, ergodic expansion pair.

2. The inversion algorithm. The inversion algorithm given below can produce expansion pairs from a summation representation of the Radon–Nikodym derivative of a λ-equivalent measure. Conditions are also given on the representation which insure that the resultant expansion pairs are valid or ergodic.

2.1. *Definitions and basic relations.* Let g be an a.e. nonnegative Lebesgue integrable function on $[0, N)$ where N is an integer $\geqq 2$ or $+\infty$. Set $G(x) = \int_0^x g(t)\,dt$ for $x \in [0, N)$ and assume $\lim_{x \to N} G(x) = 1$. Then G is an a.e. differentiable non-decreasing function from $[0, N)$ onto $[0, 1)$.

Next set $h(x) = \sum_{k=0}^{N-1} g(x+k)$ for $x \in (0, 1)$ and assume that h is positive over its domain of definition. Since $\int_0^1 h(t)\,dt = \int_0^N g(t)\,dt = 1$, h is a probability density which determines a λ-equivalent probability measure ω on (0, 1).

Now set $H(x) = \int_0^x h(t)\,dt$ for $x \in [0, 1]$ and note that H and H^{-1} are one-to-one strictly increasing a.e. differentiable transformations on [0, 1]. Finally, define

$$f_U(x) = H^{-1}(G(x)) \quad \text{for} \quad x \in [0, N) \qquad \text{and}$$

$$f_D(x) = H^{-1}(1 - G(x-1) \quad \text{for} \quad x \in [1, N+1).$$

We see immediately that f_D is a continuous a.e. differentiable non-increasing function on $[1, N+1)$ such that $f_D(1) = 1$ and $\lim_{x \to N+1} f_D(x) = 0$. Similarly, f_U is a continuous a.e. differentiable non-decreasing function on $[0, N)$ such that $f_U(0) = 0$ and $\lim_{x \to N} f_U(x) = 1$. Therefore, if we set $f_D(x) = 0$ for $x \geqq N+1$ and $f_U(x) = 1$ for $x \geqq N$, f_D and f_U satisfy A′ and B′ respectively.

In the following, let

$$\varphi_D(x) = f_D{}^{-1}(x), \qquad \varphi_U(x) = f_U{}^{-1}(x),$$

$$T_D(x) = \langle \varphi_D(x) \rangle, \qquad T_U(x) = \langle \varphi_U(x) \rangle$$

and $R(x) = H^{-1}(1 - H(x))$.

Note that φ_D, φ_U, T_D, and T_U are a.e. differentiable functions on [0, 1] and that R is a strictly decreasing a.e. differentiable function on [0, 1]. Beforeproceeding to discuss the expansion properties of f_U and f_D, we present the following lemma concerning elementary relations between them.

LEMMA 2.1. *The following relations hold*:

2.1.1. $f_U(x) = R(f_D(x+1)) \quad f_D(x) = R(f_U(x-1))$

2.1.2. $\varphi_U(x) = \varphi_D(R(x))-1 \quad \varphi_D(x) = \varphi_U(R(x))+1$

2.1.3. $T_U(x) = T_D(R(x)) \quad T_D(x) = T_U(R(x))$

2.1.4. $f_U'(x) = g(x)/h(f(x))$ a.e. $\quad f_D'(x) = -g(x-1)/h(f_D(x))$ a.e.

2.1.5. $\varphi_U'(x) = \varphi_D'(R(x))R'(x)$ a.e. $\quad \varphi_D'(x) = \varphi_D'(R(x))R'(x)$ a.e.

2.1.6. $R(x) = R^{-1}(x)$

2.1.7. $R'(x) = -h(x)/h(R(x))$ a.e.

PROOF. 2.1.1 and 2.1.6 follow directly from the definitions of f_U, f_D, and R. For 2.1.2 we have

$$\begin{aligned}\varphi_U(x) &= \text{g.l.b.}\,\{t\,|\,f_U(t) \geqq x\}\\ &= \text{g.l.b.}\,\{t\,|\,R(f_D(t+1)) \geqq x\}\\ &= \text{g.l.b.}\,\{t\,|\,f_D(t+1) \geqq R(x)\}\\ &= \text{g.l.b.}\,\{t-1\,|\,f_D(t) \geqq R(x)\}\\ &= \text{g.l.b.}\,\{t\,|\,f_D(t) \geqq R(x)\}-1\\ &= \varphi_D(R(x))-1\end{aligned}$$

and similarly for $\varphi_D(x)$. 2.1.3 follows directly from 2.1.2 since

$$\begin{aligned}T_U(x) &= \langle \varphi_U(x)\rangle = \langle \varphi_D(R(x))-1\rangle\\ &= \langle \varphi_D(R(x))\rangle = T_D(R(x))\end{aligned}$$

and similarly for $T_D(x)$. For 2.1.4 we use the differential form

$$df^{-1}(u) = du\bigg/\left(\frac{df}{du}(f^{-1}(u))\right)$$

to obtain

$$\begin{aligned}f_U'(x) = dH^{-1}(G(x)) &= g(x)\bigg/\left(\frac{dH}{dx}(H^{-1}(G(x))\right)\\ &= g(x)/h(f_U(x)) \quad \text{a.e.}\end{aligned}$$

and

$$\begin{aligned}f_D'(x) = dH^{-1}(1-G(x-1)) &= -g(x-1)\bigg/\left(\frac{dH}{dx}(H^{-1}(1-G(x-1)))\right)\\ &= -g(x-1)/h(f_D(x)) \quad \text{a.e.}\end{aligned}$$

2.1.5 is simply an application of the chain rule to 2.1.2 with the proviso that the equality holds only where both derivatives exist. Finally, 2.1.7 follows again from the above mentioned differential form since $R'(x) = dH^{-1}(1-H(x)) = -h(x)/h(R(x))$ a.e.

The validity of the inversion algorithm is shown by

THEOREM 2.2. *(f_U, h) and (f_D, h) are expansion pairs.*

PROOF. Since the inverse image of any interval is at most a countable union of intervals under either transformation, each is measurable and it is sufficient to prove that $\omega(T^{-1}(0, \alpha)) = \omega((0, \alpha))$ for $\alpha \in (0, 1)$. If we let $f = f_U$, then we have

$$\begin{aligned}\omega(T_U{}^{-1}(0,\alpha)) &= \textstyle\sum_{k=0}^{N-1} \int_{f(k)}^{f(k+\alpha)} h(t)\,dt = \sum_{k=0}^{N-1} H(f(k+\alpha)) - H(f(k)) \\ &= \textstyle\sum_{k=0}^{N-1} \left(\int_0^{k+\alpha} g(t)\,dt - \int_0^k g(t)\,dt\right) = \sum_{k=0}^{N-1} \int_k^{k+\alpha} g(t)\,dt \\ &= \textstyle\sum_{k=0}^{N-1} \int_0^{\alpha} g(t+k)\,dt = \int_0^{\alpha} \sum_{k=0}^{N-1} g(t+k)\,dt \\ &= \textstyle\int_0^{\alpha} h(t)\,dt = \omega((0,\alpha))\end{aligned}$$

so that T_U preserves ω. The proof for $f = f_D$ follows in exactly the same way.

2.2. *Conditions for valid and ergodic expansion pairs.* If we now consider the following condition on the function g:

(D1) $g(x) > 0$ a.e. and

(D2) $g(t) < \inf_{0 \leq x \leq 1} \sum_{k=0}^{N-1} g(x+k)$ for all $t \in (0, N)$; we have

THEOREM 2.3. *If g satisfies conditions* D *then (f_U, h) and (f_D, h) are valid expansion pairs.*

PROOF. Clearly f_D satisfies A1, f_U satisfies B1, and D1 implies A2 and B2 respectively. Since $|f_U'(x)| = g(x)/h(f_U(x))$ a.e. and $|f_D'(x)| = g(x-1)/h(f_D(x))$ a.e. condition D2 guarantees that $|f_U'(x)| < 1$ a.e. for $x \in (0, N)$ and $|f_D'(x)| < 1$ a.e. for $x \in (1, N+1)$. Therefore, by the mean value theorem, f_D satisfies A3 and f_U satisfies B3. Since f_D meets conditions A and f_U meets conditions B, by Theorem 1.1 both are valid for f-expansions.

To show that the pair (f_U, h) and (f_D, h) are ergodic expansion pairs, we can either show that f_U and f_D satisfy Rényi's condition C or demonstrate directly that T_U and T_D are ergodic endomorphisms. The first method is, in general, very difficult but the following lemma can be of help in some special cases.

LEMMA 2.4. *If a function f on* $[0, N)$ *satisfies*

(i) $0 < \varepsilon_1 \leq |f'(x)| \leq \varepsilon_2 < 1$ *for* $x \in [0, N)$ *and*

(ii) f' *Lipschitz of order* 1

then f satisfies condition C.

PROOF. If $0 \leqq t_1 < t_2 \leqq N$, then from (i) we have that

$$|f(t_2)-f(t_1)| \leqq \varepsilon_2 |t_2-t_1|$$

and from (ii) we have that

$$\sup_{t_1<t<t_2} f'(t) - \inf_{t_1<t<t_2} f'(t) \leqq M(t_2-t_1)$$

where M is a constant independent of t_1 and t_2. Now,

$$\begin{aligned}
&\frac{\sup_{0<t<1} f'(a_1+f(a_2+\cdots+f(a_n+t)))}{\inf_{0<t<1} f'(a_1+f(a_2+\cdots+f(a_n+t)))}-1 \\
&= \frac{\sup_{0<t<1} f'(a_1+f(a_2+\cdots+f(a_n+t))) - \inf_{0<t<1} f'(a_1+f(a_2+\cdots+f(a_n+t)))}{\inf_{0<t<1} f'(a_1+f(a_2+\cdots+f(a_n+t)))} \\
&\leqq \frac{M\,|f(a_2+f(a_3+\cdots+f(a_n+1)))-f(a_2+f(a_3+\cdots+f(a_n)))|}{\varepsilon_1} \\
&\leqq \frac{M\varepsilon_2\,|f(a_3+f(a_4+\cdots+f(a_n+1)))-f(a_3+f(a_4+\cdots+f(a_n)))|}{\varepsilon_1} \\
&\leqq \cdots \leqq \frac{M\varepsilon_2^{\,n-1}}{\varepsilon_1};
\end{aligned}$$

therefore,

$$\begin{aligned}
\frac{\sup_{0<t<1} H_n(x,t)}{\inf_{0<t<1} H_n(x,t)} &= \frac{\sup_{0<t<1} (d/dt) f(a_1(x)+f(a_2(x)+\cdots+f(a_n(x)+t)))}{\inf_{0<t<1} (d/dt) f(a_1(x)+f(a_2(x)+\cdots+f(a_n(x)+t)))} \\
&\leqq \prod_{j=1}^{n} \frac{\sup_{0<t<1} f'(a_j(x)+f(a_{j+1}(x)+\cdots+f(a_n(x)+t)))}{\inf_{0<t<1} f(a_j(x)+f(a_{j+1}(x)+\cdots+f(a_n(x)+t)))} \\
&\leqq \prod_{j=1}^{n} \left(1+\frac{M\varepsilon_2^{\,j-1}}{\varepsilon_1}\right) \\
&\leqq \prod_{j=1}^{\infty} \left(1+\frac{M\varepsilon_2^{\,j-1}}{\varepsilon_1}\right) = C < \infty,
\end{aligned}$$

since, by Theorem 8.6.1 of Hille [4], the infinite product converges if $\sum_{j=1}^{\infty} M\varepsilon_2^{\,j-1}/\varepsilon_1$ converges, which it obviously does.

By noting that $f_U'(x) = g(x)$ a.e. and $f_D'(x) = -g(x-1)$ a.e. when ω is Lebesgue measure, we see that the conditions of this lemma are reduced to conditions on the function g.

2.3. *Rokhlin's formula.* We conclude this section with

THEOREM 2.5. *If f_D and f_U satisfy* A *and* C *and* B *and* C *respectively, then* $h(T_U) = h(T_D)$.

PROOF.

$$\begin{aligned}
h(T_U) &= \textstyle\int_0^1 \log|\varphi_U'(x)|\, h(x)\,dx = \int_0^1 \log|\varphi_D'(R(x))R'(x)|\, h(x)\,dx \\
&= \textstyle\int_1^0 \log|\varphi_D'(x)R'(R(x))|\, h(R(x))R'(x)\,dx \\
&= \textstyle\int_0^1 \log|\varphi_D'(x)|\, h(x)\,dx + \int_0^1 \log|R'R(x)|\, h(x)\,dx \\
&= h(T_D) + \int_0^1 \log\left|\frac{h(R(x))}{h(x)}\right| h(x)\,dx \\
&= h(T_D) + \textstyle\int_0^1 \log|h(R(x))|\, h(x)\,dx - \int_0^1 \log|h(x)|\, h(x)\,dx \\
&= h(T_D) + \textstyle\int_1^0 \log|h(x)| h(R(x))R'(x)\,dx - \int_0^1 \log|h(x)|\, h(x)\,dx \\
&= h(T_D).
\end{aligned}$$

3. Examples. In this section we present examples of the use of the inversion algorithm which include some known expansion pairs along with some new ones.

3.1. *Lebesgue measure.* Perhaps the easiest and most interesting measure to invert is Lebesgue measure which has density function $h(x) = 1$. We shall use the notation $i = j(k)l$ to denote $i = j, j+k, j+2k, \cdots, j+mk$, where $j+mk \leqq l < j+(m+1)k$. Suppose, for example, we have nonnegative constants p_k, $k = 0(1)N-1$, such that $\sum_{k=0}^{N-1} p_k = 1$ and we set $g(x) = p_k$ for $k \leqq x < k+1$. Then, since $H(x) = H^{-1}(x) = x$, we have

$$f_U(x) = G(x) = \textstyle\int_0^x g(t)\,dt = \sum_{k=0}^{[x]-1} p_k + \langle x\rangle p_{[x]} \quad \text{and}$$
$$f_D(x) = 1 - G(x-1) = 1 - \textstyle\sum_{k=0}^{[x]-2} p_k - \langle x\rangle p_{[x]-1}.$$

A well-known special case of this expansion is obtained by setting $p_k = 1/M$ for $k = 0(1)M-1$ to get

$$f_U(x) = \frac{[x]}{M} + \frac{\langle x\rangle}{M} = \frac{x}{M} \quad \text{and}$$

$$f_D(x) = 1 - \frac{[x]-1}{M} + \frac{\langle x\rangle}{M} = 1 - \frac{x-1}{M}.$$

These are called the M-adic expansions since they yield the expansion of numbers base M.

Suppose now we insist that $0 < \varepsilon_1 \leqq p_k \leqq \varepsilon_2 < 1$. Since $f_U'(x) = g(x)$ a.e. and $f_D'(x) = -g(x-1)$ a.e., it is easily seen that g satisfies condition D and f satisfies condition C by Lemma 2.1. Therefore, f_D and f_U satisfy conditions A and B and A and C respectively. Letting $S_n = \sum_{k=0}^n p_k$ and $S_{-1} = 0$, we can compute the entropy of T_D and T_U by Rokhlin's formula as follows:

$$\begin{aligned}
h(T_D) = h(T_U) = \textstyle\int_0^1 \log|\varphi_U^{-1}(x)|\,dx &= \textstyle\sum_{n=0}^{N-1} \int_{S_{n-1}}^{S_n} \log(1/p_n)\,dx \\
&= \textstyle\sum_{n=0}^{N-1} (-\log p_n)(S_n - S_{n-1}) \\
&= -\textstyle\sum_{n=0}^{N-1} p_n \log p_n.
\end{aligned}$$

3.2. *Generalizations of the continued fraction.* Another interesting family of f-expansions is provided by a special case of a summation theorem involving the psi function. Suppose b_i, $i = 1(1)n$, are distinct constants not less than 1 and

$$U_n(x) = \frac{p(x+n)}{\prod_{i=1}^{m}(x+n+b_i)}$$

where $m \geqq 2$ and $p(x)$ is a polynomial of degree $m-2$ or less. By the partial fraction theorem, we may write $U_n(x)$ as

$$U_n(x) = \sum_{i=1}^{m} \frac{a_i}{x+n+b_i}$$

where $\sum a_i = 0$. Then by a theorem cited by Davis [2], page 39, (19) we have

$$\textstyle\sum_{n=0}^{\infty} U_n(x) = -\sum_{i=1}^{m} a_i \Psi(x+b_i)$$

where Ψ is the psi function defined by $\Psi(x) = (d/dx)\ln\Gamma(x)$ for $x > 0$. We now set

$$g(x) = U_{[x]}(\langle x\rangle) = \frac{p(x)}{\prod_{i=1}^{m}(x+b_i)}$$

and assume the $U_n(x)$ have been normalized so that

$$\begin{aligned}\textstyle\int_0^\infty g(t)\,dt = \sum_{n=0}^{\infty}\int_0^1 U_n(t)\,dt &= \textstyle -\sum_{i=1}^{m} a_i \int_0^1 \Psi(t+b_i)dt \\ &= \textstyle -\sum_{i=1}^{m} a_i(\ln\Gamma(1+b_i)-\ln\Gamma(b_i)) \\ &= \textstyle -\sum_{i=1}^{m} a_i \ln b_i = 1.\end{aligned}$$

Since

$$\textstyle h(x) = \sum_{n=0}^{\infty} g(x+n) = \sum_{n=0}^{\infty} U_n(x) = -\sum_{i=1}^{m} a_i \Psi(x+b_i)$$

we have

$$\begin{aligned}H(x) = \textstyle\int_0^x h(t)\,dt &= \textstyle -\sum_{i=1}^{m} a_i \int_0^x \Psi(t+b_i)\,dt \\ &= \textstyle -\sum_{i=1}^{m} a_i \ln(\Gamma(x+b_i)/\Gamma(b_i)).\end{aligned}$$

Now assume m is even, $b_i = b_{i-1}+1$ for $i = 2(2)m$, and that $p(x+n)$ has been chosen so that $a_i = -a_{i-1}$ for $i = 2(2)m$. Then setting $k = m/2$ and $c_i = a_{2i-1}$ and $d_i = b_{2i-1}$ for $i = 1(1)k$, we have

$$\begin{aligned}H(x) &= -\sum_{i=1}^{k} c_i \ln\left(\frac{\Gamma(x+d_i)}{\Gamma(d_i)}\frac{\Gamma(d_i+1)}{\Gamma(x+d_i+1)}\right) \\ &= -\sum_{i=1}^{k} c_i \ln\left(\frac{d_i}{x+d_i}\right).\end{aligned}$$

If $k = 1$, $b = b_1$, and $B = \ln(1+b^{-1})$ then

$$H^{-1}(x) = b\exp(Bx-1),$$

and since

$$g(x) = [B(x+b)(x+b+1)]^{-1},$$

we have

$$G(x) = 1 + B^{-1} \ln\left(\frac{x+b}{x+b+1}\right).$$

Therefore, the two functions

$$f_U(x) = H^{-1}(G(x)) = \frac{x}{x+b+1}$$

and

$$f_D(x) = H^{-1}(1-G(x-1)) = \frac{b}{x+b-1}$$

form expansion pairs with the density function

$$h(x) = \frac{1}{B(x+b)}.$$

Note that when $b = 1$ f_D yields the continued fraction expansion.

If $k = 2$, $c_1 = -c_2$, and $B = \ln(b_3(b_1+1)/(b_1(b_3+1))$ then

$$H^{-1}(x) = \frac{b_1 b_3 (1-e^{Bx})}{b_1 e^{Bx} - b_3},$$

and since

$$g(x) = \frac{(b_3-b_1)(2x+b_3+b_1+1)}{B(x+b_1)(x+b_1+1)(x+b_3)(x+b_3+1)}$$

$$= \frac{1}{B}\left(\frac{1}{x+b_1} - \frac{1}{x+b_1+1} - \frac{1}{x+b_3} + \frac{1}{x+b_3+1}\right)$$

we have

$$G(x) = \int_0^x g(x)\,dt = 1 + B^{-1} \ln\left(\frac{(x+b_1)(x+b_3+1)}{(x+b_3)(x+b_1+1)}\right).$$

Therefore the two functions

$$f_U(x) = \frac{b_1 b_3(b_3-b_1) + b_1(x+b_3)(x+b_1+1) - b_3(x+b_1)(x+b_3+1)}{(b_1-b_3) + b_1(x+b_1)(x+b_3+1) - b_3(x+b_3)(x+b_1+1)}$$

and

$$f_D(x) = \frac{b_1 b_3(b_1-b_3)}{b_1(x+b_3-1)(x+b_1) - b_3(x+b_1-1)(x+b_3)}$$

form expansion pairs with the density function

$$h(x) = \frac{b_3 - b_1}{B(x+b_1)(x+b_3)}.$$

Finally, consider the case $k = 2$, $c_1 = c_2$, and $B = \ln((b_1+1)(b_3+1)/(b_1 b_3))$. Here we have

$$H^{-1}(x) = \frac{[(b_1-b_3)^2 + 4b_1 b_3 e^{Bx}]^{\frac{1}{2}} - (b_1+b_3)}{2}$$

and since

$$g(x) = \frac{1}{B}\left(\frac{1}{x+b_1} - \frac{1}{x+b_1+1} + \frac{1}{x+b_3} - \frac{1}{x+b_3+1}\right)$$

we have

$$G(x) = 1 + B^{-1} \ln\left[\frac{(x+b_1)(x+b_3)}{(x+b_1+1)(x+b_3+1)}\right].$$

Therefore the two functions

$$f_U(x) = \tfrac{1}{2}[(b_1-b_3)^2 + 4(b_1+1)(b_3+1)(x+b_1)(x+b_3)(x+b_1+1)^{-1}(x+b_3+1)^{-1}]^{\frac{1}{2}} - \tfrac{1}{2}(b_1+b_3)$$

and

$$f_D(x) = \tfrac{1}{2}[(b_1-b_3)^2 + 4b_1 b_3(x+b_1)(x+b_3)(x+b_1+1)^{-1}(x+b_3+1)^{-1}]^{\frac{1}{2}} - \tfrac{1}{2}(b_1+b_3)$$

form expansion pairs with the density function

$$h(x) = \frac{1}{B}\left(\frac{b_1+2x+b_3}{(x+b_1)(x+b_3)}\right).$$

3.3. *Miscellaneous examples.* Using the inversion algorithm, available expansion pairs are at least as numerous as the entries in various series summation tables such as Jolly [5] or Davis [2].

For example, consider the use of the familiar exponential series

$$\frac{ae^{ax}}{e^a - 1} = \frac{a}{e^a - 1}\sum_{k=0}^{\infty}\frac{(ax)^k}{k!}$$

for $a > 0$.
Here we let

$$g(x) = \frac{(a\langle x\rangle)^{[x]}}{[x]!}\frac{a}{e^a - 1}$$

and, therefore,

$$G(x) = \frac{a}{e^a-1}\int_0^x \frac{(a\langle t\rangle)^{[t]}}{[t]!}\,dt$$
$$= \frac{1}{e^a-1}\left(\sum_{k=1}^{[x]} \frac{a^k}{k!} + \frac{(a\langle x\rangle)^{[x+1]}}{[x+1]!}\right)$$

where we ignore the summation term if $0 \leqq x < 1$. Next, since we are setting

$$h(x) = \frac{ae^{ax}}{e^a-1}$$

we have

$$H(x) = \frac{a}{e^a-1}\int_0^x e^{at}\,dt = \frac{e^{ax}-1}{e^a-1}.$$

Inverting H, we find $H^{-1}(x) = 1/a \log((e^a-1)x+1)$. Therefore, we have

$$f_U(x) = H^{-1}(G(x)) = \frac{1}{a}\log\left(\sum_{k=1}^{[x]} \frac{a^k}{k!} + \frac{(a\langle x\rangle)^{[x+1]}}{[x+1]!} + 1\right)$$

and, similarly,

$$f_D(x) = H^{-1}(1-G(x-1)) = \frac{1}{a}\log\left(e^a - \sum_{k=1}^{[x-1]} \frac{a^k}{k!} - \frac{a(a\langle x\rangle)^{[x]}}{[x]!}\right).$$

Another family of expansion pairs, which extends the above Lebesgue family, is provided by picking α_i, $i = 0(1)n$, such that $0 = \alpha_0 < \alpha_1 < \cdots < \alpha_{n-1} < \alpha_n = 1$ and setting

$$h(x) = \beta_i \quad \alpha_{i-1} \leqq x < \alpha_i, \qquad i = 1(1)n$$

such that $\sum_{i=1}^n \beta_i(\alpha_i - \alpha_{i-1}) = 1$. Then, if $p_i > 0$ for $i = 0(1)N-1$ and $\sum_{i=0}^{N-1} p_i = 1$, we set $g(x) = p_{[x]}\beta_i$ for $\alpha_{i-1} \leqq \langle x\rangle < \alpha_i$. This family inverts easily and yields monotone "broken line" functions.

4. Associated processes with finite memory. In this section we present a sufficient condition for the use of the inversion to construct an expansion pair whose associated stochastic process has the same finite dimensional distribution as a given stationary Markov process of finite multiplicity. In a special case this construction is also shown to be unique.

That there is no loss of generality in assuming the given process is stationary is shown by

THEOREM 4.1. *If $(f, d\omega/d\lambda)$ is an expansion pair, then its associated stochastic process, $[a_i, \omega, (0, 1)_f]$, is stationary.*

PROOF. It is well known that if T is an endomorphism and X is a random variable then X and $X \circ T^i$ have the same distribution for $i \geqq 0$. Therefore, letting $T = \langle f^{-1} \rangle$ and $X = [f^{-1}]$ we see that the a_i are identically distributed.

4.1. *Inversion using a Markov process.* Suppose $[x_i, P, \Omega]$ is a stationary Markov process of finite multiplicity τ and state space $S = \{0, 1, \cdots, N-1\}$ such that $P[x_j = i_j, j = 1(1)\tau] > 0$ for all $(i_1, \cdots, i_\tau) \in S^\tau$. For any $M \geqq 1$ and $(i_1, \cdots, i_M) \in S^M$ let us define

$$I_M(i_1, \cdots, i_M) = \textstyle\sum_{j=1}^{M} i_j N^{M-j} \qquad \text{and}$$

$$F(i_1, \cdots, i_M) = \textstyle\sum P[x_n = j_n, n = 1(1)M]$$

where the latter summation extends over all $(j_1, \cdots, j_M)$ for which $I_M(j_1, \cdots, j_M) \leqq I_M(i_1, \cdots, i_M)$. That is, for each M the function F is a cumulative distribution function on the lexicographically ordered M-tuples of states. For notational convenience let us further define the following "boundary conditions" on F:

$$F(i_1, \cdots, i_{M-1}, -1) = F(i_1, \cdots, i_{M-1}-1) \quad \text{and} \quad F(-1) = 0.$$

Now suppose ω is a λ-equivalent measure on (0, 1) and let $h = d\omega/d\lambda$ and $H(x) = \int_0^x h(t)\,dt$ as usual. For each $(i_1, \cdots, i_\tau) \in S^\tau$, let $J(i_1, \cdots, i_\tau)$ be the indicator function of the interval $[H^{-1}(F(i_1, \cdots, i_\tau - 1)), H^{-1}(F(i_1, \cdots, i_\tau))]$ and define

$$g(x) = \textstyle\sum_{S^\tau} P[x_1 = [x] \mid x_{j+1} = i_j, j = 1(1)\tau] J(i_1, \cdots, i_\tau)(\langle x \rangle) h(\langle x \rangle).$$

Since, for all $x \in (0, 1)$, we have

$$\begin{aligned} \textstyle\sum_{k=0}^{N-1} g(x+k) &= \textstyle\sum_{k=0}^{N-1} \sum_{S^\tau} P[x_1 = k \mid x_{j+1} = i_j, j = 1(1)\tau] J(i_1, \cdots, i_\tau)(x) h(x) \\ &= \textstyle\sum_{k=0}^{N-1} P[x_1 = k \mid x_{j+1} = i_j^*, j = 1(1)\tau] h(x) \\ &\qquad\qquad \text{for} \quad J(i_1^*, \cdots, i_\tau^*)(x) = 1 \\ &= h(x), \end{aligned}$$

g may be used in the inversion algorithm for h. Let us set $f(x) = H^{-1}(G(x))$ and denote the stochastic process associated with (f, h) by $[a_i, \omega, (0, 1)_f]$. Using this notation, we have

THEOREM 4.2. *If F, H, and f as defined above satisfy*

$$H(f(i_1 + f(i_2 + \cdots + f(i_\tau)))) = F(i_1, \cdots, i_\tau - 1)$$

for all $(i_1, \cdots, i_\tau) \in S^\tau$, then $[x_i, P, \Omega]$ and $[a_i, \omega, (0, 1)_f]$ have the same finite dimensional distributions.

PROOF. First, we have

$$\begin{aligned}
\omega[a_j = i_j, j = 1(1)\tau+1]
&= \int_{f(i_1+f(i_2+\cdots+f(i_{\tau+1})))}^{f(i_1+f(i_2+\cdots+f(i_{\tau+1}+1)))} h(t)\,dt \\
&= H(f(i_1+f(i_2+\cdots+f(i_{\tau+1}+1)))) - H(f(i_1+f(i_2+\cdots+f(i_{\tau+1})))) \\
&= G(i_1+f(i_2+\cdots+f(i_{\tau+1}+1))) - G(i_1+f(i_2+\cdots+f(i_{\tau+1}))) \\
&= \int_{i_1+f(i_2+\cdots+f(i_{\tau+1}))}^{i_1+f(i_2+\cdots+f(i_{\tau+1}+1))} g(t)\,dt \\
&= P[x_1 = i_1 \mid x_j = i_j, j = 2(1)\tau+1] \int_{f(i_2+\cdots+f(i_{\tau+1}))}^{f(i_2+\cdots+f(i_{\tau+1}+1))} h(t)\,dt \\
&= P[x_1 = i_1 \mid x_j = i_j, j = 2(1)\tau+1] \\
&\qquad \cdot (H(f(i_2+\cdots+f(i_{\tau+1}+1))) - H(f(i_2+\cdots+f(i_{\tau+1})))) \\
&= P[x_1 = i_1 \mid x_j = i_j, j = 2(1)\tau+1] P[x_j = i_{j+1}, j = 1(1)\tau] \\
&= P[x_j = i_j, j = 1(1)\tau+1].
\end{aligned}$$

Then, using induction, we assume $n \geqq \tau+2$ and $\omega[a_j = i_j, j = 1(1)n-1] = P[x_j = i_j, j = 1(1)n-1]$ for all $(i_1, \cdots, i_{n-1}) \in S^{n-1}$ and show that

$$\begin{aligned}
\omega[a_j = i_j, j = 1(1)n] &= \int_{f(i_1+f(i_2+\cdots+f(i_n)))}^{f(i_1+f(i_2+\cdots+f(i_n+1)))} h(t)\,dt \\
&= H(f(i_1+f(i_2+\cdots+f(i_n+1)))) - H(f(i_1+f(i_2+\cdots+f(i_n)))) \\
&= G(i_1+f(i_2+\cdots+f(i_n+1))) - G(i_1+f(i_2+\cdots+f(i_n))) \\
&= \int_{i_1+f(i_2+\cdots+f(i_n))}^{i_1+f(i_2+\cdots+f(i_n+1))} g(t)\,dt \\
&= P[x_1 = i_1 \mid x_j = i_j, j = 2(1)\tau+1] \omega[a_j = i_{j+1}, j = 1(1)n] \\
&= P[x_1 = i_1 \mid x_j = i_j, j = 2(1)\tau+1] P[x_j = i_j, j = 2(1)n] \\
&= P[x_j = i_j, j = 1(1)n].
\end{aligned}$$

4.2. *The uniqueness of the construction for Lebesgue measure.* In the above construction, one sees that the function g is just a "wrinkled" version of the density function h over each interval $[k, k+1)$. Furthermore, if the resultant process is to have a finite memory, the wrinkles must occur at exactly those points in the condition of Theorem 4.2. It has been conjectured that this fixed wrinkling is also necessary for the resultant process to have a finite memory. That is, if an associated stochastic process has finite memory, then the derivative of the function with which the process is associated is a fixed wrinkling of the density function of the process. That this is indeed the case when $h(x) = 1$ is shown by

THEOREM 4.3. *If $[a_i, \lambda, (0, 1)_f]$ is a stochastic process of multiplicity τ associated with a valid expansion pair $(f, 1)$, then there exists a real-valued function C on $S^{\tau+1}$ such that*

$$f'(x) = C(i_1, \cdots, i_{\tau+1})$$

for a.e. x *in* $[i_1+f(i_2+f(i_3+\cdots+f(i_{\tau+1}))), i_1+f(i_2+f(i_3+\cdots+f(i_{\tau+1}+1))))$.

PROOF. For $n \geqq 1$ and $(i_1, \cdots, i_n) \in S^n$, let us set

$$M(i_1, \cdots, i_n) = P[x_n = i_n \mid x_j = i_j, j = 1(1)n-1]$$
$$= \frac{f(i_1 + f(i_2 + \cdots + f(i_n + 1))) - f(i_1 + f(i_2 + \cdots + f(i_n)))}{f(i_1 + f(i_2 + \cdots + f(i_{n-1} + 1))) - f(i_1 + f(i_2 + \cdots + f(i_{n-1})))}$$

and

$$D(i_1, \cdots, i_n) = \frac{f(i_1 + f(i_2 + \cdots + f(i_n + 1))) - f(i_1 + f(i_2 + \cdots + f(i_n)))}{f(i_2 + f(i_3 + \cdots + f(i_n + 1))) - f(i_1 + f(i_3 + \cdots + f(i_n)))}$$

Noting that

$$D(i_1, \cdots, i_n) = \frac{M(i_1, \cdots, i_n)}{M(i_2, \cdots, i_n)} D(i_1, \cdots, i_{n-1})$$

we have by recursion that

$$D(i_1, \cdots, i_n) = D(i_1) \prod_{j=2}^{n} \frac{M(i_1, \cdots, i_j)}{M(i_2, \cdots, i_j)}.$$

Further, since $[a_i, \lambda, (0, 1)_f]$ is stationary of multiplicity τ, we know that

$$M(i_1, \cdots, i_n) = M(i_{n-\tau}, i_{n-\tau+1}, \cdots, i_n).$$

Now, since f is valid for f-expansion, we have a.e.

$$f'(x) = \lim_{n \to \infty} D([x], a_1(\langle x \rangle), \cdots, a_n(\langle x \rangle)),$$

so setting $i_1 = [x]$ and $i_j = a_{j-1}(\langle x \rangle)$ for $j = 2, 3, \cdots,$ we have

$$f'(x) = \lim_{n \to \infty} D(i_1) \prod_{j=2}^{n} \frac{M(i_1, \cdots, i_j)}{M(i_2, \cdots, i_j)}.$$

But for $j \geqq \tau + 2$, we have

$$M(i_1, \cdots, i_j) = M(i_2, \cdots, i_j) = M(i_{j-\tau}, i_{j-\tau+1}, \cdots, i_j),$$

so a.e.,

$$f'(x) = D(i_1) \prod_{j=2}^{\tau+1} \frac{M(i_1, \cdots, i_j)}{M(i_2, \cdots, i_j)} = C(i_1, \cdots, i_{\tau+1}).$$

5. Associated processes with infinite memory. In this final section we use the specialization of the inversion algorithm introduced in Chapter IV to construct a sequence of expansion pairs with Lebesgue measure which converges to a pair $(f, 1)$ whose associated stochastic process has the same finite dimensional distributions as a given stationary process.

5.1. *Approximating an arbitrary stationary process.* Let $[x_i, P, \Omega]$ be an arbitrary stationary stochastic process with finite state space $S = \{0, 1, \cdots, N-1\}$ such that

$P[x_j = i_j, j = 1(1)n] > 0$ for all $(i_1, \cdots, i_n) \in S^n$ and $n \geqq 1$. We shall call such processes finite positive.

For $\tau = 1, 2, \cdots$, define the sequence of measures P_τ on Ω by setting

$$P_\tau[x_j = i_j, j = 1(1)n] = P[x_j = i_j, j = 1(1)n] \quad \text{for} \quad n \leqq \tau$$

and

$$\begin{aligned} &P_\tau[x_j = i_j, j = 1(1)n] \\ &\qquad = P[x_j = i_j, j = 1(1)\tau] \textstyle\prod_{k=\tau+1}^{n} P[x_k = i_k \mid x_j = i_j, j = k-\tau(1)k-1] \end{aligned}$$

for $n > \tau$. From this definition, it is easily seen that for each τ $[x_i, P_\tau, \Omega]$ is a stationary Markov process of multiplicity at most τ. Furthermore, this sequence of processes is consistent in the sense that

$$P_\tau[x_j = i_j, j = 1(1)n] = P_{\tau+1}[x_j = i_j, j = 1(1)n]$$

for all $n \leqq \tau$. Let F_τ be defined for each of these processes as in 4.1.

If we use this sequence of Markov processes in the construction of Section 4.1 and take $h(x) = 1$ we have

THEOREM 5.1. *The stochastic process associated with* $(f_\tau, 1)$, $[a_{\tau,i}, \lambda, (0, 1)_{f_\tau}]$, *has the same finite dimensional distributions as* $[x_i, P_\tau, \Omega]$.

PROOF. Using Theorem 4.2 and letting $f = f_\tau$, we need only show that

$$H(f(i_1 + f(i_2 + \cdots + f(i_\tau)))) = F_\tau(i_1, \cdots, i_\tau - 1)$$

for all $(i_1, \cdots, i_\tau) \in S^\tau$. But since $H(x) = x$, this reduces to

$$f(i_1 + f(i_2 + \cdots + f(i_\tau))) = F_\tau(i_1, \cdots, i_\tau - 1)$$

and since $f(x) = G(x)$, we have

$$\begin{aligned} &f(i_1 + f(i_2 + \cdots + f(i_\tau))) \\ &\quad = \textstyle\int_0^{i_1 + f(i_2 + \cdots + f(i_\tau))} g(t)\, dt \\ &\quad = \textstyle\int_0^{i_1 + f(i_2 + \cdots + f(i_\tau))} \sum_{S^\tau} P[x_1 = [x] \mid x_{j+1} = k_j, j = 1(1)\tau] J(k_1, \cdots, k_\tau)(\langle x \rangle) \\ &\quad = \textstyle\sum_{l=0}^{i_1-1} P[x_1 = l] \\ &\qquad + \textstyle\int_0^{f(i_2 + \cdots + f(i_\tau))} \sum_{S^\tau} P[x_1 = i_1 \mid x_{j+1} = k_j, j = 1(1)\tau] J(k_1, \cdots, k_\tau)(x) \\ &\quad = \textstyle\sum_{l=0}^{i_1-1} P[x_1 = l] \\ &\qquad + \textstyle\sum_{I_{\tau-1}(k_1, \cdots, k_{\tau-1}) < I_{\tau-1}(i_2, \cdots, i_\tau)} P[x_1 = i_1 \mid x_{j+1} = k_j, j = 1(1)\tau] \\ &\quad = F_\tau(i_1, \cdots, i_\tau - 1). \end{aligned}$$

Clearly if $[x_i, P, \Omega]$ satisfies

$$E\colon\ 0 < P[x_{n+1} = i_{n+1} \mid x_j = i_j, j = 1(1)n] \leqq \varepsilon < 1$$

for all $(i_1, \cdots, i_{n+1}) \in S^{n+1}$ and $n \geqq 1$ then $[x_i, P_\tau, \Omega]$ satisfies E and $(f_\tau, 1)$ is a valid expansion pair for all τ. Suppose, on the other hand, that $[x_i, P, \Omega]$ satisfies

$$F\colon \text{if} \qquad M_n = \sup_{S^{n+1}} P[x_{n+1} = i_{n+1} \mid x_j = i_j, j = 1(1)n]$$
$$\text{and} \qquad m_n = \inf_{S^{n+1}} P[x_{n+1} = i_{n+1} \mid x_j = i_j, j = 1(1)n]$$

then there exists a constant F such that $M_n/m_n \leqq F^{1/n}$ for all $n \geqq 1$.

THEOREM 5.2. *If* $[x_i, P, \Omega]$ *satisfies condition* F, *then* $[x_i, P_\tau, \Omega]$ *satisfies* F *and* $(f_\tau, 1)$ *is an ergodic expansion pair for all* τ.

PROOF. For fixed τ and n, we have

$$\begin{aligned} M_{\tau,n} &= \sup_{S^{n+1}} P_\tau[x_{n+1} = i_{n+1} \mid x_j = i_j, j = 1(1)n] \\ &= \sup_{S^{n+1}} \frac{P_\tau[x_j = i_j, j = 1(1)n+1]}{P_\tau[x_j = i_j, j = 1(1)n]} \\ &= \sup_{S^{n+1}} P[x_{n+1} = i_{n+1} \mid x_j = i_j, j = n+1-k(1)n] \end{aligned}$$

where

$$\begin{aligned} k &= n \qquad n \leqq \tau \\ &= \tau \qquad n > \tau. \end{aligned}$$

Therefore, if $n \leqq \tau$, then $M_{\tau,n} = M_n$ and if $n > \tau$, $M_{\tau,n} = M_\tau$. A similar argument shows the same to be true for $m_{\tau,n}$ so that, in either case, we have $M_{\tau,n}/m_{\tau,n} \leqq F^{1/n}$. Hence $[x_i, P_\tau, \Omega]$ satisfies condition F.
Now, for a.e. x in $(0, 1)$ and $n \geqq 1$, we have

$$\begin{aligned} &\frac{\sup_{0<t<1} (d/dt) f(a_1(x) + f(a_2(x) + \cdots + f(a_n(x) + t)))}{\inf_{0<t<1} (d/dt) f(a_1(x) + f(a_2(x) + \cdots + f(a_n(x) + t)))} \\ &\quad = \frac{\sup_{0<t<1} \prod_{i=1}^n f'(a_i(x) + f(a_{i+1}(x) + \cdots + f(a_n(x) + t)))}{\inf_{0<t<1} \prod_{i=1}^n f'(a_i(x) + f(a_{i+1}(x) + \cdots + f(a_n(x) + t)))} \\ &\quad \leqq \prod_{i=1}^n \frac{\sup_{0<t<1} f'(a_i(x) + f(a_{i+1}(x) + \cdots + f(a_n(x) + t)))}{\inf_{0<t<1} f'(a_i(x) + f(a_{i+1}(x) + \cdots + f(a_n(x) + t)))} \\ &\quad \leqq \left(\frac{M_{\tau,n}}{m_{\tau,n}}\right)^n \leqq F. \end{aligned}$$

Hence f_τ satisfies condition C which implies $(f_\tau, 1)$ is an ergodic expansion pair.

5.2. *A representation theorem.* Suppose, once again, that $[x_i, P, \Omega]$ is an arbitrary stationary finite positive process with state space S. Let

$$B(i_1, \cdots, i_{n+1}) = [i_1 + F(i_2, \cdots, i_{n+1} - 1), i_1 + F(i_2, \cdots, i_{n+1})]$$

where F is defined relative to $[x_i, P, \Omega]$ as in Section 4.1. Set

$$B_n = \{B(i_1, \cdots, i_{n+1}) \mid (i_1, \cdots, i_{n+1}) \in S^{n+1}\}$$

and let $\mathscr{B}_n$ be the field generated by B_n. Then $([0, N), \mathscr{B}_n)$, $n = 1, 2, \cdots$, is a sequence of measurable spaces such that $\mathscr{B}_n \subset \mathscr{B}_{n+1}$ for $n \geqq 1$.

Now, for each n, let $([0, N)^n, \mathscr{B}^n)$ denote the Cartesian product $\prod_{i=1}^n ([0, N), \mathscr{B}_n)$ and let $([0, N)^\infty, \mathscr{B})$ be the Cartesian product of all of the $([0, N), \mathscr{B}_n)$. Define the probability measure P_n on $([0, N)^n, \mathscr{B}^n)$ by setting

$$P_n(B) = P[x_j = i_{n+1,j}, j = 2(1)n+1]/N i_{1,j} = i_{2.j} = \cdots = i_{j,j}, j = 1(1)n+1,$$
$$= 0 \qquad \text{otherwise},$$

for all $B = B(i_{1,1}, i_{1,2}) \times \cdots \times B(i_{n+1,1}, \cdots, i_{n+1,n+1})$ in $\prod_{i=1}^n \mathscr{B}_n$ and extending P_n to $\mathscr{B}^n$ in the natural way. It is easily seen that $P_1, P_2, \cdots$ is a consistent sequence of measures so, by the Kolmogorov consistency theorem, there exists a unique probability measure P^* on $([0, N)^\infty, \mathscr{B})$ such that $P_n(B) = P^*(B \times \prod_{i=1}^\infty [0, N))$ for all $B \in \mathscr{B}_n$.

Now let $\bar{\mathscr{B}} = \bigvee_{n=i}^\infty \mathscr{B}_n$ and define the probability measure $\bar{P}$ on $([0, N), \bar{\mathscr{B}})$ by setting

$$\bar{P}(\textstyle\bigcap_{k=1}^n B(i_{n_k, 1}, i_{n_k, 2}, \cdots, i_{n_k, n_k+1})) = P^*(\prod_{k=1}^n B(i_{n_k, 1}, i_{n_k, 2}, \cdots, i_{n_k, n_k+1})$$
$$\times \textstyle\prod_{i=1}^\infty [0, N))$$

for any $n \geqq 1$; $n_k \geqq 1$, $k = 1, 2, \cdots, n$; and $B(i_{n_k,1}, \cdots, i_{n_k,n_k+1}) \in B_{n_k}$ and extending $\bar{P}$ to $\bar{\mathscr{B}}$ in the natural way. Note that

$$\bar{P}(B(i_1, \cdots, i_{n+1})) = P[x_j = i_j, j = 2(1)n+1]/N$$

for all $B(i_1, \cdots, i_{n+1}) \in B_n$ and $n \geqq 1$.

Next, suppose f_τ', $\tau = 1, 2, \cdots$, is the sequence of a.e. derivatives of the f_τ defined relative to $[x_i, P, \Omega]$ as in 5.1. We see immediately that f_τ' is measurable with respect to $([0, N), \mathscr{B}_\tau)$ and we have already remarked that $\mathscr{B}_1 \subset \mathscr{B}_2 \subset \cdots \subset \mathscr{B}_\tau \subset \mathscr{B}_{\tau+1} \subset \cdots \subset \bar{\mathscr{B}}$.

THEOREM 5.3. *The sequence* $(f_\tau', \mathscr{B}_\tau)$, $\tau = 1, 2, \cdots$, *is a martingale.*

PROOF. First, we have

$$E(f_\tau') = \textstyle\sum_{S^{\tau+1}} P[x_1 = i_1 \mid x_j = i_j, j = 2(1)\tau+1] P[x_j = i_j, j = 2(1)\tau+1]/N$$
$$= \textstyle\sum_{S^{\tau+1}} P[x_j = u_j, j = 1(1)\tau+1]/N = 1/N$$

so that the expectation of each f_τ' is certainly finite. Secondly, for $x \in B(i_1, \cdots, i_{\tau+1})$

$$\begin{aligned}
&E(f_{\tau+1}' \mid \mathscr{B}_\tau)(x)\\
&= (P[x_j = i_j, j = 2(1)\tau+1])^{-1} \textstyle\sum_{k=0}^{N-1} P[x_1 = i_1 \mid x_j = i_j, j = 2(1)\tau+1, x_{\tau+2} = k]\\
&\qquad \times P[x_j = i_j, j = 2(1)\tau+1, x_{\tau+2} = k]\\
&= (P[x_j = i_j, j = 2(1)\tau+1])^{-1} \textstyle\sum_{k=0}^{N-1} P[x_j = i_j, j = 1(1)\tau+1, x_{\tau+2} = k]\\
&= P[x_j = i_j, j = 1(1)\tau+1]/P[x_j = i_j, j = 2(1)\tau+1]\\
&= P[x_1 = i_1 \mid x_j = i_j, j = 2(1)\tau+1]\\
&= f_\tau'(x)
\end{aligned}$$

and since B_n is a partition of $[0, N)$, the proof is complete.

Therefore, by the martingale convergence theorem, there is a function g such that $f_\tau' \to g$ a.e. Further, since $0 \leqq f_\tau' \leqq 1$, by the bounded convergence theorem we have $f_n \to \int g$ a.e. Let us define $f = \int g$.

We see immediately that if we let $T = \langle f^{-1} \rangle$ and $T_\tau = \langle f_\tau^{-1} \rangle$ we have

$$\lambda(T^{-1}([0,\alpha))) = \sum_{k=0}^{N-1} \int_{f(k)}^{f(k+\alpha)} dt = \sum_{k=0}^{N-1} \lim_{\tau\to\infty} \int_{f_\tau(k)}^{f_\tau(k+\alpha)} dt$$
$$= \lim_{\tau\to\infty} \lambda(T_\tau^{-1}([0,\alpha))) = \alpha$$

since each $(f_\tau, 1)$ is an expansion pair. Therefore, $(f, 1)$ is an expansion pair.

Further, if $[a_i, \lambda, (0, 1)_f]$ is the stochastic process associated with $(f, 1)$, then

$$\lambda[a_j = i_j, j = 1(1)n] = \int_{f(i_1+f(i_2+\cdots+f(i_n)))}^{f(i_1+f(i_2+\cdots+f(i_n+1)))} dt$$
$$= \lim_{\tau\to\infty} \int_{f_\tau(i_1+f_\tau(i_2+\cdots+f_\tau(i_n)))}^{f_\tau(i_1+f_\tau(i_2+\cdots+f_\tau(i_n+1)))} dt$$
$$= P[x_j = i_j, j = 1(1)n]$$

since

$$\int_{f_\tau(i_1+f_\tau(i_2+\cdots+f_\tau(i_n)))}^{f_\tau(i_1+f_\tau(i_2+\cdots+f_\tau(i_n+1)))} dt = P[x_j = i_j, j = 1(1)n]$$

all $\tau \geqq n$. Therefore $[a_i, \lambda, (0, 1)_f]$ has the same finite dimensional distributions as $[x_n, P, \Omega]$. As a result, we have proven

THEOREM 5.4. *If (f, h) is an expansion pair whose associated stochastic process is finite positive, then there exists an expansion pair $(f^*, 1)$ whose associated stochastic process has the same finite dimensional distributions.*

Acknowledgments. I would like to thank Professor John R. Kinney for his guidance and the Department of Statistics and Probability, Michigan State University, for its financial support during the preparation of this paper.

REFERENCES

[1] BISSINGER, B. N. (1944). A generalization of continued fractions. *Bull. Amer. Math. Soc.* **50** 868–876.

[2] DAVIS, H. T. (1962). *The Summation of Series.* Principia Press of Trinity Univ., San Antonio.

[3] EVERETT, C. J. (1946). Representations for real numbers. *Bull. Amer. Math. Soc.* **52** 861–869.

[4] HILLE, E. (1959). *Analytic Function Theory.* Blaisdell, New York.

[5] JOLLEY, L. B. W. (1961). *Summation of Series.* Dover, New York.

[6] KHINCHIN, A. YA (1964). *Continued Fractions.* Univ. of Chicago Press.

[7] KINNEY, J. and PITCHER T. (1966). The dimension of some sets defined in terms of *f*-expansions. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **4** 293–315.

[8] RÉNYI, A. (1957). Representations for real numbers and their ergodic properties. *Acta Math. Acad. Sci. Hungar.* **8** 477–493.

[9] ROKHLIN, V. A. (1961). Exact endomorphisms of a Lebesgue space. *Izv. Akad. Nauk. SSSR, Ser. Mat.* **25** 499–530. (English translation in *Amer. Math. Soc. Transl. Ser. 2* **39** (1964) 1–36.)

[10] RYLL-NARDZEWSKI, C. (1951). On the ergodic theorems (II). Ergodic theory of continued fractions. *Studia Math.* **12** 74–79.