AN INVERSION ALGORITHM FOR ONE-DIMENSIONAL F-EXPANSIONS

Thesis for the Degree of Ph. D.

MICHIGAN STATE UNIVERSITY
SCOTT BATES GUTHERY

1969

This is to certify that the

thesis entitled

AN INVERSION ALGORITHM FOR
ONE-DIMENSIONAL F-EXPANSIONS

presented by

Scott Bates Guthery

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Statistics and Probability

John R Kinney
Major professor

Date Nov 13, 1969

O-169

ABSTRACT

AN INVERSION ALGORITHM FOR
ONE-DIMENSIONAL F-EXPANSIONS

By

Scott Bates Guthery

This paper presents an algorithm for the construction of a function, whose fractional part preserves a given Lebesgue equivalent measure on (0,1), from a summation representation of the Radon-Nikodym derivative of the given measure. Examples are given and the technique is used to construct a function whose associated f-expansion stochastic process has the same finite dimensional distributions as a given stationary process.

AN INVERSION ALGORITHM FOR ONE-DIMENSIONAL F-EXPANSIONS

By

Scott Bates Guthery

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Statistics and Probability

1969

To Peter

ACKNOWLEDGMENTS

I would like to thank Professor John R. Kinney for his patient guidance before and during the preparation of this dissertation. His insights in information theory, once decoded, have served to greatly expand my appreciation of the field.

I would also like to thank Mrs. Noralee Barnes for her help in preparing the manuscript, the Department of Statistics and Probability for financial support during my graduate career, and Messrs. Allan Oaten and Michael Waterman for many helpful conversations.

TABLE OF CONTENTS

I. 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. 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 $<\ >$ the fractional part, we use the expansion algorithm:

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

and for $i \geq 1$, 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 $\rho_n(x) = f(a_1(x) + f(a_2(x) + \ldots + f(a_n(x))$

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

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

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

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, in all cases, assumed to be the Borel field $\mathcal{B}$ 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

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

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.

For example, to calculate the frequence of the digit p in the continued fraction expansion of a number x in $(0,1)_f$, one defines

$$I_p(q) = \begin{cases} 1 & q = p \\ 0 & q \neq p \end{cases}$$

and sets

$$\begin{aligned} F(p,x) &= \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} I_p(a_k(x)) \\ &= \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} I_p([1/T^k(x)]). \end{aligned}$$

Using the individual ergodic theorem, one then has

$$\begin{aligned} \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} I_p([1/T^k(x)]) &= \int_0^1 I_p([1/t])d\omega(t) \qquad \text{a.e.} \\ &= \frac{1}{\log 2} \int_0^1 \frac{I_p([1/t])}{(t+1)}\, dt \\ &= \frac{1}{\log 2} \log \frac{(1+p)^2}{p(p+2)} . \end{aligned}$$

That is, for a.e. x in $(0,1)$ the frequence of p in its continued fraction expansion is $\frac{1}{\log 2} \log \frac{(p+1)^2}{p(p+2)}$.

Then, in 1957, Renyi (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 non-negative, continuous, and strictly decreasing for $1 \le t \le N$ and $f(t) = 0$ for $t \ge N$ where $N > 2$ is an integer or $+\infty$;

A3) $|f(t_2) - f(t_1)| \le |t_2 - t_1|$ for $1 \le t_1 < t_2$ and $|f(t_2) - f(t_1)| < |t_2 - t_1|$ if $\tau - \varepsilon < t_1 < t_2$ 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 non-negative, continuous, and strictly increasing for $0 \le t \le N$ and $f(t) = 1$ for $t \ge N$ where $N > 1$ is an integer or $+\infty$;

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

and C) if $H_n(x,t) = \frac{d}{dt} f(a_1(x) + f(a_2(x) + \ldots + f(a_n(x) + t)))$ then

$$\frac{\sup\limits_{0<t<1} H_n(x,t)}{\inf\limits_{0<t<1} H_n(x,t)} \le C < +\infty$$

where the constant $C \ge 1$ depends neither on x nor on n;

Renyi proved

<u>Theorem</u> <u>1.1</u>. 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) $C^{-1} \leq \frac{d\omega}{d\lambda} \leq C$.

This theorem defines an entire class of functions whose measure theoretic f-expansion properties can be investigated using the technique introduced by Ryll-Nardzewski, viz., 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 Renyi. Defining $\varphi = f^{-1}$ and

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

he proved

<u>Theorem 1.2</u>. 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 endomorphism T and the theorem says that

$$\lambda(B_n(x)) \approx e^{-nh(T)}.$$

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.

2. Terminology

Suppose we consider the following conditions on a function f:

A') $f(1) = 1$; $f(t)$ is non-negative, continuous and decreasing for $1 \le t \le N$; and $f(t) = 0$ for $t \ge N$ where $N > 2$ is an integer or $+\infty$;

B') $f(0) = 0$; $f(t)$ is non-negative, continuous, and increasing for $0 \le t \le N$; and $f(t) = 1$ for $t \ge N$ where $N > 1$ is an integer or $+\infty$.

If f satisfies A' let us define

$$f^{-1}(x) = \mathrm{glb}\{t \mid f(t) \le x\}$$

and if f satisfies B' let us define

$$f^{-1}(x) = \mathrm{glb}\{t \mid f(t) \ge x\}$$

for all $x \in (0,1)$. With these definitions, we see that an f which satisfies A' or B' may be used in the expansion algorithm and that the set $(0,1)_f$ is well-defined. Such a function will be said to be available for f-expansions. Obviously, any function which satisfies A or B satisfies A' or B' respectively. Note also that f^{-1} is continuous at all but at most a countable set of points and that except for these points $f^{-1}(x)$ is that unique y such that $f(y) = x$.

Suppose now that f is available for f-expansions and that ω is a λ-equivalent measure on $(0,1)$. If the transformation $T = \langle f^{-1} \rangle$ is an endomorphism on $((0,1), \mathcal{B}, \omega)$; i.e., T is measurable and $\omega(T^{-1}B) = \omega(B)$ for all $B \in \mathcal{B}$; 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. In this terminology Renyi'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 $C^{-1} \leq d\omega/d\lambda \leq C$ and $(f, d\omega/d\lambda)$ is a valid, ergodic expansion pair.

It has been found that many functions may be invariant with respect to the same measure. In the following we will study this relationship by providing an inversion algorithm which produces a variety of functions which preserve a given λ-equivalent measure. Conditions will also be presented which insure that the resultant expansion pairs are ergodic and valid. As a result we will have a number of examples of the results of Renyi's theorem and, therefore, functions whose f-expansion properties can be investigated with the individual ergodic theorem. Finally, we note that just as many functions preserve the same measure, many expansion pairs may be associated with the same stochastic process. We shall close by showing that the inversion algorithm may also be of use in studying this relationship.

II. 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.

1. Definitions and Basic Relations

Let g be an a.e. non-negative Lebesgue integrable function on $[0,N)$ where N is an integer ≥ 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 nondecreasing function from $[0,N)$ onto $[0,1)$ which is absolutely continuous on every finite subinterval of $[0,N)$.

Next we 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 probability measure ω on the Borel subsets of $(0,1)$ which is equivalent to Lebesgue measure. If we now set $H(x) = \int_0^x h(t)dt$ for $x \in [0,1]$, then H and H^{-1} are 1-1 strictly increasing a.e. differentiable transformations on $[0,1]$. Finally, we define

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

and

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

We note immediately that f_U is an a.e. differentiable nondecreasing function on $[0,N)$ such that $f_U(0) = 0$ and $\lim_{x\to N} f_U(x) = 1$. On the other hand, f_D is an a.e. differentiable nonincreasing function on $[1,N+1)$ such that $f_D(1) = 1$ and $\lim_{x\to N+1} f_D(x) = 0$. Let us complete f_U and f_D by setting

$$f_U(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq N \end{cases}$$

and

$$f_D(x) = \begin{cases} 1 & x < 1 \\ 0 & x \geq N+1. \end{cases}$$

It is easily seen that f_D and f_U satisfy A' and B' respectively so that both are available for f-expansions.

In the following, we shall let

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

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

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 $[0,1]$. Before proceeding to discuss the expansion properties of f_U and f_D, we present the following lemma concerning elementary relations between them.

<u>Lemma</u> <u>2.1</u>. The following relations hold:

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

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

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

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

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

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

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

<u>Proof</u>.

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

$$\begin{aligned}
\varphi_U(x) &= \mathrm{glb}\{t \mid f_U(t) \ge x\} \\
&= \mathrm{glb}\{t \mid R(f_D(t+1)) \ge x\} \\
&= \mathrm{glb}\{t \mid f_D(t+1) \ge R(x)\} \\
&= \mathrm{glb}\{t-1 \mid f_D(t) \ge R(x)\} \\
&= \mathrm{glb}\{t \mid f_D(t) \ge R(x)\} - 1 \\
&= \varphi_D(R(x)) - 1
\end{aligned}$$

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 $df^{-1}(u) = du/(\frac{df}{du}(f^{-1}(u)))$ to obtain

$$f_U'(x) = d\,H^{-1}(G(x)) = g(x)/(\frac{dH}{dx}(H^{-1}(G(x)))$$

$$= g(x)/h(f_U(x)) \text{ a.e.}$$

and

$$f_D'(x) = d\,H^{-1}(1 - G(x-1)) = -g(x-1)/(\frac{dH}{dx}(H^{-1}(1 - G(x-1))))$$

$$= -g(x-1)/h(f_D(x)) \text{ a.e.}$$

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

$$R'(x) = d\,H^{-1}(1 - H(x)) = -h(x)/h(R(x)) \text{ a.e.}$$

That (f_U,h) and (f_D,h) are, in fact, expansion pairs is shown by

Theorem 2.2. The transformations T_U and T_D are endomorphisms on $((0,1),\mathcal{B},\omega)$.

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

$$\omega(T_U^{-1}(0,\alpha)) = \sum_{k=0}^{N-1} \int_{f(k)}^{f(k+\alpha)} h(t)dt = \sum_{k=0}^{N-1} H(f(k+\alpha)) - H(f(k))$$

$$= \sum_{k=0}^{N-1} \left(\int_0^{k+\alpha} g(t)dt - \int_0^{k} g(t)dt \right) = \sum_{k=0}^{N-1} \int_k^{k+\alpha} g(t)dt$$

$$= \sum_{k=0}^{N-1} \int_0^{\alpha} g(t+k)dt = \int_0^{\alpha} \sum_{k=0}^{N-1} g(t+k)dt$$

$$= \int_0^{\alpha} h(t)dt = \omega((0,\alpha))$$

so that T_U preserves ω.

If, on the other hand, we let $f = f_D$, we have

$$\omega(T_D^{-1}(0,\alpha)) = \sum_{k=1}^{N} \int_{f(k+\alpha)}^{f(k)} h(t)dt = \sum_{k=1}^{N} H(f(k)) - H(f(k+\alpha))$$

$$= \sum_{k=1}^{N} \left(\int_0^{k+\alpha-1} g(t)dt - \int_0^{k-1} g(t)dt \right)$$

$$= \sum_{k=0}^{N-1} \int_k^{k+\alpha} g(t)dt = \sum_{k=0}^{N-1} \int_0^{\alpha} g(t+k)dt$$

$$= \int_0^{\alpha} \sum_{k=0}^{N-1} g(t+k)dt = \int_0^{\alpha} h(t)dt = \omega((0,\alpha))$$

so that T_D preserves ω and the proof is complete.

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) $g(t) < \inf_{0 \le x \le 1} \sum_{k=0}^{N-1} g(x+k)$ 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 $|f_U'(x)| = \frac{g(x)}{h(f_U(x))}$ a.e. and $|f_D'(x)| = \frac{g(x-1)}{h(f_D(x))}$ 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 Renyi'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

E1) $0 < \varepsilon_1 \le |f'(x)| \le \varepsilon_2 < 1$ for $x \in [0,N)$

and E2) f' Lipschitz of order 1

then f satisfies condition C.

Proof.

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

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

and from (ii) we have that

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

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

$$\frac{\sup\limits_{0<t<1} f'(a_1+f(a_2+\ldots+f(a_n+t)))}{\inf\limits_{0<t<1} f'(a_1+f(a_2+\ldots+f(a_n+t)))} - 1 =$$

$$= \frac{\sup\limits_{0<t<1} f'(a_1+f(a_2+\ldots+f(a_n+t))) - \inf\limits_{0<t<1} f'(a_1+f(a_2+\ldots+f(a_n+t)))}{\inf\limits_{0<t<1} f'(a_1+f(a_2+\ldots+f(a_n+t)))}$$

$$\leq \frac{M|f(a_2+f(a_3+\ldots+f(a_n+1))) - f(a_2+f(a_3+\ldots+f(a_n)))|}{\epsilon_1}$$

$$\leq \frac{M\epsilon_2|f(a_3+f(a_4+\ldots+f(a_n+1))) - f(a_3+f(a_4+\ldots+f(a_n)))|}{\epsilon_1}$$

$$\leq \ldots \leq \frac{M\epsilon_2^{n-1}}{\epsilon_1}$$

therefore,

$$\frac{\sup\limits_{0<t<1} H_n(x,t)}{\inf\limits_{0<t<1} H_n(x,t)} = \frac{\sup\limits_{0<t<1} \frac{d}{dt} f(a_1(x) + f(a_2(x) +\ldots+ f(a_n(x)+t)))}{\inf\limits_{0<t<1} \frac{d}{dt} f(a_1(x) + f(a_2(x) +\ldots+ f(a_n(x)+t)))}$$

$$\leq \prod_{j=1}^{n} \frac{\sup\limits_{0<t<1} f'(a_j(x) + f(a_{j+1}(x) +\ldots+ f(a_n(x)+t)))}{\inf\limits_{0<t<1} f(a_j(x) + f(a_{j+1}(x) +\ldots+ f(a_n(x)+t)))}$$

$$\leq \prod_{j=1}^{n} \left(1 + \frac{M\epsilon_2^{j-1}}{\epsilon_1}\right)$$

$$\leq \prod_{j=1}^{\infty} \left(1 + \frac{M\epsilon_2^{j-1}}{\epsilon_1}\right) = C < \infty$$

since, by theorem 8.6.1 of Hille (4), the infinite product converges if $\sum\limits_{j=1}^{\infty} \frac{M\epsilon_2^{j-1}}{\epsilon_1}$ 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 input function g.

3. Rokhlin's Formula

We conclude this chapter 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.

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

III. EXAMPLES

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

1. Lebesgue Measure

Perhaps the easiest and most interesting measure to invert is Lebesgue measure which has density function $h(x) = 1$. If we, for example, have non-negative constants p_k such that $\sum_{k=0}^{N-1} p_k = 1$ then we set $g(x) = p_k$ for $k \le x < k+1$ and $k = 0,1,\ldots,N-1$. Then, since $H(x) = H^{-1}(x) = x$, we have

$$f_U(x) = G(x) = \int_0^x g(t)dt = \sum_{k=0}^{[x]-1} p_k + \langle x\rangle p_{[x]}$$

and

$$f_D(x) = 1 - G(x-1) = 1 - \sum_{k=0}^{[x]-2} p_k - \langle x\rangle p_{[x]-1}.$$

A well-known special case of this expansion is obtained by setting $p_k = \frac{1}{M}$ for $k = 0,1,\ldots,M-1$, from which we get

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

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 $0 < \epsilon_1 \le p_k \le \epsilon_2 < 1$. Then by noting that $h(x) = 1$ implies $f_U'(x) = g(x)$ 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 the conditions of Renyi's theorem. Letting $S_n = \sum_{k=0}^{n} p_k$ and $S_{-1} = 0$, we can compute their entropy by Rokhlin's formula as follows:

$$h(T_D) = h(T_U) = \int_0^1 \log\left|\varphi_U^{-1}(x)\right| dx = \sum_{n=0}^{N-1} \int_{S_{n-1}}^{S_n} \log \frac{1}{p_n}\, dx$$

$$= \sum_{n=0}^{N-1} (-\log p_n)(S_n - S_{n-1})$$

$$= -\sum_{n=0}^{N-1} p_n \log p_n .$$

2. <u>Generalizations of the Continued Fraction</u>

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 $m \ge 2$ 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) we have

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

where Ψ is the psi function defined by $\Psi(x) = \frac{d}{dx} \ln \Gamma(x)$ 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

$$\int_0^\infty g(t)dt = \sum_{n=0}^{\infty} \int_0^1 U_n(t)dt = -\sum_{i=1}^{m} a_i \int_0^1 \Psi(t+b_i)dt$$

$$= -\sum_{i=1}^{m} a_i(\ln \Gamma(1+b_i) - \ln \Gamma(b_i))$$

$$= -\sum_{i=1}^{m} a_i \ln b_i = 1.$$

Since

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

we have

$$H(x) = \int_0^x h(t)dt = -\sum_{i=1}^{m} a_i \int_0^x \Psi(t+b_i)dt$$

$$= -\sum_{i=1}^{m} a_i \ln(\Gamma(x+b_i)/\Gamma(b_i)).$$

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

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

$$= -\sum_{i=1}^{k} c_i \ln(\frac{d_i}{x+d_i})$$

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

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

and since

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

we have

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

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 $B = \ln(\frac{b_3(b_1+1)}{b_1(b_3+1)})$ then

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

and since

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

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

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

$$f_U(x) = \frac{b_1b_3(b_3-b_1) + b_1(x+b_3)(x+b_1+1) - b_3(x+b_1)(x+b_3+1)}{(b_1-b_3) + b_1(x+b_1)(x+b_3+1) - b_3(x+b_3)(x+b_1+1)}$$

and

$$f_D(x) = \frac{b_1b_3(b_1-b_3)}{b_1(x+b_3-1)(x+b_1) - b_3(x+b_1-1)(x+b_3)}$$

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 $B = \ln\left(\frac{(b_1+1)(b_3+1)}{b_1b_3}\right)$.

Here we have

$$H^{-1}(x) = \frac{\sqrt{(b_1-b_3)^2 + 4b_1b_3e^{Bx}} - (b_1+b_3)}{2}$$

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

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

and

$$f_D(x) = \frac{1}{2}\sqrt{(b_1-b_3)^2 + 4b_1b_3(x+b_1)(x+b_3)(x+b_1+1)^{-1}(x+b_3+1)^{-1}}-\frac{1}{2}(b_1+b_3)$$

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. 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

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

for $a > 0$.

Here we let

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

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 $0 \le x < 1$. Next, since we are setting

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

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

$$H^{-1}(x) = \frac{1}{a} \log((e^a - 1)x + 1).$$

Therefore, we have

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

$$= \frac{1}{a} \log(\sum_{k=1}^{[x]} \frac{a^k}{k!} + \frac{(a\langle x\rangle)^{[x+1]}}{[x+1]!} + 1)$$

and, similiarly,

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

Another family of expansion pairs, which extends the above Lebesgue family, is provided by picking $\{\alpha_i\}_{i=0}^n$ where $0 = \alpha_0 < \alpha_1 < \ldots < \alpha_{n-1} < \alpha_n = 1$ and setting

$$h(x) = \beta_i \quad \alpha_{i-1} \leq x < \alpha_i, \ i = 1, 2, \ldots, n$$

such that $\sum_{i=1}^n \beta_i(\alpha_i - \alpha_{i-1}) = 1$. Then, if $p_i > 0$ for $i = 0, 1, \ldots, N-1$ and $\sum_{i=0}^{N-1} p_i = 1$, we set

$$g(x) = p_{[x]} \beta_i$$

for $\alpha_{i-1} \leq \langle x\rangle < \alpha_i$. This family inverts easily and yields monotone "broken line" functions.

IV. 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.

The result follows directly from

Lemma 4.2. Let $(\Omega, \mathfrak{A}, P)$ be a probability space and T be an endomorphism on the space.

Then the random variables X and $X \circ T$ are identically distributed.

Proof.

If B is a Borel subset of the real line, then

$$\begin{aligned} P[X \in B] &= P(\{\omega | X(\omega) \in B\}) \\ &= P(T^{-1}\{\omega | X(\omega) \in B\}) \\ &= P(\{\omega' | T\omega' \in \{\omega | X(\omega) \in B\}\}) \\ &= P(\{\omega | X(T(\omega)) \in B\}) \\ &= P[X \circ T \in B]. \end{aligned}$$

Since $(f, d\omega/d\lambda)$ is an expansion pair the transformation $T(x) = \langle f^{-1}(x)\rangle$ is an endomorphism on $((0,1)_f, \mathcal{B}, \omega)$. Therefore, by the Lemma, $a_1(x) = [f^{-1}(x)]$ and $a_k(x) = [f^{-1}(T^k x)]$ for $k = 2,3,\ldots$ all have the same distribution.

1. Inversion Using a Markov Process

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

$$I_M(i_1,\ldots,i_M) = \sum_{j=1}^{M} i_j\, N^{M-j}$$

and

$$F(i_1,\ldots,i_M) = \Sigma\, P[x_n = j_n,\ n = 1(1)M]$$

where the latter summation extends over all $(j_1,\ldots,j_M)$ for which $I_M(j_1,\ldots,j_M) \leq I_M(i_1,\ldots,i_M)$. Let us further define $F(i_1,\ldots,i_M\text{-}1) = F(i_1,\ldots,i_{M-1},-1)$ and $F(-1) = 0$.

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 $(i_1,\ldots,i_\tau) \in S^\tau$, let $J(i_1,\ldots,i_\tau)$ be the indicator function of the interval $[H^{-1}(F(i_1,\ldots,i_\tau\text{-}1)),\ H^{-1}(F(i_1,\ldots,i_\tau))]$ and define

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

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

$$\sum_{k=0}^{N-1} g(x+k) = \sum_{k=0}^{N-1} \sum_{S^\tau} P[x_1 = k | x_{j+1} = i_j,\ j = 1(1)\tau] J(i_1,\ldots,i_\tau)(x)h(x)$$

$$= \sum_{k=0}^{N-1} P[x_1 = k | x_{j+1} = i_j^*,\ j = 1(1)\tau]h(x) \quad \text{for} \quad J(i_1^*,\ldots,i_\tau^*)(x) = 1$$

$$= h(x),$$

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

<u>Theorem</u> <u>4.3</u>. If $H(f(i_1 + f(i_2 +\ldots+ f(i_\tau))) = F(i_1,\ldots,i_\tau - 1)$ for all $(i_1,\ldots,i_\tau) \in S^\tau$, then $[x_i,P,\Omega]$ and $[a_i,\omega,(0,1)_f]$ have the same finite dimensional distributions.

<u>Proof</u>.

First, we have

$$\omega[a_j = i_j,\ j = 1(1)\tau+1] = \int_{f(i_1+f(i_2+\ldots+f(i_{\tau+1})))}^{f(i_1+f(i_2+\ldots+f(i_{\tau+1}+1)))} h(t)dt$$

$$= H(f(i_1+f(i_2+\ldots+f(i_{\tau+1}+1)))) - H(f(i_1+f(i_2+\ldots+f(i_{\tau+1}))))$$

$$= G(i_1+f(i_2+\ldots+f(i_{\tau+1}+1)))) - G(i_1+f(i_2+\ldots+f(i_{\tau+1})))$$

$$= \int_{i_1+f(i_2+\ldots+f(i_{\tau+1}))}^{i_1+f(i_2+\ldots+f(i_{\tau+1}+1))} g(t)dt$$

$$= P[x_1 = i_1 | x_j = i_j,\ j = 2(1)\tau+1] \int_{f(i_2+\dots+f(i_{\tau+1}))}^{f(i_2+\dots+f(i_{\tau+1}+1))} h(t)dt$$

$$= P[x_1 = i_1 | x_j = i_j,\ j = 2(1)\tau+1] \\ (H(f(i_2+\dots+f(i_{\tau+1}+1)))- \\ H(f(i_2+\dots+f(i_{\tau+1}))))$$

$$= P[x_1 = i_1 | x_j = i_j,\ j = 2(1)\tau+1] P[x_j = i_{j+1}, j=1(1)\tau]$$

$$= P[x_j = i_j,\ j = 1(1)\tau+1].$$

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

$$\omega[a_j = i_j,\ j = 1(1)n] = \int_{f(i_1+f(i_2+\dots+f(i_n)))}^{f(i_1+f(i_2+\dots+f(i_n+1)))} h(t)dt$$

$$= H(f(i_1+f(i_2+\dots+f(i_n+1))))-H(f(i_1+f(i_2+\dots+f(i_n))))$$

$$= G(i_1+f(i_2+\dots+f(i_n+1)))-G(i_1+f(i_2+\dots+f(i_n)))$$

$$= \int_{i_1+f(i_2+\dots+f(i_n))}^{i_1+f(i_2+\dots+f(i_n+1))} g(t)dt$$

$$= P[x_1=i_1|x_j=i_j,\ j=2(1)\tau+1]\omega[a_j=i_{j+1},\ j=1(1)n]$$

$$= P[x_1=i_1|x_j=i_j,\ j=2(1)\tau+1]P[x_j=i_j,\ j=2(1)n]$$

$$= P[x_j = i_j,\ j = 1(1)n].$$

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 $[\ell,\ell+1)$. Furthermore, if the resultant process is to have a finite memory, the wrinkles must occur at exactly those points in the condition of Theorem 4.3. 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 4.4. If $[a_i,\lambda,(0,1)_f]$ is a stochastic process of multiplicity τ associated with a valid expansion pair $(f,1)$, then for each $(i_1,\ldots,i_{\tau+1}) \in S^{\tau+1}$ we have

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

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

Proof.

For $n \ge 1$ and $(i_1,\ldots,i_n) \in S^n$, let us set

$$M(i_1,\ldots,i_n) = P[x_n=i_n|x_j=i_j,\ j=1(1)n-1]$$

$$= \frac{f(i_1+f(i_2+\ldots+f(i_n+1)))-f(i_1+f(i_2+\ldots+f(i_n)))}{f(i_1+f(i_2+\ldots+f(i_{n-1}+1)))-f(i_1+f(i_2+\ldots+f(i_{n-1})))}$$

and

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

Noting that

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

we have by recursion that

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

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

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

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

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

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

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

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

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

so a.e.,

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

V. 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 $S = \{0,1,\ldots,N-1\}$ such that $P[x_j = i_j,\ j = 1(1)n] > 0$ for all $(i_1,\ldots,i_n) \in S^n$ and $n \geq 1$. We shall call such processes finite positive.

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

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

and

$$P_\tau[x_j = i_j,\ j = 1(1)n] = P[x_j = i_j,\ j = 1(1)\tau] \prod_{k=\tau+1}^{n} P[x_k = i_k | x_j = i_j, j = k-\tau(1)k-1]$$

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 $n \le \tau$. 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 4.3 and letting $f = f_\tau$, we need only show that

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

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

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

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

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

$$+ \sum_{I_{\tau-1}(k_1,\ldots,k_{\tau-1})<I_{\tau-1}(i_2,\ldots,i_\tau)} P[x_1=i_1|x_{j+1}=k_j, j=1(1)\tau]$$

$$= F_\tau(i_1,\ldots,i_\tau-1).$$

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

E: $0 < P[x_{n+1} = i_{n+1}|x_j = i_j,\ j = 1(1)n] \le \epsilon < 1$ for all $(i_1,\ldots,i_{n+1}) \in S^{n+1}$ and $n \ge 1$

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: if

$$M_n = \sup_{S^{n+1}} P[x_{n+1} = i_{n+1}|x_j = i_j,\ j = 1(1)n]$$

and

$$m_n = \inf_{S^{n+1}} P[x_{n+1} = i_{n+1}|x_j = i_j,\ j = 1(1)n]$$

then there exists a constant F such that

$$\frac{M_n}{m_n} \le F^{1/n}$$

for all $n \ge 1$.

<u>Theorem 5.2</u>. 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.

<u>Proof</u>.

For fixed τ and n, we have

$$M_{\tau,n} = \sup_{S^{n+1}} P_\tau[x_{n+1} = i_{n+1} | x_j = i_j,\ j = 1(1)n]$$

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

$$= \sup_{S^{n+1}} P[x_{n+1} = i_{n+1} | x_j = i_j,\ j = n+1-k(1)n]$$

where

$$k = \begin{cases} n & n \le \tau \\ \tau & n > \tau \end{cases}.$$

Therefore, if $n \le \tau$, then $M_{\tau,n} = M_n$ and if $n > \tau$, $M_{\tau,n} = M_\tau$. A similiar argument shows the same to be true for $m_{\tau,n}$ so that, in either case, we have

$$\frac{M_{\tau,n}}{m_{\tau,n}} \le F^{1/n}.$$

Hence $[x_i, P_\tau, \Omega]$ satisfies condition F.

Now, for a.e. x in $(0,1)$ and $n \ge 1$, we have

$$\frac{\sup\limits_{0<t<1} \frac{d}{dt} f(a_1(x) + f(a_2(x) + \ldots + f(a_n(x)+t)))}{\inf\limits_{0<t<1} \frac{d}{dt} f(a_1(x) + f(a_2(x) + \ldots + f(a_n(x)+t)))}$$

$$= \frac{\sup\limits_{0<t<1} \prod\limits_{i=1}^{n} f'(a_i(x) + f(a_{i+1}(x) + \ldots + f(a_n(x)+t)))}{\inf\limits_{0<t<1} \prod\limits_{i=1}^{n} f'(a_i(x) + f(a_{i+1}(x) + \ldots + f(a_n(x)+t)))}$$

$$\le \prod_{i=1}^{n} \frac{\sup\limits_{0<t<1} f'(a_i(x) + f(a_{i+1}(x) + \ldots + f(a_n(x)+t)))}{\inf\limits_{0<t<1} f'(a_i(x) + f(a_{i+1}(x) + \ldots + f(a_n(x)+t)))}$$

$$\le \left(\frac{M_{\tau,n}}{m_{\tau,n}}\right)^n \le F.$$

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

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

for all $(i_1,\ldots,i_{n+1}) \in S^{n+1}$ and $n \ge 1$. Set

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

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

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

$$P_n(B) = \begin{cases} P[x_j = i_{n+1,j},\ j = 2(1)n+1]/N & i_{1,j}=i_{2,j}=\ldots=i_{j,j},\ j=1(1)n+1 \\ 0 & \text{otherwise} \end{cases}$$

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

Now let $\overline{\mathcal{B}} = \bigvee_{n=1}^{\infty} \mathcal{B}_n$ and define the probability measure $\overline{P}$ on $([0,N), \overline{\mathcal{B}})$ by setting

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

for any $n \geq 1$; $n_k \geq 1$, $k = 1,2,\ldots,n$; and $B(i_{n_k,1}, \ldots, i_{n_k,n_k+1}) \in B_{n_k}$ and extending $\overline{P}$ to $\overline{\mathcal{B}}$ in the natural way. Note that

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

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

Next, suppose f'_τ, $\tau = 1,2,\ldots$, 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 $([0,N), \mathcal{B}_\tau)$ and we have already remarked that $\mathcal{B}_1 \subset \mathcal{B}_2 \subset \ldots \subset \mathcal{B}_\tau \subset \mathcal{B}_{\tau+1} \subset \ldots \subset \overline{\mathcal{B}}$.

<u>Theorem</u> <u>5.3</u>. The sequence $(f'_\tau, \mathcal{B}_\tau)$, $\tau = 1,2,\ldots$, is a martingale.

<u>Proof</u>.

First, we have

$$E(f'_\tau) = \sum_{S^{\tau+1}} P[x_1 = i_1 | x_j = i_j, j=2(1)\tau+1] P[x_j = i_j, j=2(1)\tau+1]/N$$

$$= \sum_{S^{\tau+1}} P[x_j = u_j,\ j = 1(1)\tau+1]/N = 1/N$$

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

$$E(f'_{\tau+1}|\mathcal{B}_\tau)(x) = (P[x_j=i_j, j=2(1)\tau+1])^{-1}\sum_{k=0}^{N-1} P[x_1=i_1|x_j=i_j, j=2(1)\tau+1, x_{\tau+2}=k]$$

$$\times P[x_j=i_j, j=2(1)\tau+1, x_{\tau+2}=k]$$

$$= (P[x_j=i_j, j=2(1)\tau+1])^{-1}\sum_{k=0}^{N-1} P[x_j=i_j, j=1(1)\tau+1, x_{\tau+2}=k]$$

$$= P[x_j=i_j, j=1(1)\tau+1]/P[x_j=i_j, j=2(1)\tau+1]$$

$$= P[x_1=i_1|x_j=i_j, j=2(1)\tau+1]$$

$$= f'_\tau(x)$$

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 $0 \le f'_\tau \le 1$, 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

$$\lambda[a_j = i_j,\ j = 1(1)n] = \int_{f(i_1+f(i_2+\ldots+f(i_n)))}^{f(i_1+f(i_2+\ldots+f(i_n+1)))} dt$$

$$= \lim_{\tau\to\infty} \int_{f_\tau(i_1+f_\tau(i_2+\ldots+f_\tau(i_n)))}^{f_\tau(i_1+f_\tau(i_2+\ldots+f_\tau(i_n+1)))} dt$$

$$= P[x_j = i_j,\ j = 1(1)n]$$

since $\int_{f_\tau(i_1+f_\tau(i_2+\ldots+f_\tau(i_n)))}^{f_\tau(i_1+f_\tau(i_2+\ldots+f_\tau(i_n+1)))} dt = P[x_j = i_j,\ j = 1(1)n]$ for all $\tau \geq n$. Therefore $[a_i,\lambda,(0,1)_f]$ has the same finite dimensional distributions as $[x_n,P,\Omega]$. As a result, we have proven

<u>Theorem</u> <u>5.4</u>. 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.

BIBLIOGRAPHY

BIBLIOGRAPHY

1. Bissinger, B.H., A generalization of continued fractions. Bull. Amer. Math. Soc. 50 (1944), 868-876.

2. Davis, H.T., The Summation of Series. The Principia Press of Trinity University: San Antonio (1962).

3. Everett, C.J., Representations for real numbers. Bull. Amer. Math. Soc. 52 (1946), 861-869.

4. Hille, E., Analytic Function Theory. Blaisdell Publishing Company: New York (1959).

5. Jolley, L.B.W., Summation of Series. Dover Publications: New York (1961).

6. Khinchin, A.Ya, Continued Fractions. The University of Chicago Press: Chicago (1964).

7. Kinney, J. and T. Pitcher, The dimension of some sets defined in terms of f-expansions. Z. Wahrscheinlichkeitstheorie Verw. Geb. 4 (1966), 293-315.

8. Renyi, A., Representations for real numbers and their ergodic properties. Acta Math. Acad. Sci. Hungar. 8 (1957), 477-493.

9. Rokhlin, V.A., Exact endomorphisms of a Lebesgue space. Izv. Akad. Nauk. SSSR, Ser. Mat. 25 (1961), 499-530; English translation, Amer. Math. Soc. Transl. Series 2 39 (1964), 1-36.

10. Ryll-Nardzewski, C., On the ergodic theorems (II). Ergodic theory of continued fractions. Studia Math. 12 (1951), 74-79.