Z. Wahrscheinlichkeitstheorie verw. Gebiete 31, 303–328 (1975)

Ergodic Properties of a Class of Piecewise Linear Transformations

Keith M. Wilkinson

1. Introduction

Rényi, in his fundamental paper on ergodic properties of transformations associated with f-expansions [15], showed that the β-transformation $T\omega = \beta\omega \pmod{\text{one}}$, $\beta > 1$, of the unit interval onto itself is ergodic with respect to Lebesgue measure and possesses an invariant measure equivalent to Lebesgue measure. The actual form of the invariant measure was found in [3] and [12]. Rohlin [16] then showed that the β-transformation is exact. In [13], Parry showed that the linear mod one transformation $T\omega = \beta\omega + \alpha \pmod{\text{one}}$, $\beta > 1$, $0 \leqq \alpha < 1$, of the unit interval onto itself possesses an invariant measure and discussed when this measure is equivalent to Lebesgue measure. More recently Shiokawa [17] introduced a class of transformations which generalise the β-transformation, possess an invariant measure equivalent to Lebesgue measure and are exact.

Because of the recent work on the classification problem for measure-preserving transformations it is of interest to decide which endomorphisms are weak Bernoulli since the natural extension [16] of a weak Bernoulli endomorphism is a Bernoulli automorphism (see [7] and [11]). The β-transformation has been shown to be weak Bernoulli in [1] and [18] and the linear mod one transformation with $\beta > 2$ has been shown to be weak Bernoulli in [19].

In this paper we introduce a class $\mathscr{L}$ of piecewise linear transformations of the unit interval onto itself which contains the β-transformation, linear mod one transformation and Shiokawa's generalisation of the β-transformation as special cases. We shall show the existence of an invariant measure for $T \in \mathscr{L}$ which is equivalent to Lebesgue measure and shall discuss a subclass of $\mathscr{L}$ whose members are Markov shifts. We then demonstrate that each $T \in \mathscr{L}$ is weak Bernoulli, our proof being similar to that used in [18] for the β-transformation. We finally look at an example which generalises the linear mod one transformation in essentially the same way that Shiokawa's transformation generalises the β-transformation and discuss two examples of linear mod one transformations which are not in the class $\mathscr{L}$ but are nevertheless weak Bernoulli.

Throughout this paper $(\Omega, \mathscr{B}, \lambda)$ will represent the probability space of Ω the unit interval $[0, 1)$ with $\mathscr{B}$ the Borel subsets of Ω and λ Lebesgue measure on $\mathscr{B}$.

2. The Class $\mathscr{L}$

Throughout this paper $\mathbf{I}$ will represent either a subset of the non-negative integers of the type $\{0, 1, \ldots, N\}$ or the set of non-negative integers itself. We also let b_i, $i \in \mathbf{I}$, be a collection of real numbers satisfying

$$b_i > 0 \quad \text{and} \quad \sum_{i \in \mathbf{I}} b_i = 1.$$

Now let $a_0=0$ and define $a_{i+1}=a_i+b_i$, $i\in\mathbf{I}$. Our transformation $T\colon \Omega\to\Omega$ is then defined for $\omega\in[a_i, a_{i+1})$ by

$$T\omega=\beta_i(\omega-a_i)+\alpha_i \tag{2.1}$$

where $\beta=\inf_{i\in\mathbf{I}}\beta_i>1$ and for each $i\in\mathbf{I}$ $0\leqq\alpha_i<1$, $\gamma_i=\beta_i b_i+\alpha_i\leqq 1$.

Since T is piecewise linear it is a measurable and non-singular transformation of $(\Omega, \mathscr{B}, \lambda)$ to itself.

Define P to be the partition of Ω with atoms $P_i=[a_i, a_{i+1})$, $i\in\mathbf{I}$. If $\lambda(TP_i)=1$ (i.e. if $\alpha_i=0$, $\gamma_i=1$) we say that P_i is a full interval of rank one, otherwise P_i is said to be non-full. For $n\geqq 1$ we let $\bigvee_{k=0}^{n-1}T^{-k}P$ represent the partition of Ω with atoms

$$P_{j_1}\cap T^{-1}P_{j_2}\cap\cdots\cap T^{-(n-1)}P_{j_n}, \quad j_k\in\mathbf{I},\ 1\leqq k\leqq n,$$

and we shall use the notation $\Delta(j_1, j_2, \ldots, j_n)$ or $\Delta(J_n)$, where J_n is understood to mean the vector $(j_1, j_2, \ldots, j_n)$, to represent this atom of $\bigvee_{k=0}^{n-1}T^{-k}P$. We say that an atom $\Delta(J_n)$ of $\bigvee_{k=0}^{n-1}T^{-k}P$ is a full interval of rank n if $\lambda(T^n\Delta(J_n))=1$ and non-full otherwise. Alternatively $\Delta(J_n)\in\bigvee_{k=0}^{n-1}T^{-k}P$ is a full interval of rank n if and only if $\chi_{T^n\Delta(J_n)}(\omega)=1$ λ-a.e. where here and in the sequel $\chi_E(\omega)$ is the indicator function of the set E.

Note that by the linearity of T, if P_i, $i\in\mathbf{I}$, is a full interval of rank one,

$$\lambda(P_i)=b_i=1/\beta_i$$

whereas if P_i is non-full,

$$\lambda(P_i)=b_i<1/\beta_i.$$

Moreover, if the atom $\Delta(j_1, j_2, \ldots, j_n)$ of $\bigvee_{k=0}^{n-1}T^{-k}P$ is full,

$$\lambda(\Delta(j_1, j_2, \ldots, j_n))=1/\beta_{j_1}\beta_{j_2}\ldots\beta_{j_n}$$

whereas if $\Delta(j_1, j_2, \ldots, j_n)$ is non-full,

$$\lambda(\Delta(j_1, j_2, \ldots, j_n))<1/\beta_{j_1}\beta_{j_2}\ldots\beta_{j_n}.$$

Hence all atoms of $\bigvee_{k=0}^{n-1}T^{-k}P$ have Lebesgue measure no greater than $\theta(n)=\beta^{-n}$. This implies that P is a generator for T.

If $\Delta(j_1, j_2, \ldots, j_n)$ is an atom of $\bigvee_{k=0}^{n-1}T^{-k}P$ with $\lambda(\Delta(j_1, j_2, \ldots, j_n))>0$ then the atoms $\Delta(j_1, j_2, \ldots, j_n, i)$, $i\in\mathbf{I}$ of $\bigvee_{k=0}^{n}T^{-k}P$ are subsets of $\Delta(j_1, j_2, \ldots, j_n)$ and

$$\bigcup_{i\in\mathbf{I}}\Delta(j_1, j_2, \ldots, j_n, i)=\Delta(j_1, j_2, \ldots, j_n).$$

Among the sets $\Delta(j_1, j_2, \ldots, j_n, i)$, $i\in\mathbf{I}$, with non-zero Lebesgue measure some will be full and the remainder non-full intervals of rank $n+1$. We shall let $\mathrm{I}(j_1, j_2, \ldots, j_n)$ be the cardinality of the set of $i\in\mathbf{I}$ for which $\lambda(\Delta(j_1, j_2, \ldots, j_n, i))>0$ and $\Delta(j_1, j_2, \ldots, j_n, i)$ is non-full. That is to say $\mathrm{I}(j_1, j_2, \ldots, j_n)$ is the number of non-full intervals of rank $n+1$ which are subsets of $\Delta(j_1, j_2, \ldots, j_n)$. We then let

$$\mathrm{I}_n=\sup\mathrm{I}(j_1, j_2, \ldots, j_n)$$

where the supremum is taken over all $\Delta(j_1, j_2, \ldots, j_n) \in \bigvee_{k=0}^{n-1} T^{-k}P$ with $\lambda(\Delta(j_1, j_2, \ldots, j_n)) > 0$ and we define

$$\mathrm{I} = \sup_{n \geqq 0} \mathrm{I}_n$$

where I_0 is the number of non-full intervals of rank 1. In order to restrict the occurrence of non-full intervals we impose the following condition on our class of transformations:

$$\beta > \mathrm{I}. \tag{2.2}$$

We shall denote by $\mathscr{L}$ the class of transformations of the form (2.1) satisfying (2.2) and in the sequel, unless otherwise stated, we shall always be working with transformations in the class $\mathscr{L}$. Note that (2.2) ensures the existence of full intervals of any rank and that $T \in \mathscr{L}$ is onto.

We end this section with three examples of $T \in \mathscr{L}$. A generalisation of these examples is studied in § 8.

(1) Choose $\beta > 1$ and let $N = [\beta]$ ($[z]$ denotes the integer part of z), $\mathbf{I} = \{0, 1, \ldots, N\}$, $b_i = 1/\beta$, $0 \leqq i \leqq N-1$, $b_N = 1 - N/\beta$, $\beta_i = \beta$, $\alpha_i = 0$, $i \in \mathbf{I}$. Then defining T by (2.1) we obtain the transformation $T\colon \Omega \to \Omega$ given by

$$T\omega = \beta\omega \pmod{\text{one}}.$$

The ergodic properties of this transformation have previously been studied in [1, 3, 12, 15, 16] and [18].

Now all atoms of P with the possible exception of P_N are full. Hence if $\Delta(j_1, j_2, \ldots, j_n)$ is an atom of $\bigvee_{k=0}^{n-1} T^{-k}P$ with non-zero Lebesgue measure and if the atoms of $\bigvee_{k=0}^{n} T^{-k}P$ which are subsets of $\Delta(j_1, j_2, \ldots, j_n)$ with non-zero Lebesgue measure are

$$\Delta(j_1, j_2, \ldots, j_n, j), \quad 0 \leqq j \leqq J = J(j_1, j_2, \ldots, j_n)$$

then the only one of these atoms which may be non-full is $\Delta(j_1, j_2, \ldots, j_n, J)$. Hence $\mathrm{I} \leqq 1 < \beta$ and so $T \in \mathscr{L}$. Note that in the special case of β integral there are no non-full intervals of any rank.

(2) Choose $\beta > 1$ and $0 < \alpha < 1$ and let $N = [\beta + \alpha]$. We distinguish between the two cases (i) $\beta + \alpha = N$ and (ii) $\beta + \alpha > N$. In case (i) we let $\mathbf{I} = \{0, 1, \ldots, N-1\}$, $b_0 = (1-\alpha)/\beta$, $b_i = 1/\beta$, $1 \leqq i \leqq N-1$, $\beta_i = \beta$, $0 \leqq i \leqq N-1$, $\alpha_0 = \alpha$ and $\alpha_i = 0$, $1 \leqq i \leqq N-1$ whereas in case (ii) we let $\mathbf{I} = \{0, 1, \ldots, N\}$, $b_0 = (1-\alpha)/\beta$, $b_i = 1/\beta$, $1 \leqq i \leqq N-1$, $b_N = 1 - (N-\alpha)/\beta$, $\beta_i = \beta$, $0 \leqq i \leqq N$, $\alpha_0 = \alpha$ and $\alpha_i = 0$, $1 \leqq i \leqq N$. In either case, defining T by (2.1), we obtain the transformation T of Ω onto itself defined by

$$T\omega = \beta\omega + \alpha \pmod{\text{one}}.$$

The ergodic properties of this transformation have previously been studied in [13] and [19].

In case (i) each atom of P is full except for P_0. Hence if the atom $\Delta(j_1, j_2, \ldots, j_n)$ of $\bigvee_{k=0}^{n-1} T^{-k}P$ has non-zero Lebesgue measure and the atoms of $\bigvee_{k=0}^{n} T^{-k}P$ with non-zero Lebesgue measure which are subsets of $\Delta(j_1, j_2, \ldots, j_n)$ are

$$\Delta(j_1, j_2, \ldots, j_n, j), \quad J = J(j_1, j_2, \ldots, j_n) \leqq j \leqq N-1,$$

then the only one of these atoms which may be non-full is $\Delta(j_1, j_2, \ldots, j_n, J)$. Hence $\mathrm{I}=1$ and so $T \in \mathscr{L}$.

In case (ii) each atom of P is full except for P_0 and P_N and hence, arguing in a similar way to the above, $\mathrm{I}=2$. Hence we must assume $\beta > 2$ for $T\omega = \beta\omega + \alpha$ (mod one) to be in $\mathscr{L}$ in this case.

(3) Let b_i, $i \in \mathbf{I}$, satisfy the conditions imposed at the beginning of this section and for each $i \in \mathbf{I}$ let $\beta_i = 1/b_i$, $\alpha_i = 0$. If T is defined by (2.1) then all atoms of P are full and hence $\mathrm{I}=0$. Moreover T preserves Lebesgue measure and is isomorphic to the Bernoulli endomorphism with probabilities b_i, $i \in \mathbf{I}$. Any Bernoulli endomorphism with finite or countably infinite state space can be represented in this way.

Note that in the definition of the class $\mathscr{L}$ we restrict attention to piecewise linear transformations with positive slopes β_i, $i \in \mathbf{I}$. We could allow some or all of the β_i, $i \in \mathbf{I}$, to be negative, replacing conditions on the slope by conditions on $|\beta_i|$, $i \in \mathbf{I}$, and obtain all of the results obtained in the sequel. As there is a straightforward isomorphism between the transformation with possibly negative slopes β_i on P_i and the transformation with positive slopes $|\beta_i|$ on P_i, $i \in \mathbf{I}$, we chose not to consider this extra generality further.

3. Relationship of $\mathscr{L}$ with f-Expansions

Following Rényi's paper [15] much has been written about ergodic properties of transformations associated with f-expansions. We shall briefly introduce the terminology of f-expansions and then show that the transformations in the class $\mathscr{L}$ may be considered to be the transformation associated with particular f-expansions.

Let f be a strictly monotonic function mapping a subset D of the non-negative real numbers onto $[0, 1)$. The associated transformation T_f of $[0, 1)$ to itself is defined by

$$T_f\omega = f^{-1}(\omega) \text{ (mod one)}.$$

Because of the monotonicity of f, T_f is a measureable, non-singular transformation of $(\Omega, \mathscr{B}, \lambda)$ to itself. The choice of f is always made so as to ensure that T_f is onto. This is done by insisting that for each $\omega \in [0, 1)$ there is at least one non-negative integer k for which $k + \omega \in D$.

For $n \geqq 1$, we define a stochastic process by

$$X_n(\omega) = [f^{-1}(T_f^{n-1}\omega)],$$

this stochastic process being known as the sequence of digits in the f-expansion of ω. The state space of this stochastic process, denoted by A_f, being known as the set of admissible digits. Note that A_f is a subset of the non-negative integers.

We say that f gives rise to f-expansions with independent digits if, for any $a_1, a_2, \ldots, a_n \in A_f$, there is a point $\omega \in \Omega$ with

$$X_i(\omega) = a_i, \qquad 1 \leqq i \leqq n,$$

otherwise the digits are said to be dependent. This terminology was introduced by Rényi [15] and should not be confused with stochastic independence of the process $X_n(\omega)$, $n \geqq 1$. Note that

$$\omega = f\big(X_1(\omega) = f(X_2(\omega) + \cdots + f(X_n(\omega) + T_f^n \omega) \ldots \big)).$$

We say that f gives rise to valid f-expansions if

$$\begin{aligned}\omega &= \lim_{n\to\infty} f_n\big(X_1(\omega), X_2(\omega), \ldots, X_n(\omega)\big)\\ &\equiv \lim_{n\to\infty} \bar{f}\big(X_1(\omega) + \bar{f}(X_2(\omega) + \cdots + \bar{f}(X_n(\omega)) \ldots)\big)\end{aligned}$$

where $\bar{f}$ is the unique monotonic extension of f to the domain $[0, \infty)$ and range $[0, 1]$. Note that for some $\omega \in \Omega$ there may be $n \geqq 1$ for which $T_f^n \omega = 0$ and in this case

$$\omega = f_n\big(X_1(\omega), X_2(\omega), \ldots, X_n(\omega)\big),$$

the remainder of the digits $X_{n+k}(\omega)$, $k \geqq 1$, representing the digits in the f-expansion of 0. If this is the case we say that ω has a finite f-expansion and we always take n to be minimal.

In order to show that $T \in \mathscr{L}$ fits into the category of transformations derived from f-expansions we let D be the subset of $[0, \infty)$ defined by

$$D = \bigcup_{k \in \mathbf{I}} [k + \alpha_k, k + \gamma_k)$$

and define f on $[k + \alpha_k, k + \gamma_k)$, $k \in \mathbf{I}$, by

$$f(x) = a_k + (x - \alpha_k - k)/\beta_k. \tag{3.1}$$

Since the only possible accumulation point of the points a_k, $k \in \mathbf{I}$, is 1 f is a mapping onto $[0, 1)$. Moreover, since at least one P_k, $k \in \mathbf{I}$, is full, D satisfies the condition to ensure T_f is onto and if $\omega \in [a_k, a_{k+1})$,

$$\begin{aligned}T_f(\omega) &= f^{-1}(\omega) \text{ (mod one)}\\ &= \beta_k(\omega - a_k) + \alpha_k.\end{aligned}$$

The set of admissible digits for these f-expansions is $\mathbf{I}$ and, since for each $k \in \mathbf{I}$

$$P_k = [a_k, a_{k+1}) = \{\omega\colon X_1(\omega) = k\},$$

for any $k_1, k_2, \ldots, k_n \in \mathbf{I}$

$$\Delta(k_1, k_2, \ldots, k_n) = \{\omega\colon X_1(\omega) = k_1, X_2(\omega) = k_2, \ldots, X_n(\omega) = k_n\}.$$

Hence the fact that not all atoms of $\bigvee_{k=0}^{n-1} T^{-k} P$ need be full shows that f-expansions with f as defined in (3.1) need not have independent digits. However since P is a generator for T these f-expansions are valid.

We conclude this section with a lemma which gives a useful expression for the f-expansion of a point $\omega \in \Omega$.

(3.2) **Lemma.** *For* $\omega \in \Omega$,

$$\omega = \sum_{n=0}^{\infty} g\big(X_{n+1}(\omega)\big) B(\omega, n)$$

where $X_n(\omega)$, $n \geqq 1$, *are the digits in the f-expansion of* ω,

$$B(\omega, n) = \begin{cases} 1, & n=0, \\ 1/\beta_{X_1(\omega)}\beta_{X_2(\omega)} \dots \beta_{X_n(\omega)}, & n \geqq 1, \end{cases}$$

and for $k \in \mathbf{I}$

$$g(k) = \tilde{f}(k) - \alpha_k/\beta_k = a_k - \alpha_k/\beta_k.$$

Proof. Since, for all $n \geqq 1$,

$$\omega = f_n(X_1(\omega), X_2(\omega), \dots, X_n(\omega) + T_f^n\,\omega)$$

it will suffice to show that

$$(3.3)\qquad f_n(X_1(\omega), X_2(\omega), \dots, X_n(\omega) + T_f^n\,\omega) = \sum_{k=0}^{n-1} g(X_{k+1}(\omega))\,B(\omega, k) + B(\omega, n)\,T_f^n\,\omega.$$

We prove (3.3) by induction. First note that, by the definition of f in (3.1), (3.3) is true for $n=1$. We now assume that (3.3) is true for $n=1, 2, \dots, N-1$. In particular

$$\begin{aligned} &f_{N-1}(X_2(\omega), X_3(\omega), \dots, X_N(\omega) + T_f^N\,\omega) \\ &\quad = \beta_{X_1(\omega)} \sum_{k=0}^{N-2} g(X_{k+2}(\omega))\,B(\omega, k+1) + \beta_{X_1(\omega)} B(\omega, N)\,T_f^N\,\omega. \end{aligned}$$

Hence,

$$\begin{aligned} f_N(X_1(\omega), X_2(\omega), \dots, X_N(\omega) + T_f^N\,\omega) &= f\big(X_1(\omega) + f_{N-1}(X_2(\omega), \dots, X_N(\omega) + T_f^N\,\omega)\big) \\ &= a_{X_1(\omega)} + \frac{1}{\beta_{X_1(\omega)}}\{f_{N-1}(X_2(\omega), \dots, X_N(\omega) + T_f^N\,\omega)\} \\ &= \sum_{k=0}^{N-1} g(X_{k+1}(\omega))\,B(\omega, k) + B(\omega, N)\,T_f^N\,\omega \end{aligned}$$

and the lemma is proved.

4. Some Properties of Full Intervals

In this section we introduce some results on the way full intervals may be used to approximate other subintervals of Ω.

(4.1) **Lemma.** *For any* $\varepsilon > 0$ *we can find* $k = k(\varepsilon)$ *such that for all* $n \geqq 1$ *we can fill* Ω *to within a set of Lebesgue measure* ε *with disjoint full intervals of ranks between* n *and* $n+k$.

Proof. Fill Ω as far as possible with full intervals of rank n. Suppose that $\Delta(k_1, k_2, \dots, k_n)$ is a non-full interval of rank n. Its Lebesgue measure is smaller than $1/\beta_{k_1}\beta_{k_2} \dots \beta_{k_n}$. If we fill $\Delta(k_1, k_2, \dots, k_n)$ as far as possible with full intervals of rank $n+1$ the remainder will have Lebesgue measure smaller than

$$(\beta_{k_1}\beta_{k_2} \dots \beta_{k_n})^{-1}\,(1/\beta).$$

If we now fill this remainder with full intervals of rank $n+2$, what is left has Lebesgue measure smaller than

$$(\beta_{k_1}\beta_{k_2} \dots \beta_{k_n})^{-1}\,(1/\beta)^2.$$

Continuing, the remainder, after filling $\Delta(k_1, k_2, \ldots, k_n)$ as far as possible with full intervals of ranks $n+1, n+2, \ldots, n+k$, will have Lebesgue measure no larger than

$$(\beta_{k_1} \beta_{k_2} \ldots \beta_{k_n})^{-1} (1/\beta)^k.$$

By condition (2.2) the lemma will be proved if we show that

$$F(n) = \sum (\beta_{k_1} \beta_{k_2} \ldots \beta_{k_n})^{-1} \tag{4.2}$$

is bounded uniformly in n where the sum is over all $\Delta(k_1, k_2, \ldots, k_n)$ with non-zero Lebesgue measure. Now

$$F(n) = \sum^1 (\beta_{k_1} \beta_{k_2} \ldots \beta_{k_n})^{-1} + \sum^2 (\beta_{k_1} \beta_{k_2} \ldots \beta_{k_n})^{-1} \tag{4.3}$$

where $\sum^1$ represents the sum over $\Delta(k_1, k_2, \ldots, k_n)$ which are full and $\sum^2$ represents the sum over $\Delta(k_1, k_2, \ldots, k_n)$ which are non-full and have non-zero Lebesgue measure. But

$$\sum^1 (\beta_{k_1} \beta_{k_2} \ldots \beta_{k_n})^{-1} \leqq 1$$

and

$$\sum^2 (\beta_{k_1} \beta_2 \ldots \beta_{k_n})^{-1} \leqq F(n-1)(1/\beta). \tag{4.4}$$

Hence

$$F(n) \leqq 1 + F(n-1)(1/\beta), \qquad n \geqq 1,$$

where we put $F(0)=1$ and so

$$F(n) \leqq (1-1/\beta)^{-1}, \qquad n \geqq 1,$$

yielding the result.

The proof of this lemma in the special case $n=1$ yields

(4.5) **Corollary.** *Let B_n be the union of those atoms $\Delta(k_1, k_2, \ldots, k_n)$ of $\bigvee_{k=0}^{n-1} T^{-k} P$ which are full but none of $\Delta(k_1, k_2, \ldots, k_m)$, $1 \leqq m \leqq n-1$, is full, then*

$$\sum_{n=1}^{\infty} \lambda(B_n) = 1.$$

(4.6) **Corollary.** *Given any sub-interval E of Ω and $\varepsilon > 0$ we can find $n_1 = n_1(\varepsilon)$ and $k = k(\varepsilon)$ such that for $n \geqq n_1$ we may fill E to within a set of Lebesgue measure ε with disjoint full intervals of ranks between n and $n+k$.*

Proof. Using (4.1) we can find $k(\varepsilon)$ such that for any $n \geqq 1$ we may fill Ω to within a set of Lebesgue measure $\varepsilon/3$ by full intervals of ranks between n and $n+k$. However, the end-points of the interval E may be contained in two of these full intervals. Choosing n_1 so that $\theta(n_1) < \varepsilon/3$ yields the result.

(4.7) **Corollary.** *Any sub-interval E of Ω is an at most countable union of disjoint full intervals (to within a set of Lebesgue measure zero).*

Proof. By (4.6) we can fill E to within a set of Lebesgue measure $1/m$, $m \geqq 2$, by full intervals of ranks between $n_1(1/m)$ and $n_1(1/m) + k(1/m)$. Let the union of these full intervals be denoted by D_m. Then

$$E=\bigcup_{m=2}^{\infty} D_m$$

and each D_m is an at most countable union of full intervals.

The final lemma of this section, while being simple to state and prove is crucial to the remainder of the paper.

(4.8) **Lemma.** *If F is a full interval of rank n and $B\in\mathscr{B}$, then*

$$\frac{\lambda(F\cap T^{-n}B)}{\lambda(F)}=\lambda(B).$$

Proof. T^n *maps* F *linearly onto* Ω *and* $F\cap T^{-n}B$ *onto* B.

5. The Invariant Measure

In this section we shall show that for each $T\in\mathscr{L}$ there is a probability measure μ on $(\Omega,\mathscr{B})$ which is invariant under T and equivalent to Lebesgue measure. We then introduce a version of $d\mu/d\lambda$ and examine its discontinuities. In order to prove the ergodicity of $T\in\mathscr{L}$ with respect to Lebesgue measure we shall need the following result of Knopp [10], a proof of which may be found in [19].

(5.1) **Lemma.** *If $E\in\mathscr{B}$ with $\lambda(E)>0$ and there is a class $\mathscr{I}$ of sub-intervals of Ω such that*

(a) *Every open sub-interval of Ω is an at most countable disjoint union of these sub-intervals to within a set of Lebesgue measure zero, and*

(b) *For each $F\in\mathscr{I}$, $\lambda(E\cap F)\geqq\gamma\lambda(F)$, where $\gamma>0$ is independent of F, then*

$$\lambda(E)=1.$$

Our first theorem in this section generalises a theorem of Rényi [15] for the β-transformation.

(5.2) **Theorem.** *If $T\in\mathscr{L}$ then*

(i) *T is ergodic with respect to λ,*

(ii) *There is a probability measure μ on $(\Omega,\mathscr{B})$ which is invariant with respect to T and such that for each $F\in\mathscr{B}$*

$$(1-1/\beta)\,\lambda(F)\leqq\mu(F)\leqq(1-1/\beta)^{-1}\,\lambda(F). \tag{5.3}$$

Proof. (i) Suppose $T^{-1}E=E\in\mathscr{B}$ and $\lambda(E)>0$. By (4.7) we see that the set $\mathscr{I}$ of all full intervals satisfies (a) of (5.1). Also, by (4.8), for each $F\in\mathscr{I}$,

$$\lambda(E\cap F)=\lambda(E)\,\lambda(F)$$

and so (b) of (5.1) is satisfied with $\gamma=\lambda(E)$. Hence, by (5.1), $\lambda(E)=1$.

(ii) Let $F\in\mathscr{B}$ and $K_n=(k_1,k_2,\ldots,k_n)$ where $k_i\in\mathbf{I}$, $1\leqq i\leqq n$. If $\lambda(\Delta(K_n))>0$, then

$$\lambda(T^{-n}F\cap\Delta(K_n))\leqq\lambda(F)\,(\beta_{k_1}\beta_{k_2}\ldots\beta_{k_n})^{-1},\qquad n\geqq1.$$

Hence,

$$\lambda(T^{-n}F)\leqq\lambda(F)\,F(n)\leqq\lambda(F)\,(1-1/\beta),\qquad n\geqq1, \tag{5.4}$$

where $F(n)$ is defined by (4.2). However, if $\Delta(K_n)$ is a full interval of rank n, then

$$\lambda(T^{-n} F \cap \Delta(K_n)) = \lambda(F)(\beta_{k_1} \beta_{k_2} \dots \beta_{k_n})^{-1}.$$

Hence

$$\lambda(T^{-n} F) \geqq G(n)\lambda(F)$$

where

$$G(n) = \sum (\beta_{k_1} \beta_{k_2} \dots \beta_{k_n})^{-1}$$

and the sum is over all $\Delta(K_n)$ which are full. Now using (4.3) and (4.4) we see that

$$G(n) \geqq F(n) - (1/\beta) F(n-1).$$

Hence for $n \geqq 1$,

(5.5) $$\frac{1}{n} \sum_{k=0}^{n-1} \lambda(T^{-k} F) \geqq \frac{1}{n} \lambda(F) \left\{ 1 + \sum_{k=1}^{n-1} (F(k) - (1/\beta) F(k-1)) \right\} \geqq \lambda(F)(1 - 1/\beta).$$

Thus, combining (5.4) and (5.5), for $n \geqq 1$,

(5.6) $$(1 - 1/\beta)\lambda(F) \leqq \frac{1}{n} \sum_{k=0}^{n-1} \lambda(T^{-k} F) \leqq (1 - 1/\beta)^{-1} \lambda(F).$$

The right hand inequality allows us to use the Dunford-Miller ergodic theorem ([4]) to see that

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \chi_F(T^k \omega)$$

exists for λ-almost all $\omega \in \Omega$. We now define for $F \in \mathscr{B}$

$$\mu_n(F) = \frac{1}{n} \sum_{k=0}^{n-1} \lambda(T^{-k} F) = \int \left(\frac{1}{n} \sum_{k=0}^{n-1} \chi_F(T^k \omega) \right) d\lambda(\omega)$$

and using the Dominated Convergence Theorem of Lebesgue integration we see that

$$\mu(F) = \lim_{n \to \infty} \mu_n(F)$$

exists for all $F \in \mathscr{B}$. Moreover, by the Vitali-Hahn-Saks Theorem ([6] p. 32) μ is a probability measure on $\mathscr{B}$ which because of (5.6) satisfies (5.3). Also

$$\mu_n(T^{-1} F) = \frac{n+1}{n} \mu_{n+1}(F) - \frac{\lambda(F)}{n}$$

and so, letting $n \to \infty$,

$$\mu(T^{-1} F) = \mu(F),$$

i.e. μ is invariant under T.

We now list some straightforward corollaries of this theorem. Note that since μ and λ are equivalent all almost everywhere results can refer to either probability measure. In particular Lemma (4.1) and its corollaries may be stated in terms of the invariant measure μ rather than λ. As we repeatedly use (5.3) we shall let $C=(1-1/\beta)^{-1}$ in the sequel.

(5.7) **Corollary.** *If h is a version of $d\mu/d\lambda$ then*

$$C^{-1}\leqq h(\omega)\leqq C \qquad a.e.$$

(5.8) **Corollary.** *$T\in\mathscr{L}$ is weak mixing.*

Proof. By [9] pp. 39–41 we have to show that the Cartesian product transformation $T\times T$ of the Cartesian product space $(\Omega\times\Omega, \mathscr{B}\times\mathscr{B}, \lambda\times\lambda)$ is ergodic. The proof of this is similar to (i) of (5.2) where we take $\mathscr{I}$ as the set of products of full intervals.

This corollary is based on a remark of Parry [12] with respect to the transformation $T\omega=\beta\omega$ (mod one). The next two corollaries are similar to results proved in [2] for the continued fraction expansion.

(5.9) **Corollary.** *If f is integrable (with respect to μ or λ) then*

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1} f(T^k\omega)=\int f(\omega)\,d\mu(\omega) \qquad a.e.$$

and in particular the asymptotic relative frequency of the digit $i\in\mathbf{I}$ among $X_1(\omega), X_2(\omega), \ldots,$

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1}\chi_{P_i}(T^k\omega)$$

is a.e. equal to $\mu(P_i)$.

Proof. The proof is an application of the Birkhoff Ergodic Theorem.

(5.10) **Corollary.** *$X_n(\omega)\in\mathscr{L}_1(\Omega, \mathscr{B}, \mu)$ (equivalently $X_n(\omega)$ has finite expectation), $n\geqq 1$, if and only if*

$$\sum_{i\in\mathbf{I}} i b_i<\infty.$$

Proof. By (5.3),

$$C^{-1}\sum_{i\in\mathbf{I}} i b_i\leqq\int X_n(\omega)\,d\mu(\omega)=\sum_{i\in\mathbf{I}} i\mu(P_i)\leqq C\sum_{i\in\mathbf{I}} i b_i.$$

(5.11) **Corollary.** *T has finite entropy if and only if*

$$\sum_{i\in\mathbf{I}} b_i\log\beta_i<\infty$$

in which case the entropy is

$$h(T)=\sum_{i\in\mathbf{I}}\mu(P_i)\log\beta_i.$$

Proof. By Parry's formula [14] we see that

$$h(T)=\sum_{i\in\mathbf{I}}\mu(P_i)\log\beta_i$$

which by (5.3) is finite if and only if

$$\sum_{i\in\mathbf{I}} b_i \log \beta_i < \infty.$$

We now give an expression for a version of the Radon-Nykodym derivative $d\mu/d\lambda$. This expression is a particular case of the density of the invariant measure for a class of transformations studied by Fischer [5]. For $n \geqq 1$, we shall let D_n represent the union of all non-full intervals of rank n which are not subsets of full intervals of any lower rank, i.e. D_n is the union of atoms $\Delta(k_1, k_2, \ldots, k_n)$ of $\bigvee_{k=0}^{n-1} T^{-k} P$ such that $\Delta(k_1, k_2, \ldots, k_m)$ is non-full for each $1 \leqq m \leqq n$. We then define

$$\phi_0(\omega) \equiv 1$$

and for $n \geqq 1$

$$\phi_n(\omega) = \sum_{\Delta(K_n) \subset D_n} \chi_{T^n \Delta(K_n)}(\omega)\, B(\omega(K_n), n)$$

where $\omega(K_n) \in \Delta(K_n)$.

(5.12) **Theorem.** *The functions* $\phi_n(\omega)$, $n \geqq 0$, *and*

$$\phi(\omega) = \sum_{n=0}^{\infty} \phi_n(\omega)$$

are Lebesgue integrable and

$$h(\omega) = \phi(\omega) / \int \phi(\omega)\, d\lambda(\omega)$$

is a version of $d\mu/d\lambda$.

Proof. For $n \geqq 1$,

$$\begin{aligned} \int \phi_n(\omega)\, d\lambda(\omega) &= \int \sum_{\Delta(K_n) \subset D_n} \chi_{T^n \Delta(K_n)}(\omega)\, B(\omega(K_n), n)\, d\lambda(\omega) \\ &= \sum_{\Delta(K_n) \subset D_n} B(\omega(K_n), n) \int \chi_{T^n \Delta(K_n)}(\omega)\, d\lambda(\omega) \\ &= \sum_{\Delta(K_n) \subset D_n} \lambda(T^n \Delta(K_n))\, B(\omega(K_n), n) \end{aligned}$$

where the interchange in the order of summation and integration is justified by the Monotone Convergence Theorem of Lebesgue integration. Now there are at most l^n atoms of $\bigvee_{k=0}^{n-1} T^{-k} P$ which are subsets of D_n. Hence

$$\int \phi_n(\omega)\, d\lambda(\omega) < (l/\beta)^n$$

Moreover, again using the Monotone Convergence Theorem,

$$\int \phi(\omega)\, d\lambda(\omega) < \sum_{n=0}^{\infty} (l/\beta)^n = (1 - l/\beta)^{-1}.$$

In order to show that $h(\omega)$ is a version of $d\mu/d\lambda$ we have to show, by a result of Parry [13], that

$$\phi(\omega) = \sum_{k\in\mathbf{I}} \chi_{TP_k}(\omega)\, \phi(\bar{f}(\omega + k))\, |\bar{f}'(\omega + k)| \qquad \lambda\text{-a.e.}$$

where f is as defined by (3.1). But

$$\begin{aligned}
&\sum_{k\in \mathbf{I}} \chi_{TP_k}(\omega)\, \phi_n(\bar{f}(\omega+k))\,|\bar{f}'(\omega+k)| \\
&\quad = \sum_{k\in \mathbf{I}} \sum_{\Delta(K_n)\subset D_n} \chi_{TP_k}(\omega)\, \chi_{T^n\Delta(K_n)}(\bar{f}(\omega+k))\, B(\omega(K_n), n)\, \beta_k^{-1} \\
&\quad = \sum_{\Delta(K_{n+1})\subset D_{n+1}} \chi_{T^{n+1}\Delta(K_{n+1})}(\omega)\, B(\omega(K_{n+1}), n+1) \\
&\qquad + \sum_{\Delta(K_{n+1})\subset B_{n+1}} B(\omega(K_{n+1}), n+1) \\
&\quad = \phi_{n+1}(\omega) + \lambda(B_{n+1})
\end{aligned}$$

where B_n is as defined in (4.5). Using (4.5) now gives the result.

In the sequel it will be useful to approximate $h(\omega)$ by

$$h_m(\omega) = \sum_{n=0}^{m} \phi_n(\omega)/\int \phi(\omega)\, d\lambda(\omega).$$

Note that $\phi_n(\omega)$ is a step function with discontinuities at the end-points of the intervals $T^n\Delta(K_n)$, $\Delta(K_n)\subset D_n$. Hence $h_m(\omega)$ is a step function with at most

$$2(\mathrm{I}+\mathrm{I}^2+\cdots+\mathrm{I}^m)=\tau(m)$$

discontinuities.

Moreover

$$\begin{aligned}
h(\omega)-h_m(\omega) &= \sum_{n=m+1}^{\infty} \phi_n(\omega)/\int \phi(\omega)\, d\lambda(\omega) \\
&\leqq \sum_{n=m+1}^{\infty} (\mathrm{I}/\beta)^n/\int \phi(\omega)\, d\lambda(\omega) = \rho(m).
\end{aligned}$$

We can now extend the results on full-intervals of § 4 to full-intervals which contain no discontinuities of $h_m(\omega)$.

(5.13) **Lemma.** *For given $\varepsilon>0$ and $m\geqq 1$ we can find $k=k(\varepsilon,m)$ and $\hat{n}=\hat{n}(\varepsilon,m)$ such that for $n\geqq\hat{n}$ we can fill Ω to within a set of Lebesgue measure (equivalently μ-measure) ε with disjoint full intervals of ranks between n and $n+k$ on each of which $h_m(\omega)$ is constant.*

Proof. Choose $n_0=n_0(\varepsilon,m)$ so that $\theta(n_0)<\varepsilon/2\tau(m)$. Taking away from Ω those atoms of $\bigvee_{i=0}^{n_0-1} T^{-i}P$ which contain discontinuities in $h_m(\omega)$ leaves at most $\tau(m)+1$ intervals on each of which $h_m(\omega)$ is constant. Name them

$$A_i,\ 1\leqq i\leqq \tau(m)+1$$

where some A_i are empty if necessary. Using (4.6) we may find $n_i(\varepsilon,m)$, $k_i(\varepsilon,m)$ such that for $n\geqq n_i(\varepsilon,m)$ we may fill A_i to within a set of Lebesgue measure

$$\varepsilon/2(\tau(m)+1)$$

with disjoint full intervals of ranks between n and $n+k_i(\varepsilon,m)$. Taking

$$\hat{n}(\varepsilon,m) = \max_{0\leqq i\leqq \tau(m)+1} n_i(\varepsilon,m), \qquad k(\varepsilon,m) = \max_{1\leqq i\leqq \tau(m)+1} k_i(\varepsilon,m)$$

yields the result.

In [16] Rohlin shows that the β-transformation is exact. We extend this result to $T\in\mathscr{L}$ using the following criterion of exactness.

(5.14) **Theorem** ([16]). *Let $\mathscr{I}$ be a countable system of subsets of Ω of positive measure such that the unions of disjoint sets $A\in\mathscr{I}$ generate $\mathscr{B}$. If there exists a positive integer-valued function $n(A)$, $A\in\mathscr{I}$, and a positive number q such that $\mu(T^{n(A)}A)=1$ for all $A\in\mathscr{I}$ and for any Borel set $B\subset A$ with measurable image $T^{n(A)}B$*

$$\mu(T^{n(A)}B)\leqq q\mu(B)/\mu(A)$$

then T is an exact endomorphism.

(5.15) **Theorem.** *$T\in\mathscr{L}$ is exact.*

Proof. We use (5.14) with $\mathscr{I}$ the set of all full intervals and $n(A)$ the rank of the full interval $A\in\mathscr{I}$. Let $B\subset A$, where $A\in\mathscr{I}$, and let $B'=T^{n(A)}B\in\mathscr{B}$. Then, using (5.3) and (4.8),

$$\begin{aligned}\mu(B')&\leqq C\lambda(B')\\&=C\lambda(A\cap T^{-n(A)}B')/\lambda(A)\\&=C\lambda(B)/\lambda(A)\\&\leqq C^3\mu(B)/\mu(A).\end{aligned}$$

Hence the hypotheses of (5.14) are satisfied with $q=C^3$.

6. Markov Properties

Cigler [3] and Smorodinsky [18] have studied the Markov properties of $T\omega=\beta\omega$ (mod one) where β is of a particular form. Shiokawa [17] has used Cigler's methods to obtain similar results for his generalisation of the β-transformation. (This generalisation will be discussed in § 8.) We generalise these results to $T\in\mathscr{L}$ where restrictions are placed on the images of the end points of the atoms of P. The method employed is based on that of Cigler [3]. The first two lemmas give equivalent statements of the conditions used and also a useful consequence of these conditions.

(6.1) **Lemma.** *In the following, for fixed $i\in\mathbf{I}$ with $\alpha_i>0$,* (i) *and* (ii) *are equivalent and both imply* (iii):

(i) $T^m\alpha_i=0$, $T^k\alpha_i>0$, $1\leqq k\leqq m-1$,

(ii) $\alpha_i=f_m(X_1(\alpha_i), X_2(\alpha_i), \ldots, X_m(\alpha_i))$, $\alpha_i\neq f_k(X_1(\alpha_i), X_2(\alpha_i), \ldots, X_k(\alpha_i))$, $1\leqq k\leqq m-1$,

(iii) *α_i is the end-point of an interval of rank m and hence also the end-point of an interval of any rank greater than m.*

Proof. (i) and (ii) are equivalent from the fact that

$$\alpha_i=f_k(X_1(\alpha_i), X_2(\alpha_i), \ldots, X_k(\alpha_i)+T^k\alpha_i),\qquad k\geqq 1.$$

To show that (ii) implies (iii) we note that

$$\begin{aligned}\Delta(X_1(\alpha_1), X_2(\alpha_i), \ldots, X_m(\alpha_i))\subset[&f_m(X_1(\alpha_i), X_2(\alpha_i), \ldots, X_m(\alpha_i)),\\&f_m(X_1(\alpha_i), X_2(\alpha_i), \ldots, X_m(\alpha_i)+1).\end{aligned}$$

But since $\alpha_i \in \Delta(X_1(\alpha_i), X_2(\alpha_i), \ldots, X_m(\alpha_i))$, the left-hand endpoint of $\Delta(X_1(\alpha_i), X_2(\alpha_i), \ldots, X_m(\alpha_i))$ is $\alpha_i = f_m(X_1(\alpha_i), X_2(\alpha_i), \ldots, X_m(\alpha_i))$.

(6.2) **Lemma.** *In the following, for fixed* $i \in \mathbf{I}$ *with* $\gamma_i < 1$, (i) *and* (ii) *are equivalent and both imply* (iii):

(i) $T^n \gamma_i = 0$, $T^k \gamma_i > 0$, $1 \leqq k \leqq n-1$,

(ii) $\gamma_i = f_n(X_1(\gamma_i), X_2(\gamma_i), \ldots, X_n(\gamma_i))$, $\gamma_i \neq f_k(X_1(\gamma_i), X_2(\gamma_i), \ldots, X_k(\gamma_i))$, $1 \leqq k \leqq n-1$,

(iii) γ_i *is the end-point of an interval of rank n and hence also the end-point of an interval of any rank greater than n.*

Proof. The proof is identical to that of (6.1).

Our aim is to prove

(6.3) **Theorem.** *Suppose for each* $i \in \mathbf{I}$

(a) *Either* $\alpha_i = 0$ (*in which case we put* $m_i = 0$) *or for some* $1 \leqq m_i < \infty$ $T^{m_i} \alpha_i = 0$ *and* $T^k \alpha_i > 0$ *for* $1 \leqq k \leqq m_i - 1$, *and*

(b) *Either* $\gamma_i = 1$ (*in which case we put* $n_i = 0$) *or for some* $1 \leqq n_i < \infty$ $T^{n_i} \gamma_i = 0$ *and* $T^k \gamma_i > 0$ *for* $1 \leqq k \leqq n_i - 1$ *then, if* $m = \sup_{i \in \mathbf{I}} (m_i, n_i)$ *is finite, T is an m-step Markov chain with respect to P (a 0-step Markov chain is a Bernoulli shift).*

Note that because of (6.1) and (6.2) conditions (a) and (b) of (6.3) are equivalent to assuming that each $\alpha_i \neq 0$ has finite f-expansion of length m_i and each $\gamma_i \neq 1$ has finite f-expansion of length n_i. In order to prove (6.3) we need the following lemma.

(6.4) **Lemma.** *Under the assumptions of* (6.3), *if* $\lambda(\Delta(k_1, k_2, \ldots, k_n)) > 0$ *then for* $n \geqq m$,

(6.5) $\lambda(\Delta(k_1, k_2, \ldots, k_n)) = (\beta_{k_1} \beta_{k_2} \cdots \beta_{k_{n-m}})^{-1} \lambda(\Delta(k_{n-m+1}, k_{n-m+2}, \ldots, k_n))$

where for $n = m$ *the first term on the right hand side is equal to* 1.

Proof. We use mathematical induction. For $n = m$ (6.5) is clearly true. Suppose (6.5) is true for $n-1$. Then, in particular

$$\lambda(\Delta(k_2, k_3, \ldots, k_n)) = (\beta_{k_2} \beta_{k_3} \cdots \beta_{k_{n-m}})^{-1} \lambda(\Delta(k_{n-m+1}, k_{n-m+2}, \ldots, k_n)).$$

Hence it remains to show

(6.6) $\lambda(\Delta(k_1, k_2, \ldots, k_n)) = \beta_{k_1}^{-1} \lambda(\Delta(k_2, k_3, \ldots, k_n))$ for $n \geqq m+1$.

Suppose first of all that $\Delta(k_1)$ is a full interval. Then T maps $\Delta(k_1, k_2, \ldots, k_n)$ linearly onto $\Delta(k_2, k_3, \ldots, k_n)$ and (6.6) is proved.

Now suppose $\Delta(k_1)$ is non-full. Then $T\Delta(k_1) = [\alpha_{k_1}, \gamma_{k_1})$. But since α_{k_1} is the left-hand end-point of an interval of rank $m_{k_1} \leqq m \leqq n-1$ and γ_{k_1} is the right-hand end-point of an interval of rank $n_{k_1} \leqq m \leqq n-1$, we must have

$$\Delta(k_2, k_3, \ldots, k_n) \subset [\alpha_{k_1}, \gamma_{k_1})$$

and since T maps $\Delta(k_1, k_2, \ldots, k_n)$ linearly onto $\Delta(k_2, k_3, \ldots, k_n)$ (6.6) is proved.

Proof of (6.3). Recall that h, the density of μ, has discontinuities only at the points $T^n \alpha_i$, $T^n \gamma_i$, $n \geqq 1$, $i \in \mathbf{I}$. Hence the assumptions of the theorem ensure that h

is constant on intervals of rank no less than m. Hence for $n \geqq 0$, if $\lambda(\Delta(k_1, k_2, \ldots, k_{m+n})) > 0$,

$$\begin{aligned}\mu(X_{m+n+1}(\omega) = k_{m+n+1} \mid X_1(\omega) = k_1, X_2(\omega) = k_2, \ldots, X_{m+n}(\omega) = k_{m+n}) \\ = \mu(\Delta(k_1, k_2, \ldots, k_{m+n}, k_{m+n+1}))/\mu(\Delta(k_1, k_2, \ldots, k_{m+n})) \\ = \lambda(\Delta(k_1, k_2, \ldots, k_{m+n}, k_{m+n+1}))/\lambda(\Delta(k_1, k_2, \ldots, k_{m+n})) \\ = \lambda(\Delta(k_{n+1}, k_{n+2}, \ldots, k_{m+n}, k_{m+n+1}))/\lambda(\Delta(k_{n+1}, k_{n+2}, \ldots, k_{m+n}))\end{aligned}$$

by (6.5). Since the last term depends only on $k_{n+1}, k_{n+2}, \ldots, k_{m+n}, k_{m+n+1}$, we see that T is an m-step Markov chain.

We showed in (5.15) that $T \in \mathscr{L}$ is exact and hence $T \in \mathscr{L}$ is mixing. Since a mixing m-step Markov chain is weak Bernoulli (the case $m = 1$ is demonstrated in [7], the case $m \geqq 2$ follows similarly) we see that if the hypotheses of (6.3) are satisfied then T is weak Bernoulli. We shall show in the next section that T is weak Bernoulli for each $T \in \mathscr{L}$.

7. $T \in \mathscr{L}$ is Weak Bernoulli

In the proof that T is weak Bernoulli we shall use the approximation of $h(\omega)$ by $h_m(\omega)$. Corresponding to $h_m(\omega)$ we define an approximation to μ by

$$\mu_m(F) = \int_F h_m(\omega)\, d\lambda(\omega), \qquad F \in \mathscr{B}.$$

The first lemma provides a useful estimate involving the convergence of μ_m to μ.

(7.1) **Lemma.** *Given $\varepsilon > 0$ we can find $M = M(\varepsilon)$ such that for $m > M$*

$$\left| \frac{\mu_m(E \cap F)}{\mu_m(E)} - \frac{\mu(E \cap F)}{\mu(E)} \right| < \varepsilon \frac{\mu(E \cap F)}{\mu(E)}$$

for any $E, F \in \mathscr{B}$.

Proof. Define $\alpha(\omega, m) = (h(\omega) - h_m(\omega))/\rho(m)$ and note that $0 \leqq \alpha(\omega, m) \leqq 1$. Then

$$\begin{aligned}\left| \frac{\mu_m(E \cap F)}{\mu_m(E)} - \frac{\mu(E \cap F)}{\mu(E)} \right| &= \left| \frac{\int_{E \cap F} (h(\omega) - \alpha(\omega, m)\rho(m))\, d\lambda(\omega)}{\int_E (h(\omega) - \alpha(\omega, m)\rho(m))\, d\lambda(\omega)} - \frac{\mu(E \cap F)}{\mu(E)} \right| \\ &= \left| \frac{\mu(E \cap F) - \alpha_1 \rho(m) \lambda(E \cap F)}{\mu(E) - \alpha_2 \rho(m) \lambda(E)} - \frac{\mu(E \cap F)}{\mu(E)} \right| \\ &\qquad \text{for some } 0 \leqq \alpha_i \leqq 1,\ i = 1, 2, \\ &= \left| \frac{\alpha_2 \rho(m) \lambda(E) \mu(E \cap F) - \alpha_1 \rho(m) \lambda(E \cap F) \mu(E)}{(\mu(E) - \alpha_2 \rho(m) \lambda(E)) \mu(E)} \right| \\ &\leqq \frac{2\rho(m)}{(C^{-1} - \rho(m))} \frac{\mu(E \cap F)}{\mu(E)}\end{aligned}$$

providing m is large enough for $\rho(m) < C^{-1}$. The fact that $\rho(m) \to 0$ as $m \to \infty$ now yields the result.

In (5.2) we expressed the invariant measure of a Borel set F as the Cesàro limit of the sequence $\lambda(T^{-n}F)$. Our next result shows that ordinary convergence takes place and moreover this convergence is uniform over partitions.

(7.2) **Lemma.** *Let Q be a finite or countable partition of Ω. Given $\varepsilon>0$ we can find $L=L(\varepsilon)$ such that for $n>L$*

$$\sum_{F\in Q}|\lambda(T^{-n}F)-\mu(F)|<\varepsilon.$$

Proof. Put $L_1=L_1(\varepsilon)=M(\varepsilon/2)$, where M is as in (7.1). Let E be a full interval of rank r, say, such that h_{L_1} is constant on E. Once chosen we keep the set E fixed. Let $\mathscr{F}_m$ be the σ-algebra generated by $\bigcup_{i=m}^{\infty} T^{-i}P$. Using Doob's Martingale Theorem ([2] p. 121) and the fact that by (5.15) $\bigcap_{m=1}^{\infty}\mathscr{F}_m$ is trivial we know that

$$\mu(E|\mathscr{F}_m)(\omega)\to\mu(E) \quad \text{a.e.}$$

as $m\to\infty$. Hence, by Egoroff's Theorem ([8] p. 88), for any $\varepsilon'>0$ we can find a set D with $\mu(D)<\varepsilon'$ such that on $\Omega\setminus D$, $\mu(E|\mathscr{F}_m)(\omega)$ converges uniformly, i.e. we can find $L_2=L_2(\varepsilon')$ such that for $m>L_2$,

$$|\mu(E|\mathscr{F}_m)(\omega)-\mu(E)|<\varepsilon' \qquad \text{for } \omega\notin D.$$

Now, since $h_{L_1}(\omega)$ is constant on E, by (4.8)

$$\mu_{L_1}(E\cap T^{-r-n}F)/\mu_{L_1}(E)=\lambda(E\cap T^{-r-n}F)/\lambda(E)=\lambda(T^{-n}F)$$

for all $F\in\mathscr{B}$. Moreover,

$$\mu(E\cap T^{-r-n}F)/\mu(E)=\frac{1}{\mu(E)}\int_{T^{-r-n}F}\mu(E|\mathscr{F}_{r+n})(\omega)h(\omega)\,d\lambda(\omega).$$

Hence,

$$\sum_{F\in Q}|\lambda(T^{-n}F)-\mu(F)|\leqq\sum_{F\in Q}\left|\frac{\mu_{L_1}(E\cap T^{-r-n}F)}{\mu_{L_1}(E)}-\frac{\mu(E\cap T^{-r-n}F)}{\mu(E)}\right|$$
$$+\sum_{F\in Q}\left|\frac{1}{\mu(E)}\int_{T^{-r-n}F}\mu(E|\mathscr{F}_{r+n})(\omega)h(\omega)\,d\lambda(\omega)-\mu(F)\right|.$$

Now, by the choice of L_1 and (7.1), the first term on the right hand side is smaller than $\varepsilon/2$. The second term is no larger than

$$\frac{1}{\mu(E)}\sum_{F\in Q}\int_{T^{-r-n}F}|\mu(E|\mathscr{F}_{r+n})(\omega)-\mu(E)|\,h(\omega)\,d\lambda(\omega)$$
$$=\frac{1}{\mu(E)}\sum_{F\in\Omega}\int_{T^{-r-n}F\cap(\Omega\setminus D)}|\mu(E|\mathscr{F}_{r+n})(\omega)-\mu(E)|\,h(\omega)\,d\lambda(\omega)$$
$$+\frac{1}{\mu(E)}\sum_{F\in Q}\int_{T^{-r-n}F\cap D}|\mu(E|\mathscr{F}_{r+n})(\omega)-\mu(E)|\,h(\omega)\,d\lambda(\omega)$$
$$\leqq\varepsilon'\mu(\Omega\setminus D)/\mu(E)+\mu(D)/\mu(E)$$
$$\leqq 2\varepsilon'/\mu(E)$$

for $n>L_2$. Hence, putting $\varepsilon'=\varepsilon\mu(E)/4$ and $L=\max(L_1,L_2)$ yields the result.

We now prove the main theorem of the paper.

(7.3) **Theorem.** *$T\in L$ is weak Bernoulli.*

Proof. We need to show that for any $\varepsilon>0$,

$$D\left(\bigvee_{i=q+n}^{q+2n}T^{-i}P,\ \bigvee_{i=0}^{n}T^{-i}P\right)<\varepsilon \tag{7.4}$$

where q depends only on ε and for any two partitions Q, R of $(\Omega, \mathscr{B}, \mu)$, $D(Q,R)$ is defined by

$$D(Q,R) = \sum_{\substack{A \in Q \\ B \in R}} |\mu(A \cap B) - \mu(A)\mu(B)|.$$

In order to establish (7.4) we shall use the approximation of μ discussed in (7.1) and the approximation of intervals by unions of full intervals.

Put $\varepsilon' = \varepsilon/6$ and put $M = M(\varepsilon')$, where M is a defined in (7.1). This fixes our approximation of μ.

Using (5.13) we can find $k = k(\varepsilon', M)$ and $\hat{n} = \hat{n}(\varepsilon', M)$ such that for $n \geqq \hat{n}$ we can fill Ω to within a set of μ-measure ε' with full intervals of ranks between n and $n+k$ on each of which $h_M(\omega)$ is constant. The set of these full intervals will be denoted by $\mathscr{F}_{n,k}$. Clearly we only need to prove (7.4) for $n \geqq \hat{n}$.

For each $n \geqq \hat{n}$ we approximate $\bigvee_{i=0}^{n} T^{-i} P$ by a partition Q_n formed in the following way. If R is an atom of $\bigvee_{i=0}^{n} T^{-i} P$ then the corresponding atom $\hat{R}$ of Q_n is formed by taking the union of all the sets in $\mathscr{F}_{n,k}$ which are subsets of R. The remainder of Ω when an atom $\hat{R}$ has been formed corresponding to each $R \in \bigvee_{i=0}^{n} T^{-i} P$ will also be an atom of Q_n. This last atom, because of the nature of $\mathscr{F}_{n,k}$, will have μ-measure less than ε'.

Now let $F \in \mathscr{F}_{n,k}$, of rank $n + j(F)$, say, $0 \leqq j(F) \leqq k$ and let $B = T^{-n-q} B'$, where $B' \in \bigvee_{i=0}^{n} T^{-i} P$ and q is to be chosen later. Then (4.8) implies that

$$\frac{\mu_M(F \cap B)}{\mu_M(F)} = \frac{\lambda(F \cap T^{-n-q} B')}{\lambda(F)} = \lambda(T^{-q+j(F)} B').$$

Hence

$$D\left(Q_n, \bigvee_{i=q+n}^{q+2n} T^{-i} P\right) \leqq \sum \mu(F) \left| \frac{\mu(F \cap B)}{\mu(F)} - \frac{\mu_M(F \cap B)}{\mu_M(F)} \right| + \sum \mu(F) |\lambda(T^{-q+j(F)} B') - \mu(B')| + 2\varepsilon'$$

where the sums on the right hand side are over $F \in \mathscr{F}_{n,k}$ and $B \in \bigvee_{i=q+n}^{q+2n} T^{-i} P$ (equivalently $B' \in \bigvee_{i=0}^{n} T^{-i} P$). Now the first term on the right hand side of the above inequality is smaller than ε' by the choice of M and the second term is smaller than ε' provided q is larger than $L(\varepsilon') + k$ where L is as defined in (7.3). Now

$$\begin{aligned}
D\left(\bigvee_{i=q+n}^{q+2n} T^{-i} P, \bigvee_{i=0}^{n} T^{-i} P\right) &= \sum |\mu(R \cap B) - \mu(R)\,\mu(B)| \\
&\leqq \sum |\mu(R \cap B) - \mu(\hat{R} \cap B)| \\
&\quad + \sum |\mu(\hat{R} \cap B) - \mu(\hat{R})\,\mu(B)| \\
&\quad + \sum |\mu(\hat{R})\,\mu(B) - \mu(R)\,\mu(B)| \\
&\leqq D\left(Q_n, \bigvee_{i=q+n}^{q+2n} T^{-i} P\right) \\
&\quad + \sum |\mu(R \cap B) - \mu(\hat{R} \cap B)| \\
&\quad + \sum |\mu(\hat{R})\,\mu(B) - \mu(R)\,\mu(B)| \\
&\leqq 6\varepsilon' = \varepsilon
\end{aligned}$$

by the choice of Q_n. Each of the sums on the right hand side is over

$$B \in \bigvee_{i=q+n}^{q+2n} T^{-i} P \quad \text{and} \quad R \in \bigvee_{i=0}^{n} T^{-i} \bar{P}$$

(with the corresponding $\hat{R} \in Q_n$). Hence the theorem is proved.

8. An Example

In this section we shall introduce an example of a transformation in $\mathcal{L}$ which is a generalisation of the following transformations:

(i) The β-transformation, $T\omega = \beta\omega$ (mod one), $\beta > 1$, whose ergodic properties have been studied in [1, 3, 12, 15, 16] and [18].

(ii) Shiokawa's generalisation of the β-transformation [17].

(iii) Linear mod one transformations, $T\omega = \beta\omega + \alpha$ (mod one), $\beta > 1$, $0 \leqq \alpha < 1$, whose ergodic properties have been studied in [13] and [19].

In this example we take $\mathbf{I} = \{0, 1, 2, \ldots, N\}$ and choose $a_i, b_i, i \in \mathbf{I}$, as outlined in § 2. For $1 \leqq i \leqq N-1$, we set $\beta_i = 1/b_i, \alpha_i = 0$ (and hence $\gamma_i = 1$), for $i = 0$ we set $\alpha_0 = \alpha$ $(0 \leqq \alpha < 1)$ and $\beta_0 = (1-\alpha)/b_0$ (hence $\gamma_0 = 1$) and for $i = N$ we set

$$\alpha_N = 0, \qquad \beta_N \leqq 1/b_N$$

(hence $\gamma = \gamma_N \leqq 1$). Then, defining T by (2.1) we obtain

$$T\omega = \begin{cases} \beta_0\,\omega + \alpha, & \omega \in P_0 \\ \beta_i(\omega - a_i), & \omega \in P_i,\ 1 \leqq i \leqq N, \end{cases} \tag{8.1}$$

where $P_i = [a_i, a_{i+1})$.

We obtain the β-transformation by setting $\alpha = 0$ and choosing

$$b_i = b, \quad i = 0, 1, \ldots, N-1, \qquad b_N \leqq b.$$

Then we put $\beta_i = \beta = 1/b$ for each $i \in \mathbf{I}$. We obtain Shiokawa's generalisation of the β-transformation just by setting $\alpha = 0$ and we obtain linear mod one transformations by setting $b_i = b$, $1 \leqq i \leqq N-1$, $b_0 \leqq b$, $b_N \leqq b$ and letting $\beta_i = \beta = 1/b$ for each $i \in \mathbf{I}$.

The intervals P_i are clearly full for $1 \leqq i \leqq N-1$, P_0 is full if and only if $\alpha = 0$ and P_N is full if and only if $\gamma = 1$. Hence we find that

$$\mathrm{l} = \begin{cases} 0 & \alpha = 0,\ \gamma = 1, \\ 1 & \alpha = 0,\ \gamma < 1 \text{ or } \alpha > 0,\ \gamma = 1, \\ 2 & \alpha > 0,\ \gamma < 1, \end{cases} \tag{8.2}$$

and so $T \in \mathcal{L}$ provided $\beta = \min_{i \in I} \beta_i > 1$ except when both $\alpha > 0$ and $\gamma < 1$ when we must insist $\beta > 2$. Wherever we discuss T as defined by (8.1) in the sequel we shall assume that the conditions for $T \in \mathcal{L}$ are satisfied. By (7.3) T as defined in (8.1) is weak Bernoulli. This has already been noted for the β-transformation in [1] and [18] and for linear mod one transformations in [19]. In the last paper it was assumed that $\beta > 2$ whereas as we have seen we only need to assume $\beta > 1$ if $\gamma = 1$.

Note that if $\alpha = 0$, $\gamma = 1$, T is a Bernoulli shift. If $\alpha = 0$, $\gamma < 1$ and $T^m \gamma = 0$, $T^j \gamma > 0$, $1 \leqq j \leqq m-1$, or if $\alpha > 0$, $\gamma = 1$ and $T^m \alpha = 0$, $T^j \alpha > 0$, $1 \leqq j \leqq m-1$, then T

is an m-step Markov shift. If $\alpha>0$ and $\gamma<1$ then T is an m-step Markov shift if $T^k\alpha=0$, $T^j\alpha>0$, $1\leqq j\leqq k-1$, $T^l\gamma=0$, $T^j\gamma>0$, $1\leqq j\leqq l-1$ and $m=\max(k,l)$.

A version of the density of the invariant measure for the β-transformation was introduced in [3] and [12] and that for linear mod one transformations in [13]. In [17] Shiokawa gives a version of the density of the invariant measure for his generalisation of the β-transformation. We give here a version of the density of the invariant measure for the class under consideration.

(8.3) **Theorem.** *Let T be as defined in* (8.1) *and if $\alpha=0$, $\gamma=1$, let*

$$h(\omega)=1,$$

if $\alpha=0$, $\gamma<1$, let

$$h(\omega)=1+\frac{1}{\beta_N}\sum_{n=0}^{\infty}B(\gamma,n)\,\chi_{[0,T^n\gamma)}(\omega),$$

if $\alpha>0$, $\gamma=1$, let

$$h(\omega)=1+\frac{1}{\beta_0}\sum_{n=0}^{\infty}B(\alpha,n)\,\chi_{[T^n\alpha,1)}(\omega),$$

and if $\alpha>0$, $\gamma<1$, let

$$h(\omega)=1+\frac{D_1}{\beta_N}\sum_{n=0}^{\infty}B(\gamma,n)\,\chi_{[0,T^n\gamma)}(\omega)+\frac{D_2}{\beta_0}\sum_{n=0}^{\infty}B(\alpha,n)\,\chi_{[T^n\alpha,1)}(\omega) \tag{8.4}$$

where D_1, D_2 are the solutions of

$$D_1=1+\frac{D_2}{\beta_0}B(\alpha),\qquad D_2=1+\frac{D_1}{\beta_N}B(\gamma) \tag{8.5}$$

and

$$B(\alpha)=\sum_{n=0}^{\infty}B(\alpha,n),\qquad B(\gamma)=\sum_{n=0}^{\infty}B(\gamma,n).$$

Then

$$\mu(F)=\int_F h(\omega)\,d\lambda(\omega),\qquad F\in\mathscr{B},$$

defines a finite measure (not necessarily normalised) which is equivalent to λ and invariant under T.

Proof. We first note that since

$$0<B(\gamma)/\beta_N<1/(\beta-1)\quad\text{and}\quad 0<B(\alpha)/\beta_0<1/(\beta-1)$$

and since in the case $\alpha>0$, $\gamma<1$ we assume $\beta>2$, the solutions of (8.5)

$$D_1=\frac{1+B(\alpha)/\beta_0}{1-B(\gamma)\,B(\alpha)/\beta_0\,\beta_N},\qquad D_2=\frac{1+B(\gamma)/\beta_N}{1-B(\gamma)\,B(\alpha)/\beta_0\,\beta_N}$$

are both positive and bounded. Hence

$$1\leqq h(\omega)\leqq D_1\,\beta/(\beta-1)$$

ensures the finiteness of μ and the equivalence of μ and λ in this case.

The equivalence of μ and λ in the other cases is immediate. We shall prove the invariance of μ under T only in the case $\alpha>0$, $\gamma<1$, the other cases being proved similarly. We use the relationship of Parry [13] already introduced in

(5.12). We first look at

$$\sum_{k=0}^{N} \chi_{[0,\, T^n\gamma)}(\bar{f}(\omega+k))\, \chi_{T P_k}(\omega)\, \frac{1}{\beta_k} = \sum_{k=0}^{N} \chi_{[0,\, T^n\gamma)\cap P_k}(\bar{f}(\omega+k))\, \frac{1}{\beta_k}.$$

Now

$$[0, T^n\gamma)\cap P_k = \begin{cases} P_k, & 0 \leqq k < X_{n+1}(\gamma), \\ [a_{X_{n+1}(\gamma)}, T^n\gamma), & k = X_{n+1}(\gamma), \\ \Phi, & X_{n+1}(\gamma) < k \leqq N, \end{cases}$$

where Φ represents the empty set.

Hence

$$\sum_{k=0}^{N} \chi_{[0,\, T^n\gamma)}(\bar{f}(\omega+k))\, \chi_{T P_k}(\omega)\, \frac{1}{\beta_k} = \frac{1}{\beta_{X_{n+1}(\gamma)}}\, \chi_{[0,\, T^{n+1}\gamma)}(\omega) + \frac{1}{\beta_0}\, \chi_{[\alpha,\, 1)}(\omega) + \sum_{k=1}^{X_{n+1}(\gamma)} \frac{1}{\beta_k} - \frac{1}{\beta_{X_{n+1}(\gamma)}}.$$

Similarly

$$\sum_{k=0}^{N} \chi_{[T^n\alpha,\, 1)}(\bar{f}(\omega+k))\, \chi_{T P_k}(\omega)\, \frac{1}{\beta_k} = \frac{1}{\beta_{X_{n+1}(\alpha)}}\, \chi_{[T^{n+1}\alpha,\, 1)}(\omega) + \frac{1}{\beta_N}\, \chi_{[0,\, \gamma)}(\omega) + \sum_{k=X_{n+1}(\alpha)}^{N-1} \frac{1}{\beta_k} - \frac{1}{\beta_{X_{n+1}(\alpha)}}.$$

Hence, with h as in (8.4),

$$\begin{aligned}
&\sum_{k=0}^{N} h(\bar{f}(\omega+k))\, \chi_{T P_k}(\omega)\, \frac{1}{\beta_k} \\
&= \sum_{k=0}^{N} \frac{1}{\beta_k}\, \chi_{T P_k}(\omega) + \frac{D_1}{\beta_N} \sum_{n=0}^{\infty} B(\gamma, n) \\
&\quad \cdot \left\{ \frac{1}{\beta_{X_{n+1}(\gamma)}}\, \chi_{[0,\, T^{n+1}\gamma)}(\omega) + \frac{1}{\beta_0}\, \chi_{[\alpha,\, 1)}(\omega) + \sum_{k=1}^{X_{n+1}(\gamma)} \frac{1}{\beta_k} - \frac{1}{\beta_{X_{n+1}(\gamma)}} \right\} \\
&\quad + \frac{D_2}{\beta_0} \sum_{n=0}^{\infty} B(\alpha, n) \left\{ \frac{1}{\beta_{X_{n+1}(\alpha)}}\, \chi_{[T^{n+1}\alpha,\, 1)}(\omega) + \frac{1}{\beta_N}\, \chi_{[0,\, \gamma)}(\omega) \right. \\
&\quad \left. + \sum_{k=X_{n+1}(\alpha)}^{N-1} \frac{1}{\beta_k} - \frac{1}{\beta_{X_{n+1}(\alpha)}} \right\} \\
&= \frac{1}{\beta_0}\, \chi_{[\alpha,\, 1)}(\omega) + \frac{1}{\beta_N}\, \chi_{[0,\, \gamma)}(\omega) + \sum_{k=1}^{N-1} \frac{1}{\beta_k} \\
&\quad + \frac{D_1}{\beta_N} \sum_{n=1}^{\infty} B(\gamma, n)\, \chi_{[0,\, T^n\gamma)}(\omega) + \frac{D_1 B(\gamma)}{\beta_N \beta_0}\, \chi_{[\alpha,\, 1)}(\omega) \\
&\quad + \frac{D_1}{\beta_N} \sum_{n=0}^{\infty} B(\gamma, n) \left\{ \sum_{k=1}^{X_{n+1}(\gamma)} \frac{1}{\beta_k} - \frac{1}{\beta_{X_{n+1}(\gamma)}} \right\} \\
&\quad + \frac{D_2}{\beta_0} \sum_{n=1}^{\infty} B(\alpha, n)\, \chi_{[T^n\alpha,\, 1)}(\omega) + \frac{D_2 B(\alpha)}{\beta_0 \beta_N}\, \chi_{[0,\, \gamma)}(\omega) \\
&\quad + \frac{D_2}{\beta_0} \sum_{n=0}^{\infty} B(\alpha, n) \left\{ \sum_{k=X_{n+1}(\alpha)}^{N-1} \frac{1}{\beta_k} - \frac{1}{\beta_{X_{n+1}(\alpha)}} \right\}.
\end{aligned}$$

Now
$$\sum_{n=0}^{\infty} B(\gamma,n)\left\{\sum_{k=1}^{X_{n+1}(\gamma)} \frac{1}{\beta_k}-\frac{1}{\beta_{X_{n+1}(\gamma)}}\right\}=\sum_{n=0}^{\infty} B(\gamma,n)\left\{g\big(X_{n+1}(\gamma)\big)-\frac{1-\alpha}{\beta_0}\right\}$$
$$=\gamma-\frac{1-\alpha}{\beta_0}B(\gamma)$$
and
$$\sum_{n=0}^{\infty} B(\alpha,n)\left\{\sum_{k=X_{n+1}(\alpha)}^{N-1} \frac{1}{\beta_k}-\frac{1}{\beta_{X_{n+1}(\alpha)}}\right\}$$
$$=\sum_{n=0}^{\infty} B(\alpha,n)\left\{\frac{1-\alpha}{\beta_0}+\sum_{k=1}^{N-1}\frac{1}{\beta_k}-g\big(X_{n+1}(\alpha)\big)-\frac{1}{\beta_{X_{n+1}(\alpha)}}\right\}$$
$$=-\frac{\gamma}{\beta_N}B(\alpha)-\alpha+1.$$
Hence
$$\sum_{k=0}^{N} h\big(\bar f(\omega+k)\big)\,\chi_{TP_k}(\omega)\frac{1}{\beta_k}$$
$$=\sum_{k=1}^{N-1}\frac{1}{\beta_k}+\frac{D_1}{\beta_N}\left\{\gamma-\frac{1-\alpha}{\beta_0}B(\gamma)\right\}+\frac{D_2}{\beta_0}\left\{1-\frac{\gamma}{\beta_N}B(\alpha)-\alpha\right\}$$
$$+\frac{1}{\beta_0}\chi_{[\alpha,1)}(\omega)\left\{1+\frac{D_1}{\beta_N}B(\gamma)\right\}+\frac{1}{\beta_N}\chi_{[0,\gamma)}(\omega)\left\{1+\frac{D_2}{\beta_0}B(\alpha)\right\}$$
$$+\frac{D_1}{\beta_N}\sum_{n=1}^{\infty}B(\gamma,n)\,\chi_{[0,T^n\gamma)}(\omega)$$
$$+\frac{D_2}{\beta_0}\sum_{n=1}^{\infty}B(\alpha,n)\,\chi_{[T^n\alpha,1)}(\omega)$$
$$=h(\omega)$$
using (8.5) and the fact that the constant term is equal to
$$\sum_{k=1}^{N-1}\frac{1}{\beta_k}+\frac{\gamma}{\beta_N}\left\{D_1-\frac{D_2}{\beta_0}B(\alpha)\right\}+\frac{1-\alpha}{\beta_0}\left\{D_2-\frac{D_1}{\beta_N}B(\gamma)\right\}=1$$
as required.

Note that in the proof of (8.3) in the case $\alpha>0$, $\gamma<1$, we only needed the assumption $\beta>2$ in order to prove that μ is equivalent to λ. In fact the proof that μ is invariant with respect to T depends only on the existence of a solution of Eqs. (8.5). A solution always exists with D_1 and D_2 positive if $\beta>2$. However a solution may exist for $1<\beta\leqq 2$, with possibly one or both of D_1, D_2 negative (for instance if $\beta_i=\beta$, $0\leqq i\leqq N$, $\beta\neq 2$, the solution is $D_1=D_2=(\beta-1)/(\beta-2)$ which is negative for $1<\beta<2$), which still gives a measure μ which is invariant with respect to T. The question then arises as to whether this measure μ is equivalent to λ. The following lemma, which was proved for the special case $T\omega=\beta\omega+\alpha$ (mod one) in [13], gives a partial answer to this question.

(8.6) **Lemma.** *Suppose there is a solution to Eqs.* (8.5) *and define* $h(\omega)$ *as in* (8.4). *If* $h(\omega)\not\equiv 0$ *and* T *is strongly ergodic, then*
$$\mu(F)=\int_F h(\omega)\,d\lambda(\omega)\qquad F\in\mathscr{B},$$
defines a finite measure which is equivalent to λ.

Proof. As $h(\omega)\not\equiv 0$ there are three cases to consider:

(i) $h(\omega)\geqq 0$, λ-a.e.,

(ii) $h(\omega)\leqq 0$, λ-a.e., and

(iii) $h(\omega)$ assumes both positive and negative values on sets of positive Lebesgue measure.

Case (ii) reduces to case (i) if we replace h by $-h$. Hence we only have to consider cases (i) and (iii). In each of these two cases there is a point $\omega_0\in\Omega$ such that $h(\omega_0)>0$. We shall first show that h is right continuous at each point of Ω and hence in particular at ω_0. Hence if $E=\{\omega\in\Omega: h(\omega)\leqq 0\}$, $\lambda(E)<1$. The remainder of the proof will show that $T^{-1}E\subset E$ (λ-a.e.) and hence, since T is assumed to be strongly ergodic, $\lambda(E)=0$, proving the lemma.

Define for $m\geqq 1$,

$$h_m(\omega)=1+\frac{D_1}{\beta_N}\sum_{n=0}^{m}B(\gamma,n)\chi_{[0,\,T^n\gamma)}(\omega)$$
$$+\frac{D_2}{\beta_0}\sum_{n=0}^{m}B(\alpha,n)\chi_{[T^n\alpha,\,1)}(\omega).$$

Then

$$|h(\omega)-h_m(\omega)|\leqq\frac{|D_1|}{\beta_N}\sum_{n=m+1}^{\infty}\frac{1}{\beta^n}+\frac{|D_2|}{\beta_0}\sum_{n=m+1}^{\infty}\frac{1}{\beta^n}$$
$$=\delta(m),\quad\text{say.}$$

Given $\varepsilon>0$, choose M so large that $\delta(M)<\varepsilon/2$ and for any $\omega\in\Omega$, let ω_1 be the smallest member of the set $\{1, T^k\alpha, T^k\gamma, 1\leqq k\leqq M\}$ which is strictly greater than ω. Then, for $\omega\leqq\omega'<\omega_1$,

$$|h(\omega)-h(\omega')|\leqq|h(\omega)-h_M(\omega)|+|h_M(\omega)-h_M(\omega')|+|h_M(\omega')-h(\omega')|<\varepsilon$$

by the choice of M and the fact that h_M is constant on $[\omega,\omega_1)$. Hence h is right continuous.

Now define the linear operator U on $\mathscr{L}_1(\Omega,\mathscr{B},\lambda)$ by

$$U\tau(\omega)=\sum_{k\in\mathbf{I}}\chi_{TP_k}(\omega)\tau(\bar{f}(k+\omega))|\bar{f}'(k+\omega)|.$$

Note that by Parry's relationship [13] a necessary and sufficient condition for τ to be the density of a measure invariant with respect to T is that τ is a fixed point under U (i.e. $U\tau=\tau$, λ-a.e.). Moreover U is $\mathscr{L}_1$ norm-preserving since

$$\int U\tau(\omega)d\lambda(\omega)=\int\sum_{k\in\mathbf{I}}\chi_{TP_k}(\omega)\tau(\bar{f}(k+\omega))|\bar{f}'(k+\omega)|d\lambda(\omega)$$
$$=\sum_{k\in\mathbf{I}}\int\chi_{TP_k}(\omega)\tau(\bar{f}(k+\omega))|\bar{f}'(k+\omega)|d\lambda(\omega)$$
$$=\sum_{k\in\mathbf{I}}\int_{P_k}\tau(\omega)d\lambda(\omega)$$
$$=\int\tau(\omega)d\lambda(\omega)$$

for any $\tau\in\mathscr{L}_1(\Omega,\mathscr{B},\lambda)$.

Now let

$$h(\omega)=h^+(\omega)-h^-(\omega),$$

where

$$h^+(\omega)=\max(0, h(\omega)) \quad \text{and} \quad h^-(\omega)=-\min(0, h(\omega)).$$

We also put $\tau^+=Uh^+$. Then, since $Uh=h$, $\tau^+(\omega)\geqq h^+(\omega)$, $\omega\in\Omega$. However, since U is norm-preserving,

$$\int \tau^+(\omega)\,d\lambda(\omega)=\int h^+(\omega)\,d\lambda(\omega).$$

Hence, $\tau^+(\omega)=h^+(\omega)$ λ-a.e., i.e. $Uh^+=h^+$ λ-a.e. and so

$$\mu^+(F)=\int_F h^+(\omega)\,d\lambda(\omega), \qquad F\in\mathscr{B},$$

defines a measure which is invariant with respect to T. Since $E=\{\omega: h^+(\omega)=0\}$,

$$\int_E h^+(\omega)\,d\lambda(\omega)=\int_{T^{-1}E} h^+(\omega)\,d\lambda(\omega)=0.$$

Hence $T^{-1}E\subset E$ (λ-a.e.), which completes the proof.

9. Concluding Remarks

The proof of (7.3) depended essentially on the facts that there is an invariant measure for T which is equivalent to λ and that there are sufficiently many full intervals to approximate any sub-interval of Ω arbitrarily closely by full intervals. The following special cases of $T\omega=\beta\omega+\alpha$ (mod one) demonstrate that these conditions are not always satisfied for $1<\beta\leqq 2$ and also that these conditions are not necessary to ensure that T is weak Bernoulli. The first example was discussed in [13].

(9.1) **Example.** Let β be the positive solution of $\beta^2=\beta+1$ and let $\alpha=(3-\beta)/2$. Then $T\omega=\beta\omega+\alpha$ (mod one) is not strongly ergodic since $T^{-1}[(\beta-1)/2, (3-\beta)/2)=[1/\beta^2, 1/\beta)$ and $(\beta-1)/2<1/\beta^2<1/\beta<(3-\beta)/2$. Hence (8.6) is not applicable. In fact, since $T\alpha=\beta/2$, $T^2\alpha=0$, $T^3\alpha=\alpha, \ldots$, etc. and $T\gamma=(2-\beta)/2$, $T^2\gamma=0$, $T^3\gamma=\alpha, \ldots$, etc. where $\gamma=\beta+\alpha-2=(\beta-1)/2$, the density of the invariant (probability) measure given by (8.4) is

$$h(\omega)=\begin{cases} 1/(7-4\beta), & \omega\in[0, 1/2\beta^2)\cup[\beta/2, 1), \\ 1/(3\beta-4), & \omega\in[1/2\beta^2, (\beta-1)/2)\cup[(3-\beta)/2, \beta/2), \\ 0, & \omega\in[(\beta-1)/2, (3-\beta)/2) \end{cases}$$

and hence $\mu(F)=\int_F h(\omega)\,d\lambda(\omega)$, $F\in\mathscr{B}$, defines a measure with respect to which λ is not absolutely continuous.

Now the non-empty rank two intervals are

$$\begin{aligned} &\Delta(0,1)=[0, 1/2\beta^4), \\ &\Delta(0,2)=[1/2\beta^4, 1/2\beta^2), \\ &\Delta(1,0)=[1/2\beta^2, 1/2\beta), \\ &\Delta(1,1)=[1/2\beta, (3-\beta)/2), \\ &\Delta(1,2)=[(3-\beta)/2, \beta/2), \\ &\Delta(2,0)=[\beta/2, 3/2\beta), \\ &\Delta(2,1)=[3/2\beta, 1). \end{aligned}$$

Hence we see that the set on which $h(\omega)=0$ is $\Delta(1,1)$ and so regarding $(1,1)$ as a state with zero probability the techniques of § 6 may be used to see that $T\omega=\beta\omega+\alpha$ (mod one) where β and α are as defined above is a 2-step Markov shift whose associated 1-step Markov shift has state space $a=(0,1)$, $b=(0,2)$, $c=(1,0)$, $d=(1,2)$, $e=(2,0)$, $f=(2,1)$. Moreover the rank three intervals with non-zero μ-measure are

$$\begin{aligned}
&\Delta(0,1,2)=[0,1/2\beta^4),\\
&\Delta(0,2,0)=[1/2\beta^4,1/\beta^4),\\
&\Delta(0,2,1)=[1/\beta^4,1/2\beta^2),\\
&\Delta(1,0,1)=[1/2\beta^2,1/\beta^3),\\
&\Delta(1,0,2)=[1/\beta^3,1/2\beta),\\
&\Delta(1,2,0)=[(3-\beta)/2,2/\beta^2),\\
&\Delta(1,2,1)=[2/\beta^2,\beta/2),\\
&\Delta(2,0,1)=[\beta/2,(3-\beta)/\beta),\\
&\Delta(2,0,2)=[(3-\beta)/\beta,3/2\beta),\\
&\Delta(2,1,0)=[3/2\beta,1).
\end{aligned}$$

Hence the one-step Markov shift associated with T has stationary distribution $p_a=p_b/\beta=p_c=p_d=p_e/\beta=p_f=1/2\beta^4(7-4\beta)$ and transition matrix π given by

$$\pi=\begin{pmatrix}
0 & 0 & 0 & 1 & 0 & 0\\
0 & 0 & 0 & 0 & 1/\beta & 1/\beta^2\\
1/\beta^2 & 1/\beta & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 1/\beta & 1/\beta^2\\
1/\beta^2 & 1/\beta & 0 & 0 & 0 & 0\\
0 & 0 & 1 & 0 & 0 & 0
\end{pmatrix}.$$

Since π^5 is a matrix with all entries non-zero we see that the one-step Markov shift associated with T is mixing and so T is weak Bernoulli.

(9.2) **Example.** If we choose $\beta>1$ and $0<\alpha<1$ such that $\beta+\alpha<2$, then there are no full intervals of rank one and consequently no full intervals of any rank. In particular let β again be the positive solution of $\beta^2=\beta+1$ and let $\alpha=1/2\beta^2$. Then $T\alpha=1/2=(1-\alpha)/\beta$, $T^2\alpha=0$, $T^3\alpha=\alpha$, $T^4\alpha=1/2,\ldots$, etc. and $T\gamma=1/2$, $T^2\gamma=0$, $T^3\gamma=\alpha$, $T^4\gamma=1/2,\ldots$, etc., where $\gamma=\beta+\alpha-1=\beta/2$. Hence the density of the invariant (probability) measure given by (8.4) is

$$h(\omega)=\begin{cases}\beta^2/(\beta^2+1), & \omega\in[0,1/2\beta^2)\cup[\beta/2,1),\\ \beta^3/(\beta^2+1), & \omega\in[1/2\beta^2,\beta/2),\end{cases}$$

which defines a measure equivalent to λ. Moreover T is a 2-step Markov shift whose associated one-step Markov shift has state space $a=(0,0)$, $b=(0,1)$,

$c=(1,0)$, $d=(1,1)$ and stationary distribution

$$\begin{aligned}
p_a&=\mu(\Delta(0,0))=\mu([0,1/2\beta^2))=1/2(\beta^2+1),\\
p_b&=\mu(\Delta(0,1))=\mu([1/2\beta^2,1/2))=\beta^2/2(\beta^2+1),\\
p_c&=\mu(\Delta(1,0))=\mu([1/2,\beta/2))=\beta^2/2(\beta^2+1),\\
p_d&=\mu(\Delta(1,1))=\mu([\beta/2,1))=1/2(\beta^2+1).
\end{aligned}$$

Since the non-empty rank three intervals are

$$\begin{aligned}
&\Delta(0,0,1)=[0,1/2\beta^2),\\
&\Delta(0,1,0)=[1/2\beta^2,1/\beta^2),\\
&\Delta(0,1,1)=[1/\beta^2,1/2),\\
&\Delta(1,0,0)=[1/2,1/\beta),\\
&\Delta(1,0,1)=[1/\beta,\beta/2),\\
&\Delta(1,1,0)=[\beta/2,1),
\end{aligned}$$

the one-step Markov shift associated with T has transition matrix π given by

$$\pi=\begin{pmatrix} 0 & 1 & 0 & 0\\ 0 & 0 & 1/\beta & 1/\beta^2\\ 1/\beta^2 & 1/\beta & 0 & 0\\ 0 & 0 & 1 & 0\end{pmatrix}.$$

Since π^4 is a matrix with all entries non-zero, T is weak Bernoulli.

Since we have examples of piecewise linear transformations which do not satisfy condition (2.2) but which are nevertheless weak Bernoulli with respect to the time-one partition we must ask whether the class $\mathscr{L}$ may be widened to define a class which includes these transformations. Because of the remarks preceding (9.1) it is clear that techniques quite different from those used in (7.3) would have to be used to show that such a class was weak Bernoulli. In particular, is $T\omega=\beta\omega+\alpha$ (mod one) weak Bernoulli for $1<\beta\leqq 2$, $\beta+\alpha$ not integral?

Acknowledgement. This paper constitutes part of the author's Ph.D. thesis for which the supervision of S. M. Rudolfer is gratefully acknowledged.

References

1. Adler, R., Weiss, B.: Notes on the β-transformation. (To appear)
2. Billingsley, P.: Ergodic Theory and Information. New York: Wiley 1965
3. Cigler, J.: Ziffernverteilung in θ-adischen Brüchen. Math. Z. **75**, 8–13 (1961)
4. Dunford, N., Miller, D.S.: On the Ergodic Theorem. Trans. Amer. Math. Soc. **60**, 538–549 (1946)
5. Fischer, R.: Ergodische Theorie von Ziffernentwicklungen in Wahrscheinlichkeitsräumen. Math. Z. **128**, 217–230 (1972)
6. Friedman, N.A.: Introduction to Ergodic Theory. New York: van Nostrand 1970
7. Friedman, N.A., Ornstein, D.S.: On isomorphism of weak Bernoulli transformations, Advances in Math. **5**, 365–394 (1970)
8. Halmos, P.R.: Measure Theory. New York: van Nostrand 1950
9. Halmos, P.R.: Lectures on Ergodic Theory. New York: Chelsea 1956

10. Knopp, K.: Mengentheoretische Behandlung einiger Probleme der diophantischen Approximationen und der transfiniten Wahrscheinlichkeiten. Math. Ann. **95**, 409–426 (1926)
11. Ornstein, D.S.: Ergodic theory, randomness and dynamical systems. New Haven: Yale Univ. Press 1973
12. Parry, W.: On the β-expansions of real numbers. Acta Math. Acad. Sci. Hungar. **11**, 401–416 (1960)
13. Parry, W.: Representations for real numbers. Acta Math. Acad. Sci. Hungar. **15**, 95–105 (1964)
14. Parry, W.: On Rohlin's formula for entropy. Acta Math. Acad. Sci. Hungar. **15**, 107–113 (1964)
15. Rényi, A.: Representations for real numbers and their ergodic properties. Acta Math. Acad. Sci. Hungar. **8**, 477–493 (1957)
16. Rohlin, V.A.: Exact endomorphisms of a Lebesgue space. Izv. Akad. Nauk SSSR Ser. Mat. **25**, 499–530 (1961) = Amer. Math. Soc. Transl. (2) **39**, 1–36 (1964)
17. Shiokawa, I.: Ergodic Properties of piecewise linear transformations. Proc. Japan Acad. **46**, 1122–1125 (1970)
18. Smorodinsky, M.: β-automorphisms are Bernoulli Shifts. Acta Math. Acad. Sci. Hungar. **24**, 273–278 (1973)
19. Wilkinson, K.M.: Ergodic properties of certain linear mod one transformations, Advances in Math. **14**, 64–72 (1974)

K. M. Wilkinson
Department of Mathematics
University of Nottingham
University Park
Nottingham NG7 2RD
England

(Received May 9, 1974)