TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 214, 1975

RUELLE'S OPERATOR THEOREM AND g-MEASURES

BY

PETER WALTERS

ABSTRACT. We use g-measures to give a proof of a convergence theorem of Ruelle. The method of proof is used to gain information about the ergodic properties of equilibrium states for subshifts of finite type.

0. In 1968 D. Ruelle proved a convergence theorem (Theorem 3 of [8]) which he used to obtain a large class of interactions, for an infinite one-dimensional classical lattice gas, which have no phase transitions. The equilibrium state of such a system was shown to have strong ergodic properties. In 1973 R. Bowen remarked that Ruelle's proof could be extended to show the convergence of the powers of a certain operator acting on the space of all real-valued continuous functions on a one-sided shift space [2]. In the paper [2] he used this result and the theory of Markov partitions, due to Ya. G. Sinai and himself, to show that if Ω_s is a basic set of an Axiom A diffeomorphism f then, with respect to $f|_{\Omega_s}$, each Hölder continuous ϕ: $\Omega_s \longrightarrow R$ has a unique equilibrium state and if $f|_{\Omega_s}$ is topologically mixing then with respect to the equilibrium state $f|_{\Omega_s}$ is a Bernoulli shift.

This was done by using Markov partitions to move the problem to one about a two-sided subshift of finite type ([1], [10]). It can then be reduced to a problem about a one-sided subshift of finite type ([2], [11, p. 28]) and then Ruelle's theorem can be used.

In this paper we connect these results with the idea of a g-measure, studied by M. Keane [5]. We shall give a proof of Ruelle's operator theorem using the notion of g-measure. The structure of this proof allows us to deduce elementary proofs of several results about equilibrium states. There is a dense subset V of $C(X)$ whose members each have a unique equilibrium state, the natural extension of the one-sided shift with respect to these measures are Bernoulli shifts, and two members of V have the same equilibrium state if and only if they differ by a function of the form $f \circ T - f + c$ where $f \in C(X)$ and $c \in R$ (here T denotes the one-sided shift). In §4 we extend these results to transformations more general than one-sided subshifts of finite type.

Received by the editors December 10, 1974.
AMS (MOS) subject classifications (1970). Primary 28A65, 58F15.

1. Equilibrium states. Let X be a compact metrizable space and $T: X \longrightarrow X$ continuous. The collection $M(X)$, of all probability measures on the σ-algebra $\mathcal{B}$ of Borel subsets of X is a convex set which is compact in the weak*-topology and the subset $M(T)$ of all T-invariant members of $M(X)$ is a closed subset of $M(X)$. If $\mu \in M(T)$ we denote the entropy of T relative to μ by $h_\mu(T)$ [9]. If $\phi \in C(X)$, the Banach space of real-valued continuous functions on X with the supremum norm $\| \ \|$, and $\mu \in M(X)$ we let $\mu(\phi)$ denote the integral of ϕ relative to μ. $\mu \in M(T)$ is an equilibrium state for $\phi \in C(X)$ if

$$h_\mu(T) + \mu(\phi) = \sup_{m \in M(T)} [h_m(T) + m(\phi)].$$

If we denote the pressure of T by $P_T: C(X) \longrightarrow R \cup \{+\infty\}$ then this is equivalent to requiring $h_\mu(T) + \mu(\phi) = P_T(\phi)$ [12]. If the mapping $\mu \longrightarrow h_\mu(T)$ of $M(T)$ to R is upper semicontinuous, then each $\phi \in C(X)$ has equilibrium states. The transformations T studied in this paper have this property and we shall be mainly concerned with which $\phi \in C(X)$ have a unique equilibrium state. Let E_ϕ denote the collection of equilibrium states for ϕ. If E_ϕ is nonempty it is a convex subset of $M(T)$ and if $\mu \longrightarrow h_\mu(T)$ is an upper semicontinuous mapping, then E_ϕ is closed. In this case any extreme point of E_ϕ is also an extreme point of $M(T)$ and so is an ergodic measure for T. As we shall see later, for certain transformations T and certain $\phi \in C(X)$ there is a unique number of E_ϕ and it has very strong ergodic properties.

In studying equilibrium states the following simple result is useful.

LEMMA 1.1. *Let $T: X \longrightarrow X$ be continuous and $\phi, \psi \in C(X)$. If $\phi - \psi = f \circ T - f + c$, for some $f \in C(X)$ and some $c \in R$, then $E_\phi = E_\psi$.*

PROOF. This follows because for each $\mu \in M(T)$

$$h_\mu(T) + \mu(\phi) = h_\mu(T) + \mu(\psi) + c. \square$$

The same conclusion holds if we assume that, for some $c \in R$, $\phi - \psi - c$ belongs to the closure of $\{f \circ T - f \mid f \in C(X)\}$.

We shall be interested in the case where $T: X \longrightarrow X$ is a one-sided subshift of finite type. This means there is a finite set S, with $|S|$ members, and a $|S| \times |S|$ matrix A whose entries are zeros and ones so that X is the subset of S^{Z^+} (where $Z^+ = \{0, 1, 2, \ldots\}$) defined by $x = \{x_n\}_0^\infty \in X$ if and only if $A_{x_n x_{n+1}} = 1$ for all $n \geqslant 0$. If S is given the discrete topology and S^{Z^+} the product topology, then X is a closed subset of the compact space S^{Z^+}. $T: X \longrightarrow X$ is defined by $(Tx)_n = x_{n+1}$. T is continuous. We shall denote by d the metric on S^{Z^+} defined by $d(x, y) = 1/(k + 1)$ where k is the greatest integer with $x_i = y_i$ for all $i < k$. A basis for the topology of X is given by the collection of all finite-dimensional cylinders; i.e. sets of the form $\{x \in X \mid x_i = a_i, r \leqslant i \leqslant s\}$ where $r < s$ and $a_r, \ldots, a_s \in S$. For $n \geqslant 0$ a measure of the oscillation of $\phi \in C(X)$

is given by $\mathrm{var}_n(\phi) = \sup\{|\phi(x) - \phi(y)| \mid d(x, y) \leq 1/(n+1)\}$. Since ϕ is continuous $\mathrm{var}_n(\phi) \longrightarrow 0$ as $n \longrightarrow \infty$, and we shall be mainly interested in those $\phi \in C(X)$ which satisfy the stronger requirement that $\Sigma_{n=0}^{\infty} \mathrm{var}_n(\phi) < \infty$. Such functions are dense in $C(X)$ because all functions which only depend on a finite number of coordinates (i.e. for some k $\phi(x) = \phi(y)$ if $d(x, y) \leq 1/(k+1)$) satisfy the condition.

If $x \in X$ and $p \in S^k$ $(k > 0)$ then px will denote the member y of S^{Z^+} where $y_i = p_i$, $0 \leq i \leq k-1$ and $y_j = x_{j-k} k \leq j$. Note that $px \in X$ if and only if $A_{p_{k-1}x_0} = 1$ and $A_{p_ip_{i+1}} = 1$, $0 \leq i \leq k-2$. T is topologically transitive (i.e. for all nonempty open sets U, V $T^{-n}U \cap V \neq \emptyset$ for some n) if and only if for each pair $i, j \in S$ there exists $k > 0$ with $(A^k)_{i,j} = 1$. T is topologically mixing (i.e. for all nonempty sets U, V there exists N with $T^{-n}U \cap V \neq \emptyset$ for all $n \geq N$) if and only if there exists $M > 0$ so that A^M has all entries positive. Clearly if T is topologically mixing $\bigcup_{n=0}^{\infty} \{T^{-n}x\}$ is dense in X for each $x \in X$.

As mentioned in §0 Bowen has shown how to relate $f|_{\Omega_k}$, where Ω_k is a basic set of the Axiom A diffeomorphism f, to a two-sided subshift of finite type, and Sinai has shown how to reduce the study of equilibrium states for a large class of functions for a two-sided subshift of finite type to the case of a one-sided subshift of finite type [11, p. 28]. In this situation one can use Ruelle's operator theorem (Theorem 3.3) to get information about equilibrium states.

2. g-**measures.** Let $T: X \longrightarrow X$ be a one-sided subshift of finite type. For $\phi \in C(X)$ we can define the Ruelle operator $L_\phi: C(X) \longrightarrow C(X)$ by $(L_\phi f)(x) = \Sigma_{y \in T^{-1}x} e^{\phi(y)} f(y)$. Of special inportance are the functions $\phi = \log g$ where $g \in G = \{g \in C(X) \mid g > 0$ and $\Sigma_{y \in T^{-1}x} g(y) = 1$ for all $x\}$. Then $(L_{\log g} f)(x) = \Sigma_{y \in T^{-1}} g(y) f(y)$ and $L_{\log g} U_T f = f$ where $U_T f = f \circ T$.

We shall use the following result, where the dual of $L: C(X) \longrightarrow C(X)$ is denoted by L^* and $E_m(f/T^{-1}\mathcal{B})$ denotes the conditional expectation of $f \in L^1(m)$ relative to the σ-algebra $T^{-1}\mathcal{B}$, where $\mathcal{B}$ denotes the Borel σ-algebra of X.

THEOREM 2.1 (Ledrappier [6]). *Let $g \in G$ and $m \in M(X)$. If L denotes $L_{\log g}$ the following are equivalent*:

(i) $L^* m = m$.

(ii) *$m \in M(T)$ and $E_m(f/T^{-1}\mathcal{B})(x) = \Sigma_{z \in T^{-1}Tx} g(z) f(x)$ a.e. (m) for $f \in L^1(m)$.*

(iii) *$m \in M(T)$ and m is an equilibrium state for* $\log g$.

PROOF. We include Ledrappier's proof for completeness.

(i) $\longrightarrow$ (ii). If $f \in C(X)$ $m(fT) = m(L(f \circ T)) = m(f)$, so $m \in M(T)$. If $f \in L^1(m)$,

$$m(f) = m(Lf) = m((Lf) \circ T) = \int \sum_{y \in T^{-1}Tx} g(y) f(y) dm(x),$$

so

$$E_m(f/T^{-1}\mathcal{B})(x) = \sum_{y\in T^{-1}Tx} g(y)f(y) \quad \text{a.e. } (m).$$

(ii) $\longrightarrow$ (iii). Since the partition of X into the sets $\{x \mid x_0 = i\}$, $i \in S$, is a one-sided generator for any measure $\mu \in M(T)$ we have $h_\mu(T) = H_\mu(\mathcal{B}/T^{-1}\mathcal{B})$, (see [9, p. 27]). If $g_\mu\colon X \longrightarrow R$ denotes the function defined a.e. (μ) by $E_\mu(f/T^{-1}\mathcal{B})(x) = \Sigma_{y\in T^{-1}Tx} g_\mu(y)f(y)$, then

$$h_\mu(T) = -\int \sum_{y\in T^{-1}Tx} g_\mu(y)\log g_\mu(y)\,d\mu(x) = -\int \log g_\mu(x)\,d\mu(x) = -\mu(\log g_\mu).$$

Therefore

$$\begin{aligned} h_\mu(T) + \mu(\log g) &= \mu(\log(g/g_\mu)) \leqslant \mu(g/g_\mu - 1) \quad \text{since } \log x \leqslant x - 1 \\ &= \int \sum_{y\in T^{-1}Tx} g_\mu(y)\left(\frac{g(y)}{g_\mu(y)} - 1\right) d\mu(x) \\ &= \int \sum_{y\in T^{-1}Tx} (g(y) - g_\mu(y))\,d\mu(x) = 0, \end{aligned}$$

and $h_\mu(T) + \mu(\log g) = 0$ (i.e. μ is an equilibrium state for $\log g$) if and only if $\log(g/g_\mu) = g/g_\mu - 1$ a.e. (μ), i.e. $g = g_\mu$ a.e. (μ). But if m satisfies (ii) then $g = g_m$ a.e. (m) so m is an equilibrium measure for $\log g$.

(iii) $\longrightarrow$ (i). From the above reasoning $m \in M(T)$ is an equilibrium state for $\log g$ if and only if $g = g_m$ a.e. (m). Let $f \in C(X)$ and we want to show $m(f) = m(Lf)$. But

$$m(Lf) = m((Lf)\circ T) = \int \sum_{y\in T^{-1}Tx} g(y)f(y)\,dm(x) = m(f).$$

We call a measure m a g-measure if it satisfies the equivalent conditions of Theorem 2.1. By the Schauder-Tychonoff theorem [3, p. 456] L^* always has a fixed point in $M(X)$ so g-measures always exist. Note for a g-measure μ $h_\mu(T) + \mu(\log g) = 0$.

LEMMA 2.1. (a) *If $g \in G$ then each g-measure gives positive measure to each nonempty open set.*

(b) *If $g_1, g_2 \in G$ and some g_1-measure coincides with some g_2-measure, then $g_1 = g_2$.*

PROOF. (a) Let m be a g-measure. It suffices to show each cylinder set of the form $B = [a_0, \ldots, a_k] = \{x \in X \mid x_i = a_i, 0 \leqslant i \leqslant k\}$ has positive measure. Choose $c > 0$ so that $g > c$. Then

$$m(B) = \int E_m(\chi_B/T^{-(k+1)}\mathcal{B})\,dm \quad \text{where } \chi_B \text{ is the characteristic function of } B,$$

$$= \int g(ax)g(T(ax)) \cdots g(T^k(ax))\,dm(x) \quad \text{where } a = (a_0, \ldots, a_k) \in S^{k+1},$$

$$\geqslant c^{k+1}.$$

(b) Let $m_1 = m_2$ where m_i is a g_i-measure, $i = 1, 2$.
By (ii) of Theorem 2.1, $g_1 = g_2$ a.e. m_1. Hence by (a) $g_1 = g_2$.

3. Convergence theorems and equilibrium states.

THEOREM 3.1. *Let $T: X \to X$ be a topologically mixing one-sided subshift of finite type and let $g \in G$. Denote $L_{\log g}$ by L. If $\Sigma_{n=0}^{\infty} \operatorname{var}_n(\log g) < \infty$, then $L^n f$ converges uniformly to a constant $\mu(f)$ for all $f \in C(X)$. μ is a g-measure and is the only one.*

PROOF. The proof is that of Keane [5] but we have adapted it slightly to our assumptions. Let $f \in C(X)$.

We first show the set $\{L^n f \mid n \geqslant 0\}$ is an equicontinuous subset of $C(X)$. Suppose k is a positive integer and $d(x, y) \leqslant 1/(k+1)$. Then $x_0 = y_0$ so if $p \in S^n$ $px \in X$ if and only if $py \in X$. Let $S^n(x)$ denote these $p \in S^n$ with $px \in X$. Then $S^n(x) = S^n(y)$ and

$$\begin{aligned}
&|L^n f(x) - L^n f(y)| \\
&\quad = \Bigg| \sum_{p \in S^n(x)} g(px)g(T(px)) \cdots g(T^{n-1}(px)) f(px) \\
&\qquad - \sum_{p \in S^n(x)} g(py)g(T(py)) \cdots g(T^{n-1}(py)) f(py) \Bigg| \\
&\quad \leqslant \Bigg| \sum_{p \in S^n(x)} g(px)g(T(px)) \cdots g(T^{n-1}(px)) [f(px) - f(py)] \Bigg| \\
&\qquad + \Bigg| \sum_{p \in S^n(x)} f(py) [g(px)g(T(px)) \cdots g(T^{n-1}(px)) \\
&\qquad\qquad - g(py)g(T(py)) \cdots g(T^{n-1}(py))] \Bigg| \\
&\quad \leqslant \sup_{p \in S^n(x)} |f(px) - f(py)| \\
&\qquad + \|f\| \sum_{p \in S^n(x)} |g(px) \cdots g(T^{n-1}(px)) - g(py) \cdots g(T^{n-1}(py))|.
\end{aligned}$$

But

$$
\begin{aligned}
&|g(px)\cdots g(T^{n-1}(px)) - g(py)\cdots g(T^{n-1}(py))| \\
&\quad = g(py)\cdots g(T^{n-1}(py))\left|\frac{g(px)\cdots g(T^{n-1}(px))}{g(py)\cdots g(T^{n-1}(py))} - 1\right| \\
&\quad \leqslant [g(py)\cdots g(T^{n-1}(py))] \\
&\qquad \cdot \max\big(e^{\operatorname{var}_k(\log g)+\operatorname{var}_{k+1}(\log g)+\cdots+\operatorname{var}_{k+n}(\log g)} - 1, \\
&\qquad\qquad 1 - e^{-\operatorname{var}_k(\log g)-\cdots-\operatorname{var}_{k+n}(\log g)}\big) \\
&\quad \leqslant [g(py)\cdots g(T^{n-1}(py))]\ \max\big(e^{\sum_{r=k}^{\infty}\operatorname{var}_r(\log g)} - 1,\ 1 - e^{-\sum_{r=k}^{\infty}\operatorname{var}_r(\log g)}\big) \\
&\quad = g(py)\cdots g(T^{n-1}(py))C_k \quad \text{where } C_k \to 0 \text{ as } k \to \infty.
\end{aligned}
$$

Therefore

$$
\begin{aligned}
|L^n f(x) - L^n f(y)| &\leqslant \sup_{p\in S^n(x)} |f(px) - f(py)| + C_k\|f\| \\
&\leqslant \sup\{|f(w) - f(z)| \mid d(w,z) \leqslant 1/(k+1)\} + C_k\|f\|,
\end{aligned}
$$

if $d(x, y) \leqslant 1/(k+1)$ and so $\{L^n f \mid n \geqslant 0\}$ is an equicontinuous subset of $C(X)$. Since $\|L^n f\| \leqslant \|f\|$ we know by the Arzela-Ascoli theorem that the closure of $\{L^n f \mid n \geqslant 0\}$ is compact. Hence there exists $f_* \in C(X)$ and a sequence n_i of positive integers such that $L^{n_i} f \rightrightarrows f_*$. If $\alpha(h)$ and $\beta(h)$ denote the minimum value and maximum value of $h \in C(X)$ then

$$
\alpha(f) \leqslant \alpha(Lf) \leqslant \cdots \leqslant \alpha(f_*) \leqslant \beta(f_*) \leqslant \cdots \leqslant \beta(Lf) \leqslant \beta(f).
$$

Note that $\alpha(f_*) = \alpha(Lf_*)$ and if $\alpha(f_*) = \alpha(Lf_*) = (Lf_*)(z)$ then $f_*(y) = \alpha(f_*)$ for $y \in T^{-1}z$; similarly if $\alpha(f_*) = \alpha(L^k f_*) = L^k f_*(w)$, then $f_*(y) = \alpha(f_*)$ for $y \in T^{-k}w$. By using topological mixing we see that every cylinder set contains a point where f_* attains its minimum. Therefore f_* is a constant and then clearly $L^n f \rightrightarrows f_*$. Write $\mu(f)$ instead of f_*. By the Riesz representation theorem μ is a probability measure on X and it is clear that $L^*\mu = \mu$. If m is another g-measure then integrating $L^n f \rightrightarrows \mu(f)$ with respect to m gives $m = \mu$.

The first part of the proof shows that instead of assuming $\sum_{n=0}^{\infty} \operatorname{var}_n(\log g) < \infty$ we could assume the weaker condition

$$
\sup_n \left[\sup \left\{ \sum_{p\in S^n(x)} |g(px)\cdots g(T^{n-1}(px)) - g(py)\cdots g(T^{n-1}(py))| \ \Big|\ d(x,y) \leqslant \frac{1}{k+1} \right\} \right] \to 0 \quad \text{as } k \to \infty.
$$

For the notions of natural extension, Bernoulli shift, exact endomorphism and strong mixing we refer to [9].

THEOREM 3.2. *Let $T: X \longrightarrow X$ be a topologically mixing one-sided subshift of finite type and let $g \in G$. If $\Sigma_{n=0}^{\infty} \operatorname{var}_n(\log g) < \infty$, then the unique g-measure μ given by Theorem 3.1 has a Bernoulli natural extension. Hence relative to the measure μ, T is an exact endomorphism and is strong mixing.*

PROOF. This proof is an adaptation of that of Bowen [2]. Another proof is in [6].

Let $(\widetilde{X}, \widetilde{B}, \widetilde{\mu}, \widetilde{T})$ denote the natural extension dynamical system. If γ denotes the partition of X into the sets $\{x \in X \mid x_0 = i\}$ as i runs through S let $\widetilde{\gamma}$ denote its extension to $\widetilde{X}$. We shall show $\widetilde{\gamma}$ is a weak Bernoulli partition for $\widetilde{T}$ [4]. We shall show for every $\epsilon > 0$ there exists N such that for $t \geqslant N$ and for all positive integers r and s $|\widetilde{\mu}(P \cap \widetilde{T}^{(s+t)}Q) - \widetilde{\mu}(P) \cdot \widetilde{\mu}(Q)| \leqslant \epsilon\widetilde{\mu}(P)\widetilde{\mu}(Q)$ for all atoms P of $\bigvee_{i=0}^{s} \widetilde{T}^{-i}\widetilde{\gamma}$ and atoms Q of $\bigvee_{i=0}^{r} \widetilde{T}^{-i}\widetilde{\gamma}$. This clearly implies $\widetilde{\gamma}$ is weak Bernoulli. The condition stated is equivalent to showing

$$|\mu(P \cap T^{-(s+t)}Q) - \mu(P) \cdot \mu(Q)| \leqslant \epsilon\mu(P)\mu(Q) \tag{1}$$

for all atoms P of $\bigvee_{i=0}^{s} T^{-i}\gamma$ and atoms Q of $\bigvee_{i=0}^{r} T^{-i}\gamma$.

Let $P = \{x \in X \mid x_i = a_i, 0 \leqslant i \leqslant s\}$ and $Q = \{x \in X \mid x_k = b_k, 0 \leqslant k \leqslant r\}$. Suppose both are nonempty. Choose δ so that $1/(1-\epsilon) < (1+\delta)^3 < 1 + \epsilon$. Choose m so that if $s > m$ and $x_i = y_i$, $0 \leqslant i \leqslant s$, then

$$(1+\delta)^{-1} < g(y) \cdots g(T^{s-m-1}y)/g(x) \cdots g(T^{s-m-1}x) < 1 + \delta.$$

m depends only on ϵ and clearly it suffices to prove (1) for $s > m$, since an atom of $\bigvee_{i=0}^{s'} T^{-i}\gamma$ is a union of atoms of $\bigvee_{j=0}^{s} T^{-j}\gamma$ if $s' < s$. Therefore assume $s > m$. Note that

$$\begin{aligned}\mu(P \cap T^{-(s+t)}Q) &= \mu(\chi_P \cdot \chi_Q \circ T^{s+t}) \\ &= \mu(L^{s+t}(\chi_P \cdot \chi_Q \circ T^{s+t})) = \mu(\chi_Q \cdot L^{s+t}\chi_P).\end{aligned}$$

Consider

$$(L^{s-m}\chi_P)(z) = \sum_{y \in T^{-(s-m)}z} g(y) \cdots g(T^{s-m-1}y)\chi_P(y). \tag{2}$$

Note that $P \cap T^{-(s-m)}z \neq \varnothing$ if and only if $z \in P' = \{x \in X \mid x_i = a_{s-m+i}, 0 \leqslant i \leqslant m\} = [a_{s-m}, \ldots, a_s]$. Therefore $(L^{s-m}\chi_P)(z) \neq 0$ if and only if $z \in P'$ and then the sum in (2) only has one nonzero term, when $y = a_0 \cdots a_{s-m-1}z$. Hence if $z, w \in P'$,

$$L^{s-m}\chi_P(z)/L^{s-m}\chi_P(w) = g(y) \cdots g(T^{s-m-1}y)/g(u) \cdots g(T^{s-m-1}u)$$

where $y = a_0 \cdots a_{s-m-1}z$ and $u = a_0 \cdots a_{s-m-1}w$, and so $(1+\delta)^{-1} < L^{s-m}\chi_P(z)/L^{s-m}\chi_P(w) < 1 + \delta$ by choice of m. Fix some $w \in P'$ and let $c = L^{s-m}\chi_P(w)$. Then $(1+\delta)^{-1}c\chi_{P'}(z) \leqslant L^{s-m}\chi_P(z) \leqslant (1+\delta)c\chi_{P'}(z)$ for all

$z \in X$. Integrating this gives $(1+\delta)^{-1}c\mu(P') \leq \mu(P) \leq (1+\delta)c\mu(P')$, and using this to eliminate c gives

$$(1+\delta)^{-2}\mu(P)\chi_{P'}(z)/\mu(P') \leq L^{s-m}\chi_P(z) \leq (1+\delta)^2\mu(P)\chi_{P'}(z)/\mu(P').$$

Therefore

$$\begin{aligned}(1+\delta)^{-2}\frac{\mu(P)}{\mu(P')}(L^{t+m}\chi_{P'})(z) &\leq (L^{t+s}\chi_P)(z)\\ &= (L^{t+m}(L^{s-m}\chi_P))(z) \leq (1+\delta)^2\frac{\mu(P)}{\mu(P')}(L^{t+m}\chi_{P'})(z).\end{aligned}$$

Using Theorem 3.1 choose N so that $t \geq N$ implies $|L^{t+m}\chi_{P'}(z) - \mu(P')| < \mu(P')\delta/(1+\delta)$ for all $z \in X$. Since m depends only on ϵ and since there are at most $|S|^m$ possible choices for P' we can choose N, depending only on ϵ, to work for all choices of P'. If $t \geq N$ then $(1+\delta)^{-3}\mu(P) \leq (L^{t+s}\chi_p)(z) \leq (1+\delta)^3\mu(P)$ and therefore

$$\chi_Q(z)(1+\delta)^{-3}\mu(P) \leq (L^{t+s}\chi_P)(z)\chi_Q(z) \leq (1+\delta)^3\mu(P)\chi_Q(z), \qquad z \in X.$$

Integrating this gives

$$(1+\delta)^{-3}\mu(P)\mu(Q) \leq \mu(P \cap T^{-(s+t)}Q) \leq (1+\delta)^3\mu(P)\mu(Q)$$

and so

$$\begin{aligned}|\mu(P \cap T^{-(s+t)}Q) - \mu(P)\mu(Q)| &\leq \mu(P)\mu(Q)\cdot\max((1+\delta)^3 - 1, 1 - 1/(1+\delta)^3)\\ &\leq \mu(P)\mu(Q)\epsilon.\end{aligned}$$

One can prove directly that (T, μ) is an exact endomorphism without the use of the Bernoulli property. Since $L^n f \rightrightarrows \mu(f), f \in C(X)$, we know $\int |L^n f - \mu(f)|\,d\mu \longrightarrow 0$ as $n \longrightarrow \infty$ if $f \in L'(\mu)$. If $E_\mu(f/\mathcal{Q})$ denotes the conditional expectation of $f \in L'(\mu)$ relative to the σ-algebra $\mathcal{Q}$ we have to show $E_\mu(f/\bigcap^\infty T^{-n}\mathcal{B}) = \mu(f)$ a.e. for all $f \in L'(\mu)$. Note that since μ is a g-measure $E_\mu(f/T^{-n}\mathcal{B}) = (L^n f)(T^n x)$ a.e. and therefore, using the martingale theorem,

$$\begin{aligned}\int\left|E_\mu\left(f\Big/\bigcap_0^\infty T^{-n}\mathcal{B}\right) - \mu(f)\right|d\mu &= \lim_{n\to\infty}\int|E_\mu(f/T^{-n}\mathcal{B}) - \mu(f)|\,d\mu\\ &= \lim_{n\to\infty}\int|L^n f - \mu(f)|\,d\mu = 0.\end{aligned}$$

In this proof we used the assumption $\Sigma_{n=0}^\infty \operatorname{var}_n(\log g) < \infty$ in getting the convergence $L^n f \rightrightarrows \mu(f)$ (Theorem 3.1) and in the definition of m. Taking into account the remark after the proof of Theorem 3.1 we could assume the following condition instead of $\Sigma_0^\infty \operatorname{var}_n(\log g) < \infty$:

$$\sup_{s>m}\left[\sup\left\{\left|\frac{g(y)\cdots g(T^{s-m-1}y)}{g(x)\cdots g(T^{s-m-1}x)} - 1\right| \,\middle|\, d(x,y) \leq \frac{1}{s+1}\right\}\right] \longrightarrow 0 \quad \text{as } m \longrightarrow \infty.$$

This condition implies the one in the remark after the proof of Theorem 3.1. The following theorem is the main tool for studying equilibrium states for one-sided subshifts of finite type. It was stated in the form given here by Bowen [2]. We shall prove it using Theorem 3.1. We use the symbol $\underset{\longrightarrow}{\longrightarrow}$ to denote convergence in $C(X)$.

THEOREM 3.3 (RUELLE'S OPERATOR THEOREM). *Let T be a topologically mixing one-sided subshift of finite type. Let $\phi \in C(X)$ satisfy $\Sigma_{n=0}^{\infty} \operatorname{var}_n(\phi) < \infty$. There exists a number $\lambda > 0$, $h \in C(X)$, and $\nu \in M(X)$ such that $h > 0$, $\nu(h) = 1$, $L_\phi h = \lambda h$, $L_\phi^* \nu = \lambda \nu$, and $L_\phi^n f/\lambda^n \underset{\longrightarrow}{\longrightarrow} \nu(f) \cdot h$.*

PROOF. The construction of λ, h and ν will be taken from Ruelle's proof [8], and we shall use g-measures to get the convergence property. We write L instead of L_ϕ.

1. To obtain λ and ν consider the mapping of $M(X)$ to itself given by $m \longrightarrow L^* m/(L_m^*)1$. The Schauder-Tycharoff fixed point theorem [3, p. 456] gives a fixed point $\nu \in M(X)$ for this map. Therefore $L^*\nu = \lambda\nu$ if $\lambda = (L_\nu^*)1$.

2. To obtain h, put $B_k = \exp \Sigma_{k+1}^{\infty} \operatorname{var}_j(\phi)$ and consider the set

$$\Lambda = \{f \in C(X) \mid \nu(f) = 1, f > 0, \text{ and for all } k \geq 1\ f(x) \leq B_k f(y) \text{ if } d(x, y) \leq 1/(k+1)\}.$$

We want to show Λ is a convex compact subset of $C(X)$. It is convex because if $f, h \in \Lambda$ and $\alpha > 0$, $\beta = 1 - \alpha > 0$, then

$$\alpha f(x) + \beta h(x) \leq \max\left(\frac{f(x)}{f(y)}, \frac{h(x)}{h(y)}\right)(\alpha f(y) + \beta h(y))$$

when $d(x, y) \leq 1/(k+1)$. To show Λ is compact we show it is bounded, closed and equicontinuous and then apply the Arzela-Ascoli theorem. To prove boundedness consider any points $x, z \in X$ and choose $v \in T^{-M}z$ with $v_0 = x_0$ (here $A^M > 0$ by topological mixing). If $f \in \Lambda$ then

$$(L^M f)(z) = \sum_{y \in T^{-M}z} e^{\phi(y) + \phi(Ty) + \cdots + \phi(T^{M-1}y)} f(y)$$

$$\geq e^{\phi(v) + \phi(Tv) + \cdots + \phi(T^{M-1}v)} f(v) \geq e^{-M\|\phi\|} B_1^{-1} f(x).$$

Therefore $f(x) \leq (L^M f(z)/\lambda^M) \cdot \lambda^M e^{M\|\phi\|} B_1$ for all z and so

$$f(x) \leq \lambda^M e^{M\|\phi\|} B_1 \nu(L^M f/\lambda^M) = \lambda^M e^{M\|\phi\|} B_1 = K.$$

Hence f is bounded. If $f \in \Lambda$ and $d(x, y) \leq 1/(k+1)$, then $f(y) - f(x) \leq f(x)[B_k - 1] \leq K[B_k - 1]$ and $B_k \longrightarrow 1$ as $k \longrightarrow \infty$. Therefore Λ is equicontinuous. Since Λ is clearly a closed subset of $C(X)$ we have that Λ is compact.

Define $\mathcal{L}$: $C(X) \longrightarrow C(X)$ by $\mathcal{L} = \lambda^{-1} L$. We want to show $\mathcal{L}\Lambda \subset \Lambda$ and then apply the Schauder-Tychanoff theorem. Let $f \in \Lambda$. $\nu(Lf) = \nu(f) = 1$ by the construction of ν. Clearly $Lf \geqslant 0$. If $d(x, y) \leqslant 1/(k+1)$, then

$$(Lf)(x) = \lambda^{-1} \sum_{u \in T^{-1}x} e^{\phi(u)} f(u) = \lambda^{-1} \sum_{i: A_{ix_0} = 1} e^{\phi(ix)} f(ix)$$

and

$$(Lf)(y) = \lambda^{-1} \sum_{v \in T^{-1}y} e^{\phi(v)} f(v) = \lambda^{-1} \sum_{i: A_{iy_0} = 1} e^{\phi(iy)} f(iy)$$

and each sum runs over the same values of i since $x_0 = y_0$. Since ix and iy have their first $k+1$ coordinates equal (i.e. $d(ix, iy) \leqslant 1/(k+2)$) we have

$$(Lf)(x) \leqslant B_{k+1} e^{\operatorname{var}_{k+1}(\phi)} (Lf)(y) = B_k (Lf)(y).$$

Therefore $L\Lambda \subset \Lambda$ and let h be a fixed point of L. We wish to show $h > 0$. If $h(x) = 0$ for some x then since $Lh = h$ we have $h(y) = 0$ if $y \in T^{-1}x$. Similarly h vanishes on the dense set $\bigcup_0^\infty \{T^{-n}x\}$ and so $h = 0$, contradicting $\nu(h) = 1$.

3. It remains to show the convergence property. Put $g(x) = e^{\phi(x)} h(x)/\lambda h(Tx)$. Then $g > 0$, g is continuous and $\Sigma_{y \in T^{-1}x}\, g(y) = 1$. Also $(L^n f)(x)/\lambda^n = h(x)(L^n_{\log g}(f/h))(x)$. If we can show g satisfies the condition in Theorem 3.1 then we shall get $L^n_\phi f/\lambda^n \longrightarrow h \cdot \mu(f/h)$ where μ is the unique g-measure. However the measure m defined by $m(f) = \nu(f \cdot h)$, $f \in C(X)$, is a g-measure because it is a probability measure and $m(L_{\log g} f) = \nu(h \cdot L_{\log g} f) = \nu(L_\phi(f \cdot h))\lambda^{-1} = \nu(f \cdot h) = m(f)$. Therefore $\mu(f/h) = \nu(f)$ and $L^n_\phi f/\lambda^n \rightrightarrows \nu(f) \cdot h$. We show g satisfies the condition at the foot of p. 382. Let $d(x, y) \leqslant 1/(s+1)$ and $s > m$. Then

$$\left| \frac{g(y) \cdots g(T^{s-m-1}y)}{g(x) \cdots g(T^{s-m-1}x)} - 1 \right|$$

$$= \left| \exp(\phi(x) - \phi(y) + \cdots + \phi(T^{s-m-1}x) - \phi(T^{s-m-1}y)) \frac{h(y)}{h(x)} \frac{h(T^{s-m}x)}{h(T^{s-m}y)} - 1 \right|$$

$$\leqslant \max \left[\exp\left(\sum_{i=m}^{s} \operatorname{var}_i(\phi) + \sum_{i=s+1}^{\infty} \operatorname{var}_i(\phi) + \sum_{i=m}^{\infty} \operatorname{var}_i(\phi) \right) - 1, \right.$$
$$\left. 1 - \exp\left(-\sum_{i=m}^{s} \operatorname{var}_i(\phi) - \sum_{i=s+1}^{\infty} \operatorname{var}_i(\phi) - \sum_{i=m}^{\infty} \operatorname{var}_i(\phi) \right) \right]$$

$$= \max \left[\exp\left(2 \sum_{i=m}^{\infty} \operatorname{var}_i(\phi) \right) - 1,\ 1 - \exp\left(-2 \sum_{i=m}^{\infty} \operatorname{var}_i(\phi) \right) \right].$$

$$\therefore \sup_{s>m}\left[\sup\left\{\left|\frac{g(y)\cdots g(T^{s-m-1}y)}{g(x)\cdots g(T^{s-m-1}x)}-1\right| \,\Big|\, d(x,y)\leqslant\frac{1}{s+1}\right\}\right]\longrightarrow 0 \quad \text{as } m\longrightarrow\infty.$$

COROLLARY 3.3(i). *Suppose $T: X \longrightarrow X$ and $\phi \in C(X)$ are as in Theorem 3.3. Then ϕ has unique equilibrium state μ_ϕ and $\mu_\phi(f) = \nu(h \cdot f), f \in C(X)$. μ_ϕ is the unique g-measure for $g = e^\phi \cdot h/\lambda h(T)$. With respect to μ_ϕ T is an exact endomorphism (and hence strong mixing) and its natural extension is a Bernoulli automorphism. μ_ϕ is positive on nonempty open sets. $P_T(\phi) = \log \lambda$ and λ is the spectral radius of L_ϕ: $C(X) \longrightarrow C(X)$. Also* $(\log L_\phi^n 1)/n \rightrightarrows P_T(\phi)$.

PROOF. We have $\phi = \log g + \log h \circ T - \log h + \log \lambda$ so ϕ and $\log g$ have the same equilibrium states by Lemma 1.1. Since we know $L_{\log g}^n f \rightrightarrows \nu(f \cdot h)$ we know $\log g$ has a unique equilibrium state which is given by $\mu_\phi(f) = \nu(f \cdot h)$, $f \in C(X)$. Therefore μ_ϕ is the unique equilibrium state for ϕ. We saw in the proof of Theorem 3.3 that is it a g-measure. By Theorem 3.2 we know that the natural extension of T relative to μ_ϕ is a Bernoulli automorphism and that (T, μ_ϕ) is exact. μ_ϕ is positive on open sets by Lemma 2.1.

Since μ_ϕ is the equilibrium state for $\phi = \log g + \log h \circ T - \log h + \log \lambda$,

$$P_T(\phi) = h_{\mu_\phi}(T) + \mu_\phi(\phi) = h_{\mu_\phi}(T) + \mu_\phi(\log g) + \log \lambda = \log \lambda,$$

since μ_ϕ is a g-measure.

The spectral radius of L_ϕ is given by $\lim_{n\to\infty} \|L_\phi^n\|^{1/n} = \lim_{n\to\infty} \|L_\phi^n 1\|^{1/n} = \lambda$ by Theorem 3.3. Since $L_\phi^n 1/\lambda^n \rightrightarrows h$ we get $(\log L_\phi^n 1)/n \longrightarrow P_T(\phi)$.

Note that λ, ν and $h > 0$ are all uniquely determined by $L_\phi^n f/\lambda^n \rightrightarrows h \cdot \nu(f)$ since $\log \lambda = \lim_{n\to\infty} n^{-1} \log L_\phi^n 1$, and $h = \lim_{n\to\infty} L_\phi^n 1/\lambda^n$.□

The above corollary generalises the result of W. Parry [7] that a topologically mixing subshift of finite type has a unique measure which maximises entropy. This is the case $\phi \equiv 0$ since then $P_T(0) = h(T)$, the topological entropy of T and an equilibrium measure for $\phi \equiv 0$ is a member μ of $M(T)$ with $h_\mu(T) = h(T)$. Let μ_0 denote the unique measure with $h_{\mu_0}(T) = h(T)$. Sinai has shown how to define a class of measures he calls Gibbs measures from a given $\mu \in M(T)$ and a given $\phi \in C(X)$ [11]. If $\Sigma_{n=0}^\infty \operatorname{var}_n(\phi) < \infty$ it can be shown that there is a unique Gibbs measure corresponding to μ_0 and ϕ, and this measure is μ_ϕ. The next corollary has been proved by Sinai from the point of view of Gibbs measures [11, p. 30].

COROLLARY 3.3(ii). *Let $T: X \longrightarrow X$ be as in Theorem 3.3 and $\phi, \psi \in C(X)$ satisfy $\Sigma_0^\infty \operatorname{var}_n(\phi) < \infty$, $\Sigma_0^\infty \operatorname{var}_n(\psi) < \infty$. Then $\mu_\phi = \mu_\psi$ if and only if $\phi - \psi = f \circ T - f + c$ for some $f \in C(X)$ and some $c \in R$.*

PROOF. If $\psi - \psi = f \circ T - f + c$ then $\mu_\phi = \mu_\psi$ by Lemma 1.1. Suppose $\mu_\phi = \mu_\psi$. By Corollary 3.3(i) μ_ϕ is the unique g_1-measure for some $g_1 =$

$e^{\phi}h_1/(\lambda_1 \cdot h_1 \circ T)$ and is the unique g_2-measure for some $g_2 = e^{\phi}h_2/(\lambda_2 \cdot h_2 \circ T)$. By Lemma 2.1 $g_1 = g_2$. Therefore

$$\phi - \psi = \log \frac{h_1 \circ T}{h_2} - \log \frac{h_1}{h_2} + \log \frac{\lambda_1}{\lambda_2}.$$

4. **Generalisations.** Suppose (X, d) is a compact metric space and T is a continuous transformation of X onto X which satisfies

(a) there exists k such that $\{T^{-1}x\}$ has less then k members, for all x,

(b) T is a local homeomorphism, and

(c) for sufficiently small $\delta > 0$, $d(x, y) = \delta$ implies $d(Tx, Ty) \geqslant \delta$.

If $\phi \in C(X)$ then L_ϕ: $C(X) \longrightarrow C(X)$ can be defined by $(L_\phi f)(x) = \Sigma_{y \in T^{-1}x} e^{\phi}(y) f(y)$. If $\epsilon > 0$, then let

$$\operatorname{var}_n(\phi, \epsilon) = \sup\{|\phi(x) - \phi(y)| \mid d(T^i x, T^i y) \leqslant \epsilon, 0 \leqslant i \leqslant n-1\}.$$

Let $G(T) = \{g \in C(X) | g > 0$ and $\Sigma_{y \in T^{-1}x}\, g(y) = 1$ for all $x \in X\}$. A similar proof to that of Theorem 3.1, 3.3 and yields

THEOREM 4.1. *Let* T: $X \longrightarrow X$ *satisfy* (a), (b), (c) *above and also,* (d) $\forall \epsilon > 0$ $\exists N$ *such that* $\forall x \in X$ $T^{-N}x$ *is* ϵ*-dense. Then if* $\Sigma_{n=0}^{\infty} \operatorname{var}_n(\phi, \epsilon) \longrightarrow 0$ *as* $\epsilon \longrightarrow 0$ *there exists a number* $\lambda > 0$, $\nu \in M(X)$, $h \in C(X)$ *such that* $h > 0$, $\nu(h) = 1$, $L_\phi^* \nu = \lambda\nu$, $L_\phi h = \lambda h$, *and* $L_\phi^n f/\lambda^n \xrightarrow{=} h \cdot \nu(f)$ *for all* $f \in C(X)$. *When* $\phi = \log g$, $g \in G(T)$, *then* $\lambda = 1$ *and* $h \equiv 1$.

A continuous transformation S: $X \longrightarrow X$ is positively expansive if there exists $\delta > 0$ so that $d(S^n x, S^n y) \leqslant \delta$ for all $n \geqslant 0$ implies $x = y$. Then any finite partition of X into Borel sets of diameter less than δ is a one-sided generator for all $\mu \in M(S)$ and so $h_\mu(S) = H_\mu(\mathcal{B}/S^{-1}\mathcal{B})$ for all $\mu \in M(S)$. This allows us to use the proof of Theorem 2.1 to prove the following result.

THEOREM 4.2. *Let* T *be a positively expansive transformation satisfying* (a), (b) *and* (c). *Suppose* $g \in G(T)$ *and let* L *denote* $L_{\log g}$.

If $m \in M(X)$ *the following are equivalent*:

(i) $L^* m = m$,

(ii) $m \in M(T)$ *and* $E_m(f/T^{-1}\mathcal{B})(x) = \Sigma_{z \in T^{-1}Tx}\, g(z) f(z)$ *a.e.* m *for* $f \in L^1(m)$,

(iii) $m \in M(T)$ *and* m *is an equilibrum state for* $\log g$.

A measure satisfying these conditions is called a g-measure.

Using this we can deduce

THEOREM 4.3. *Let* T *be positively expansive and satisfy* (a), (b), (c), (d). *Let* ϕ *be as in Theorem* 4.1. *Then* ϕ *has a unique equilibrium state* μ_ϕ *and* $\mu_\phi(f) = \nu(h \cdot f), f \in C(X)$. μ_ϕ *is the unique* g*-measure for* $g = e^{\phi}h/(\lambda \cdot h \circ T)$. *With* *respect to μ_ϕ the natural extension of T is a Bernoulli shift. μ_ϕ is positive on open sets. Also $P_T(\phi) = \log \lambda$ and λ is the spectral radius of L_ϕ.*

When proving the Bernoulli property one uses for γ any partition of X into Borel sets of diameter less than δ. The proof is slightly more complicated than the proof in Theorem 3.2 because the characteristic function of a set of the form $\bigcap_{i=0}^n T^{-1}B_i$, $B_i \in \gamma$, may not be continuous.

REFERENCES

1. R. Bowen, *Markov partitions for Axiom* A *diffeomorphisms*, Amer. J. Math. 92 (1970), 725–747. MR 43 #2740.

2. ———, *Bernoulli equilibrium states for Axiom* A *diffeomorphisms*, Math. Systems Theory 8 (1975), 289–294.

3. N. Dunford and J. T. Schwartz, *Linear operators.* I: *General theory*, Pure and Appl. Math., vol. 7, Interscience, New York, 1958. MR 22 #8302.

4. N. A. Friedman and D. S. Ornstein, *On the isomorphism of weak Bernoulli transformations*, Advances in Math. 5 (1970), 365–394. MR 43 #478c.

5. M. Keane, *Strongly mixing g-measures*, Invent. Math. 16 (1972), 309–324. MR 46 #9295.

6. F. Ledrappier, *Principe variationnel et systemes dynamiques symboliques*, preprint, 1974.

7. W. Parry, *Intrinsic Markov chains*, Trans. Amer. Math. Soc. 114 (1965), 55–66.

8. D. Ruelle, *Statistical mechanics of a one-dimensional lattice gas*, Comm. Math. Phys. 9 (1968), 267–278. MR 38 #3013.

9. V. A. Rohlin, *Lectures on the entropy theory of transformations with invariant measure*, Uspehi Mat. Nauk 22 (1967), no. 5 (137), 3–56 = Russian Math. Surveys 22 (1967), 1–52. MR 36 #349.

10. Ja. G. Sinaĭ, *Markov partitions and Y-diffeomorphisms*, Funkcional. Anal. i Priložen. 2 (1968), no. 1, 64–89 = Fuctional Anal. Appl. 2 (1968), 61–82. MR 38 #1361.

11. ———; *Gibbs measures in ergodic theory*, Uspehi Mat. Nauk 27 (1972), no. 4 (166), 21–64 = Russian Math. Surveys 27 (1972), no. 4, 21–70.

12. P. Walters, *A variational principle for the pressure of continuous transformations*, Amer. J. Math. (to appear).

MATHEMATICS INSTITUTE, UNIVERSITY OF WARWICK, COVENTRY, WARWICKSHIRE CV4 7AL, ENGLAND