CHAPTER 12

STATISTICAL PROPERTIES IN ERGODIC THEORY

12.1 Exact endomorphisms

DEFINITION. We call a measure preserving transformation $T : X \to X$ on a probability space $(X, \mathcal{B}, \mu)$ an *exact endomorphism* if $\cap_{n=0}^{\infty} T^{-n}\mathcal{B} = \{X, \emptyset\}$ up to a set of zero measure (i.e. if $B \in T^{-n}\mathcal{B}$, for every $n \geq 0$, then $\mu(B) = 0$ or $\mu(B) = 1$).

PROPOSITION 12.1. *$T : X \to X$ is exact if for any positive measure set A with $T^n A \in \mathcal{B}(n \geq 0), \mu(T^n(A)) \to 1$ as $n \to +\infty$.*

It is easy to see that this sufficient condition for exactness is also necessary [2] (although we will not need this here).

PROOF. First we remark that T is exact if every measurable set A satisfying for arbitrary n the relationship $A = T^{-n}(T^n A)$ is of either measure zero or measure 1. For such a set A, it is clear that $\mu(A) = 1$ if $\mu(A) > 0$, as $\mu(T^n A) = \mu(A)$ and so $\lim_{n\to\infty} \mu(T^n A) = \mu(A) = 1$ if $\mu(A) > 0$.

■

PROPOSITION 12.2. *If T is exact then it is strong-mixing.*

PROOF. Consider the sub-sigma-algebras $\mathcal{B} \supset T^{-1}\mathcal{B} \supset T^{-2}\mathcal{B} \supset \ldots \supset \{X, \emptyset\}$. We can associate the nested subspaces $L^2(\mathcal{B}) \supset L^2(T^{-1}\mathcal{B}) \supset L^2(T^{-2}\mathcal{B}) \supset \ldots \supset \mathbb{C}$ and for each $n \neq 0$ we can choose an orthonormal basis $\{k_i \circ T^n\}_{i=0}^{j_n}$ for $L^2(T^{-n}\mathcal{B}) \ominus L^2(T^{-(n+1)}\mathcal{B})$. It follows that $\{k_i \circ T^n\}_{i=0}^{j_n}{}_{n=0}^{\infty}$ is an orthonormal basis for $L^2(X, B, \mu)$. Two functions $f, g \in L^2(X, B, \mu) \ominus \mathbb{R}$ can be written in the form

$$\begin{cases} f = \sum_{n=0}^{\infty} \sum_i a_{n,i} k_i \circ T^n + \left(\int f d\mu\right), \\ g = \sum_{n=0}^{\infty} \sum_i b_{n,i} k_i \circ T^n + \left(\int g d\mu\right), \end{cases}$$

where $a_{n,i}, b_{n,i} \in \mathbb{R}$. In particular,

$$\int f \circ T^N g d\mu = \sum_{n=0}^{\infty} \sum_i a_{n,i} b_{n+N,i} + \int f d\mu \int g d\mu \to \int f d\mu \int g d\mu$$

as $N \to +\infty$, i.e. T is strong-mixing.

■

EXAMPLE 1 (ONE-SIDED APERIODIC MARKOV SHIFTS). We can modify the definition of the Markov shift and define

$$X_A^+ = \{x \in \prod_{n\in\mathbb{N}^+} \{0, \dots, k-1\} : A(x_n, x_{n+1}) = 1, n \in \mathbb{Z}^+\}$$

and $\sigma : X_A^+ \to X_A^+$ by $(\sigma x)_n = x_{n+1}$. For the stochastic matrix P (with entries $P(i,j) = 0$ iff $A(i,j) = 0$) letting p be its left eigenvector we define the measure on a cylinder

$$[i_0, \dots, i_{l-1}] = \{x \in X_A^+ : x_j = i_j, 0 \le j \le l-1\},$$

$$\mu[i_0, \dots, i_{l-1}] = p(i_0)P(i_0, i_1)\dots P(i_{l-2}, i_{l-1}).$$

Let A be aperiodic. Then the argument for the (two sided) Markov shift still applies and we see that T is strong-mixing; moreover, $\forall \epsilon > 0, \forall$ cylinders C, $\exists N > 0$ such that $\forall n \ge N$ and any cylinder D we have $|\mu(C \cap T^{-n}D) - \mu(C)\mu(D)| \le \epsilon\mu(C)\mu(D)$. By approximating an arbitrary set $B \in \mathcal{B}$ by a cylinder D we see that the same result holds on replacing D by B.

Assume that $E \in \cap_{n=0}^{\infty} T^{-n}\mathcal{B}$ and write $E = T^{-n}E_n$. For any cylinder C we see from the above observations that

$$\mu(C \cap E) = \mu(C \cap T^{-n}E_n) \ge (1-\epsilon)\mu(E_n)\mu(C) = (1-\epsilon)\mu(E)\mu(C);$$

since $\epsilon > 0$ is arbitrary we see that $\mu(C \cap E) \ge \mu(E)\mu(C)$ for all cylinders C. By approximation by disjoint unions of cylinders we can replace this by $\mu(B \cap E) \ge \mu(E)\mu(B)$, $\forall B \in \mathcal{B}$. If we take $B = X - E$ we see that $\mu(E)\mu(X - E) = 0$. This completes the proof that T is exact.

12.2 Statistical properties of piecewise expanding Markov maps

Consider a piecewise expanding C^2 surjective Markov map $T : I \to I$ for which there exists $\beta > 1$ with $\inf_{x\in I} |T'(x)| \ge \beta$. We can define an operator $\mathcal{L} : L^1(I) \to L^1(I)$ as follows.

DEFINITION. Given $f \in L^1(I)$ we define the *Perron-Frobenius operator* by

$$\mathcal{L}f(x) = \sum_{y\in T^{-1}x} \frac{f(y)}{|T'(y)|} \left(= \sum_{i=1}^{k} f(\psi_i x)|\psi_i'(x)|\chi_{TI_i}(x) \right)$$

(where ψ_i denotes the inverse of $T|I_i$).

LEMMA 12.3. *For any $f \in L^1(I)$ satisfying $(\mathcal{L}f)(x) = f(x)$ the measure μ defined by $f = \frac{d\mu}{dx}$ is T-invariant.*

PROOF. This follows from the change of variables formula since we have $\mu(T^{-1}A) = \int_{T^{-1}A} f(x)dx = \sum_{i=1}^{k} \int_{TI_i \cap A} |\psi_i'(x)| f \circ \psi_i(x)dx = \int_A \mathcal{L}f(x)dx = \mu(A)$.

■

We have the following result.

PROPOSITION 12.4 (SMOOTH INVARIANT MEASURES FOR PIECEWISE EXPANDING MARKOV MAPS). *There exists an invariant probability measure μ which is absolutely continuous with respect to the (normalized) Haar-Lebesgue measure λ (i.e. there exists $f \in L^1(I)$ such that $\mu(B) = \int_B f(x)d\lambda(x)$ for every Borel set $B \in \mathcal{B}$).*

PROOF. By Lemma 12.3, to construct μ it suffices to find such a function f satisfying $\mathcal{L}f = f$. We first choose a point $x \in I$ and for any $n \geq 1$ we look at the families $T^{-n}x$ of all n-iterate pre-images of x.

It is easy to see from the chain rule that

$$\mathcal{L}^n 1(x) = \sum_{y \in T^{-1}x} \frac{\mathcal{L}^{n-1}1(y)}{|T'(y)|} = \sum_{y \in T^{-n}x} \frac{1}{|T^{n\prime}(y)|}.$$

We denote the inverse of $T^n | \cap_{j=0}^{n-1} T^{-j} I_{i_{j+1}}$ by $\psi_{i_1 \dots i_n}$. Let $\mathcal{V}$ be the partition generated by $\{T(I_i) : 1 \leq i \leq k\}$. Then for $x, x' \in V \in \mathcal{V}$ we can compare

$$\begin{aligned} |\mathcal{L}^n 1(x) - \mathcal{L}^n 1(x')| &= \Big| \sum_{y \in T^{-n}x} \frac{1}{|T^{n\prime}(y)|} - \sum_{y' \in T^{-n}x'} \frac{1}{|T^{n\prime}(y')|} \Big| \\ &= \sum_{i_1, \dots, i_n} \big| |\psi_{i_1 \dots i_n}'(x)| - |\psi_{i_1 \dots i_n}'(x')| \big| \chi_{T^n I_{i_1 \dots i_n}}(x) \end{aligned}$$

where $I_{i_1 \dots i_n} = \cap_{j=0}^{n-1} T^{-j} I_{i_{j+1}}$. Observe that

$$\begin{aligned} \log \Big| \frac{\psi_{i_1 \dots i_n}'(x')}{\psi_{i_1 \dots i_n}'(x)} \Big| &= \sum_{j=1}^{n} \log \Big| \frac{\psi_{i_j}'\left(\psi_{i_{j+1} \dots i_n} x'\right)}{\psi_{i_j}'\left(\psi_{i_{j+1} \dots i_n} x\right)} \Big| \\ &= \sum_{j=1}^{n} \log \Big| \frac{T'\left(\psi_{i_j \dots i_n} x'\right)}{T'\left(\psi_{i_j \dots i_n} x\right)} \Big| \\ &\leq \sum_{j=1}^{n} \log \left(1 + D \frac{|x - x'|}{\beta^{n-j}}\right) \end{aligned}$$

where D bounds $\frac{|T''|}{|T'|}$ on I. Then we have a constant $C > 1$ such that

$$\frac{\sup_{x \in TI_{i_n}} |\psi'_{i_1 \dots i_n}(x)|}{\inf_{x \in TI_{i_n}} |\psi'_{i_1 \dots i_n}(x)|} \leq C, \quad \forall i_1, \dots, i_n, n > 0.$$

The property allows us to find a constant $K < +\infty$ such that

$$\sum_{i_1, \dots, i_n} |\psi'_{i_1 \dots i_n}(x)| \chi_{TI_{i_n}}(x) \leq K, \quad \forall x.$$

We conclude that there exists $D' > 0$ such that $|\mathcal{L}^n 1(x) - \mathcal{L}^n 1(x')| \leq K \left(|x-y| \frac{D'}{1 - \frac{1}{\beta}} \right)$

(where none of the bounds on the right hand side depends on n). We conclude that $\forall n \geq 1$

(1) the functions $\mathcal{L}^n 1$ are bounded in the supremum norm,
(2) the functions $\mathcal{L}^n 1$ are an equicontinuous family.

We construct a new family of averages

$$F_n(x) = \frac{1}{n} \sum_{k=0}^{n-1} \mathcal{L}^k 1(x), \qquad n \geq 0.$$

We again see that

(1) the functions F_n are bounded in the supremum norm,
(2) the functions F_n are an equicontinuous family.

By the Ascoli theorem, there must be a limit point $F_{n_r} \to f \quad (\geq 0)$ in the continuous functions on each component of $[0,1] - \{x_0, \dots, x_k\}$ and since $\int \mathcal{L}^n 1 dx = 1$ we have $\int f d\lambda = \lim_{n \to \infty} \int F_{n_r} d\lambda = 1$. Moreover, we see that

$$\begin{aligned}
\mathcal{L}F_{n_r}(x) = \sum_{y \in T^{-1}x} \frac{F_{n_r}(y)}{|T'(y)|} &= \sum_{y \in T^{-1}x} \frac{1}{n_r} \sum_{k=0}^{n_r - 1} \frac{\mathcal{L}^k 1(y)}{|T'(y)|} \\
&= \frac{1}{n_r} \sum_{k=0}^{n_r - 1} \sum_{y \in T^{-1}x} \frac{\mathcal{L}^k 1(y)}{|T'(y)|} \\
&= \frac{1}{n_r} \sum_{k=0}^{n_r - 1} \mathcal{L}^{k+1} 1(x) \\
&= F_{n_r}(x) - \frac{1}{n_r} \left(1(x) - \mathcal{L}^{n_r} 1(x) \right).
\end{aligned}$$

Letting $r \to +\infty$ we see that

$$\mathcal{L}f(x) = \sum_{y \in T^{-1}x} \frac{f(y)}{|T'(y)|} = f(x),$$

completing the proof.

■

DEFINITION. We say that T is aperiodic if there exists a positive number m such that $\lambda(T^{-m}I_i \cap I_j) > 0$, $\forall i, j > 0$.

THEOREM 12.5. *The absolutely continuous invariant measure μ in Proposition 12.4 is exact if $T : I \to I$ is aperiodic.*

PROOF. By Proposition 12.1 it suffices to show that for any set $A \in \mathcal{B}$ with $\mu(A) > 0$ and for which $T^n A \in \mathcal{B}$ for all $n \geq 0$ we have that $\lim_{n\to+\infty} \mu(T^n A) = 1$.

Given $\underline{i} = (i_1, \dots, i_n)$ we write $I_{\underline{i}} = \cap_{j=1}^{n} T^{-j+1} I_{i_j}$ if $\mathrm{int}\left(\cap_{j=1}^{n} T^{-j+1} I_{i_j}\right) \neq \emptyset$. As T is piecewise invertible on each atom I_i, we know that $T^n|_{I_{\underline{i}}}$ is a C^1-diffeomorphism. For all $I_{\underline{i}}$ and for all $n > 0$ we write $(T^n|_{I_{\underline{i}}})^{-1} = \psi_{\underline{i}}$. Let $x, y \in T^n I_{\underline{i}} \quad (= T I_{i_n})$; then it follows from the mean value theorem that

$$|\psi_{\underline{i}}(x) - \psi_{\underline{i}}(y)| = |\psi'_{\underline{i}}(\theta)||x - y|$$

for some $\theta \in I_{\underline{i}}$. From the above equality and the condition (i) in page 39, the diameter diam $\left(I_{\underline{i}}\right)$ of $I_{\underline{i}}$ decays exponentially fast (i.e., $\mathrm{diam}(I_{\underline{i}}) \leq \frac{1}{\beta^n}$). This implies that the partition $\mathcal{I} = \{I_i\}$ is a " generating partition". In particular, for any $\epsilon > 0$ we can choose a finite disjoint set of cylinders $\{I_{\underline{j}} : \underline{j} = (j_1, \dots, j_l)\}$, with $\mu\left(\left(\cup_{\underline{j}} I_{\underline{j}}\right) \Delta A\right) < \epsilon$. The following estimates will be useful in the rest of the proof.

(a) Given $\delta > 0$ there exists at least one cylinder $I_{\underline{j}}$ (where $\underline{j} = (j_1 \dots j_l)$, say) for which

$$\lambda(A \cap I_{\underline{j}}) \geq (1 - \delta)\, \lambda(I_{\underline{j}}). \tag{12.1}$$

Assume for a contradiction that this is not the case, then for all cylinders $I_{\underline{j}}$ we would have $\lambda(A \cap I_{\underline{j}}) \leq (1 - \delta)\lambda(I_{\underline{j}})$. We can extend this inequality to disjoint unions of cylinders, and then by approximation to arbitrary sets $B \in \mathcal{B}$ to get $\mu(A \cap B) \leq (1 - \delta)\mu(B)$. However, if we take $B = A$, then we get $\mu(A) \leq (1 - \delta)\mu(A)$, which contradicts $\mu(A) > 0$.

(b) We observe that there is a constant $C \geq 1$ such that for any cylinder $I_{\underline{i}}$

$$\sup_{x,y \in T^n I_{\underline{i}}} \frac{|{\psi_{\underline{i}}}'(x)|}{|{\psi_{\underline{i}}}'(y)|} \leq C. \tag{12.2}$$

(This is usually referred to as *Renyi's condition.*)

From the change of variables formula we see that

$$\begin{aligned}
\lambda(T^l I_{\underline{j}} \cap (T^l A)^c) &\leq \int_{I_{\underline{j}} \cap A^c} |(T^l)'(x)| d\lambda(x) \\
&\leq \left(\sup_{y \in I_{\underline{j}}} |(T^l)'(y)| \right) \lambda(I_{\underline{j}} \cap A^c) \\
&\leq C \left(\inf_{y \in I_{\underline{j}}} |(T^l)'(y)| \right) \lambda(I_j \cap A^c) \qquad \text{(using (12.2))} \\
&\leq C \frac{\int_{I_j} |(T^l)'(x)| d\lambda(x)}{\lambda(I_j)} \lambda(I_j \cap A^c) \\
&\leq C\delta\lambda(T^l I_j) \qquad \text{(using (12.1))}.
\end{aligned}$$

If $T^l(I_{\underline{j}}) = I$, then we could proceed directly to the end of the proof. However, since this need not be the case, we require the following sublemma.

SUBLEMMA 12.5.1. *There exist $S > 0$ and a subset I' of $T^l(I_{\underline{j}})$ which is a finite disjoint union of elements of $\vee_{i=0}^{S-1} T^{-i}\{I_1 \dots I_k\}$ and satidfies $T^S(I') = I$.*

PROOF. Let $\{U_1, \dots, U_N\} = \{TI_1, \dots, TI_k\}$, where $N \leq k$, denote the collection of images under T of the original intervals. The aperiodicity assumption implies that for each $1 \leq j \leq N$ there exists $0 < s_j < +\infty$ such that each U_i, $i = 1, \dots, N$, contains a cylinder $I^{(i,j)}_{m_1,\dots,m_{s_j}}$ satisfying $T^{s_j} I^{(i,j)}_{m_1,\dots,m_{s_j}} = U_j$. In particular, we see that $T^{s_i} U_i \supset T^{s_i} I^{(i,j)}_{m_1,\dots,m_{s_j}} = U_i$.

Let $T^l(I_{\underline{j}}) = U_i$. Setting $S = \prod_{j=1}^N s_j$ and $I' = \cup_{j=1}^N I^{(i,j)}_{m_1,\dots,m_{s_j}}$ allows us to have that $I' \subset T^l(I_{\underline{j}})$ and $T^S I' \supset \cup_{j=1}^N U_j = X$.

■

We need only modify the previous argument to write

$$\lambda(T^S(I' \cap (T^l A)^c)) \leq D\delta$$

for some uniform constant $D > 0$. Since $\lambda(T^S(I' \cap (T^l A)^c)) \geq 1 - \lambda(T^S(I' \cap T^l A))$, we see that

$$\lambda(T^{l+S} A) > \lambda(T^S(I' \cap T^l A)) > 1 - D\delta.$$

Since μ is absolutely continuous with respect to λ we conclude that $\mu(T^n A) \to 1$ as $n \to +\infty$.

■

COROLLARY 12.5.1. *If $T : I \to I$ is aperiodic, then it is strong-mixing with respect to any absolutely continuous invariant measure. In particular, there exists a unique absolutely continuous invariant probability measure.*

PROOF. By Proposition 12.1 the exact measure μ is also strong mixing. By Proposition 11.2 it is also ergodic, and since no two distinct ergodic measures can be equivalent to Lebesgue measure (and thus each other) uniqueness follows.

PROPOSITION 12.6. *μ is equivalent to λ.*

PROOF. First we show the following fact:

$$\forall \epsilon > 0, \exists N(\epsilon) > 0 \text{ such that for each } x \in I, T^{-N(\epsilon)}x \text{ is } \epsilon\text{-dense in } I. \quad (12.3)$$

As we have already observed in Theorem 12.5, for $\forall I_{j_1 \dots j_l}$ there exist a set of cylinders $\{I^{(i)}_{m_1 \dots m_{s_i}} : i = 1, \dots N\}$ and $S > 0$ satisfying $T^S(\cup_{i=1}^N I^{(i)}_{m_1 \dots m_{s_i}}) = I$. Let $x \in I_{h_1 \dots h_t}$. Then $\exists i$ s.t. $m_1 \dots m_{s_i} h_1 \dots h_t$ is an admissible sequence and so $\psi_{m_1 \dots m_{s_i}}(x) \in I_{m_1 \dots m_{s_i}} \subset T^l I_{j_1 \dots j_l}$. Hence we have that $T^{-(l+s_i+t)}x \cap I_{j_1 \dots j_l} \neq \emptyset$. Here we take $t = S - s_i$. Let $l = l(\epsilon)$ be a positive integer such that $\sup_{I_{j_1 \dots j_l}} \operatorname{diam} I_{j_1 \dots j_l} < \epsilon$. Then, each $I_{j_1 \dots j_l}$ contains at least a point belonging to $T^{-(l+S)}x$. Choosing $N(\epsilon) = l(\epsilon) - S$, we have the fact (12.3).

It remains to show that f is bounded away from zero. Assume for a contradiction that $f(x) = 0$. Then since for all $n \geq 1, \mathcal{L}^n f(x) = \sum_{y \in T^{-n}x} \frac{f(y)}{|T^{n\prime}(y)|} = 0$, we see that $f(y) = 0$ whenever $T^n y = x$. By the property (12.3) the set of such points is dense. The continuity of f implies that f is identically zero, contradicting $\int f d\lambda = 1$.

■

PROPOSITION 12.7. *For irreducible piecewise expanding Markov maps $T : I \to I$ the following condition is equivalent to strong mixing:*

$$\lambda \circ T^{-n}(A) \to \mu(A), \text{ as } n \to +\infty \quad (\forall A \in \mathcal{B}),$$

where λ is Lebesgue measure.

PROOF. It is enough to observe that

$$\begin{aligned}\lambda(T^{-n}A) &= \int_I \chi_{T^{-n}A}(x) d\lambda(x) = \int \chi_A(T^n x) f(x)^{-1} d\mu(x) \\ &\to (\int \chi_A(x) d(\mu(x)) \cdot (\int d\lambda(x)) = \mu(A).\end{aligned}$$

■

REMARK. Under the generating condition we can extend these results to multi-dimensional piecewise expanding Markov maps with countable infinite partitions.

Since the invariant density f is strictly positive, we can make the following definition.

DEFINITION. We define an operator $\hat{\mathcal{L}} : L^1(I) \to L^1(I)$ by $\hat{\mathcal{L}}(h) = \frac{1}{f}\mathcal{L}(fh)$ where $h \in L^1(X)$.

PROPOSITION 12.8. *$\hat{\mathcal{L}}^*(\mu) = \mu$, i.e. the dual operator $\hat{\mathcal{L}}^*$ acting on measures (defined by $(\mathcal{L}^*\mu)(A) = \int \mathcal{L}\chi_A d\mu$) fixes μ.*

PROOF. It is an immediate consequence of Sublemma 14.2.3 and the definition.

■

THEOREM 12.9 (CONVERGENCE TO INVARIANT DENSITY). *$\mathcal{L}^n(h) \to f\left(\int h d\lambda\right)$ uniformly for $h \in C^0(I)$.*

PROOF. Define $g = \frac{f(x)}{f(Tx)|T'(x)|}$. From Renyi's condition we have that there exists a uniform constant $D \geq 1$ such that $\forall x, x' \in U_k$

$$D(x, x') = \sup_{n \geq 1} \sup_{y \in T^{-n}x, y' \in T^{-n}x'} \prod_{i=1}^{n-1} \frac{|g(T^i y)|}{|g(T^i y')|}$$

is bounded above by D and furthermore

$$D(x, x') \to 1 \text{ as } |x - x'| \to 0.$$

An easy calculation shows that $\{\hat{\mathcal{L}}^n h : n \geq 0\}$ is equicontinuous on each component of $I - \partial\mathcal{V}$ for $\forall h \in C_0(I - \partial\mathcal{V})$. It follows from the definition of $\hat{\mathcal{L}}^n$ that $||\hat{\mathcal{L}}^n h||_\infty$ is bounded by $||h||_\infty$ and so the closure of $\{\hat{\mathcal{L}}^n h : n \geq 0\}$ in $C(I - \partial\mathcal{V})$ is compact. Hence there are a subsequence $\{n_i\} \to \infty$ $(i \to \infty)$ and $h^* \in C^0(I - \partial\mathcal{V})$ such that $\hat{\mathcal{L}}^{n_i} h \to h^*$ uniformly.

We can now show that any limit point of the sequence is a constant which, in particular, shows that the limit exists. Notice that $\min_{x \in I}(\hat{\mathcal{L}}^k h^*(x)) = \min_{x \in I}(h^*(x))$ for all $k \geq 0$. For any $k \geq 0$ choose $z \in I$ such that $\hat{\mathcal{L}}^k h^*(z) = \min_{x \in I} h^*(x)$. Then for all $y \in T^{-k}z$ we have that $h^*(y) = \min_{x \in I} h^*(x)$. In fact,

$$\hat{\mathcal{L}}^k h^*(z) = \sum_{T^k y = z} \left(g(y) \ldots g(T^{k-1}y)\right) h^*(y) \geq \min_{x \in I} h^*(x)$$

with equality if and only if $h^*(y) = \min_{x \in I} h^*(x)$, $\forall y \in T^{-n}x$. By (12.3) we see that the set of y such that $\exists k \geq 1$ with $T^k y = x$ is dense. Thus h^* is a constant function with value $\min_{x \in I} h^*(x)$ on a dense set, and thus by piecewise continuity is constant almost everywhere.

Moreover, this constant takes the value $\lim_{n \to +\infty} \int \hat{\mathcal{L}}^n h d\mu = \int h d\mu$. Replacing h by $\frac{h}{f}$ for $h \in C^0(I)$ and appealing to the definition of $\hat{\mathcal{L}}$ we get that

$$\mathcal{L}^n(h) = \mathcal{L}^n(f \cdot \frac{h}{f}) = f \cdot \hat{\mathcal{L}}^n(\frac{h}{f}) \to f\left(\int \frac{h}{f} \cdot f d\lambda\right) = f \cdot \left(\int h d\lambda\right)$$

uniformly as $n \to +\infty$.

■

12.3 Rohlin's entropy formula

In this section we want to give a formula for the entropy of an irreducible Markov piecewise expanding interval map $T : I \to I$ with respect to the unique absolutely continuous probability measure μ.

THEOREM 12.10 (ROHLIN ENTROPY FORMULA).

$$h_\mu(T) = \int \log |T'(T^i x)| d\mu(x).$$

PROOF. The proof follows immediately from the string of statements (i)-(iv) below.

(i) By the chain rule we can write $\log |(T^N)'(x)| = \sum_{i=0}^{N-1} \log |T'(T^i x)|$ for each $x \in I$, $N \geq 1$. Since the measure μ is ergodic (even exact) we can apply the Birkhoff ergodic theorem to deduce that

$$\frac{1}{N} \log |(T^N)'(x)| \to \int \log |T'(x)| d\mu(x) \text{ as } N \to +\infty.$$

(ii) Let $x \in I_{i_1, \dots, i_N} = \cap_{j=1}^N T^{-(j-1)} I_{i_j}$; then using Renyi's condition we can estimate

$$\begin{aligned} \lambda(I_{i_1 \dots i_N}) &= \int_{T^N I_{i_1 \dots i_N}} \frac{1}{|(T^N)'(\psi_{i_1 \dots i_N} z)|} d\lambda(z) \\ &\leq C \left(\inf_{x \in I_{i_1 \dots i_N})} \frac{1}{|(T^N)'(x)|} \right) \lambda(T I_{i_N}) \\ &\leq C \left(\frac{1}{(T^N)'(x)} \right) \end{aligned}$$

and

$$\begin{aligned} \lambda(I_{i_1 \dots i_N}) &\geq \frac{1}{C} \left(\sup_{z \in I_{i_1 \dots i_N}} \frac{1}{|(T^N)'(z)|} \right) \lambda(T(I_{i_N})) \\ &\geq \frac{1}{C} \left(\min_{1 \leq i \leq k} \lambda(I_i) \right) \frac{1}{|(T^N)'(x)|}. \end{aligned}$$

Thus we see that for any $x \in I$

$$- \lim_{N \to +\infty} \frac{1}{N} \log \lambda(I_{i_1 \dots i_N}) = \lim_{N \to +\infty} \frac{1}{N} \log |(T^N)'(x)|$$

(where $x \in I_{i_1 \dots i_N}$).

(iii) Since the density f of the invariant measure is bounded from below and away from infinity, we see that

$$-\lim_{N \to +\infty} \frac{1}{N} \log \lambda(I_{i_1 \dots i_N}) = -\lim_{N \to +\infty} \frac{1}{N} \log \mu(I_{i_1 \dots i_N}).$$

(iv) Finally, we claim that

$$-\lim_{N \to +\infty} \frac{1}{N} \log \mu(I_{i_1 \dots i_N}) = h_\mu(T).$$

This is an application of the Shannon-McMillan-Brieman theorem to interval maps, whose proof we present in the next section.

■

12.4 The Shannon-McMillan-Brieman theorem

We now give an application of entropy to describe the asymptotic size of elements in partitions.

Let $\alpha = \{A_1, A_2, \dots\}$ be a measurable partition of the space $(X, \mathcal{B})$, i.e. $X = \cup_{i=1}^n A_i$ and $A_i \cap A_j = \emptyset$ for $i \neq j$ (up to a set of zero μ-measure).

For each $n \geq 1$ we consider the new partition $\alpha_n = \vee_{i=0}^{n-1} T^{-i}\alpha$. For almost all $x \in X$ we can choose a unique element $A_n(x) \in \alpha_n$ with $x \in A_n(x)$.

THEOREM 12.11 (SHANNON-MCMILLAN-BREIMAN THEOREM). *Let $T : X \to X$ be a measure preserving transformation of a probability space $(X, \mathcal{B}, \mu)$. Let α be a partition. For almost all $x \in X$ we have that*

$$-\frac{\log \mu\left(A_n(x)\right)}{n} \to E(f|\mathcal{I})(x)$$

as $n \to +\infty$, where $f(x) = I(\alpha| \vee_{n=1}^{\infty} T^{-n}\alpha)(x)$ and $\mathcal{I}$ is the sigma-algebra generated by the T-invariant sets $T^{-1}B = B$.

COROLLARY 12.11.1. *If T is ergodic then for almost all $x \in X$*

$$-\frac{\log \mu\left(A_n(x)\right)}{n} \to h(T, \alpha) \text{ as } n \to +\infty.$$

If α is a generating partition then

$$-\frac{\log \mu\left(A_n(x)\right)}{n} \to h_\mu(T) \text{ as } n \to +\infty.$$

PROOF. Assuming the theorem, the ergodicity of the measure and the T-invariance of the limit imply that it is a constant. Integrating therefore gives that the limit is

$$\int E(f|\mathcal{I})d\mu = \int f d\mu = H(\alpha| \vee_{n=1}^{\infty} T^{-n}\alpha) = h(T, \alpha).$$

PROOF OF THEOREM 12.11. We first observe that

$$I(\vee_{i=0}^{n-1} T^{-i}\alpha)(x) = -\log \mu\left(A_n(x)\right)$$

Using the basic identities for the information function we see that

$$\begin{aligned} & I(\vee_{i=0}^{n-1} T^{-i}\alpha) \\ &= I(\alpha| \vee_{i=1}^{n-1} T^{-i}\alpha) + I(\vee_{i=1}^{n-1} T^{-i}\alpha) \\ &= I(\alpha| \vee_{i=1}^{n-1} T^{-i}\alpha) + I(\alpha| \vee_{i=1}^{n-2} T^{-i}\alpha)T \\ &\quad + \ldots + I(\alpha|T^{-1}\alpha)T^{n-2} + I(\alpha)T^{n-1}. \end{aligned} \tag{12.4}$$

We see from (12.4) that (almost everywhere)

$$\begin{aligned} & \limsup_{n\to+\infty} \frac{1}{n}|I(\vee_{i=0}^{n-1} T^{-i}\alpha) - E(f|\mathcal{I})| \\ & \leq \limsup_{n\to+\infty} \frac{1}{n}|I(\vee_{i=0}^{n-1} T^{-i}\alpha) - \sum_{i=0}^{n-1} fT^i| \\ & + \limsup_{n\to+\infty} |\frac{1}{n}\sum_{i=0}^{n-1} fT^i - E(f|\mathcal{I})| \end{aligned} \tag{12.5}$$

(using the triangle inequality). By the Birkhoff ergodic theorem (Theorem 10.6) we know that

$$\lim_{n\to+\infty} \frac{1}{n}|\sum_{i=0}^{n-1} fT^i - E(f|\mathcal{I})| = 0$$

(almost everywhere) and thus the second term on the right hand side of (12.5) vanishes.

We can next write from (12.5) that

$$\begin{aligned} & \frac{1}{n}\left| I(\vee_{i=0}^{n-1-i} T^{-i}\alpha) - \sum_{i=0}^{n-1} fT^i \right| \\ & \leq \frac{1}{n}\sum_{i=0}^{n-1} |I(\alpha| \vee_{j=1}^{n-1-i} T^{-j}\alpha)T^i - I(\alpha| \vee_{j=1}^{\infty} T^{-j}\alpha)T^i| \end{aligned}$$

(using also the definition of f). For $N \geq 1$ we define

$$F_N(x) := \sup_{N \leq i \leq n} |I(\alpha| \vee_{j=1}^{n-i} T^{-j}\alpha)(x) - I(\alpha| \vee_{j=1}^{\infty} T^{-j}\alpha)(x)|$$

and then upon fixing $N \geq 1$ we see that

$$\begin{aligned}
&\frac{1}{n}\left| I(\vee_{i=0}^{n-1} T^{-i}\alpha) - \sum_{i=0}^{n-1} fT^i \right| \\
&\leq \left(\frac{F_N T^n + F_N T^{n-1} + \ldots + F_N T^{n-N}}{n} \right) \\
&\quad + \left(\frac{\sum_{i=0}^{N-1} |I(\alpha| \vee_{j=1}^{n-i} T^{-j}\alpha)T^i - I(\alpha| \vee_{j=1}^{\infty} T^{-j}\alpha)T^i|}{n} \right)
\end{aligned} \tag{12.6}$$

We can bound the second term on the right hand side of (12.6) by

$$\begin{aligned}
&\frac{1}{n}\left| \sum_{i=0}^{N-1} |I(\alpha| \vee_{j=1}^{n-i} T^{-j}\alpha)T^i - I(\alpha| \vee_{j=1}^{\infty} T^{-j}\alpha)T^i| \right| \\
&\leq \frac{N}{n}\left(\sup_{k \geq N} \left(I(\alpha| \vee_{j=1}^{\infty} T^{-j}\alpha) + I(\alpha| \vee_{j=1}^{k} T^{-j}\alpha) \right) \right)
\end{aligned}$$

which tends to 0 (almost everywhere) as $n \to +\infty$.

We now turn to the first term on the right hand side of (12.6). We observe that by the Birkhoff ergodic theorem

$$\limsup_{n \to +\infty} \left(\frac{F_N T^n + F_N T^{n-1} + \ldots + F_N T^{n-N}}{n} \right) = E(F_N|\mathcal{I}).$$

Notice that $F_N \geq F_{N+1}$ and so

$$E(F_N|\mathcal{I}) \geq E(F_{N+1}|\mathcal{I}) \geq 0$$

(since $E(.|\mathcal{I})$ is a positive operator). Since $E(F_N|\mathcal{I}) \to 0$ (and is dominated by an integrable function) then

$$\lim_{N \to +\infty} \int E(F_N|\mathcal{I}) d\mu = \lim_{N \to +\infty} \int F_N d\mu = 0.$$

This completes the proof.

■

12.5 Comments and references

A good reference for more information on exactness is Rohlin's original paper [2].

Without the Markov assumption (but still assuming the uniform expansion property) the existence of an absolutely continuous invariant measure follows from the work of Lasota and Yorke [1].

There is an alternative proof of the Shannon-McMillan-Brieman theorem given in [3, 5.2]

References

1. A. Lasota and J. Yorke, *On the existence of invariant measures for piecewise monotonic transformations*, Trans. Amer. Math. Soc. **86** (1973), 481-488.
2. V. Rohlin, *Exact endomorphisms of lebesgue space*, Amer. Math. Soc. Transl. (2) **39** (1964), 1-36.
3. D. Rudolph, *Ergodic Theory on Lebesgue spaces*, O.U.P., Oxford, 1994.