Journal of Statistical Physics, Vol. 98, Nos. 1/2, 2000

Dimensions and Waiting Times for Gibbs Measures

Jean-René Chazottes[1]

Received June 16, 1999

For shifts of finite type, we relate the waiting time between two different orbits, one chosen according to an ergodic measure, the other according to a Gibbs measure, to Billingsley dimensions of generic sets. This is achieved by computing Billingsley dimensions of saturated sets in terms of a relative entropy which satisfies a pointwise ergodic result. As a by-product, a similar result is obtained for match lengths that are dual quantities of waiting times.

KEY WORDS: Gibbs measure; thermodynamic formalism; recurrence; waiting time; relative entropy; Hausdorff dimension; Billingsley dimension.

1. INTRODUCTION AND SET-UP

The concept of waiting-time we deal with in the present work as well as a related recurrence-time concept was introduced by Wyner and Ziv.[9] Ornstein and Weiss proved in ref. 5 a characterization of the metric entropy $h(\mu)$ of an ergodic measure μ in terms of recurrence times: if R_n is the first time at which the first n symbols of a sequence appear again, then they proved that $(1/n) \log R_n$ converges to $h(\mu)$, μ-almost everywhere. One can expect a similar result for the waiting time W_n (see Section 4 for a precise definition), that is the time needed before the n first symbols of a sequence appear in another sequence picked independently with the same ergodic measure. Shields[8] provided a counter-example to show that in the general ergodic case we do not have $(1/n) \log W_n \to h(\mu)$. He proved this result when μ is weak and very weak Bernoulli.[4]

A natural question arises: what does happen when one picks the second sequence with a different measure from the first one. One expect this waiting time to be generally asymptotically much longer. We show that

[1] CPT-CNRS, Luminy Case 907, F-13288 Marseille Cedex 9, France; e-mail: jeanrene@cptsg2.univ-mrs.fr.

0022-4715/00/0100-0305$18.00/0

one gets the entropy divided by a suitable dimension of a set of generic points.

This dimension appeared naturally as an extension of Hausdorff dimension in the context of sets of real numbers characterized by digit properties of their g-adic representations (see ref. 1 and ref. therein). These sets are typical examples of saturated sets. In the context of dynamical systems, typical saturated sets are sets of generic points of invariant measures.

Our work is organised as follows. The remaining part of the present section sets basic notations, definitions and a discussion of a strong mixing property. Section 2 develops the notion of relative entropy, which satisfies a pointwise ergodic Theorem, and relates it to Billingsley dimensions of saturated sets through an explicit formula generalizing previous results. The following section is devoted to the main results, that is the link between asymptotic waiting times and these dimensions. As a by-product, we deduce similar results for match lengths which is in a certain sense the dual quantity of waiting times. The last section contains the proofs of two intermediate results needed to establish the main results.

Shifts of Finite Type. Let $\mathscr{A} = \{1, 2,..., k\}$ be a finite set and $A = (A_{ij})_{i,\, j=1,..., k}$ be an irreducible and aperiodic $k \times k$ matrix with entries 0 or 1. Define the space Ω by

$$\Omega := \{\omega = \{\omega_i\}_{i=0}^{\infty} \in \mathscr{A}^{\mathbb{N}} : A_{\omega_i \omega_{i+1}} = 1 \ \forall i \in \mathbb{N}\}$$

For each $n \geqslant 1$, define $\pi_n : \Omega \to \Omega_n := \{1, 2,..., k\}^n$. The n-cylinder determined by ω is written as $[\pi_n.\omega]$, sometimes $[\omega_1 \cdots \omega_n]$. The collection of all cylinders generates the Borel σ-algebra $\mathscr{C}$.

For a fixed $0 < \alpha < 1$, define the usual metric d_α on Ω by

$$d_\alpha(\omega, \xi) = \alpha^{\nu(\omega, \xi)}$$

$$\text{where} \quad \nu(\omega, \xi) = \begin{cases} \max\{n \in \mathbb{N} : \pi_n.\omega = \pi_n.\xi\} & \text{if} \quad \omega \neq \xi \\ +\infty & \text{if} \quad \omega = \xi \end{cases}$$

We denote the shift on Ω by T:

$$(T\omega)_i = \omega_{i+1}, \quad \forall \omega$$

$\mathscr{I}_T$ will denote the σ-algebra of T-invariant Borel sets of Ω and $\mathscr{M}_T$ the set of all T-invariant Borel probability measures on Ω.

Saturated Sets. For any $\omega \in \Omega$, let $\Delta(\omega)$ be the set of accumulation points (w.r.t. the weak* topology) of the sequence of empirical measures

$$L_n(\omega) := \frac{1}{n} \sum_{j=0}^{n-1} \delta_{T^j \omega}$$

where $\delta_\omega(B) = 1$ if $\omega \in B$, 0 if $\omega \notin B$. $\Delta(\omega)$ is a non-empty compact connected subset of $\mathscr{M}_T$. One defines an equivalence relation on Ω, by setting:

$$\forall (\omega, \omega') \in \Omega \times \Omega, \qquad \omega \sim \omega' \Leftrightarrow \Delta(\omega) = \Delta(\omega') \tag{1}$$

The subset M of Ω is said to be saturated if it is saturated in the class of the equivalence relation $\sim$. For every $H \subset \mathscr{M}_T$, let us denote by $\nabla(H) := \{\omega : \Delta(\omega) = H\}$. The so-called smallest saturated sets are of the form $\nabla\Delta(\omega)$ for some $\omega \in \Omega$ i.e., the equivalence class of ω. The set $G(\nu)$ of generic points of any T-invariant measure ν provides the simplest example of smallest saturated set. Recall that

$$G(\nu) = \{\omega \in \Omega : \Delta(\omega) = \{\nu\}\}$$

with $\nu(G(\nu)) = 1$ if, and only if, ν is ergodic. If not, $\nu(G(\nu)) = 0$.

Thermodynamic Formalism (See ref. 6 for details). Let $\mathscr{F}_\alpha := \mathscr{F}_\alpha(\Omega)$ the space of Hölder continuous functions on Ω (w.r.t. d_α) and define the norm on $\mathscr{F}_\alpha$ by setting

$$\|g\| = \|g\|_\infty + \|g\|_\alpha$$

where $\|g\|_\infty$ is the supremum norm on the set of continuous functions $\mathscr{C}(\Omega)$ and $\|g\|_\alpha$ is the constant:

$$\|g\|_\alpha = \sup \left\{ \frac{\operatorname{var}_n(g)}{\alpha^n} : n \geqslant 0 \right\}$$

with $\operatorname{var}_n(g) = \sup \{ |g(\omega) - g(\xi)| : \pi_n . \omega = \pi_n . \xi \}$.

For $\varphi \in \mathscr{F}_\alpha$, define the Ruelle–Perron–Frobenius operator $\mathscr{L} = \mathscr{L}_\varphi$: $\mathscr{C}(\Omega) \to \mathscr{C}(\Omega)$ by

$$\mathscr{L} f(\omega) = \sum_{\omega' : T\omega' = \omega} e^{\varphi(\omega')} f(\omega')$$

We assume that

$$\mathscr{L}1 = 1 \tag{2}$$

that is φ is a normalised potential. If not, one can always obtain (2) by replacing φ with $\varphi' = \varphi + \log h - \log(h \circ T) - P(\varphi)$, where $P(\varphi)$ is the topological pressure of φ and h the eigenfunction of $\mathscr{L}$ corresponding to the maximal eigenvalue $e^{P(\varphi)}$. Hence hypothesis (2) can be made without loss of generality.

Let μ_φ be the Gibbs measure for the potential φ. In this setting, this measure is also the equilibrium state that is to say it satisfies the following variational principle,

$$P(\varphi) = h(\mu_\varphi) + \int \varphi \, d\mu_\varphi = 0 \tag{3}$$

where $h(\mu_\varphi)$ is the (metric) entropy of μ_φ. $P(\varphi) = 0$ follows from assumption (2) and $h(\mu_\varphi) > 0$.

A Strong Mixing Property. Let $(X, \mathscr{B}, T, \mu)$ a dynamical system, $\mathscr{Q}$ a finite or countable partition of X. As usually, we write $\mathscr{Q}_k := \bigvee_{j=0}^{k-1} T^{-j}\mathscr{Q}$ for the dynamical refinement of order k of the partition $\mathscr{Q}$ and denote by $\sigma(\mathscr{Q}_k)$ the σ-algebra generated by $\mathscr{Q}_k$.

The partition $\mathscr{Q}$ is said to be Ψ-mixing with speed Ψ if the sequence

$$\Psi(n) := \sup_{k,\, l} \sup_{\substack{R \in \sigma(\mathscr{Q}_k) \\ S \in T^{-(n+k)}\mathscr{Q}_l}} \left| \frac{\mu(R \cap S)}{\mu(R)\, \mu(S)} - 1 \right|, \qquad \forall n \geqslant 1$$

is such that $\lim_{n \to \infty} \Psi(n) = 0$. It is easy to check that the Ψ-mixing implies weak-Bernoulli.

Lemma 1.1. If μ_φ is a Gibbs measure with φ a Hölder continuous function, then the partition consisting in one-cylinders $[j]$, $j = 1,..., k$ is exponentially Ψ-mixing.

We give the proof for the sake of completness. In fact, this strong mixing property remains valid when the potential is of summable variations, that is if $\sum_{n=1}^{\infty} \mathrm{var}_n(\varphi) < \infty$. Moreover, the normalisation procedure can again be applied.

Proof. Take $A \in \mathscr{C}_k$, $B \in \mathscr{C}_l$. Then

$$\left| \frac{\mu_\varphi(A \cap T^{-(n+k)}B)}{\mu_\varphi(A)\,\mu_\varphi(B)} - 1 \right|$$

$$\leqslant \frac{1}{\mu_\varphi(A)\,\mu_\varphi(B)} \left| \int_B \mathscr{L}^{n+k}(\chi_A - \mu_\varphi(A))\, d\mu_\varphi \right|$$

$$\leqslant \frac{1}{\mu_\varphi(A)} \|\mathscr{L}^n(\mathscr{L}^k \chi_A - \mu_\varphi(A))\|_\infty \leqslant \theta^n \left(\frac{\|\mathscr{L}^k \chi_A\|}{\mu_\varphi(A)} + 1 \right)$$

where $0 < \theta < 1$ is the spectral radius of $\mathscr{L}$. But $(\|\mathscr{L}^k \chi_A\| / \mu_\varphi(A))$ is bounded for any k (see ref. 6). ▮

2. RELATIVE ENTROPY AND BILLINGSLEY DIMENSIONS OF SATURATED SETS

To ease notations we set μ for μ_φ in this section.

2.1. Relative Entropy and Information

Let γ be the σ-algebra generated by one-cylinders $[j]$, $j = 1,..., k$. Then for a probability measure ρ on Ω define the information of T w.r.t. γ given $\mathscr{C}$:

$$I_\rho(\omega) := I_\rho(\gamma \mid T^{-1}\mathscr{C})(\omega) = -\sum_{i=1}^{k} \chi_{[i]}(\omega) \log \rho[i \mid T^{-1}\omega]$$

where $\rho[i \mid T^{-1}\omega]$ is obtained as the limit *almost everywhere* of the probability distribution on the alphabet $\mathscr{A}$:

$$\rho_n[i \mid T^{-1}\omega] := \frac{\rho[i\omega_1 \cdots \omega_n]}{\rho[\omega_1 \cdots \omega_n]} = \rho([i] \mid T^{-1}\gamma_n)(\omega)$$

(the limit exists by the Increasing Martingale Theorem). The information is defined in such a way that $h(\rho) = \int I_\rho \, d\rho$.

Suppose now that $\mu := \mu_\varphi$ a normalized Gibbs measure. Define, for any $n \geqslant 1$,

$$I_{\mu_n}(\omega) := -\log \frac{\mu[\pi_n.\omega]}{\mu[\pi_{n-1}.T\omega]}$$

that is the information w.r.t. the σ-algebra formed from n-cylinders. The function $\varphi_n\colon \Omega \to \mathbb{R}$ such that $\varphi_n(\omega) := \mu[\pi_n.\omega]/\mu[\pi_{n-1}.T\omega]$ clearly defines an Hölder continuous potential which depends only on the n first symbols of ω. The unique Gibbs measure associated to φ_n is called the n-step Markov approximation of μ_φ because μ_n converges weakly to μ_φ. We have the following Lemma:

Lemma 2.1. If $\mu := \mu_\varphi$ is a normalized Gibbs measure, then I_μ^n uniformly converges to $-\varphi$ in the uniform norm as n goes to infinity.

The proof is easy (see ref. 6 for instance). Consequently, we can define the canonical information of μ:

Definition 2.2. The continuous realisation of I_μ is called the canonical information of μ: $I_\mu^c := -\varphi$.

Before stating the next proposition, we introduce the notion of relative entropy.

Let η and μ be two T-invariant measures on Ω. If μ is supported by Ω then,

$$H_n(\eta \mid \mu) := \int \log \frac{\eta[\pi_n.\omega]}{\mu[\pi_n.\omega]}\, d\eta(\omega) = \sum_{a \in \Omega_n} \eta[a] \log \frac{\eta[a]}{\mu[a]} \tag{4}$$

Definition 2.3. The relative entropy of η with respect to μ is the positive quantity

$$h(\eta \mid \mu) := \limsup_{n \to +\infty} \frac{1}{n} H_n(\eta \mid \mu) \tag{5}$$

Proposition 2.4. Let $\mu = \mu_\varphi$ be a Gibbs measure and η a T-invariant measure. Then

(1) $\lim_{n\to\infty} -(1/n)\log(\eta[\pi_n.\omega]/\mu[\pi_n.\omega]) = \mathbb{E}(I_\mu^c - I_\eta \mid \mathscr{I}_T)(\omega)$ η-a.e. and is an element of $L^1(\eta)$.

(2) $h(\eta \mid \mu) = \int (I_\mu^c - I_\eta)\, d\eta$.

Proof. (1) Let μ_n the n-step Markov approximation of μ. If follows from the definition of $I_{\mu_n}^c$ that $\log \mu[\pi_n.\omega] = \sum_{k=0}^{n-1} I_{\mu_{n-k+1}}^c(T^k\omega)$. Therefore,

$$\left| \frac{1}{n} \sum_{k=0}^{n-1} I_\mu^c(T^k\omega) + \frac{1}{n} \log \mu[\pi_n.\omega] \right| \leqslant \frac{1}{n} \sum_{k=0}^{n-1} \|I_\mu^c - I_{\mu_k}^c\|_\infty$$

Lemma (2.1) and Cesaro Lemma both imply:

$$\lim_{n \to +\infty} \left(\frac{1}{n} \sum_{k=0}^{n-1} I_\mu^c(T^k\omega) + \frac{1}{n} \log \mu[\pi_n.\omega] \right) = 0 \tag{6}$$

For any T-invariant measure η, Birkhoff ergodic Theorem ensures that:

$$\lim_{n \to +\infty} \frac{1}{n} \sum_{k=0}^{n-1} I_\mu^c(T^k\omega) = \mathbb{E}(I_\mu^c \mid \mathscr{I}_T)(\omega) \qquad \eta\text{-a.e. and is a element of } L^1(\eta) \tag{7}$$

By Shannon–McMillan Theorem:

$$\lim_{n \to +\infty} \frac{1}{n} \log \eta[\pi_n.\omega] = \mathbb{E}(I_\eta \mid \mathscr{I}_T)(\omega) \qquad \eta\text{-a.e. and is a element of } L^1(\eta) \tag{8}$$

Combining (6), (7), (8), one gets assertion (1).

(2) Using (5), one can deduce from the pointwise convergence given in Statement (1) that

$$h(\eta \mid \mu) = \limsup_{n \to +\infty} \frac{1}{n} \int \log \frac{\eta[\pi_n.\omega]}{\mu[\pi_n.\omega]} \, d\eta(\omega) = \int \mathbb{E}(I_\mu^c - I_\eta \mid \mathscr{I}_T) \, d\eta = \int (I_\mu^c - I_\eta) \, d\eta$$

∎

Corollary 2.5. If η is ergodic and μ as in Proposition 2.4, then

$$\lim_{n \to \infty} -\frac{1}{n} \log \frac{\eta[\pi_n.\omega]}{\mu[\pi_n.\omega]} = h(\eta \mid \mu) = -\int \varphi \, d\eta - h(\eta) \ \eta\text{-a.e.}$$

The proof of Corollary 2.5 is immediate by η-integration. (Remark that if μ is not normalised, then $h(\eta \mid \mu) = P(\varphi) - \int \varphi \, d\eta - h(\eta)$.) It is clear that $h(\mu \mid \mu) = 0$, but in general, $h(\eta \mid \mu) = 0$ does not implies $\eta = \mu$. However, if μ is a Gibbs measure, the variational principle (3) immediately implies that $h(\eta \mid \mu) = 0$ if, and only if, $\eta = \mu$.

This corollary suggests that there is an exponential splitting of Poincaré recurrences for orbits that are generic for different ergodic measures. The notion of waiting time can be seen as a rigourous way of making precise this rough idea.

2.2. Billingsley Dimensions of Saturated Sets

We define Billingsley dimensions by means of a Caratheodory construction (see ref. 7 for more details about Caratheodory constructions).

Given an arbitrary nonatomic probability measure μ on Ω, a Borel set $M \subset \Omega$ and a strictly positive number ε, $\mathscr{R}_\varepsilon(M)$ is by definition the set of all covers of M with cylinders of μ-measure less than ε. For $\beta \in [0, 1]$, define

$$\mathbf{H}_\varepsilon^\beta(M; \mu) := \inf_{R \in \mathscr{R}_\varepsilon(M)} \sum_{c \in R} \mu(c)^\beta$$

Since the map $\varepsilon \mapsto \mathbf{H}_\varepsilon^\beta(M; \mu)$ is monotone, we set:

$$\mathbf{H}^\beta(M; \mu) := \lim_{\varepsilon \to 0^+} \mathbf{H}_\varepsilon^\beta(M; \mu)$$

It is easy to check that by construction:

$$\beta < \beta' \quad \text{and} \quad \mathbf{H}^\beta(M; \mu) < \infty \Rightarrow \mathbf{H}^{\beta'}(M; \mu) = 0$$

$$\beta < \beta' \quad \text{and} \quad \mathbf{H}^{\beta'}(M; \mu) > 0 \Rightarrow \mathbf{H}^\beta(M; \mu) = \infty$$

This leads to the following definition:

Definition 2.6. Let μ be a non-atomic probability measure on Ω, M an arbitrary Borel subset of Ω and $\mathbf{H}^\beta(M; \mu)$ defined above. Then the Billingsley dimension of M w.r.t. μ is

$$\begin{aligned} \dim_\mu M &:= \inf\{\beta \in [0, 1] : \mathbf{H}^\beta(M; \mu) = 0\} \\ &= \sup\{\beta \in [0, 1] : \mathbf{H}^\beta(M; \mu) = \infty\} \end{aligned}$$

Remark that if (Ω, T) is the full shift and λ the Parry measure (the unique measure of maximal entropy), then $\dim_\lambda$ is nothing but the usual Hausdorff dimension. (Take $\alpha = k^{-1}$, where α gives the distance d_α on Ω of Section 1, k is the cardinal of the alphabet $\mathscr{A}$.)

For two given probability measures μ and η on Ω, define the singularity function of η with respect to μ by setting

$$\forall \omega \in \Omega, \qquad \frac{\eta}{\mu}(\omega) := \liminf_{n \to +\infty} \frac{\log \eta[\pi_n.\omega]}{\log \mu[\pi_n.\omega]} \quad (\geqslant 0)$$

with the classical conventions for the undertemined ratios of the form $\log(x)/\log(y)$. If μ is non-atomic, then it is proved in ref. 1 that for every subset M of Ω,

$$\dim_\mu M = \inf_\eta \sup \left\{ \frac{\eta}{\mu}(\omega) : \omega \in M \right\}$$

where the infimum is taken over all probability measures η on Ω. Define the quasi-metric q on $\mathscr{M}_T$ by setting:

$$q(\mu, \eta) := \sup_{\omega \in \Omega} \max \left\{ \left| \log \frac{\mu}{\eta}(\omega) \right|, \left| \log \frac{\eta}{\mu}(\omega) \right| \right\}$$

This quasi-metric is constructed in such a way that for any Borel subset M of Ω, if μ is a q-limit of the sequence $(\mu_n)_n$ of non-atomic probability measures (i.e., $\lim_n q(\mu, \mu_n) = 0$), then

$$\lim_{n \to +\infty} \dim_{\mu_n} M = \dim_\mu M \tag{9}$$

Notice that in general the topology induced by q is not the weak* topology. We need the following Lemma.

Lemma 2.7. Any Gibbs measure μ on Ω is the q-limit of the sequence $(\mu_n)_n$ of its Markov approximations.

Proof. Fix an arbitrary $\omega \in \Omega$. Then there exists a strictly increasing function $\theta: \mathbb{N} \to \mathbb{N}$ such that:

$$\frac{\mu}{\mu_n}(\omega) = \lim_{n \to \infty} \frac{\log \mu[\pi_{\theta(n)} . \omega]}{\log \mu_n[\pi_{\theta(n)} . \omega]}$$

where μ_n is the n-step Markov approximation of μ defined in Section 3.1.

Weak* compacity of the set of probability measures on Ω implies that the empirical measure $L_{\theta(k)}(\omega)$ weakly converges to a probability measure ρ_ω.

It is well-known (see ref. 6 for instance) that there exists a constant $C > 0$ such that for all $k \geqslant 1$,

$$C^{-1} \exp\left(- \sum_{j=0}^{\theta(k)-1} \varphi(T^j \omega) \right) \leqslant \mu[\pi_{\theta(k)} . \omega] \leqslant C \exp\left(- \sum_{j=0}^{\theta(k)-1} \varphi(T^j \omega) \right)$$

A similar inequality is valid for μ_n *mutatis mutandis*. As a consequence, we get:

$$\frac{\mu}{\mu_n}(\omega) = \frac{\rho_\omega(I_\mu^c)}{\rho_\omega(I_{\mu_n}^c)} \tag{10}$$

Lemma 2.1 ensures that there exists a compact interval $J \subset (0, +\infty)$ which contains both $\rho_\omega(I^c_\mu)$ and $\rho_\omega(I^c_{\mu_n})$. From (10) one gets

$$\left|\log \frac{\mu}{\mu_n}(\omega)\right| = \left|\log \frac{\rho_\omega(I^c_\mu)}{\rho_\omega(I^c_{\mu_n})}\right| \leqslant K\,|\rho_\omega(I^c_\mu - I^c_{\mu_n})| \leqslant K\,\|I^c_\mu - I^c_{\mu_n}\|_\infty$$

where K is a positive constant. We thus obtain

$$q(\mu, \mu_n) \leqslant K\,\|I^c_\mu - I^c_{\mu_n}\|_\infty$$

which concludes the proof by applying again Lemma 2.1. ∎

Stated in our notations, Cajar[1] proves that for a finite step Markov measure μ on Ω, one has

$$\forall \omega \in \Omega, \qquad \dim_\mu \nabla\Delta(\omega) = \inf_{\eta \in \Delta(\omega)} \left\{\frac{h(\eta)}{\eta(I^c_\mu)}\right\} \tag{11}$$

The next Lemma establishes the same formula for Gibbs measures.

Lemma 2.8. If μ is a Gibbs measure, then

$$\forall \omega \in \Omega, \qquad \dim_\mu \nabla\Delta(\omega) = \inf_{\eta \in \Delta(\omega} \left\{\frac{h(\eta)}{\eta(I^c_\mu)}\right\}$$

Proof. Let μ_n be the n-step Markov approximation of the Gibbs measure μ. A simple reformulation of (11) gives for all integers n:

$$\dim_{\mu_n} \nabla\Delta(\omega) = \inf_{\eta \in \Delta(\omega)} \{\gamma(\eta \mid \mu_n)\} \tag{12}$$

where we put $\gamma(\eta \mid \mu_n) := h(\eta)/\eta(I_{\mu^c_n})$ to ease notations. By Lemma 2.7, μ is q-limit of the Markov measures μ_n. It follows from (9) that

$$\dim_\mu \nabla\Delta(\omega) = \lim_{n \to +\infty} \dim_{\mu_n} \nabla\Delta(\omega) = \lim_{n \to +\infty} \inf_{\eta \in \Delta(\omega)} \{\gamma(\eta \mid \mu_n)\} \tag{13}$$

In addition, there exists a constant K such that for any $\eta \in \mathscr{M}_T$,

$$|\gamma_n(\eta \mid \mu) - \gamma(\eta \mid \mu_n)| = h(\eta) \left|\frac{1}{\eta(I^c_\mu)} - \frac{1}{\eta(I^c_{\mu_n})}\right| \leqslant h_{top} K\,\|I^c_\mu - I^c_{\mu_n}\|_\infty$$

which implies that $\gamma(\cdot\,|\,\mu_n)$ converges uniformly in $\mathscr{M}_T$ to $\gamma(\cdot\,|\,\mu)$ by Lemma 2.1. Therefore, the following commutation of symbols arises,

$$\lim_{n\to+\infty}\inf_{\eta\in\Delta(\omega)}\{\gamma(\eta\,|\,\mu_n)\}=\inf_{\eta\in\Delta(\omega)}\{\lim_{n\to+\infty}\gamma(\eta\,|\,\mu_n)\}=\inf_{\eta\in\Delta(\omega)}\{\gamma(\eta\,|\,\mu)\}$$

and by comparison with (13), the proof is complete. ∎

2.3. Relationship Between Relative Entropy and Dimensions

Proposition 2.9. Let ν be a T-invariant measure and $\mu=\mu_\varphi$ a Gibbs measure. Then

$$\dim_\mu G(\nu)=\frac{h(\nu)}{h(\nu)+h(\nu\,|\,\mu)}$$

Proof. It is an immediate corollary of Lemma 2.8 by taking the saturated set $G(\nu)$. ∎

3. MAIN RESULTS

Let define precisely the waiting time.

Definition 3.1. For any $n\geqslant 1$, the waiting time $W_n\colon\Omega\times\Omega\to\mathbb{N}\backslash\{0\}$ is defined as follows:

$$W_n(\omega,\xi):=\inf\{k\geqslant 1:\pi_n.T^k\xi=\pi_n.\omega\}$$

The conditional measure w.r.t. the set U of positive measure will be denoted in the sequel by μ_U:

$$\mu_U(A):=\frac{\mu(A\cap U)}{\mu(U)}$$

We can now give the main result.

Theorem 3.2. Let ν be an ergodic measure, μ_φ Gibbs measure and $W_n(\omega,\xi)$ the waiting time given by Definition 3.1. Denote by $\nu\times\mu_\varphi$ the product measure on $\Omega\times\Omega$. Then,

$$\lim_{n\to\infty}\frac{1}{n}\log W_n(\omega,\xi)=\frac{h(\nu)}{\dim_{\mu_\varphi}G(\nu)},\qquad \nu\times\mu_\varphi\text{-a.e.}$$

The proof makes use of two ingredients. The first one is the proposition below, whose proof is postponed to the last section, the second one is the ergodic result for relative entropy, that is Proposition 2.4.

Proposition 3.3. If η is an ergodic measure and $c(n)$ a sequence of positive constants satisfying $\sum_{n\geqslant 1} ne^{-c(n)} < \infty$, then, for n sufficiently large,

(1) $\log(W_n(\omega, \xi)\, \eta([\pi_n.\omega])) \geqslant -c(n)$ for $(\eta \times \eta)$-almost all (ω, ξ),

(2) if η is ψ-mixing, $\log(W_n(\omega, \xi)\, \eta([\pi_n.\omega])) \leqslant c(n)$ for $(\eta \times \eta)$-almost all (ω, ξ).

From this proposition one can easily deduce the following lemma (whose proof is also postponed to the last section):

Lemma 3.4. Let η be an ergodic measure and μ_φ a Gibbs measure. Under the same hypothesis of Proposition 3.3 for the sequence of constants $c(n)$, one gets

$$-c(n) \leqslant \log(W_n(\omega, \xi)\, \mu_\varphi([\pi_n.\omega])) \leqslant c(n) \qquad \text{for } (\eta \times \mu_\varphi)\text{-almost all } (\omega, \xi)$$

Proof of Theorem 3.2. Choose $c(n) = \varepsilon n$, $\varepsilon > 0$ arbitrary, which of course satisfies $\sum ne^{-c(n)} < \infty$. By Proposition 3.3 and Lemma 3.4, we get, for n sufficiently large and $\nu \times \mu_\varphi$ almost everywhere:

$$-\varepsilon \leqslant \frac{1}{n} \log W_n(\omega, \xi) + \frac{1}{n} \log \nu[\pi_n.\omega] - \frac{1}{n} \log \frac{\nu[\pi_n.\omega]}{\mu_\varphi[\pi_n.\omega]} \leqslant \varepsilon$$

Now apply Shannon–McMillan Theorem, Corollary 2.5 and Proposition 2.9. The desired conclusion follows from arbitrariness of ε. ∎

Remark 3.5. If the measure ν of Theorem 3.2 is T-invariant but non ergodic, it is easy to check that we get:

$$\lim_{n\to\infty} \frac{1}{n} \log W_n(\omega, \xi) = \mathbb{E}(I_\mu^c \mid \mathscr{I}_T)(\omega), \qquad \nu \times \mu_\varphi\text{-a.e.}$$

Theorem 3.2 allows to interpretate the Billingsley dimension of generic sets as a correction to entropy to get the asymptotic waiting time between two orbits picked according to distinct ergodic measures.

Define the dual quantity of W_n as follows:

Definition 3.6. Let $\omega, \xi \in \Omega$. For any $m \geqslant 1$, define

$$M_m(\omega, \xi) := \inf\{n \geqslant 1 : \pi_n.\omega \neq \pi_n.\sigma^{j-1}\xi,\ j = 1, 2,..., m\} \tag{14}$$

So $M_m(\omega, \xi)$ is the length n of the shortest block $\pi_n.\omega$ of ω which does not appear in ξ starting anywhere in $\pi_n.\xi$. From this definition, one deduces easily that $W_n > m$ if, and only if, $M_m \leqslant n$. This property yields the following proposition:

Proposition 3.7. Let ν be an ergodic measure, μ_φ a Gibbs measure and $M_m(\omega, \xi)$ the matching time given by Definition 14. Denote by $\nu \times \mu_\varphi$ the product measure on $\Omega \times \Omega$. Then,

$$\lim_{n \to \infty} \frac{M_m(\omega, \xi)}{\log m} = \frac{\dim_{\mu_\varphi} G(\nu)}{h(\nu)}, \qquad \nu \times \mu_\varphi\text{-a.e.}$$

4. PROOF OF PROPOSITION 3.3 AND LEMMA 3.4

Proof Proposition 3.3. (1) Consider the dynamical system on $\Omega \times \Omega$ equiped with the invariant measure $\mathbf{P} := \eta \times \eta$. For any cylinder $[\pi_n.\omega]$ with non-zero η-measure, we have for an arbitrary $K > 1$:

$$\begin{aligned}\mathbf{P}_{[\pi_n.\omega]}\{(\omega, \xi) : W_n(\omega, \xi) < K\} &\leqslant \sum_{j=1}^{\lfloor K \rfloor} \mathbf{P}_{[\pi_n.\omega]}\{(\omega, \xi) : W_n(\omega, \xi) = j\} \\ &\leqslant \sum_{j=1}^{\lfloor K \rfloor} \eta\{\xi : \pi_n T^j.\xi = \pi_n.\omega\} \\ &\leqslant K\eta[\pi_n.\omega]\end{aligned}$$

The last inequality comes from the invariance of the measure η. Setting

$$K = \frac{e^{-c(n)}}{\eta([\pi_n.\omega])}$$

gives

$$\mathbf{P}_{[\pi_n.\omega]}\{(\omega, \xi) : \log(W_n(\omega, \xi)\, \eta([\pi_n.\omega])) < -c(n)\} \leqslant e^{-c(n)}$$

Remark that if $K=(e^{-c(n)}/\eta([\pi_n.\omega]))\leqslant 1$, then $\mathbf{P}_{(\omega,\xi):[\pi_n.\omega]}\{W_n(\omega,\xi)<K\}=0$, since $W_n\geqslant 1$ by definition, so the above bound also trivially holds. Since the previous inequality is independent of $[\pi_n.\omega]$, the Borel–Cantelli lemma gives (1). Notice that only the invariance of the measure is used in the proof, not the ergodicity.

(2) Let $\delta\in(0,1)$ be arbitrary and choose d such that $\Psi(d)<\delta$. This is always possible by Lemma 1.1. Fix an integer N_1 large enough so that $e^{c(n)}\geqslant 2(n+d)$ for all $n\geqslant N_1$. Fix $n\geqslant N_1$ and let $K\geqslant 2(n+d)$ arbitrary. Then, for each cylinder $[\pi_n.\omega]$ with noh-zero measure, we have:

$$
\begin{aligned}
&\mathbf{P}_{[\pi_n.\omega]}\{(\omega,\xi):W_n(\omega,\xi)>K\}\\
&\quad=\eta\{\xi:\pi_n.\xi=\pi_n.\omega,\ \pi_n.T\xi=\pi_n.\omega,...,\ \pi_n.T^{\lfloor K\rfloor}\xi=\pi_n.\omega\}\\
&\quad\leqslant(1-\eta([\pi_n.\omega]))\prod_{j=1}^{\lfloor K/(n+d)\rfloor-1}\eta_{\{\xi:\pi_n.T^{i(n+d)}\xi=\pi_n\omega,\ 0\leqslant i<j\}}\\
&\qquad\times\{\xi:\pi_n.T^{j(n+d)}\xi\neq\pi_n.\omega\}\\
&\quad=(1-\eta[\pi_n.\omega])\prod_{j=1}^{\lfloor K/(n+d)\rfloor-1}(1-\eta_{S_j}(R_j))
\end{aligned}
$$

where

$$R_j:=\{\xi:\pi_n.T^{i(n+d)}\xi\neq\pi_n.\omega,\ i=0,1,...,j-1\}\in\sigma(\mathcal{Q}_{j(n+d)-d})$$

and

$$S_j:=\{\xi:\pi_n.T^{j(n+d)}\xi=\pi_n.\omega\}\in T^{-(j(n+d)+1)}\mathcal{Q}_j$$

The particular choice of d and the invariance of η both imply

$$\eta_{S_j}(R_j)\leqslant(1-\delta)\,\eta(S_j)=(1-\delta)\,\eta([\pi_n.\omega])$$

for all j, and then:

$$
\begin{aligned}
\mathbf{P}_{[\pi_n.\omega]}\{(\omega,\xi):W_n(\omega,\xi)>K\}&\leqslant(1-(1-\delta)\,\eta[\pi_n.\omega])^{\lfloor K/(n+d)\rfloor}\\
&\leqslant\frac{1}{\delta}(1-(1-\delta)\,\eta[\pi_n.\omega])^{K/(n+d)}
\end{aligned}
$$

For any $n \geqslant N$, let $K = e^{c(n)}/\eta[\pi_n.\omega] \geqslant 2(n+d)$. Then

$$\mathbf{P}_{[\pi_n.\omega]}\{(\omega, \xi) : \log(W_n(\omega, \xi)\, \eta[\pi_n.\omega]) > c(n)\}$$
$$\leqslant \frac{1}{\delta}\,(1-(1-\delta)\,\eta[\pi_n.\omega])^{(1/\eta[\pi_n.\omega])(e^{c(n)}/(n+d))}$$
$$\leqslant \frac{1}{\delta}\,\rho^{((1-\delta)\,e^{c(n)})/(n+d)}$$

where $\rho := \sup\{(1-z)^{1/z} : 0 < z \leqslant 1-\delta\}$.

Since $\sum_{n \geqslant 1} ne^{-c(n)} < \infty$, $(1/n)\, e^{c(n)} \to \infty$ as n diverges, and we can choose N_2 large enough to ensure that

$$(\rho^{(1-\delta)})^{e^{c(n)}/(n+d))} \leqslant 2ne^{-c(n)}, \quad \forall n \geqslant N_2$$

Consequently,

$$\mathbf{P}_{[\pi_n.\omega]}\{(\omega, \xi) : \log(W_n(\omega, \xi)\, \eta([\pi_n.\omega])) > c(n)\}$$
$$\leqslant \frac{2}{\delta}\, ne^{-c(n)}, \quad \forall n \geqslant N_3 := \max(N_1, N_2)$$

Since this bound is independent of the cylinder $[\pi_n.\omega]$,

$$\sum_{n \geqslant 1} \mathbf{P}_{[\pi_n.\omega]}\{(\omega, \xi) : \log(W_n(\omega, \xi)\, \eta([\pi_n.\omega])) > c(n)\}$$
$$\leqslant M + \frac{2}{\delta} \sum_{n \geqslant N_3} ne^{-c(n)} < \infty$$

and the Borel–Cantelli lemma gives the desired result. So the Proposition is now proved. ▮

Proof of Lemma 3.4. The proof is analogous to the proof of Proposition 3.3 by replacing $\eta \times \eta$ by $\eta \times \mu_\varphi$, since the measure μ_φ is Ψ-mixing by Lemma 1.1. Notice that the μ_φ-measure of any allowed cylinder is strictly positive. ▮

ACKNOWLEDGMENTS

Xavier Bressaud and Ricardo Lima are gratefuly acknowledged for fruitful discussions. It was brought to my attention by one of the referees that there is some overlap of my results with those in ref. 3, though they are stated in a different context (that is Markov chains) and without the dimensional aspect.

REFERENCES

1. H. Cajar, *Billingsley Dimension in Probability Spaces*, Lecture Notes in Math., Vol. 892 (Springer-Verlag, Berlin, 1981).
2. M. Denker, C. Grillenberger, and K. Sigmund, *Ergodic Theory on Compact Spaces*, Lecture Notes in Math., Vol. 527 (Springer-Verlag, 1976).
3. I. Kontoyiannis, Asymptotic recurrence and waiting times for stationary processes, *J. Theor. Prob.* **11**:795–811 (1998).
4. K. Marton and P. C. Shields, Almost-sure waiting time results for weak and very weak Bernoulli processes, *Ergod. Th. Dynam. Sys.* **15**:951–960 (1995).
5. D. S. Ornstein and B. Weiss, Entropy and data compression schemes, *IEEE Transactions on Information Theory* **39**(1) (1993).
6. W. Parry and M. Pollicott, Zeta functions and the periodic orbit structure of hyperbolic dynamics, *Astérisque* **187–188**, SMF (1990).
7. Ya. B. Pesin, *Dimension Theory in Dynamical Systems, Contemporary Views and Applications* (Chicago Lectures in Mathematics, 1997).
8. P. C. Shields, Waiting times: positive and negative results on the Wyner–Ziv problem, *J. Theor. Prob.* **6**:499–519 (1993).
9. A. Wyner and J. Ziv, Some asymptotic properties of the entropy of a stationary ergodic data source with applications to data compression, *IEEE Trans. Inform. Th.* **IT-35**: 1250–1258 (1989).