

ASYMPTOTIC LAWS FOR SYMBOLIC DYNAMICAL SYSTEMS

ZAQUEU COELHO

ABSTRACT. These are notes of a mini-course given in the Summer School on Symbolic Dynamics, Temuco (Chile), January 1997. The aim is to present some results concerning asymptotic limit laws for the occurrence of asymptotically rare events associated with some symbolic dynamical systems. The results will be mainly presented for shifts of finite type, but there will be left questions of whether these results remain valid for more general or possibly intrinsically different subshifts.

CONTENTS

Introduction	1
1. Preliminaries	2
1.1. Bernoulli trials	2
1.2. Occurrence and Waiting Times	4
2. General Setup and Motivation	8
2.1. Return Maps and Expected Return Times	8
2.2. Asymptotically Rare Events	10
2.3. Known Results and Motivation	10
3. Shifts of Finite Type and Equilibrium States	16
3.1. Hölder Potentials	17
3.2. Entropy and Pressure	18
3.3. Ruelle-Perron-Frobenius Operator	18
3.4. The Central Limit Theorem	19
3.5. Pianigiani-Yorke Measure	21
4. Point Processes and Convergence in Law	24
4.1. Convergence of Point Processes	25
4.2. Entrance Times and Visiting Times	26
4.3. Final Remarks	29
4.4. Some Questions	30
References	31

INTRODUCTION

It is natural in Probability Theory to study the asymptotic behaviour of random events, when these events are subject to conditions which make them vary in probability. Special concern has been given to studying how long one has to wait in order to see the occurrence of these events. The typical example of this approach, is a simple random walk on the states $\{0, 1, \dots, N\}$ with a reflecting barrier at N and an absorbing barrier at 0 . In this case, one is interested to know how the law of the random time

Date: September 1997.

1991 *Mathematics Subject Classification.* Primary 58F11, 58F03; Secondary 28D05, 58F13, 60G55, 60G10.

for absorption behaves asymptotically when N diverges. Rescaling these times by $N \log(N)$ gives rise to a limit law of an exponential random variable (cf. [Ald]). Hence, in general, the usual question is to decide which scale is best suited to describe the asymptotic behaviour, and then to show whether there is convergence in law to some well known distribution.

These problems have inspired the general study of asymptotic limit laws for the occurrence times of certain events which have asymptotically zero probability. These events are commonly known as asymptotically rare events. In these notes, we give a summary of recent developments in this subject within the dynamical systems framework. Later we specialise to symbolic dynamical systems, where there is a sketch of the technique involved in obtaining certain limit laws of occurrence times of the rare event corresponding to visiting a subsystem of a symbolic dynamical system. The latter is work in progress of P. Collet and the author [CC2].

These notes are organised in the following way. In the Preliminaries we show how these problems arise in the study of Markov Chains. Then we introduce a general setup in which all of these problems can be phrased in Ergodic Theory, and we give some motivation coming from the consideration of different types of ergodic dynamical systems. In Section 3 we summarise some important properties of shifts of finite type and Hölder equilibrium states, which will be used in Section 4 for the study of asymptotic law of visiting times to a subsystem of finite type. In the Final Remarks we leave some questions of whether these results can be extended to different types of symbolic dynamical systems.

The present notes are summaries of the four lectures given by the author in the Summer School on Symbolic Dynamics, Temuco (Chile), January 1997. Due to other heavy commitments, the author was unable to prepare this final version prior to the course, for which he apologises to the organisers of the School for any inconvenience caused.

1. PRELIMINARIES

We show a series of examples in which the concept of asymptotically rare event appears naturally. The reader should be familiar with the classical result that the Binomial distribution (subject to a combined asymptotic limit) gives rise to the Poisson distribution. We make a brief exposition of this fact in what follows and later on, we relate this initial step with more general asymptotic limits.

1.1. Bernoulli trials.

Let $X_i \in \{0, 1\}$, for $i = 1, 2, \dots$, denote a sequence of independent random variables defined on a standard probability space (Ω, \mathbb{P}) . We suppose X_i corresponds to the independent repetition of a particular random experiment where the possible observed outcomes are failure in the i -th attempt $X_i = 0$, or success in the i -th attempt $X_i = 1$. By repetition of the experiment we mean that the distributions of the X_i 's are identical, therefore there is a probability $0 < p < 1$ of success and consequently a probability $1 - p$ of failure. For fixed $n > 1$, let Y_n be the number of successes in n successive repetitions of the experiment, i.e.

$$Y_n = X_1 + \dots + X_n .$$

Using a simple counting method of the number of possibilities we see that Y_n has the Binomial distribution, i.e.

$$\mathbb{P}(Y_n = k) = \binom{n}{k} p^k (1-p)^{n-k},$$

for each $k = 0, 1, \dots, n$. Now it is clear that if we let n diverge to infinity and maintain k fixed, the probability of having exactly k successes in n attempts of the experiment tends to zero. However, if we allow the probability of successes p to converge to zero with n , then it becomes “asymptotically rare” to have successes when n diverges and therefore we may have a chance of getting positive asymptotic probability of having exactly k successes in n attempts. This is in fact the classical result that if we take $p = p_n$ such that $p_n \sim \lambda/n$ for some $\lambda > 0$, then indeed

$$\begin{aligned} \mathbb{P}(Y_n = k) &= \binom{n}{k} p_n^k (1-p_n)^{n-k} \\ &\sim \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} \longrightarrow \frac{\lambda^k}{k!} e^{-\lambda}, \end{aligned}$$

which is the Poisson distribution of parameter λ .

We mentioned this classical result to illustrate the concept of asymptotically rare event. However, from the point of view of our study in these notes, this example has a more intriguing structure since in order to get convergence of the distribution of Y_n , the random variables Y_n needed to be defined in different probability spaces. To be more precise, we introduce the language of shift spaces which will be needed later and translate this result in the framework of symbolic dynamics.

Let $\mathcal{A} = \{0, 1, \dots, \ell - 1\}$ be an alphabet of ℓ numbers (corresponding maybe to different states of a random discrete dynamical system, or the finite possible outcomes of a random experiment). The forward sequence of outcomes is the space

$$\mathcal{A} = \{(x_0, x_1, x_2, \dots) : x_i \in \mathcal{A}, \forall i \in \mathbb{N}\}.$$

There is a natural map (the shift) $\sigma: \mathcal{A} \rightarrow \mathcal{A}$ acting on the sequences by

$$\sigma(x_0, x_1, x_2, \dots) = (x_1, x_2, \dots).$$

We put a topology on this space by defining the cylinder sets

$$C[i_0, i_1, \dots, i_m]_s = \{(x_0, x_1, x_2, \dots) \in \mathcal{A} : x_j = i_j, \text{ for } j = 0, 1, \dots, m\}$$

to be a base of open sets. This is the Tychonov product topology on \mathcal{A} induced by the discrete topology on \mathcal{A} . The space \mathcal{A} is a compact metrisable space in this topology and σ is a continuous surjective map. A finite Borel measure μ on \mathcal{A} is uniquely determined by its values on the cylinders sets $C[i_0, i_1, \dots, i_m]_s$.

Now we interpret the classical result mentioned earlier. In this case $\mathcal{A} = \{0, 1\} = \{\text{“failure”}, \text{“success”}\}$, and μ is the product measure on \mathcal{A} generated by the measure $(1-p, p)$ on $\{0, 1\}$, which will be denoted by $(1-p, p)$. The measure of the cylinder sets are

$$\mu(C[i_0, i_1, \dots, i_m]_s) = p^{\sum_{j=0}^m i_j} (1-p)^{m+1 - \sum_{j=0}^m i_j}.$$

Defining $f: \mathcal{A} \rightarrow \{0, 1\}$ by $f(x_0, x_1, x_2, \dots) = x_0$, we have

$$X_i(x) = f(\sigma^i x),$$

for $i = 0, 1, \dots$, is a sequence of i.i.d. random variables defined on the probability space (\mathcal{A}, μ) and the random variables

$$Y_n = X_0 + X_1 + \dots + X_{n-1} = \sum_{i=0}^{n-1} f(\sigma^i x) = x_0 + x_1 + \dots + x_{n-1}$$

have the Binomial distribution $B(n, p)$, i.e.

$$\mu\{x \in \mathcal{A} : Y_n(x) = k\} = \binom{n}{k} p^k (1-p)^{n-k}.$$

Hence the classical result corresponds to considering a sequence of probability measures μ_n on \mathcal{A} as $(1-p_n, p_n)$ and it states that, for all fixed $k = 0, 1, \dots$,

$$\lim_{n \rightarrow \infty} \mu_n\{x \in \mathcal{A} : Y_n(x) = k\} = \frac{\lambda^k}{k!} e^{-\lambda},$$

whenever $p_n \sim \lambda/n$ for some $\lambda > 0$.

Here both measure and random variable are varying with n . We will consider in the sequel asymptotic limit laws, where the probability space (\mathcal{A}, μ) remains fixed and the problem will be to find rescalings of a sequence of random variables of type Y_n in order to guarantee convergence in distribution.

1.2. Occurrence and Waiting Times.

We begin with a simple example. Let us record the times of occurrence of success in the independent repetition of an experiment. So given a random sequence of outcomes

$$x = (x_0, x_1, x_2, \dots) \in \mathcal{A} \quad (\mathcal{A} = \{0, 1\})$$

we are interested in n such that $x_n = 1$. Let

$$\tau(x) = \tau^{(1)}(x) = \inf\{n \geq 0 : x_n = 1\}$$

and define, for $j \geq 2$,

$$\tau^{(j)}(x) = \inf\{n > \tau^{(j-1)}(x) : x_n = 1\}.$$

If μ denotes the measure $(1-p, p)$ on \mathcal{A} then

$$\mathbb{P}(\tau = k) = \mu\{x \in \mathcal{A} : \tau(x) = k\} = p(1-p)^k,$$

for all $k \geq 0$. Also it is an exercise to show that, for fixed $k > 0$,

$$\begin{aligned} \mathbb{P}(\tau^{(j)} - \tau^{(j-1)} = k \mid \tau^{(j-1)} = k_{j-1}, \dots, \tau^{(1)} = k_1) \\ = \mathbb{P}(\tau^{(j)} - \tau^{(j-1)} = k) = p(1-p)^{k-1}, \end{aligned}$$

for all $k_1 < \dots < k_{j-1}$. Therefore the *inter-occurrence* times $1 + \tau^{(1)}, \tau^{(2)} - \tau^{(1)}, \tau^{(3)} - \tau^{(2)}, \dots$ are independent and identically distributed with geometric distribution of parameter p . Consequently $\tau^{(j)}$ has the Negative Binomial distribution $BN(j, p)$.

Suppose now we have two random sequences of outcomes corresponding to the repetition of an experiment with ℓ possible different results. Let $x = (x_0, x_1, x_2, \dots)$ and $y = (y_0, y_1, y_2, \dots)$ denote the sequence of outcomes (x independent of y). We say there is a *match* at time n if $x_n = y_n$. We may think of matching at time n as a success and no matching is failure. We then note that

$$\mathbb{P}(x_n = y_n) = \mathbb{P}(x_0 = y_0) = p,$$

for all $n > 0$. Suppose $0 < p < 1$ and write $q = 1 - p$. Now consider the problem of having a successive sequence of matchings. Define $\Phi: \mathcal{A} \times \mathcal{A} \rightarrow \mathbb{N}$ by

$$\Phi(x, y) = k \quad \text{if} \quad x_0 = y_0, x_1 = y_1, \dots, x_{k-1} = y_{k-1}, x_k \neq y_k,$$

and consider the random variables $Z_n = \Phi(\sigma^n x, \sigma^n y)$.

Exercise. Show that $Z = \{Z_n : n \geq 0\}$ is a Markov chain on the countable states $\{0, 1, \dots\}$ with transition probabilities

$$\begin{aligned} \mathbb{P}(Z_n = k - 1 \mid Z_{n-1} = k) &= 1 \quad \text{if } k \geq 1, \\ \mathbb{P}(Z_n = k \mid Z_n = 0) &= p^k q \quad \text{if } k \geq 0. \end{aligned}$$

The Markov chain Z has a graph representation given by Figure 1. This Markov chain is positive recurrent with stationary distribution given by $\pi_j = p^j q$.

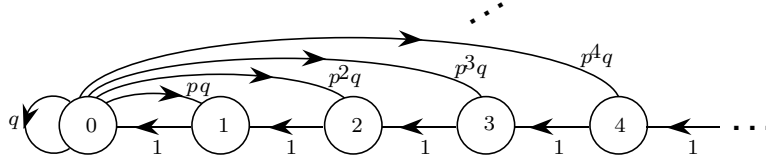


FIGURE 1. The Markov chain Z

Consider now the waiting time W_k to have k matchings in a row. Then $W_k = (k - 1) + \tau_k$, where

$$\tau_k = \inf\{n \geq 0 : Z_n \geq k\}.$$

The conditional distribution of τ_k given that we start with no matching is

$$\begin{aligned} \mathbb{P}(\tau_k = m \mid Z_0 = 0) &= \mathbb{P}(Z_m \geq k, Z_{m-1} < k, \dots, Z_1 < k \mid Z_0 = 0) \\ &= \sum_{j \geq k} \mathbb{P}(Z_m = j, Z_{m-1} < k, \dots, Z_1 < k \mid Z_0 = 0) \\ &= \sum_{j \geq k} \mathbb{P}(Z_m = j, Z_{m-1} = 0, Z_{m-2} < k, \dots, Z_1 < k \mid Z_0 = 0). \end{aligned}$$

For $m = 1$ we have

$$\mathbb{P}(\tau_k = 1 \mid Z_0 = 0) = \sum_{j \geq k} \mathbb{P}(Z_1 = j \mid Z_0 = 0) = \sum_{j \geq k} p^j q = p^k,$$

for $m = 2$,

$$\mathbb{P}(\tau_k = 2 \mid Z_0 = 0) = p^k q,$$

and for $m > 2$,

$$\begin{aligned} &\mathbb{P}(\tau_k = m \mid Z_0 = 0) \\ &= \sum_{j \geq k} (p^j q) \mathbb{P}(Z_{m-1} = 0, Z_{m-2} < k, \dots, Z_1 < k \mid Z_0 = 0) \\ &= p^k \mathbb{P}(Z_{m-1} = 0, Z_{m-2} < k, \dots, Z_1 < k \mid Z_0 = 0). \end{aligned}$$

Hence if we define the constants $\nu_0(k, 1) = 1$, $\nu_0(k, 2) = q$ and for $m > 2$,

$$\nu_0(k, m) = \mathbb{P}(Z_{m-1} = 0, Z_{m-2} < k, \dots, Z_1 < k \mid Z_0 = 0),$$

we obtain in short

$$\mathbb{P}(\tau_k = m \mid Z_0 = 0) = p^k \nu_0(k, m).$$

The expected time to have k successes in a row is then

$$\begin{aligned} \mathbb{E}(W_k \mid Z_0 = 0) &= (k - 1) + \mathbb{E}(\tau_k \mid Z_0 = 0) \\ &= (k - 1) + \sum_{m \geq 1} m p^k \nu_0(k, m). \end{aligned}$$

Here we note that $\mathbb{E}(\tau_k \mid Z_0 = 0)$ diverges with k , since the event corresponding to the random path Z_m visiting one of the states $\{k, k+1, \dots\}$ is “asymptotically rare” with k . Therefore in order to have a chance to get convergence in distribution of the sequence of random variables τ_k we rescale it by its expectation, so we consider

$$\begin{aligned} \mathbb{P}\left(\frac{\tau_k}{\mathbb{E}(\tau_k)} \leq t \mid Z_0 = 0\right) &= \sum_{m \leq t \mathbb{E}(\tau_k)} \mathbb{P}(\tau_k = m \mid Z_0 = 0) \\ &= \sum_{m \leq t \mathbb{E}(\tau_k)} p^k \nu_0(k, m). \end{aligned}$$

Now we see that for $m > 2$ we have

$$\nu_0(k, m) = \sum_{i=0}^{m-2} \sum_{\substack{0=j_0 < j_1 < \dots < j_i \leq m-2 \\ j_s - j_{s-1} < k}} p^{m-2-i} q^i.$$

The latter corresponds to fixing the number of visits to the state 0 and then summing over the possible different number of loops. The following result can be proved using renewal theory.

Theorem 1 (*Asymptotic Exponential Law for τ_k*). *For every $t > 0$ we have*

$$\lim_{k \rightarrow \infty} \mathbb{P}(\tau_k > t \mathbb{E}(\tau_k) \mid Z_0 = 0) = e^{-t}.$$

In fact,

$$\lim_{k \rightarrow \infty} \mathbb{P}(\tau_k > t \mathbb{E}(\tau_k)) = e^{-t}.$$

Moreover, the scaling $\mathbb{E}(\tau_k)$ satisfies $\mathbb{E}(\tau_k) \sim \lambda p^{-k}$ with $0 < \lambda < 1$ depending only on p .

We will give a sketch proof of a much stronger result in the more general setting of equilibrium states on shifts of finite type, and the above result will be a consequence of that.

To finish this section we mention the problem of subsequent k matchings (for an example see Figure 2).

$$\begin{array}{l} \text{Matching Chain:} \qquad \qquad \qquad (k = 3) \\ \qquad \qquad \qquad (01001011000010 \overbrace{111} \ 00101100 \overbrace{111} 1100 \dots) \\ \text{Z-Chain:} \qquad \qquad \qquad \tau_3^{(3)} = 27 \quad \dots \\ \qquad \qquad \qquad (01001021000010 \overbrace{3} 2100102100 \overbrace{54} 32100 \dots) \\ \qquad \qquad \qquad \tau_3^{(1)} = 15 \qquad \tau_3^{(2)} = 26 \end{array}$$

FIGURE 2. Example of a Matching Chain and translation into the Z-Chain

How about the asymptotic distribution of the inter-occurrence times $\tau_k^{(2)} - \tau_k^{(1)}$, $\tau_k^{(3)} - \tau_k^{(2)}$, \dots , in this context? Intuitively, they should not be asymptotically exponential (after rescaling), since when we “hit” k matchings for the first time we may have hit more than k matchings in a row.

The situation would be different if we were interested in the occurrence of say no match, k matchings and no match in succession (i.e. studying the times when the

string

$$(1) \quad 0 \underbrace{11 \cdots 1}_k 0$$

occurs). In the latter case, for fixed k , every hit does not mean hitting more than k successes in a row, and then the inter-occurrence times rescaled by expectation should converge to an exponential law of rate 1. (This can be shown to be true also using renewal theory.) If we denote by $T_k^{(j)}$ the times of occurrence of the string in (1) then the following result holds. (Define $T_k^{(0)} \equiv 0$.)

Theorem 2 (*Asymptotic Exponential Law for Inter-occurrence Times*). *For every $t > 0$ and $j = 1, 2, \dots$, we have*

$$\lim_{k \rightarrow \infty} \mathbb{P} \left(T_k^{(j)} - T_k^{(j-1)} > t \mathbb{E}(T_k^{(j)} - T_k^{(j-1)}) \mid Z_0 = 0 \right) = e^{-t}.$$

In fact,

$$\lim_{k \rightarrow \infty} \mathbb{P} \left(T_k^{(j)} - T_k^{(j-1)} > t \mathbb{E}(T_k^{(j)} - T_k^{(j-1)}) \right) = e^{-t}.$$

Moreover, the scaling $\mathbb{E}(T_k^{(j)} - T_k^{(j-1)})$ satisfies $\mathbb{E}(T_k^{(j)} - T_k^{(j-1)}) \sim q^{-2} p^{-k}$ for all j , and the random variables $D_k^{(j)} = T_k^{(j)} - T_k^{(j-1)}$ are asymptotically independent when k diverges.

The above result is a particular case of a more general result which we will discuss in later sections of these notes.

In order to relate this problem with the classical result mentioned in Section 1.1 we will compute the asymptotic law associated with the number of hits in the event no match, n matchings in a row and no match. Let, for fixed $n > 0$, $\Psi_n: \mathcal{A} \times \mathcal{A} \rightarrow \{0, 1\}$ be defined by

$$\Psi_n(x, y) = \omega = (\omega_0, \omega_1, \dots),$$

where $\omega_s = 1$ iff $x_s \neq y_s$, $x_{s+1} = y_{s+1}$, \dots , $x_{s+n} = y_{s+n}$, and $x_{s+n+1} \neq y_{s+n+1}$. Then the functions $X_i: \mathcal{A} \times \mathcal{A} \rightarrow \{0, 1\}$ defined by $X_i(x, y) = \omega_i$ form a sequence of identically distributed random variables on the probability space $(\mathcal{A} \times \mathcal{A}, \mu \times \mu)$ with distribution $(1 - p_n, p_n)$, where p_n is the probability of the event no match, n matchings in a row and no match, i.e. $p_n = p^n q^2$. (We note that X_i is *not* an independent sequence!) Let

$$Y_s = X_1 + \dots + X_s.$$

Defining $N_n = [t q^{-2} p^{-n}]$, where $[t]$ denotes the integer part of t , we obtain as a direct consequence of Theorem 2.

Corollary 3. *Given $t > 0$, for every fixed $k \geq 0$, we have*

$$\lim_{n \rightarrow \infty} (\mu \times \mu) \{ (x, y) \in \mathcal{A} \times \mathcal{A} : Y_{N_n}(x, y) = k \} = \frac{t^k}{k!} e^{-t}.$$

Therefore, an analogue of the classical result also exists in this context, since the above gives an asymptotic Poisson limit law for the number of hits to the event no match, n matchings in a row and no match, in time $N_n \sim t q^{-2} p^{-n}$. (This is *not true* if we considered the event n matchings in a row!)

The study of sequence matching is far reaching and has been an active research field in recent years in the case of weakly dependent random sequences. In the independent case, there are also many interesting questions one can ask, for instance, what is the asymptotic distribution of the maximum number of matchings between two samples

of the same length when the length diverges. An example of this approach is given in Arratia et al. [AGW], which contains a refinement of earlier results by Erdős & Rényi [ER], and uses the famous Chen-Stein method to prove Poisson approximation results.

In the following sections we will rephrase the initial problem in a general setup. Then we will mention some results from different dynamical systems which will serve as motivation for the consideration of these problems. Next we will specialise the problem to shifts of finite type endowed with a Hölder equilibrium state, where we will spend some time explaining the ingredients behind the technique which allows us to prove these asymptotic distribution results for “well”-behaved symbolic dynamical systems. In the Final Remarks we discuss how the general results of Section 4 prove the comments in this Section.

2. GENERAL SETUP AND MOTIVATION

Let $(\Omega, \mathfrak{B}, \mu)$ be a standard probability space and let $f: \Omega \rightarrow \Omega$ be a *measure-preserving* map (i.e. $\mu(f^{-1}A) = \mu(A)$, $\forall A \in \mathfrak{B}$). This means that given a point $\omega \in \Omega$ randomly chosen according to μ , the probability to visit a set $A \in \mathfrak{B}$ in n -steps is $\mu(A)$. Recall the Poincaré Recurrence

$$\mu\{x \in A: f^n x \in A, \text{ for some } n > 0\} = \mu(A),$$

i.e. μ -almost every point in A returns to A . In these notes we will assume that μ is *ergodic* for f (i.e. $\nexists A \in \mathfrak{B}$, $0 < \mu(A) < 1$ such that $\mu(A \Delta f^{-1}A) = 0$). This is equivalent to every fixed $A \in \mathfrak{B}$ with $\mu(A) > 0$ satisfies

$$\mu\{\omega \in \Omega: f^n \omega \in A, \text{ for some } n > 0\} = 1,$$

i.e. μ -almost every point in the whole space visits A . So it makes sense to talk about visiting times to a fixed set A of positive measure for μ -almost every point in Ω . Hence fix $A \in \mathfrak{B}$ with $\mu(A) > 0$, and define $\tau: \Omega \rightarrow \mathbb{N}$ by

$$\tau(x) = \tau^{(1)}(x) = \inf\{n > 0: f^n x \in A\}.$$

The function $\tau = \tau^{(1)}$ will be called the *first visiting time* of A . (Note that $\tau(x) < \infty$, μ -a.e. x .) Since by Poincaré Recurrence μ -a.e. x in A returns to A we can also define the *j -th visiting time* of A , which is a function $\tau^{(j)}: \Omega \rightarrow \mathbb{N}$ given by

$$\tau^{(j)}(x) = \inf\{n > \tau^{(j-1)}(x): f^n x \in A\},$$

for $j \geq 2$.

2.1. Return Maps and Expected Return Times.

Let $\mathbf{T}_A: A \rightarrow A$ be the induced map of f on A , i.e.

$$\mathbf{T}_A(x) = f^{(\tau(x))}(x)$$

is the *first return map* to the set A . We note that \mathbf{T}_A preserves the restriction of the probability measure μ on A . Therefore, \mathbf{T}_A preserves the conditional probability measure μ_A defined by

$$\mu_A(B) = \frac{\mu(B \cap A)}{\mu(A)}, \quad \forall B \in \mathfrak{B}.$$

In this context we recall the following result which shows that the expected return time to A given that we start in A is exactly $1/\mu(A)$.

Kac's Theorem. *Let $(\Omega, \mathcal{B}, \mu, f)$ be an ergodic dynamical system and let $A \in \mathcal{B}$ be a set of positive measure, then*

$$\mathbb{E}_A(\tau) = \int_A \tau(x) d\mu_A(x) = \frac{1}{\mu(A)} .$$

We will give an idea of the proof in the case f has an inverse μ -a.e.. (In the general case one should use the natural extension of f .)

Proof. Define, for $i \geq 1$, the sets $A_i = \{x \in A : \tau(x) = i\}$. Then $\mu(A) = \mu(\cup_{i \geq 1} A_i)$ and we may consider the Rokhlin Tower depicted in Figure 3. The map f takes a point in A and moves one step up in the ladder, until it gets to the top and then it returns back to A . In every step, f preserves the measure μ of any of the sets A_i on the base. Defining the sets $B_0 = A$, and for $j > 0$,

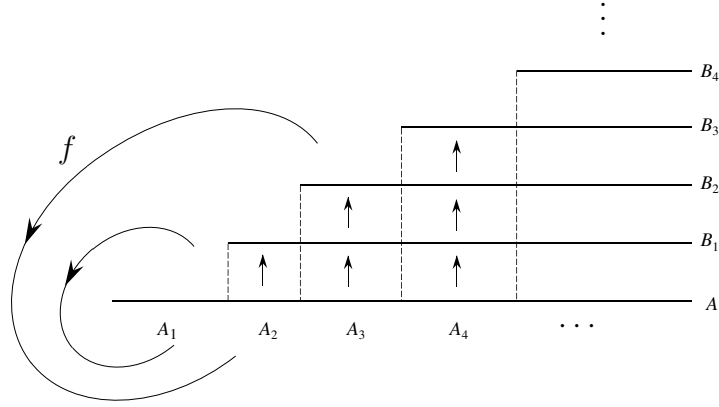


FIGURE 3. The Rokhlin Tower of f induced on A

$$B_j = f(B_{j-1}) \setminus \left(\bigcup_{i=0}^{j-1} B_i \right) ,$$

and using the fact that μ is ergodic for f^{-1} we have

$$\mu\left(\bigcup_{j \geq 0} B_j\right) = \mu\left(\bigcup_{n \geq 1} f^n A\right) = \mu(\Omega) .$$

Note also that the restriction of \mathbf{T}_A to each A_i satisfies $\mathbf{T}_A|_{A_i} = f^i|_{A_i}$. Finally, we have

$$\begin{aligned} \int_A \tau(x) d\mu(x) &= \sum_{m \geq 1} m \mu\{x \in A : \tau(x) = m\} \\ &= \sum_{m \geq 1} m \mu(A_m) = \mu\left(\bigcup_{j \geq 0} B_j\right) = 1 . \end{aligned}$$

Therefore,

$$\int_A \tau(x) d\mu_A(x) = \frac{1}{\mu(A)} .$$

□

2.2. Asymptotically Rare Events.

Suppose we are given a sequence of subsets $A_k \in \mathfrak{B}$ with $\mu(A_k) > 0$, but such that

$$\lim_{k \rightarrow \infty} \mu(A_k) = 0 .$$

We will say in this case that $\{A_k\}$ is *asymptotically rare*, since the expected time to wait for the occurrence of A_k is in general divergent with k . However, *it is not always* the case that

$$\lim_{k \rightarrow \infty} \mathbb{E}(\tau_k) \mu(A_k) = 1 ,$$

where τ_k denotes the first visiting time of A_k . We are interested in the asymptotic distribution of these visiting times, i.e. in the non-negative functions

$$\tilde{F}_k(t) = \mu \left\{ x \in \Omega : \frac{\tau_k(x)}{\mathbb{E}(\tau_k)} \leq t \right\} .$$

Since the measure of the events A_k are usually easier to compute than $\mathbb{E}(\tau_k)$, it is usually taken as a scale c_k for these times. Therefore the general problem is then to find a sequence of positive numbers $c_k \rightarrow 0$ such that there exists the limit

$$F(t) = \lim_{k \rightarrow \infty} F_k(t) = \lim_{k \rightarrow \infty} \mu \{ x \in \Omega : c_k \tau_k(x) \leq t \} ,$$

for all $t > 0$. We note that necessarily we must have

$$\int_0^\infty t dF(t) = \lim_{k \rightarrow \infty} c_k \mathbb{E}(\tau_k) .$$

Therefore if we wish to have finite expectation of the limit distribution, we should choose c_k such that the latter limit exists.

2.3. Known Results and Motivation.

In this section we will discuss special sequences of asymptotically rare events in certain ergodic dynamical systems, where the limit distribution $F(t)$ is known to exist. The idea here is only to motivate the study of this problem for symbolic dynamical systems, which will be done in the later sections. Therefore, we will not attempt to be exhaustive. (For instance, there are many results related to this subject in the framework of continuous time Markov chains and certain Particle Systems, which we will not mention cf. [Ald], [LLR]. Also missing is probably the first result on Poisson limit laws associated to dynamical systems, due to Doeblin ([Doe], 40), who considers certain random variables associated to the continued fraction expansion.)

2.3.1. Toral Automorphisms.

What we discuss here is essentially contained in Pitskel ([Pit], 91). Let $\mathcal{T}^n = \mathbb{R}^n / \mathbb{Z}^n = S^1 \times \cdots \times S^1$ denote the n -torus with its natural structure as cartesian n -product of circles $S^1 = \{z \in \mathbb{C} : |z| = 1\}$ viewed as a Lie group. Every continuous group endomorphism $f : \mathcal{T}^n \rightarrow \mathcal{T}^n$ has the form

$$f(z_1, \dots, z_n) = (z_1^{\alpha_{11}} \cdots z_n^{\alpha_{1n}}, \dots, z_1^{\alpha_{n1}} \cdots z_n^{\alpha_{nn}}) ,$$

where the constants $\alpha_{ij} \in \mathbb{Z}$. Let \mathbf{A} denote the matrix with entries $[\alpha_{ij}]$. The map f is an automorphism if and only if $\det(\mathbf{A}) = 1$ or -1 . Additionally, if \mathbf{A} has no eigenvalues of modulus one, then f is called a *hyperbolic automorphism*. The standard example is given by the matrix

$$\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} ,$$

in this case $f(z, w) = (z^2 w, z w)$. Every surjective continuous endomorphism of \mathcal{T}^n preserves μ the product measure induced by the Lebesgue measure on each of the circles S^1 , i.e. μ is the normalised Haar measure on \mathcal{T}^n .

Let now f be a hyperbolic automorphism. Fixing a base point $x \in \mathcal{T}^n$ and given $\varepsilon > 0$ we consider the first visiting time to the ball $B_\varepsilon(x)$ of radius ε centred at x , i.e. we define $\tau: \mathcal{T}^n \rightarrow \mathbb{N}$ by

$$\tau_\varepsilon(\omega) = \inf\{m > 0: f^m \omega \in B_\varepsilon(x)\}.$$

Define $A_k = B_{\varepsilon_k}(x)$ for some sequence $\varepsilon_k \rightarrow 0$. (Also write τ_k for τ_{ε_k} .) Pitskel shows the following result.

Theorem 4. *There exists $\varepsilon_k \rightarrow 0$ such that*

$$F(t) = \lim_{k \rightarrow \infty} F_k(t) = \lim_{k \rightarrow \infty} \mu\{\omega \in \mathcal{T}^n: \mu(A_k) \tau_k(\omega) \leq t\} = 1 - e^{-t},$$

for μ -a.e. $x \in \mathcal{T}^n$ and all $t > 0$.

Using Markov partitions (cf. [Bow]), this result has a consequence for finite state Markov Chains. We recall that given a general ergodic dynamical system $(\Omega, \mathfrak{B}, \mu, f)$, a finite partition of Ω into the disjoint subsets $B_1, \dots, B_\ell \in \mathfrak{B}$, defines a map $\varphi: \Omega \rightarrow \mathcal{A}$ by

$$\varphi(\omega) = (\omega_0, \omega_1, \dots),$$

where $\mathcal{A} = \{1, \dots, \ell\}$ and $\omega_n = j$ if $f^n \omega \in B_j$. We then have the following commutative diagram

$$\begin{array}{ccc} \Omega & \xrightarrow{f} & \Omega \\ \varphi \downarrow & & \downarrow \varphi \\ \mathcal{A} & \xrightarrow{\sigma} & \mathcal{A} \end{array}$$

where σ denotes the shift map. Projecting the measure μ on \mathcal{A} by $\mu \mapsto \nu = \mu \circ \varphi^{-1}$ we obtain an ergodic shift-invariant measure on \mathcal{A} .

Remark. In the precise construction of Markov partitions there is a need to consider \mathcal{A} , i.e. the space of “double-sided” sequences in ℓ -symbols. This is because in general one would like φ to be invertible ν -a.e. Here we will only consider the one-sided case, since we are only interested in the motivation of the result of Pitskel for Markov Chains. In fact, Pitskel’s approach is to consider the above result for finite state Markov Chains and transfer the result for automorphisms of \mathcal{T}^n . For the special case considered here, when $\Omega = \mathcal{T}^n$, f being a hyperbolic automorphism, and \mathcal{A} replacing \mathcal{A} , there exists a finite partition of \mathcal{T}^n such that φ is invertible ν -a.e. and ν is a Markov measure on \mathcal{A} (cf. [AW]).

Rephrasing the result of Pitskel we have the following consequence for Markov Chains. Let μ be a Markov measure on \mathcal{A} , i.e. there exists a $\ell \times \ell$ stochastic matrix $P = [P(i, j)]$ with stationary probability vector $p = (p_1, \dots, p_\ell)$ such that

$$\mu(C[i_0, \dots, i_m]_s) = p_{i_0} P(i_0, i_1) \cdots P(i_{m-1}, i_m),$$

for every cylinder $C[i_0, \dots, i_m]_s$ in \mathcal{A} . Suppose the matrix P is irreducible and aperiodic. Fixing $x = (x_0, x_1, \dots) \in \mathcal{A}$, consider the sequence of cylinders

$$A_k = C[x_0, \dots, x_k]_0,$$

then this can be interpreted as a sequence of asymptotically rare neighbourhoods of $x \in \mathcal{A}$. So if τ_k denotes the first time we see the string $[x_0 \cdots x_k]$ then Pitskel’s result implies

Theorem 5. For μ -a.e. $x \in \mathcal{A}$, there exists a subsequence $k_j \rightarrow \infty$ such that

$$F_{k_j}(t) = \mu\{\omega \in \mathcal{A} : \mu(A_{k_j}) \tau_{k_j}(\omega) \leq t\}$$

has the limit $F(t) = 1 - e^{-t}$, for every $t > 0$, when k_j goes to infinity.

In fact the above result holds for the sequence k as we will see later. Note that for $x = (1, 1, 1, \dots)$ and

$$A_k = C \underbrace{[1, \dots, 1]}_{(k+1) \text{ times}}_0,$$

the above result would partially solve one of the matching problems mentioned in Section 1.2. However, since the above result holds only for μ -a.e. x we cannot apply it just yet.

2.3.2. Hyperbolic Diffeomorphisms.

Pitskel's result was generalised by Hirata ([Hir], 93) for Hyperbolic Dynamical Systems. (The latter has also a consequence for shifts of finite type, which we will talk about later.) Let M be a smooth compact Riemannian manifold and $f: M \rightarrow M$ a diffeomorphism. Let $\Omega \in M$ denote the *non-wandering set* of f , i.e. the set of points $x \in M$ such that $\mathcal{U} \cap (\cup_{n>0} f^n \mathcal{U}) \neq \emptyset$ for every neighbourhood \mathcal{U} of x . We say that Ω is (uniformly) *hyperbolic* if there exist $C > 0$, $0 < \rho < 1$ and a continuous splitting of the tangent bundle $T_x M = E_x^s \oplus E_x^u$ into the so-called stable E^s and unstable E^u bundles, respectively, satisfying

- (i) $Df_x(E_x^s) = E_{f_x}^s$, $Df_x(E_x^u) = E_{f_x}^u$;
- (ii) $\|Df_x^n(v)\| \leq C \rho^n \|v\|$, $\|Df_x^{-n}(w)\| \leq C \rho^n \|w\|$,

for all $n > 0$, and all $v \in E_x^s$, $w \in E_x^u$, at all points $x \in \Omega$. We say that f is an *Axiom A diffeomorphism* if Ω is hyperbolic and the set of periodic points for f are dense in Ω . The map f is said to be an *Anosov diffeomorphism* if M is hyperbolic. We note here that the conditions defining Axiom A diffeomorphisms do not depend on the Riemannian metric chosen on M . However, the constants C and ρ do depend on the metric.

Under the hypothesis of topological mixing of f restricted to Ω (i.e. for every pair of non-empty open sets U, V in Ω , with respect to the induced topology of M in Ω , there exists $N > 0$ such that $U \cap f^{-n}V \neq \emptyset$, for all $n > N$), and assuming f is Axiom A, there exists a unique probability measure μ supported on Ω such that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \text{vol}(f^{-i}B) = \mu(B),$$

for every Borel set B in M such that $\text{vol}(\partial B) = 0$. Here $\text{vol}(B)$ denotes the normalised volume measure on M and ∂B is the boundary of B . The measure μ is f -invariant and it is called the *Bowen-Sinai-Ruelle* measure of f . (See Bowen [Bow] for the ergodic properties of these maps and for the construction of Markov Partitions.)

With the assumptions of topological mixing and Axiom A on f , Hirata proves a Poisson limit law for visiting times which we now describe. With respect to the Riemannian structure on M , consider $B_\varepsilon(x)$, the ε -ball centred at x , and define the visiting times to this ball

$$\tau_\varepsilon(\omega) = \tau_\varepsilon^{(1)}(\omega) = \inf\{n > 0 : f^n \omega \in B_\varepsilon(x)\},$$

and for $j \geq 2$,

$$\tau_\varepsilon^{(j)}(\omega) = \inf\{n > \tau_\varepsilon^{(j-1)}(\omega) : f^n \omega \in B_\varepsilon(x)\}.$$

Hirata's result is the following.

Theorem 6. *For μ -a.e. $x \in \Omega$, as $\varepsilon \rightarrow 0$ we have*

$$F_\varepsilon(t) = \mu\{\omega \in \Omega: \mu(B_\varepsilon(x)) \tau_\varepsilon(\omega) \leq t\} \rightarrow 1 - e^{-t},$$

and also for $j \geq 2$,

$$F_\varepsilon^{(j)}(t) = \mu\{\omega \in \Omega: \mu(B_\varepsilon(x)) (\tau_\varepsilon^{(j)}(\omega) - \tau_\varepsilon^{(j-1)}(\omega)) \leq t\} \rightarrow 1 - e^{-t}.$$

Moreover, the random variables $\tau_\varepsilon^{(1)}, \tau_\varepsilon^{(2)} - \tau_\varepsilon^{(1)}, \dots, \tau_\varepsilon^{(j)} - \tau_\varepsilon^{(j-1)}, \dots$ are asymptotically independent.

As we will see in Section 4.1, the above result means that the visiting times to an ε -ball $B_\varepsilon(x)$ centred at a fixed μ -generic point x , converges in law, as $\varepsilon \rightarrow 0$, to a Poisson point process of rate 1, when the times are rescaled by $\mu(B_\varepsilon(x))^{-1}$.

2.3.3. Piecewise-Expanding Maps of an Interval.

Here we describe the result of Collet and Galves ([CG], 93), which is also a Poisson limit law in the context of piecewise-expanding maps. Let $f: [0, 1) \rightarrow [0, 1)$ be a measurable map. We say that f is *piecewise-expanding* if there exists a partition $0 = a_0 < a_1 < \dots < a_{\ell-1} < a_\ell = 1$ such that f is C^∞ restricted to each (a_{j-1}, a_j) , f has a C^∞ -extension to $[a_{j-1}, a_j]$, and there exists $\rho > 1$ such that $|f'(x)| > \rho$ everywhere f is differentiable.

Assuming f to be eventually piecewise-expanding (i.e. there exists $m > 0$ such that f^m is piecewise-expanding) and topologically mixing, there exists a unique probability measure μ on $[0, 1)$ which is f -invariant and absolutely continuous with respect to Lebesgue measure. Let $h(x)$ be the Radon-Nikodym derivative $d\mu(x)/dx$. We know that $h(x) > 0$ for Lebesgue a.e. x . Here we will also assume that $\text{ess inf}\{h(x)\} > 0$.

Let I_k be a sequence of intervals satisfying $I_k \cap f^s I_k = \emptyset$, for $s = 1, \dots, m_k$, where m_k diverges as $k \rightarrow \infty$. This condition guarantees that I_k does not contain periodic points with uniform bounded period on k . Under these hypotheses Collet and Galves prove

Theorem 7. *The visiting times to I_k , rescaled by $\mu(I_k)^{-1}$, converge in law to a Poisson point process of rate 1.*

We say f is a *Markov map* if the image of each interval (a_{j-1}, a_j) is a union of these intervals. Such maps have a natural coding into the symbolic dynamical systems using the map φ described in Section 2.3.1, and using $B_j = [a_{j-1}, a_j)$, $j = 1, \dots, \ell$ as the finite partition of $\Omega = [0, 1)$. It can be shown that the measure μ projects onto \mathcal{A} as a Markov measure. Also any finite state Markov Chain can be obtained in this manner by considering some Markov map f . Therefore the above result has a direct consequence for finite state Markov Chains as follows.

Let μ be a Markov measure on \mathcal{A} associated to an irreducible and aperiodic matrix P . Let σ be the shift on \mathcal{A} . Consider a sequence of cylinders

$$A_k = C[x_0^k, \dots, x_k^k]_0,$$

where $x_i^j \in \{1, \dots, \ell\}$ are chosen such that

$$A_k \cap \sigma^s A_k = \emptyset,$$

for $s = 1, \dots, m_k$, where m_k diverges with k . (For instance, the sequence

$$A_1 = C[0, 1, 0]_0, A_2 = C[0, 1, 1, 0]_0, \dots, A_k = C[0, 1, \underbrace{\dots, 1}_k, 0]_0, \dots,$$

which appeared in Section 1.2 related to the occurrence of the event “no match, k matchings in a row, and no match”, satisfies this condition.) Therefore the following result proves Theorem 2.

Corollary 8. *The visiting times to A_k , rescaled by $\mu(A_k)^{-1}$, converge in law to a Poisson point process of rate 1.*

2.3.4. The Approach of Two Trajectories for Piecewise-Expanding Maps.

Let $f: [0, 1) \rightarrow [0, 1)$ be a piecewise-expanding map of the circle $S^1 \equiv [0, 1)/ \sim$, i.e. identifying 0 and 1 through the covering map $\theta \mapsto e^{2\pi i\theta}$. Suppose f is topologically mixing and let $d\mu(x) = h(x)dx$ be its absolutely continuous invariant measure. We will also assume $\text{ess inf}\{h(x)\} > 0$.

Given $\varepsilon > 0$ and two points $x, y \in S^1$, consider the times $n > 0$ such that

$$|f^n x - f^n y| \leq \varepsilon ,$$

i.e. the times for which the f -orbits of x and y get ε -close. Coelho and Collet ([CC1], 94) find a scale in which the distribution of these times (after rescaling) converge in law. First we will rephrase this problem in the general setup. Consider the 2-torus $\Omega = \mathcal{T}^2 = S^1 \times S^1$ and the map $T = f \times f: \mathcal{T}^2 \rightarrow \mathcal{T}^2$ given by $T(x, y) = (fx, fy)$. The map T preserves the product measure $m = \mu \times \mu$. Writing

$$\Delta_\varepsilon = \{(x, y) \in \mathcal{T}^2: |x - y| \leq \varepsilon\} ,$$

that is, Δ_ε is an ε -neighbourhood of the diagonal in $S^1 \times S^1$, we define the visiting times

$$\tau_\varepsilon(\omega) = \tau_\varepsilon^{(1)}(\omega) = \inf\{n > 0: T^n \omega \in \Delta_\varepsilon\} ,$$

and for $j \geq 2$,

$$\tau_\varepsilon^{(j)}(\omega) = \inf\{n > \tau_\varepsilon^{(j-1)}(\omega): T^n \omega \in \Delta_\varepsilon\} .$$

Coelho and Collet show that the distribution of the rescaled random variable τ_ε , i.e. the function

$$F_\varepsilon(t) = m\{\omega \in \mathcal{T}^2: m(\Delta_\varepsilon) \tau_\varepsilon(\omega) \leq t\} ,$$

satisfies

$$\lim_{\varepsilon \rightarrow 0} F_\varepsilon(t) = 1 - e^{-\lambda t} ,$$

for some $\lambda > 1$. We note that this provides an example where $m(\Delta_\varepsilon) \mathbb{E}(\tau_\varepsilon)$ does not tend to 1 when ε goes to zero. Also, if we consider the inter-occurrence times $\tau_\varepsilon^{(j)} - \tau_\varepsilon^{(j-1)}$, rescaled by $m(\Delta_\varepsilon)^{-1}$, then we *do not* have convergence of

$$F_\varepsilon^{(j)}(t) = m\{\omega \in \mathcal{T}^2: m(\Delta_\varepsilon) (\tau_\varepsilon^{(j)}(\omega) - \tau_\varepsilon^{(j-1)}(\omega)) \leq t\}$$

to $1 - e^{-\lambda t}$, as ε tends to zero. However, we know that there is convergence to some non-negative function, due to the convergence of the corresponding point process (see Section 4.1). Also, after rescaling, $\tau_\varepsilon^{(1)}$, $\tau_\varepsilon^{(2)} - \tau_\varepsilon^{(1)}$, \dots , $\tau_\varepsilon^{(j)} - \tau_\varepsilon^{(j-1)}$, \dots are *not* asymptotically independent. (Compare this with Theorem 2.)

Nevertheless, if we consider “entrance times” instead of visiting times, in other words, we define $N_\varepsilon^{(1)}(\omega) = \tau_\varepsilon^{(1)}(\omega)$, and for $j \geq 2$,

$$N_\varepsilon^{(j)}(\omega) = \inf\{n > N_\varepsilon^{(j-1)}(\omega): T^{n-1} \omega \notin \Delta_\varepsilon, T^n \omega \in \Delta_\varepsilon\} .$$

then the random variables $N_\varepsilon^{(1)}$, $N_\varepsilon^{(2)} - N_\varepsilon^{(1)}$, $N_\varepsilon^{(3)} - N_\varepsilon^{(2)}$, \dots are asymptotically independent (after rescaling). Moreover, rescaling these by $m(\Delta_\varepsilon)^{-1}$, more precisely, considering $m(\Delta_\varepsilon) N_\varepsilon^{(1)}$, $m(\Delta_\varepsilon) (N_\varepsilon^{(2)} - N_\varepsilon^{(1)})$, \dots , then these random variables are asymptotically identically distributed with exponential limit law of parameter $\lambda > 1$.

Using the terminology of point processes (to be introduced in Section 4.1), what Coelho and Collet prove are the following results.

Theorem 9. *The rescaled entrance times to an ε -neighbourhood of the diagonal converge in law to a Poisson point process of constant parameter $\lambda > 1$.*

Theorem 10. *The rescaled visiting times to an ε -neighbourhood of the diagonal converge in law to a marked Poisson point process of constant parameter measure.*

This problem translated into symbolic dynamics reads: given $\varepsilon > 0$ and two independent sequences $x = (x_0, x_1, x_2, \dots) \in \mathcal{A} = \{1, \dots, \ell\}$ and $y = (y_0, y_1, y_2, \dots) \in \mathcal{A}$, record the times $n > 0$ such that $\sigma^n x, \sigma^n y$ are ε -close. Here, to be ε -close means that $x_0 = y_0, x_1 = y_1, \dots, x_k = y_k$ for some integer $k = k(\varepsilon) > 0$. Therefore, this means to wait until observing k matchings of the sequences x and y . When x and y are given randomly according to different Markov measures m_1, m_2 on \mathcal{A} , then considering the shift space $(\mathcal{A} \times \mathcal{A})$ and letting the pair (x, y) denote

$$(x, y) = ((x_0, y_0), (x_1, y_1), (x_2, y_2), \dots),$$

we see that the product measure $m_1 \times m_2$ is Markov with respect to the shift transformation on $(\mathcal{A} \times \mathcal{A})$. Hence this problem may be studied by renewal theory (as outlined in Section 1.2).

2.3.5. Homeomorphisms of the circle.

Let $f: S^1 \rightarrow S^1$ be an orientation-preserving homeomorphism with irrational rotation number $\rho \in (0, 1)$. We recall that f is semi-conjugate to the rotation by ρ , i.e. there exists an orientation-preserving continuous surjective map $h: S^1 \rightarrow S^1$ such that $f \circ h = h \circ R_\rho$, where $R_\rho(x) = x + \rho \pmod{\mathbb{Z}}$ is the rotation by ρ . Here we are thinking of $S^1 = \mathbb{R}/\mathbb{Z}$ with an additive group structure. The map f has a unique invariant probability measure μ satisfying $\mu[0, x) = h(x)$ for all $x \in [0, 1)$. This is equivalent to the fact that h carries μ to Lebesgue measure on S^1 . Therefore, from the point of view of Ergodic Theory, the pair (f, μ) is isomorphic to (R_ρ, dx) .

We will mention the main result of Coelho and de Faria ([CF], 96) on the subject of visiting times of an f -orbit to a particular decreasing sequence of intervals. From the isomorphism mentioned above, the result should be read as if it is for rotations. (However, in [CF] there is also an extension of the result for the case when f is a diffeomorphism of the circle which is C^1 -conjugate to R_ρ , and the visiting times are studied using a random point with respect to Lebesgue measure.)

In what follows we will assume $f = R_\rho$ and μ is Lebesgue measure on S^1 . Consider the continued fraction expansion of ρ :

$$\rho = \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{\dots}}},$$

where the a_i 's are positive integers. We will use the notation $\rho = [a_1, a_2, \dots]$. We recall that $a_n = [G^n(\rho)]$, where $G(\rho) = 1/\rho \pmod{\mathbb{Z}}$ is the Gauss map on $[0, 1)$, and $[t]$ denotes the integer part of t . In order to describe the result we will also need to define the so-called double Gauss map $\Gamma: \mathcal{T}^2 \rightarrow \mathcal{T}^2$ given by

$$\Gamma(\rho, y) = (G\rho, [a_1, b_1, b_2, \dots]),$$

where $y = [b_1, b_2, \dots]$.

Fixing a base point $z \in S^1$. Consider the sequence of best approximations of z by its f -orbit (fz, f^2z, \dots) , i.e. find the sequence of times k such that $|f^k z - z| < |f^j z - z|$ for all $j = 1, 2, \dots, k-1$. It can be shown that this sequence is given by q_n , where

$$\frac{p_n}{q_n} = [a_1, a_2, \dots, a_n] = \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{\dots + \frac{1}{a_n}}}}$$

denotes the truncated continued fraction expansion of ρ in its irreducible form. The points z and $\bar{z} = z + 1/2 \pmod{\mathbb{Z}}$ divide the circle into two connected components, and for every $n > 1$ the points $f^{q_{n-1}}z$ and $f^{q_n}z$ lie on opposite components relative to z . Let I_n denote the interval with endpoints $\{f^{q_n}z, z\}$ contained in the same component and define $J_n = I_n \cup I_{n-1}$. Coelho and de Faria [CF] prove the following

Theorem 11. *The visiting times to J_n , rescaled by $\mu(J_n)^{-1}$, converge in law under a subsequence $\{n_i\}$ of n , if and only if, either $G^{n_i}(\rho)$ converges to zero or $\Gamma^{n_i}(\rho, \rho)$ converges to (θ, β) , for some $0 < \theta < 1$ and $\beta \in [0, 1)$. Moreover, for each choice of (θ, β) , the corresponding limit point process are pairwise different.*

The above result has the consequence

Corollary 12. *The visiting times to J_n , rescaled by $\mu(J_n)^{-1}$, do not converge in law for Lebesgue a.e. rotation number $\rho \in [0, 1)$.*

In a way, for the purpose of building a general theory of asymptotic limit laws, this result is disappointing, since if we consider visiting times to neighbourhoods of length ε around a base point, it shows that there is no universal limit law for rotations, i.e. no analog to Hirata's type of result. Nevertheless, this gives incentive to find out what are the ergodic reasons preventing the existence of universal laws in this context. We will return to these questions in the Section 4.4.

3. SHIFTS OF FINITE TYPE AND EQUILIBRIUM STATES

Let $\mathcal{A} = \{1, \dots, \ell\}$ be a finite set of symbols (represented by the number of each state). Let M be an $\ell \times \ell$ matrix of 0's and 1's representing the allowable transitions in a directed graph G of ℓ -vertices. There exists an edge of G joining i to j if and only

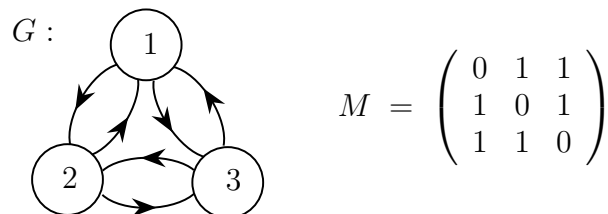


FIGURE 4. An example of G and M .

if $M(i, j) = 1$. The matrix M is called the *incidence matrix* of the directed graph G . We consider the subset $\Omega = \Omega_G \subseteq \mathcal{A}^\infty$ of infinite sequences corresponding to forward infinite paths in the graph G , i.e.

$$\Omega = \{w = (w_0, w_1, \dots) \in \mathcal{A}^\infty : M(w_{n-1}, w_n) = 1, \forall n > 0\}.$$

If $\sigma: \mathcal{A} \rightarrow \mathcal{A}$ denotes the shift map

$$\sigma(w_0, w_1, \dots) = (w_1, w_2, \dots),$$

then we will also denote by σ its restriction to Ω . We note that $\sigma(\Omega) \subseteq \Omega$, and with respect to the product topology on \mathcal{A} induced by the discrete topology on \mathcal{A} , Ω is closed. The pair (Ω, σ) is a (one-sided) *shift of finite type*. (Considering \mathcal{A} and doubly infinite allowable paths in G , we obtain the *two-sided shift of finite type*). We will assume throughout that the incidence matrix M is irreducible and aperiodic, in other words, it is the support of an irreducible and aperiodic Markov process on the states \mathcal{A} .

Let the cylinders in Ω be the intersection of the cylinders of \mathcal{A} with Ω ,

$$C[i_0, \dots, i_m]_s = \{x = (x_0, x_1, \dots) \in \Omega: x_{j+s} = i_s, \forall j = 0, \dots, m\}.$$

Recall that the Markov measures with support on Ω are the probability measures of the type:

$$\mu(C[i_0, \dots, i_m]_s) = p_{i_0} P(i_0, i_1) \cdots P(i_{m-1}, i_m),$$

where P is the transition probability matrix respecting the allowable transitions, in short

$$M(i, j) = 0 \iff P(i, j) = 0.$$

Here $p = (p_1, \dots, p_\ell)$ is the stationary vector representing the stationary distribution associated to P (i.e. $pP = p$).

3.1. Hölder Potentials.

The consideration of equilibrium states gives a richer class of shift-invariant probability measures which strictly contains the Markov measures. These are defined by a “driving” potential φ , which we will now describe.

Let $\varphi: \Omega \rightarrow \mathbb{R}$ be a bounded function. Define the n -th variation of φ by

$$\text{var}_n(\varphi) = \sup\{|\varphi(x) - \varphi(y)|: x_0 = y_0, \dots, x_n = y_n\}.$$

The function φ is continuous if and only if we have $\text{var}_n(\varphi) \rightarrow 0$, as $n \rightarrow \infty$. We say that φ has *summable variation* if

$$\sum_{n>0} \text{var}_n(\varphi) < \infty,$$

and φ is called *Hölder continuous* if there exists $0 < \theta < 1$ such that $\{\text{var}_n(\varphi) \theta^{-n}\}_{n>0}$ is a bounded sequence; that is, $\text{var}_n(\varphi)$ goes to zero exponentially fast with n .

We refer to φ as a *potential* on Ω , which could be interpreted as a source of interaction of the paths in the graph G . For instance, suppose φ satisfies

$$\sum_{\{i: M(i, x_0)=1\}} e^{\varphi(i, x_0, x_1, \dots)} = 1, \quad \forall x = (x_0, x_1, \dots) \in \Omega.$$

Then φ is said to be *normalised* in this case. We may think of $e^{\varphi(i, x_0, x_1, \dots)}$ as the transition probability to move from (x_0, x_1, \dots) to (i, x_0, x_1, \dots) . This defines a Markov process on Ω and the stationary distributions of this process are the equilibrium states of φ (which will be defined later in another manner).

Consider also the case when φ depends only on the first $k + 1$ coordinates, i.e. if $x = (x_0, x_1, \dots)$ and $y = (y_0, y_1, \dots)$ satisfy $x_0 = y_0, \dots, x_k = y_k$ then $\varphi(x) = \varphi(y)$. Let k be smallest integer satisfying this property. In this case, a normalised potential φ defines a k -step Markov process on the symbols $\{1, \dots, \ell\}$, since the transition probabilities in Ω would depend only on (x_0, x_1, \dots, x_k) .

When φ is not normalised we will use the Ruelle operator and its spectral properties to study equilibrium states, which will be the stationary distributions of the Markov process defined above associated to a normalised φ' canonically associated to φ .

3.2. Entropy and Pressure.

Let μ be a shift-invariant probability measure on Ω . Recall that the entropy of μ is given by the limit

$$h(\mu) = \lim_{n \rightarrow \infty} -\frac{1}{n} \sum_{C \in B_n(\Omega)} \mu(C) \log \mu(C),$$

where $B_n(\Omega)$ denotes the cylinders in Ω of type $C[x_0, \dots, x_n]_0$. The pressure of the potential $\varphi: \Omega \rightarrow \mathbb{R}$ is defined by

$$\mathcal{P}(\varphi) = \sup \left\{ h(\mu) + \int \varphi d\mu \right\},$$

where the supremum is taken over all shift-invariant probability measures μ on Ω . We say that a shift-invariant probability measure m is an *equilibrium state* of φ if m satisfies

$$\mathcal{P}(\varphi) = h(m) + \int \varphi dm.$$

For every continuous potential φ , there exists at least one equilibrium state m . If φ has summable variation then there exists a unique equilibrium state (see [Bow]). Note that when $\varphi = 0$ its equilibrium state m is the measure of maximal entropy, this is a one-step Markov measure known as the Parry measure on the shift of finite type (Ω, σ) (cf. [Par]). For general properties of the pressure for continuous dynamical systems see Walters [Wal].

We will be mostly interested when φ is Hölder continuous because of the nice spectral properties of the Ruelle operator. Here it is appropriate to mention that with the hypothesis of φ being normalised, there was a belief that φ continuous would imply uniqueness of its equilibrium state. This is shown to be false with an example given in [BK]. Recently, Coelho and Quas ([CQ], 97) show that if the variations of φ satisfy

$$\sum_{n=r}^{\infty} \prod_{i=r}^n (1 - a_i) = \infty,$$

for some $r \geq 1$, where $a_i = \frac{\ell}{2} \text{var}_i(\varphi)$, and ℓ is the cardinality of \mathcal{A} , then φ has a unique equilibrium state m and the natural extension of (Ω, σ, m) is Bernoulli (i.e. metrically isomorphic to a Bernoulli scheme).

In the next section we will drop the condition of φ being normalised.

3.3. Ruelle-Perron-Frobenius Operator.

For φ continuous, define an operator $L = L_\varphi: C(\Omega) \rightarrow C(\Omega)$ on the space of continuous functions $C(\Omega)$, $\psi \mapsto L_\varphi(\psi)$, by

$$L_\varphi(\psi)(x) = \sum_{M(i, x_0)=1} e^{\varphi(ix)} \psi(ix),$$

where $ix = (i, x_0, x_1, \dots)$ and $x = (x_0, x_1, \dots)$. The operator L has important spectral properties when restricted to certain subspaces of $C(\Omega)$. For instance, if φ is Hölder, L_φ acts as an operator on the Hölder continuous functions, and restricted to this subspace, it has a simple and isolated eigenvalue $e^{\mathcal{P}(\varphi)}$. The rest of the spectrum is contained in a disc of radius strictly smaller than $e^{\mathcal{P}(\varphi)}$. There exists a strictly positive

eigenfunction associated to $e^{\mathcal{P}(\varphi)}$. Therefore there exists w Hölder continuous such that $w(x) > 0$ everywhere and

$$L_\varphi(w)(x) = e^{\mathcal{P}(\varphi)}w(x),$$

for all $x \in \Omega$. This implies that

$$\sum_{M(i,x_0)=1} e^{\varphi(ix)}w(ix) = e^{\mathcal{P}(\varphi)}w(x),$$

and then

$$\sum_{M(i,x_0)=1} e^{\varphi(ix)-\mathcal{P}(\varphi)+\log w(ix)-\log w(x)} = 1.$$

Therefore, defining $\varphi'(x) = \varphi(x) - \mathcal{P}(\varphi) - \log w(\sigma x) + \log w(x)$ we would have φ' is Hölder and normalised. The equilibrium state of φ is the same equilibrium state of φ' . In fact, m the equilibrium state of φ is the unique shift-invariant probability measure which is fixed by the dual operator L_φ^* , i.e.

$$\int L_\varphi(\psi) dm = \int \psi dm, \quad \forall \psi \in C(\Omega).$$

Another way to characterise the equilibrium state m is a *homogeneous* property: there exists $K > 1$ such that

$$\frac{1}{K} < \frac{m(C[x_0, x_1, \dots, x_n]_0)}{\exp\left\{- (n+1)\mathcal{P}(\varphi) + \sum_{j=0}^n \varphi(\sigma^j x)\right\}} < K,$$

for all $x = (x_0, x_1, \dots) \in \Omega$. Replacing φ by the normalised φ' , we may assume φ is normalised and Hölder. (Recall that this implies $\mathcal{P}(\varphi) = 0$ and the Ruelle operator L_φ will have 1 as a simple isolated eigenvalue, with the rest of the spectrum contained in a disc of radius strictly smaller than 1, when operating on the space of Hölder continuous functions.) Therefore, if φ is Hölder then there exist $0 < \rho < 1$ and $c > 0$ (independent of ψ) such that

$$\|L_\varphi^n(\psi) - \int \psi dm\|_\Omega \leq c\rho^n.$$

This implies that, if ψ is continuous then

$$L_\varphi^n(\psi)(x) \xrightarrow{n \rightarrow \infty} \int \psi dm,$$

uniformly in $x \in \Omega$ (this is done by approximating continuous functions by Hölder continuous functions).

On what follows we will assume (Ω, σ) is an irreducible and aperiodic shift of finite type with incidence matrix M , defining graph G , and with equilibrium state m of a Hölder normalised potential φ .

3.4. The Central Limit Theorem.

Let $\psi: \Omega \rightarrow \mathbb{R}$ be a Hölder continuous function such that $\int \psi dm = 0$ (mean zero). Define the *variance* of ψ by

$$\underline{\sigma}^2 = \underline{\sigma}^2(\psi) = \lim_{n \rightarrow \infty} \frac{1}{n} \int \left(\sum_{j=0}^{n-1} \psi(\sigma^j x) \right)^2 dm(x),$$

where we think of $\psi(x), \psi(\sigma x), \psi(\sigma^2 x), \dots$ being a sequence of random variables defined on the probability space (Ω, m) . This sequence is stationary since m is shift-invariant. If ψ does not have zero mean, define $\underline{\sigma}^2(\psi) = \underline{\sigma}^2(\tilde{\psi})$, where $\tilde{\psi} = \psi - \int \psi dm$. The following result shows a connection between the pressure and the above statistical constants (see Ruelle [Rue] for a proof). Let \mathcal{H} denote the space of Hölder continuous functions defined on Ω .

Theorem 13. *Consider the pressure map $\mathcal{P}: \mathcal{H} \rightarrow \mathbb{R}$ given by $\psi \mapsto \mathcal{P}(\psi)$. For every fixed function $\psi \in \mathcal{H}$, and s a real parameter, we have*

- (a) $s \mapsto \mathcal{P}(\varphi + s\psi)$ is a real analytic function ;
- (b) $\frac{d}{ds} \{ \mathcal{P}(\varphi + s\psi) \} \Big|_{s=0} = \int \psi dm$;
- (c) $\frac{d^2}{ds^2} \{ \mathcal{P}(\varphi + s\psi) \} \Big|_{s=0} = \underline{\sigma}^2(\psi)$.

The variance $\underline{\sigma}^2(\psi)$ is zero if and only if there exists a Hölder function $h: \Omega \rightarrow \mathbb{R}$ satisfying $h(\sigma x) - h(x) = \psi(x)$ for all $x \in \Omega$. It can be shown that if there exists a measurable function h satisfying the latter property for m -a.e. x , then there exists a Hölder continuous function satisfying this property. This is sometimes referred to as *rigidity* of the equilibrium state m .

A good exposition of the subject in the present Section, which contains an elegant proof of the next couple of results can be found in [CP], see also [PP].

Theorem 14 (Central Limit Theorem). *Suppose that $\underline{\sigma}(\psi) \neq 0$ and $\int \psi dm = 0$. Define*

$$(S_n \psi)(x) = \sum_{j=0}^{n-1} \psi(\sigma^j x) .$$

Then $\frac{S_n \psi}{\underline{\sigma} \sqrt{n}}$ is asymptotically Normal with zero mean and variance 1, which means that

$$F_n(t) = m \left\{ x \in \Omega : \frac{1}{\underline{\sigma} \sqrt{n}} \sum_{j=0}^{n-1} \psi(\sigma^j x) \leq t \right\} \xrightarrow{n \rightarrow \infty} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-s^2/2} ds = N(t) .$$

Furthermore, $\|F_n - N\|_\infty = O(1/\sqrt{n})$, i.e. $\{\sqrt{n} \|F_n - N\|_\infty\}$ is a bounded sequence.

This is the Normal asymptotic law for Hölder equilibrium states. When ψ satisfies an extra condition which prevents $S_n \psi$ to be asymptotically lattice distributed, then there is a refinement in the Central Limit Theorem. More precisely, we say that ψ defines a *non-lattice distribution* if, for all $a \in \mathbb{R}$, the values $S_n \psi(x) - na$ for x such that $\sigma^n x = x$, generate a dense additive subgroup of \mathbb{R} . This is equivalent to the condition that whenever

$$h(\sigma x) = \alpha e^{is\psi(x)} h(x) ,$$

for some complex constant α and measurable function h , then necessarily $s = 0$ and $h(x)$ is a constant m -a.e. x .

Theorem 15. *Let $\psi \in \mathcal{H}$ be given such that $\underline{\sigma}(\psi) \neq 0$ and $\int \psi dm = 0$. If ψ defines a non-lattice distribution then*

$$F_n(t) - N(t) = \frac{\mathcal{P}^{(3)}(0)}{6 \underline{\sigma} \sqrt{2\pi n}} \left(1 - \frac{t^2}{\underline{\sigma}^2} \right) e^{-t^2/2\underline{\sigma}^2} + o\left(\frac{1}{\sqrt{n}}\right) ,$$

where $\mathcal{P}^{(3)}(0)$ denotes the third derivative of the pressure map $s \mapsto \mathcal{P}(\varphi + s\psi)$ at $s = 0$.

Higher-order asymptotics can also be obtained under stronger conditions on ψ (see [CP]).

3.5. Pianigiani-Yorke Measure.

In this section we will find the correct scale to study the convergence of asymptotic visiting times to a subsystem of finite type inside the shift of finite type (Ω, σ) .

Suppose for a moment that we have two equilibrium states m_1 of φ_1 and m_2 of φ_2 (φ_1, φ_2 normalised and Hölder continuous on Ω).

Let $x = (x_0, x_1, \dots)$ be a generic point for m_1 , i.e. the shift orbit of x determines the statistical behaviour of m_1 , or more precisely, for every continuous function $g: \Omega \rightarrow \mathbb{R}$ we have

$$\frac{1}{n} \sum_{j=0}^{n-1} g(\sigma^j x) \xrightarrow{n \rightarrow \infty} \int g dm_1 .$$

Let $y = (y_0, y_1, \dots)$ be a generic point for m_2 . We may think of the pair (x, y) as a point in (\mathcal{A}^2) by writing

$$(x, y) = ((x_0, y_0), (x_1, y_1), \dots) .$$

Defining an incidence matrix on the states \mathcal{A}^2 by

$$\widetilde{M}((i_1, i_2), (j_1, j_2)) = M(i_1, j_1) M(i_2, j_2) ,$$

we obtain the product shift of finite type $\widetilde{\Omega} \subseteq (\mathcal{A}^2)$. The corresponding shift map $\tilde{\sigma}$ is clearly conjugate to $\sigma \times \sigma$. The product measure $m_1 \times m_2$ is canonically isomorphic to the equilibrium state of the potential $\tilde{\varphi}: \widetilde{\Omega} \rightarrow \mathbb{R}$ given by

$$(2) \quad \tilde{\varphi}((x_0, y_0), (x_1, y_1), \dots) = \varphi_1(x) + \varphi_2(y) .$$

For future reference we will denote $\tilde{\varphi}$ by $\varphi_1 \otimes \varphi_2$.

Now let $\Delta \subseteq \{1, \dots, \ell\}^2$ be the subset of \mathcal{A}^2 consisting of pairs $\{(j, j)\}_{j=1, \dots, \ell}$. Then (x, y) match on the first $k+1$ coordinates if

$$(x_i, y_i) \in \Delta, \quad \forall i = 0, 1, \dots, k .$$

Therefore, studying the asymptotic properties of the waiting time for k matchings in a row for two sequences x and y , is equivalent to visiting a subset of states Δ in the product symbol set, consecutively $k+1$ times, in the product shift of finite type.

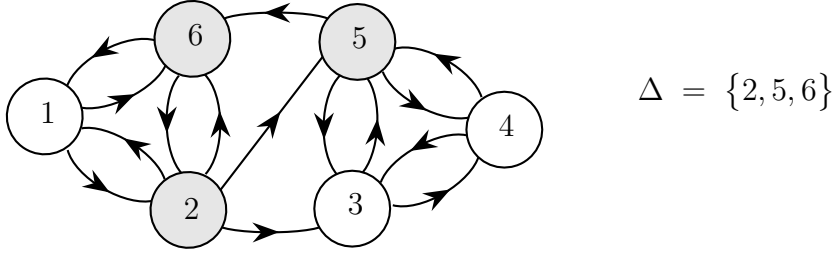
This leads us to the consideration of a general case, where $\Delta \subseteq \{1, \dots, \ell\}$ is a subset of states, $x = (x_0, x_1, \dots)$ is a random point in (Ω, m) (i.e. generic with respect to m) and we study the times of k successive hits in Δ . (Here we are identifying Δ with the union of cylinders $\cup_{i \in \Delta} C[i]_0$ in Ω .) Let χ_Δ indicate a hit in Δ at time 0, i.e. the characteristic function of Δ :

$$\chi_\Delta(x) = \begin{cases} 1 & \text{if } x_0 \in \Delta; \\ 0 & \text{if } x_0 \notin \Delta. \end{cases}$$

Then hitting $k+1$ times in succession means that

$$\prod_{j=r}^{r+k} \chi_\Delta(\sigma^j x) = 1 ,$$

for some $r \geq 0$. Suppose that the incidence matrix M restricted to the states of Δ is irreducible and aperiodic. (Therefore, thinking of the states of Δ as painted states, if we forget the transitions in Ω which passes through non-painted states, then we obtain an irreducible and aperiodic subsystem of finite type Ω_Δ in (Ω, σ) . See Figure 5 for an example.)

FIGURE 5. An example of Δ .

Let Δ_k denote the union of cylinders of length $k + 1$ with symbols belonging to Δ , that is

$$\Delta_k = \{x = (x_0, x_1, \dots) \in \Omega : x_i \in \Delta, \forall i = 0, 1, \dots, k\}.$$

The visiting times to Δ_k of the shift orbit of a point in Ω , corresponds to the times of k -consecutive hits in Δ . Define as before

$$\tau_k(x) = \inf\{n > 0 : \sigma^n x \in \Delta_k\}$$

as the first visiting time to Δ_k . (In the next chapter, we will consider the whole process of subsequent visits to Δ_k).

In order to study the asymptotic properties of these times, as mentioned in Section 2, we need to find a scale $c_k \rightarrow 0$ such that the distribution

$$F_k(t) = m\{x \in \Omega : c_k \tau_k(x) \leq t\}$$

converges as $k \rightarrow \infty$. As we saw, the natural scale here is $\mathbb{E}(\tau_k)$. However, as we also noticed, the study of $m(\Delta_k)$ could be a source of inspiration, and indeed we will see that it will provide a good scale.

Recall some properties of the operator L_φ :

- (i) $\int \psi dm = \int L_\varphi(\psi) dm, \quad \forall \psi \in C(\Omega);$
- (ii) $\int \psi_1(\psi_2 \circ \sigma) dm = \int \psi_2 L_\varphi(\psi_1) dm, \quad \forall \psi_1, \psi_2 \in C(\Omega).$

First we note that

$$\chi_{\Delta_k} = \prod_{j=0}^k \chi_{\Delta \circ \sigma^j} = \chi_{\Delta} \cdot \left[\left(\prod_{j=0}^{k-1} \chi_{\Delta \circ \sigma^j} \right) \circ \sigma \right].$$

Therefore, using the properties (i) and (ii) we obtain

$$\begin{aligned} \int_{\Delta_k} \psi dm &= \int \psi \cdot \chi_{\Delta_k} dm \\ &= \int \psi \cdot \chi_{\Delta} \cdot (\chi_{\Delta_{k-1}} \circ \sigma) dm \\ &= \int L_\varphi(\psi \cdot \chi_{\Delta}) \cdot \chi_{\Delta_{k-1}} dm \\ &= \int_{\Delta_{k-1}} L_\varphi(\psi \cdot \chi_{\Delta}) dm. \end{aligned}$$

Thus, in short we have

$$(3) \quad \int_{\Delta_k} \psi \, dm = \int_{\Delta_{k-1}} L_\varphi(\psi \cdot \chi_\Delta) \, dm, \quad \forall \psi \in C(\Omega).$$

Define the restricted Ruelle operator $L_\Delta: C(\Omega) \rightarrow C(\Omega)$ by

$$\begin{aligned} L_\Delta(\psi)(x) &= L_\varphi(\psi \cdot \chi_\Delta)(x) \\ &= \sum_{\substack{M(i,x_0)=1 \\ i \in \Delta}} e^{\varphi(ix)} \psi(ix) \chi_\Delta(ix) \\ &= \sum_{\substack{M(i,x_0)=1 \\ i \in \Delta}} e^{\varphi(ix)} \psi(ix). \end{aligned}$$

With this definition we obtain from (3)

$$(4) \quad \int_{\Delta_k} \psi \, dm = \int_{\Delta_{k-1}} L_\Delta(\psi) \, dm = \dots = \int L_\Delta^{k+1}(\psi) \, dm,$$

for all $\psi \in C(\Omega)$.

Knowing some spectral properties of L_Δ would be useful here. We recall a recent result by Collet et al. ([CMS], 97) about Pianigiani-Yorke measures for aperiodic subsystems of finite type. This proves to be the key point here.

Consider the restricted subsystem Ω_Δ (suppose that $\Delta \subsetneq \{1, \dots, \ell\}$). The restriction of φ to Ω_Δ defines a Hölder potential φ_Δ . Consider the pressure \mathcal{P}_Δ of φ_Δ as a function defined on the subsystem Ω_Δ . The Ruelle operator $L_{\varphi_\Delta}: C(\Omega_\Delta) \rightarrow C(\Omega_\Delta)$ restricted to Hölder continuous functions defined on Ω_Δ , has a simple isolated eigenvalue $e^{\mathcal{P}_\Delta}$. Thus, there exists a positive Hölder function $\tilde{w}: \Omega_\Delta \rightarrow \mathbb{R}^+$ such that

$$\begin{aligned} L_{\varphi_\Delta}(\tilde{w})(x) &= \sum_{\substack{M(i,x_0)=1 \\ i \in \Delta}} e^{\varphi(ix)} \tilde{w}(ix) \\ &= e^{\mathcal{P}_\Delta} \tilde{w}(x), \end{aligned}$$

for all $x \in \Omega_\Delta$. We note that, since $\Delta \neq \{1, \dots, \ell\}$ then necessarily $\mathcal{P}_\Delta < 0$ (i.e. $e^{\mathcal{P}_\Delta} < 1$). This is because we are assuming φ to be normalised. (In the general case we would have always $\mathcal{P}(\varphi) > \mathcal{P}_\Delta$.)

Collet, Martinez and Schmitt [CMS] proves

Theorem 16. *There exists a strictly positive function $w_\Delta: \Omega \rightarrow \mathbb{R}^+$ such that*

$$L_\Delta(w_\Delta) = e^{\mathcal{P}_\Delta} w_\Delta,$$

and $w_\Delta|_{\Omega_\Delta} \equiv \tilde{w}$. Furthermore,

$$\left\| e^{-n\mathcal{P}_\Delta} L_\Delta^n(\psi) - w_\Delta \int_{\Omega_\Delta} \psi \, dm_\Delta \right\|_\Omega \xrightarrow{n \rightarrow \infty} 0,$$

for all $\psi \in C(\Omega)$. (Here m_Δ is the unique equilibrium state of φ_Δ on the subsystem Ω_Δ .)

Let \mathcal{B} denote the Borel σ -algebra on Ω . The measure μ_{PY} on (Ω, \mathcal{B}) defined by

$$\mu_{PY}(B) = \int_B w_\Delta \, dm, \quad \forall B \in \mathcal{B},$$

is called the *Pianigiani-Yorke measure* of the subsystem (Ω_Δ, σ) . This is a quasi-stationary measure related to the asymptotic approach to the subsystem (Ω_Δ, σ)

(i.e. $\mu_{PY} = \alpha \mu_{PY} \circ \sigma^{-1}$, for some $\alpha > 1$). In fact, in our context $\alpha = e^{-\mathcal{P}\Delta}$. This measure satisfies

$$\frac{\mu_{PY}(B)}{\mu_{PY}(\Omega)} = \lim_{k \rightarrow \infty} m(\sigma^{-k} B \mid \Delta_{k-1}),$$

for all $B \in \mathcal{B}$. The Pianigiani-Yorke measure is not shift-invariant and carries an information of how subsystems of finite type are embedded into larger systems. It would be nice if there would be a non-trivial relation between these measures and certain classification problems of shifts of finite type, but this is subject for further developments.

To finish this section we state the next result, which shows how fast the measure of Δ_k goes to zero and also gives the implied constant.

Lemma 17. *We have*

$$\lim_{k \rightarrow \infty} e^{-k\mathcal{P}\Delta} m(\Delta_{k-1}) = \mu_{PY}(\Omega).$$

Coming back to our original problem of visiting times, the above result means that $c_k = e^{k\mathcal{P}\Delta}$ is a suitable scale for rescaling the visiting times to Δ_k . In the next Section we will recall and give a proof of Lemma 17. There will be also a characterisation of the asymptotic limit law of visiting times to Δ_k .

4. POINT PROCESSES AND CONVERGENCE IN LAW

In this section we discuss some general properties of convergence in law for point processes. This is the natural terminology to treat the problems mentioned earlier concerning visiting times, matching times, etc.

If (Ω, μ) is a probability space, a point process with values in $[0, +\infty)$ defined on Ω is a map $\tau: \Omega \rightarrow \mathcal{M}_\sigma[0, +\infty)$, where $\mathcal{M}_\sigma[0, +\infty)$ denotes the σ -finite measures on $[0, +\infty)$ such that the support, $\text{supp}(\tau(w)) \subseteq [0, +\infty)$, of the measure $\tau(w)$ is discrete for m -a.e. $w \in \Omega$. For every continuous function $g: [0, +\infty) \rightarrow \mathbb{R}$ of compact support we may consider the random variable

$$N(g)(w) = \int_0^\infty g d\tau(w).$$

The statistical properties of τ can then be studied, varying the functions g . The typical choice for g is the indicator function of an interval $[a, b)$, and in this case, the above integral defines a random variable corresponding to the number of hits of the point process in the interval $[a, b)$.

An example is the *Poisson point process of rate $\lambda > 0$* , which is the point process defined by

$$\tau(w) = \sum_{n>0} \delta_{X_n(w)},$$

where δ_t denotes the Dirac measure at the point t , and the random variables $X_1, X_2 - X_1, X_3 - X_2, \dots$ are independent identically distributed with exponential law of parameter $\lambda > 0$, i.e. defining $X_0 \equiv 0$ we have

$$\mathbb{P}(X_i - X_{i-1} \leq t) = 1 - e^{-\lambda t},$$

for all $i \geq 1$, and the joint distributions are products of the marginals,

$$\mathbb{P}(X_1 \leq t_1, X_2 - X_1 \leq t_2, \dots, X_j - X_{j-1} \leq t_j) = \prod_{i=1}^j (1 - e^{-\lambda t_i}),$$

for every $j \geq 1$ and every choice of $t_i \in [0, +\infty)$. Let $g : [0, +\infty) \rightarrow \mathbb{R}$ be given by $g(t) = \chi_{[a,b)}(t)$. Then

$$N(g)(w) = \sum_{n>0} \chi_{[a,b)}(X_n(w))$$

i.e. $N(g)$ counts the number of hits at $[a, b)$ for the random impulses given by the Poisson process. Another interesting random variable is $N(g_s)$, where $g_s(t) = \chi_{[0,s)}(t)$, since it has the Poisson distribution of parameter λs :

$$\mathbb{P}(N(g_s) = k) = \frac{(\lambda s)^k}{k!} e^{-\lambda s},$$

for all $k = 0, 1, \dots$. We will also consider another naturally defined point process in the next Section.

4.1. Convergence of Point Processes.

We say that a sequence of point processes τ_n converges in law to a point process τ , if for every continuous $g : [0, +\infty) \rightarrow \mathbb{R}$ with compact support, the random variables

$$N_n(g)(w) = \int g d\tau_n(w)$$

converge in distribution to some random variable $N(g)(w)$. This is equivalent to having convergence of the moment generating function

$$\mathbb{E}(e^{z N_n(g)}) \xrightarrow{n \rightarrow \infty} \mathbb{E}(e^{z N(g)}),$$

for z in some open neighbourhood of the origin in the complex plane \mathbb{C} . In the case $\tau : \Omega \rightarrow \mathcal{M}_\sigma[0, +\infty)$ is a Poisson point process of parameter $\lambda > 0$, it is not difficult to prove that (cf. [DV])

$$(5) \quad \mathbb{E}(e^{z N(g)}) = \exp \left\{ \lambda \int_0^\infty (e^{z g(y)} - 1) dy \right\}.$$

Therefore, to show that a sequence of point processes, representing the visits to an asymptotically rare sequence of events, converges to a Poisson point process, it suffices to show that the moment generating functions converge to the expression in (5).

Actually, we will show convergence of a sequence of point processes to a marked Poisson point process, which we now describe. This consists of a point process with

$$\tau(w) = \sum_{n>0} K_n(w) \delta_{X_n(w)}$$

where $K_n(w) \in \mathbb{N}$ are i.i.d. random variables, $X_1, X_2 - X_1, X_3 - X_2, \dots$ are real positive i.i.d. random variables and K_i, X_j are independent for any collection of i 's and j 's. Moreover, the sequence $X_1, X_2 - X_1, X_3 - X_2, \dots$ have exponential distribution of parameter $\lambda > 0$. If π denotes the distribution of each K_n , i.e. $\pi = (\pi_k)_{k \geq 1}$, and $\pi_k = \mathbb{P}(K_n = k)$, then τ is called a *marked Poisson point process with constant parameter measure* $\lambda \pi$.

The moment generating function of such a process is given by

$$\mathbb{E}(e^{z N(g)}) = \exp \left\{ \lambda \sum_{k=1}^\infty \pi_k \int_0^\infty (e^{z k g(y)} - 1) dy \right\}.$$

Thus, one way to prove the convergence to a marked Poisson point process is to prove convergence of the moment generating function and identifying the limit with the above expression in a neighbourhood of the origin in \mathbb{C} . This was the method used

by Coelho and Collet [CC1] for the process of visits to the diagonal for piecewise expanding maps of the circle (see Section 2.3.4).

4.2. Entrance Times and Visiting Times.

What will be done here on what follows is also joint work with P. Collet [CC2] (under construction). We recall the setup done in Section 3 regarding the construction of a particular sequence of asymptotically rare events Δ_k in an irreducible and aperiodic shift of finite type. For the sake of clarity in the exposition, we will repeat some of the definitions.

Let $\mathcal{A} = \{1, \dots, \ell\}$ be the states of an irreducible and aperiodic shift of finite type (Ω, σ) , $\Omega \subseteq \mathcal{A}$. Let $\varphi \in \mathcal{H}$ be a Hölder potential on Ω defining an equilibrium state m on Ω . We consider the painting of some states of Ω by choosing $\Delta \subseteq \{1, \dots, \ell\}$, where $\Delta \neq \{1, \dots, \ell\}$. As we mentioned before, Δ defines a subsystem of finite type in Ω , which we called Ω_Δ . We assume that this subsystem is irreducible and aperiodic.

By changing the potential φ to another Hölder potential defining the same equilibrium state m , we may suppose φ is normalised, i.e.

$$\sum_{M(i, x_0)=1} e^{\varphi(ix)} = 1,$$

for all $x \in \Omega$, where M denotes the incidence matrix of Ω . We also defined the restricted Ruelle operator

$$L_\Delta(\psi)(x) = L_\varphi(\psi \cdot \chi_\Delta)(x) = \sum_{\substack{M(i, x_0)=1 \\ i \in \Delta}} e^{\varphi(ix)} \psi(ix).$$

If $e^{\mathcal{P}_\Delta}$ is the simple isolated eigenvalue of the Ruelle operator L_{φ_Δ} corresponding to the restriction φ_Δ of φ to the subsystem Ω_Δ , then by Theorem 16, there exists $w_\Delta: \Omega \rightarrow \mathbb{R}$ Hölder continuous and strictly positive such that

$$(6) \quad L_\Delta(w_\Delta) = e^{\mathcal{P}_\Delta} w_\Delta, \quad \text{and} \quad \|e^{-n\mathcal{P}_\Delta} L_\Delta^n(\psi) - w_\Delta \int_{\Omega_\Delta} \psi dm_\Delta\|_\Omega \xrightarrow{n \rightarrow \infty} 0,$$

where m_Δ denotes the equilibrium state of φ_Δ with support on the subsystem Ω_Δ . We defined the Pianigiani-Yorke measure

$$\mu_{PY}(B) = \int_B w_\Delta dm,$$

for any Borel set $B \subseteq \Omega$. Let $\Delta_k = \{x \in \Omega : x_i \in \Delta, i = 1, \dots, k\}$.

Proof of Lemma 17. We want to show that

$$\lim_{k \rightarrow \infty} e^{-k\mathcal{P}_\Delta} m(\Delta_{k-1}) = \mu_{PY}(\Omega) = \int w_\Delta dm.$$

First we recall from (4) that

$$\int_{\Delta_{k-1}} \psi dm = \int L_\Delta^k(\psi) dm,$$

for all $\psi \in C(\Omega)$. We also have from (6)

$$e^{-k\mathcal{P}_\Delta} L_\Delta^k(\psi) = w_\Delta \int_{\Omega_\Delta} \psi dm_\Delta + o(1),$$

which implies that

$$e^{-k\mathcal{P}_\Delta} \int_{\Delta_{k-1}} \psi \, dm = \left(\int w_\Delta \, dm \right) \left(\int_{\Omega_\Delta} \psi \, dm_\Delta \right) + o(1) .$$

Therefore, for every $\psi \in C(\Omega)$ we obtain

$$\lim_{k \rightarrow \infty} e^{-k\mathcal{P}_\Delta} \int_{\Delta_{k-1}} \psi \, dm = \left(\int w_\Delta \, dm \right) \left(\int_{\Omega_\Delta} \psi \, dm_\Delta \right) .$$

Putting $\psi \equiv 1$ gives the desired result

$$e^{-k\mathcal{P}_\Delta} m(\Delta_{k-1}) \xrightarrow[k \rightarrow \infty]{} \int w_\Delta \, dm . \quad \square$$

Now we turn our attention to the times when we see blocks of length k consisting of painted symbols. For each fixed k , consider the *visiting time point process* $\tau_k: \Omega \rightarrow \mathcal{M}_\sigma[0, +\infty)$ defined by

$$\tau_k(x) = \sum_{j \geq 0} \chi_{\Delta_{k-1}}(\sigma^j x) \delta_{j \cdot c_k} ,$$

where we introduced the scale $c_k = e^{k\mathcal{P}_\Delta}$. (Recall that $P_\Delta < 0$.)

Consider also the *entrance time point process* of Δ_{k-1} rescaled by c_k ,

$$\tau_k^e(x) = \sum_{j \geq 1} \chi_{\Delta_{k-1}}(\sigma^j x) \chi_{\Delta_{k-1}^c}(\sigma^{j-1} x) \delta_{j \cdot c_k} .$$

The following results summarise the asymptotic behaviour of these processes.

Theorem 18. *The entrance time point process τ_k^e converges in law to a Poisson point process of parameter $\lambda > 0$.*

Theorem 19. *The visiting time point process τ_k converges in law to a marked Poisson point process with constant parameter measure $\lambda\pi$, where $\pi = (\pi_j)_{j \geq 1}$ is geometrically distributed.*

Intuitively, we may interpret this result by thinking of the subsystem Ω_Δ in Ω as an attractor in probability, in the sense that considering a sequence of open neighbourhoods \mathcal{U}_k in Ω such that $\Omega_\Delta = \bigcap_k \mathcal{U}_k$, we must have necessarily $m(\mathcal{U}_k) \rightarrow 0$, and for very large k , a m -generic point $x \in \Omega$ will take (after rescaling by the proper power of the restricted pressure of Ω_Δ) an exponential time to enter \mathcal{U}_k . When it is in there, it takes an asymptotically geometric time to leave this neighbourhood, and as soon as it is out, the process regenerates itself.

Turning to a more technical discussion, we continue to give an idea of how one can prove the above results. First we note that these results are consequences of the following

Theorem 20. *There exists $\lambda > 0$ and a probability measure $\pi = (\pi_j)_{j \geq 1}$ on $\{1, 2, \dots\}$ such that, for every $g: [0, +\infty) \rightarrow \mathbb{R}$ continuous with compact support, the random variable*

$$N_k(g)(x) = \sum_{j \geq 0} \chi_{\Delta_{k-1}}(\sigma^j x) g(j \cdot c_k) ,$$

satisfies

$$(7) \quad \mathbb{E}(e^{zN_k(g)}) = \int N_k(g)(x) \, dm(x) \xrightarrow[k \rightarrow \infty]{} \exp \left\{ \lambda \sum_{j=1}^{\infty} \pi_j \int_0^{\infty} (e^{zjg(y)} - 1) \, dy \right\} .$$

In fact the constants λ and π_j are given by the following expressions

$$\lambda = (1 - e^{\mathcal{P}_\Delta}) \int w_\Delta dm, \quad \pi_j = (1 - e^{\mathcal{P}_\Delta}) e^{(j-1)\mathcal{P}_\Delta}.$$

Idea of the Proof of Theorem 20.

Considering the factorial moments of $N_k(g)$, i.e. $\mathbb{E}(N_k(g)^n)$, we show that the limit $\nu_n = \lim_{k \rightarrow \infty} \mathbb{E}(N_k(g)^n)$ exists. Then we identify the function

$$F_g(z) = \sum_{n \geq 0} \frac{z^n}{n!} \nu_n$$

with the complex function in (7), for z in a small neighbourhood of the origin in \mathbb{C} . This is similar to what is done in [CC1], so we illustrate only the main steps.

We see that

$$\begin{aligned} \mathbb{E}(N_k(g)^n) &= \mathbb{E} \left[\left(\sum_{j \geq 0} \chi_{\Delta_{k-1}}(\sigma^j x) g(j \cdot c_k) \right)^n \right] \\ &= \sum_{0 \leq j_1, \dots, 0 \leq j_n} \prod_{s=1}^n g(j_s \cdot c_k) \mathbb{E} \left(\prod_{s=1}^n \chi_{\Delta_{k-1}} \circ \sigma^{j_s} \right). \end{aligned}$$

Let us compute this for $n = 1$. In this case,

$$\begin{aligned} \mathbb{E}(N_k(g)) &= \mathbb{E} \left(\sum_{j \geq 0} \chi_{\Delta_{k-1}}(\sigma^j x) g(j \cdot c_k) \right) \\ &= \sum_{j \geq 0} g(j \cdot c_k) \mathbb{E}(\chi_{\Delta_{k-1}} \circ \sigma^j) \\ &= \frac{m(\Delta_{k-1})}{c_k} \sum_{j \geq 0} c_k g(j \cdot c_k) \end{aligned}$$

As $c_k = e^{k\mathcal{P}_\Delta}$, and $e^{-k\mathcal{P}_\Delta} m(\Delta_{k-1}) \rightarrow \int w_\Delta dm$, we obtain

$$\mathbb{E}(N_k(g)) \xrightarrow{k \rightarrow \infty} \left(\int w_\Delta dm \right) \int_0^\infty g(y) dy.$$

In the case $n \geq 2$, as k diverges, there is a need to study the asymptotic behaviour of

$$\mathbb{E} \left(\prod_{s=1}^n (\chi_{\Delta_{k-1}} \circ \sigma^{j_s}) \right) = \int \left(\prod_{s=1}^n (\chi_{\Delta_{k-1}} \circ \sigma^{j_s}) \right) dm,$$

for $j_0 = 0 < j_1 < \dots < j_n$. In order to do this, we are able to show the following result.

Lemma 21. *The limit*

$$C_n = \lim_{k \rightarrow \infty} m(\Delta_{k-1})^{-1} \sum_{\substack{0=q_0 < q_1 < \dots < q_{n-1} \\ q_s - q_{s-1} \leq k}} \mathbb{E} \left(\prod_{s=0}^{n-1} \chi_{\Delta_{k-1}} \circ \sigma^{q_s} \right),$$

exists for every $n \geq 1$. Furthermore, for $n \geq 1$,

$$C_n = (e^{-\mathcal{P}_\Delta} - 1)^{-(n-1)}.$$

We then use the above constants to show the next preliminary result.

Proposition 22. *For each fixed n , the limit $\nu_n = \lim_{k \rightarrow \infty} \mathbb{E}(N_k(g)^n)$ exists, and it is the n -th derivative at the origin of the complex function*

$$F_g(z) = \exp \left\{ \sum_{j=1}^{\infty} C_j \int_0^{\infty} (e^{zg(y)} - 1)^j dy \right\},$$

which is an analytic function on a disc around the origin in \mathbb{C} .

Using the above Proposition, Theorem 20 follows by solving the equation

$$\exp \left\{ \sum_{j=1}^{\infty} C_j \int_0^{\infty} (e^{zg(y)} - 1)^j dy \right\} = \exp \left\{ \lambda \sum_{j=1}^{\infty} \pi_j \int_0^{\infty} (e^{zjg(y)} - 1) dy \right\}$$

in the constants λ and π_j , and showing that both functions define an analytic function in a neighbourhood of the origin in \mathbb{C} . Therefore the limiting process is a marked Poisson point process with constant parameter measure $\lambda\pi$. \square

We finish this section with the ingredients involved in the proof of Proposition 22. In order to prove that $\lim_{k \rightarrow \infty} \int N_k(g)^n dm$ exists, we use a sort of relativised decay of correlations of the equilibrium state m . It can be stated as

Lemma 23. *There exist $K > 0$ and $0 < \gamma < 1$ such that*

$$\left| \mathbb{E}(\chi_{\Delta_{k-1}} \cdot \chi_B \circ \sigma^{k+r-1}) - m(\Delta_{k-1})m(B) \right| \leq K\gamma^r e^{kP_{\Delta}} m(B),$$

for every $k \geq 1$, $r \geq 1$, and for every Borel set $B \subseteq \Omega$.

The above Lemma allows us to control the summations involving the expectation of the various products of characteristic functions, which appear as “clusters” of successive visits to Δ_{k-1} , when dealing with the quantities

$$\sum_{0 \leq j_1, \dots, 0 \leq j_p} \mathbb{E} \left(\prod_{s=1}^p (\chi_{\Delta_{k-1}} \circ \sigma^{j_s}) \right).$$

More details will be given in the final version of [CC2].

4.3. Final Remarks.

Here we discuss how the results of the previous section can be used in the examples given in the Preliminaries, and the various comments made after the motivating examples coming from different ergodic dynamical systems of Section 2.

First we note that the consideration of Δ being a subset of the symbol space \mathcal{A} is general enough to obtain all irreducible and aperiodic subsystems of finite type in (Ω, σ) . This is because of the higher-block representation of any subshift of finite type (see M. Boyle’s lectures in this volume), which could be represented as Ω_{Δ} for Δ being the choice of blocks which define the subshift of finite type.

The main result of the last Section, applied to products of shifts of finite type, has a consequence for the waiting time for matchings of two random sequences $x = (x_0, x_1, \dots)$ and $y = (y_0, y_1, \dots)$, when say, x is randomly chosen according to an equilibrium state m_1 , and y randomly chosen according to the equilibrium state m_2 . In this case, the scale we need to apply in order to study the asymptotic behaviour of these matching times is the exponential of the pressure on the restricted system defined by the diagonal in $\{1, \dots, \ell\}^2$.

We have already mentioned in (2) that, if φ_1 is the Hölder potential defining m_1 (and φ_2 defining m_2), then $m_1 \times m_2$ is the equilibrium state of the potential $\varphi_1 \otimes \varphi_2$. Denote the restriction of $\varphi_1 \otimes \varphi_2$ to the diagonal subsystem by $\varphi_{\Delta}(x, x) = \varphi_1(x) + \varphi_2(x)$. We

may think of $\varphi_\Delta(x, x) = \varphi_\Delta(x)$ as a function defined on (Ω, σ) . Therefore the pressure \mathcal{P}_Δ will be given by

$$\mathcal{P}_\Delta = \sup \left\{ h(\mu) + \int (\varphi_1 + \varphi_2) d\mu \right\},$$

where the supremum is taken over all shift-invariant probability measures μ on Ω . The latter is true because the the diagonal subsystem is canonically conjugate to (Ω, σ) . The above results prove our comments in Section 1, when we talked about the inter-occurrence times for matchings.

We also return to Hirata's type of result. Given a generic $x \in \Omega$ consider the cylinders $C[x_0, x_1, \dots, x_k]_0$ and study the visiting times to these cylinders asymptotically with k . This corresponds to defining $\Delta_i = \{x_i\}$ and writing

$$\Delta^{(k)} = \left\{ w \in \Omega : w_i \in \Delta_i, \forall i = 0, 1, \dots, k \right\}.$$

If x is not eventually periodic, then it can be shown, using a similar technique sketched in the last Section, that the visiting time process, rescaled by $m(C[x_0, x_1, \dots, x_k]_0)^{-1}$, converges in law to a Poisson point process with rate 1, which we mentioned earlier. Hopefully, all of these results are going to appear in the forthcoming paper [CC2].

4.4. Some Questions.

One can also consider the case when $\Delta \subsetneq \{1, \dots, \ell\}$ is chosen such that the subsystem Ω_Δ is not irreducible (or if it is irreducible, but maybe not aperiodic). It is expected that periodicity may imply no convergence of the rescaled process. Something like the example given in Figure 6 could also prove to be intriguing.

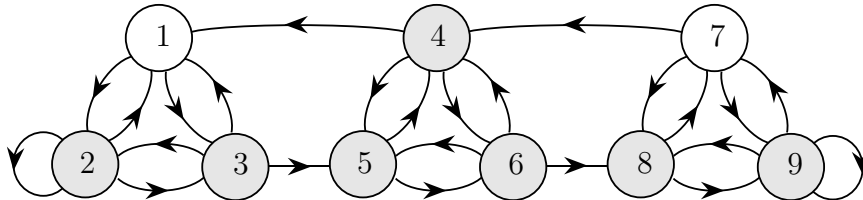


FIGURE 6. Example of a non-irreducible Ω_Δ .

Intuition may imply that asymptotically, as we approach Ω_Δ , a generic point x will prefer to see the irreducible subshift on the symbols $\{8, 9\}$, since the pair of subsystems, one determined by the symbols $\{2, 3\}$, and the other determined by $\{4, 5, 6\}$, have a drift in the bigger space to the right of Figure 6. Finding a scale to study the convergence in law of asymptotic visiting times in this case is not an easy task. (We know the answer in this particular case and it turns out that the exponential of the pressure, relative to the restriction of the potential on the subsystem determined by $\{8, 9\}$ is the correct scale.)

How about other subshifts? For instance, substitution systems like the Morse sequence (see B. Host's lectures in this volume)

$$a \mapsto ab, \quad \text{and} \quad b \mapsto ba.$$

The corresponding subshift of $\{a, b\}$ is defined as the orbit closure (under the shift) of the infinite word

$$w_\infty = (abbabaabbaababbabaababbaabbabaab \dots),$$

where w_∞ is obtained as the direct limit of a sequence of substitutions starting with the letter a , i.e. $a \mapsto ab \mapsto abba \mapsto abbabaab \mapsto \dots$. This system is uniquely ergodic, i.e. there exists a unique shift-invariant probability measure μ and

$$\mu(C[x_0, x_1, \dots, x_k]_0) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=0}^{n-1} \chi_C(\sigma^j w_\infty),$$

where $C = C[x_0, x_1, \dots, x_k]_0$. Therefore the measure of C is the relative frequency in which C appears in w_∞ . How the times to visit C distribute as $k \rightarrow \infty$ if $x = (x_0, x_1, \dots)$ is a generic point for μ ?

Due to some correspondence between certain substitution systems (for instance, the Fibonacci subshift) and irrational rotations of the circle, there may be a way to study these asymptotic processes using the technique in [CF]. We mentioned in Section 2.3.5 that for Lebesgue almost every rotation number the process of visits to a neighbourhood of a base point does not converge, if the process is rescaled by any scale. However, when the rotation number is the golden mean, then *there is convergence* and the limiting process is not Poisson and the inter-occurrence times are not asymptotically independent.

It seems to us that the Poisson asymptotic law is associated to Bernoulli systems. The consideration of rotations shows that these processes do not converge in general. Therefore, we are tempted to believe that there is a need of a higher order of mixing to get convergence of these processes. Since some substitution systems are weakly-mixing (although never strongly mixing), it would be interesting to find any analog of the discussions of our lectures in this context.

We should also mention that it is not clear what happens if we consider the same type of problems for irrational rotations of the n -torus. In [CF] we used renormalisation techniques to prove our results and these are not totally well developed for the n -torus.

Finally, we propose a general question. Let $F(t)$ be a distribution function of a positive random variable and consider the full shift on ℓ symbols, i.e. $(\{1, \dots, \ell\}, \sigma)$. Are there an ergodic shift-invariant measure μ on $\{1, \dots, \ell\}$ and a sequence of positive measure sets A_n with $\mu(A_n) \rightarrow 0$ as $n \rightarrow \infty$, such that

$$F_n(t) = \mu\{x: \mu(A_n) \tau_n(x) \leq t\} \xrightarrow{n \rightarrow \infty} F(t) ?$$

(where as usual $\tau_n(x)$ denotes $\inf\{s > 0 : \sigma^s x \in A_n\}$.)

ACKNOWLEDGEMENTS: The author would like to thank François Blanchard for the invitation to give these lectures. Thanks are also due to the organisers and participants of the Summer School for providing an excellent atmosphere during the meeting. The present notes are based on transparencies which were handwritten by the author for the School and copies were given to the participants. The author thanks Rodrigo Davila who did a superb job typing the contents of the transparencies. Rodrigo's source file was edited and modified later by the author to produce the present notes.

REFERENCES

- [AW] R. Adler & B. Weiss, *Similarity of automorphisms of the torus*, *Memoirs of the AMS* **98**, 1970.
- [Ald] D. Aldous, *Probability Approximations via the Poisson Clumping Heuristic*, Springer-Verlag, NY (1989).
- [AGW] R. Arratia, L. Gordon & M.S. Waterman, The Erdős-Rényi law in distribution, for coin tossing and sequence matching, *Ann. Statist.* **18** (1990), 539–570.

- [Bow] R. Bowen, *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, Springer Lect. Notes in Maths **470**, 1975.
- [BK] M. Bramson & S.A. Kalikow, Non-uniqueness in g -functions, *Israel J. Math.* **84** (1993), 153–160.
- [Bre] L. Breiman, *Probability theory*, Addison-Wesley Series in Statistics (1968).
- [Col] P. Collet, *Some ergodic properties of maps of the interval*, in Dynamical Systems & Frustrated Systems. R. Bamon, J.-M. Gambaudo & S. Martinez editors (to appear).
- [CC1] Z. Coelho & P. Collet, Limit law for the close approach of two trajectories of expanding maps of the circle, *Prob. Th. and rel. fields* **99** (1994), 237–250.
- [CC2] Z. Coelho & P. Collet, Asymptotic limit law for subsystems of shifts of finite type. In preparation.
- [CF] Z. Coelho & E. de Faria, Limit laws of entrance times for homeomorphisms of the circle, *Israel J. Math.* **93** (1996), 93–112.
- [CP] Z. Coelho & W. Parry, Central limit asymptotics for shifts of finite type, *Israel J. Math.* **69**, (1990), 235–249.
- [CQ] Z. Coelho & A. Quas, Criteria for \bar{d} -Continuity, *Trans. of Amer. Math. Soc.* **350**, (1998) 3257–3268.
- [CG] P. Collet & A. Galves, Asymptotic distribution of entrance times for expanding maps of the interval, *Dynamical Systems and Applications* 139–152, World Sci. Ser. Appl. Anal., **4**, World Sci. Publishing, 1995.
- [CMS] P. Collet, S. Martínez & B. Schmitt, The Pianigiani-Yorke measure for topological Markov chains, *Israel J. Math.* **97** (1997), 61–70.
- [Doe] W. Doeblin, Remarques sur la théorie métrique des fractions continues, *Compositio Math.* **7** (1940), 353–371.
- [Don] W.F. Donoghue Jr, *Monotone matrix functions and analytic continuation*, Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen Band **207**, Springer-Verlag (1974).
- [DV] D.J. Daley & D. Vere-Jones, *An introduction to the theory of point processes*, Springer Series in Statistics, Springer-Verlag (1988).
- [ER] P. Erdős & A. Rényi, On a new law of large numbers, *J. Analyse Math.* **23** (1970), 103–111.
- [GS] G.R. Grimmett & D.R. Stirzaker, *Probability and Random Processes*, Clarendon Press, Oxford (1992).
- [Hir] M. Hirata, Poisson law for Axiom A diffeomorphisms, *Erg. Th. & Dyn. Sys* **13** (1993), 533–556.
- [LLR] M.R. Leadbetter, G. Lindgreen & H. Rootzén, *Extremes and Related Properties of Random Sequences and Processes*, Springer-Verlag, NY (1983).
- [Nev] J. Neveu, *Processus pontuels*, in Springer Lecture Notes in Maths **598** (1976), 249–445.
- [Pit] B. Pitskel, Poisson limit law for Markov chains, *Erg. Th. & Dyn. Sys.* **11** (1991), 501–513.
- [Par] W. Parry, Intrinsic Markov chains, *Trans. of Amer. Math. Soc.* **112** (1964), 55–65.
- [PP] W. Parry & M. Pollicott, *Zeta functions and the periodic orbit structure of hyperbolic dynamical systems*, Astérisque (Soc. Math. de France) **187–188** (1990).
- [Rue] D. Ruelle, *Thermodynamic formalism*, Addison-Wesley, Reading, Mass., 1978.
- [Wal] P. Walters, *An introduction to ergodic theory*, Graduate Texts in Maths **79**, Springer Verlag, New York, 1982.

DEPARTAMENTO DE MATEMÁTICA APLICADA, FACULDADE DE CIÊNCIAS, UNIVERSIDADE DO PORTO, RUA DAS TAIPAS, 135, P-4050 PORTO, PORTUGAL

E-mail address: zaqueu@maths.warwick.ac.uk