Nonlinearity **22** (2009) 371–379 doi:10.1088/0951-7715/22/2/007

Return times in a process generated by a typical partition

Paulina Grzegorek[1] **and Michal Kupsa**[2]

[1] Institute of Mathematics and Computer Science, Wroclaw University of Technology, Wybrzeże Wyspiańskiego 27, 50–370 Wrocław, Poland
[2] Institute of Information Theory and Automation, Czech Academy of Science, Prague, Czech Republic

E-mail: paulina.grzegorek@pwr.wroc.pl and kupsa@utia.cas.cz

Received 13 June 2008, in final form 30 November 2008
Published 29 December 2008
Online at stacks.iop.org/Non/22/371

Recommended by L Bunimovich

Abstract
In Downarowicz and Lacroix (2006 Law of series) and Downarowicz *et al* (2007 ESAIM P&S), the authors show that for every ergodic aperiodic dynamical system, the process generated by a typical partition has the following property: the zero function is a pointwise limit, along a subsequence of lengths n_k of upper density 1 and with probabilities increasing to 1, of the distribution functions of the normalized (i.e. appropriately scaled) hitting times to cylinder sets of lengths n_k. Of course, this is the smallest possible limit distribution. We indicate two classes of systems where at least one more limit distribution coexists, and occurs with the same 'strength' (i.e. for every typical process, along a subsequence of lengths of upper density 1 and with probabilities increasing to 1): in α-mixing systems this is the exponential limit distribution (the largest possible in systems of positive entropy), and in adding machines this is the distribution $L(t) = \max\{0, \min\{1, t\}\}$ (the largest possible distribution for hitting times).

Mathematics Subject Classification: 37A25, 37A05, 60G10

1. Introduction

Since the late seventies, when Lempel and Ziv introduced an algorithm for data compression based on recurrence of cylinder sets, the statistics of return times in dynamical systems, especially for cylinder sets, has been intensively studied. The Ornstein–Weiss theorem [3] establishes that if μ is an ergodic measure and $\tau_n(x)$ denotes the first return time of x to its own cylinder of length n then

$$\frac{1}{n}\log \tau_n(x) \to h_\mu(T)$$

μ-almost surely, where h_μ denotes the entropy of μ. But the logarithmic expression in the above statement is insensitive to subexponential variations of $\tau_n(x)$. Thus, this theorem says nothing about the 'shape' of the distribution function of the return times. If we restrict our attention to one cylinder set B then the Kac theorem [4] assures that the expected value of the return time to B equals precisely $1/\mu(B)$, but still the *proportions* between the return times at various points of B remain unregulated. To capture them one has to study the distribution functions of the *normalized* return time to B, $\tilde{\tau}_B = \mu(B)\tau_B$, a random variable defined on the cylinder B, with expected value always equal to 1.

One of the natural questions asked in this context is whether the distribution functions of the normalized return times to cylinder sets converge as the length of the cylinders increases to infinity. Is coexistence of several limit distributions (achieved on different sets of cylinders or along different subsequences n_k of lengths) possible in an ergodic process? What limit distributions can appear in this role? Of course, we are only interested in 'observable' limit distributions, i.e. achieved on collections of n_k-cylinders whose joint measure does not decay to 0 as $n_k \to \infty$. The simple example of an independent process (the full shift) shows that without such restriction all imaginable limit distributions are possible in one system.

There is a simple approximative integral relation between the distribution of the normalized return time and that of the normalized *hitting time* (see next section for precise definitions; see [5] for the proof of the relation), which becomes strict as the lengths of cylinders grow to infinity. So, the results concerning limit distributions for return times have equivalent translations in terms of analogous distributions for the hitting times. Both approaches are mixed in the literature depending on the convenience of formulation.

Quite a long and impressive list of works on the subject of limit distributions demonstrates two major kinds of behaviour:

First, the exponential distribution $E(t) = \max\{0, 1 - \mathrm{e}^{-t}\}$ was proved to be a limit distribution of the hitting (and also return) times in Markov chains by Pitskel in [6], and then the same limit distribution was found in the classes of ϕ-mixing and exponentially α-mixing processes by Galves and Schmitt [7] and Abadi [8]. In the above classes, the fast correlation decay is a reason why these processes possess the limit distribution characteristic for i.i.d. processes. For a more complete list of papers on the exponential limit distribution of the return times see the survey [9]. The limit distribution E in these systems is unique for the process, i.e. it is achieved almost everywhere and along the full sequence n of lengths, leaving no room for any other (observable) limit distributions. Probably the most general result concerning limit distributions in all processes of positive entropy was obtained by Downarowicz and Lacroix [1]: all observable limit distributions of the hitting times to cylinders are bounded from above by the exponential distribution E.

A second group of results were obtained for zero-entropy processes. In this case the correlation between events in the past and in the future seems to cause that limit distributions of the return times are distributions of discrete finitely valued random variables. This behaviour is characteristic for periodic systems. Coelho and de Faria find such limit distributions in homeomorphisms of the unit circle [10], in particular in irrational rotations and Sturmian shifts (for circle maps intervals were considered instead of cylinders). Durand and Maass have obtained similar results in minimal low-complexity Cantor systems [11]. Nevertheless, also other examples of limit distributions of the return times were found in the class of zero-entropy processes [12–14]. In [14], it is even proved that every distribution of a non-negative random variable with expected value less than or equal to one can be obtained as a limit distribution of the return times in a process generated by a partition of a rank-one system along a sequence of cylinders. Unfortunately, all the results demonstrating the existence of the limit distributions in certain processes either deal with non-observable limits or observability is not proved. The only exception is a process constructed in [12], where the characteristic function $1_{[0,\infty)}$ is the unique limit distribution for the return times, i.e. piecewise linear distribution $L(t) = \max\{0, \min\{1, t\}\}$ is the unique distribution for the hitting times.

Now consider a measure preserving transformation on the standard probability space. This object gives rise to many processes, each generated by some finite partition of the space. Downarowicz and Lacroix asked about a limit distribution of the hitting times, which would be common to all 'typical' processes, i.e. processes generated by 'typical' partitions in a given ergodic system (here 'typical' means belonging to a residual set in the complete metric on the set of partitions into at most m elements). They proved [1] that for every non-periodic ergodic system of positive entropy, the zero function is a limit distribution of the hitting times to cylinders achieved along a subsequence of lengths of upper density 1 and with probabilities increasing to 1, in every typical process defined on this system. Recently the authors have extended their result to arbitrary aperiodic ergodic systems (also of entropy zero; [2]). The zero limit distribution for hitting times has an interesting interpretation: very long cylinders (of the selected lengths) are visited in 'clusters' (with relatively very short gaps inside each cluster) separated by huge long pauses. For this reason, results of this kind are given the nickname 'the law of series'.

In this paper we follow the point of view of Downarowicz and Lacroix and we focus on limit distributions appearing with probability increasing to 1 in all typical processes on a given dynamical system. Because the zero limit distribution is achieved only along a subsequence of lengths, there is still room for other limit distributions. Moreover, because several disjoint subsequences may have upper density one, the other limit distributions may be achieved along sequences of lengths with the same upper density as the zero limit distribution. In theorem 1, we prove that the exponential distribution E is another limit distribution (for both the return and hitting times) achieved in this manner in every typical process generated on any α-mixing system (i.e. in a system having an α-mixing generating partition). Recall that positive entropy and the Downarowicz–Lacroix theorem imply that E is the largest limit distribution function that could be expected in this case. In other words, in this class, in a typical process we find both extremities (appearing with equal strength): the smallest and largest possible limit distribution functions in positive entropy.

The above result is obtained with the help of another one, which has its own interest; we prove that for the α-mixing process itself (i.e. without perturbing the generating partition) the exponential distribution is the unique limit distribution (i.e. achieved along the full sequence of lengths; see theorem 4). This generalizes the aforementioned analogous results obtained by Abadi, Galves and Schmitt for ϕ-mixing and exponentially α-mixing processes.

In the second part of our paper, we perform a similar search for alternative typically occurring limit distributions in certain zero-entropy systems, namely, in adding machines. In theorem 2, we prove that in this case the piecewise linear distribution function L is such a limit distribution, again achieved along a sequence of lengths of upper density 1 and with probability increasing to 1. This is the largest distribution function admissible for the hitting time in any process (the corresponding distribution of return times is concentrated at 1 and corresponds to visits in the cylinder with approximately equal gaps of time). So, once again, here, in a typical process, we have the coexistence of two extremities: the smallest and the largest of all possible limit distributions.

2. Preliminaries and the main theorems

Let $(X, \mathcal{B}, \mu, T)$ be a dynamical system on a standard probability space. For $U \subset X$ with $\mu(U) > 0$, *Poincaré recurrence theorem* states that the variable

$$\tau_U(x) = \inf\{k \geqslant 1 : \ T^k x \in U\}$$

is μ_U-a.e. well defined, where μ_U is the conditional measure μ on U. The corresponding random variable (on U) is called the *return time*. In ergodic systems the same formula is well defined μ almost everywhere and the corresponding random variable (on X) is called the *hitting time*. Because, by the Kac theorem, the expected value of τ_U with respect to μ_U equals $1/\mu(U)$, the random variable $\mu(U)\tau_U$ has expected value 1. This variable will be called the *normalized return time*. By analogy, the same variable considered on X will be called the *normalized hitting time*.

We define

$$\tilde{F}_U(t) = \frac{1}{\mu(U)}\mu(\{x \in U : \ \mu(U) \cdot \tau_U(x) \leqslant t\})$$

and

$$F_U(t) = \mu(\{x \in X : \ \mu(U) \cdot \tau_U(x) \leqslant t\}).$$

Since there is an explicit relation between the distribution functions of the return and the hitting times, which can be deduced from the observation of the skyscraper above the set U (see also [5]), we will use either one, depending on which one is more convenient in a given context.

Before we state the main theorems, we introduce some notation. Let m denote a fixed integer. We consider a family $\mathfrak{P}$ of at most m-element partitions $\mathcal{P}$ of X. There is a natural pseudometric on $\mathfrak{P}$, called the *Rokhlin metric*

$$d(\mathcal{P}, \mathcal{R}) = 1 - \sup\left\{\sum_{P \in \mathcal{P}'} \mu(P \cap f(P))\right\},$$

where the supremum is taken over all one-to-one functions f with a domain $\mathcal{P}' \subseteq \mathcal{P}$ and range $f(\mathcal{P}') \subseteq \mathcal{R}$. This pseudometric identifies partitions modulo measure zero and it is a metric on the corresponding factor space (we will also denote it by $\mathfrak{P}$). Rokhlin metric endows $\mathfrak{P}$ with a topology of a Polish space.

There is also a lattice structure $(\mathfrak{P}, \prec)$ where $\prec$ is defined as

$$\mathcal{P} \prec \mathcal{R} \qquad \text{iff } \forall_{R \in \mathcal{R}} \exists_{P \in \mathcal{P}} \quad R \subset P \ (\text{mod } \mu).$$

We then say that $\mathcal{P}$ is coarser than $\mathcal{R}$. Every two partitions have a supremum $\mathcal{P} \vee \mathcal{R} = \{P \cap R : P \in \mathcal{P}, \ R \in \mathcal{R}\}$ which is called the *join* of partitions $\mathcal{P}$ and $\mathcal{R}$. We use the following notation

$$\mathcal{P}^n = \bigvee_{i=0}^{n-1} T^{-i}\mathcal{P}.$$

Elements of $\mathcal{P}^n$ are called cylinders of lengths n (with respect to the partition $\mathcal{P}$). For $x \in X$ by $\mathcal{P}^n(x)$ we understand a cylinder from $\mathcal{P}^n$ that contains a point x.

For a process $(X, \mathcal{B}, \mu, T, \mathcal{P})$, a right-continuous non-decreasing function $F : \mathbb{R} \to [0, 1]$ will be called a *substantial* limit distribution for the return, respectively, hitting times, if there exists a set $\mathcal{N}(\mathcal{P}) \subset \mathbb{N}$ of upper density one and families of cylinders $\mathcal{R}_n \subseteq \mathcal{P}^n$, $n \in \mathcal{N}(\mathcal{P})$, satisfying the following conditions:

- $(\mu(\bigcup \mathcal{R}_n))_{n\in\mathcal{N}(\mathcal{P})}$ converges to 1,
- for every sequence of cylinders $(B_n)_{n\in\mathcal{N}(\mathcal{P})}$, $B_n \in \mathcal{R}_n$, the distribution functions $\tilde{F}_{B_n}$, respectively, F_{B_n} converge weakly to F.

In the definition, we use the notion of weak convergence of distribution functions, which is defined in the standard way: the distribution functions $F_k : \mathbb{R} \to [0, 1]$, $k \in \mathbb{N}$, *converge weakly* to a non-decreasing right-continuous function $F : \mathbb{R} \to [0, 1]$ if $\lim_{k\to\infty} F_k(t) = F(t)$ for all continuity points $t \in \mathbb{R}$ of F. A useful fact about this convergence is that we can metrize it. Let us denote the set of all non-decreasing right-continuous functions from $\mathbb{R}$ to $[0, 1]$ by $\mathcal{F}$. Fix a Borel probability measure ν on $\mathbb{R}$ which is equivalent to the Lebesgue measure (the measures have the same Borel sets of measure 0). The metric is then generated by the $\mathcal{L}^1_\nu(\mathbb{R})$-norm

$$d_\nu(F, G) = \int_{-\infty}^{\infty} |F(t) - G(t)|\, \mathrm{d}\nu(t).$$

This metric is, of course, smaller than the uniform distance.

We note that the most commonly used metric to metrize weak convergence in the space of distribution functions is the Prokhorov metric. Unfortunately, this metric extended to $\mathcal{F}$ does not metrize weak convergence any more. The convergence in that metric is stronger than weak convergence because it requires pointwise convergence in both infinities, i.e. the limit must again be a distribution function. This is the reason why we use another metric, more appropriate for our purpose.

A partition $\mathcal{P}$ is α-mixing if the function $\alpha(l)$ defined below converges to 0 as l goes to infinity

$$\alpha(l) = \sup_{n\geqslant 1} \sup_{B\in\mathcal{P}^n} \sup_{C\in\mathcal{P}^\infty} |\mu_B(T^{-(n+l+1)}C) - \mu(C)|,$$

where $\mathcal{P}^\infty = \sigma(\bigcup_{n=1}^\infty \mathcal{P}^n)$ is the smallest σ-algebra containing all cylinders of all lengths and the supremum is taken over the sets B, such that $\mu(B) > 0$.

In order to state our main theorems in a more readable form, we introduce a notion of 'typicality' in the Polish space $\mathfrak{P}$. We say that a function $F \in \mathcal{F}$ is a substantial limit distribution for the return times in a process of system X *generated by a typical partition* $\mathcal{P} \in \mathfrak{P}$ if there exists a residual subset $\mathfrak{P}' \subseteq \mathfrak{P}$ such that, for every partition $\mathcal{P} \in \mathfrak{P}'$, F is a substantial limit distribution for return times in the process $(X, \mathcal{B}, \mu, T, \mathcal{P})$.

Theorem 1. *Let* $(X, \mathcal{B}, \mu, T)$ *possess a finite generating* α*-mixing partition. Then* $E(t) = \max(0, 1 - \mathrm{e}^{-t})$ *is a substantial limit distribution for the return times in a process generated by a typical partition.*

Theorem 2. *Let* $(X, \mathcal{B}, \mu, T)$ *be an adding machine. Then the function* $1_{[1,\infty)}$ *is a substantial limit distribution for the return times in a process generated by a typical partition.*

In the case of an adding machine the corresponding substantial limit distribution of the hitting times is equal to the piecewise linear function L.

3. The exponential limit distribution

In [8, 9], Abadi shows that E is a unique limit distribution for every ϕ-mixing and exponentially α-mixing process. We generalize this result for every α-mixing process.

First, let us mention that every α-mixing process has positive entropy. This is a direct consequence of the upper estimate of measures of cylinders given in [8, lemma 1]. The α-mixing property ensures ergodicity of the process.

The main part of the proof that E is the unique limit distribution for all α-mixing processes is contained in the proof of proposition 3 which states that the difference between the distribution of the return times and the distribution of the hitting times to a cylinder is small, when the length of the cylinder is big enough.

Proposition 3. *Let* $(X, \mathcal{A}, \mu, T, \mathcal{P})$ *be an* α*-mixing process. Then there exist families of cylinders* $\mathcal{R}_n \in \mathcal{P}^n$*,* $n \in \mathbb{N}$*, increasing in measure to* 1*, such that* $\sup_{B\in\mathcal{R}_n} ||F_B - \tilde{F}_B||_{\sup}$ *tends to zero, when n goes to infinity.*

Proof. We must prove that for every $\varepsilon > 0$, there is $n_0 \in \mathbb{N}$ such that for every $n \geqslant n_0$ there is a family of cylinders $\mathcal{R}_n \in \mathcal{P}^n$ which satisfies the conditions $\mu(\bigcup \mathcal{R}_n) \geqslant 1 - \varepsilon$ and $||F_B - \tilde{F}_B||_{\sup} < \varepsilon$, for every $B \in \mathcal{R}_n$.

Fix $\varepsilon > 0$. Let k be such that $\alpha(k) < \varepsilon/3$. For $n \in \mathbb{N}$ and $x \in X$ denote the element in $\mathcal{P}^n$ which contains x by $\mathcal{P}^n(x)$, and

$$X_n = \{x \in X;\ \tau_{\mathcal{P}^n(x)}(x) \leqslant n + k\},$$

$$\mathcal{R}'_n = \{B \in \mathcal{P}^n;\ 0 < (n+k)\mu(B) < \varepsilon/3\},$$

$$\mathcal{R}_n = \{B \in \mathcal{R}'_n;\ \mu_B(X_n) < \varepsilon/3\}.$$

By the Shannon–MacMillan–Breiman theorem (see, e.g., [15]), the measure of the union $\bigcup \mathcal{R}'_n$ tends to 1, whenever n increases to infinity. Thus, there is $n_1 \in \mathbb{N}$ such that for every $n \geqslant n_1$, the measure of $\bigcup \mathcal{R}'_n$ is bigger than $1 - \varepsilon/2$.

In addition, by the Ornstein–Weiss theorem [3], the measure of X_n tends to zero, when n increases to infinity. Hence, there is $n_0 \geqslant n_1$, such that for every $n \geqslant n_0$, $\mu(X_n) < \varepsilon^2/6$. Let $n \geqslant n_0$. Then

$$\begin{aligned}
\mu\left(\bigcup \mathcal{R}_n\right) &> 1 - \varepsilon/2 - \sum_{B\in\mathcal{R}'_n\setminus\mathcal{R}_n} \mu(B) = 1 - \varepsilon/2 - \sum_{B\in\mathcal{R}'_n\setminus\mathcal{R}_n} \frac{\mu(B\cap X_n)}{\mu_B(X_n)} \\
&\geqslant 1 - \varepsilon/2 - \sum_{B\in\mathcal{R}'_n\setminus\mathcal{R}_n} \frac{\mu(B\cap X_n)}{\varepsilon/3} \geqslant 1 - \varepsilon/2 - \sum_{B\in\mathcal{P}^n} \frac{\mu(B\cap X_n)}{\varepsilon/3} \\
&= 1 - \varepsilon/2 - \mu(X_n)\cdot 3/\varepsilon > 1 - \varepsilon.
\end{aligned}$$

It remains to prove that $||F_B - \tilde{F}_B||_{\sup}$ is less than ε, for every $B \in \mathcal{R}_n$. Let $B \in \mathcal{R}_n$. Since the functions F_B and $\tilde{F}_B$ are constant on every interval $[m\mu(B), (m+1)\mu(B))$, $m \in \mathbb{N}$, for every $t < (n+k+1)\mu(B)$ we have

$$\begin{aligned}
|F_B(t) - \tilde{F}_B(t)| &\leqslant |F_B(t)| + |\tilde{F}_B(t)| \leqslant F_B((n+k)\mu(B)) + \tilde{F}_B((n+k)\mu(B)) \\
&\leqslant \mu\left(\bigcup_{i=1}^{n+k} T^{-i}B\right) + \mu_B\{\tau_B \leqslant (n+k)\} \\
&\leqslant (n+k)\mu(B) + \mu_B\{\tau_B \leqslant (n+k)\} < \varepsilon/3 + \varepsilon/3 < \varepsilon.
\end{aligned}$$

If $(n+k+1)\mu(B) \leqslant t$, then there is a natural m such that $(n+k+m)\mu(B) \leqslant t < (n+k+m+1)\mu(B)$. Using the following equality

$$\{\tau_B \leqslant n+k+m\} = \{\tau_B \leqslant (n+k)\} \cup \bigcup_{j=n+k+1}^{n+k+m} T^{-j}B,$$

we get

$$\begin{aligned}
|F_B(t) - \tilde{F}_B(t)| &= |F_B((n+k+m)\mu(B)) - \tilde{F}_B((n+k+m)\mu(B))| \\
&\leqslant \left|\mu\{\tau_B \leqslant (n+k+m)\} - \mu\left(\bigcup_{j=n+k+1}^{n+k+m} T^{-j}B\right)\right| \\
&\quad + \left|\mu\left(\bigcup_{j=n+k+1}^{n+k+m} T^{-j}B\right) - \mu_B\left(\bigcup_{j=n+k+1}^{n+k+m} T^{-j}B\right)\right| \\
&\quad + \left|\mu_B\left(\bigcup_{j=n+k+1}^{n+k+m} T^{-j}B\right) - \mu_B\{\tau_B \leqslant (n+k+m)\}\right| \\
&\leqslant \mu\{\tau_B \leqslant (n+k)\} + \alpha(k) + \mu_B\{\tau_B \leqslant (n+k)\} \\
&\leqslant (n+k)\mu(B) + \alpha(k) + \mu_B\{\tau_B \leqslant (n+k)\} < 3 \cdot \varepsilon/3 = \varepsilon. \qquad \square
\end{aligned}$$

Theorem 4. *Let $(X, \mathcal{A}, \mu, T, \mathcal{P})$ be an α-mixing process. Then E is the unique limit distribution for the hitting times.*

Proof. We need to prove that for every $\varepsilon > 0$, there is $n_0 \in \mathbb{N}$ such that for every $n \geqslant n_0$ there is a family of cylinders $\mathcal{R}_n \in \mathcal{P}^n$ which satisfies the conditions $\mu(\bigcup \mathcal{R}_n) \geqslant 1 - \varepsilon$ and $d_v(F_B, E) < \varepsilon$, for every $B \in \mathcal{R}_n$.

Let us mention the paper [5]. There Haydn *et al* proved that in every aperiodic ergodic system, $||G_B - F_B||_{\sup} \leqslant \mu(B)$, where B is any set of positive measure and G_B is a real function defined as follows

$$G_B(t) = \begin{cases} 0, & \text{iff } t < 0, \\ \displaystyle\int_0^t 1 - \tilde{F}_B(s)\,\mathrm{d}s, & \text{iff } t \geqslant 0. \end{cases}$$

In other words, G_B is right differentiable everywhere on $\mathbb{R}$ and $(G_B)'_+ = 1 - \tilde{F}_B$. If we denote $f = G_B - E$ and use the equality $E' = 1 - E$, we get

$$f + f'_+ = G_B - E + 1 - \tilde{F}_B - (1 - E) = G_B - F_B + F_B - \tilde{F}_B$$

Since f is bounded, continuous, equal to zero for negative numbers and $\lim_{t\to\infty} f(t) = 0$, so f has a global maximum and minimum. Denote the point where the maximum is by x_0. There is a sequence $x_n \nearrow x_0$, such that $f'_+(x_n) \geqslant 0$ (if it was not, there would be an interval (a, x_0) on which the right derivative of f would be strictly negative and then x_0 would not be a maximum). Hence

$$\sup_{t\in\mathbb{R}}(f(t) + f'_+(t)) \geqslant \sup_{n\in\mathbb{N}}(f(x_n) + f'_+(x_n)) \geqslant \lim_{n\to\infty} f(x_n) = f(x_0) = \sup_{t\in\mathbb{R}} f(t).$$

Thus, for $B \in \mathcal{R}_n$ we have

$$\begin{aligned}
\|F_B - E||_{\sup} &\leqslant \|f\|_{\sup} + \|F_B - G_B\|_{\sup} \leqslant \|f + f'_+\|_{\sup} + \|F_B - G_B\|_{\sup} \\
&\leqslant 2\|G_B - F_B\|_{\sup} + \|F_B - \tilde{F}_B\|_{\sup} \leqslant 2\mu(B) + \|F_B - \tilde{F}_B\|_{\sup}.
\end{aligned}$$

The previous proposition concludes the proof. $\square$

Proof of theorem 1. We use the ideas from [1]. For a given distribution function F, $\varepsilon > 0$ and natural N, we define the following set of partitions

$$\mathfrak{B}_{N,\varepsilon}(F) = \Big\{\mathcal{P} \in \mathfrak{P} : \text{ for every } n \in [N, N^2],$$
$$\mu\Big(\bigcup\{B \in \mathcal{P}^n; d_\nu(F_B, F) < \varepsilon\}\Big) > 1 - \varepsilon\Big\}$$

It is easy to see that for every distribution function F, $\mathfrak{B}_{N,\varepsilon}(F)$ is open in $\mathfrak{P}$. Of course, the sum $\mathfrak{B}_\varepsilon(F) = \bigcup_{N\geqslant 1} \mathfrak{B}_{N,\varepsilon}(F)$ is also open.

The next step of the proof is to show that if a system $(X, \mathcal{B}, \mu, T)$ possesses a generating α-mixing partition $\mathcal{M}$, then the sets $\mathfrak{B}_\varepsilon(E)$ are dense, for all $\varepsilon > 0$. Because $\mathcal{M}^n$, $n \in \mathbb{N}$, increase to $\mathcal{B}$, every partition from $\mathfrak{P}$ can be arbitrarily well approximated by the m-element partition coarser than some $\mathcal{M}^n$. In other words, the partitions which are coarser than some $\mathcal{M}^n$ form a dense set in $\mathfrak{P}$. Every partition consisting of elements which are sums of sets of an α-mixing partition is also α-mixing. By the previous proposition, E is the unique limit distribution for the hitting times in such processes. Thus, such partitions belong to $\mathfrak{B}_\varepsilon(E)$, for every $\varepsilon > 0$. We get that $\mathfrak{B}_\varepsilon(E)$ are open dense sets and $\cap_{n\in\mathbb{N}}\mathfrak{B}_{\varepsilon_n}(E)$ is a G_δ dense set, for any sequence $(\varepsilon_n)_{n\in\mathbb{N}}$, decreasing to zero.

At the end, consider a sequence $(\varepsilon_n)_{n\in\mathbb{N}}$, decreasing to zero, and some partition $\mathcal{P} \in \cap_{n\in\mathbb{N}}\mathfrak{B}_{\varepsilon_n}(E)$. For a fixed $\varepsilon > 0$, we will prove that there are infinitely many $N \in \mathbb{N}$, such that $\mathcal{P} \in \mathfrak{B}_{N,\varepsilon}(E)$. Assume that there are only finitely many integers $N \in \mathbb{N}$, such that $\mathcal{P} \in \mathfrak{B}_{N,\varepsilon}(E)$. Denote the finite set of all such integers N by K. Let ε' be a minimum of $d_\nu(F_B, E)$ taken over all $B \in \mathcal{P}^n$, $n \in [N, N^2]$, $N \in K$. Since all these distances are positive (F_B is a distribution of a discrete random variable, whereas E is continuous), so ε' is. But from the definition, we get that $\mathcal{P}$ does not belong to any $\mathfrak{B}_{\varepsilon_n}(E)$, for $\varepsilon_n < \varepsilon'$. This is a contradiction. The fact that the union of infinitely many intervals $[N, N^2]$ has upper density 1 concludes the proof. □

4. Adding machines

Let $(p_i)_{i\in\mathbb{N}}$ be an increasing sequence of natural numbers such that $p_0 \geqslant 2$ and p_i divides p_{i+1}, for every $i \in \mathbb{N}$,

$$X = \left\{(x_i)_{i\in\mathbb{N}} \in \prod_{i=0}^{\infty}\{0, 1, \ldots, p_i - 1\}; \quad x_{i+1} = x_i \bmod p_i, \text{ for every } i \in \mathbb{N}\right\}.$$

This set becomes an Abelian group under the operation $(x + y)_i = x_i + y_i \bmod p_i$, $x, y \in X$. The operation is continuous with respect to the product of discrete topologies on sets $\{0, 1, \ldots, p_i - 1\}$, $i \in \mathbb{N}$. In this setting, X is a topological Abelian group. We denote the family of Borel subsets of X by $\mathcal{B}$ and Haar measure on X by μ_H. Haar measure is, of course, invariant under the translation $L : X \to X$, defined by the equation $L(x) = x + (1, 1, \ldots)$. Since the orbit of $(1, 1, \ldots)$ under the translation L is dense in X, the dynamical system $(X, \mathcal{B}, \mu_H, L)$ is ergodic (see [16]). The system is called an adding machine and $(p_i)_{i\in\mathbb{N}}$ is its periodic structure.

For $i \in \mathbb{N}$, let us define the partition $\mathcal{P}_i$ of X as follows

$$\mathcal{P}_i = \{P_{i,s}; s \in \{0, 1, \ldots, p_i - 1\}\}; \qquad P_{i,s} = \{(x_j)_{j\in\mathbb{N}} \in X;\ x_i = s\}.$$

By the definition, $\mathcal{P}_{i+1}$ is finer than $\mathcal{P}_i$, for every $i \in \mathbb{N}$. Moreover, union of all these partitions generates the σ-field $\mathcal{B}$. Thus, every partition can be approximated in the Rokhlin metric by a partition coarser than some $\mathcal{P}_i$.

Proof of theorem 2. The proof is similar to the proof of theorem 1. Let $(X, \mathcal{B}, \mu_H, L)$ be an adding machine with a periodic structure $(p_i)_{i\in\mathbb{N}}$. Let $\mathcal{R}$ be an m-element partition of X, which is coarser than some $\mathcal{P}_i$. In other words, every set in partition $\mathcal{R}$ is a union of sets $P_{i,s}$, $0 \leqslant s < p_i$. By the definition, $L^{-p_i}(P_{i,s}) = P_{i,s}$ for every $0 \leqslant s < p_i$. Hence, $L^{-p_i}(\mathcal{R}) = \mathcal{R}$ and $\mathcal{R}^n = \mathcal{R}^{p_i}$, for every $n \geqslant p_i$. Let $n \geqslant p_i$. Since the adding machine is ergodic, $\mathcal{R}^n$ can be written as a cycle of iterated images

$$\mathcal{R}^n = \{L^i(B); 1 \leqslant i \leqslant \#\mathcal{R}^n\}, \qquad \text{for any } B \in \mathcal{R}^n.$$

Thus, the distribution for the return times to any $B \in \mathcal{R}^n$ is equal to $1_{[1,\infty)}$ and $\mathcal{R} \in \mathfrak{B}_\varepsilon(1_{[1,\infty)})$. Since the partitions $\mathcal{P}_i$, $i \in \mathbb{N}$ increase to $\mathcal{B}$, partitions $\mathcal{R}$'s are dense in $\mathfrak{P}$ and so $\mathfrak{B}_\varepsilon(1_{[1,\infty)})$ is, for every $\epsilon > 0$.

Let us consider a residual set $\bigcap_{n\in\mathbb{N}}\mathfrak{B}_{\varepsilon_n}(1_{[1,\infty)})$, where $(\varepsilon_n)_{n=1}^\infty$ is an arbitrary sequence decreasing to zero. A partition $\mathcal{P}$ from this set generates either a periodic or an aperiodic process. In the periodic case, there is some n_0, such that for every $n \geqslant n_0$, $\mathcal{P}^n$ can be written as a cycle of iterated images of a set $B \in \mathcal{P}^n$. Therefore, $\mathcal{P} \in \mathfrak{B}_{N,\varepsilon}(1_{[1,\infty)})$ for every $\varepsilon > 0$ and $N \geqslant n_0$. Let $\mathcal{P}$ generate an aperiodic process. Then F_B differs from $1_{[1,\infty)}$, for every cylinder B, but in a similar way as in the proof of theorem 1 we obtain that $\mathcal{P}$ belongs to $\mathfrak{B}_{N_n,\varepsilon_n}(1_{[1,\infty)})$ for some increasing sequence $(N_n)_{n=1}^\infty$ of natural numbers and sequence $(\varepsilon_n)_{n=1}^\infty$ of positive real numbers decreasing to zero. $\square$

Acknowledgment

The first author's research was supported by MNiSW Grant N N201 387434. The second author's research was supported by GA ASCR under the grant KJB100750901.

References

[1] Downarowicz T and Lacroix Y 2006 Law of series *Preprint* arXiv:0810.4504v1
[2] Downarowicz T, Lacroix Y and Leandri D 2007 Spontaneous clustering in theoretical and some empirical stochastic processes *ESAIM P&S* submitted (arXiv:0810.4509v1)
[3] Ornstein D S and Weiss B 1993 Entropy and data compression *IEEE Trans. Inform. Theory* **39** 78–83
[4] Kac M 1947 On the notion of recurrence in discrete stochastic processes *Bull. Am. Math. Soc.* **53** 1002–10
[5] Haydn N, Lacroix Y and Vaienti S 2005 Hitting and return times in ergodic dynamical systems *Ann. Probab.* **33** 2043–50
[6] Pitskel B 1991 Poisson limit law for Markov chains *Ergodic Theory Dyn. Syst.* **11** 501–13
[7] Galves A and Schmitt B 1997 Inequalities for hitting times in mixing dynamical systems *Random Comput. Dyn.* **5** 337–47
[8] Abadi M 2001 Exponential approximation for hitting times in mixing processes *Math. Phys. Electron. J.* **7** 2
[9] Abadi M 2003 Statistics and error terms of occurrence times in mixing processes *Preprint*
[10] Coelho Z and de Faria E 1996 Limit laws of entrance times for homeomorphisms of the circle *Israel J. Math.* **93** 93–112
[11] Durand F and Maass A 2001 Limit laws of entrance times for low-complexity Cantor minimal systems *Nonlinearity* **14** 683–700
[12] Chaumoitre V and Kupsa M 2005 Asympotics for return times of rank-one systems *Stochastics Dyn.* **5** 65–73
[13] Lacroix Y 2002 Possible limit laws for entrance times of an ergodic aperiodic dynamical system *Israel J. Math.* **132** 253–64
[14] Chaumoitre V and Kupsa M 2006 k-limit laws of return and hitting times *Discrete Contin. Dyn. Syst.* A **15** 73–86
[15] Petersen K 1983 *Ergodic Theory* (*Cambridge Studies in Advanced Mathematics* vol 2) (Cambridge: Cambridge University Press)
[16] Walters P 1982 *An Introduction to Ergodic Theory* (*Graduate Texts in Mathematics* vol 79) (New York: Springer) p 30