# Periodicity distribution of cylinder sets in a product measure


## Julia Codas & Miguel Abadi (Advisor)

Universidade de São Paulo

{codas,leugim}@ime.usp.br

Instituto de Matemática Pura e Aplicada


## Introduction

We consider a dynamical system $(\Omega, \mathcal{F}, \mu, f)$, with a finite or countable generating partition and a stationary ergodic product measure $\mu$. We define the first hitting time $T_{\mathcal{C}}(x)$ of a point $x$ in $\Omega$ to a cylinder set $\mathcal{C}$ as the smallest positive integer $n$ such that $f^n x \in \mathcal{C}$, or infinite if the orbit of $x$ never enters in $\mathcal{C}$. This is a classical well studied quantity. Recent results show that its statistical properties depends on a structural quantity called the first possible return of $\mathcal{C}$, which depends only on $\mathcal{C}$ and not on $x$, and is defined as $\tau_n(\mathcal{C}) = \inf\{T_{\mathcal{C}}(x) : x \in \mathcal{C}\}$. In this work, we present results concerning the distribution function of this quantity $\tau_n(\mathcal{C})$.

## First possible return of a cylinder

Let $\mathcal{A}$ be a finite or countably infinte generating partition of $\Omega$. The elements of the $n$th join $\mathcal{A}^n \doteq \bigvee_{j=0}^{n-1} f^{-1}\mathcal{A}$, $n = 1, 2, \cdots$ are called $n$-cylinders. Since $\mathcal{A}$ is generating, the atoms of $\mathcal{A}^{\mathbb{N}}$ are single points of $\Omega$, that is, the points $y$ in $\Omega$ are of the form $y = (y_0 y_1 y_2 \cdots)$, where $y_j \in \mathcal{A}$. Note that a $n$-cylinder $\mathcal{C} \in \mathcal{A}^n$ can be uniquely identified with a finite sequence $x_0^{n-1} \doteq (x_0, \cdots, x_{n-1})$ of elements of $\mathcal{A}$ such that $\mathcal{C} = \{y \in \Omega : y_0 \cdots y_{n-1} = x_0 \cdots x_{n-1}\}$. In this case, we say that the $n$-cylinder $\mathcal{C} \in \mathcal{A}^n$ is defined by the finite sequence $x_0^{n-1}$.

**Definition.** *For a given $n$-cylinder $\mathcal{C} \in \mathcal{A}^n$ defined by $x_0^{n-1}$, we define the hitting time of a point $y \in \Omega$ into $\mathcal{C}$ as*

$$T_{\mathcal{C}}(y) = \inf\{j \geqslant 1 : y_j \cdots y_{j+n-1} = x_0 \cdots x_{n-1}\}$$

*(and infinite otherwise).*

**Definition.** *The first possible return of the $n$-cylinder $\mathcal{C}$ is defined as the infimum of the hitting times $T_{\mathcal{C}}(y)$ over all $y \in \mathcal{C}$. That is,*

$$\tau_n(\mathcal{C}) = \inf\{T_{\mathcal{C}}(y) : y \in \mathcal{C}\}.$$

**Example.** *Consider the 6-cylinder $\mathcal{C}$ defined by $x_0^5 = (1, 0, 1, 0, 1, 0)$. The first possible return of $\mathcal{C}$ is the number of shifts necessary to occur the first overlaping of the sequence $x_0^5$ with a translated copy of itself:*

```
            1 0 1 0 1 0
1st shift     1 0 1 0 1 0
2nd shift       1 0 1 0 1 0
```

*Thus we have $\tau_n(\mathcal{C}) = 2$.*

It was proved in [2] and [4] that, asymptotically, $\tau_n$ grows with the same velocity as $n$. Thus to study the behavior of the distribution of the function $\tau_n$ as $n$ diverges, we need to reescale $\tau_n$ as $S_n = n - \tau_n$. Since we consider a stationary product measure, all cylinders have positive measure. Therefore for a given $n$-cylinder $\mathcal{C} \in \mathcal{A}^n$ defined by $x_0^{n-1}$, $S_n(\mathcal{C})$ represents the size of the overlap between two copies of $x_0^{n-1}$, where one of them is shifted $\tau_n(\mathcal{C})$ times.

**Example.** *Consider again the 6-cylinder $\mathcal{C}$ defined by $x_0^5 = (1, 0, 1, 0, 1, 0)$. Then we have $S_6(\mathcal{C}) = 4$:*

$$\begin{array}{cccccccc} 1 & 0 & 1 & 0 & 1 & 0 & & \\ & & 1 & 0 & 1 & 0 & 1 & 0 \end{array}$$

$$\underbrace{1\ 0}_{\tau_6(\mathcal{C})}\ \underbrace{1\ 0\ 1\ 0}_{S_6(\mathcal{C})}$$

## Auxiliary sets

**Definition.** *Let $j$ and $n$ be positive integers such that $j < n$. The set $B_n(j)$ is defined as: if a $n$-cylinder $\mathcal{C} \in \mathcal{A}^n$, defined by the sequence $x_0^{n-1}$, is in $B_n(j)$, then $x_0^{n-1}$ can be written in the following form*

$$x_0^{n-1} = (\underbrace{\underbrace{x_0, \cdots, x_{j-1}}_{j}, \cdots, x_0, \cdots, x_{j-1}}_{\lfloor n/j \rfloor}, x_0, \cdots, x_{r-1}),$$

*with $n = j\lfloor n/j \rfloor + r$, $0 \leqslant r < j$. If $\lfloor n/j \rfloor = n/j$, then $r = 0$ and we have that $x_0^{r-1}$ is the empty set.*

The importance of the $B_n(j)$ sets comes from the following equality:

$$\{\mathcal{C} \in \mathcal{A}^n : S_n(\mathcal{C}) > k\} = \bigcup_{j=\lceil \frac{n-k}{2} \rceil}^{n-k} B_n(j).$$

Then

$$\mu(\{S_n > k\}) = \mu\left(\bigcup_{j=\lceil \frac{n-k}{2} \rceil}^{n-k} B_n(j)\right).$$

## Periodicity distribution

To study the distribution of the random variable $S_n$, we can use the inclusion-exclusion formula, that is,

$$\mu\left(\bigcup_{j=\lceil \frac{n-k}{2} \rceil}^{n-k} B_n(j)\right) = \sum_{r=1}^{\lfloor \frac{n-k}{2} \rfloor + 1} (-1)^{r-1} \sum_{J \subseteq \mathcal{I} : |J| = r} \mu\left(\bigcap_{j \in J} B_n(j)\right),$$

where $\mathcal{I} = \{\lceil \frac{n-k}{2} \rceil, \lceil \frac{n-k}{2} \rceil + 1, \cdots, n-k\}$.

**Notation:** $m_k^n \doteq \sum_{\mathcal{C} \in \mathcal{A}^n} \mu(\mathcal{C})^k$

**Theorem.** *Let $n$, $k$ be positive integers such that $n \geqslant 4k$, and let $\mu$ be a stationary product measure. Then*

$$\sum_{j=\lceil \frac{n-k}{2} \rceil}^{n-k} \mu(B_n(j)) = \frac{m_2^k}{1 - m_2} + \epsilon(n, k),$$

*where $\epsilon(n, k)$ is a function that depends on $n$ and $k$, and such that*

$$\epsilon(n, k) \to 0 \text{ when } n \to \infty.$$

**Corollary.** *Let $\mu$ be a stationary product measure. Fixed $n$ and $k$ positive integers such that $n \geqslant 4k$. Let $\delta$ be a real number such that $\frac{1}{2} < \delta \leqslant \frac{3}{4}$. Then*

$$\sum_{j=\lceil \frac{n-k}{2} \rceil}^{\lceil \delta n \rceil} \mu(B_n(j)) = \lambda(n, k),$$

*where $\lambda(n, k)$ is a function that depends on $n$ and $k$, and such that*

$$\lambda(n, k) \to 0 \text{ when } n \to \infty.$$

**Proposition.** *Let $\mu$ be a stationary product measure. For every fixed $r \geqslant 2$, the respective term of the inclusion-exclusion formula*

$$\sum_{J \subseteq \mathcal{I} : |J| = r} \mu\left(\bigcap_{j \in J} B_n(j)\right),$$

*has a lower bound strictly bigger than zero that does not depend on $n$.*

### References


[1] M. Abadi and R. Lambert. The distribution of the short-return function. *Nonlinearity*, 26(5):1143–1162, 2013.

[2] S. Troubetzkoy B. Saussol and S. Vaienti. Recurrence, dimensions, and Lyapunov exponents. *J. Statist. Phys.*, 106(3-4):623–634, 2002.

[3] S. Gallo M. Abadi and E.A. Rada Mora. The shortest possible return time of $\beta$-mixing processes. *IEEE Trans. Inform. Theory*, 64(7):4895–4906, 2018.

[4] J.R. Chazottes V. Afraimovich and B. Saussol. Pointwise dimensions for Poincare recurrences associated with maps and special flows. *Discrete Contin. Dyn. Syst.*, 9(2):263–280, 2003.



### Acknowledgements

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.