QUANTITATIVE COVERING AND EQUIDISTRIBUTION

NATALIA JURGA

Abstract. Covering (density of orbits) and equidistribution are fundamental properties of an ergodic dynamical system. In this expository note, we investigate the quantitative aspects of these properties for a simple model. In particular we investigate how much of a single orbit we must observe before we can expect it to reach a certain density in the state space and for it to be almost equidistributed, in a certain sense. We explore more broadly how these questions regarding the statistics of individual dynamical orbits are closely related to the multifractal properties of the underlying ergodic measure.

1. Stopping times: local statistical properties in dynamical systems

We begin this note by discussing an emerging research direction in the study of the statistical properties of dynamical systems. We use the term dynamical systems in the broad modern sense, to include maps of metric spaces, random walks, iterated function systems, (semi-)group actions etc, all of which can be studied from a dynamical perspective.

The *global* statistical properties of dynamical systems is a well-established and thriving topic of study across the field. This includes the study of mixing rates (how fast the system forgets its initial state, on average) and statistical limit theorems (how time averages distribute in the long run, across the entire system). The techniques involved in studying these kinds of properties are usually purely analytical and probabilistic, and can often be reduced to studying the spectral properties of an appropriate transition operator.

This note will be concerned with *local* statistical properties, that is, the statistical properties of *individual trajectories* of the system. Such properties can be investigated from the point of view of *stopping times*. A stopping time is a function (on the state space or space of possible orbits) which records how long it takes for the corresponding orbit to meet a predetermined condition. Examples of stopping times are:

- hitting times (the first time the dynamics accesses a certain region of the phase space),
- cover times (the first time the dynamics covers the phase space to some resolution),
- blanket times (the first time the empirical distribution along an orbit resembles the underlying stationary or invariant measure in a particular sense) etc.

For systems with discrete state spaces e.g. random walks on graphs, finite state Markov chains etc, stopping times have been thouroughly investigated in the literature, and often require both analytical and combinatorial techniques. Unlike global statistical properties, stopping times have been shown to be a lot more sensitive to geometry. For instance in the setting of random walks on graphs, stopping times can vary wildly depending on the geometry of the graph (such as bottlenecks, symmetries and poorly connected parts of the graph), whereas global statistics are more robust to phenomena such as rare long paths and capture a smoother average behaviour.

Now let us consider the setting of dynamical systems where the orbits belong to a continuous state space. In this setting the hitting time has been intensely studied, largely due to its connections with inducing/renormalisation techniques which were developed to study parabolic/slowly-chaotic systems. However, more general stopping times had not received any attention until [JM, BJK, JT]. From our initial investigations it seems that in this setting stopping times also exhibit more sensitivity to the geometry, in particular the geometric measure theoretic properties of the system. However the continuity of the state space makes this sensitivity more complex and subtle to characterise than the discrete case. Therefore a key aim of this program is to explore the interplay between the local statistical properties of dynamical systems (such as stopping times) and the multifractal properties of the invariant/stationary measure.

In this expository note we will estimate the cover and blanket times for a simple system in which the arguments can be presented in a very streamlined way. This note contains an estimate for the cover time for this example, which is a simplified version of the arguments in [JM]. The second result is an estimate on the blanket time for this system. This is original, but due to the simplicity of the system we will show that the result can be reduced clearly and rather effortlessly to an analogous result for Markov chains. Although the very good regularity of the measure in our system will mean that few ideas from multifractal analysis will be required, our aim is to illuminate places where a future more general theory would benefit from a more refined application of multifractal analysis. We will do this through discussion and open questions. Our intended audience for this note is a fractal geometer, perhaps with some expertise in multifractal analysis, who would like a straightforward entry to the basic combinatorial and probablistic techniques which are involved in this line of research.

In §2 we introduce the system which we will restrict our attention to, introduce the cover and blanket times formally in this context and state our main results. We also discuss more general versions of these results in the literature and frame them from the multifractal perspective. In §3 we introduce symbolic dynamics and appropriate 'symbolic stopping times' and demonstrate their relationship to the stopping times for our system. In §4 we use the symbolic dynamics to model an approximation of our system at any scale by an appropriate Markov chain, and connect its stopping times to the symbolic stopping times. §5 and 6 contain the proofs of our main results. Finally in §7 we discuss possible future directions of study.

1.1. **Notation.** We write $f(x) \asymp g(x)$ if there exists a positive finite constant C such that $C^{-1}g(x) \leq f(x) \leq Cg(x)$ for all x in the domain, which will be clear from the context. We use big O notation and write $f(x) = O(g(x))$ if if there exists a positive finite constant C such that $f(x) \leq Cg(x)$ for all x in the domain.

2. Quantitative covering and equidistribution for the chaos game

The study of stopping times is often grounded in a computational/algorithmic perspective: "How many steps does my algorithm require to meet this condition?". In fact this perspective is inherent to the particular system we would like to study, since the 'dynamics' will arise via an algorithm known as the Chaos Game.

Let $e_1 = (0,0), e_2 = (1,0), e_3 = (\frac{1}{2}, \frac{\sqrt{3}}{2})$ and consider the two-dimensional simplex

$$\Delta = \{\alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3 : \alpha_i \geq 0, \alpha_1 + \alpha_2 + \alpha_3 = 1\} \subset \mathbb{R}^2.$$

Consider the iterated function system (IFS) $\{f_i : \Delta \to \Delta\}_{i=1}^3$ where the f_i are the similarity mappings

$$f_1(x) = \frac{1}{2}x$$
$$f_2(x) = \frac{1}{2}x + \left(\frac{1}{2}, 0\right)$$
$$f_3(x) = \frac{1}{2}x + \left(\frac{1}{4}, \frac{\sqrt{3}}{4}\right).$$

Note that e_i is the unique fixed point of f_i. By Hutchinson's Theorem [H], the IFS has a unique attractor, i.e. there is a unique, non-empty compact set $F \subset \Delta$ such that $F = \bigcup_{i=1}^3 f_i(F)$. The attractor F is the well-known Sierpinski triangle which is a classical example of a self-similar set.

In [Ba] Barnsley introduced the Chaos Game, which is an algorithm for constructing attractors of iterated function systems. In the case of the Sierpinski triangle, this algorithm takes a particularly neat form.

Fix any initial point $x_0 \in F$. Then pick some index $i_1 \in \{1,2,3\}$ uniformly at random. Let x_1 be the point which lies halfway between x_0 and e_{i_1}. Next, pick $i_2 \in \{1,2,3\}$ uniformly at random, independently of the first choice. Let x_2 be the point which lies halfway between x_1 and e_{i_2}. We continue to generate a random sequence of points $(x_n)_{n=1}^\infty$ in this way. The reader may check that this random sequence is contained in F and that for any $n \in \mathbb{N}$, $x_n = f_{i_n \dots i_1}(x_0) := f_{i_n} \circ \cdots \circ f_{i_1}(x_0)$. It will be notationally convenient to denote $f_{i_0} = \mathrm{id}$.

Let $B(y,r)$ denote an open ball in $\mathbb{R}^2$. Let $s = \dim F = \frac{\log 3}{\log 2}$ and $\mathcal{H}^s$ denote s-dimensional Hausdorff measure [F], restricted to F and normalised so that $\mathcal{H}^s(F) = 1$. Barnsley showed that almost surely:

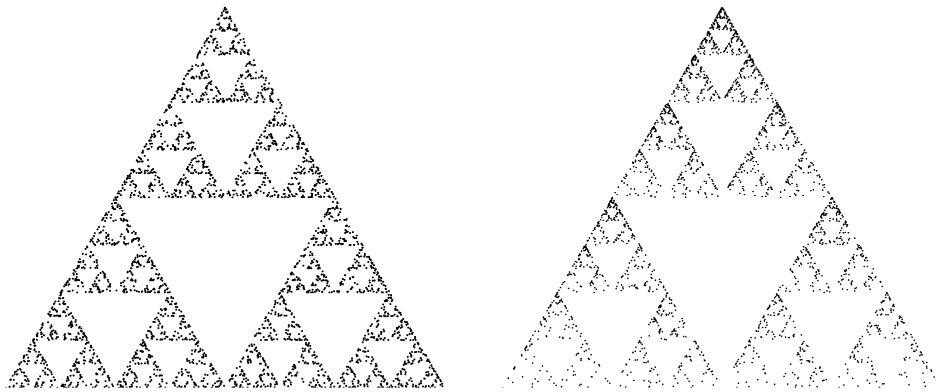
(A) the random sequence $(x_n)_{n\in\mathbb{N}}$ is dense in F i.e. $\overline{\{x_n\}_{n\in\mathbb{N}}} = F$
(B) the random sequence $(x_n)_{n\in\mathbb{N}}$ equidistributes to the normalised s-dimensional Hausdorff measure: for all $y \in F$, $r > 0$,

$$\lim_{N\to\infty} \frac{\#\{0 \leq n \leq N : x_n \in B(y,r)\}}{N+1} \to \mathcal{H}^s(B(y,r))). \tag{1}$$

In this note, we will investigate quantitative versions of statements (A) and (B). To this end we introduce two stopping times (non-negative integer valued random variables which record the first time that a prescribed event occurs). The first of these stopping times records at what time the sequence $(x_n)_{n=0}^\infty$ reaches a certain density in F. We will say that a set E is r-dense in F if for all $x \in F$ there exists $y \in E$ such that $|x - y| < r$.

Definition 2.1 (Cover time)**.** Given $x_0 \in F$ and $r \in (0,1)$ define the cover time at scale r to be the random variable $\tau_{x_0}^r$ given by

$$\tau_{x_0}^r = \inf\{n \geq 0 : \{x_0, \dots, x_n\} \text{ is } r\text{-dense in } F\}.$$



(I) Random output of Chaos Game with uniform weights.

(II) Random output of Chaos Game adapted to non-uniform weights. Here the probability to choose the map f_1, f_2, f_3 was $\frac{1}{4}, \frac{1}{4}, \frac{1}{2}$ respectively.

FIGURE 1. Each of these diagrams shows a randomly-generated sequence $(x_n)_{n=0}^N$ starting at $x_0 = 0$ and terminating when every point of the Sierpinski triangle has been approached to within distance $r = 2^{-6}$.

Note that $\tau_{x_0}^r$ can be understood as a function on the space of sequences $\Sigma = \{(i_1 i_2 \dots) : i_j \in \{1,2,3\}\}$ since

$$\tau_{x_0}^r(\mathtt{i}) = \inf\{n \geq 0 : \{x_0, f_{i_1}(x_0), \dots, f_{i_n \dots i_1}(x_0)\} \text{ is } r\text{-dense in } F\}.$$

This can be thought of as the first time the Chaos Game algorithm has obtained a representation of the attractor to a prescribed resolution.

The second stopping time records at what time the sequence $(x_n)_{n=0}^\infty$ is almost equidistributed, in a certain sense.

Definition 2.2 (Blanket time)**.** Given $x_0 \in F$ and $\epsilon, r \in (0,1)$ define the ϵ-approximate blanket time at scale r to be the random variable $\tau_{x_0}^{r,\epsilon}$ given by

$$\tau_{x_0}^{r,\epsilon} = \inf\left\{N \geq 0 : \forall y \in F, \frac{\#\{0 \leq n \leq N : x_n \in B(y,r)\}}{N+1} \geq (1-\epsilon)\mathcal{H}^s(B(y,r))\right\}.$$

Note that $\tau_{x_0}^{r,\epsilon}$ can be viewed as a function $\tau_{x_0}^{r,\epsilon} : \Sigma \to \mathbb{N} \cup \{\infty\}$ given by

$$\tau_{x_0}^{r,\epsilon}(\mathtt{i}) = \inf\left\{N \geq 0 : \forall y \in F, \frac{1}{N+1}\sum_{k=0}^{N} \mathbf{1}_{B(y,r)}(f_{i_k \dots i_0}(x_0)) \geq (1-\epsilon)\mathcal{H}^s(B(y,r))\right\}.$$

The heuristic behind this stopping time is that we don't only want the 'resolution' of the output of the Chaos Game to be correct, but we also want the distribution to look 'roughly uniform'.

The first basic and natural questions to ask about these stopping times are: what can be said asymptotically about the almost sure or expected behaviour of $\tau_{x_0}^r$ and $\tau_{x_0}^{r,\epsilon}$? In this note we focus on the expectation.

2.1. **Cover time estimates.**

We will prove the following estimate for the cover time:

Theorem 2.3. *There exists $C < \infty$ such that for any $x_0 \in F$ and $r > 0$,*

$$\frac{1}{C}\left(\frac{1}{r}\right)^{\frac{\log 3}{\log 2}} \log\left(\frac{1}{r}\right) \leq \mathbb{E}\tau_{x_0}^r \leq C\left(\frac{1}{r}\right)^{\frac{\log 3}{\log 2}} \log\left(\frac{1}{r}\right).$$

2.1.1. *Discussion.* More generally, one could consider an IFS $\{S_i : \mathbb{R}^d \to \mathbb{R}^d\}_{i=1}^N$ of strict contractions and let β be an ergodic measure on the symbolic space $\{1, \ldots, N\}^{\mathbb{N}}$. Let F be the attractor of this IFS and, choosing an initial point $x_0 \in F$, consider the random sequence of points $(x_n)_{n\in\mathbb{N}}$ generated by choosing $x_n = S_{i_n \ldots i_1}(x_0)$ with probability $\beta([i_1 \ldots i_n])$. We call this the generalised Chaos Game. Analogous statements to (A) and (B) can be shown to hold in this more general setting by letting $\mathcal{H}^s$ be replaced by the measure μ supported on F which is obtained by pushing forward the measure $\nu([i_1 \ldots i_n]) := \beta([i_n \ldots i_1])$ to F. (Note that $\beta = \nu$ in the case that β is Bernoulli).

Quantitative versions of (A) were considered in [JM, BJK] when β is sufficiently mixing. Analogues of Theorem 2.3 were obtained where the expected cover time was shown to satisfy

$$\lim_{r\to 0} \frac{\log \mathbb{E}\tau_{x_0}^r}{-\log r} = d$$

where d is the Minkowski dimension of the measure μ (where it exists)

$$d = \dim_{\mathrm{M}} \mu = \lim_{r\to 0} \frac{\log \min_{x\in \operatorname{supp}\mu} \mu(B(x, r))}{\log r}. \tag{2}$$

The Minkowski dimension can be interpreted as the $L^{-\infty}$ dimension of μ, which may be instructive to keep in mind since this provides a connection to the coarse multifractal spectrum (which roughly speaking describes, for any $\alpha > 0$ how many disjoint r-balls centred in $\operatorname{supp}\mu$ there are with measure approximately r^α) via the Legendre transform.

In [JM], where the Chaos Game was studied for self-similar sets F and β Bernoulli, the precise asymptotic for $\mathbb{E}\tau_{x_0}^r$ was determined exactly up to a uniform constant (i.e. the logarithmic factor was determined as in Theorem 2.3) in some, but not all cases. For example, when β is the unique Bernoulli measure β which pushes forward to the measure of maximal dimension μ supported on F, it was shown that $\mathbb{E}\tau_{x_0}^r \asymp r^{-d}\log(1/r)$.

Comparing this with Theorem 2.3 provides the right lens to interpret the asymptotic in Theorem 2.3; while the cover time is primarily determined by the time to hit the least accessible ball at a given scale, since $\mathcal{H}^s$ is an extremely regular measure it follows that all balls have roughly the same measure (thus are equally accessible) and indeed the Minkowski dimension of $\mathcal{H}^s$ coincides with all other notions of dimension of $\mathcal{H}^s$ (as well as the dimension of F).

Question 2.4. *Consider the generalised Chaos Game on a general self-similar set. Is it possible to prove that*

$$\mathbb{E}\tau_{x_0}^r \asymp r^{-d} \log N(r)$$

where $N(r)$ can be characterised in terms of the coarse multifractal spectrum? Heuristically one would expect $N(r)$ to be the number of disjoint balls at scale r whose measure is approximately minimal for that scale.

The simplest unknown case is when the S_i are similarities with contraction ratios r_i, β is a Bernoulli measure associated to the probability vector $(p_1, \ldots, p_N)$ and the maximum in the ratio $\max_{1 \leq i \leq N} \frac{\log p_i}{\log r_i}$ is attained at a unique value of $i \in \{1, \ldots, N\}$. While lower and upper estimates were obtained for this case in [JM, Theorem 1.1]*, they differ by a logarithmic factor.*

The Minkowski dimension was introduced by Falconer, Fraser and Kaenmaki in [FFK]. It does not always exist (more generally one should take limsups and liminfs in (2) to consider upper/lower Minkowski dimensions $\overline{\dim}_{\mathrm{M}}\mu$ and $\underline{\dim}_{\mathrm{M}}\mu$). The name derives from the fact that the Minkowski (box-counting) dimension of F, whenever it exists, satisfies $\dim_{\mathrm{M}} F = \min\{\overline{\dim}_{\mathrm{M}}\mu\}$ where the minimum is taken over fully supported finite Borel measures on F. It is a global property of the measure rather than a local property, for example it was shown in [BJK, Remark 5.2] that there exist self-affine measures on Bedford-McMullen carpets whose Minkowski dimension is strictly bigger than the maximum of the local dimension spectrum

$$\max_{x \in \mathrm{supp}\mu} \limsup_{r \to 0} \frac{\log \mu(B(x,r))}{\log r}.$$

We note that if the setup were to be relaxed even further to include IFS containing either infinitely many maps or maps which were not strict contractions, then the Chaos Game yields natural examples of measures μ whose Minkowski dimension is infinite. In this case the expected cover time at scale r satisfies a stretched exponential law in r rather than a power law, as seen in [JT, Examples 7.3 and 7.4].

2.2. **Blanket time estimates.** We will prove the following result for the blanket time:

Theorem 2.5. *There exists $C < \infty$ such that for any $r, \epsilon \in (0,1)$, and any $x_0 \in F$,*

$$\frac{1}{C}\left(\frac{1}{r}\right)^{\frac{\log 3}{\log 2}} \log\left(\frac{1}{r}\right) \leq \mathbb{E}\tau_{x_0}^{r,\epsilon} \leq \frac{C}{\epsilon^3}\left(\frac{1}{r}\right)^{\frac{\log 3}{\log 2}} \log\left(\frac{1}{r\epsilon}\right).$$

For a fixed ϵ, this result can be interpreted as comparability of the cover and blanket times. Namely, once the Chaos Game covers the fractal at a certain scale, one can expect to only wait a constant times that amount of time longer (where the constant is independent of the scale but depends on the notion of 'approximate') before the Chaos Game has approximately equidistributed at that scale.

2.2.1. *Discussion.* The blanket time has previously not been studied and it would be interesting to see to what extent a generalisation or refinement of Theorem 2.5 could be obtained.

The lower bound in Theorem 2.5 follows trivially from the fact that the cover time is a lower bound on the blanket time. Therefore a natural first question is:

Question 2.6. *For the standard Chaos Game on the Sierpinski triangle, is it possible to obtain a lower bound which also depends on ϵ? Is it possible to determine the exact dependence on ϵ i.e. find $c(\epsilon)$ such that $\mathbb{E}\tau_{x_0}^{r,\epsilon} \asymp c(\epsilon) \cdot \mathbb{E}\tau_{x_0}^{r}$?*

We will see in Proposition 3.2 that the regularity of the measure $\mathcal{H}^s$ is key to a clean translation of this problem into a symbolic setting. Even extending Theorem 2.5 to the Chaos Game on the Sierpinski triangle with non-uniform weights would require non-trivial arguments to handle the ensuing loss of regularity of the measure μ.

Question 2.7. *Consider the generalised Chaos Game on a general self-similar set. Is it true that*

$$\mathbb{E}\tau_{x_0}^{r,\epsilon} \leq c(\epsilon) \cdot \mathbb{E}\tau_{x_0}^{r}$$

and how does $c(\epsilon)$ depend on the regularity of the measure?

3. Symbolic cover and blanket times

In this section we recast the stopping times in the language of the symbolic space Σ, which will eventually allow us to relate the cover and blanket times to the classical cover and blanket times of certain Markov chains.

Given $n \in \mathbb{N}$, let $\Sigma_n = \{1,2,3\}^n$. Given $\mathtt{i} \in \Sigma_n$ we call

$$[\mathtt{i}] = \{\mathtt{j} \in \Sigma : \mathtt{j} = \mathtt{ik} \text{ for some } \mathtt{k} \in \Sigma\}$$

a cylinder set (of length n), where $\mathtt{ik}$ denotes the concatenated infinite word. Let $\Pi : \Sigma \to F$ be the coding map

$$\Pi(i_1 i_2 \ldots) = \lim_{n\to\infty} f_{i_1\ldots i_n}(\Delta)$$

where we write $f_{i_1\ldots i_n} = f_{i_1} \circ \cdots \circ f_{i_n}$. The initial point $x_0 \in F$ will be considered as fixed throughout this section. Fix $\mathtt{i}_0 \in \Sigma$ such that $\Pi(\mathtt{i}_0) = x_0$, noting that this is not a unique choice due to the overlap of the maps in the IFS. Let β denote the uniform Bernoulli measure on Σ, noting that $\mathcal{H}^s = \beta \circ \Pi^{-1}$.

It will be convenient to consider the following symbolic version of the cover time. Consider the function $\tau_{\mathtt{i}_0}^r : \Sigma \to \mathbb{N} \cup \{\infty\}$ given by

$$\tau_{\mathtt{i}_0}^m(\mathtt{i}) = \inf\{n \geq 0 : \{\mathtt{i}_0, i_1\mathtt{i}_0, \ldots, i_n \ldots i_1\mathtt{i}_0\} \text{ visits every cylinder of length } m\}.$$

Proposition 3.1. *For any $m \in \mathbb{N}$,*

$$\tau_{x_0}^{\frac{1}{2^{m-1}}} \leq \tau_{\mathtt{i}_0}^m \leq \tau_{x_0}^{\frac{1}{2^{m+1}}}.$$

Proof. Fix $\mathtt{i} \in \Sigma$ and let $n = \tau_{\mathtt{i}_0}^m(\mathtt{i})$. For the first inequality, notice that since projected cylinders of length m have diameter $\frac{1}{2^m}$, $\{x_0, f_{i_1}(x_0), \ldots, f_{i_n\ldots i_1}(x_0)\}$ must necessarily be $\frac{1}{2^{m-1}}$-dense in F.

For the second inequality, let $n = \tau_{x_0}^{\frac{1}{2^{m+1}}}(\mathtt{i})$ and fix arbitrary $\mathtt{j} \in \{1,2,3\}^m$. Since $\{x_0, f_{i_1}(x_0), \ldots, f_{i_n\ldots i_1}(x_0)\}$ is $\frac{1}{2^{m+1}}$- dense, $f_{\mathtt{j}}(F) \setminus \{f_{\mathtt{j}}(e_1), f_{\mathtt{j}}(e_2), f_{\mathtt{j}}(e_3)\}$ must contain a point from this sequence. Moreover, the first m digits of the coding of any point in $f_{\mathtt{j}}(F) \setminus \{f_{\mathtt{j}}(e_1), f_{\mathtt{j}}(e_2), f_{\mathtt{j}}(e_3)\}$ must necessarily be $\mathtt{j}$. This implies $\{\mathtt{i}_0, i_1\mathtt{i}_0, \ldots, i_n \ldots i_1\mathtt{i}_0\} \cap [\mathtt{j}] \neq \emptyset$. ◻

For convenience of notation let i_0 denote the empty word. We now define a symbolic version of the blanket time: $\tau_{\mathtt{i}_0}^{m,\epsilon}:\Sigma\to\mathbb{N}\cup\{\infty\}$ given by

$$\tau_{\mathtt{i}_0}^{m,\epsilon}(\mathtt{i})=\inf\left\{N\geq 0:\forall\mathtt{j}\in\{1,2,3\}^m,\frac{1}{N+1}\sum_{k=0}^{N}\mathbf{1}_{[\mathtt{j}]}(i_k\dots i_0\mathtt{i}_0)\geq(1-\epsilon)\frac{1}{3^m}\right\}.$$

Proposition 3.2. *For all $\epsilon\in(0,1)$ there exists $m_\epsilon\in\mathbb{N}$ with $m_\epsilon=O\left(\log\left(\frac{1}{\epsilon}\right)\right)$ such that for all $m\in\mathbb{N}$*

$$\tau_{x_0}^{\frac{1}{2^m},\epsilon}\leq\tau_{\mathtt{i}_0}^{m+m_\epsilon,\frac{\epsilon}{2}}.$$

Proof. Note that $\mathcal{H}^s$ is Ahlfors regular, i.e. there exist $0<a<b<\infty$ such that for all $y\in F$ and all $r>0$,

$$ar^s\leq\mathcal{H}^s(B(y,r))\leq br^s.$$

Let c be a constant such that for all $0<\delta,r<1$, the annulus $B(0,r)\setminus B(0,r(1-\delta))$ can be covered by $\frac{c}{\delta}$ balls of radius $r\delta$. Let $m_\epsilon\in\mathbb{N}$ be sufficiently large that $m_\epsilon>\frac{1}{s-1}\log_2\left(\frac{6bc}{a\epsilon}\right)$.

Let $n=\tau_{\mathtt{i}_0}^{m+m_\epsilon,\frac{\epsilon}{2}}$ and fix $y\in F$. Let $\mathcal{C}\subset\Sigma_{m+m_\epsilon}$ be the collection of $\mathtt{j}$ for which $f_{\mathtt{j}}(F)\subset B(y,\frac{1}{2^m})$.

By definition of the symbolic blanket time we have

$$\frac{1}{n+1}\sum_{k=0}^{n}\mathbf{1}_{B(y,r)}(f_{i_k\dots i_0}(x_0))\geq\frac{1}{n+1}\sum_{\mathtt{j}\in\mathcal{C}}\sum_{k=0}^{n}\mathbf{1}_{[\mathtt{j}]}(i_k\dots i_0\mathtt{i}_0)\geq\left(1-\frac{\epsilon}{2}\right)\sum_{\mathtt{j}\in\mathcal{C}}\beta([\mathtt{j}]).$$

By definition of c, the annulus $B(y,2^{-m})\setminus B(y,2^{-m}(1-2^{-m-m_\epsilon}))$ can be covered by $c2^{m_\epsilon}$ balls of radius 2^{-m-m_ϵ}. However we are interested in covering the intersection of F with the annulus $(B(y,2^{-m})\setminus B(y,2^{-m}(1-2^{-m-m_\epsilon})))\cap F$ by balls which are centred in F. By replacing the balls appearing in the cover of the annulus with appropriate balls centred in F it is not difficult to see that the intersection of the annulus with F can be covered by $3c2^{m_\epsilon}$ balls (each of which is centred in F) of radius 2^{-m-m_ϵ}. Therefore

$$\begin{aligned}\frac{\sum_{\mathtt{j}\in\mathcal{C}}\beta([\mathtt{j}])}{\mathcal{H}^s(B(y,2^{-m}))}&\geq 1-\frac{\mathcal{H}^s(B(y,2^{-m})\setminus B(y,2^{-m}(1-2^{-m_\epsilon})))}{\mathcal{H}^s(B(y,2^{-m}))}\geq 1-\frac{2^{m_\epsilon}3bc(2^{-m-m_\epsilon})^s}{a(2^{-m})^s}\\&=1-\frac{3bc}{2^{m_\epsilon(s-1)}a}>1-\frac{\epsilon}{2}.\end{aligned}$$

Therefore

$$\frac{1}{n+1}\sum_{k=0}^{n}\mathbf{1}_{B(y,r)}(f_{i_k\dots i_0}(x_0))\geq(1-\frac{\epsilon}{2})^2\mathcal{H}^s(B(y,2^{-m}))>(1-\epsilon)\mathcal{H}^s(B(y,2^{-m}))$$

which completes the proof. □

Alfhors regularity of $\mathcal{H}^s$ was crucial to the last proposition. A highly fluctuating measure may support too much measure near the boundary of a ball, which would prevent us from discarding the annulus in the way we did without too much loss of measure. For a less regular measure, this is a key part of the argument which would need to be refined using more subtle multifractal analysis.

4. Markov chains theory

Given an irreducible Markov chain $(X_n)_{n\in\mathbb{N}}$ on a state space V and stationary probability $\pi = (\pi(v))_{v\in V}$, the cover time [LP] of the Markov chain is defined as the random variable

$$\tau_{\mathrm{cov}} = \inf\{n \geq 0 : \forall v \in V, \exists 0 \leq k \leq n \text{ s.t. } X_k = v\}$$

and the ϵ-approximate blanket time [WZ] of the Markov chain is defined as the random variable

$$\tau_{\mathrm{bl}}^{\epsilon} = \inf\left\{n \geq 0 : \forall v \in V, \frac{\#\{0 \leq k \leq n : X_k = v\}}{n+1} \geq (1-\epsilon)\pi(v)\right\}.$$

For each $m \in \mathbb{N}$ we now define an irreducible Markov chain $(X_n^{(m)})_{n\in\mathbb{N}}$ whose cover and blanket times will coincide with an appropriate symbolic cover and blanket time introduced in the previous section.

Proposition 4.1. *Let $m \in \mathbb{N}$. We may build an irreducible Markov chain $(X_n^{(m)})_{n\in\mathbb{N}}$ on the state space Σ_m as follows. Define a $3^m \times 3^m$ square matrix $A_m = [a_{\mathtt{i},\mathtt{j}}]_{\mathtt{i},\mathtt{j}\in\Sigma_m}$ by*

$$a_{\mathtt{i},\mathtt{j}} := \begin{cases} \frac{1}{3} & \text{if } \exists i \text{ such that } [i\mathtt{i}] \subset [\mathtt{j}] \\ 0 & \text{otherwise,} \end{cases}$$

and define a vector $\pi_m \in \mathbb{R}^{3^m}$ by $\pi_{\mathtt{i}} := \frac{1}{3^m}$. Then:

(a) A_m is a row stochastic matrix,
(b) π_m is a left stationary vector, i.e. $\pi_m A_m = \pi_m$,
(c) A_m is irreducible.

Finally, letting $\tau_{\mathrm{cov},m}$ and $\tau_{\mathrm{bl},m}^{\epsilon}$ denote its cover and blanket times, for $\mathtt{i}_0 \in \Sigma^m$, $\mathbb{E}_{\mathtt{i}_0}\tau_{\mathrm{cov},m} = \mathbb{E}\tau_{\mathtt{i}_0}^{2^{-m}}$ and $\mathbb{E}_{\mathtt{i}_0}\tau_{\mathrm{bl},m}^{\epsilon} = \mathbb{E}\tau_{\mathtt{i}_0}^{2^{-m},\epsilon}$.

Proof. To see that A_m is row stochastic, fix $\mathtt{i} = i_1 \dots i_m \in \Sigma_m$ and note that $a_{\mathtt{i},\mathtt{j}} > 0$ for $\mathtt{j} = j_1 \dots j_m$ if and only if $j_1 j_2 \dots j_m = i i_1 \dots i_{m-1}$ for some $i \in \{1,2,3\}$. There will be exactly three such words $\mathtt{j}$ given by $1i_1 \dots i_{m-1}, 2i_1 \dots i_{m-1}, 3i_1 \dots i_{m-1}$. Hence $\sum_{\mathtt{j}\in\Sigma_m} a_{\mathtt{i},j} = 1$.

Since $\#\Sigma_m = 3^m$, π_m is clearly stochastic. To see that $\pi_m A_m = \pi_m$, fix $\mathtt{j} = j_1 \dots j_m \in \Sigma_m$ and note that $a_{\mathtt{i},\mathtt{j}} > 0$ for $\mathtt{i} = i_1 \dots i_m$ if and only if $i i_1 \dots i_{m-1} = j_1 \dots j_m$ for some $i \in \{1,2,3\}$. This will be the case for exactly three words $\mathtt{i}$, given by $j_2 \dots j_m 1, j_2 \dots j_m 2, j_2 \dots j_m 3$. Therefore the dot product of π_m and the $\mathtt{j}$ column of A_m is $3 \cdot \frac{1}{3^m} \cdot \frac{1}{3} = \frac{1}{3^m}$ which confirms stationarity.

To see that A_m is irreducible, we will see that $A_m^{3^m} = [a_{\mathtt{i},\mathtt{j}}^m]_{\mathtt{i},\mathtt{j}\in\Sigma_m}$ is a positive matrix. Indeed, $a_{\mathtt{i},\mathtt{j}}^m > 0$ if and only if there exist $k_1, \dots, k_m \in \{1,2,3\}$ such that $[k_1 \dots k_m \mathtt{i}] \subset [\mathtt{j}]$, which is clear by taking $k_1 = j_1, \dots, k_m = j_m$.

The coincidence of the expected cover and blanket times is clear. □

5. Proof of Theorem 2.3

For $\mathtt{i} \in \Sigma_m$ we define

$$\tau_{\mathtt{i}} := \min\{t \geq 0 : X_t^{(m)} = \mathtt{i}\}$$

which we call the hitting time to the state $\mathtt{i}$.

The motivation behind relating the cover time of the Chaos Game to the cover time of a Markov chain is that we can now appeal to the wealth of literature on the cover time of Markov chains. To estimate the cover time we will use a classical bounds of Matthews' (see [LP, Theorem 11.2] based on [M]). This result is proved using a 'coupon-collector' combinatorial type of argument which is excellently presented in [LP, Theorem 11.2]. This result captures the key heuristic upon which a possible affirmative answer to Question 2.4 might be built.

Proposition 5.1. *For any $m \in \mathbb{N}$,*

$$\min_{\mathtt{i},\mathtt{j} \in \Sigma_m} \mathbb{E}_{\mathtt{i}} \tau_{\mathtt{j}} \left(1 + \frac{1}{2} + \cdots + \frac{1}{3^m}\right) \leq \min_{\mathtt{i} \in \Sigma_m} \mathbb{E}_{\mathtt{i}} \tau_{\mathrm{cov},m}$$

and

$$\max_{\mathtt{i} \in \Sigma_m} \mathbb{E}_{\mathtt{i}} \tau_{\mathrm{cov},m} \leq \max_{\mathtt{i},\mathtt{j} \in \Sigma_m} \mathbb{E}_{\mathtt{i}} \tau_{\mathtt{j}} \left(1 + \frac{1}{2} + \cdots + \frac{1}{3^m}\right).$$

We now introduce a few more pieces of notation which will be used in the proof of Theorem 2.3. We define

$$\tau_{\mathtt{i}}^{+} := \min\{t \geq 1 : X_t^{(m)} = \mathtt{i}\},$$

which we call the first return time to the state $\mathtt{i}$. The famous Kac lemma establishes that $\mathbb{E}_{\mathtt{i}} \tau_{\mathtt{i}}^{+} = 3^m$.

Finally let $w_{\mathtt{i}}$ denote the first time that $\mathtt{i}$ appears using all new bits, that is, with no overlap with the initial state. This random variable is easier to study than the hitting time $\tau_{\mathtt{i}}$ since it does not depend on the initial state. There are two trivial observations which will be useful: (i) $w_{\mathtt{i}} \geq \tau_{\mathtt{i}}$ for all $\mathtt{i} \in \Sigma_m$ and (ii) since $w_{\mathtt{i}}$ does not depend on the initial state, $\mathbb{E} w_{\mathtt{i}} \geq \mathbb{E}_{\mathtt{i}} \tau_{\mathtt{i}}^{+}$, where we have purposefully suppressed the dependence on the initial state in $\mathbb{E} w_{\mathtt{i}}$.

We are now ready to prove Theorem 2.3.

Proof of Theorem 2.3. It will be sufficient to prove the bounds along the sequence of scales $r = \frac{1}{2^m}$. We begin by proving the upper bound. By the lower bound in Proposition 3.1, Proposition 4.1, the upper bound in Proposition 5.1 and since $1 + \frac{1}{3} + \cdots + \frac{1}{3^m} \asymp \log(3^m)$, it remains to estimate $\max_{\mathtt{i},\mathtt{j} \in \Sigma_m} \mathbb{E}_{\mathtt{i}} \tau_{\mathtt{j}}$ from above. Moreover since

$$\max_{\mathtt{i},\mathtt{j} \in \Sigma_m} \mathbb{E}_{\mathtt{i}} \tau_{\mathtt{j}} \leq \max_{\mathtt{j} \in \Sigma_m} \mathbb{E} w_{\mathtt{j}}$$

it is sufficient to prove $\max_{\mathtt{j} \in \Sigma_m} \mathbb{E} w_{\mathtt{j}} = O(3^m)$. Let $\mathtt{j} = j_m \dots j_1$. Now,

$$\mathbb{E} w_{j_m \dots j_1} \leq \frac{1}{3}(\mathbb{E} w_{j_{m-1} \dots j_1} + 1) + \frac{2}{3}(\mathbb{E} w_{j_{m-1} \dots j_1} + 1 + \mathbb{E} w_{j_m \dots j_1})$$

hence

$$\frac{1}{3} \mathbb{E} w_{j_m \dots j_1} \leq \mathbb{E} w_{j_{m-1} \dots j_1} + 1.$$

Proceeding inductively we obtain

$$\frac{1}{3^m} \mathbb{E}_{w_{j_m \dots j_1}} \leq \mathbb{E} w_{j_1} + \frac{3}{2}.$$

In particular

$$\mathbb{E} w_{j_m \dots j_1} \leq C 3^m$$

where $C = \frac{3}{2} + \mathbb{E} w_{j_1} = \frac{3}{2} + \frac{1}{3} \sum_{n=0}^{\infty} (\frac{2}{3})^n (n+1)$. This completes the proof of the upper bound.

By the upper bound in Proposition 3.1, Proposition 4.1 and the lower bound in Proposition 5.1 it remains to estimate $\min_{\mathtt{i},\mathtt{j}\in\Sigma_m}\mathbb{E}_{\mathtt{i}}\tau_{\mathtt{j}}$ from below. Let $\theta_{\mathtt{i},\mathtt{j}} = \mathbb{P}_{\mathtt{i}}(\tau_{\mathtt{j}} < m)$. Then

$$w_{\mathtt{j}} \leq \tau_{\mathtt{j}} + \mathbf{1}_{\{\tau_{\mathtt{j}}<m\}}(m + w^*_{\mathtt{j}})$$

where $w^*_{\mathtt{j}}$ is the amount of time required to build $\mathtt{j}$ from new bits after the mth bit has been added. Since $w^*_{\mathtt{j}}$ is independent of the event $\{\tau_{\mathtt{j}} < m\}$ and $w_{\mathtt{j}}$ has the same distribution as $w^*_{\mathtt{j}}$, taking expectations we get

$$\mathbb{E}w_{\mathtt{j}} \leq \mathbb{E}_{\mathtt{i}}\tau_{\mathtt{j}} + \theta_{\mathtt{i},\mathtt{j}}(m + \mathbb{E}w_{\mathtt{j}})$$

which we rearrange to obtain

$$\mathbb{E}_{\mathtt{i}}\tau_{\mathtt{j}} \geq \mathbb{E}w_{\mathtt{j}}(1 - \theta_{\mathtt{i},\mathtt{j}}) - \theta_{\mathtt{i},j}m. \tag{3}$$

Now,

$$\begin{aligned}\theta_{\mathtt{i},\mathtt{j}} = \mathbb{P}_{\mathtt{i}}(\tau_{\mathtt{j}} < m) &= \mathbb{P}_{\mathtt{i}}(\tau_{\mathtt{j}} = 1) + \cdots + \mathbb{P}_{\mathtt{i}}(\tau_{\mathtt{j}} = m-1)\\ &\leq \frac{1}{3} + \cdots + \frac{1}{3^{m-1}} \leq \frac{1}{2}\end{aligned}$$

where the second line follows by considering the number of correct digits that need to be appended for the event $\{\tau_{\mathtt{j}} = k\}$ to occur.

Combining with (3) gives

$$\mathbb{E}_{\mathtt{i}}\tau_{\mathtt{j}} \geq \frac{1}{2}\mathbb{E}w_{\mathtt{j}} - \frac{m}{2} \geq \frac{1}{2}\mathbb{E}\tau^+_{\mathtt{j}} - \frac{m}{2} \geq \frac{3^m}{2} - \frac{m}{2}$$

where we have used Kac lemma in the final inequality. The lower bound follows since $\frac{3^m}{2} - \frac{m}{2} \asymp 3^m$. □

6. Proof of Theorem 2.5

We will require the following result of Winkler and Zuckerman [WZ] about the blanket time of Markov chains. A key tool in the proof of this result is in proving estimates on sets of large deviations for multiple return times to a given region of the phase space.

Theorem 6.1. *Consider a collection $\mathcal{X}$ of Markov chains, each of which satisfies $\max_{v\in V}\mathbb{E}_v\tau_{\mathrm{cov}} = O(\max_{v,w}\mathbb{E}_v\tau_w \log \#V)$ (where the implied constants are uniform over $\mathcal{X}$). Then for each Markov chain in $\mathcal{X}$,*

$$\max_{v\in V}\mathbb{E}_v\tau^{\epsilon}_{\mathrm{bl}} = O\left(\frac{\max_{v,w}\mathbb{E}_v\tau_w \log \#V}{\epsilon^2}\right),$$

where again implied constants are uniform over $\mathcal{X}$.

Proof of Theorem 2.5. We will prove the result along the sequence of scales $r = \frac{1}{2^m}$, the proof for intermediate scales is similar.

For the lower bound note that $\mathbb{E}\tau^{r,\epsilon}_{x_0} \geq \mathbb{E}\tau^r_{x_0} \geq \frac{1}{C}r^{\frac{\log 3}{\log 2}}\log(\frac{1}{r})$ by Theorem 2.3.

For the upper bound, notice that by Theorem 2.3 (or indeed its proof), the hypothesis of Theorem 6.1 holds for the collection of Markov chains defined in Proposition 4.1. Therefore by Propositions 3.2, 4.1 and Theorem 6.1,

$$\mathbb{E}\tau_{x_0}^{\frac{1}{2^m},\epsilon} \leq \mathbb{E}\tau_{\mathtt{i}_0}^{m+m_\epsilon,\frac{\epsilon}{2}} = \mathbb{E}_{\mathtt{i}_0}\tau_{\mathrm{bl},m+m_\epsilon}^{\frac{\epsilon}{2}} = O\left(\frac{3^{m+m_\epsilon}\log(3^{m+m_\epsilon})}{\epsilon^2}\right) = O\left(\frac{3^m\log(\frac{3^m}{\epsilon})}{\epsilon^3}\right)$$

where the final equality is because $m_\epsilon = O(\log(\frac{1}{\epsilon}))$. □

6.1. **Conclusion.** We have shown that for the (uniform weights) Chaos Game on the Sierpinski triangle, the question of quantitative covering and equidistribution can be cleanly translated into the classical problems of cover and blanket times for Markov chains. This was due to the simplicity of the model we considered: the Sierpinski triangle is a homogeneous self-similar set and therefore r- neighbourhoods admit a straightforward symbolic coding, $\mathcal{H}^s$ can be obtained as the pushforward of a Bernoulli measure and therefore the Chaos Game can be modelled by a Markov chain, and since $\mathcal{H}^s$ is a very regular measure, "closeness to stationarity" for the Markov chains can be closely related to approximate equidistribution for the Chaos Game.

7. Further directions

We end with a couple of more speculative questions of broader scope.

Ding, Lee and Peres [DLP] proved that for random walks on undirected graphs, cover times and blanket times are uniformly comparable (i.e. for any $\epsilon > 0$ there exists a constant C_ϵ such that for *any* graph, its ϵ-approximate blanket time $\mathbf{B}^\epsilon$ and its cover time $\mathbf{C}$ satisfy the estimate $\mathbf{C} \leq \mathbf{B}^\epsilon \leq C_\epsilon \mathbf{C}$. This suggests the following question:

Question 7.1. *Can a uniform comparability result be established for the cover and blanket times within a natural class of dynamical systems?*

To be more precise, let us fix $\epsilon > 0$ and consider a dynamical system $\mathcal{X} = (X, T, \mu)$ (where $T : X \to X$ is a map of a metric space and μ is ergodic). We can naturally extend the definitions of cover and blanket times to this setting; let $\mathbb{E}\tau_{\mathrm{cov}}^{\mathcal{X}}$ denote its expected cover time and $\mathbb{E}\tau_{\mathrm{bl},\epsilon}^{\mathcal{X}}$ denote its expected ϵ-approximate blanket time. Given a family of dynamical systems $\mathcal{D}$ we say that the blanket times and cover times are uniformly comparable within $\mathcal{D}$ if there exists a constant $C = C(\mathcal{D}, \epsilon)$ such that $\mathbb{E}\tau_{\mathrm{bl},\epsilon}^{\mathcal{X}} \leq C \cdot \mathbb{E}\tau_{\mathrm{cov}}^{\mathcal{X}}$ for all $\mathcal{X} \in \mathcal{D}$.

What are the natural classes of dynamical systems on which the cover and blanket times are uniformly comparable?

Due to the duality between random walks on undirected graphs and electrical networks [DS], tools from electrical network theory have been successfully leveraged to solve various problems for random walks on graphs e.g. the notion of effective resistance in electrical networks theory is useful in bounding cover times of graphs and was fundamental to the proof of Ding, Lee and Peres [DLP] mentioned above.

Question 7.2. *Can the principles and tools from electrical network theory be recast in the language of dynamical systems to yield a compelling formalism?*

References

[Ba] M. Barnsley, *Fractals everywhere*, Academic Press Professional, Inc., 1988.

[BJK] B. Bárány, N. Jurga, I. Kolossváry, *On the convergence rate of the chaos game* International Mathematics Research Notices, 2023(5), 4456-4500.

[DLP] J. Ding, J. R. Lee, Y. Peres *Cover times, blanket times, and majorizing measures.* Annals of mathematics, 175(3) (2012) 1409-1471.

[DS] P. Doyle and J. Snell. *Random walks and electric networks.* Vol. 22. American Mathematical Soc., 1984.

[F] K. Falconer. *Fractal geometry: mathematical foundations and applications.* John Wiley & Sons. (2013)

[FFK] K. J. Falconer, J. M. Fraser, A. Käenmäki, *Minkowski dimension for measures,* Proceedings of the American Mathematical Society, 151(02), pp.779-794.

[H] J. Hutchinson *Fractals and self similarity.* Indiana University Mathematics Journal, 30(5) (1981) 713-747.

[JM] N. Jurga, I.D. Morris, *How long is the Chaos Game?* Bull. Lond. Math. Soc. **53** (2021) 1749–1765.

[JT] N. Jurga, M. Todd, *Cover times in dynamical systems* To appear in Israel Journal of Mathematics.

[LP] D. Levin, Y. Peres, *Markov chains and mixing times*, Second Edition, American Mathematical Society, Providence, RI, 2017.

[M] P. Matthews. *Covering problems for brownian motion on spheres*, Ann. Probab. **16** (1988) 189–199.

[WZ] P. Winkler and D. Zuckerman. *Multiple cover time.* Random Structures & Algorithms 9.4 (1996): 403-411.

Data Statement

This work does not rely on any publicly available datasets. All mathematical results are theoretical and reproducible from the definitions and proofs given within the paper.

Funding and conflict of interest

There are no known conflicts of interest. This work was partly funded by a Leverhulme ECR fellowship ECF-2021-385.

Mathematical Institute, University of St Andrews, Scotland, KY16 9SS

Email address: naj1@st-andrews.ac.uk