Uniform scaling limits for ergodic measures

Jonathan M. Fraser
School of Mathematics, The University of Manchester, Manchester, M13 9PL, UK
Email: jonathan.fraser@manchester.ac.uk

Mark Pollicott
Mathematics Institute, Zeeman Building, University of Warwick, Coventry, CV4 7AL, UK
Email: mpollic@maths.warwick.ac.uk

September 29, 2015

Abstract

We provide an elementary proof that ergodic measures on one-sided shift spaces are 'uniformly scaling' in the following sense: at almost every point the scenery distributions weakly converge to a common distribution on the space of measures. Moreover, we show how the limiting distribution can be expressed in terms of, and derived from, a 'reverse Jacobian' function associated with the corresponding measure on the space of left infinite sequences. Finally we specialise to the setting of Gibbs measures, discuss some statistical properties, and prove a Central Limit Theorem for ergodic Markov measures.

Mathematics Subject Classification 2010: 37B10, 28A33, 37A99, 60F05.
Key words and phrases: Ergodic measure, uniformly scaling measure, Gibbs measure.

1 Introduction

Given a set or measure one is often interested in studying the fine structure, i.e., properties derived from infinitesimal behaviour. As such it is important to understand 'tangents' and what is currently emerging in the literature on geometric measure theory, ergodic theory and fractal geometry is that understanding the dynamics of the process of 'zooming-in' to the tangents is even more valuable. Some of these ideas go back a long way, in particular to Furstenberg's work in the 60s and 70s, see [F1], but the techniques and philosophies have recently been applied to great effect, for example see [F2, G, H1, H2, HS1]. First one defines a process of zooming-in at a point in the support of a given measure. This may not converge

but one is interested in weak accumulation points of this process in the appropriate space of measures. One drawback of this approach is that one may obtain too many 'tangent measures' and not be able to relate them sensibly back to the original measure. As such one looks to define a measure on the space of measures (commonly referred to as a *distribution*) which best describes which measures are most prevalent during the process of zooming-in. We will make this precise in the context of measures on shift spaces in Section 1.1. Gavish [G] introduced the concept of a measure on Euclidean space being 'uniformly scaling' if at almost every point in the support of the measure, the (continuous) zooming-in process generates the same distribution. As such, uniformly scaling measures are very homogeneous and it turns out that one can make very strong statements about their geometry. In particular, see [H2, H1, HS1, HS2, KSS] for recent and influential developments in this direction.

In this paper we study the process described above abstractly in the context of ergodic measures on shift spaces and, in particular, the appropriate analogue of Gavish's notion of 'uniformly scaling'. Our main contributions are an elementary proof that shift ergodic measures are uniformly scaling and a demonstration of the relationship between the generated distribution in terms of the 'reverse Jacobian' of the naturally associated measure on the space of left infinite sequences. Our main result will be given in Section 2 and proved in Section 3. We also show that the stronger form of uniform scaling (which requires convergence of the q-sparse scenery distributions) also holds for strongly mixing measures, but not for all ergodic measures. In Section 4 we discuss the simpler setting of Gibbs measures, where the reverse Jacobian is the classical g-function. We consider some simple examples and finish by proving a Central Limit Theorem for the scaling scenery of ergodic Markov measures.

The fact that ergodic measures are uniformly scaling is at least implicit in other works, in particular [H1, F2], and even explicit in [H2, Section 3] in the setting of interval maps. The main difficulties in most of these examples are structural complications corresponding to the geometry of the specific setting. We consider the problem in symbolic space, where these complications do not occur, and aim to give a clear exposition of why the necessary convergence occurs relying only on classical ergodic theory.

1.1 Scaling scenery for measures on shift spaces

Let $\mathcal{I} = \{1, \dots, k\}$ be a finite alphabet, $\Sigma^+ = \prod_0^\infty \mathcal{I}$ be the space of one-sided sequences over $\mathcal{I}$ and σ denote the usual (left) shift map. Abusing notation slightly we write $x = (x_0, \dots, x_{n-1}) \in \prod_0^{n-1} \mathcal{I}$ and $x = (x_0, x_1, \dots) \in \Sigma^+$. We equip Σ^+ with the standard metric defined by

$$d(x, y) = 2^{-n(x,y)}$$

where for $x \neq y$, $n(x, y) = \max\{n \in \mathbb{N} : x_k = y_k \text{ for all } k = 0, \dots, n\}$ or -1 if no such n exists. Write $\mathcal{P}(\Sigma^+)$ for the space of Borel probability measures on Σ^+ and $\mathcal{P}_\sigma(\Sigma^+)$ for the space of shift invariant measures from $\mathcal{P}(\Sigma^+)$. Equip both these spaces of measures with the weak topology, which can be metrised with either the

Levy-Prokhorov or Wasserstein metric for example, and write $\mathrm{spt}(\mu)$ for the support of a measure μ. Let $x = (x_0, x_1, x_2, \cdots) \in \Sigma^+$ and $n \geqslant 1$ and define "blow up" maps $T_{n,x} : \Sigma^+ \to \Sigma^+$ by

$$T_{n,x}(y_0, y_1, y_2, \cdots) = (x_0, \cdots, x_{n-1}, y_0, y_1, y_2, \cdots)$$

i.e., inserting the first n terms from the sequence x at the front of y. We define *cylinder sets* as

$$[x_0, \ldots, x_{n-1}]_0^{n-1} \;=\; T_{n,x}\Sigma^+ \;=\; \{y \in \Sigma^+ : y_i = x_i \text{ for } 0 \leqslant i \leqslant n-1\}$$

and the following definition allows us to blow up μ on the cylinders containing x.

Definition 1.1. *The maps $T_{n,x}$ induce a sequence of new measures $\mu_{x,n} \in \mathcal{P}(\Sigma^+)$, which are called* minimeasures*, defined by*

$$\mu_{x,n}(A) = \frac{\mu\left(T_{n,x}A\right)}{\mu\left(T_{n,x}\Sigma^+\right)}$$

for measurable $A \subset \Sigma^+$, provided $\mu\left(T_{n,x}\Sigma^+\right) > 0$. This sequence of minimeasures is called the scaling scenery *of μ at x and any weak-$*$ accumulation point of the scaling scenery is called a* micromeasure *of μ at x.*

Recently, there has been considerable interest in understanding the limiting behaviour of the scaling scenery and many closely related concepts. It is perhaps unreasonable to expect the scaling scenery to converge, but one is interested in which minimeasures are most prevalent in the scaling scenery and to this end we define a sequence of *measures on the space of measures* by taking Cesàro averages of Dirac measures on the minimeasures along the scaling scenery and then hope that this converges. Let $\mathcal{D}(\Sigma^+) = \mathcal{P}(\mathcal{P}(\Sigma^+))$ be the space of Borel measures on $\mathcal{P}(\Sigma^+)$, which we call the space of *distributions*.

Definition 1.2. *The Nth* scenery distribution *of μ at $x \in \mathrm{spt}(\mu)$ is*

$$\frac{1}{N}\sum_{n=0}^{N-1} \delta_{\mu_{x,n}} \;\in\; \mathcal{D}(\Sigma^+)$$

and any weak-$*$ accumulation point of the sequence of Nth scenery distributions is called a micromeasure distribution*.*

It is straightforward to see that any micromeasure distribution at x is supported on the set of micromeasures of μ at x. Motivated by Gavish [G], we define the analogue of uniformly scaling for shift spaces as follows.

Definition 1.3. *A measure $\mu \in \mathcal{P}(\Sigma^+)$ is a* uniformly scaling measure *if there exists a distribution $Q \in \mathcal{D}(\Sigma^+)$ such that at μ almost every $x \in \Sigma^+$*

$$\frac{1}{N}\sum_{n=0}^{N-1} \delta_{\mu_{x,n}} \;\to_{w^*}\; Q.$$

In this case we say that μ generates *the distribution Q.*

One is often interested in a stronger notion of uniform scaling. Motivated by [HS1], we make the following definitions.

Definition 1.4. *Given $q \in \mathbb{N}$, the* q-sparse Nth scenery distribution *of μ at $x \in \mathrm{spt}(\mu)$ is*

$$\frac{1}{N}\sum_{n=0}^{N-1} \delta_{\mu_{x,qn}} \ \in \ \mathcal{D}(\Sigma^+).$$

We say that a measure $\mu \in \mathcal{P}(\Sigma^+)$ is a strongly uniformly scaling measure *if there exists a family of distributions $\{Q_q\}_{q\in\mathbb{N}}$ from $\mathcal{D}(\Sigma^+)$ such that at μ almost every $x \in \Sigma^+$ and for every $q \in \mathbb{N}$*

$$\frac{1}{N}\sum_{n=0}^{N-1} \delta_{\mu_{x,qn}} \ \to_{w^*} \ Q_q.$$

We note that for applications it is not usually important if the Q_q vary with q or not. However, from a philosophical point of view it seems sensible to highlight measures for which Q_q is constant and indeed this will be the case in our Theorem 2.2.

1.2 Ergodic measures and the reverse Jacobian

Let $\Sigma = \prod_{-\infty}^{\infty} \mathcal{I}$ be the space of infinite two-sided sequences where we write $x = (x_m, \dots, x_n) \in \prod_m^n \mathcal{I}$ (with $m < n$) and $x = (\dots, x_{-1}; x_0, x_1, \dots) \in \Sigma$. We also write σ for the (invertible) left shift map on Σ given by

$$\sigma(\cdots, x_{-2}, x_{-1}; x_0, x_1, x_2, \cdots) = (\cdots, x_{-1}, x_0; x_1, x_2, x_3, \cdots)$$

and let $\mathcal{P}_\sigma(\Sigma)$ denote the space of shift invariant Borel probability measures on Σ.

Lemma 1.5. *There is a natural bijection between the spaces $\mathcal{P}_\sigma(\Sigma^+)$ and $\mathcal{P}_\sigma(\Sigma)$. Moreover, this map is also a bijection between ergodic measures on Σ^+ and Σ.*

Proof. For the first part we use the (unique) extension of an invariant measure μ on Σ^+ to Σ given by $\mu([x_m, \cdots, x_n]_m^n) := \mu([x_m, \cdots, x_n]_0^{n-m})$ (with $m < n$). The fact that ergodic measures are paired with ergodic measures is straightforward and omitted. □

Given an ergodic measure $\mu \in \mathcal{P}_\sigma(\Sigma)$, define a sequence of functions $g_n : \Sigma \to [0, 1]$ by

$$g_n(x) = \frac{\mu\big([x_{-n}, x_{-(n-1)}, \cdots, x_{-1}]_{-n}^{-1}\big)}{\mu\big([x_{-n}, x_{-(n-1)}, \cdots, x_{-2}]_{-n}^{-2}\big)}$$

for $x = (x_l)_{l=-\infty}^{\infty} \in \mathrm{spt}(\mu)$ and $g_n(x) = 0$ for $x \in \Sigma \setminus \mathrm{spt}(\mu)$. From this sequence of functions we are able to define the reverse Jacobian g which we will need to state our main result, Theorem 2.1.

Lemma 1.6. *The limit $g(x) := \lim_{n\to+\infty} g_n(x)$ exists for μ almost every $x \in \Sigma$ and, moreover, $g \in L^1(\Sigma, \mu)$.*

Proof. Consider the space of left infinite sequences Σ^- and let μ^- be the push forward of μ to Σ^- via the natural restriction. Let σ^{-1} be the associated right shift and note that μ^- need not be σ^{-1} invariant. Observe that g_n only depends on past coordinates, i.e. $g_n(x) = g_n(x')$ if x and x' are such that $x_n = x'_n$ for all $n < 0$, and so for $x \in \mathrm{spt}(\mu)$

$$g_n(x) = \frac{\mu^-\big([x_{-n}, x_{-(n-1)}, \cdots, x_{-1}]_{-n}^{-1}\big)}{\mu^- \circ \sigma^{-1}\big([x_{-n}, x_{-(n-1)}, \cdots, x_{-1}]_{-n}^{-1}\big)}$$

which is the reciprocal of the Radon-Nikodym derivative $\mathrm{d}(\mu^- \circ \sigma^{-1})/\mathrm{d}\mu^-$ with respect to the σ-algebra generated by the cylinders of length n in Σ^-. Even though μ^- may be singular, the Radon-Nikodym derivative is well-defined because σ^{-1} is countable to one, see [P, Section 10-1] and also [PW]. It follows from [Pa, Proposition 48.1] that g_n converges almost surely to an L^1 function g, which is the Jacobian of $\mathrm{d}(\mu^- \circ \sigma^{-1})/\mathrm{d}\mu^-$ with respect to the full Borel σ-algebra. □

2 Scaling scenery for ergodic measures

We now wish to make more precise statements about the scaling scenery and to do so we need to introduce some more notation. Given any word $e = (e_0, e_1, \ldots, e_{m-1}) \in \prod_0^{m-1} \mathcal{I}$, and $b > a > 0$ the open sets

$$\mathcal{U}^e(a, b) := \big\{\nu \in \mathcal{P}(\Sigma^+) : \nu([e_0 \ldots e_{m-1}]_0^{m-1}) \in (a, b)\big\}$$

generate the weak-$*$ topology on $\mathcal{P}(\Sigma^+)$ and so determining the value of a distribution on these generating sets determines it uniquely.

Theorem 2.1. *Every ergodic measure $\mu \in \mathcal{P}_\sigma(\Sigma^+)$ is a uniformly scaling measure. Moreover, for a given ergodic $\mu \in \mathcal{P}_\sigma(\Sigma^+)$ the generated distribution $Q \in \mathcal{D}(\Sigma^+)$ is characterised as follows. Also write μ for the associated two-sided ergodic measure from Lemma 1.5 and let $g \in L^1(\Sigma, \mu)$ be given by Lemma 1.6. Then*

$$Q\left(\mathcal{U}([e_0 \ldots e_{m-1}]_0^{m-1}, a, b)\right) \;=\; \mu\left(\left\{y \in \Sigma \,:\, a < \prod_{k=1}^{m} g\big(\sigma^k(y^- e)\big) < b\right\}\right)$$

for any cylinder $[e_0 \cdots e_{m-1}]_0^{m-1}$ and $a < b$ and where we write $y^- e = (\cdots, y_{-2}, y_{-1}; e_0 \cdots e_{m-1}, \cdots)$ observing that since g only depends on past coordinates it does not matter how we complete the sequence to the right.

This can be compared with the result for interval maps which appears in [H2, Proposition 3.6]. The relevance of extensions to bi-infinite sequences in the context of blowing up a set or measure has also been observed before. In particular, see Sullivan's limit diffeomorphisms [S] and subsequent developments and applications of these ideas [CJQ, BF, HS1, H2]. One heuristic justification is that the positive coordinates give location and the negative coordinates give distortion as one zooms in at that location. Recall that a measure $\mu \in \mathcal{P}_\sigma(\Sigma^+)$ is *completely ergodic* if it is σ^q-ergodic for all $q \in \mathbb{N}$ and observe that any strongly mixing measure is completely ergodic.

Theorem 2.2. *Every completely ergodic measure $\mu \in \mathcal{P}_\sigma(\Sigma^+)$ is a strongly uniformly scaling measure. Indeed, for all $q \in \mathbb{N}$ and almost every $x \in \Sigma^+$, the q-sparse scenery distributions converge to the distribution $Q \in \mathcal{D}(\Sigma^+)$ described in Theorem 2.1. In particular, this holds if μ is strongly mixing.*

Finally we point out that it is easy to construct invariant measures which are not uniformly scaling, see Example 4.6. For such examples the scenery distributions almost surely converge to a common distribution within each ergodic component, but the distributions can vary between component. Also, it is easy to construct ergodic measures which are not strongly uniformly scaling, see Example 4.7.

3 Proof of Theorems 2.1 and 2.2

Throughout this section we will write μ both for the original ergodic measure in $\mathcal{P}_\sigma(\Sigma^+)$ and for the associated ergodic measure in $\mathcal{P}_\sigma(\Sigma)$ from Lemma 1.5. Given a word $e = (e_0, e_1, \dots, e_{m-1}) \in \prod_0^{m-1} \mathcal{I}$, let us define a sequence of functions $g_n^e : \Sigma \to [0,1]$ by

$$g_n^e(x) = \frac{\mu\big([x_{-n}, x_{-(n-1)}, \cdots, x_{-1}; e_0, e_1, \dots, e_{m-1}]_{-n}^{m-1}\big)}{\mu\big([x_{-n}, x_{-(n-1)}, \cdots, x_{-1}]_{-n}^{-1}\big)}$$

for $x = (x_l)_{l=-\infty}^{\infty} \in \operatorname{spt}(\mu)$ and $g_n^e(x) = 0$ for $x \in \Sigma \setminus \operatorname{spt}(\mu)$.

Lemma 3.1. *For μ almost every $x \in \Sigma$, the sequence $g_n^e(x)$ converges and*

$$\lim_{n\to+\infty} g_n^e(x) \;=\; \prod_{k=1}^{m} g\big(\sigma^k(x^- e)\big) \;=:\; g^e(x)$$

where $x^- e = (\dots, x_{-n}, x_{-(n-1)}, \cdots, x_{-1}; e_0, e_1, \dots, e_{m-1}, \dots)$ recalling that g only depends on the past coordinates and so it does not matter how $x^- e$ is filled in to the right.

Proof. We assume that $\mu\big([x_{-n}, x_{-(n-1)}, \cdots, x_{-1}; e_0, e_1, \dots, e_{m-1}]_{-n}^{m-1}\big) > 0$ for all $n \in \mathbb{N}$ and that $x \in \operatorname{spt}(\mu)$. If this is not the case then the result is trivial and $g^e(x) = 0$. We have

$$\begin{aligned}
g_n^e(x) &= \frac{\mu\big([x_{-n}, x_{-(n-1)}, \cdots, x_{-1}; e_0, e_1, \dots, e_{m-1}]_{-n}^{m-1}\big)}{\mu\big([x_{-n}, x_{-(n-1)}, \cdots, x_{-1}]_{-n}^{-1}\big)} \\
&= \prod_{k=1}^{m} \frac{\mu\big([x_{-n}, x_{-(n-1)}, \cdots, x_{-1}; e_0, e_1, \dots, e_{k-1}]_{-n}^{k-1}\big)}{\mu\big([x_{-n}, x_{-(n-1)}, \cdots, x_{-1}; e_0, e_1, \dots, e_{k-2}]_{-n}^{k-2}\big)} \\
&= \prod_{k=1}^{m} g_{n+k}\big(\sigma^k(x^- e)\big) \\
&\to \prod_{k=1}^{m} g\big(\sigma^k(x^- e)\big)
\end{aligned}$$

for μ almost every $x \in \Sigma$ as $n \to +\infty$ by Lemma 1.6. □

Lemma 3.2. *Let $e \in \prod_0^{m-1} \mathcal{I}$. Then for any $\epsilon, \delta > 0$ we can choose a measurable set $B \subset \Sigma$ with $\mu(B) < \delta$ and n_0 such that for $n \geqslant n_0$ we have*

$$\sup_{x \in \Sigma \setminus B} |g_n^e(x) - g^e(x)| < \epsilon.$$

Proof. This is an immediate consequence of Egorov's Theorem. □

Lemma 3.3. *Fix a measurable set $B \subseteq \Sigma$ and $a, b \in \mathbb{R}$ with $a < b$. Then for μ almost every $x \in \Sigma$, as $N \to +\infty$ we have*

$$\frac{1}{N} \# \{0 \leqslant n \leqslant N-1 \,:\, \sigma^n x \in B\} \to \mu(B)$$

and

$$\frac{1}{N} \# \{0 \leqslant n \leqslant N-1 \,:\, g^e(\sigma^n x) \in (a,b)\} \to \mu(\{y \in \Sigma \,:\, g^e(y) \in (a,b)\}).$$

Proof. This follows immediately by applying the Birkhoff ergodic theorem for $\sigma : \Sigma \to \Sigma$ and μ. □

Observe that $\mu_{x,n}([e]_0^{m-1})$ is defined for all $x \in \mathrm{spt}(\mu)$ and so we can extend it to a function of $x = (x_l)_{l=-\infty}^{\infty} \in \Sigma$ by setting it to zero whenever $(x_l)_{l=0}^{\infty} \notin \mathrm{spt}(\mu) \subseteq \Sigma^+$. This has the advantage that

$$\mu_{x,n}([e_0, \ldots, e_{m-1}]_0^{m-1}) = \frac{\mu[x_0, x_1, \cdots, x_{n-1}, e_0, \ldots, e_{m-1}]_0^{m+n-1}}{\mu[x_0, x_1, \cdots, x_{n-1}]_0^{n-1}} = g_n^e(\sigma^n x). \quad (3.1)$$

In particular, for $x \in \Sigma$ the terms $\mu_{x,n}([e]_0^{m-1})$ depend only on the future coordinates. We are now in position to prove Theorem 2.1.

Proof. Fix $e = (e_0, e_1, \ldots, e_{m-1}) \in \prod_0^{m-1} \mathcal{I}$ and $a, b \in \mathbb{R}$ with $a < b$. We will estimate the measure of $\mathcal{U}^e(a,b)$ for scenery distributions at generic $x \in \Sigma^+$. The following fact is stated merely for clarity.

Fact 3.4. *If $\left(\frac{1}{N}\sum_{n=0}^{N-1} \delta_{\mu_{x,n}}\right)(\mathcal{U}^e(a,b))$ converges for μ almost every $x \in \Sigma$ to a constant then $\left(\frac{1}{N}\sum_{n=0}^{N-1} \delta_{\mu_{x,n}}\right)(\mathcal{U}^e(a,b))$ converges for μ almost every $x \in \Sigma^+$ to the same constant.*

Let $B \subseteq \Sigma$ and n_0 be taken from Lemma 3.2 and observe that for $N > n_0$ and for all $x \in \Sigma$ we have

$$\begin{aligned}
\left(\frac{1}{N}\sum_{n=0}^{N-1} \delta_{\mu_{x,n}}\right)(\mathcal{U}^e(a,b)) &= \frac{1}{N} \# \left\{0 \leqslant n \leqslant N-1 \,:\, \mu_{x,n}([e_0, \ldots, e_{m-1}]_0^{m-1}) \in (a,b)\right\} \\
&= \frac{1}{N} \# \{0 \leqslant n \leqslant N-1 \,:\, g_n^e(\sigma^n x) \in (a,b)\} \qquad \text{by (3.1)} \\
&\leqslant \frac{n_0}{N} \;+\; \frac{1}{N} \# \{0 \leqslant n \leqslant N-1 \,:\, g^e(\sigma^n x) \in (a-\epsilon, b+\epsilon)\} \\
&\quad + \; \frac{1}{N} \# \{0 \leqslant n \leqslant N-1 \,:\, \sigma^n x \in B\}.
\end{aligned}$$

Letting $N \to +\infty$ we can apply Lemma 3.3 and deduce that for μ almost every $x \in \Sigma^+$ if $Q \in \mathcal{D}(\Sigma^+)$ is an accumulation point of the scenery distributions at x, then

$$Q\left(\mathcal{U}^e(a,b)\right) \leqslant \mu(\{y \in \Sigma : g^e(y) \in (a-\epsilon, b+\epsilon)\}) + \delta.$$

A similar argument shows that

$$Q\left(\mathcal{U}^e(a,b)\right) \geqslant \mu(\{y \in \Sigma : g^e(y) \in (a+\epsilon, b-\epsilon)\}) - \delta.$$

Since $\epsilon, \delta > 0$ are arbitrary and $\mathcal{D}(\Sigma^+)$ is sequentially compact by Prokhorov's Theorem, we deduce that for μ almost every $x \in \Sigma^+$ the scenery distributions at x converge to a common distribution $Q \in \mathcal{D}(\Sigma^+)$ satisfying

$$\begin{aligned} Q\left(\mathcal{U}^e(a,b)\right) &= \mu(\{y \in \Sigma : g^e(y) \in (a,b)\}) \\ &= \mu\left(\left\{y \in \Sigma : a < \prod_{k=1}^{m} g\big(\sigma^k(y^- e)\big) < b\right\}\right) \end{aligned}$$

which completes the proof. □

Theorem 2.2 follows by an almost identical argument to that used to prove Theorem 2.1 and so we omit the details. Observe that

$$\left(\frac{1}{N}\sum_{n=0}^{N-1} \delta_{\mu_{x,qn}}\right)\left(\mathcal{U}^e(a,b)\right) = \frac{1}{N} \#\left\{0 \leqslant n \leqslant N-1 : g^e_{qn}\big((\sigma^q)^n x\big) \in (a,b)\right\}$$

and then apply the Birkhoff ergodic theorem for $\sigma^q : \Sigma \to \Sigma$ in the same way as before. We use the fact that if $\mu \in \mathcal{P}_\sigma(\Sigma^+)$ is completely ergodic, then the associated measure $\mu \in \mathcal{P}_\sigma(\Sigma)$ from Lemma 1.5 is also completely ergodic.

4 Scaling scenery for Gibbs measures

In this section we specialise to the setting of Gibbs measures and consider some simple examples. Let $\phi : \Sigma_A^+ \to \mathbb{R}$ be a Hölder continuous potential for a subshift of finite type Σ_A^+. A measure $\mu \in \mathcal{P}(\Sigma^+)$ supported on Σ_A^+ is called a *Gibbs measure* for ϕ if there exists constants $C_1, C_2 > 0$ such that

$$C_1 \leqslant \frac{\mu\big([x_0, \ldots, x_{n-1}]_0^{n-1}\big)}{\exp\left(\sum_{k=0}^{n-1} \phi(\sigma^k x) - nP(\phi)\right)} \leqslant C_2 \tag{4.1}$$

for all $x \in \Sigma_A^+$ and all $n \in \mathbb{N}$ and where $P(\phi)$ is the pressure of ϕ, see [B]. If Σ_A^+ is topologically mixing, then there is a unique shift invariant Gibbs measure $\mu = \mu_\phi \in \mathcal{P}_\sigma(\Sigma^+)$ and this Gibbs measure is strongly mixing and thus completely ergodic. Two very simple examples of shift invariant Gibbs measures are *Bernoulli measures* and *Markov measures*. We will use these as examples and so briefly recall their definitions. Let $(p_i)_{i \in \mathcal{I}}$ be a strictly positive probability vector associated to $\mathcal{I}$. Given the potential $\phi(x) = \log p_{x_0}$ for the full shift, the unique invariant Gibbs measure satisfies

$$\mu\big([x_0, \ldots, x_{n-1}]_0^{n-1}\big) = p_{x_0} \cdots p_{x_{n-1}}$$

and is called a *Bernoulli measure*. One more level of complexity yields Markov measures. Given a right stochastic matrix $P = \{p_{i,j}\}_{i,j\in\mathcal{I}}$ and a left invariant (stationary) probability vector $(\pi_i)_{i\in\mathcal{I}}$, the associated *Markov measure* is defined by

$$\mu\big([x_0,\dots,x_{n-1}]_0^{n-1}\big) \;=\; \pi_{x_0}p_{x_0,x_1}\cdots p_{x_{n-2},x_{n-1}}.$$

Markov measures are invariant Gibbs measures for the potential $\phi(x) = \log p_{x_0,x_1}$ and are supported on the subshift of finite type given by the transition matrix formed by replacing all non-zero entries in P with 1s. If P is irreducible, then there is a unique stationary probability vector and thus a unique associated Markov measure and this measure is ergodic. Furthermore, if P is aperiodic, then the associated subshift of finite type is topologically mixing and the Markov measure is strongly mixing. We will also utilise the theory of Gibbs measures on the two-sided shift space Σ which are defined similarly, see [B].

Lemma 4.1. *If μ is an invariant Gibbs measure for a Hölder potential, then $\psi := \log g$ is a Hölder potential for the corresponding invariant Gibbs measure on Σ given by Lemma 1.5 where $g : \Sigma \to [0,1]$ is the (almost everywhere defined) reverse Jacobian function given by Lemma 1.6.*

Proof. This is a standard result in the general theory of g-measures, beginning with Keane in the 70s [K]. The fact that ψ is a potential for the two-sided Gibbs measure μ is due to Ledrappier [L], see also [W, Theorem 2.1]. □

Theorem 2.1 and Lemma 4.1 combine to yield the following result for Gibbs measures.

Corollary 4.2. *Let $\mu \in \mathcal{P}_\sigma(\Sigma^+)$ be an ergodic Gibbs measure for a Hölder continuous potential ϕ defined on a subshift of finite type Σ_A^+. Then μ is uniformly scaling generating a distribution $Q \in \mathcal{D}(\Sigma^+)$. Moreover, there exists a Hölder potential $\psi : \Sigma_A \to \mathbb{R}$ for the associated two-sided Gibbs measure from Lemma 1.5 (which is supported on the corresponding two-sided subshift of finite type Σ_A) such that*

$$Q\left(\mathcal{U}([e_0\dots e_{m-1}]_0^{m-1},a,b)\right) \;=\; \mu\left(\left\{y\in\Sigma\,:\,a<\exp\left(\sum_{k=1}^{m}\psi\big(\sigma^k(y^-e)\big)\right)<b\right\}\right)$$

for any cylinder $[e_0\cdots e_{m-1}]_0^{m-1}$ and $a < b$ and where we write $y^-e = (\cdots,y_{-2},y_{-1};e_0\cdots e_{m-1},\cdots)$ observing that since ψ depends only on past coordinates it does not matter how we complete the sequence to the right. Moreover, if Σ_A^+ is topologically mixing, then μ is strongly uniformly scaling and generates distributions Q_q which are all equal to Q above.

The following proposition shows that for an ergodic Gibbs measure the support of the distribution Q is very homogeneous in that all measures in the support are uniformly equivalent to some first level blow up. A similar observation was made in [FP], where the uniform equivalence was needed to pursue geometric applications. For $i \in \mathcal{I}$, let $\mu_i \in \mathcal{P}(\Sigma^+)$ be defined by

$$\mu_i(A) = \frac{\mu\left(iA\right)}{\mu\left([i]_0^0\right)}$$

for a measurable set $A \subseteq \Sigma^+$.

Proposition 4.3. *Let $\mu \in \mathcal{P}(\Sigma^+)$ be a Gibbs measure for a Hölder continuous potential ϕ defined on a subshift of finite type Σ_A^+. Then there exists a uniform constant $C \geqslant 1$ depending only on the potential such that for all measurable $A \subseteq \Sigma^+$ and all mini- or micromeasures ν, for some $i \in \mathcal{I}$ we have*

$$C^{-1}\,\mu_i(A) \;\leqslant\; \nu(A) \;\leqslant\; C\,\mu_i(A).$$

Proof. It suffices to prove the result for minimeasures because the bounds are clearly preserved under weak convergence to any micromeasure. Let $\nu = \mu_{x,n}$ be a minimeasure of μ at $x \in \Sigma_A^+$ and depth $n \in \mathbb{N}$ and let $i = x_{n-1}$. It suffices to estimate the measure only for cylinders, so let $y \in \Sigma^+$ and $m \in \mathbb{N}$ define an arbitrary cylinder $[y_0, \ldots, y_{m-1}]_0^{m-1} \subseteq \Sigma^+$. Let the kth variation of the potential ϕ be defined by

$$\operatorname{var}_k(\phi) \;=\; \sup_{x,y\in\Sigma^+} \Big\{ |\phi(x) - \phi(y)| : x_0 = y_0, \ldots, x_{k-1} = y_{k-1} \Big\}.$$

A simple consequence of ϕ being Hölder is that it has *summable variations*, i.e.

$$V(\phi) := \sum_{k=0}^{\infty} \operatorname{var}_k(\phi) < \infty.$$

The choice of i implies that $\mu_i\big([y_0, \ldots, y_{m-1}]_0^{m-1}\big) = 0$ if and only if $\nu\big([y_0, \ldots, y_{m-1}]_0^{m-1}\big) = 0$ and so assume without loss of generality that these measures are not zero. We have

$$\begin{aligned}
&\frac{\nu\big([y_0, \ldots, y_{m-1}]_0^{m-1}\big)}{\mu_i\big([y_0, \ldots, y_{m-1}]_0^{m-1}\big)} \\
&\quad= \quad \frac{\mu\big([x_0, \ldots, x_{n-1}, y_0, \ldots, y_{m-1}]_0^{m+n-1}\big)\mu\big([x_{n-1}]_0^0\big)}{\mu\big([x_0, \ldots, x_{n-1}]_0^{n-1}\big)\mu\big([x_{n-1}y_0, \ldots, y_{m-1}]_0^m\big)} \\
&\quad\leqslant \quad \frac{C_2^2 \exp\Big(\sum_{k=0}^{m+n-1} \phi\big(\sigma^k(x_0, \ldots, x_{n-1}\,y)\big) - (m+n)P(\phi)\Big) \exp\Big(\phi\big((x_{n-1}\,y)\big) - P(\phi)\Big)}{C_1^2 \exp\Big(\sum_{k=0}^{n-1} \phi(\sigma^k(x)) - nP(\phi)\Big) \exp\Big(\sum_{k=0}^{m} \phi(\sigma^k(x_{n-1}\,y)) - (m+1)P(\phi)\Big)} \\
&\quad= \quad \frac{C_2^2}{C_1^2}\; \exp\left(\sum_{k=0}^{n-1} \phi\big(\sigma^k(x_0, \ldots, x_{n-1}\,y)\big) - \sum_{k=0}^{n-1} \phi(\sigma^k(x))\right) \\
&\quad\leqslant \quad \frac{C_2^2}{C_1^2}\; \exp\left(\sum_{k=0}^{n-1} \operatorname{var}_{n-k}(\phi)\right) \;\leqslant\; \frac{C_2^2}{C_1^2} \exp\big(V(\phi)\big) \;<\; \infty.
\end{aligned}$$

A similar argument going in the opposite direction yields

$$\frac{\nu\big([y_0, \ldots, y_{m-1}]_0^{m-1}\big)}{\mu_i\big([y_0, \ldots, y_{m-1}]_0^{m-1}\big)} \;\geqslant\; \frac{C_1^2}{C_2^2} \exp\big(-V(\phi)\big) \;>\; 0$$

completing the proof. □

Note that if Σ_A^+ is the full shift, then all of the μ_i are equivalent to μ. Thus Proposition 4.3 shows that all micromeasures of a fully supported Gibbs measure

are themselves Gibbs measures for the same potential and a pleasant consequence of this is that there is at most one *invariant* micromeasure for any (invariant or non-invariant) fully supported Gibbs measure.

In the simpler setting of locally constant potentials, one can say even more. In fact, an explicit expression for the generated distribution Q can be derived easily from the definitions.

Example 4.4. *Let $\mu \in \mathcal{P}_\sigma(\Sigma^+)$ be a Bernoulli measure. Then all minimeasures and micromeasures at any point are equal to μ itself and so μ is strongly uniformly scaling and generates the distribution $\delta_\mu \in \mathcal{D}(\Sigma^+)$.*

Proof. This follows immediately from the definitions. □

The situation for Markov measures is only slightly more complicated. Here there are k different measures one can find in the scaling scenery, corresponding to the first level blow ups μ_i.

Example 4.5. *Let $\mu \in \mathcal{P}_\sigma(\Sigma_A^+)$ be an ergodic Markov measure for a subshift of finite type Σ_A^+. Then μ is uniformly scaling and generates the distribution*

$$\sum_{i\in\mathcal{I}} \pi_i \, \delta_{\mu_i} \in \mathcal{D}(\Sigma^+).$$

Moroever, if Σ_A^+ is topologically mixing, then μ is strongly uniformly scaling and all of the generated distributions Q_q are equal to the above.

Proof. Let $i \in \mathcal{I}$, $x \in \Sigma^+$ such that $x_{n-1} = i$ and let $y \in \Sigma^+$ and $m \in \mathbb{N}$ be arbitrary. Then

$$\begin{aligned}
\mu_{x,n}([y_0,\ldots,y_{m-1}]_0^{m-1}) &= \frac{\mu\big([x_0,\ldots x_{n-1},y_0,\ldots,y_{m-1}]_0^{m+n-1}\big)}{\mu\big([x_0,\ldots,x_{n-1}]_0^{n-1}\big)} \\
&= \frac{\pi_{x_0}\, p_{x_0x_1}\cdots p_{x_{n-2}x_{n-1}} p_{x_{n-1}y_0}\cdots p_{y_{m-2}y_{m-1}}}{\pi_{x_0}\, p_{x_0x_1}\cdots p_{x_{n-2}x_{n-1}}} \\
&= p_{x_{n-1}y_0}\cdots p_{y_{m-2}y_{m-1}} \\
&= \mu_i([y_0,\ldots,y_{m-1}]_0^{m-1})
\end{aligned}$$

and so for such x and n, $\mu_{x,n} = \mu_i$. This observation combined with the Birkhoff ergodic theorem implies that for μ almost all x we have

$$\begin{aligned}
\frac{1}{N}\sum_{n=0}^{N-1} \delta_{\mu_{x,n}} &= \sum_{i\in\mathcal{I}} \left(\frac{1}{N}\sum_{n=0}^{N-1} \mathbf{1}_{[i]_0^0}\big(\sigma^n(x)\big)\right)\delta_{\mu_i} \\
&\to_{w^*} \sum_{i\in\mathcal{I}} \left(\int_{\Sigma^+} \mathbf{1}_{[i]_0^0}\,\mathrm{d}\mu\right)\delta_{\mu_i} \\
&= \sum_{i\in\mathcal{I}} \pi_i\,\delta_{\mu_i}
\end{aligned}$$

completing the proof. The topologically mixing case follows similarly. □

These results can easily be extended to "generalised Markov measures", i.e., the Gibbs measures for locally constant functions. The following examples are a simple demonstration of the sharpness of Theorems 2.1 and 2.2.

Example 4.6. *The Markov measure associated with the stochastic matrix*

$$\begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

and left invariant probability vector $(1/3, 1/3, 1/3)$ *is* ***invariant but not uniformly scaling*** *as the scenery distributions converge to different limits almost surely within the two ergodic components.*

Example 4.7. *The unique Markov measure associated with the stochastic matrix*

$$\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1/2 & 1/2 & 0 \end{pmatrix}$$

is ***ergodic but not strongly uniformly scaling****. In particular, for even* q *it is not* σ^q *ergodic and the* q*-sparse scenery distributions converge to different limits almost surely within the* σ^q *ergodic components.*

Note that both of these (counter) examples have positive entropy. The simplicity of the result in the case of Markov measures allows us to make more precise statements about the statistical behaviour of the scenery distributions. For example, we have the following Central Limit Theorem (CLT).

Corollary 4.8 (Central Limit Theorem)**.** *Let* μ *be an ergodic Markov measure and* $Q = \sum_{i\in\mathcal{I}} \pi_i \, \delta_{\mu_i}$ *be the distribution it generates. Fix a cylinder* $[e_0 \cdots e_{m-1}]_0^{m-1}$ *and* $a < b$ *and write* $\mathcal{U} = \mathcal{U}([e_0 \dots e_{m-1}]_0^{m-1}, a, b)$*. If* $Q(\mathcal{U}) \in (0,1)$*, then letting* $\sigma^2 = Q(\mathcal{U}) - Q(\mathcal{U})^2 > 0$ *we have*

$$\frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \delta_{\mu_{x,n}}(\mathcal{U}) \;-\; \sqrt{N}\, Q(\mathcal{U}) \;\Rightarrow\; \mathcal{N}(0, \sigma^2)$$

where $\Rightarrow$ *denotes convergence in distribution.*

Proof. Write $\mathcal{I}_{\mathcal{U}} = \{i \in \mathcal{I} : \mu_i \in \mathcal{U}\}$. Example 4.5 and the classical CLT yield

$$\begin{aligned} \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \delta_{\mu_{x,n}}(\mathcal{U}) \;-\; \sqrt{N}\, Q(\mathcal{U}) \;&=\; \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} \left(\sum_{i\in\mathcal{I}_{\mathcal{U}}} \mathbf{1}_{[i]_0^0}\big(\sigma^n(x)\big) \right) \;-\; \sqrt{N}\, Q(\mathcal{U}) \\ &\Rightarrow\; \mathcal{N}(0, \sigma^2) \end{aligned}$$

which completes the proof. We have used the fact that $X_n := \sum_{i\in\mathcal{I}_{\mathcal{U}}} \mathbf{1}_{[i]_0^0}\big(\sigma^n(x)\big)$ is an i.i.d. sequence taking the value 1 with probability $Q(\mathcal{U}) = \sum_{i\in\mathcal{I}_{\mathcal{U}}} \pi_i$ and 0 otherwise. □

Other related statistical results follow similarly, but we omit further details. If one was interested in obtaining a CLT for general Gibbs measures then, inspecting the proof of Theorem 2.1, one obtains

$$\frac{1}{\sqrt{N}}\sum_{n=0}^{N-1}\delta_{\mu_{x,n}}(\mathcal{U}) \;-\; \sqrt{N}\,Q(\mathcal{U}) \;=\; \frac{1}{\sqrt{N}}\sum_{n=0}^{N-1}\mathbf{1}_{(g_n^e)^{-1}(a,b)}\big(\sigma^n(x)\big) \;-\; \sqrt{N}\,Q(\mathcal{U}).$$

This expression is more difficult to handle for two reasons. The first is that it involves an ergodic sum for a sequence of functions and so one needs an analogue of the CLT for Maker's ergodic theorem. The second and more important reason is that the sequence $X_n := \mathbf{1}_{(g_n^e)^{-1}(a,b)}\big(\sigma^n(x)\big)$ is not i.i.d. and moreover we cannot guarantee that the functions $\mathbf{1}_{(g_n^e)^{-1}(a,b)}$ or even $\mathbf{1}_{(g^e)^{-1}(a,b)}$ are Hölder continuous, despite the fact we know (in the Gibbs setting) that g_n^e and g^e are Hölder. This prevents us from using several standard results on CLTs for ergodic sums, see [PP, CP]. In the setting of ergodic non-Gibbs μ, a CLT appears even harder to achieve because we can only guarantee g^e is L^1.

Acknowledgements

This work began while JMF was a PDRA of MP at the University of Warwick. JMF and MP were financially supported in part by the EPSRC grant EP/J013560/1. The authors thank Pablo Shmerkin and Tuomas Sahlsten for providing helpful references.

References

[BF] T. Bedford and A. M. Fisher. Ratio geometry, rigidity and the scenery process for hyperbolic Cantor sets, *Ergodic Theory Dynam. Systems*, **17** (1997), 531–564.

[B] R. Bowen. *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, Lecture Notes in Math. 470. Berlin: Springer, 1975.

[CP] Z. Coelho and W. Parry. Central limit asymptotics for shifts of finite type, *Israel J. Math.*, **69**, (1990), 235–249.

[CJQ] G. Cui, Y. Jiang and A. Quas. Scaling functions, Gibbs measures, and Teichmüller spaces of circle endomorphisms, *Discrete Contin. Dynam. Systems*, **5**, (1999), 535–552.

[FP] J. M. Fraser and M. Pollicott. Micromeasure distributions and applications for conformally generated fractals, *Math. Proc. Camb. Phil. Soc. (to appear)*, available at http://arxiv.org/abs/1502.05609.

[F1] H. Furstenberg. Intersections of Cantor sets and transversality of semigroups, *Problems in Analysis (Princeton Mathematical Series, 31).* (1970), 41–59.

[F2] H. Furstenberg. Ergodic fractal measures and dimension conservation, *Ergodic Theory Dynamical Systems*, **28**, (2008), 405–422.

[G] M. Gavish. Measures with uniform scaling scenery, *Ergodic Theory Dynam. Systems*, **31**, (2011), 33–48.

[H1] M. Hochman. Dynamics on fractals and fractal distributions, *preprint*, 2010, available at http://arxiv.org/abs/1008.3731.

[H2] M. Hochman. Geometric rigidity of times-m invariant measures, *J. Eur. Math. Soc.*, **14**, (2012), 1539–1563.

[HS1] M. Hochman and P. Shmerkin. Local entropy averages and projections of fractal measures, *Ann. Math.*, **175**, (2012), 1001–1059.

[HS2] M. Hochman and P. Shmerkin. Equidistribution from fractal measures, *Invent. Math. (to appear)*, available at http://arxiv.org/abs/1302.5792.

[KSS] A. Käenmäki, T. Sahlsten and P. Shmerkin. Dynamics of the scenery flow and geometry of measures, *Proc. London Math. Soc. (to appear)*, available at http://arxiv.org/abs/1401.0231.

[K] M. Keane. Strongly mixing g-measures, *Invent. Math.*, **16**, (1972), 309–324.

[L] F. Ledrappier. Principe variationnel et systèmes dynamiques symboliques, *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, **30**, (1974), 185–202.

[P] W. Parry. *Entropy and generators in ergodic theory*, W. A. Benjamin, Inc., New York-Amsterdam 1969 xii+124 pp.

[PP] W. Parry and M. Pollicott. *Zeta functions and periodic orbit structure of hyperbolic dynamics*, Astérisque, (1990), 1–268.

[PW] W. Parry and P. Walters. Endomorphisms of a Lebesgue space, *Bull. Amer. Math. Soc.*, **78**, (1972), 272–276.

[Pa] K. R. Parthasarathy. *Introduction to probability and measure*, The Macmillan Co. of India, Ltd., Delhi, 1977. xii+312 pp.

[S] D. Sullivan. Differentiable structures on fractal-like sets, determined by intrinsic scaling functions on dual Cantor sets, *The Mathematical Heritage of Hermann Weyl* (Durham, NC, 1987), *Proc. Sympos. Pure Math.*, **48**, *Amer. Math. Soc.*, Providence, RI, (1988), 15–23.

[W] P. Walters. Ruelle's operator theorem and g-measures, *Trans. Amer. Math. Soc.*, **214**, (1975), 375–387.