MANCHESTER
1824

The University of Manchester

Distributional and local limit laws for a class of iterated maps that contract on average

Santos, Sara I and Walkden, Charles

2011

MIMS EPrint: **2011.39**

Manchester Institute for Mathematical Sciences
School of Mathematics

The University of Manchester

Reports available from: http://eprints.maths.manchester.ac.uk/
And by contacting: The MIMS Secretary
School of Mathematics
The University of Manchester
Manchester, M13 9PL, UK

ISSN 1749-9097

Distributional and local limit laws for a class of iterated maps that contract on average

Sara I. Santos[*†] and Charles Walkden[*]

26th March 2011

Abstract

We consider iterated function schemes that contract on average with place-dependent probabilities. We are interested in generalisations of the central limit theorem, particularly to observations with infinite variance. By studying the spectral properties of an associated one-parameter family of transfer operators acting on an appropriate function space, we prove both a distributional and local limit law with convergence to a stable distribution.

§1 Introduction

The study of the limiting behaviour of the sum of a sequence of observations of random variables is a key problem in dynamical systems and probability theory. For example, the ergodic theorem describes the average behaviour of such sums. In the case where the observation has finite variance, the central limit theorem then describes the how these sums are distributed around their expected value, namely convergence in distribution to a normal distribution. More generally, if the observation does not have a finite variance, then one can ask about convergence in distribution to a stable law. Stable limit laws have been well-understood for i.i.d. random variables [IL, for example], however there has been much recent interest in analogues of such results in dynamical systems, particularly in hyperbolic and non-uniformly hyperbolic systems [AD2, Gou, for example] and for random walks on the affine group of the real line [GP]. In this note we study iterated function schemes (IFS) with place-dependent probabilities that satisfy a 'contraction-on-average' condition. Central limit theorems, and generalisations thereof,

[*]School of Mathematics, The University of Manchester, Oxford Road, Manchester, M13 9PL, U.K.

[†]An earlier version of the results in this paper appeared in the first author's PhD thesis. Financially supported by Fundação para a Ciência e Tecnologia, Portugal, grant SFRH/BD/6921/2001.

for IFSs that contract-on-average have been studied in [Pe, HH2, W]. Such systems are of interest in a wide range of situations, see [DF] and the references cited therein. We discuss stable limit laws in the context of IFSs that contract on average. We also prove a local version of the stable limit law. The methodology uses the spectral properties of a one-parameter family of transfer operators P_t. By studying P_t on an appropriate function space, motivated by [HH1, HH2, GP], we prove that P_t is a quasi-compact operator with a simple maximal eigenvalue λ_t. We then apply a result from [AD1, AD2] to study the asymptotic expansion of λ_t, allowing us to relate λ_t to the sums in question.

§2 Statement of results

Let (X, d) be a locally compact (but not necessarily compact) second countable metric space. Consider a finite or countable family of Lipschitz maps $T_j : X \to X$, $1 \leq j < M$ (where $M \leq \infty$). We are interested in studying the statistical properties of the iterated function scheme (IFS) formed by applying the maps T_j chosen at random according to place-dependent Markov transition probabilities.

Let $p_j : X \to [0, 1]$ be continuous maps such that $\sum_j p_j(x) = 1$ for each $x \in X$. Define a Markov transition probability by

$$p(x, A) = \sum_j p_j(x)\chi_A(T_j x)$$

for each Borel subset $A \subset X$. (Here χ_A denotes the characteristic function of A.)

We say that the system *contracts on average after* 1 *step* if there exists $r \in (0, 1)$ such that

$$\sup_{x,y,z\in X, y\neq z} \sum_j p_j(x)d(T_j y, T_j z) \leq rd(y, z). \tag{1}$$

Remark. More precisely, we could refer to (1) as *arithmetic* contraction on average. It is strictly stronger than a *logarithmic contraction on average* condition, namely

$$\sup_{x,y,z\in X, y\neq z} \sum_j p_j(x) \log d(T_j y, T_j z) \leq r \log d(y, z),$$

as assumed in [BDEG].

More generally, we will consider IFSs that contract on average after n_0 steps; see §3.1.

We also assume that the $p_j \geq 0$ are continuous and satisfy a Dini condition (cf. [E1]).

With these assumptions, together with the mild technical assumptions in §3.1, it is known [BDEG, Pe] that there exists a unique attractive stationary Borel probability measure ν on X, i.e. for all Borel sets A

$$\int p(x, A)\, d\nu(x) = \nu(A). \tag{2}$$

Moreover, for any $x_0 \in X$, we have that $\int d(x, x_0)\, d\nu(x) < \infty$. Let

$$\theta_0 = \sup \left\{ \theta > 0 \mid \int d(x, x_0)^\theta\, d\nu(x) < \infty \text{ for all } x_0 \in X \right\}.$$

Note that $\theta_0 > 1$ and that $\int d(x, x_0)^\theta\, d\nu(x) < \infty$ for all $\theta < \theta_0$.

Let $\Sigma = \{\mathbf{j} = (j_1, j_2, \ldots) \mid 1 \leq j_k < M\}$ denote the one-sided full-shift. Define a cylinder set by $[j_1, j_2, \ldots, j_n] = \{\mathbf{i} = (i_k) \in \Sigma \mid i_k = j_k, 1 \leq k \leq n\}$. For each $x \in X$ we define a probability measure μ_x on Σ by defining μ_x on cylinder sets by

$$\mu_x[j_1, j_2, \ldots, j_n] = p_{j_1}(x) p_{j_2}(T_{j_1} x) \cdots p_{j_n}(T_{j_{n-1}} \cdots T_{j_1} x). \tag{3}$$

For each $x \in X$ and $\mathbf{j} = (j_1, j_2, \ldots) \in \Sigma$ we define

$$Z_n(x, \mathbf{j}) = T_{j_n} \cdots T_{j_1}(x)$$

and set $Z_0(x, \mathbf{j}) = x$. Then $Z_n(x, \mathbf{j})$ is an X-valued Markov chain with respect to μ_x, with initial state x and transition probability p. For convenience if $\mathbf{j} = (j_1, j_2, \ldots) \in \Sigma$ then we shall often write $T_n(\mathbf{j}) = T_{j_n} \cdots T_{j_1}$.

We can relate μ_x and ν as follows [E2]. Define $\pi_x(\mathbf{j}) = \lim_{n\to\infty} T_{j_1} T_{j_2} \cdots T_{j_n}(x)$ for μ_x-a.e. $\mathbf{j} \in \Sigma$. Then for all $x \in X$ we have $\pi_x^* \mu_x = \nu$.

Let $f : X \to \mathbb{R}$ be a continuous function on X. We are interested in the distribution of the sequence of observations

$$S_n f(x, \mathbf{j}) = \sum_{k=1}^{n} f(Z_k(x, \mathbf{j})). \tag{4}$$

It is known [E1] that $S_n f$ satisfies a pointwise ergodic theorem: for all $x \in X$ and μ_x-a.e. $\mathbf{j} \in \Sigma$,

$$\lim_{n\to\infty} \frac{1}{n} S_n f(x, \mathbf{j}) = \nu(f). \tag{5}$$

Under the mild technical hypotheses stated in §3.1, a central limit theorem is also known to hold [Pe, HH2]. Let $f : X \to \mathbb{R}$ be a bounded Lipschitz function and fix $x \in X$. Then

$$\frac{1}{\sqrt{n}} S_n f(x, \cdot) \to_d \mathcal{N}_{\nu(f), \sigma^2(f)}, \tag{6}$$

provided that the variance $\sigma^2(f) > 0$. Here $\mathcal{N}_{\nu(f),\sigma^2(f)}$ denotes the normal distribution with mean $\nu(f)$ and variance $\sigma^2(f)$ and $\to_d$ denotes convergence in distribution. The variance is given by

$$\sigma^2(f) = \lim_{n\to\infty} \frac{1}{n} \int (S_n f(x,\cdot) - n\nu(f))^2 \, d\mu_x.$$

If f is a bounded Lipschitz function then $\sigma^2(f) < \infty$.

The space X is typically not compact and so it makes sense to consider functions f which satisfy some degree of regularity (a Hölder condition, for example) but which are not in L^2 and which do not have a finite variance. In this case, it is natural to conjecture that the sequence of observations (4), when normalised by a sequence that grows like $n^{1/p}$ for a suitable parameter $p \in (0, 2)$ (called the *order*), converges in distribution to a stable distribution $Y_{p,\beta,b,c}$ (where $p \in (0, 2), \beta \in [-1, 1], b \in \mathbb{R}, c > 0$ are parameters described in §5). Stable laws can be characterised as being generalisations of the Gaussian distribution that keep the stability property: if X and Y are two random variables with the same stable distribution (up to an affine rescaling) then $X + Y$ has the same distribution (up to an affine rescaling). Stable laws of order $p = 1$ are technically more difficult to deal with [AD1] and for simplicity we concentrate on the case $p \in (0, 1) \cup (1, 2)$. We give a brief introduction to stable laws and their properties that we shall need in §5.

Let $f : X \to \mathbb{R} \in L^1(\nu)$ be continuous. We assume that f satisfies a Hölder condition that we make precise in §3.2. Assume in addition that for some $p \in (0, 2)$

$$\nu\{x \mid f(x) > t\} = \frac{1}{t^p}(C_1 + o(1)), \quad \nu\{x \mid f(x) < -t\} = \frac{1}{t^p}(C_2 + o(1)), \quad (7)$$

for constants $C_1, C_2 > 0$. (This condition can be weakened to include a slowly varying function—see §§5,7,8.)

Our main result is the stable limit theorem.

Theorem 2.1 (Distributional Stable Limit Theorem)
Suppose that the IFS (T_j, p_j) contracts on average and satisfies the technical hypotheses in §3.1. Suppose that $f : X \to \mathbb{R}$ satisfies a Hölder condition stated in §3.2 and that, for some $p \in (0, 2)$, (7) holds. Then for all $x \in X$,

$$\mu_x \left\{ \mathbf{j} \in \Sigma \;\middle|\; \frac{S_n f(x, \mathbf{j}) - a_n}{n^{1/p}} < t \right\} \to \int_{-\infty}^{t} dY_p,$$

as $n \to \infty$, for some stable law Y_p of order p, where

$$a_n = \begin{cases} 0 & \text{if } p < 1, \\ n\nu(f), & \text{if } p > 1. \end{cases} \quad (8)$$

Given an observation f, the key to our results is an analysis of the spectral properties of a one-parameter family of transfer operators P_t acting on a certain Banach space of functions. The expansion of the maximal eigenvalue λ_t of P_t as a function of t is intimately related to the characteristic function of the respective stable distribution. An observation f that essentially does not take values in a lattice will produce a non-periodic perturbation of the spectrum of the transfer operator P. In this case the only t for which P_t has an eigenvalue of modulus 1 is $t = 0$. Lattice-valued observations f give rise to periodicity in t of the spectrum of P_t.

For a non-arithmetic observation f we have the following local limit theorem.

Theorem 2.2 (Non-arithmetic local stable limit theorem) *Suppose that the IFS (T_j, p_j) contracts on average and satisfies the technical hypotheses in §3.1. Suppose that $f : X \to \mathbb{R}$ satisfies a Hölder condition given in §3.2 and the non-arithmeticity assumption in §8.1, and that, for some $p \in (0, 2)$, (7) holds. Then there exists a stable law Y_p of order p with density y_p such that for any $a, b \in \mathbb{R}$, $a < b$ and any $x \in X$ we have*

$$\lim_{n\to\infty} \left| n^{1/p} \mu_x\{\mathbf{j} \in \Sigma \mid S_n f(x, \mathbf{j}) - a_n \in z + [a, b]\} - y_p(z/n^{1/p})(b - a) \right| = 0$$

uniformly in $z \in \mathbb{R}$, where a_n is given by (8).

In the case where the observation f has finite variance, the same method of proof also provides a local central limit theorem.

Theorem 2.3 (Non-arithmetic local central limit theorem) *Suppose that the IFS (T_j, p_j) contracts on average and satisfies the technical hypotheses in §3.1. Suppose that $f : X \to \mathbb{R}$ satisfies a Hölder condition given in §3.2 and the non-arithmeticity assumption in §8.1, and $0 < \sigma^2(f) < \infty$. Then for any $a, b \in \mathbb{R}$, $a < b$ and any $x \in X$ we have*

$$\lim_{n\to\infty} \Big| (2\pi\sigma^2(f)n)^{1/2} \mu_x\{\mathbf{j} \in \Sigma \mid S_n f(x, \mathbf{j}) - n\nu(f) \in z + [a, b]\} - e^{\frac{-z^2}{2\sigma^2(f)n}} (b - a) \Big| = 0$$

uniformly in $z \in \mathbb{R}$.

§3 Assumptions and function spaces

§3.1 Technical hypotheses

Let (X, d) be a locally compact second countable metric space. Choose and fix a choice of origin $x_0 \in X$.

Let $T_j : X \to X$ be a finite or countable family of Lipschitz maps. If $T : X \to X$ is Lipschitz then we define

$$\|T\| = \sup_{x,y \in X, x \neq y} \frac{d(Tx, Ty)}{d(x,y)}.$$

Let $p_j : X \to [0,1]$ be a countable family of probability functions such that $p_j(x) \geq 0$ for all $x \in X, j \in \mathbb{N}$ and are Dini continuous. Define the probability measure μ_x on Σ by (3). Define $m(p_j) = \sup\{|p_j(x) - p_j(y)| \mid x, y \in X, d(x,y) \leq 1\}$.

Definition. The IFS *contracts on average after n_0 steps* if there exists $r \in (0,1)$ such that

$$\sup_x \mathbb{E}_x \left(\|T_{n_0}(\mathbf{j})\| \right) \leq r. \tag{9}$$

We remark that there exist examples of IFSs (T_j, p_j) that contract on average after n_0 steps, but which do not contract after 1 step and for which none of the T_j are strict contractions [Pe].

We will assume the following technical conditions hold.

(i) We have

$$\sup_{x,y \in X} \mathbb{E}_x \left(\frac{d(T_j(y), x_0)}{1 + d(y, x_0)} \right) < \infty. \tag{10}$$

(ii) We have

$$\sup_{x \in X} \mathbb{E}_x \left(\max\{1, \|T_j\|\} + d(T_j x_0, x_0) \right) m(p_j) < \infty. \tag{11}$$

(iii) We assume that for each $x, y \in X$ there exists $\mathbf{i} = (i_1, i_2, \ldots) \in \Sigma$ such that for each m,

$$\mu_x[i_1, i_2, \ldots, i_m], \mu_y[i_1, i_2, \ldots, i_m] \neq 0. \tag{12}$$

Assumption (10) can be viewed as a moment assumption on the T_j. Note that (10), (11) are automatically satisfied if there are finitely many maps. Assumption (12) is an irreducibility assumption that ensures that the transfer operator below has 1 as the unique simple maximal eigenvalue.

§3.2 Function spaces

We assume that the continuous observation $f : X \to \mathbb{R}$ satisfies a uniform Hölder condition, namely for some $\alpha \in (0,1]$:

$$|f|_{(\alpha)} = \sup_{x,y \in X, x \neq y} \frac{|f(x) - f(y)|}{d(x,y)^\alpha} < \infty. \tag{13}$$

For example, if $X = \mathbb{R}$ then $f(x) = x^\alpha$ satisfies (13). In this case, $\sigma^2(f) = \infty$ if $2\alpha > \theta_0$.

Choose $\varepsilon, \beta, \gamma > 0$ such that

$$0 < \varepsilon < \varepsilon + \beta < \gamma < \varepsilon + 2\beta < 1. \tag{14}$$

For $\lambda \in [0,1]$ we define $d_\lambda(x) = 1 + \lambda d(x, x_0)$.

We will fix a choice of λ below. For a continuous function $w : X \to \mathbb{C}$ define

$$|w|_\gamma = \sup_{x \in X} \frac{|w(x)|}{d_\lambda(x)^\gamma}$$

and

$$|w|_{\varepsilon,\beta} = \sup_{x,y \in X, x \neq y} \frac{|w(x) - w(y)|}{d(x,y)^\varepsilon d_\lambda(x)^\beta d_\lambda(y)^\beta}.$$

Then $|\cdot|_\gamma$ is a norm and $|w|_{\varepsilon,\beta}$ is a semi-norm. We define a norm by

$$\|w\|_{\varepsilon,\beta,\gamma} = |w|_\gamma + |w|_{\varepsilon,\beta}.$$

Then the spaces $C_\gamma = \{w : X \to \mathbb{R} \mid w \text{ is continuous and } |w|_\gamma < \infty\}$ and $C_{\varepsilon,\beta,\gamma} = \{w \in C_\gamma \mid \|w\|_{\varepsilon,\beta,\gamma} < \infty\}$ are Banach spaces with respect to the norms $|\cdot|_\gamma$ and $\|\cdot\|_{\varepsilon,\beta,\gamma}$, respectively. Such function spaces have been well-used in the study of IFSs and random walks on groups [GP, HH1, HH2, Pe]. Note that if $\lambda_1, \lambda_2 \in (0,1]$ then $d_{\lambda_1}(x) \leq \lambda_2^{-1} d_{\lambda_2}(x)$. Although different values of $\lambda \in (0,1)$ give different norms $|\cdot|_\gamma$ and $\|\cdot\|_{\varepsilon,\beta,\gamma}$, it follows that the norms are equivalent.

Let $T : X \to X$ be Lipschitz. Define

$$\delta_\lambda(T) = \sup_{x \in X} \frac{d_\lambda(Tx)}{d_\lambda(x)}.$$

Lemma 3.1
If $T : X \to X$ is Lipschitz then $\delta_\lambda(T) \leq \max\{1, \|T\|\} + \lambda d(Tx_0, x_0)$.

Proof. Observe that

$$\begin{aligned}
\frac{d_\lambda(Tx)}{d_\lambda(x)} &\leq \frac{1 + \lambda d(Tx, Tx_0)}{1 + \lambda d(x, x_0)} + \frac{\lambda d(Tx_0, x_0)}{1 + \lambda d(x, x_0)} \\
&\leq \frac{1 + \lambda \|T\| d(x, x_0)}{1 + \lambda d(x, x_0)} + \lambda d(Tx_0, x_0) \\
&\leq \max\{1, \|T\|\} + \lambda d(Tx_0, x_0).
\end{aligned}$$

❑

We now choose λ. Note that, as $\varepsilon + 2\beta < 1$ and taking $n = qn_0$, it follows from (9) that

$$\begin{aligned}
&\sup_x \mathbb{E}_x \left(\|T_{qn}(\cdot)\|^\varepsilon (1 + \|T_{qn}(\cdot)\|)^{\varepsilon+2\beta} \right) \\
&\leq \sup_x \mathbb{E}_x \left(\|T_{qn}(\cdot)\|^\varepsilon \right) + \sup_x \mathbb{E}_x \left(\|T_{qn}(\cdot)\|^{\varepsilon+2\beta} \right) \\
&\leq r^{q\varepsilon} + r^{q(\varepsilon+2\beta)}.
\end{aligned}$$

By fixing q sufficiently large we can ensure that $r^{q\varepsilon} + r^{q(\varepsilon+2\beta)} < 1$. Let $n_0' = qn_0$ for this value of q. As $d_\lambda(T_n(\mathbf{j}))^{2\beta} < 1 + \|T_n(\mathbf{j})\|^{2\beta} + \lambda^{2\beta} d(T_n(\mathbf{j})x_0, x_0)^{2\beta}$, by choosing λ sufficiently small we can ensure that

$$\sup_x \mathbb{E}_x \left(\|T_{n_0'}(\cdot)\|^\varepsilon d_\lambda(T_{n_0'}(\cdot))^{2\beta} \right) < r' < 1 \tag{15}$$

for some $r' < 1$. We now fix λ as in (15).

The following result is well-known.

Lemma 3.2
The space $C_{\varepsilon,\beta,\gamma}$ is a Banach space with respect to the norm $\|\cdot\|_{\varepsilon,\beta,\gamma}$.

Motivated by [HH1, HH2], we introduce the following norm: for $w \in C_{\varepsilon,\beta,\gamma}$ define

$$\|w\|^{(1)}_{\varepsilon,\beta} = \int |w|\, d\nu + |w|_{\varepsilon,\beta}.$$

We shall write $|w|^{(1)} = \int |w|\, d\nu$.

Lemma 3.3
The space $C_{\varepsilon,\beta,\gamma}$ is a Banach space when equipped with the norm $\|\cdot\|^{(1)}_{\varepsilon,\beta}$. Moreover, the norms $\|\cdot\|^{(1)}_{\varepsilon,\beta}$ and $\|\cdot\|_{\varepsilon,\beta,\gamma}$ are equivalent.

Proof. We prove that $C_{\varepsilon,\beta,\gamma}$ is a Banach space with respect to $\|\cdot\|^{(1)}_{\varepsilon,\beta}$. Let $w_n \in C_{\varepsilon,\beta,\gamma}$ be a Cauchy sequence with respect to $\|\cdot\|^{(1)}_{\varepsilon,\beta}$. Let $y_0 \in X$ and define $v_n(x) = w_n(x) - w_n(y_0)$. Then

$$\begin{aligned}
|v_n(x) - v_m(x)| &= |(w_n(x) - w_m(x)) - (w_n(y_0) - w_m(y_0))| \\
&\leq |w_n - w_m|_{\varepsilon,\beta} d(x, y_0)^\varepsilon d_\lambda x^\beta d_\lambda(y_0)^\beta
\end{aligned}$$

so that

$$\nu(|v_n - v_m|) \leq |w_n - w_m|_{\varepsilon,\beta} d_\lambda(y_0)^\beta \int d(x, y_0)^\varepsilon d_\lambda(x, x_0)\, d\nu. \tag{16}$$

As the integrand in (16) is $O(d(x, y_0)^{\varepsilon+\beta})$ and $\varepsilon + \beta < \theta_0$, it follows that $\nu(|v_n - v_m|) \leq C|w_n - w_m|_{\varepsilon,\beta}$ for some constant $C > 0$. Hence v_n is a

Cauchy sequence in the Banach space $L^1(\nu)$ of L^1 functions with respect to ν. Hence v_n converges in $L^1(\nu)$ to, say, $v \in L^1(\nu)$ and $\nu(v_n) \to \nu(v)$.

As w_n is a Cauchy sequence with respect to $\| \cdot \|^{(1)}_{\varepsilon,\beta}$, we see that $\nu(w_n)$ is a Cauchy sequence of complex numbers, and so converges. Note that for all x, $w_n(y_0) = w_n(x) - w_n(x)$. Integrating this with respect to ν we obtain $w_n(y_0) = \nu(w_n) - \nu(v_n)$. Hence for each y_0, $w_n(y_0)$ converges. Hence w_n converges pointwise to some function, say w. As $\nu(w_n)$ converges, it follows from the Dominated Convergence Theorem that $\nu(w_n) \to \nu(w)$.

It remains to check that $w \in C_{\varepsilon,\beta,\gamma}$. Given $\epsilon > 0$, choose N such that if $n, m \geq N$ then $|w_n - w_m|_{\varepsilon,\beta} < \epsilon$. Letting $m \to \infty$ implies that $|w_n - w|_{\varepsilon,\beta} \leq \epsilon$. Also note that

$$|w_n(x) - w_n(x_0)| \leq |w_n|_{\varepsilon,\beta} d(x, x_0)^\varepsilon d_\lambda(x)^\beta;$$

letting $n \to \infty$ and dividing by $d_\lambda(x)^\gamma$ it follows that

$$\frac{|w(x)|}{d_\lambda(x)^\gamma} \leq \frac{|w(x_0)|}{d_\lambda(x)^\gamma} + |w|_{\varepsilon,\beta} \frac{d(x, x_0)^\varepsilon d_\lambda(x)^\beta}{d_\lambda(x)^\gamma}.$$

As $\varepsilon + \beta < \gamma$, we have that $d(x, x_0)^\varepsilon d_\lambda(x)^\beta / d_\lambda(x)^\gamma \leq C$, for some constant $C > 0$. Hence $|w|_\gamma < \infty$.

We prove that the two norms $\| \cdot \|_{\varepsilon,\beta,\gamma}$, $\| \cdot \|^{(1)}_{\varepsilon,\beta}$ on $C_{\varepsilon,\beta,\gamma}$ are equivalent. For $w \in C_{\varepsilon,\beta,\gamma}$ note that

$$\begin{aligned} |w(y)| &\leq |w(x)| + |w|_{\varepsilon,\beta} d(x, y)^\varepsilon d_\lambda(x)^\beta d_\lambda(y)^\beta \\ &\leq |w(x)| + |w|_{\varepsilon,\beta} d(x, y)^\varepsilon (d(x, x_0)^\varepsilon d_\lambda(x)^\beta d_\lambda(y)^\beta \\ &\quad + d(y, x_0)^\varepsilon d_\lambda(x)^\beta d_\lambda(y)^\beta). \end{aligned}$$

Hence there exists a constant $C > 0$ such that

$$|w(y)| \leq \nu(|w|) + |w|_{\varepsilon,\beta} d_\lambda(y)^\beta C \left(\nu(d(x, x_0)^{\varepsilon+\beta} + d(x, x_0)^\beta) \right).$$

As $\beta < \gamma$, it follows that

$$\frac{|w(y)|}{d_\lambda(y)^\gamma} \leq \frac{|w(y)|}{d_\lambda(y)^\beta} \leq \nu(|w|) + C' |w|_{\varepsilon,\beta}$$

for some constant $C' > 0$. Hence $\|w\|_{\varepsilon,\beta,\gamma} \leq \nu(|w|) + (C' + 1)|w|_{\varepsilon,\beta} \leq (C' + 1)\|w\|^{(1)}_{\varepsilon,\beta}$.

As $C_{\varepsilon,\beta,\gamma}$ is a Banach space with respect to both $\| \cdot \|_{\varepsilon,\beta,\gamma}$ and $\| \cdot \|^{(1)}_{\varepsilon,\beta}$ it follows from the Open Mapping Theorem that there exists a constant $C'' > 0$ such that $\|w\|^{(1)}_{\varepsilon,\beta} \leq C'' \|w\|_{\varepsilon,\beta,\gamma}$. ❑

We shall need the following result.

Lemma 3.4
The inclusion $\iota : (C_{\varepsilon,\beta,\gamma}, \|\cdot\|^{(1)}_{\varepsilon,\beta}) \hookrightarrow (C_{\varepsilon,\beta,\gamma}, |\cdot|^{(1)})$ *is compact.*

Proof. Let $w_n \in C_{\varepsilon,\beta,\gamma}$ and suppose that $\|w_n\|^{(1)}_{\varepsilon,\beta} \leq 1$. As $|w_n|_{\varepsilon,\beta} < 1$, it follows that w_n is equicontinuous on every compact subset of X. By a diagonalisation argument, there exists a subsequence $n_k \to \infty$ such that w_{n_k} converges uniformly to $w \in C_{\varepsilon,\beta,\gamma}$, $\|w\|^{(1)}_{\varepsilon,\beta} \leq 1$, on every compact set. As $|w(x)| \leq d_\lambda(x)^\gamma$, it follows from the Dominated Convergence Theorem that $|w - w_{n_k}|^{(1)} = \nu(|w_{n_k} - w|) \to 0$. ❏

§4 Spectral properties of a family of transfer operators

Define the operator P on continuous functions w by

$$Pw(x) = \sum_j p_j(x) w(T_j x).$$

Then P maps the space of continuous functions to itself and $P1 = 1$, so that 1 is an eigenvalue for P. For stronger spectral properties of P to hold we need to restrict P to $C_{\varepsilon,\beta,\gamma}$. The following result is proved in [Pe], albeit on a slightly different function space; we sketch the argument in §4.1 below.

Proposition 4.1 ([Pe])
Under the technical hypotheses in §3.1, the operator P maps $C_{\varepsilon,\beta,\gamma}$ to itself, has 1 as a simple maximal eigenvalue with associated eigenprojection ν, and the remainder of the spectrum is contained within a disc of radius $\rho < 1$.

We will need to study the spectral properties of the following one-parameter family of perturbed transfer operators. Fix a continuous function f with $|f|_{(\alpha)} < \infty$. For each $t \in \mathbb{R}$ define

$$P_t w = P(e^{itf} w) = \sum_j p_j(x) e^{itf(T_j x)} w(T_j x).$$

We shall see below that P_t maps $C_{\varepsilon,\beta,\gamma}$ to itself. The relevance of P_t to the sums of observations (4) is given by observing that

$$P_t^n w(x) = \mathbb{E}_x \left(e^{itS_n f(x,\cdot)} w(Z_n(x,\cdot)) \right). \tag{17}$$

We will prove that, for sufficiently small t, P_t has a unique simple maximal eigenvalue λ_t with corresponding eigenprojection π_t and a spectral gap so that the remainder of the spectrum is contained within a disc of radius ρ_t. We will do this by establishing a Lasota-Yorke inequality for P_t and citing a result of Hennion [H]. We also want to determine the continuity properties of λ_t, π_t, ρ_t, etc, as t varies. To do this, we will apply a theorem of Keller and Liverani [KL].

§4.1 A Lasota-Yorke inequality

We shall need the following estimates.

Lemma 4.2
(i) Let $\beta \in (0,1)$. Then

$$B_n(\beta) = \sup_{x,y} \mathbb{E}_x \left(\frac{d_\lambda(Z_n(y,\cdot))^\beta}{d_\lambda(y)^\beta} \right) < \infty.$$

(ii) We have

$$A(\gamma,\varepsilon) = \sup_{x,y\in X, x\neq y} \sum_j \frac{d_\lambda(T_j x)^\gamma}{d_\lambda(x)^\gamma} \frac{|p_j(x) - p_j(y)|}{d(x,y)^\varepsilon} < \infty.$$

Proof. We prove (i). Let n_0 be as in (9) and let

$$R = \sup_{\ell=0,\ldots,n_0-1} \sup_{x,y,z,y\neq z} \mathbb{E}_x \left(\frac{d(Z_\ell(y,\cdot), Z_\ell(z,\cdot))}{d(y,z)} \right) < \infty.$$

Then $\mathbb{E}_x(d(Z_n(y,\cdot), Z_n(z,\cdot))) \le R r^{\left\lfloor \frac{n}{n_0} \right\rfloor} = R\rho^n$ for some $\rho \in (0,1)$, increasing R slightly if necessary. Then

$$\begin{aligned}
&\mathbb{E}_x \left(d_\lambda(Z_k(y,\cdot))^\beta \right) \\
&\le 1 + \lambda^\beta\, \mathbb{E}_x \left(d(Z_k(y,\cdot), Z_k(x_0,\cdot)) \right)^\beta + \lambda^\beta\, \mathbb{E}_x \left(d(Z_k(x_0,\cdot), x_0) \right)^\beta . \quad (18)
\end{aligned}$$

Now

$$\begin{aligned}
&\mathbb{E}_x \left(d(Z_k(x_0,\cdot), x_0) \right)^\beta \\
&\le \sum_{\ell=1}^k \sum_{j_\ell,\cdots,j_k} \mu_{T_{j_{\ell-1}}\cdots T_{j_1}(x_0)}[j_\ell,\ldots,j_k] d(T_{j_k}\cdots T_{j_{\ell+1}}(T_\ell x_0), T_{j_k}\cdots T_{j_{\ell+1}}(x_0))^\beta \\
&\le \sum_{\ell=1}^k R\rho^{k-\ell} \sup_x \mathbb{E}_x d(Z_1(x_0,\cdot), x_0),
\end{aligned}$$

which is bounded by a constant, by (10). The middle term in the right-hand side of (18) is bounded by $\lambda^\beta R^\beta \rho^{n\beta} d(y,x_0)^\beta$. It follows that $B_n(\beta) < \infty$.

We prove (ii). If $d(x,y) \ge 1$ then $1/d(x,y) \le 1$ so that

$$\begin{aligned}
A(\gamma,\varepsilon) &\le \sup_{x,y\in X, x\neq y} \sum_j \frac{d_\lambda(T_j x)^\gamma}{d_\lambda(x)^\gamma} p_j(x) + \frac{d_\lambda(T_j x)^\gamma}{d_\lambda(x)^\gamma} p_j(y) \\
&\le \sup_{x,y} \mathbb{E}_x(\delta(T_j)^\gamma) + \mathbb{E}_y(\delta(T_j)^\gamma) \\
&\le 2B_1(\gamma).
\end{aligned}$$

If $d(x,y) \leq 1$ then, recalling that $m(p_j) = \sup_{x,y \in X, d(x,y) \leq 1} |p_j(x) - p_j(y)|/d(x,y)$, we have

$$\begin{aligned}
\sum_j \frac{d_\lambda(T_j x)^\gamma}{d_\lambda(x)^\gamma} \frac{|p_j(x) - p_j(y)|}{d(x,y)^\varepsilon} &= \sum_j \frac{d_\lambda(T_j x)^\gamma}{d_\lambda(x)^\gamma} \frac{|p_j(x) - p_j(y)|}{d(x,y)} d(x,y)^{1-\varepsilon} \\
&\leq \sum_j \delta_\lambda(T_j)^\gamma m(p_j).
\end{aligned}$$

Hence (ii) follows from (i) and (11). ❑

We can now prove a Lasota-Yorke inequality for P_t.

Proposition 4.3
There exist constants $R_n > 0$ such that for all $w \in C_{\varepsilon,\beta,\gamma}$ we have

$$|P_t^n w|_{\varepsilon,\beta} \leq \sup_x \mathbb{E}_x \left(\|T_n(\cdot)\|^\varepsilon \delta_\lambda(T_n(\cdot))^{2\beta} \right) |w|_{\varepsilon,\beta} + R_n |t|^{\varepsilon/\alpha} |w|_\gamma.$$

Proof. Let $w \in C_{\varepsilon,\beta,\gamma}$.

First note that

$$\begin{aligned}
|P_t^n w(x)| &\leq \sum_{j_1,\dots,j_n} \mu_x[j_1,\dots,j_n] |w(Z_n(x,\mathbf{j}))| \\
&\leq \sum_{j_1,\dots,j_n} \mu_x[j_1,\dots,j_n] |w|_\gamma d_\lambda(Z_n(x,\mathbf{j}))^\gamma \\
&\leq \sum_{j_1,\dots,j_n} \mu_x[j_1,\dots,j_n] |w|_\gamma \delta_\lambda(T_n(\mathbf{j}))^\gamma d_\lambda(x)^\gamma
\end{aligned}$$

so that

$$|P_t^n w|_\gamma \leq \mathbb{E}_x \left(\delta_\lambda(T_n(\cdot))^\gamma \right) |w|_\gamma.$$

Note that

$$\mathbb{E}_x \left(\delta_\lambda(T_n(\cdot))^\gamma \right) \leq 1 + \mathbb{E}_x(\|T_n(\cdot)\|)^\gamma + \lambda^\gamma \, \mathbb{E}_x \left(d(T_j x_0, x_0)^\gamma \right) < \infty$$

as $\gamma < 1$ and (10). Hence there exists $M > 0$ such that $|P_t^n|_\gamma \leq M^n$.

Let $x, y \in X$ and assume, without loss of generality, that $d_\lambda(y) \leq d_\lambda(x)$. We can write

$$\begin{aligned}
&|P_t^n w(x) - P_t^n w(y)| \\
&= \left| \sum_{j_1,\dots,j_n} \left(\mu_x[j_1,\dots,j_n] e^{itS_n f(x,\mathbf{j})} w(Z_n(x,\mathbf{j})) - \mu_y[j_1,\dots,j_n] e^{itS_n f(y,\mathbf{j})} w(Z_n(y,\mathbf{j})) \right) \right| \\
&\leq \Sigma_w^{(n)} + \Sigma_f^n + \Sigma_\mu^n
\end{aligned}$$

where

$$\begin{aligned}
\Sigma_w^{(n)} &= \left| \sum_{j_1,\ldots,j_n} \mu_x[j_1,\ldots,j_n] e^{itS_nf(x,\mathbf{j})} \left(w(Z_n(x,\mathbf{j})) - w(Z_n(y,\mathbf{j}))\right) \right|, \\
\Sigma_f^{(n)} &= \left| \sum_{j_1,\ldots,j_n} \mu_x[j_1,\ldots,j_n] \left(e^{itS_nf(x,\mathbf{j})} - e^{itS_n(y,\mathbf{j})}\right) w(Z_n(y,\mathbf{j})) \right|, \\
\Sigma_\mu^{(n)} &= \left| \sum_{j_1,\ldots,j_n} \left(\mu_x[j_1,\ldots,j_n] - \mu_y[j_1,\ldots,j_n]\right) e^{itS_nf(y,\mathbf{j})} w(Z_n(y,\mathbf{j})) \right|.
\end{aligned}$$

Now

$$\begin{aligned}
\Sigma_w^{(n)} &\leq \sum_{j_1,\ldots,j_n} \mu_x[j_1,\ldots,j_n] |w(Z_n(x,\mathbf{j})) - w(Z_n(y,\mathbf{j}))| \\
&\leq \sum_{j_1,\ldots,j_n} \mu_x[j_1,\ldots,j_n] |w|_{\varepsilon,\beta} d(Z_n(x,\mathbf{j}), Z_n(y,\mathbf{j}))^\varepsilon d_\lambda(Z_n(x,\mathbf{j}))^\beta d_\lambda(Z_n(y,\mathbf{j}))^\beta \\
&\leq |w|_{\varepsilon,\beta}\, \mathbb{E}_x \left(\|T_n(\cdot)\|^\varepsilon \delta_\lambda(T_n(\cdot))^{2\beta}\right) d(x,y)^\varepsilon d_\lambda(x)^\beta d_\lambda(y)^\beta. \qquad (19)
\end{aligned}$$

We can write

$$\begin{aligned}
\Sigma_f^{(n)} &\leq \sum_{j_1,\ldots,j_n} \mu_x[j_1,\ldots,j_n] \left| e^{itS_nf(x,\mathbf{j})} - e^{itS_nf(y,\mathbf{j})} \right| |w(Z_n(x,\mathbf{j}))| \\
&\leq |w|_\gamma \sum_{j_1,\ldots,j_n} \mu_x[j_1,\ldots,j_n] \left| e^{it(S_nf(x,\mathbf{j}) - S_nf(y,\mathbf{j}))} - 1 \right| d_\lambda(Z_n(y,\mathbf{j}))^\gamma \\
&\leq |w|_\gamma \sum_{k=1}^n \Sigma_f^{(n),k}
\end{aligned}$$

where

$$\Sigma_f^{(n),k} = \sum_{j_1,\ldots,j_n} \mu_x[j_1,\ldots,j_n] \left| e^{it(f(Z_k(x,\mathbf{j})) - f(Z_k(y,\mathbf{j})))} - 1 \right| d_\lambda(Z_n(y,\mathbf{j}))^\gamma.$$

Recall that for any $\eta > 0$ we have that $|e^{ix} - 1| < \max\{2, |x|^\eta\} \leq 2|x|^\eta$. Recalling that $|f|_{(\alpha)} < \infty$, by taking $\eta = \varepsilon/\alpha$ we can bound

$$\begin{aligned}
\Sigma_f^{(n),k} &\leq 2|t|^{\varepsilon/\alpha} |f|_{(\alpha)}^{\varepsilon/\alpha} \sum_{j_1,\ldots,j_n} \mu_x[j_1,\ldots,j_n] d(Z_k(x,\mathbf{j}), Z_k(y,\mathbf{j}))^\varepsilon d_\lambda(Z_n(x,\mathbf{j}))^\gamma \\
&\leq 2|t|^{\varepsilon/\alpha} |f|_{(\alpha)}^{\varepsilon/\alpha} \sum_{j_1,\ldots,j_k} \mu_x[j_1,\ldots,j_k] d(Z_k(x,\mathbf{j}), Z_k(y,\mathbf{j}))^\varepsilon d_\lambda(Z_k(x,\mathbf{j}))^\gamma \\
&\qquad \times \sum_{j_{k+1},\ldots,j_n} \mu_{Z_k(x,\mathbf{j})}[j_{k+1},\ldots,j_n] \frac{d_\lambda(T_{j_n}\cdots T_{j_{k+1}}(Z_k(y,\mathbf{j})))^\gamma}{d_\lambda(Z_k(y,\mathbf{j}))^\gamma} \\
&\leq 2|t|^{\varepsilon/\alpha} |f|_{(\alpha)}^{\varepsilon/\alpha} B_{n-k}(\gamma)\, \mathbb{E}_x \left(d(Z_k(x,\cdot), Z_k(y,\cdot))^\varepsilon d_\lambda(Z_k(y,\mathbf{j}))^\gamma\right) \\
&\leq 2|t|^{\varepsilon/\alpha} |f|_{(\alpha)}^{\varepsilon/\alpha} B_{n-k}(\gamma)\, \mathbb{E}_x \left(\|T_k(\cdot)\|^\varepsilon \delta_\lambda(T_k(\cdot))^\gamma\right) d(x,y)^\varepsilon d_\lambda(y)^\gamma.
\end{aligned}$$

As $d_\lambda(y) \le d_\lambda(x)$ and $\gamma < 2\beta$ it follows that

$$\Sigma_f^{(n),k} \le 2|t|^{\varepsilon/\alpha}|f|_{(\alpha)}^{\varepsilon/\alpha} B_{n-k}(\gamma)\,\mathbb{E}_x\left(\|T_k(\cdot)\|^\varepsilon \delta_\lambda(T_k(\cdot))^\gamma\right) d(x,y)^\varepsilon d_\lambda(x)^\beta d_\lambda(y)^\beta. \tag{20}$$

We can write

$$\begin{aligned}
\Sigma_\mu^{(n)} &\le \sum_{j_1,\ldots,j_n} |\mu_x[j_1,\ldots,j_n] - \mu_y[j_1,\ldots,j_n]|\,|w(Z_n(y,\mathbf{j}))| \\
&\le |w|_\gamma \sum_{k=1}^n \Sigma_\mu^{(n),k}
\end{aligned}$$

where

$$\begin{aligned}
\Sigma_\mu^{(n),k} &= \sum_{j_1,\ldots,j_n} \mu_{Z_k(x,\mathbf{j})}[j_{k+1},\ldots,j_n]\,|p_{j_k}(Z_{k-1}(x,\mathbf{j})) - p_{j_k}(Z_{k-1}(y,\mathbf{j}))| \\
&\quad \times \mu_y[j_1,\ldots,j_{k-1}] d_\lambda(Z_n(y,\mathbf{j}))^\gamma \\
&= \sum_{j_1,\ldots,j_{k-1}} \left(\sum_{j_{k+1},\ldots,j_n} \mu_{Z_k(x,\mathbf{j})}[j_{k+1},\ldots,j_n] \frac{d_\lambda(T_{j_n}\cdots T_{j_{k+1}}(Z_k(x,\mathbf{j})))^\gamma}{d_\lambda(Z_k(x,\mathbf{j}))^\gamma}\right) \\
&\quad \times \sum_{j_k} \frac{d_\lambda(T_{j_k}(Z_{k-1}(x,\mathbf{j}))^\gamma)}{d_\lambda(Z_{k-1}(x,\mathbf{j}))^\gamma} \frac{|p_{j_k}(Z_{k-1}(x,\mathbf{j})) - p_{j_k}(Z_{k-1}(y,\mathbf{j}))|}{d(Z_{k-1}(x,\mathbf{j}), Z_{k-1}(y,\mathbf{j}))^\varepsilon} \\
&\quad \mu_y[j_1,\ldots,j_{k-1}] d(Z_{k-1}(x,\mathbf{j}), Z_{k-1}(y,\mathbf{j}))^\varepsilon d_\lambda(Z_{k-1}(x,\mathbf{j}))^\gamma \\
&\le B_{n-k}(\gamma) A(\gamma,\varepsilon) \sum_{j_1,\ldots,j_{k-1}} \mu_y[j_1,\ldots,j_{k-1}] d(Z_{k-1}(x,\mathbf{j}), Z_{k-1}(y,\mathbf{j}))^\varepsilon d_\lambda(Z_{k-1}(x,\mathbf{j}))^\gamma \\
&\le B_{n-k}(\gamma) A(\gamma,\varepsilon)\,\mathbb{E}_y\left(\|T_{k-1}(\cdot)\|^\varepsilon \delta_\lambda(T_{k-1}(\cdot))^\gamma\right) d(x,y)^\varepsilon d_\lambda(y)^\gamma.
\end{aligned}$$

As $d_\lambda(y) \le d_\lambda(x)$ and $\gamma < 2\beta$ it follows that

$$\Sigma_\mu^{(n),k} \le B_{n-k}(\gamma) A(\gamma,\varepsilon)\,\mathbb{E}_y\left(\|T_{k-1}(\cdot)\|^\varepsilon \delta_\lambda(T_{k-1}(\cdot))^\gamma\right) d(x,y)^\varepsilon d_\lambda(x)^\beta d_\lambda(y)^\beta. \tag{21}$$

As $\varepsilon + \gamma < \varepsilon + 2\beta < 1$ and by Lemma 4.2, the right-hand sides of (19), (20), (21) are finite. Hence there exists a constant $R_n > 0$ such that

$$|P_t^n w|_{\varepsilon,\beta} \le \sup_x \mathbb{E}_x\left(\|T_n(\cdot)\|^\varepsilon \delta_\lambda(T_n(\cdot))^{2\beta}\right) |w|_{\varepsilon,\beta} + R_n |t|^{\varepsilon/\alpha} |w|_\gamma.$$

❏

Remark. By taking $n = 1$ in Proposition 4.3 we see that P_t maps $C_{\varepsilon,\beta,\gamma}$ into itself.

Consider the case $t = 0$. Taking $n = n_0'$ in Proposition 4.3 we have that

$$|P_t^{n_0'} w|_{\varepsilon,\beta} \le r'|w|_{\varepsilon,\beta} + R'|t|^{\varepsilon/\alpha}|w|_\gamma \tag{22}$$

for some $R' > 0$. It follows from Hennion's improvement [H] of the classical Ionescu-Marinescu-Tulcea theorem that $P^{n_0'}$, and hence P, is a quasi-compact operator. Hence we can decompose

$$P = \sum_{\lambda \in G} \lambda \pi_\lambda + Q \tag{23}$$

where G is a finite group of eigenvalues of modulus 1, π_λ is the eigenprojection onto the corresponding eigenspace, Q is the eigenprojection onto the remainder of the spectrum with $\rho(Q) < \rho$ for some $\rho \in (0,1)$, and $\pi_\lambda Q = Q\pi_\lambda = 0$. The following result shows that, under (12), 1 is the only eigenvalue of modulus 1 for P.

Lemma 4.4
The only eigenvalue λ of modulus 1 of $P : C_{\varepsilon,\beta,\gamma} \to C_{\varepsilon,\beta,\gamma}$ is $\lambda = 1$ and the only eigenfunctions are constants.

Proof. This is proved in [Pe], albeit on a slightly different function space. For completeness we sketch the argument. Clearly $P1 = 1$, so that 1 is an eigenvalue of P and the constants are eigenfunctions. We decompose P as in (23) and we show that $\pi_1(w) = \int w\, d\nu$ and $\pi_\lambda = 0$ for $\lambda \neq 1$.

Suppose that $Pw = w$, where $w \in C_{\varepsilon,\beta,\gamma}$ is bounded.

Sub-lemma 4.5
For each $x \in X$ we claim that there exists a set $\Sigma_x \subset \Sigma$ such that for all $\mathbf{j} \in \Sigma_x$ we have:

(i) $\lim_{n\to\infty} w(Z_n(x,\mathbf{j}))$ *exists,*

(ii) $\liminf_{n\to\infty} d(x_0, Z_n(x,\mathbf{j})) < \infty$,

(iii) for all $m \geq 1$ and all $i_1,\ldots,i_m$, we have that $w(Z_n(x,\mathbf{j})) = w(T_{i_m}\cdots T_{i_1} Z_n(x,\mathbf{j}))$ whenever $\mu_{Z_n(x,\mathbf{j})}[i_1,\ldots,i_m] \neq 0$.

Let $x, y \in X$. For each $\mathbf{j} \in \Sigma_x$, by (ii) we can choose a sequence $n_k = n_k(\mathbf{j})$ such that $Z_{n_k}(x,\mathbf{j}) \to x_\mathbf{j}$ for some $x_\mathbf{j} \in X$. As w is continuous, by (iii) we have that $w(x_\mathbf{j}) = w(T_{i_m}\cdots T_{i_1} x_\mathbf{j})$ whenever $\mu_{x_\mathbf{j}}[i_1,\ldots,i_m] \neq 0$. Similarly, for each $\mathbf{j}' \in \Sigma_y$, we can find $y_{\mathbf{j}'} \in X$ such that $w(y_{\mathbf{j}'}) = w(T_{i_m}\cdots T_{i_1} y_{\mathbf{j}'})$ whenever $\mu_{y_{\mathbf{j}'}}[i_1,\ldots,i_m] \neq 0$.

By (12), choose $\mathbf{i} = (i_1, i_2, \ldots) \in \Sigma$ such that for all m, $\mu_{x_\mathbf{j}}[i_1,\ldots,i_m]$, $\mu_{y_{\mathbf{j}'}}[i_1,\ldots,i_m] \neq 0$. Then, as the p_j are continuous, (iii) is satisfied for sufficiently large n_k. Hence

$$\begin{aligned}
&|w(x_\mathbf{j}) - w(y_j)| \\
&\quad = |w(T_{i_m}\cdots T_{i_1} x_\mathbf{j}) - w(T_{i_m}\cdots T_{i_1} y_{\mathbf{j}'})| \\
&\quad \leq |w|_{\varepsilon,\beta} \|T_{i_m}\cdots T_{i_1}\|^\varepsilon d_\lambda(T_{i_m}\cdots T_{i_1})^{2\beta} d(x_\mathbf{j}, y_{\mathbf{j}'})^\varepsilon d_\lambda(x_\mathbf{j})^\beta d_\lambda(y_{\mathbf{j}'})^\beta \\
&\quad \leq C(x_\mathbf{j}, y_{\mathbf{j}'})\, \mathbb{E}_x \left(\|T_m(\mathbf{i})\|^\varepsilon \delta_\lambda(T_m(\mathbf{i}))^{2\beta} \right).
\end{aligned}$$

Taking $m = qn_0'$ where n_0' is given by (15) we see that this expression is bounded by $C(x_{\mathbf{j}}, y_{\mathbf{j}'})(r')^q$. Letting $q \to \infty$ shows that $w(x_{\mathbf{j}}) = w(y_{\mathbf{j}'})$ whenever $\mathbf{j} \in \Sigma_x$ and $\mathbf{j}' \in \Sigma_y$. As $w(x_{\mathbf{j}}) = \lim_{n_k} w(Z_{n_k}(x, \mathbf{j}))$ and $w(x) = \mathbb{E}_x(w(x_{x_{\mathbf{j}}}))$, and similarly for y, it follows that $w(x) = w(y)$. Hence w is constant.

Let $M_n = n^{-1} \sum_{k=0}^{n-1} P^k$. Then $\|M_n - \pi_1\| \to 0$ as $n \to \infty$ (here $\|\cdot\|$ denotes the operator norm on $C_{\varepsilon,\beta,\gamma}$). As $|M_n w|_\infty \leq |w|_\infty$, it follows that $\pi_1 w$ is bounded. By the above, $\pi_1 w$ is constant.

If $Pw = w$ and $w \in C_{\varepsilon,\beta,\gamma}$ is not bounded, then define $w_c(x) = w(x)$ if $|w(x)| \leq c$ and $w_c(x) = cw(x)/|w(x)|$ if $|w(x)| > c$. Then $w_c \in C_{\varepsilon,\beta,\gamma}$ is bounded. Hence $\pi_1 w_c$ is constant. Note that $|M_n w_c(x) - \pi_1 w_c(x)| \leq \|M_n - \pi_1\| \|w_c\|_{\varepsilon,\beta,\gamma} d_\lambda(x)^\gamma$. By dominated convergence, $M_n w_c \to M_n w$ as $c \to \infty$. Hence $\pi_1 w_c \to \pi_1 w$ as $c \to \infty$, so that $\pi_1 w$ is constant.

It remains to prove Sublemma 4.5.

Proof of Sublemma 4.5. Let $\mathcal{B}_n$ denote the sub-σ-algebra of Σ generated by cylinder sets of length n. Then it is straightforward to check from the definitions that if $Pw = w$ then

$$\mathbb{E}_x(w(Z_n(x, \mathbf{j})) \mid \mathcal{B}_n) = P(w(Z_n(x, \mathbf{j}))) = w(Z_n(x, \mathbf{j})).$$

Hence $w(Z_n(x, \cdot))$ is a martingale with respect to $\mathcal{B}_n$. Property (i) then follows from standard properties of martingales.

Property (ii) is a simple consequence of the contraction on average assumption and (10).

Note that

$$\begin{aligned}
&\mathbb{E}_x\left(\sum_{n=0}^{\infty} \sum_{i_1,\ldots,i_m} \mu_{Z_n(x,\cdot)}[i_1, \ldots, i_m] \left(w(T_{i_m} \cdots T_{i_1} Z_n(x, \cdot)) - w(Z_n(x, \cdot))\right)^2\right) \\
&= \sum_{n=0}^{\infty} \mathbb{E}_x\left(\sum_{i_1,\ldots,i_m} \mu_{Z_n(x,\cdot)}[i_1, \ldots, i_m] \left(w(T_{i_m} \cdots T_{i_1} Z_n(x, \cdot)) - w(Z_n(x, \cdot))\right)^2\right) \\
&= \sum_{n=0}^{\infty} \mathbb{E}_x(w(Z_{n+m}(x, \cdot))^2) - \mathbb{E}_x(w(Z_n(x, \cdot))^2) \\
&\leq 2m|w|_\infty
\end{aligned}$$

where we have used the fact that $\mathbb{E}_x((w(Z_{n+m}(x, \cdot)) - w(Z_n(x, \cdot)))^2) = \mathbb{E}_x(w(Z_{n+m}(x, \cdot))^2) - \mathbb{E}_x(w(Z_n(x, \cdot))^2)$. Hence the first integrand in the expression above is finite μ_x-a.e., hence the summand converges to zero μ_x-a.e. ❑

Now suppose that $Pw = \lambda w$ where $\lambda \in G$, $w \in C_{\varepsilon,\beta,\gamma}$. Introduce a new operator $\hat{P}$ defined on $X \times \mathbb{N}$ by

$$\hat{P}W(x, n) = \sum_j p_j(x) W(T_j x, n + 1).$$

By taking $W(x,n) = \lambda^{-n} w(x)$, we see that $\hat{P}W = W$ if and only if $Pw = \lambda w$. A similar argument to that above then shows that w is zero. ❑

§4.2 Perturbation of the spectrum of P_t

We want to study how the spectrum of P_t behaves for t in a neighbourhood of 0. To do this, we use the following perturbation result of Keller and Liverani:

Theorem 4.6 ([KL])
Suppose that $(B, \|\cdot\|)$ is a Banach space equipped with a second norm $|\cdot| \leq \|\cdot\|$ (we do not require $(B, |\cdot|)$ to be complete). Let P_t be a one-parameter family of bounded linear operators. Suppose that

(i) the inclusion $\iota : (B, \|\cdot\|) \hookrightarrow (B, |\cdot|)$ is compact;

(ii) there exists an interval J', $0 \in J'$, and $n_0 > 0$ such that for $t \in J'$, $P_t^{n_0}$ satisfies a uniform Lasota-Yorke condition: there exists $\rho \in (0,1)$ and $R > 0$ such that for all $w \in B$, $t \in J$,

$$\|P_t^{n_0} w\| \leq \rho \|w\| + R|w|;$$

(iii) $|||P_t - P_0||| \to 0$ as $t \to 0$ (here, if $Q : B \to B$ is a bounded linear operator, then $|||Q||| = \sup_{w \in B} |Qw|/\|w\|$).

Then there exists an interval $J \subset J'$ containing 0 such that if $t \in J$ then P_t is quasi-compact. Suppose in addition that P_0 has a unique simple maximal eigenvalue at 1. Then P_t has a unique simple maximal eigenvalue λ_t with corresponding eigenprojection π_t. For $t \in J$, the dependence $t \mapsto \lambda_t$ is continuous, and $|||\pi_t - \pi_0||| \to 0$. Moreover, there exists $\rho_0 < 1$ such that if Q_t denotes the projection onto the remainder of the spectrum of P_t, then Q_t has spectral radius $\rho(Q_t) \leq \rho_0$, for all $t \in J$.

To apply Theorem 4.6, we need to study how $P_t - P$ varies in an appropriate norm. As is observed in a similar context in [HH2],

$$\begin{aligned} |P_t w(x) - w(x)| &\leq \sum_j p_j(x) |e^{itf(T_j x)} - 1| |w(T_j x)| \\ &\leq \sum_j p_j(x) |t|^{\varepsilon} |f(T_j x)|^{\varepsilon} |w|_{\gamma} \delta_{\lambda}(T_j)^{\gamma} d_{\lambda}(x)^{\gamma}. \end{aligned}$$

If X is not compact, then $d_\lambda(x)$ is unbounded. Hence $t \mapsto P_t$ is not a priori continuous in the $|\cdot|_\gamma$-topology. For this reason we work with the norms $\|\cdot\|_{\varepsilon,\beta}^{(1)}$ and $|\cdot|^{(1)}$.

By Lemma 3.4, hypothesis (i) of Theorem 4.6 holds.

For an operator $Q : C_{\varepsilon,\beta,\gamma} \to C_{\varepsilon,\beta,\gamma}$ we define

$$|||Q||| = \sup\left\{\frac{|Qw|^{(1)}}{\|w\|^{(1)}_{\varepsilon,\beta}} \mid w \in C_{\varepsilon,\beta,\gamma}\right\}.$$

Lemma 4.7
There exists a constant $C > 0$ such that

$$|||P_t - P||| \leq C|t|^{\varepsilon}.$$

Proof. Note that if $w \in C_{\varepsilon,\beta,\gamma}$ then

$$\begin{aligned} |P_t w - Pw| &\leq \sum_j p_j(x)|e^{itf(T_j x)} - 1||w(T_j x)| \\ &\leq \sum_j p_j(x)|t|^{\varepsilon}|f(T_j x)|^{\varepsilon}|w|_{\gamma}\delta_{\lambda}(T_j)^{\gamma}d_{\lambda}(x)^{\gamma}. \end{aligned}$$

From (13) we have that

$$|f(T_j x)|^{\varepsilon} \leq |f(x_0)|^{\varepsilon} + |f|^{\varepsilon}_{(\alpha)} d(T_j x, x_0)^{\varepsilon\alpha} \leq C\delta_{\lambda}(T_j)^{\varepsilon\alpha} d_{\lambda}(x)^{\varepsilon\alpha}$$

for some constant $C > 0$. Hence

$$\nu(|P_t w - w|) \leq C|w|_{\gamma}|t|^{\varepsilon}\,\mathbb{E}_x\left(\delta(T_j x)^{\varepsilon\alpha+\gamma}\right)\int d_{\lambda}(x)^{\varepsilon\alpha+\gamma}\,d\nu. \tag{24}$$

By definition, $|w|_{\gamma} < \|w\|_{\varepsilon,\beta,\gamma}$. By Lemma 3.3 and the fact that (14) implies that $\varepsilon\alpha + \gamma < 1$ so that the two integrals in (24) are finite, it follows that $\nu(|P_t w - w|) \leq C'|t|^{\varepsilon}\|w\|^{(1)}_{\varepsilon,\beta}$. ❑

Hence hypothesis (ii) of Theorem 4.6 holds.

By Lemma 3.3 there exists $C > 0$ such that $|\cdot|_{\gamma} \leq C(|\cdot|_{\varepsilon,\beta} + |\cdot|^1)$. From (22) we have that

$$|P_t^{n_0} w|_{\varepsilon,\beta} \leq (r' + CR'|t|^{\varepsilon/\alpha})|w|_{\varepsilon,\beta} + CR'|t|^{\varepsilon/\alpha}|w|^{(1)}.$$

Choose ρ with $r' < \rho < 1$. Then there exists an interval J' containing 0 such that if $t \in J'$ then $r' + CR'|t|^{\varepsilon/\alpha} < \rho$. Hence hypothesis (iii) of Theorem 4.6 holds.

By Lemma 4.4, P has a simple maximal eigenvalue at 1. Hence by Theorem 4.6 there exists an interval J, $0 \in J$, and $\rho_0 < 1$ such that for $t \in J$ we can write

$$P_t = \lambda_t \pi_t + Q_t$$

where $t \mapsto \lambda_t$ is continuous, π_t is a one-dimensional projection operator with $\lim_{t\to 0}|||\pi_t - \nu||| = 0$, and Q_t is the projection onto the remainder of the spectrum and has spectral radius at most ρ_0. As $\pi_t Q_t = Q_t \pi_t = 0$, it follows that $P_t^n = \lambda_t^n \pi_t + Q_t^n$.

§5 Stable laws

Stable distributions are generalisations of the Gaussian distribution. Given a sequence of normally distributed independent random variables X_i it is well-known that the partial sums $S_n = X_0 + \cdots + X_{n-1}$ have the property that $n^{-1/2}S_n$ is also normally distributed. The stable laws are characterised by this behaviour: a suitable rescaling of independent partial sums has the same distribution. More precisely, we have the following definition:

Definition. A distribution function F is called *stable* if, for any a_1, $a_2 > 0$ and any b_1, b_2, there exist constants $a > 0$ and b such that $F(a_1x + b_1) * F(a_2x + b_2) = F(ax + b)$, where $*$ denotes convolution.

It is known that four parameters completely determine a stable law [IL]: $Y_{p,\beta,b,c}$ denotes the stable law with *order* $p \in (0, 2)$, *symmetry* $\beta \in [-1, 1]$, *origin* $b \in \mathbb{R}$ and *scaling factor* $c > 0$. We often identify stable distributions up to translation and scaling, denoting them simply by $Y_{p,\beta}$.

The tail behaviour of a stable law is encoded in its order p. For $p \in (0, 1) \cup (0, 2)$, there are constants $c_1, c_2 \geq 0$ (not both zero) such that $P(Y_{p,\beta} > t) = (c_1 + o(1))t^{-p}$ and $P(Y_{p,\beta} < t) = (c_2 + o(1))t^{-p}$.

Recall that if a random variable has distribution F then the *characteristic function* ϕ of this distribution is given by

$$\phi(t) = \int_{-\infty}^{\infty} e^{itx}\, dF(x).$$

In general, an explicit formula for the density of a stable distribution is only known in a handful of special cases. However, explicit formulæ for their characteristic functions are known.

Theorem 5.1 ([IL])
A distribution Y is stable of order $p \in (0, 1) \cup (1, 2)$ if and only if its characteristic function $\phi_Y(t)$ can be written in the form

$$\phi_Y(t) = \exp\left(ibt - c|t|^p\left(1 - i\beta\,\mathrm{sign}(t)\tan\left(\frac{\pi}{2}p\right)\right)\right), \tag{25}$$

where β, b and c are constants.

There is a corresponding formula for stable laws of order 1 [AD1].

We will also need the notion of slowly varying function.

Definition. A function $l : \mathbb{R} \to \mathbb{R}$ is said to be *slowly varying* if for every $t \in \mathbb{R}$

$$\lim_{x\to\infty} \frac{l(tx)}{l(x)} = 1.$$

Definition. We say that a function f (with distribution F) is in the *domain of attraction* of a stable law of order $p \in (0, 2)$ if there is a slowly varying function $l : \mathbb{R} \to \mathbb{R}$ and constants $c_1, c_2 \geq 0$, not both zero, such that

$$1 - F(t) = \frac{1}{t^p}(c_1 + o(1))l(t), \; F(-t) = \frac{1}{t^p}(c_2 + o(1))l(t) \tag{26}$$

as $t \to \infty$.

The expansion of the characteristic function of a function in the domain attraction of a stable law was studied by Ibragimov and Linnik for $p \neq 1$ [IL] and by Aaronson and Denker [AD1] for $p = 1$ in dimension one and for all p for multidimensional distributions. For simplicity, in what follows we will focus on real-valued functions and the case $p \in (0, 1) \cup (0, 2)$.

In §6 we relate the characteristic function of an observation f defined on the IFS to the expansion of the maximal eigenvalue of P_t. We shall use the following criterion.

Theorem 5.2 ([IL])
Let F be a distribution with characteristic function $\phi_F(t)$. Suppose that $p \neq 1$ and let $Y = Y_{p,\beta,b,c}$ be the stable law with characteristic function given by (25). Then a necessary and sufficient condition for F to be in the domain of attraction of the stable law Y is that in the neighbourhood of the origin

$$\phi_F(t) = \exp\left(ibt - c|t|^p l(t)\left(1 - i\beta \operatorname{sign}(t) \tan\left(\frac{\pi}{2}p\right)\right)\right),$$

where $l(t)$ is slowly varying as $t \to 0$.

§6 Expansion of the maximal eigenvalue

When $\sigma^2(f) < \infty$ then one has the following expansion of the maximal eigenvalue λ_t of P_t:

$$\lambda_t = \nu(f)t + \frac{\sigma^2(f)}{2}t^2 + o(t^2)$$

from which the Central Limit Theorem then follows. Under additional hypotheses, the $o(t^2)$ term can be improved and Berry-Esseen bounds can be proved. This is discussed in [HH2] in the case of random walks on semigroups of Lipschitz mappings that contract on average, with a place-independent probability.

When $\sigma^2(f)$ is infinite we have the following asymptotic expansion of λ_t.

Theorem 6.1 (Expansion of the maximal eigenvalue)
For $p \in (0, 1) \cup (1.2)$ suppose that f has distribution F and is in the domain of attraction of a stable law $Y = Y_{p,\beta,b,c}$ of order p. Let λ_t be the maximal

eigenvalue of the perturbed operator P_t. Then there is a slowly varying function l such that

$$\operatorname{Re}\log\lambda_t = -\operatorname{sign}(t)|t|^p c l(t^{-1}) + o(|t|^p l(t^{-1}))$$

and

$$\operatorname{Im}\log\lambda_t = tb - \operatorname{sign}(t)|t|^p c\beta \tan\left(\frac{p\pi}{2}\right) + o(t^p l(t^{-1})),$$

where $\beta = (c_2 - c_1)/(c_1 + c_2)$ with c_1, c_2 as in (26), b is given by

$$b = \begin{cases} 0 & \text{for } p \in (0,1), \\ \displaystyle\int_{-\infty}^{\infty} x\, dF(x) & \text{for } p \in (1,2) \end{cases}$$

and $c = (c_1 + c_2)\pi/2$.

Proof (sketch). The proof follows closely ideas in [AD2] (which in turn makes use of ideas in [N]). We indicate the modifications required.

Since f, with distribution F, is in the domain of attraction of a stable law $Y_{p,\beta}$ there is a slowly varying function $l : \mathbb{R} \to \mathbb{R}$ and constants $c_1, c_2 \geq 0$, not both zero, such that (26) holds.

The proof consists of estimating $1 - \lambda_t$, where λ_t is the maximal eigenvalue of P_t, and then using the fact that $\log(\lambda_t) = \log(1 + (\lambda_t - 1)) = (\lambda_t - 1) + O(|\lambda_t - 1|^2)$.

Let w_t denote the maximal eigenfunction of P_t corresponding to the eigenvalue λ_t so that $P_t w_t = \lambda_t w_t$. We normalise w_t so that $\int w_t\, d\nu = 1$. Let $\mathcal{F}$ denote $f^{-1}(\mathcal{B})$ where $\mathcal{B}$ denotes the σ-algebra of Borel subsets of $\mathbb{R}$. We define $\tilde{w}_t$ by $\tilde{w}_t \circ f = \mathbb{E}(w_t \mid \mathcal{F})$. Note that

$$\begin{aligned} \lambda_t &= \lambda_t \nu(w_t) = \nu(\lambda_t w_t) = \nu(P_t w_t) = \nu(P(e^{itf} w_t)) \\ &= \nu(e^{itf} w_t) = \nu(e^{itf}\, \mathbb{E}\,(w_t \mid \mathcal{F})) = \int e^{itx} \tilde{w}_t\, dF(x). \end{aligned}$$

Similarly,

$$1 = \int \tilde{w}_t\, dF(x)$$

so that $dF_t = \tilde{w}_t\, dF$ is a probability measure on $\mathbb{R}$. Hence

$$1 - \lambda_t = \int_{-\infty}^{\infty} (1 - e^{itx})\, dF_t(x).$$

The proof now proceeds as in [AD2, N]. ❏

§7 Distributional stable limit theorems

We are now in position to prove our main result.

Theorem 7.1
For $p \in (0,1) \cup (1,2)$, suppose that f is continuous, $|f|_{(\alpha)} < \infty$ and is in the domain of attraction of a stable law Y of order p, satisfying (26). Then for any $x \in X$,

$$\mu_x \left\{ \mathbf{j} \in \Sigma \;\middle|\; \frac{S_n f(x, \mathbf{j}) - a_n}{b_n} < t \right\} \to \int_{-\infty}^{t} dY,$$

as $n \to \infty$ with $a_n = 0$ if $p < 1$ and $a_n = n\nu(f)$ if $p > 1$, and $b_n^p = nl(b_n)$.

Proof. By replacing f by $f - \nu(f)$ there is no loss in generality in assuming that $a_n = 0$.

It is well-known that a sequence of random variables converges in distribution if and only if their corresponding characteristic functions converge pointwise at continuity points. The characteristic function of $Y = Y_{p,\beta}$ is

$$\mathbb{E}(e^{itY}) = \exp\left(-c|t|^p \left(1 - i\beta \operatorname{sign}(t) \tan\left(\frac{p\pi}{2}\right)\right)\right).$$

Hence it is sufficient to prove that

$$\mathbb{E}_x \left(\exp\left(it \frac{S_n f(x, \cdot)}{b_n} \right) \right) \to \exp\left(-c|t|^p \left(1 - i\beta \operatorname{sign}(t) \tan\left(\frac{p\pi}{2}\right)\right)\right)$$

as $n \to \infty$.

By §4.2, we can write $P_s^n = \lambda_s^n \pi_s + Q_s^n$ for $s \in J$, a neighbourhood of 0. Fix t. Then $t/b_n \in J$ for sufficiently large n. Recalling (17), we note that

$$\begin{aligned}
&\left| P^n_{\frac{t}{b_n}} 1(x) - \exp\left(-c|t|^p \left(1 - i\beta \operatorname{sign}(t) \tan\left(\frac{p\pi}{2}\right)\right)\right) \right| \\
&= \left| \lambda^n_{\frac{t}{b_n}} \pi_{\frac{t}{b_n}}(1)(x) + Q^n_{\frac{t}{b_n}}(1)(x) - \exp\left(-c|t|^p \left(1 - i\beta \operatorname{sign}(t) \tan\left(\frac{p\pi}{2}\right)\right)\right) \right| \\
&\leq \left| \lambda^n_{\frac{t}{b_n}} \pi_{\frac{t}{b_n}}(1)(x) - \exp\left(-c|t|^p \left(1 - i\beta \operatorname{sign}(t) \tan\left(\frac{p\pi}{2}\right)\right)\right) \right| + \left| Q^n_{\frac{t}{b_n}}(1)(x) \right| \\
&\leq \left| \lambda^n_{\frac{t}{b_n}} - \exp\left(-c|t|^p \left(1 - i\beta \operatorname{sign}(t) \tan\left(\frac{p\pi}{2}\right)\right)\right) \right| \\
&\qquad + \left| \lambda^n_{\frac{t}{b_n}} \right| \cdot \left| \pi_{\frac{t}{b_n}}(1)(x) - 1 \right| + \left| Q^n_{\frac{t}{b_n}}(1)(x) \right|. \qquad (27)
\end{aligned}$$

As $|\lambda^n_{t/b_n}|$ is bounded above by 1 and $|||\pi_t - \nu||| \to 0$, we have

$$\left| \lambda^n_{\frac{t}{b_n}} \right| \cdot \left| \pi_{\frac{t}{b_n}}(1)(x) - 1 \right| \to 0$$

as $n \to \infty$. Moreover,

$$\left| Q^n_{\frac{t}{b_n}}(1)(x) \right| \to 0$$

as $n \to \infty$ as $\|Q^n_{\frac{t}{b_n}}\| \le \rho_0^n$ provided n is sufficiently large. It remains to show that the first term in (27) converges to 0.

It follows from Theorem 6.1 that

$$\operatorname{Re}\log\lambda_{\frac{t}{b_n}} = -n\,\operatorname{sign}(t)c\left|\frac{t}{b_n}\right|^p l\left(\frac{b_n}{t}\right) + o\left(\left|\frac{t}{b_n}\right|^p l\left(\frac{b_n}{t}\right)\right).$$

and

$$\operatorname{Im}\log\lambda_{\frac{t}{b_n}} = -n\,\operatorname{sign}(t)c\left|\frac{t}{b_n}\right|^p \left(\beta l\left(\frac{b_n}{t}\right)\tan\left(\frac{p\pi}{2}\right)\right) + o\left(\left|\frac{t}{b_n}\right|^p l\left(\frac{b_n}{t}\right)\right).$$

Note that

$$o\left(\left|\frac{t}{b_n}\right|^p l\left(\frac{b_n}{t}\right)\right) = o\left(\frac{t^p}{n}\frac{l(b_n/t)}{l(b_n)}\right). \tag{28}$$

As l is slowly varying, $l(b_n/t)/l(b_n) \to 1$ as $n \to \infty$. Hence (28) converges to 0 as $n \to \infty$.

Similarly,

$$-n\,\operatorname{sign}(t)c\left|\frac{t}{b_n}\right|^p l\left(\frac{b_n}{t}\right) = -\operatorname{sign}(t)c|t|^p\frac{l(b_n/t)}{l(b_n)} \to -\operatorname{sign}(t)c|t|^p$$

and

$$\begin{aligned} -n\,\operatorname{sign}(t)c\left|\frac{t}{b_n}\right|^p \beta l\left(\frac{b_n}{t}\right)\tan\left(\frac{p\pi}{2}\right) &= -\operatorname{sign}(t)c|t|^p\beta\frac{l(b_n/t)}{l(b_n)}\tan\left(\frac{p\pi}{2}\right) \\ &\to -\operatorname{sign}(t)c|t|^p\beta\tan\left(\frac{p\pi}{2}\right) \end{aligned}$$

as $n \to \infty$. This completes the proof. ❏

§8 Local stable limit theorems

§8.1 Arithmeticity and cohomology of observations

The local distributional limiting behaviour of the sequence of observations $f(Z_n(x,\cdot))$ depends on whether f essentially takes values in a lattice.

Definition. A function $f : X \to \mathbb{R}$ is said to be *arithmetic* if there exists $t_0 \neq 0$ such that P_{t_0} has an eigenvalue of modulus 1, i.e. there exists $r \in [0, 2\pi)$, $w \in C_{\varepsilon,\beta,\gamma}$, $w \neq 0$, such that $P_{t_0}w = e^{ir}w$. Otherwise, we say that f is *non-arithmetic*.

We have the following characterisation of arithmeticity. We say that two functions f_1, f_2 are *cohomologous* if there exists a function h such that $f_1T_j = f_2T_j + h - hT_j$ for all j.

Proposition 8.1
Suppose that f is continuous and that $\int f\,d\nu = 0$. Then the following are equivalent:

(i) f is arithmetic.

(ii) f is cohomologous to a function taking values in a lattice.

Proof. If f is cohomologous to a lattice-valued function then there exists $t_0 \neq 0$ such that $t_0 f T_j = r + kT_j = h - hT_j$ where $r \in \mathbb{R}$ and k is $2\pi\mathbb{Z}$-valued. It follows that $P_{t_0}(e^{ih}) = e^{ir} e^{ih}$ so that f is arithmetic.

The converse is a standard convexity argument. Suppose that for some $t_0 \neq 0$ there exists a non-zero $w \in C_{\varepsilon,\beta,\gamma}$ such that $P_{t_0} w = e^{ir} w$. Then $|w| \leq P(|w|)$. As $\nu(|w|) \leq \nu(P(|w|)) = \nu(|w|)$ it follows that $|w|$ is a constant, which we may take to be 1, ν-a.e. Writing $w = e^{ih}$ we have that $P_{t_0}(e^{ih}) = \sum_j p_j(x) e^{it_0 f(T_j x) + ih(T_j x)} = e^{ir} e^{ih}$, a convex sum of complex numbers of modulus 1. Hence this sum is trivial, i.e. $fT_j = r + kT_j + h - hT_j$ for some lattice valued function k. ❏

§8.2 Local stable limit theorems

In this section we prove the local stable limit theorem in the case where f is in the domain of attraction of a stable law of order $p \in (0,1) \cup (1,2)$.

Theorem 8.2
Let f be a continuous non-arithmetic function with $|f|_{(\alpha)} < \infty$ and $\nu(f) = 0$. Suppose that f satisfies (26) and is in the domain of attraction of a stable law Y_p with density y_p. Fix $x \in X$. Then for any $a, b \in \mathbb{R}$, $a < b$ we have

$$\lim_{n\to\infty} |b_n \mu_x\{\mathbf{j} \in \Sigma \mid S_n f(x, \mathbf{j}) \in z + [a,b]\} - y_p(z/b_n)(b-a)| = 0$$

uniformly in $z \in \mathbb{R}$, where $b_n^p = bl(b_n)$.

Proof. Define the sequence of measures m_n on $\mathbb{R}$ by defining how they integrate continuous functions: if $w : \mathbb{R} \to \mathbb{R}$ is continuous

$$\int w \, dm_n = \mathbb{E}_x \left(\frac{b_n}{y_p(z/b_n)} \int w(-z + S_n f(x, \cdot)) \right)$$

for a continuous function $w : \mathbb{R} \to \mathbb{R}$. To prove the theorem it is sufficient to prove that m_n weak$*$ converges to Lebesgue measure on $\mathbb{R}$. Let H denote the space of continuous functions $w : \mathbb{R} \to \mathbb{R}$ such that the Fourier transform $\hat{w}$ is compactly supported. To prove that m_n weak$*$ converges to Lebesgue measure, it is sufficient to prove that $m_n(w) \to \int w(t)\, dt$ for all $w \in H$ [Br]. Define

$$A_n(z) = b_n \, \mathbb{E}_x \left(w(-z + S_n f(x, \cdot)) \right) - y_p(z/b_n) \int_{-\infty}^{\infty} w(t)\, dt.$$

Thus it is sufficient to prove that $|A_n(z)| \to 0$ as $n \to \infty$ uniformly in $z \in \mathbb{R}$.

We denote the Fourier transform of w by

$$\hat{w}(s) = \int_{-\infty}^{\infty} w(t) e^{-ist}\, dt.$$

Then by the inversion formula we have

$$w(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{w}(s) e^{ist}\, ds.$$

Moreover, by Theorem 5.1 we have that

$$y_p(z) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi_p(t) e^{-itz}\, dt$$

where

$$\phi_p(t) = \exp\left(-c|t|^p(1 - i\beta\, \mathrm{sign}(t) \tan(p\pi/2))\right).$$

Suppose that $\hat{w}$ is supported on $[-\delta, \delta]$. Then for any $\alpha \in (0, \delta)$ we can write

$$\begin{aligned}
&2\pi A_n(z) \\
&= b_n\, \mathbb{E}_x \left(\int_{-\infty}^{\infty} \hat{w}(t) e^{it(-z+S_n f(x,\cdot))}\, dt \right) - \hat{w}(0) \int_{-\infty}^{\infty} \phi_p(t) e^{-\frac{itz}{b_n}}\, dt \\
&= b_n \int_{-\delta}^{\delta} \hat{w}(t) \int_{\Sigma} e^{itS_n f(x,\cdot)}\, d\mu_x\, e^{-itz}\, dt - \hat{w}(0) \int_{-\infty}^{\infty} \phi_p(t) e^{-\frac{itz}{b_n}}\, dt \\
&= \int_{-\delta b_n}^{\delta b_n} \hat{w}\left(\frac{t}{b_n}\right) P^n_{\frac{t}{b_n}} 1(x) e^{-\frac{itz}{b_n}}\, dt - \hat{w}(0) \int_{-\infty}^{\infty} \phi_p(t) e^{-\frac{itz}{b_n}}\, dt \\
&= \int_{|t| \le \alpha b_n} \left(\hat{w}\left(\frac{t}{b_n}\right) P^n_{\frac{t}{b_n}} 1(x) - \hat{w}(0)\phi_p(t) \right) e^{-\frac{itz}{b_n}}\, dt && (29) \\
&\quad + \int_{\alpha b_n \le |t| \le \delta b_n} \hat{w}\left(\frac{t}{b_n}\right) P^n_{\frac{t}{b_n}} 1(x) e^{-\frac{itz}{b_n}}\, dt && (30) \\
&\quad + \int_{|t| \ge \alpha b_n} \hat{w}(0) \phi_p(t) e^{-\frac{itz}{b_n}}\, dt. && (31)
\end{aligned}$$

Recalling that $P_t^n = \lambda_t^n \pi_t + Q_t^n$ we see that

$$\begin{aligned}
(29) \;=\; & \int_{|t| \le \alpha b_n} \left(\hat{w}\left(\frac{t}{b_n}\right) \lambda^n_{\frac{t}{b_n}} \pi_{\frac{t}{b_n}} 1(x) - \hat{w}(0)\phi_p(t) \right) e^{-\frac{itz}{b_n}}\, dt && (32) \\
& + \int_{|t| \le \alpha b_n} \hat{w}\left(\frac{t}{b_n}\right) Q^n_{\frac{t}{b_n}} 1(x) e^{-\frac{itz}{b_n}}\, dt. && (33)
\end{aligned}$$

By §4.2 we can choose $\alpha \in (0, \delta)$ sufficiently small so that there exists $\rho < 1$ and a constant $C > 0$ such that $\|Q^n_{t/b_n}\| \le C\rho_0^n$ and $|\lambda_{t/b_n}| \in (\rho, 1]$ for all $|t| \le \alpha b_n$. Notice that from the proof of Theorem 7.1 that $\lambda^n_{t/b_n} \to \phi_p(t)$ as $n \to \infty$. It follows from §4.2 that $\pi_{t/b_n} 1 \to \nu(1) = 1$ as $n \to \infty$. By continuity, $\hat{w}(t/b_n) \to \hat{w}(0)$ as $n \to \infty$. Hence the integrand in (32) converges to 0 as $n \to \infty$. As, moreover, the integrand is bounded in modulus by an integrable function, it follows from the Dominated Convergence Theorem that (32) converges to 0 as $n \to \infty$.

Noting that as $b_n^p = nl(b_n)$ for a slowly varying function l, it follows that $b_n = O(n^{\frac{1}{p}+\epsilon})$ for any $\epsilon > 0$. Hence

$$|(33)| \leq C\rho_0^n b_n \to 0$$

as $n \to \infty$.

Similarly, we can estimate (30). For each $t \in [\alpha, \delta]$, P_t does not have 1 as an eigenvalue. Hence there exists $\eta < 1$ and a constant $C > 0$ such that $\|P_t^n\| \leq C\eta^n$. Hence

$$|(30)| \leq C\eta^n b_n \to 0$$

as $n \to \infty$.

Clearly (31) converges to 0 as $n \to \infty$ as $b_n \to \infty$ and $\phi_p \in L^1$.

Noting that none of the constants above depend on z we see that $A_n(z) \to 0$ as $n \to \infty$, uniformly in $z \in \mathbb{R}$. ❏

Remark. The proof of the local central limit theorem is similar, cf. [BPD], with $e^{-t^2/2}$ replacing $y_p(t)$.

Remark. One can also formulate and prove a stable local limit theorem in the case of an arithmetic Hölder function f; cf. [AD2].

§9 Random difference equations

Let $a_j \in \mathbb{R}, a_j > 0$, $b_j \in \mathbb{R}$. random difference equation is determined by the equation

$$z_{j+1} = a_{j+1}z_j + b_{j+1} \tag{34}$$

where the (a_j, b_j) are i.i.d. pairs of real numbers. There is a great deal of literature on the solution to such equations and their applications ([Ke, DF], for example). In the case where the (a_j, b_j) are chosen independently from a finite or countable set then (34) can be viewed as an IFS by taking $T_j(x) = a_j x + b$, chosen with probability p_j.

The affine IFS (T_j, p_j) satisfies (1) precisely when $\sum p_j a_j < 1$. (The technical hypothesis (10) holds provided $\sum p_j b_j < \infty$.) In this case, the tail behaviour of the invariant measure ν is well-known.

Theorem 9.1 ([Ke, Gol])
Suppose the affine IFS (T_j, p_j) contracts on average and that the closed subgroup of $\mathbb{R}$ generated by $\log a_j$ is $\mathbb{R}$. Define θ_0 by $\sum p_j a_j^{\theta_0} = 1$. Suppose that for some $\epsilon > 0$ we have $\sum p_j a_j^{\theta_0+\epsilon}$, $\sum p_j b_j^{\theta_0+\epsilon} < \infty$. Then $\nu([t, \infty)) \sim \frac{1}{t^{\theta_0}}$ as $t \to \infty$.

Taking $f(x) = x^\alpha$ where $\alpha \in (0,1]$ we see that f satisfies (7) with $p = \theta_0/\alpha$, and so the distributional and local limit Theorems (when f is non-arithmetic) above hold in this case.

More generally, the case of a random walk on the affine group of $\mathbb{R}$ has been studied [GP]. In this case the probability used to choose the T_j need not be supported on a discrete set of maps. In the case where $f(x) = x$, a complete analysis of the expansion of the maximal eigenvalue of P_t, and an precise identification of the stable limit law can be achieved [GP].

References

[AD1] J. Aaronson and M. Denker, *Characteristic functions of random variables attracted to 1-stable laws*, Ann. Prob., **26** (1998), 399–415.

[AD2] J. Aaronson and M. Denker, *Local limit theorems for partial sums of stationary sequences generated by Gibbs-Markov maps*, Stoch. Dyn., **1** (2001), 193–237.

[BDEG] M.F. Barnsley, S.G. Demko, J.H. Elton and J.S. Geronimo, *Invariant measures for Markov processes arising from iterated function systems with place-dependent probabilities.* Ann. Inst. H. Poincaré Probab. Statist., **24** (1988), 367–394.

[Br] L. Breiman, Probability, Addison-Wesley, Reading, Mass., 1968.

[BPD] A. Broise, M. Peigné and F. Dal'bo, *Études spectrales d'opérateurs de transfert et applications*, Astérisque **238** (1996), Soc. Math. de France.

[DF] P. Diaconis and D. Freedman, *Iterated random functions*, SIAM Rev., **41** (1999), 45–76.

[E1] J.H. Elton, *An ergodic theorem for iterated maps*, Ergod. Th. & Dynam. Sys., **7** (1987), 481–488.

[E2] J.H. Elton, *A multiplicative ergodic theorem for Lipschitz maps*, Stoch. Proc. Appl., **34** (1990), 39–47.

[Gol] C.M. Goldie, *Implicit renewal theory and tails of solutions of random equations*, Ann. Appl. Prob., **1** (1991), 126–166.

[Gou] S. Gouëzel, *Central limit theorem and stable laws for intermittent maps*, Prob. Th. Rel. Fields, **128** (2004), 82-122.

[GP] Y. Guivarc'h and E. le Page, *On spectral properties of a family of transfer operators and convergence to stable laws for affine random walks*, Ergod. Th. & Dynam. Sys., **28** (2008), 423–446.

[H] H. Hennion, *Décomposition spectrale des opérateurs de Doeblin-Fortet*, Proc. Amer. Math. Soc., **69** (1993), 627–634.

[HH1] H. Hennion and L. Hervé, *Limit theorems for Markov chains and stochastic properties of dynamical systems by quasi-compactness*, Springer Lecture Notes in Mathematics, **1766** (2001).

[HH2] H. Hennion and L. Hervé, *Central limit theorems for iterated random Lipschitz mappings*, Ann. Prob., **32** (2004), 1934–1984.

[IL] I. Ibragimov and Yu. Linnik, *Independent and Stationary Sequences of Random Variables*, Wolters-Noordhodd Publishing, Groningen, 1971.

[KL] G. Keller and C. Liverani, *Stability of the spectrum for transfer operators*, Ann. Sc. Norm. Super. Pisa, **28** (1999), 141–152.

[Ke] H. Kesten, *Random difference equations and renewal theory for products of random matrices*, Acta Math, **131** (1973), 207–248.

[N] S.V. Nagaev, *Some limit theorems for stationary Markov chains*, Theor. Probab. Appl. **2** (1957), 378–406.

[Pe] M. Peigné, *Iterated function systems and spectral decomposition of the associated Markov operator*, Publ. Inst. Rech. Math. Rennes 1993, Univ. Rennes I, Rennes, 1993.

[W] C.P. Walkden, *Invariance principles for iterated maps that contract on average*, Trans. Amer. Math. Soc., **359** (2007), 1081-1097.

Sara Santos, School of Mathematics, The University of Manchester, Oxford Road, Manchester M13 9PL, U.K., email: `sara.i.santos@gmail.com`

Charles Walkden, School of Mathematics, The University of Manchester, Oxford Road, Manchester M13 9PL, U.K., email: `charles.walkden@manchester.ac.uk`