ON THE RATE OF CONVERGENCE TO EQUILIBRIUM FOR COUNTABLE ERGODIC MARKOV CHAINS

Stefano Isola

Dipartimento di Matematica, Università degli Studi di Bologna,
piazza di Porta S.Donato 5, I-40127 Bologna, Italy.
e-mail: isola@dm.unibo.it

Abstract. Using elementary methods, we prove that for a countable Markov chain P of ergodic degree $\ell \geq 1$ the rate of convergence towards the stationary distribution is $o(n^{-(\ell-1)})$, provided the initial distribution satisfies certain conditions of asymptotic decay. An example, modelling temporal intermittency in dynamical system theory, is worked out in detail, illustrating the relationships between convergence behaviour, analytic properties of the generating functions associated to transition probabilities and spectral properties of the Markov operator P on the Banach space ℓ_1. Explicit conditions allowing to obtain either optimal bounds or the actual asymptotics for the rate of convergence are also discussed.

AMS 1991 Subject Classification: Primary 60J10

Keywords: Countable Markov chains, convergence to equilibrium, ergodic degreee, intermittent dynamical systems.

1. Preliminaries and statement of the results.

Let S be a countable set and $P : S \times S \to [0,1]$ be a transition probability matrix. With no loss we may set $S = \mathbb{N}$. We shall assume that P governs an irreducible, recurrent and aperiodic Markov chain $X = (x_n)_0^\infty$ with state space S. To be more precise, we set $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$ and let Ω denote the subset of $S^{\mathbb{N}_0}$ given by all sequences $\omega = (\omega_i)_{i \in \mathbb{N}_0}$ which satisfy for any integer i: $p_{\omega_i \omega_{i+1}} \equiv P(\omega_i, \omega_{i+1}) > 0$. Let moreover $\mathbf{P}_\nu$ be the probability measure with initial distribution ν (that of x_0) on Ω, i.e.

$$\mathbf{P}_\nu\{x_n(\omega) = j\} = \sum \nu_i p_{ij}^n = (\nu P^n)_j$$

where $p_{ij}^n \equiv P^{(n)}(i,j)$. In particular, if $\nu = \delta_i$, where i is some reference state chosen from the outset, we have

$$\mathbf{P}_i\{x_n(\omega) = j\} = \mathbf{P}\{x_n(\omega) = j | x_0(\omega) = i\} = p_{ij}^n.$$

Our sample space will be Ω equiped with the restriction of the product σ-field and with probability measure $\mathbf{P}_\nu$ for some initial distribution ν. We shall denote by $\mathbf{E}_\nu$ the expectation w.r.t. $\mathbf{P}_\nu$. Finally, for any $n \in \mathbb{N}_0$ we let x_n be the projection on the n^{th} coordinate, i.e. $x_n(\omega) = \omega_n$.

It is well known that P has a (unique) stationary probability distribution if and only if it is ergodic or, equivalently, positive recurrent (see e.g. [C], [KSK]). This means that for some (and hence for all) $i \in S$, the $\mathbf{P}_i$-expectation m_i of $\min\{n \in \mathbb{N}_0, x_n(\omega) = i\}$, the time of the first visit at i, is finite. In this case, the distribution π on S given by $\pi = (\pi_i)_1^\infty = (1/m_i)_1^\infty$ is a solution to $\pi = \pi P$ and thus defines a stationary distribution.

In the sequel we identify sequences $\nu = (\nu_i)_1^\infty \in \ell_1(S)$, the corresponding row vectors $\nu = (\nu_1, \nu_2, \ldots)$, and finite signed measures on S, and define

$$\|\nu\| = \sum_{i=1}^{\infty} |\nu_i|.$$

A signed measure ν satisfying $\sum \nu_l = 1$ will be called a signed distribution in the sequel. Similarly, we shall identify sequences $\mathbf{u} = (u_i)_1^\infty \in \ell_\infty(S)$, the corresponding column vectors $\mathbf{u} = (u_1, u_2, \ldots)^t$, and bounded functions on S. We introduce the classical *taboo* quantities:

$$\begin{aligned}
f_{ij}^n &= \mathbf{P}\{x_l(\omega) \neq j, 0 < l < n, x_n(\omega) = j \mid x_0(\omega) = i\}, \\
{}_kp_{ij}^n &= \mathbf{P}\{x_l(\omega) \neq k, 0 < l < n, x_n(\omega) = j \mid x_0(\omega) = i\}, \\
f_{ij}^* &= \sum_{n=1}^\infty f_{ij}^n, \qquad {}_kp_{ij}^* = \sum_{n=1}^\infty {}_kp_{ij}^n.
\end{aligned}$$

Clearly ${}_jp_{ij}^n = f_{ij}^n$. Since we have a unique recurrent class, $f_{ij}^* = 1$ for all $i, j \in S$. Moreover, for an ergodic chain we have ([C], Chap. I.9, Thm. 5)

$$\lim_{n\to\infty} \frac{\sum_{k=0}^n p_{ij}^n}{\sum_{k=0}^n p_{ii}^n} = \frac{\pi_j}{\pi_i} = {}_ip_{ij}^*. \tag{1.1}$$

The last quantity can also be viewed as the $\mathbf{P}_1$-mean number of visits to the state j before return to i. The relation between the f_{ij}'s and the transition probabilities p_{ij} is given by ([C], Chap. I.5, Thm. 2)

$$p_{ij}^0 = \delta_{ij}, \quad f_{ij}^0 = 0$$

and

$$p_{ij}^n = \sum_{k=1}^n f_{ij}^k p_{jj}^{n-k}. \tag{1.2}$$

We also have

$$\pi_i = \frac{1}{\sum_{n=1}^\infty n\, f_{ii}^n}. \tag{1.3}$$

For $k \in I\!N$, $i \in S$, let $t_k^{(i)}$ be the time of the k-th entrance into state i, and let

$$r_k^{(i)}(\omega) = t_{k+1}^{(i)} - t_k^{(i)}, \quad k \geq 0 \tag{1.4}$$

be the sequence of times between returns (set $t_0^{(i)} = 0$). Clearly we have $r_0^{(i)} \geq 0$ and $r_k^{(i)} > 0$ for $k \geq 1$. Moreover, the state i being recurrent, $r_1^{(i)}, r_2^{(i)}, \ldots$ are i.i.d. random variables under the probability $\mathbf{P}_i$. Their common distribution is given by

$$\mathbf{P}_i\{r_k^{(i)}(\omega) = n\} = f_{ii}^n \qquad k \geq 1. \tag{1.5}$$

On the other hand, having fixed an initial distribution ν and a reference state i, the random variable $r_0^{(i)}$ (the delay in the embedded renewal process) is distributed according to $\mathbf{P}_\nu$. More specifically,

$$\mathbf{P}_\nu\{r_0^{(i)} = n\} = \nu_i\, \delta_{n0} + \sum_{l \neq i} \nu_l f_{li}^n.$$

For $\ell \in I\!N_0$, $i, j \in S$ (and $k \geq 1$), we set

$$M_{ij}^{(\ell)} := \mathbf{E}_i(|r_k^{(j)}|^\ell) = \sum_{n=1}^\infty n^\ell f_{ij}^n. \tag{1.6}$$

Notice that $m_i \equiv M_{ii}^{(1)}$. Given a signed distribution ν on S, we also set,

$$\hat{M}_{\nu i}^{(\ell)} := \mathbf{E}_\nu(|r_0^{(i)}|^\ell) = \sum_{l \neq i} \nu_l \sum_{n=1}^\infty n^\ell f_{li}^n = \sum_{l \neq i} \nu_l\, M_{li}^{(\ell)}. \tag{1.7}$$

Lemma 1. *If $\ell \geq 1$, then $M_{ii}^{(\ell)} < \infty$ if and only if $\hat{M}_{\pi i}^{(\ell-1)} < \infty$.*

Proof. We have

$$\begin{aligned}
\hat{M}_{\pi i}^{(\ell-1)} &= \sum_{l \neq i} \pi_l \sum_{n=1}^\infty n^{\ell-1} f_{li}^n = \pi_i \sum_{l \neq i} {}_ip_{il}^* \sum_{n=1}^\infty n^{\ell-1} f_{li}^n \\
&= \pi_i \sum_{n=1}^\infty n^{\ell-1} \sum_{m=0}^\infty \sum_{l \neq i} {}_ip_{il}^m\, {}_ip_{li}^n = \pi_i \sum_{n=1}^\infty n^{\ell-1} \sum_{m=0}^\infty {}_ip_{ii}^{n+m} \\
&= \pi_i \sum_{n=1}^\infty n^{\ell-1} \sum_{m=n}^\infty f_{ii}^m = \pi_i \sum_{n=1}^\infty n^\ell f_{ii}^n = \pi_i\, M_{ii}^{(\ell)}
\end{aligned}$$

where we have used the last identity in (1.1) along with the properties ${}_ip_{il}^0 = 0$ and ${}_ip_{ii}^{n+m} = \sum_{l \neq i} {}_ip_{il}^m\, {}_ip_{li}^n$. $\diamondsuit$

Furthermore, it is well known that, for a recurrent chain, if $M_{ii}^{(\ell)} < \infty$ for some state i then $M_{jj}^{(\ell)} < \infty$ for all $j \in S$ (see, e.g., [C], Chap. I.11, Cor. 1). We now state the following definition (see [KSK] for related notions),

Definition 1. *The recurrent chain P has* ergodic degree ℓ *if $M_{ii}^{(\ell)} < \infty$ but $M_{ii}^{(\ell+1)} = \infty$ for some (and hence for all) $i \in S$.*

Notice that $M_{ii}^{(0)} = 1$ so that the degree is always defined. Thus, null chains have ergodic degree 0 whereas ergodic chains have ergodic degree at least 1. If $M_{ii}^{(\ell)} < \infty$ for every ℓ, one says that P has infinite ergodic degree. This happens for instance if the coefficients f_{ii}^n decay geometrically with n. In this case the corresponding chain is accordingly called geometrically ergodic. We refer to [FMM] for related convergence results in the geometrically ergodic case.

Finally, the preceeding observations motivate the next definition.

Definition 2. *Given an ergodic chain P with state space S and a signed distribution ν on S, we say that ν has P-order ℓ, if $\hat{M}_{\nu i}^{(\ell)} < \infty$ but $\hat{M}_{\nu i}^{(\ell+1)} = \infty$ for some (and hence for all) $i \in S$.*

In particular, from Lemma 1 it follows that if P has ergodic degree ℓ then π has P-order $\ell - 1$, and viceversa. We now state the main result of this Section.

Theorem 1. *Suppose P has ergodic degree $\ell \geq 1$. Then, for any initial signed distribution ν of P-order at least $\ell - 1$, we have*

$$||\nu P^n - \pi|| = o(n^{-(\ell-1)}).$$

Remark. In the next Section we shall prove Theorem 1 in several Lemmas using elementary arguments. Some partial results in our direction were obtained in [F1]. Related results can be obtained using the coupling method (see [L]). Another method, making use of matrix-valued analytic functions and allowing to sharpen the above result under suitable conditions, is proposed in Section 3 for a specific example.

We let τ be the shift transformation on Ω, that is $x_k \circ \tau(\omega) = \omega_{k+1}$. With P and π one can define a τ-invariant Markov random field $\mu = \mu(P, \pi)$ supported by Ω as follows:

$$\mu(\{x_k(\omega) = \xi_0, \ldots, x_{k+n}(\omega) = \xi_n\}) = \pi_{\xi_0} \prod_{j=1}^{n} p_{\xi_{j-1}\xi_j} \tag{1.8}$$

We then have the following,

Corollary 1. *Suppose $M_{ii}^{(\ell)} < \infty$ for some $\ell \geq 1$. Then, for any pair of bounded vectors $\mathbf{u}, \mathbf{v} : S \to I\!R$,*

$$|\, \mu(\mathbf{u}(x_n)\mathbf{v}(x_0)) - \mu(\mathbf{u}(x_0))\, \mu(\mathbf{v}(x_0))\,| = o(n^{-(\ell-1)})$$

Acknowledgements: I would like to thank Lai-Sang Young for interesting conversations at the origin of this research.

2. Proofs of the main results.

First, we recall the following result, which will be used several times in the sequel and whose proof can be found in ([C], Chap. I.5, Lemma A):

Lemma 2. *Let $\{a_n\}_{n\geq 0}$ be a sequence of nonnegative numbers not all vanishing. If*

$$\lim_{n\to\infty} \frac{a_n}{\sum_{m=0}^{n} a_m} = 0$$

then, whenever the sequence $\{b_n\}_{n\geq 0}$ of real numbers has a limit, we have

$$\lim_{n\to\infty} \frac{\sum_{m=0}^{n} a_m b_{n-m}}{\sum_{m=0}^{n} a_m} = \lim_{n\to\infty} b_n.$$

The following three Lemmas are supposed to hold under the hypotheses of Theorem 1.

Lemma 3. *For each $i \in S$, we have*

$$|p_{ii}^n - \pi_i| = o(n^{-(\ell-1)}).$$

Proof. We introduce the generating functions

$$P_{ij}(z) = \sum_{n=0}^{\infty} p_{ij}^n\, z^n, \qquad F_{ij}(z) = \sum_{n=0}^{\infty} f_{ij}^n\, z^n \tag{2.1}$$

and from (1.2) we get the relations (we set $f_{ij}^n = 0$ for $n = 0$)

$$P_{ii}(z) = \frac{1}{1 - F_{ii}(z)}, \quad P_{ij}(z) = F_{ij}(z)P_{jj}(z), \quad i \neq j. \tag{2.2}$$

We first show that the function $P_{ii}(z)$ is analytic in $|z| < 1$ and converges at every point of the unit circle besides $z = 1$. Indeed, recurrence of the state i implies $F_{ii}(1) = 1$, so that $|F_{ii}(z)| < 1$ for $|z| < 1$ because $f_{ii}^n \geq 0$. Moreover, $|F_{ii}(z)| < 1$ also for $|z| = 1$, $z \neq 1$. This follows from the fact that, since the chain is aperiodic, g.c.d.$\{n, f_{ii}^n \neq 0\} = 1$. Now set

$$D_{ii}(z) = \sum_{n=0}^{\infty} d_n^{(i)} z^n, \quad d_n^{(i)} := \sum_{k>n} f_{ii}^k$$

and notice that $D_{ii}(z)$ converges absolutely in $|z| \leq 1$ and has no zeros on $|z| = 1$. In addition $\sum_{n=0}^{\infty} d_n^{(i)} = m_i$. It then follows (see [Z], p.140) that the function

$$\frac{1}{D_{ii}(z)} = (1 - z)P_{ii}(z)$$

has a power series expansion which converges absolutely in the closed unit disk and, moreover, its value at $z = 1$ is $m_i^{-1} = \pi_i$. As a consequence of Abel's theorem (see [T], p.229) we then find

$$\lim_{n\to\infty} p_{ii}^n = \pi_i.$$

We now study the rate of convergence to zero of

$$\mu_n^{(i)} := m_i \, p_{ii}^n - 1$$

as $n \to \infty$. To this end, we first observe that $\mu_n^{(i)}$ is the coefficient of z^n in

$$H_{ii}(z) := m_i \, P_{ii}(z) - \frac{1}{1 - z} = \frac{E_{ii}(z)}{D_{ii}(z)}$$

where

$$E_{ii}(z) = \sum_{n=0}^{\infty} e_n^{(i)} z^n, \quad e_n^{(i)} := \sum_{k>n} d_k^{(i)}.$$

Also observe that

$$\partial F_{ii}(1) = \sum_{n=1}^{\infty} n f_{ii}^{n} = M_{ii}^{(1)} = m_i = D_{ii}(1).$$

where the notation $\partial^m G(z_0) = d^m G(z)/dz^m\big|_{z=z_0}$ has been used. Let us write $H_{ii}(z)$ as

$$H_{ii}(z) = P_{ii}(z) \cdot \{\, (1-z) E_{ii}(z) \,\} = P_{ii}(z) \cdot \{\, m_i - D_{ii}(z) \,\}.$$

It then follows that under the only assumption that $M_{ii}^{(1)} < \infty$ the power series of the term within braces converges absolutely for $z = 1$ and its value is 0. This fact and the above discussion entail that $\mu_n^{(i)} \to 0$ as $n \to \infty$. Next, assume $M_{ii}^{(2)} < \infty$. In this case we have

$$\partial^2 F_{ii}(1) = \sum_{n=2}^{\infty} n(n-1) f_{ii}^{n} = M_{ii}^{(2)} - M_{ii}^{(1)} = 2\, \partial D_{ii}(1) = 2\, E_{ii}(1).$$

Then notice that $n\, \mu_n^{(i)}$ is the coefficient of z^{n-1} in

$$\partial H_{ii}(z) = P_{ii}(z) \cdot \left\{ E_{ii}(z) - m_i \, \frac{\partial D_{ii}(z)}{D_{ii}(z)} \right\}$$

and by the above identities the power series of the term within curly brackets converges absolutely for $z = 1$ and its value is 0. We then have that $n\, \mu_n^{(i)}$ tends to zero.

More generally, if $M_{ii}^{(\ell)} < \infty$, one has the identities

$$\begin{aligned} \partial^\ell F_{ii}(1) &= \sum_{n=\ell}^{\infty} n(n-1)\dots(n-\ell+1) f_{ii}^{n} \\ &= \sum_{j=1}^{\ell} S_\ell^{(j)} M_{ii}^{(j)} = \ell\, \partial^{\ell-1} D_{ii}(1) = \ell(\ell-1)\, \partial^{\ell-2} E_{ii}(1) \end{aligned}$$

where the $S_\ell^{(j)}$ are Stirling numbers of the first kind defined by the relation

$$x(x-1)\dots(x-\ell+1) = \sum_{j=1}^{\ell} S_\ell^{(j)} x^j.$$

For $k \geq 0$, set

$$J_k(z) := \frac{1}{P_{ii}(z)} \cdot \partial^k H_{ii}(z)$$

so that

$$J_{k+1}(z) = \partial J_k(z) + J_k(z) \cdot \partial \log P_{ii}(z).$$

Using these relations and proceeding inductively in ℓ we find that if $M_{ii}^{(\ell)} < \infty$ then the power series of

$$\frac{1}{P_{ii}(z)} \cdot \partial^{\ell-1} H_{ii}(z)$$

converges absolutely for $z = 1$ and its value is 0. The above and the fact that $n(n-1)\ldots(n-\ell+2)\mu_n^{(i)}$ is the coefficient of $z^{n-\ell+1}$ in $\partial^{\ell-1}H_{ii}(z)$ then complete the proof. $\diamondsuit$

Lemma 4. *For all $j \in S$,*

$$|p_{ij}^n - \pi_j| = o(n^{-(\ell-1)})$$

Proof. First, using Lemma 2, it follows from (1.2) and $\sum_n f_{ij}^n = 1$ that $p_{ij}^n \to \pi_j$ as $n \to \infty$. Moreover, as already remarked, the condition $M_{ii}^{(\ell)} < \infty$ implies that $\sum_{n=1}^{\infty} n^\ell f_{ij}^n < \infty$, for all pairs (distinct or not) $i, j \in S$. This implies $\sum_{k=n}^{\infty} f_{ij}^k = o(n^{-\ell})$. Putting together these observations, Lemma 2 and Lemma 3, along with the inequality

$$|p_{ij}^n - \pi_j| \leq \sum_{k=1}^{n} f_{ij}^k |p_{jj}^{n-k} - \pi_j| + \pi_j \sum_{k=n+1}^{\infty} f_{ij}^k,$$

we then have that the rate of convergence is the same as in Lemma 3. $\diamondsuit$

These properties entail that P^n tends to the matrix whose rows are $(\pi_1, \pi_2, \ldots)$. In addition, we have the following

Lemma 5.

$$\sum_{j \in S} |p_{ij}^n - \pi_j| = o(n^{-(\ell-1)}).$$

Proof. This follows by first noting that the proofs of Lemmas 2 and 3 actually imply that $|p_{ij}^n/\pi_j - 1| = o(n^{-(\ell-1)})$ and then using the fact that π is a probability vector. $\diamondsuit$

Proof of Theorem 1. Lemma 5 implies that $||\delta_i P^n - \pi|| = o(n^{-(\ell-1)})$. Putting $\nu = \sum \nu_l \, \delta_l$ and using the fact that ν is normalized, i.e. $\sum \nu_l = 1$, we write

$$\begin{aligned} \nu P^n - \pi &= \sum_l \nu_l \, (\delta_i P^n - \pi) + \sum_{l \neq i} \nu_l \, (\delta_l P^n - \delta_i P^n) \\ &= (\delta_i P^n - \pi) + \sum_{l \neq i} \nu_l \, (\delta_l P^n - \delta_i P^n) \end{aligned}$$

The ℓ_1-norm of the first term in the r.h.s. is then estimated by Lemma 5. For the second term we have $\|\delta_l P^n - \delta_i P^n\| = \sum_j |p_{lj}^n - p_{ij}^n|$. Using the decompositions $p_{lj}^n = {}_i p_{lj}^n + \sum_{k=1}^{n-1} f_{li}^k p_{ij}^{n-k}$ ([C], Chap. I.9, Thm. 1) and $p_{ij}^n = \sum_{k=1}^{n-1} f_{li}^k \, p_{ij}^n + \sum_{k=n}^{\infty} f_{li}^k \, p_{ij}^n$, and noting that $\sum_j {}_i p_{lj}^n = \sum_{k \geq n} f_{li}^k$ and $\sum_j p_{ij}^n = 1$, we obtain

$$\|\delta_l P^n - \delta_i P^n\| \leq 2 \sum_{k=n}^{\infty} f_{li}^k + \sum_{k=1}^{n-1} f_{li}^k \sum_j |p_{ij}^{n-k} - p_{ij}^n|.$$

Now, using Lemma 2, Lemma 5 and $f_{li}^* = 1$, one readily shows that the second term above gives a contribution $o(n^{-(\ell-1)})$ to $\|\nu P^n - \pi\|$. The same holds true for the first term, since the assumption that ν has P-order at least $\ell - 1$ (w.r.t to the reference state i) implies that

$$\sum_{l \neq i} \nu_l \sum_{k \geq n} f_{li}^k = o(n^{-(\ell-1)}). \qquad \diamondsuit$$

Remark 1. As we made no use of the fact that $M_{ii}^{(\ell+1)} = \infty$, the statement of Theorem 1 actually holds under the only assumption that $M_{ii}^{(\ell)} < \infty$. We shall see later an example where the assumption on the ergodic degree can be made effective to yield optimal bounds on the rate of convergence.

Remark 2. The proof given above brings out the meaning of the condition on the P-order of the initial distribution ν. This is related to the fact that the behaviour of $|p_{ij}^n - \pi_j|$ and hence of $||\delta_i P^n - \pi||$ is necessarily not uniform in the departing state index i. Indeed, according to the above discussion, such uniformity would imply the existence of two positive constant C_1, C_2 and an integer n_0, which *do not depend* on i and l, such that, for all $n \geq n_0$

$$C_1 \leq \frac{\sum_{k \geq n} f_{li}^k}{\sum_{k \geq n} f_{il}^k} \leq C_2.$$

This, in turn, would imply that the ratio $M_{li}^{(1)}/M_{il}^{(1)}$ satisfies a similar bound. On the other hand it is well known that $\lim_{l \to \infty} (M_{li}^{(1)}/M_{il}^{(1)}) = 0$, for all $i \in S$ ([C], Chap. I.11, Thm. 6; see also [H]).

Proof of Corollary 1. For any pair $\mathbf{u} \in \ell_\infty(S)$, $\rho \in \ell_1(S)$ we define $\overline{\rho \mathbf{u}} = (\rho(1)u(1), \rho(2)u(2), \ldots)$ and $\rho \cdot \mathbf{u} = \sum_{i \in S} \rho(i)u(i)$. Thus $\overline{\rho \mathbf{u}} \cdot \mathbf{1} = \rho \cdot \mathbf{u}$, and the unit column vector $\mathbf{1} = (1, 1, \ldots)^t$ satisfies $P\mathbf{1} = \mathbf{1}$. For definiteness and without loss, suppose that $\mu(\mathbf{u})\,\mu(\mathbf{v}) \neq 0$. Then we have

$$\begin{aligned} |\,\mu(\mathbf{u}(x_n)\mathbf{v}(x_0)) - \mu(\mathbf{u}(x_0))\,\mu(\mathbf{v}(x_0))\,| &= |\,\overline{\pi \mathbf{v}}\,P^n \cdot \mathbf{u} - (\pi \cdot \mathbf{v})(\pi \cdot \mathbf{u})\,| \\ &= |\,(\overline{\pi \mathbf{v}}\,P^n - \pi\,(\overline{\pi \mathbf{v}} \cdot \mathbf{1})\,) \cdot \mathbf{u}\,| \\ &\leq \|\mathbf{u}\|_\infty\,\|\mathbf{v}\|_\infty\,\|\,\nu\,P^n - \pi\,\| \end{aligned}$$

where ν denotes the normalized ℓ_1 row vector $\overline{\pi \mathbf{v}}/(\pi \cdot \mathbf{v})$. The result now follows putting together Lemma 1 and Theorem 1. $\diamondsuit$

3. Convergence vs analytic and spectral properties. An example.

As we have seen, the dependence on the departing state i of the behaviour of $\|\delta_i P^n - \pi\|$, although not explicitly indicated in Lemma 5, is what makes our assumptions on the P-order of the initial distribution ν necessary.

Moreover, from our discussion it follows that the rate of convergence to zero of $\|\delta_i P^n - \pi\|$ is connected with the analytic properties of the generating functions $P_{ij}(z)$ in the vicinity of the singular point $z = 1$.

If we now consider P as a bounded linear Markov operator acting on the Banach space $\ell_1(S)$, its adjoint P^* is represented by the transposed matrix acting on the dual space $\ell_1^* = \ell_\infty$. The resolvent $R_\lambda(P) := (\lambda I - P)^{-1}$ admits, for $|\lambda| > \|P\|$, the expansion

$$\lambda\, R_\lambda(P) = I + \sum_{n=1}^{\infty} \left(\frac{P}{\lambda}\right)^n$$

which shows that $1 - \delta_{ij} + P_{ij}(z)$ is the (i,j)-element of $\lambda\, R_\lambda(P)$, with the identification $z = 1/\lambda$. This, in turn, indicates that the convergence properties of $\|\delta_i P^n - \pi\|$, the analytic properties of the functions $P_{ij}(z)$, and the spectral properties of P in $\ell_1(S)$ are intimately connected items. In particular, the dependence of the first two from the state index i plays an important role in determining nature of the latter, as we shall see in the following example[1].

Example. Suppose that $S = I\!N$ and the transition matrix is

$$P = \begin{pmatrix} p_1 & p_2 & p_3 & \dots \\ 1 & 0 & 0 & \dots \\ 0 & 1 & 0 & \dots \\ 0 & 0 & 1 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

The space Ω is then given by all sequences ω satisfying the following condition: given ω_i then either $\omega_{i-1} = \omega_i + 1$ or $\omega_{i-1} = 1$. We shall assume that the probability vector $p = (p_1, p_2, \ldots)$ has the property g.c.d.$\{n : p_n > 0\} = 1$. It then follows that the corresponding chain is aperiodic and recurrent. Let the coefficients d_n be defined by $d_n := \sum_{i>n} p_i$, $(n \geq 0)$. The steady-state equation is $\pi_n = \sum_{i \in S} \pi_i\, p_{in}$ and is formally solved by $\pi_n = \pi_1\, d_{n-1}$, $(n \geq 1)$. We also have $f_{11}^n = p_n$. Consequently, the chain is positive-recurrent if and only if $\sum d_n < \infty$, null-recurrent in the opposite case. In the former case, we have $\pi_1 = (\sum_{n=1}^{\infty} n p_n)^{-1} = (\sum_{n=0}^{\infty} d_n)^{-1}$. Notice that the two probability vectors π and p coincide if and only if $p_n = 2^{-n}$. On the other hand, if $p_n \sim n^{-(\ell+2)}\, L(n)$ with $L(n)$ slowly varying at infinity, then the chain has ergodic degree ℓ.

Remark. This example has interesting applications in modelling temporal intermittency in dynamical systems theory (see, e.g., [W]). A brief discussion on

[1] We shall adopt the convention that a matrix (t_{ij}) representing an operator T acts from the right, that is through the equations $(Tx)_j = \sum_{i \in S} x_i\, t_{ij}$.

the consequences of the results stated below in the context of intermittent interval maps is given in the Appendix at the end of the paper.

Proposition 1. *Suppose that the chain P defined above has finite ergodic degree $\ell \geq 1$. Then,*

I. The generating functions $P_{ij}(z)$ defined in (2.1) are analytic in the open unit disk with no singularities on the unit circle besides a non-polar singular point at $z = 1$;

II. The spectrum $\sigma(P)$ of the Markov operator P acting on $\ell_1(\mathbb{N})$ coincides with the closed unit disk and decomposes as follows: $\sigma_p(P) = \{\lambda : |\lambda| < 1\} \cup \{1\}$ and $\sigma_c(P) = \{\lambda : |\lambda| = 1, \lambda \neq 1\}$;

III. for any bounded vector $\mathbf{u}$ and any initial distribution $\nu = (\nu_i)_1^\infty \in \ell_1(S)$ s.t. $\nu_i = O(\pi_i)$, the quantity $(\nu P^n - \pi) \cdot \mathbf{u}$ decays as $o(n^{-(\ell-1)})$.

Assume furthermore that $p_n \sim n^{-(\ell+2)} L(n)$ with $L(n)$ slowly varying at infinity. Then, for n large enough,

$$|(\nu P^n - \pi) \cdot \mathbf{u}| \leq C\, n^{-\ell} L(n),$$

where $C = (\pi \cdot \mathbf{u})(\nu \cdot \mathbf{1})/(\ell(\ell+1)m_1)$. Under the additional hypotheses that $u_i = o(1)$ and $\nu_i = o(\pi_i)$ we have

$$(\nu P^n - \pi) \cdot \mathbf{u} \sim C\, n^{-\ell} L(n) \quad \text{as} \quad n \to \infty.$$

The proof of Proposition 1 will follow from the discussion given hereafter.

I. Generating functions and their analytic properties.

First, it is easy to check that all entries of the first n rows of P^n are positive, the i-th row of P being the $(i + n - 1)$-th of P^n. More specifically, one sees inductively that for $n > 1$, $i > 1$, $j \in \mathbb{N}$,

$$P^n(i, j) = P^{n-1}(i - 1, j). \tag{3.1}$$

For the generating functions of the $P^n(i,j)$'s we then obtain the relations

$$\begin{aligned} P_{ij}(z) &= \delta_{ij} + z^{i-1} P_{1j}(z), \quad j \geq i > 1 \\ P_{i1}(z) &= z^{i-1} P_{11}(z), \quad i \geq 1 \\ P_{ij}(z) &= z^{i-j} + z^{i-1} P_{1j}(z), \quad i > j > 1. \end{aligned} \tag{3.2}$$

It then suffice to study the behaviour of the entries of the first row. They satisfy the recurrence relations (recall that $P^0(i,j) = \delta_{ij}$),

$$P^n(1,j) = P^{n-1}(1,1)\, P(1,j) + P^{n-1}(1,j+1), \qquad j \geq 1.$$

This yields

$$P^n(1,j) = \sum_{k=1}^{n-1} P(1,k)\ P^{n-k}(1,j) + P(1,j+n-1). \tag{3.3}$$

Putting $j = 1$ and recalling that $P(1,k) = p_k = f_{11}^k$ one gets a particular case of equation (1.2),

$$P^n(1,1) = \sum_{k=1}^{n} P(1,k)\ P^{n-k}(1,1).$$

We then find for the generating function of the $P^n(1,1)$'s,

$$P_{11}(z) = \frac{1}{1 - \sum_{n=1}^{\infty} p_n z^n} = \frac{1}{(1-z)D(z)} \tag{3.4}$$

where $D(z) = \sum_{n=0}^{\infty} d_n z^n$. More generally, we get for $j > 1$

$$P_{1j}(z) = z^{1-j}\, P_j(z)\, P_{11}(z) \tag{3.5}$$

where $P_j(z) = \sum_{n=j}^{\infty} p_n z^n$. Finally, using (1.1)-(1.2) along with (3.2), (3.4) and (3.5) we obtain

$$\begin{aligned} F_{ij}(z) &= z^{i-j}, & i > j, \\ F_{ij}(z) &= \frac{z^{i-j} P_j(z)}{1 - \sum_{0<n<j} p_n z^n}, & j \geq i. \end{aligned} \tag{3.6}$$

It is easy to show that the function $D(z)$ has no zeros for $|z| \leq 1$. Indeed, since $\sum_{n=1}^{\infty} p_n = 1$ and $p_n \geq 0$ we have $|\sum_{n=1}^{\infty} p_n z^n| < 1$ for $|z| < 1$. Furthermore, we know that $D(1) = m_1 = \sum_n n p_n$. Now suppose that $D(e^{i\theta}) = 0$ with $0 < \theta < 2\pi$. This implies $\sum_n e^{in\theta} p_n = 1$, or $\cos(n\theta) = 1$, $\forall n \geq 1$ s.t. $p_n > 0$. But this is impossibile because of our assumption on the probability vector p. Therefore the functions $P_{ij}(z)$ are analytic in the open unit disk with only one singularity at $z = 1$.

The precise nature of the latter depends on the asymptotic behaviour of the coefficients p_n. For example, if p_n decreases exponentially so does d_n and hence $P_{ij}(z)$ can be extended analytically to an open domain containing the unit disk, with the exception of a simple pole at $z = 1$. More generally, one sees that if the chain has finite ergodic degree, that is $\sum n^\ell p_n < \infty$ but $\sum n^{\ell+1} p_n = \infty$, for some $\ell \geq 0$, then $z = 1$ is a non-polar singular point for $P_{ij}(z)$. To see this, define the family of formal tail sequences $\{d_n^{(m)}\}$, $m \geq 1$, as $d_n^{(1)} = d_n$ and $d_n^{(m)} = \sum_{l>n} d_l^{(m-1)}$ for $m > 1$, and formal power series $D^{(m)}(z) = \sum d_n^{(m)} z^n$, so that, in particular, $D^{(1)}(z) = D(z)$. It is easy to realize that if P is ergodic with ergodic degree $\ell \geq 1$ then $\{d_n^{(m)}\}$ is defined for $m \leq \ell + 1$, $D^{(\ell)}(z)$ is absolutely convergent at $z = 1$, with $D^{(\ell)}(1) = d_0^{(\ell+1)}$, but $D^{(\ell+1)}(z)$ diverges at $z = 1$. Now, if $\ell = 0$ then clearly $(1 - z)P_{11}(z) \to 0$ but $P_{11}(z) \to \infty$ as $z \uparrow 1$. If, instead, $\ell \geq 1$, we define

$$H^{(\ell)}(z) = P_{11}(z) \, (d_0^{(\ell+1)} - D^{(\ell)}(z)) = \frac{D^{(\ell+1)}(z)}{D^{(1)}(z)}.$$

Again, we have $(1 - z)H^{(\ell)}(z) \to 0$ but $H^{(\ell)}(z) \to \infty$ as $z \uparrow 1$. This shows that for each of these functions, but in particular for $P_{11}(z)$, the point $z = 1$ is a non-polar singular point. Due to (3.2) and (3.5), similar statements hold for $P_{ij}(z)$ with arbitrary i, j.

II. Spectral properties of $P : \ell_1(I\!N) \to \ell_1(I\!N)$.

¿From (3.1)-(3.2) we have that the rate of convergence of $P^n(i,j)$ to π_j is not uniform in the departing state i (see also the Remark after the proof of Theorem 1). We are now going to see how this fact reflects in the nature of the spectrum of P in ℓ_1. In particular, the eigenvalue 1 is not isolated, even in the case where the p_n's are exponentially decreasing.

We study the structure of the spectrum of P using the method of generating functions (see, e.g., [VJ]). The vector equation $y = Px$, with $y = (y_1, y_2, \ldots)$ and $x = (x_1, x_2, \ldots)$, takes the form

$$Y(w) = \frac{X(w) - x_1 w(1-w)D(w)}{w}$$

where $Y(w) = \sum_{n=1}^{\infty} y_n w^n$ and $X(w) = \sum_{n=1}^{\infty} x_n w^n$. In a similar fashion, the formal solutions of the equations

$$(\lambda I - P)x = 0, \quad (\lambda I - P^*)\mathbf{x} = 0 \quad \text{and} \quad (\lambda I - P)x = y$$

can be written as

$$X(w) = \frac{x_1 w(1-w)D(w)}{1-\lambda w}, \tag{3.7}$$

$$X(w) = p \cdot \mathbf{x}\left(\frac{w}{\lambda - w}\right) \tag{3.8}$$

and

$$X(w) = \frac{x_1 w(1-w)D(w) - wY(w)}{1-\lambda w}, \tag{3.9}$$

respectively, where $p \cdot \mathbf{x} = \sum_{n \geq 1} x_n p_n$. The equation $1 - \lambda w = 0$ (and its reciprocal $\lambda - w = 0$) entails that the boundary of $\sigma(P)$ (and of $\sigma(P^*)$) is the unit circle. Let us first consider the point $\lambda = 1$. The formal expressions in (3.7) and (3.8) become

$$X(w) = x_1 w D(w) \quad \text{and} \quad X(w) = p \cdot \mathbf{x}\left(\frac{w}{1-w}\right).$$

The latter has the solution $X(w) = w/(1-w)$ which is the generating function of the unit vector in ℓ_∞. On the other hand, the former is the generating function of an ℓ_1-vector if and only if $D(1) < \infty$. Hence, we have that in the positive-recurrent case 1 lies in $\sigma_p(P)$ (for the null-recurrent chain it lies in $\sigma_r(P)$). More generally, one sees that the open unit disc $\{\lambda : |\lambda| < 1\}$ is always in the point spectrum, whereas any λ s.t. $|\lambda| = 1$, $\lambda \neq 1$ lies in $\sigma_c(P)$. Indeed, take $\lambda = e^{i\theta}$ with $0 < \theta < 2\pi$. Then the equation in (3.8) gives for the coefficients x_n

the relation $x_n = e^{-i(n+1)\theta}\, p\cdot x$. Multiplying by p_n and summing over n we then get $1 = \sum_n p_n e^{-i(n+1)\theta}$ which is impossible. Finally, one can easily check using (3.9) that any point λ with $|\lambda| > 1$ is in the resolvent set. Indeed, one can always produce a choice of x_1 in such a way that $x_1 w_0(1 - w_0)D(w_0) - w_0 Y(w_0) = 0$ where w_0 satisfies $1 - \lambda w_0 = 0$, thus letting (3.9) be the generating function of an ℓ_1-vector.

III. Convergence properties.

Next, we discuss the convergence properties of this chain under the hypothesis that it is positive-recurrent. In particular we shall prove statement 3) of Proposition 1 (the first part of which is a version of Theorem 1), using a different, more direct, method, which appears to be interesting in its own, for it may be extended to some more general (i.e. non-markovian) mixing Gibbs random fields [I]. In addition, this method allows one to obtain optimal bounds for the rate of convergence.

For $z \in \mathbb{C}$, consider the matrix L_z given by

$$L_z = \begin{pmatrix} p_1 z & p_2 z & p_3 z & \cdots \\ p_1 z^2 & p_2 z^2 & p_3 z^2 & \cdots \\ p_1 z^3 & p_2 z^3 & p_3 z^3 & \cdots \\ \vdots & \vdots & \vdots & \end{pmatrix}$$

For $z = 1$ the matrix L_z can be viewed as the transition matrix of the process $r_0^{(1)}, r_1^{(1)}, \ldots$ given by the sequence of times between returns to the state 1 (see (1.4)). The vector equation $y = L_z x$, takes the generating function form $Y(w) = p \cdot \mathbf{1}_w\, X(z)$ where $\mathbf{1}_w := (w, w^2, w^3, \ldots)^t$. We then have that, for z in the open unit disk, L_z defines a bounded operator-valued analytic function acting on $\ell_1(S)$, with bounded continuous extension to the boundary $|z| = 1$. We also have that, for each z in the closed unit disk, $\sigma_p(L_z) = \{\lambda_z\}$ and the eigenvalue λ_z is simple. To see this, it suffices to observe that the eigen-equation $(\lambda_z I - L_z)x = 0$ takes the form $p \cdot \mathbf{1}_w\, X(z) = \lambda_z\, X(w)$ which has the unique solution $\lambda_z = p \cdot \mathbf{1}_z$ and $X(w) = \mathrm{const.}\, p\cdot \mathbf{1}_w$, that is $x = \mathrm{const.}\, p$. The dual equation $(\lambda_z I - L_z^*)\mathbf{x} = 0$ gives the eigenvector $X(w) = zw/(1 - zw)$, that is $\mathbf{x} = \mathbf{1}_z$.

There is a simple algebraic relation between the matrices L_z and P: let Q be the transient chain given by the matrix

$$Q = \begin{pmatrix} 0 & 0 & 0 & \dots \\ 1 & 0 & 0 & \dots \\ 0 & 1 & 0 & \dots \\ 0 & 0 & 1 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

An easy calculation shows that

$$(I - zQ)(I - L_z) = (I - zP). \tag{3.10}$$

This relation entails that if u is an eigenvector of P with eigenvalue $1/z$, then $v = u(I - zQ)$ is an eigenvector of L_z with eigenvalue 1. On the other hand we already know that P, when acting on ℓ_1, has spectral radius equal to 1 and no eigenvalues on the unit circle besides eventually 1. The choice $z = 1$ gives $u = \pi$ and $v = \pi(I - Q) = \pi_1\, p$, as expected.

Let now $\mathbf{u} : S \to R$ be a bounded vector and ν an initial distribution on S, which will be assumed to decay not slower than π at infinity. The latter condition is equivalent to the assumption made in Theorem 1: if the P has ergodic degree $\ell \geq 1$ then π (ν) has P-order (at least) $\ell - 1$. In fact, ν can be viewed as the distribution of the random variable $1 + \min\{n : n \geq 0,\, x_n = 1\}$.

We then consider the following generating function,

$$S(z) = \sum_{n=0}^{\infty} z^n \left(\nu P^n - \pi\right) \cdot \mathbf{u}.$$

Using (3.10) and $\nu\, L_z = (\nu \cdot \mathbf{1}_z)\, p$ we have, for $|z| < 1$,

$$\begin{aligned} \sum_{n=0}^{\infty} z^n \,\nu P^n \cdot \mathbf{u} &= \nu(I - zP)^{-1} \cdot \mathbf{u} \\ &= \nu(I - L_z)^{-1}(I - zQ)^{-1} \cdot \mathbf{u} \\ &= \frac{(\nu \cdot \mathbf{1}_z)\, p\, (I - zQ)^{-1} \cdot \mathbf{u}}{1 - \lambda_z} + \nu\, (I - zQ)^{-1} \cdot \mathbf{u} \\ &= \frac{(\nu \cdot \mathbf{1}_z)\, (m_1\, \pi_z \cdot \mathbf{u})}{1 - \lambda_z} + \nu_z \cdot \mathbf{u} \end{aligned}$$

where $m_1 = \pi_1^{-1} = D(1)$, $\nu_z = \nu\,(I - zQ)^{-1}$ and $\pi_z = \pi_1 p\,(I - zQ)^{-1}$ (in particular $\pi_z|_{z=1} \equiv \pi$). Therefore

$$\begin{aligned} S(z) &= \frac{(m_1\,\pi_z\cdot\mathbf{u})\,(\nu\cdot\mathbf{1}_z)}{1-\lambda_z} - \frac{(\pi\cdot\mathbf{u})\,(\nu\cdot\mathbf{1})}{1-z} + \nu_z\cdot\mathbf{u} \\ &= \frac{(m_1\,\pi\cdot\mathbf{u})\,(\nu\cdot\mathbf{1})}{1-\lambda_z} - \frac{(\pi\cdot\mathbf{u})\,(\nu\cdot\mathbf{1})}{1-z} + R(z) \\ &= (\pi\cdot\mathbf{u})\,(\nu\cdot\mathbf{1})\,H(z) + R(z) \end{aligned}$$

where we have used the notation

$$H(z) = \frac{m_1}{1-\lambda_z} - \frac{1}{1-z} = \frac{\sum_{n=0}^{\infty} e_n\,z^n}{\sum_{n=0}^{\infty} d_n\,z^n}, \quad \text{with} \quad e_n = \sum_{k>n} d_k,$$

and

$$R(z) = \frac{(m_1\,\pi_z\cdot\mathbf{u})\,(\nu\cdot\mathbf{1}_z) - (m_1\,\pi\cdot\mathbf{u})\,(\nu\cdot\mathbf{1})}{1-\lambda_z} + \nu_z\cdot\mathbf{u}.$$

Reasoning as in the proof of Lemma 3, one then shows that if $M_{11}^{(\ell)} < \infty$ then the coefficients of $H(z)$ decay as $o(n^{-(\ell-1)})$. It remains to examine the behaviour of $R(z)$. We first rewrite it as

$$R(z) = \frac{(m_1\pi\cdot\mathbf{u})\,(\nu\cdot\mathbf{1}_z - \nu\cdot\mathbf{1})}{1-\lambda_z} + \frac{(\nu\cdot\mathbf{1}_z)\,(m_1\pi_z\cdot\mathbf{u} - m_1\pi\cdot\mathbf{u})}{1-\lambda_z} + \nu_z\cdot\mathbf{u}.$$

We have

$$\frac{\nu\cdot\mathbf{1} - \nu\cdot\mathbf{1}_z}{1-\lambda_z} = \frac{\sum_{n=0}^{\infty} \eta_n\,z^n}{\sum_{n=0}^{\infty} d_n\,z^n}, \quad \text{with} \quad \eta_n = \sum_{k>n} \nu_k.$$

Moreover, a straightforward calculation yields

$$m_1\pi_z\cdot\mathbf{u} = \sum_{n=0}^{\infty} z^n \left(\sum_{k=1}^{\infty} u_k\,p_{n+k}\right)$$

and therefore

$$\frac{m_1\pi\cdot\mathbf{u} - m_1\pi_z\cdot\mathbf{u}}{1-\lambda_z} = \frac{\sum_{n=0}^{\infty} \xi_n z^n}{\sum_{n=0}^{\infty} d_n\,z^n}, \quad \text{with} \quad \xi_n = \sum_{k=1}^{\infty} u_k\,d_{k+n}.$$

In a similar way, we get

$$\nu_z\cdot\mathbf{u} = \sum_{n=0}^{\infty} \gamma_n\,z^n = \frac{\sum_{n=0}^{\infty} \left(\sum_{l=0}^{n} \gamma_l\,d_{n-l}\right) z^n}{\sum_{n=0}^{\infty} d_n\,z^n}, \quad \text{with} \quad \gamma_n = \sum_{k=1}^{\infty} u_k\,\nu_{n+k}.$$

In order to compare the term $\nu_z \cdot \mathbf{u}$ with $H(z)$, we rewrite it as

$$\nu_z \cdot \mathbf{u} = \frac{\sum_{n=0}^{\infty} \left(\sum_{l=0}^{n} \gamma_l \, d_{n-l}\right) z^n}{\sum_{n=0}^{\infty} d_n \, z^n}.$$

Lemma 2 then entails that the coefficients of the power series at the numerator behave asymptotically as $m_1 \, \gamma_n$. In the same way, the coefficients of the product $\sum_{n=0}^{\infty} \xi_n z^n \cdot \sum_{n=0}^{\infty} \nu_n z^n$ have the same behaviour as $(\nu \cdot \mathbf{1}) \, \xi_n$. Moreover we have

$$|\xi_n| \le \|u\|_\infty \sum_{k>n} d_k = \|u\|_\infty \, e_n, \qquad |\gamma_n| \le \|u\|_\infty \sum_{k>n} \nu_k = \|u\|_\infty \, \eta_n.$$

Finally, comparing all the above terms, we have found that under our assumptions on the distribution ν and the vector $\mathbf{u}$, and under the hypothesis that $M_{11}^{(\ell)} < \infty$, the quantity $(\nu P^n - \pi) \cdot \mathbf{u}$ decays as $o(n^{-(\ell-1)})$.

We conclude by showing how this bound can be sharpened under the hypothesis that $p_n \sim n^{-(\ell+2)} \, L(n)$ with $L(n)$ slowly varying at infinity. We start by showing that the coefficients of the power series of $H(z)$ are asymptotically equal to $\ell^{-1}(\ell+1)^{-1} D(1)^{-1} n^{-\ell} \, L(n)$. To see this, suppose first that $\ell = 1$. Then $D(1) < \infty$ and we have $d_n \sim (\ell+1)^{-1} n^{-(\ell+1)} \, L(n)$, $e_n \sim \ell^{-1}(\ell+1)^{-1} n^{-\ell} \, L(n)$. Notice that the power series $\sum_{n=0}^{\infty} e_n z^n$ is divergent at $z = 1$. Lemma 2, or, alternatively, a repeated application of a Tauberian theorem for power series (see, e.g., [F2], Chap. XIII.5, Thm. 5) then yield the claim. More generally, if $\ell > 1$ we have (with the notation of I) that the numbers $\{d_n^{(m)}\}$ are defined for $m \le \ell + 1$, $D^{(\ell)}(1) = d_0^{(\ell+1)}$, but $D^{(\ell+1)}(z)$ diverges at $z = 1$. We can then write

$$\begin{aligned} H(z) \cdot \sum_{n=0}^{\infty} d_n z^n = (z-1)^{\ell-1} \, D^{(\ell+1)}(z) + \\ + \sum_{m=3}^{\ell+1} (z-1)^{m-3} \left(d_0^{(m-1)} + d_0^{(m)} \right) \end{aligned}$$

and the claim follows as above. Finally, under the additional hypotheses that $u_i = o(1)$ and $\nu_i = o(\pi_i)$ we have $\xi_n = o(e_n)$ and $\eta_n = o(e_n)$. This prevents from possible cancellations among the various coefficients introduced above and yields the asymptotic behaviour claimed in Proposition 1.

Appendix. *Denumerable Markov chains and intermittent interval maps.*

The example discussed in Section 3 is isomorphic (mod 0) to the iteration process of the piecewise affine 'intermittent' map $f : [0,1] \to [0,1]$ given by

$$f(x) = \begin{cases} (x-d_1)/\alpha_1, & \text{if } d_1 \le x \le d_0 \\ d_{i-1} + (x-d_i)/\alpha_i, & \text{if } d_i \le x < d_{i-1},\ i \ge 2 \end{cases} \tag{A.1}$$

Here the numbers $d_i = \sum_{l>i} p_l$ are supposed to be all distinct, and $\alpha_i = p_i/p_{i-1}$, $i \ge 1$ (with $p_0 = 1$). In what follows we shall always assume that $\sum d_i < \infty$. This map is named 'intermittent' for, if $\overline{\lim}\, \alpha_i = 1$, then f can be viewed as a piecewise affine approximation of a smooth transformation of $[0,1]$ which is expanding everywhere but at the fixed point in the origin, where the derivative is equal to one (see [W], [LSV], [M]).

The partition $\mathcal{I}$ of $[0,1]$ into the intervals $I_i = [d_i, d_{i-1}]$, $i \ge 1$ is a Markov partition for f. Let Ω, π and P be as in the Example of Section 3. One then sees that the map $\phi : \Omega \to [0,1]$ defined by: $\phi(\omega) = x$ according to $f^j(x) \in I_{\omega_j}$, $j \ge 0$, is a bijection between Ω and the residual set of points in $(0,1]$ which are not preimages of 1 w.r.t. the map f. Clearly ϕ conjugates f with the shift τ on Ω. Moreover, let μ be the τ-invariant Markov probability measure on Ω defined in (1.8) (with π and P as above). Then $\rho = \mu \circ \phi^{-1}$ is f-invariant and it is easy to see that $P(i,j) = \rho\,(I_i \cap f^{-1} I_j)/\,\rho(I_i)$. Finally, if f is the piecewise affine approximation of a smooth transformation of $[0,1]$ having a tangency at $x = 0^+$ of order $1+1/\eta$, with $\eta > 0$, then $p_i \sim i^{-(1+\eta)}$ and hence $\alpha_i \sim 1-(1+\eta)/i$. Thus, in order to have $\sum d_i < \infty$ it is necessary that $\eta > 1$, and the corresponding Markov chain P has ergodic degree $\ell = \eta - 1$.

It turns out that an effective tool for studying the statistical properties of the iterates of a (non-invertible) transformation f on $[0,1]$ is given by Perron-Frobenius operator $M : L^1([0,1], dx) \to L^1([0,1], dx)$, which is defined by the validity of (see, e.g., [LM])

$$\int_0^1 u \circ f^n(x)\, v(x)\, dx = \int_0^1 u(x)\, M^n v(x)\, dx \tag{A.2}$$

for all pairs $u, v \in L^1$. For f as in (A.1), $n = 1$ and $d_i \leq x < d_{i-1}$, (A.2) yields

$$(M\,v)(x) = \alpha_{i+1}\, v\,(d_{i+1} + \alpha_{i+1}(x - d_i)) + \alpha_1\, v\,(d_1 + \alpha_1 x). \qquad (A.3)$$

Let us note that the space $\ell_1(S, p)$ of vectors $u : I\!\!N \to I\!\!R$ such that

$$\|u\|_{1,p} := \sum_{i \in S} |u_i|\, p_i < \infty$$

is left invariant by the operator M, which takes on the matrix representation

$$M(i,j) = \frac{p_i}{p_j}\, P(i,j), \qquad i, j \geq 1. \qquad (A.4)$$

The eigenequation $M\,h = h$ has a solution $h \in \ell_1(S,p)$ given by $h_i = h_1\, p_1\, d_{i-1}/p_i$. Moreover, one easily check that the vector p satisfies $M^* p = p$. Therefore, recalling that $\pi_i = \pi_1\, d_{i-1}$, and putting $h_1 = \pi_1/p_1$, we get $\pi_i = h_i\, p_i$. One then sees that the vector $h \in \ell_1(S,p)$ corresponds to the (locally constant) density of the absolutely continuous f-invariant probability measure $\rho(dx) = h(x)\, dx$, with $h \in L^1([0,1], dx)$ and $h(x) \equiv h_i$ for $d_i \leq x < d_{i-1}$. Observe that $\rho(I_i) = \pi_i$. Now, using (A.2) we find

$$\rho(u \circ f^n\, v) - \rho(u)\, \rho(v) = \int_0^1 u(x)\,[\,(M^n\, v h)(x) - \rho(v)\, h(x)]\, dx \qquad (A.5)$$

Suppose that u and v are bounded L^1-functions taking constant values u_i and v_i on the elements I_i of the Markov partition $\mathcal{I}$. We shall denote by $\mathbf{u} = (u_i)_{i=1}^\infty$ and $\mathbf{v} = (v_i)_{i=1}^\infty$ the corresponding vectors in $\ell_\infty(I\!\!N)$. Using (A.4), (A.5) and the above observations we get (the notation is as in the proof of Corollary 1),

$$\rho(u \circ f^n\, v) - \rho(u)\, \rho(v) = (\overline{\pi \mathbf{v}}\, P^n - (\pi \cdot \mathbf{v})\, \pi) \cdot \mathbf{u}. \qquad (A.6)$$

Now set $u_\infty = \overline{\lim}\, u_i$, $v_\infty = \overline{\lim}\, v_i$, and suppose that $u_\infty \neq 0$ or $v_\infty \neq 0$. Then, setting $\hat{\mathbf{u}} = \mathbf{u} - u_\infty \mathbf{1}$ and $\hat{\mathbf{v}} = \mathbf{v} - v_\infty \mathbf{1}$ have that $(\pi \cdot \hat{\mathbf{u}})(\pi \cdot \hat{\mathbf{v}}) \neq 0$ provided $\pi \cdot \mathbf{u} \neq u_\infty$ and $\pi \cdot \mathbf{v} \neq v_\infty$. Moreover $\overline{\lim}\, \hat{u}_i = 0$ and $\overline{\lim}\, \hat{v}_i = 0$. On the other hand we plainly have $(\overline{\pi \hat{\mathbf{v}}}\, P^n - (\pi \cdot \hat{\mathbf{v}})\, \pi) \cdot \hat{\mathbf{u}} = (\overline{\pi \mathbf{v}}\, P^n - (\pi \cdot \mathbf{v})\, \pi) \cdot \mathbf{u}$. We then see

that the conditions $\rho(u) \equiv \pi \cdot \mathbf{u} \neq u_\infty$ and $\rho(v) \equiv \pi \cdot \mathbf{v} \neq v_\infty$ are equivalent to the conditions $u_i = o(1)$ and $\nu_i = o(\pi_i)$ (along with $(\pi \cdot \mathbf{u})(\nu \cdot 1) \neq 0$) assumed in the last statement of Proposition 1, with the identification $\nu = \overline{\pi \mathbf{v}}/\pi \cdot \mathbf{v}$.
The following result is now a direct consequence of Proposition 1 (see [LSV], [M] for related results):

Corollary 2. *Let $f : [0,1] \to [0,1]$ be as in (A.1) and assume that $\alpha_i \sim 1 - (1+\eta)/i$ for some $\eta > 1$. Then, for any pair of bounded L^1-functions $u, v : [0,1] \to I\!R$, locally constant on the Markov partition $\mathcal{I}$, there is a positive constant $C = C(u,v)$ such that, for n large enough,*

$$| \, \rho(u \circ f^n \, v) - \rho(u) \, \rho(v) \, | \leq C \, n^{-(\eta-1)}.$$

Assume furthermore that $\rho(u) \neq u_\infty$ and $\rho(v) \neq v_\infty$. Then we have

$$\rho(u \circ f^n \, v) - \rho(u) \, \rho(v) \sim C \, n^{-(\eta-1)} \quad \text{as} \quad n \to \infty.$$

We conclude with a final remark. ¿From the proof of Proposition 1 it follows that if the conditions $\rho(u) \neq u_\infty$ and $\rho(v) \neq v_\infty$ are violated, then cancellations may take place to accelerate the convergence rate. A trivial example is obtained by taking u, v constant on $[0,1]$. Conversely, one may argue as follows: take $a > 0$ and let $t_a(x)$ be the first entrance time into the set $[a, 1]$. When an orbit falls in a small (compared to a) neighbourhood of 0 it stays there for a time which can be arbitrarily large before reaching again $[a, 1]$. More precisely, from the above discussion one readily finds that, under the assumptions of Corollary 2,

$$\rho\{x \in [0,1] \, : \, t_a(x) > n\} \sim C(a) \, n^{-(\eta-1)} \quad \text{as} \quad n \to \infty.$$

Thus, if the condition is satisfied, namely if the average value of the test functions is reached away from the origin, then the term $\rho(u \circ f^n \, v)$ cannot approach its asymptotic value $\rho(u) \, \rho(v)$ at a rate faster than that given by the statistics of first entrance times given above.

References.

[C] K L Chung: *Markov chains with stationary transition probabilities*, Springer-Verlag Berlin Heidelberg New York 1967.

[F1] W Feller: Fluctuation theory of recurrent events, *TAMS* **67** (1949), 99-119.

[F2] W Feller: *An Introduction to Probability Theory and Its Applications*, Volume 2, J.Wiley and Sons, New York 1970.

[FMM] G Fayolle, V A Malyshev and M V Menshikov: *Topics in the constructive theory of countable markov chains*, Cambridge University Press, Cambridge 1992.

[H] T E Harris: First passage and recurrence distribution, *TAMS* **73** (1952), 471-486.

[I] S Isola: Dynamical zeta functions and correlation functions for intermittent interval maps, Preprint 1996.

[KSK] J G Kemeny, J L Snell and A W Knapp: *Denumerable Markov Chains*, Springer-Verlag, 1976.

[L] T Lindvall: *Lectures on the Coupling Method*, John Wiley and Sons, Inc. 1992.

[LM] A Lasota and M C Mackey: *Probabilistic properties of deterministic systems*, Cambridge University Press, 1985.

[LSV] A Lambert, S Siboni and S Vaienti, Statistical properties of a non-uniformly hyperbolic map of the interval, *J. Stat. Phys.* **72** (1993), 1305-1330.

[M] M Mori, On the intermittency of a piecewise linear map, *Tokyo J. Math.* **16** (1993), 411-428.

[T] E C Titchmarsh: *The Theory of Functions*, 2nd ed., Oxford University Press, 1939.

[VJ] D Vere-Jones: On the spectra of some linear operators associated with queueing systems, *Z. Wahrsch.* **2** (1963), 12-21.

[W] X J Wang, Statistical physics of temporal intermittency, *Phys. Rev.* **A40** (1989), 6647.

[Z] A Zygmund: *Trigonometrical series*, Cambridge at the University Press, 1959.