A CONTRACTION THEOREM FOR MARKOV CHAINS ON GENERAL STATE SPACES

THOMAS KAIJSER

Communicated by Lucian Beznea

Let $\{X_n, n = 0, 1, 2, ...\}$ denote a Markov chain on a general state space and let f be a nonnegative function. The purpose of this paper is to present conditions which will imply that $f(X_n)$ tends to 0 a.s., as n tends to infinity. As an application we obtain a result on synchronisation for random dynamical systems. At the end of the paper, we also present a result on convergence in distribution for Markov chains on general state spaces, thereby generalising a similar result for Markov chains on compact metric spaces.

AMS 2010 Subject Classification: 60J05, 60F15, 60F05.

Key words: functions of Markov chains, synchronisation, convergence in distribution, random dynamical systems, iterated function systems.

1. INTRODUCTION

Let $(K, \mathcal{E})$ be a separable, measurable, space. Let $P : K \times \mathcal{E} \to [0, 1]$ be a transition probability function (tr.pr.f) on $(K, \mathcal{E})$ and, for $x \in K$, let $\{X_n(x), n = 0, 1, 2, ...\}$ denote the Markov chain generated by the starting point x and the tr.pr.f $P : K \times \mathcal{E} \to [0, 1]$.

Next, let $f : K \to [0, \infty)$ be a nonnegative, measurable function. For $a > 0$ we define

$$K(a) = \{x \in K : 0 < f(x) < a\},$$

and we define

$$K(0) = \{x \in K : f(x) = 0\}.$$

If we want to emphasize the dependence of f we may write $K_f(a)$ instead of $K(a)$. We denote the complement of $K(0)$ by K'. Thus $K' = \{x \in K : f(x) > 0\}$.

If the set $K_f(0)$ is such that

$$x \in K(0) \Rightarrow f(X_1(x)) \in K(0) \;\; a.s. \;,$$

then we say that $K_f(0)$ is *closed* under P.

We say that the set $K_f(0)$ is *absorbing* with respect to the tr.pr.f P, if for all $x \in K$

$$f(X_n(x)) \to 0 \;\; a.s. \;. \tag{1}$$

The *purpose* of this paper is to introduce conditions which will imply that the set $K(0)$ is absorbing.

We shall first introduce the following *regularity* condition.

Definition 1.1. If the equality

$$Pr[f(X_1(x)) > 0] = 1,$$

holds for all $x \in K'$, we say that Condition R holds and that the couple (f, P) is *regular*.

We shall next present three conditions which together with Condition R will imply that $K(0)$ is absorbing with respect to P.

The first condition is the most important one.

Definition 1.2. We say that the pair (f, P) has the *geometric mean contraction (GMC) property*, if there exist a number $\epsilon_0 > 0$, a number $\kappa_0 > 0$ and an integer N_0, such that, if $x \in K(\epsilon_0)$, then

$$E[\log(f(X_{N_0}(x)))] < -\kappa_0 + \log f(x). \tag{2}$$

Before continuing with the next two conditions, let us consider the following, somewhat stronger condition for comparison.

Definition 1.3. We say that the pair (f, P) has the *arithmetic mean contraction (AMC) property*, if there exist a number $\epsilon_0 > 0$, a number ρ_0, $0 < \rho_0 < 1$, and an integer N_0 such that, if $x \in K(\epsilon_0)$, then

$$E[f(X_{N_0}(x))] < \rho_0 f(x).$$

Clearly the AMC-property implies the GMC-property because of Jensen's inequality.

In order for the GMC-property to be useful it is necessary that the Markov chain $\{X_n(x), n = 0, 1, 2, ...\}$ - for every $\epsilon > 0$ and every $x \in K$ - sooner or later enters the set $K(\epsilon)$. The following condition is introduced for this reason.

Definition 1.4. We say that the pair (f, P) satisfies *Condition C* if for every $\epsilon > 0$, every $\xi > 0$ and every $x \in K'$ we can find an integer N such that

$$Pr[f(X_n(x)) \geq \epsilon, \;\; n = 0, 1, 2, ..., N] < \xi.$$

Our third condition is a second order moment condition.

Definition 1.5. We say that *Condition B* holds if there exists a constant $\epsilon_1 > 0$ such that for $n = 1, 2, ...$

$$\sup_{x \in K(\epsilon_1)} E[|\log f(X_n(x)) - \log f(x)|^2] < \infty. \tag{3}$$

THEOREM 1.1. *Let $(K, \mathcal{E})$ be a measurable space, and let $P : K \times \mathcal{E} \to [0,1]$ be a tr.pr.f on $(K, \mathcal{E}, \delta)$. Let $f : K \to [0, \infty)$ be a nonnegative, measurable function, and suppose that (f, P) has the GMC-property. Suppose also that Condition C and Condition B are satisfied, that $K(0)$ is closed under P, and that (f, P) is regular. Then $K(0)$ is absorbing.*

The plan of the paper is as follows. In the next section, we give some background and show how Theorem 1.1 can be used to prove *synchronisation.* In Section 3, we give a very brief sketch of the proof of Theorem 1.1 and in Section 4 we give the details. In Section 5 finally, we prove a theorem on *convergence in distribution* related to Theorem 1.1.

2. BACKGROUND AND MOTIVATION

Let $(S, \mathcal{F}, \delta_0)$ be a separable, measurable space with metric δ_0, let $(A, \mathcal{A})$ be another measurable space, let $h : S \times A \to S$ be a measurable function and let μ be a probability measure on $(A, \mathcal{A})$. The triple $\{(S, \mathcal{F}, \delta_0), (A, \mathcal{A}, \mu), h\}$ is often called a *random dynamical system* (r.d.s) or an *iterated function system* (i.f.s).

Let $(A^n, \mathcal{A}^n)$, $n = 1, 2...$ be defined recursively by $(A^1, \mathcal{A}^1) = (A, \mathcal{A})$, $A^{n+1} = A^n \times A$, $\mathcal{A}^{n+1} = \mathcal{A}^n \otimes \mathcal{A}$ and define $h^n : S \times A^n \to S$ recursively by $h^1 = h$ and

$$h^{n+1}(s, a^{n+1}) = h(h^n(s, a^n), a_{n+1}), \tag{4}$$

where thus $(a_1, a_2, ..., a_n) = a^n$ denotes a generic element in A^n. Let us also assume that for each $a \in A$ we have

$$\gamma(h, a) = \sup\{\frac{\delta_0(h(s, a), h(t, a))}{\delta_0(s, t)} : s \neq t\} < \infty. \tag{5}$$

If (5) holds for all $a \in A$ we say that *Condition L* holds.

In the paper [3] from 1978, the following three conditions were introduced (formulated slightly different than the formulations used below). First though a few more notations.

Let $\{Z_n, n = 1, 2, ...\}$ be a sequence of independent stochastic variables with values in $(A, \mathcal{A})$ and distribution μ. In agreement with the terminology in [4] we call $\{Z_n, n = 1, 2, ...\}$ the *index sequence* associated to the r.d.s $\{(S, \mathcal{F}, \delta_0), (A, \mathcal{A}, \mu), h\}$. We write $Z^N = (Z_1, Z_2, ..., Z_N)$. For $s \in S$ and $\epsilon > 0$ we define $B(s, \epsilon) = \{t \in S : \delta_0(s, t) < \epsilon\}$. If $g : S \to S$ we define

$$r_\epsilon g : S \to [0, \infty]$$

by

$$r_\epsilon g(s) = \sup\{\frac{\delta_0(g(t'), g(t''))}{\delta_0(t', t'')} : t','' \in B(s, \epsilon), t' \neq t''\}.$$

Definition 2.1. Suppose there exist an integer N, a number $\epsilon > 0$, and a number $\kappa_0 > 0$ such that for all $s \in S$

$$E[\log r_\epsilon h^N(s, Z^N)] \leq -\kappa_0$$

then we say that *Condition G′* holds.

If furthermore there exists a constant C such that for all $s \in S$ also

$$E[(\log^+ r_\epsilon h^N(s, Z^N))^2] \leq C$$

then we say that *Condition B′* holds.

Finally, if also, to every $\epsilon > 0$, we can find an integer M and a number $\alpha_0 > 0$, such that, if $(Z_1, Z_2, ..., Z_M) = Z^M$ and $(W_1, W_2, ..., W_M) = W^M$ are two independent sequences of independent stochastic variables with distribution μ, it follows that for any two points s_1 and s_2 in S we have

$$Pr[\delta(h^M(s_1, Z^M), h^M(s_2, W^M)) < \epsilon] \geq \alpha_0$$

then we say that *Condition C′* holds.

The following theorem was proved in [3].

THEOREM 2.1. *Let $\{(S, \mathcal{F}, \delta_0), (A, \mathcal{A}, \mu), h\}$ be a r.d.s such that $(S, \mathcal{F}, \delta_0)$ is a compact metric space. Suppose that Condition G′, Condition B′, Condition C′ and Condition L hold. Then there exists a unique probability measure ν on $(S, \mathcal{F}, \delta_0)$, such that, for each $s \in S$, the distribution $\mu_{n,s}$ of $h^n(s, Z^n)$ converges in distribution towards ν.*

The motivation for Theorem 2.1 was that it could be applied in order to show that the so called *angle process* associated to *products of random matrices* converges to a unique limit measure. (Today this measure is often called the Furstenberg measure, see *e.g.* [1].)

In the last two decades there has been much interest in the problem of *synchronisation* for random dynamical systems (see e.g the reference lists in [7] and [8]). The question, that one is interested in, is the following:
If $\{(S, \mathcal{F}, \delta_0), (A, \mathcal{A}, \mu), h\}$ is a r.d.s, the sequence $\{Z_n, n = 1, 2, ...\}$ is the associated index sequence and $s_1, s_2 \in S$, when does it hold that

$$\lim_{n\to\infty} \delta_0(h^n(s_1, Z^n), h^n(s_2, Z^n)) = 0, \ a.s. ?$$

As a corollary to Theorem 1.1 we almost immediately obtain the following result regarding synchronisation of random dynamical systems.

First though, some further definitions. The following condition is slightly weaker than Condition G'.

Definition 2.2. Let $\{(S, \mathcal{F}, \delta_0), (A, \mathcal{A}, \mu), h\}$ be a r.d.s and let $\{Z_n, n = 1, 2, ...\}$ denote the associated index sequence. Suppose there exist a number

$\epsilon_0 > 0$, a number $\kappa_0 > 0$ and an integer N_0 such if s_1 and s_2 in S and $\delta_0(s_1, s_2) < \epsilon_0$ then

$$E[\log \frac{\delta_0(h^{N_0}(s_1, Z^{N_0}), h^{N_0}(s_2, Z^{N_0}))}{\delta_0(s_1, s_2)}] \leq -\kappa_0.$$

We then say that *Condition G1* is satisfied.

Our next condition is essentially the same as Condition B above. Let us first set

$$D(\epsilon) = \{(s,t) \in K \times K : 0 < \delta_0(s,t) < \epsilon\}.$$

Definition 2.3. Let $\{(S, \mathcal{F}, \delta_0), (A, \mathcal{A}, \mu), h\}$ be a r.d.s and let $\{Z_n, n = 1, 2, ...\}$ denote the associated index sequence. We say that *Condition B1* holds, if there exists a constant $\epsilon_1 > 0$ such that

$$\sup_{(s_1,s_2) \in D(\epsilon_1)} E[|\log \frac{\delta_0(h^n(s_1, Z^n), h^n(s_2, Z^n))}{\delta_0(s_1, s_2)}|^2] < \infty.$$

Our third condition is essentially the same as Condition C above.

Definition 2.4. Let $\{(S, \mathcal{F}, \delta_0), (A, \mathcal{A}, \mu), h\}$ be a r.d.s and let $\{Z_n, n = 1, 2, ...\}$ denote the associated index sequence. We say that *Condition C1* holds if for every $\epsilon > 0$, every $\xi > 0$ and every pair $s_1, s_2 \in S$ there exists an integer N, such that

$$Pr[\inf\{\delta_0(h(s_1, Z^n), h(s_2, Z^n)), n = 0, 1, 2, ..., N\} \geq \epsilon] < \xi.$$

THEOREM 2.2. *Let $\{(S, \mathcal{F}, \delta_0), (A, \mathcal{A}, \mu), h\}$ be a r.d.s such that $(S, \mathcal{F}, \delta_0)$ is separable, let $\{Z_n, n = 1, 2, ...\}$ denote the associated index sequence, and suppose that Condition G1, Condition B1 and Condition C1 hold. Suppose also that*

$$Pr[\delta_0(h(s, Z_1), h(t, Z_1)) > 0] = 1 \;\; if \;\; \delta_0(s,t) > 0. \tag{6}$$

Let $s_1, s_2 \in S$. Then

$$\lim_{n \to \infty} \delta_0(h^n(s_1, Z^n), h^n(s_2, Z^n)) = 0$$

almost surely.

Proof. Let $K = S \times S$, let $\mathcal{E} = \mathcal{F} \otimes \mathcal{F}$, define $\tilde{h} : K \times A \to K$ by

$$\tilde{h}((s_1, s_2), a) = (h(s_1, a), h(s_2, a)).$$

Since S is separable and h is $\mathcal{F}$ - measurable it follows that $\tilde{h}$ is $\mathcal{E}$ - measurable.

Next define the tr.pr.f $P : K \times \mathcal{E} \to [0, 1]$ by

$$P[(s_1, s_2), E] = \mu(A((s_1, s_2), E))$$

where

$$A((s_1, s_2), E) = \{a \in A : \tilde{h}((s_1, s_2), a) \in E\}.$$

Since $\tilde{h}$ is $\mathcal{E} - measurable$ it is well-known that P is a tr.pr.f. Finally define $f : K \to [0, \infty)$ by

$$f((s_1, s_2)) = \delta_0(s_1, s_2). \tag{7}$$

Evidently f is a continuous function and hence measurable.

From Condition G1 follows that (f, P) has the GMC-property, from Condition B1 follows that (f, P) satisfies Condition B and from Condition C1 follows that (f, P) satisfies Condition C. Since $(S, \mathcal{F}, \delta_0)$ is a separable, metric space, it follows that $K = S \times S$ is separable. Since also $K(0) = \{x \in K : f(x) = 0\} = \{(s, t) \in S \times S : \delta(s, t) = 0\} = \{(s, t) \in S \times S : s = t\}$, it follows that the set $K(0)$ is closed under P. Finally from (6) follows that (f, P) is regular. Hence all hypotheses of Theorem 1.1 are fulfilled. Hence, if we as usual let $\{X_n(x), n = 0, 1, 2, ...\}$ denote the Markov chain generated by the tr.pr.f P and the initial value $x \in K$, it follows from Theorem 1.1 that for all $x \in K$

$$f(X_n(x)) \to 0 \ \ a.s. \tag{8}$$

as $n \to \infty$.

Since, for $n = 1, 2, ...$ and $(s_1, s_2) \in S \times S$

$$X_n((s_1, s_2))) = (h^n(s_1, Z^n), h^n(s_2, Z^n)) \ a.s.$$

it follows from (8) that

$$\delta_0(h^n(s_1, Z^n), h^n(s_2, Z^n)) \to 0, \quad a.s.$$

as $n \to \infty$, which was what we wanted to prove. □

Before ending this section let me mention one reason for formulating Theorem 1.1 using an unspecified function f, instead of just letting f be defined as in the proof of Theorem 2.2. (See (7).) When writing up this paper I thought that Theorem 1.1 could be of use when trying to prove that the area of the nth normalised random subdivision of a convex polygon tends to zero by letting f be defined as the area of the convex polygon. However it turned out to be more difficult to do this than I had anticipated. (See [10] and [9] for further details on random subdivisions of convex polygons.)

All the same, I am quite sure that other situations will arise where the problem will be to prove that a pair (f, P), consisting of a tr.pr.f P on a measurable space and a nonnegative function f on this space, is such, that the set $\{f(x) = 0\}$ is *absorbing* with respect to P.

3. SKETCH OF PROOF OF THEOREM 1.1

The proof of Theorem 1.1 is in principle quite easy.

Thus, let $\epsilon > 0$ and $\eta > 0$ be given. The idea is simply to show that if ϵ' is chosen sufficiently small - and much smaller then ϵ, and we start our Markov chain in $K(\epsilon')$ then the probability that the Markov chain ever leaves the set $K(\epsilon)$ is less then say $\eta/2$. In order to find such an ϵ' we shall – first of all – use Condition G, but shall also need Condition B.

We use Condition B in two ways. First of all, it allows us to use Chebychev's inequality. Secondly, it allows us to have control over $f(X_n(x))$ for $n = 1, 2, ..., N_0 - 1$, where thus N_0 is the integer in the definition of the GCM-property (see Definition 1.2).

From Condition C, it follows that for every $x \in K$ we can find an integer N such that the probability that the Markov chain when starting at x has not entered the set $K(\epsilon')$ before time N is less then $\eta/2$. By combining this fact with the fact that the probability the Markov chain ever leaves the set $K(\epsilon)$ once it has entered the set $K(\epsilon')$ is less then $\eta/2$, it follows that lim sup of $f(X_n(x))$ is less then η, which implies that $f(X_n(x))$ tends to 0 almost surely.

This is the strategy of the proof. To fill in the details is by no means difficult, but requires a few pages.

4. PROOF OF THEOREM 1.1

That (1) holds, if $x \in K(0)$, is a trivial consequence of the fact that we have assumed that $K_f(0)$ is closed with respect to the tr.pr.f P.

In order to prove (1) for $x \in K'$ we have to show that for every $L > 0$ and every $\eta > 0$ we, for each $x \in K'$, can find an integer N such that

$$Pr[\sup\{\log f(X_n(x)) : n \geq N\} > -L] < \eta. \tag{9}$$

Thus let $\eta > 0$ and $L > 0$ be given. In order to show that we for each $x \in K$ can find an integer N such that (9) holds, we shall use arguments quite similar to arguments used when proving for example Hajek-Renyi's theorem. (See e.g [2], Satz 36.2.)

Let ϵ_0, κ_0 and N_0, be such that (2) holds if $x \in K(\epsilon_0)$ and let ϵ_1 be such that (3) holds. Set

$$\beta = \min\{\epsilon_0, \epsilon_1, e^{-L}\}. \tag{10}$$

Next set

$$c_1 = \sup_{x \in K(\beta)} E[|\log f(X_1(x)) - \log f(x)|^2], \tag{11}$$

set

$$c_2 = \sup_{x \in K(\beta)} E[|\log f(X_{N_0}(x)) - \log f(x)|^2] \tag{12}$$

and define

$$b = \sup_{x \in K(\beta)} E[\log f(X_1(x)) - \log f(x)].$$

That $-\infty < b < \infty$ follows from (11) and Schwartz inequality.

The following lemma will be used repeatedly.

LEMMA 4.1. *Let $x \in K(\beta)$, set $y_0 = \log f(x)$ and set $Y_n = \log f(X_n(x))$ for $n = 1, 2, \ldots$.*

a) If $t > b$ then

$$Pr[Y_1 > y_0 + t] < \frac{c_1}{(t-b)^2}.$$

b) If $t > -\kappa_0$ then

$$Pr[Y_{N_0} > y_0 + t] < \frac{c_2}{(t+\kappa_0)^2}.$$

c) If $t > -\kappa_0$ and $y_0 + t < \log \beta$ then for $k = 1, 2, \ldots,$

$$Pr[\sup\{Y_{nN_0}, n = 1, 2, \ldots, k\} > y_0 + t] < c_2 \sum_{n=1}^{k} (\frac{1}{t + n\kappa_0})^2. \tag{13}$$

d) If $t > 0$ and $y_0 + t < \log \beta$ then

$$Pr[\sup\{Y_{nN_0}, n = 1, 2, \ldots\} > y_0 + t] < (\frac{c_2}{\kappa_0})\frac{1}{t}.$$

Proof. We first prove a). Obviously $E[Y_1]$ exists, because of Condition B, and satisfies

$$E[Y_1] \leq b + y_0.$$

Hence

$$\begin{aligned} Pr[Y_1 > y_0 + t] &= Pr[Y_1 - E[Y_1] > t - (E[Y_1] - y_0)] \\ &\leq Pr[Y_1 - E[Y_1] > t - b] \leq \frac{E[(Y_1 - E[Y_1])^2]}{(t-b)^2} \\ &\leq \frac{c_1}{(t-b)^2}. \end{aligned}$$

The proof of b) is identical. Just replace the constant b by $-\kappa_0$ and replace the constant c_1 by c_2.

The proof of c) is a little more complicated. For $n = 1, 2, \ldots.$ we define

$$M_n = \sup\{Y_{kN_0}, k = 1, 2, \ldots, n\},$$

we define

$$B_1 = A_1 = \{Y_{N_0} > y_0 + t\}$$

and, for $n = 2, 3, ...$, we define

$$B_n = \{Y_{nN_0} > y_0 + t, M_{n-1} \leq y_0 + t\}$$

and

$$A_n = \{M_n > y_0 + t\}.$$

Clearly

$$Pr[A_n] = \sum_{k=1}^{n} Pr[B_k].$$

Also define

$$M = \sup\{Y_{kN_0}, k \geq 1\}$$

and

$$A = \{M \geq t + y_0\}.$$

From part b) we already know that (13) holds for $k = 1$. In order to prove (13) for $k \geq 2$ we proceed as follows. We first note that

$$Pr[B_k] = Pr[Y_{kN_0} > y_0 + t | M_{k-1} \leq t + y_0] Pr[M_{k-1} \leq t + y_0]. \tag{14}$$

(That $Pr[M_{k-1} \leq t + y_0] > 0$ will follow by induction.) Now

$$\begin{aligned} &Pr[Y_{kN_0} > t + y_0 | M_{k-1} \leq t + y_0] \\ &= Pr[Y_{kN_0} - E[Y_{kN_0} | M_{k-1} \leq t + y_0] \\ &> t + y_0 - E[Y_{kN_0} | M_{k-1} \leq t + y_0] | M_{k-1} \leq t + y_0]. \end{aligned}$$

Since $t + y_0 < \log \beta$ it follows from (12) that

$$E[(Y_{kN_0} - E[Y_{kN_0}] | M_{k-1} \leq t + y_0])^2 | M_{k-1} \leq t + y_0] \leq c_2.$$

Hence by Chebyshev's inequality we find

$$Pr[Y_{kN_0} > t + y_0 | M_{k-1} \leq t + y_0]$$

$$\leq \frac{c_2}{(t + y_0 - E[Y_{kN_0} | M_{k-1} \leq t + y_0])^2}. \tag{15}$$

It is now time to estimate $E[Y_{kN_0} | M_{k-1} \leq t + y_0]$. We have

$$\begin{aligned} &E[Y_{kN_0} | M_{k-1} \leq t + y_0] \\ &= E[Y_{kN_0} - E[Y_{(k-1)N_0} | M_{k-1} \leq t + y_0] | M_{k-1} \leq t + y_0] \\ &\quad + E[Y_{(k-1)N_0} | M_{k-1} \leq t + y_0] \\ &\leq E[Y_{kN_0} - Y_{(k-1)N_0} | M_{k-1} \leq t + y_0] + E[Y_{(k-1)N_0} | M_{k-2} \leq t + y_0] \\ &\leq -\kappa_0 + E[Y_{(k-1)N_0} | M_{k-2} \leq t + y_0]. \end{aligned}$$

(That $E[Y_{(k-1)N_0} | M_{k-1} \leq t + y_0] \leq E[Y_{(k-1)N_0} | M_{k-2} \leq t + y_0]$ is a consequence of the fact that for any real stochastic variable ξ for which $E[\xi]$ exists and any real number α for which $Pr[\xi < \alpha] > 0$ we have $E[\xi | \xi < \alpha] \leq E[\xi]$.)

Since $E[Y_{N_0}] \leq y_0 - \kappa_0$ it follows by induction that

$$E[Y_{kN_0}|M_{k-1} \leq t + y_0] \leq -k\kappa_0 + y_0$$

and if we insert this estimate into (15) we obtain from (15) and (14) that

$$Pr[B_k] \leq Pr[Y_{kN_0} > t + y_0|M_{k-1} \leq t + y_0] \leq \frac{c_2}{(t + k\kappa_0)^2}$$

and hence

$$Pr[A_n] \leq \sum_{k=1}^{n} \frac{c_2}{(t + k\kappa_0)^2}$$

and thereby part c) is proved.

Finally, using the function $g : [0, \infty) \to \mathtt{R}$ defined by

$$g(s) = \frac{1}{(t + s\kappa_0)^2}$$

and the integral

$$\int_0^\infty g(s)\mathrm{d}s = \int_0^\infty \frac{\mathrm{d}s}{(t + s\kappa_0)^2} = \frac{1}{t\kappa_0}$$

as an upper bound of

$$\sum_{k=1}^{\infty} \frac{1}{(t + k\kappa_0)^2},$$

we find that

$$Pr[A] = Pr[M > t + \log(f(x_0)] = \sum_{k=1}^{\infty} Pr[B_k] < \frac{c_2}{\kappa_0}\frac{1}{t}$$

and thereby also part d) of Lemma 4.1 is proved. □

Next, set

$$\eta_1 = \frac{\eta}{4(N_0)^2},$$

define

$$t_1 = \sqrt{\frac{\eta_1}{c_1}} + b + 1 \tag{16}$$

and define ϵ_2 by the equation

$$\log \epsilon_2 + N_0|t_1| + 1 = \log \beta. \tag{17}$$

Lemma 4.2. *Suppose* $x \in K(\epsilon_2)$ *and set* $y_0 = \log f(x)$ *and* $Y_n = \log f(X_n(x))$, $n = 1, 2, \ldots$. *Then, for* $m = 1, 2, 3, ..., N_0 - 1$,

$$Pr[Y_m > y_0 + mt_1] < m\eta_1. \tag{18}$$

Proof. Since $t_1 > b$ and $\epsilon_2 < \beta$, it follows from part a) of Lemma 4.1 and (16) that

$$Pr[Y_1 > y_0 + t_1] < \frac{c_1}{(t_1 - b)^2} = \frac{c_1}{(\sqrt{\frac{\eta_1}{c_1}} + b + 1 - b)^2} < \eta_1$$

and hence (18) holds for $m = 1$.

Next suppose that (18) holds for $m = m_0$. Then for $m = m_0 + 1$ we obtain

$$\begin{aligned} &Pr[Y_{m_0+1} > y_0 + (m_0+1)t_1] \\ &= Pr[\{Y_{m_0+1} > y_0 + (m_0+1)t_1\} \cap \{Y_{m_0} \le y_0 + m_0 t_1\}] \\ &+ Pr[\{Y_{m_0+1} > y_0 + (m_0+1)t_1\} \cap \{Y_{m_0} > y_0 + m_0 t_1\}] \\ &\le Pr[Y_{m_0+1} > y_0 + (m_0+1)t_1 | Y_{m_0} \le y_0 + m_0 t_1] + Pr[Y_{m_0} > y_0 + m_0 t_1] \\ &Pr[Y_{m_0+1} > y_0 + (m_0+1)t_1 | Y_{m_0} \le y_0 + m_0 t_1] + m_0 \eta_1. \end{aligned} \tag{19}$$

Since $y_0 + m_0 t_1 < y_0 + N_0|t_1|$, if $m_0 < N_0$, and $N_0|t_1| + y_0 < \log \beta$ because of (17) and the fact that $x \in K(\epsilon_2)$, and also $t_1 > b$ (see (16)), it follows from part a) of Lemma 4.1 and (16) that

$$\begin{aligned} &Pr[Y_{m_0+1} > y_0 + m_0 t_1 + t_1 | Y_{m_0} \le y_0 + m_0 t_1] \\ &= Pr[Y_{m_0+1} - Y_{m_0} > y_0 - Y_{m_0} + m_0 t_1 + t_1 | Y_{m_0} \le y_0 + m_0 t_1] \\ &\le Pr[Y_{m_0+1} - Y_{m_0} > t_1 | Y_{m_0} \le y_0 + m_0 t_1] < \frac{c_1}{(t_1 - b)^2} = \frac{c_1}{(\sqrt{\frac{\eta_1}{c_1}} + b + 1 - b)^2} < \eta_1 \end{aligned}$$

and hence by (4) follows that

$$Pr[Y_{m_0+1} > y_0 + (m_0+1)t_1] \le \eta_1 + m_0 \eta_1 = (m_0+1)\eta_1.$$

That (18) holds for $m = 1, ..., N_0 - 1$ thus follows by induction. □

We shall next use the estimate in part d) of Lemma 4.1 to determine an even smaller number than ϵ_2. First, we set

$$\eta_2 = \frac{\eta}{4N_0} \tag{20}$$

then we define ϵ_3 by the equation

$$\log \epsilon_3 = \log \epsilon_2 - \frac{c_2}{\kappa_0 \eta_2} - 1 \tag{21}$$

and define t_2 by

$$t_2 = \frac{c_2}{\kappa_0 \eta_2}. \tag{22}$$

Having defined ϵ_3 we can formulate the following proposition from which Theorem 1.1 follows easily if we use Condition C.

PROPOSITION 4.1. *Suppose $x \in K(\epsilon_3)$. Then*

$$Pr[\sup\{\log f(X_n(x)) : n \geq 1\} > -L] < \eta/2. \tag{23}$$

Proof. Let $x \in K(\epsilon_3)$, set $y_0 = \log f(x)$ and for $n = 0, 1, 2, ...,$ set $Y_n = \log f(X_n(x))$. In order to prove Proposition 4.1 we first introduce some further notations. For $m = 0, 1, 2, ..., N_0 - 1$ we define $M^{(m)}$ by

$$M^{(m)} = \sup\{Y_{m+nN_0}, n = 0, 1, ...\}$$

and

$$A^{(m)} = \{M^{(m)} > -L\}.$$

Clearly $\{\sup\{\log f(X_n(x)) : n \geq 1\} > -L\} \subset \cup_{m=0}^{N_0-1} A^m$ and therefore, in order to prove (23), it suffices to show that

$$Pr[A^{(m)}] < \eta/2N_0 \tag{24}$$

for $m = 0, 1, 2, ..., N_0 - 1$.

Let us first consider the case when $m = 0$. Since $x \in K(\epsilon_3)$ it follows easily from (21), (22), (17) and (10) that $y_0 + t_2 < \log \beta < -L$ and hence $Pr[M^{(0)} > -L] \leq Pr[M^{(0)} > y_0 + t_2]$. Since $t_2 > 0$ it follows by part d) of Lemma 4.1 and the definition of η_2 (see(20)) that

$$Pr[M^{(0)} > -L] \leq (\frac{c_2}{\kappa_0})\frac{1}{t_2} = \eta_2 = \frac{\eta}{4N_0} < \frac{\eta}{2N_0};$$

hence $Pr[A^{(0)}] < \frac{\eta}{2N_0}$ and hence (24) holds for $m = 0$.

Now, let m satisfy $1 \leq m \leq N_0 - 1$. We then find that

$$\begin{aligned} Pr[M^{(m)} > -L] &= Pr[\sup\{Y_{m+nN_0} : n = 0, 1, 2, ...\} > -L] \\ &< Pr[\sup\{Y_{m+nN_0} : n = 1, 2, ...\} > -L | Y_m \leq y_0 + mt_1] \\ &\quad + Pr[Y_m > y_0 + mt_1]. \end{aligned} \tag{25}$$

Since $\epsilon_3 < \epsilon_2$ it follows from Lemma 4.2 that

$$Pr[Y_m > y_0 + mt_1] < m\eta_1 = \frac{m\eta}{4N_0^2} < \frac{\eta}{4N_0}. \tag{26}$$

Furthermore, since 1) $\log \epsilon_3 + mt_1 < \log \epsilon_3 + N_0|t_1|$, 2) $t_2 > 0$ and 3) $\log \epsilon_3 + N_0|t_1| + t_2 < \log \epsilon_2 + N_0|t_1| < \log \beta < -L$ we find, by using d) of Lemma 4.1 and the definition of η_2 that

$$\begin{aligned} &Pr[\sup\{Y_{m+nN_0} : n = 1, 2, ...\} > -L | Y_m \leq y_0 + mt_1] \\ &< Pr[\sup\{Y_{m+nN_0} : n = 1, 2, ...\} > Y_m + t_2 | Y_m \leq y_0 + mt_1] \\ &\leq (\frac{c_2}{\kappa_0})\frac{1}{t_2} = \eta_2 = \frac{\eta}{4N_0} \end{aligned}$$

which combined with the inequalities (26) and (25) implies that

$$Pr[A^{(m)}] < \frac{\eta}{2N_0}$$

and hence (24) holds for $m = 0, 1, 2, ..., N_0 - 1$ and thereby Proposition 4.1 is proved. □

Now in order to conclude the proof of Theorem 1.1 we need to show that we for each $x \in K'$ can find an integer N such that (9) holds. Thus let $x_0 \in K'$ be given. Let ϵ_3 be defined as above. From Condition C (see Definition 1.4) it follows that there exists an integer N_0 such that

$$Pr[X_n(x_0) \notin K(\epsilon_3), n = 0, 1, 2, ..., N_0] < \eta/2. \tag{27}$$

Then by using the (strong) Markov property, Proposition 4.1 and inequality (27), it is easily proved that

$$Pr[\sup\{\log f((X_n(x_0))\} \geq -L] < \eta$$

from which follows that $f(X_n(x_0)) \to 0$ $a.s$ as $n \to \infty$. □

5. A CONVERGENCE THEOREM

For sake of completeness let us also prove a modified version of Theorem 2.1.

THEOREM 5.1. *Let $\{(S, \mathcal{F}, \delta_0), (A, \mathcal{A}, \mu), h\}$ be a r.d.s such that $(S, \mathcal{F}, \delta_0)$ is a complete, separable, metric space. Let $\{Z_n, n = 1, 2, ...\}$ denote the associated index sequence, and for $s \in S$ and $n = 0, 1, 2, ...,$ let $\mu_{n,s}$ denote the distribution of $h^n(s, Z^n)$, where thus h^n is defined by (4). Suppose that Condition L, Condition G1, Condition B1, and Condition C′ hold. Suppose also that a) there exists an element s_0 such that $\{\mu_{n,s_0}, n = 0, 1, 2, ...\}$ is a* tight *sequence, b)*

$$Pr[\delta_0(h(s, Z_1), h(t, Z_1)) > 0] = 1 \;\; if \;\; \delta_0(s, t) > 0, \tag{28}$$

c)

$$\int_S \gamma(h, a)\mu(\mathrm{d}a) < \infty \tag{29}$$

where thus $\gamma(h, a)$ is defined by (5).

Then there exists a unique probability measure ν such that for every $s \in S$

$$\mu_{n,s} \to \nu$$

in distribution, as $n \to \infty$.

Proof. Define $P : S \times \mathcal{F} \to [0,1]$ by

$$P(s,F) = \mu(A(s,F))$$

where $A(s,F)$ is defined by

$$A(s,F) = \{a : h(s,a) \in F\}.$$

Let $Lip[S]$ denote the set of real, bounded, Lipschitz-continuous functions on $(S,\mathcal{F})$. From the definition of P it is easily seen that

$$E[u(h^n(s,Z^n))] = \int_S u(t)P^n(s,\mathrm{d}t)$$

for $n = 1,2,...$, if u is a uniformly bounded, continuous, function on $(S,\mathcal{F},\delta_0)$, where thus P^n denotes the nth iteration of P.

In order to prove the theorem it is well-known that it suffices to prove that there exists a unique probability measure ν such that for every $s \in S$ and every $u \in Lip[S]$

$$\lim_{n\to\infty} \int_S u(t)\mu_{n,s}(\mathrm{d}t) = \int_S u(t)\nu(\mathrm{d}t)$$

From hypothesis c) follows that

$$u \in Lip[S] \Rightarrow \int_S u(h(\cdot,a))\mu(da) \in Lip[S]. \tag{30}$$

Using hypothesis a), the implication (30) and a classical argument due to Krylov and Bogolyubov (see *e.g.* [6], Section 32.2), it is not difficult to prove that there exists at least one invariant measure, ν say. (See *e.g.* [5], Proposition 5.16, for details.)

Suppose next that there exist two invariant measures ν and τ. Let $\{V_n, n = 0,1,2,...\}$ and $\{W_n, n = 0,1,2,...\}$ be two independent Markov chains, the first generated by the initial distribution ν and the tr.pr.f P, the second generated by the initial distribution τ and the tr.pr.f P. For a bounded real valued function u write $||u|| = \sup\{|u(s)| : s \in S\}$.

Next, choose $u \in Lip[S]$ such that $0 < \gamma(u) \leq 1$, $||u|| \leq 1$ and

$$0 < \int_S u(s)\nu(\mathrm{d}s) - \int_S u(s)\tau(\mathrm{d}s) = a.$$

Since $\nu \neq \tau$ such a function exists.

Further set

$$\eta = a/8$$

and define L by

$$-L = \log(a/8).$$

Note that $a \leq 2$ since $||u|| \leq 1$. Define $K = S \times S$, $\mathcal{E} = \mathcal{F} \otimes \mathcal{F}$ and define $f : K \to [0,\infty)$ by

$$f((s,t)) = \delta_0(s,t). \tag{31}$$

Let ϵ_2 and ϵ_3 be defined as in the proof of Theorem 1.1.

Further, let $\{Z_n, n = 1, 2, ...\}$ and $\{U_n, n = 1, 2, ...\}$ denote two independent index sequences and finally define M so large that

$$\sup_{s,t \in S} Pr[\delta_0(h^n(s, Z^n), h^n(t, U^n))] \geq \epsilon_3, n = 1, 2, ..., M] < a/8. \tag{32}$$

That we can find such a number M follows easily from Condition C'.

Next, define the stochastic variable T by

$$T = \min\{n : \delta_0(V_n, W_n) < \epsilon_3\}.$$

From hypothesis b), Condition G1 and Condition B it follows that we can apply Proposition 4.1 to the Markov chain

$$\{(s, t), (h^n(s, Z^n), h^n(t, Z^n)), n = 1, 2...\}$$

and the function f defined by (31).

Therefore, if $n > M$ it follows from Proposition 4.1 and the definition of M (see (32)) that

$$E[u(V_n)] - E[u(W_n)] \leq \gamma(u)(a/8) \sum_{k=1}^{M} Pr[T = k]$$

$$+2||u|| \sum_{k=1}^{M} Pr[T = k](a/8) + (a/8)2||u|| \leq 5a/8 < a$$

which gives rise to a contradiction, since both $\{V_n, n = 0, 1, 2, ...\}$ and $\{W_n, n = 0, 1, 2, ...\}$ are stationary sequences. Hence there is only one invariant probability measure ν (say) associated to P, and therefore, if $\{X_n, n = 0, 1, 2, ...\}$ denotes the Markov chain generated by P and the starting point s_0, it follows that

$$\lim_{n \to \infty} E[u(X_n)] = \int_S u(t)\nu(\mathrm{d}t)$$

for all real, bounded, uniformly continuous functions u.

Next, let $t_0 \in S$ be fixed but arbitrary, let $\{X_n, n = 0, 1, 2, ...\}$ and $\{X'_n, n = 0, 1, 2...\}$ be two independent Markov chains, the first generated by P and the initial point s_0, the other by P and the initial point t_0. Let $a > 0$ be chosen arbitrary, define L by $-L = \log(\min\{a/8, 1/2\})$ and define $\eta = \min\{a/8, 1/2\}$. Let ϵ_2 and ϵ_3 be defined as above and let also the integer M be defined as above. (See (32).) Let again $u \in Lip[S]$ satisfy $0 < \gamma(u) \leq 1$ and $||u|| \leq 1$.

By using the same kind of arguments as above it follows easily that

$$|E[u(X_n)] - E[u(X'_n)]| \leq 5a/8 < a,$$

for all $n \geq M$ and since $a > 0$ was chosen arbitrarily, it follows that

$$\lim_{n\to\infty}(E[u(X_n)] - E[u(X'_n)]) = 0.$$

Hence

$$\lim_{n\to\infty}(E[u(X_n)] - E[u(X'_n)]) = 0$$

for all $u \in Lip[S]$ and since

$$\lim_{n\to\infty} E[u(X_n)] = \int_S u(t)\nu(\mathrm{d}t)$$

it follows that also

$$\lim_{n\to\infty} E[u(X'_n)] = \int_S u(t)\nu(\mathrm{d}t)$$

for all $u \in Lip[S]$ and thereby the theorem is proved. □

REFERENCES

[1] B. Bárány, M. Pollicott and K. Simon, *Stationary measures for projective transformations: the Blackwell and Furstenberg measures.* J. Stat. Phys. **148** (2012), 393–421.

[2] H. Bauer, *Wahrscheinlichkeitstheorie und Grundzüge der Masstheorie, 3. Auflage.* De Gruyter, Berlin, 1978.

[3] T. Kaijser, *A limit theorem for Markov chains in compact metrics spaces with applications to products of random matrices.* Duke Math. J. **45** (1978), 311–349.

[4] T. Kaijser, *On a new contraction condition for random systems with complete connections.* Rev. Roumaine Math. Pures Appl. **26** (1981), 1075–1117.

[5] T. Kaijser, *On convergence in distribution of the Markov chain generated by the filter kernel induced by a fully dominated Hidden Markov Model.* Dissertationes Math. **514** (2016).

[6] M. Loève, *Probability Theory.* Third Edition. Van Nostrand, 1963.

[7] D. Malicet, *Random walks on Homeo*(S^1). arxiv:1412.8618, 2014.

[8] J. Newman, *Synchronisation in Invertible Random Dynamical Systems on the Circle.* arXiv:1502.07618, 2015.

[9] T.M. Nguyen and S. Volkov, *A universal result for consecutive random subdivision of polygons.* Random Structures Algorithms. doi.10.1002/rsa.20702, 2016.

[10] S. Volkov *Random geometric subdivisions.* Random Structure Algorithms **43** (2012), 115–130.

Received 1 October 2015

Linköping University,
Department of Mathematics,
SE-581 83 Linköping,
Sweden
thomas.kaijser@liu.se