ANGEWANDTE MATHEMATIK

UND

INFORMATIK

Limit Theorems for Iterated Random Functions by Regenerative Methods

Gerold Alsmeyer und Cheng-Der Fuh

Universität Münster und Academia Sinica
e-mail: gerolda@math.uni-muenster.de, stcheng@ccvax.sinica.edu.tw

Bericht Nr. 11/00-S (korrigierte Version)

UNIVERSITÄT MÜNSTER

Limit Theorems for Iterated Random Functions by Regenerative Methods

GEROLD ALSMEYER

Institut für Mathematische Statistik
Fachbereich Mathematik
Westfälische Wilhelms-Universität Münster
Einsteinstraße 62
D-48149 Münster, Germany

CHENG-DER FUH [1]

Institute of Statistical Science
Academia Sinica
Taipei, Taiwan, ROC

Let $(\mathbb{X}, d)$ be a complete separable metric space and $(F_n)_{n\geq 0}$ a sequence of i.i.d. random functions from $\mathbb{X}$ to $\mathbb{X}$ which are uniform Lipschitz, that is, $L_n = \sup_{x\neq y} d(F_n(x), F_n(y))/d(x,y) < \infty$ a.s. Providing the mean contraction assumption $\mathbb{E}\log^+ L_1 < 0$ and $\mathbb{E}\log^+ d(F_1(x_0), x_0) < \infty$ for some $x_0 \in \mathbb{X}$, it is known (see [4]) that the forward iterations $M_n^x = F_n \circ ... \circ F_1(x)$, $n \geq 0$, converge weakly to a unique stationary distribution π for each $x \in \mathbb{X}$. The associated backward iterations $\hat{M}_n^x = F_1 \circ ... \circ F_n(x)$ are a.s. convergent to a random variable $\hat{M}_\infty$ which does not depend on x and has distribution π. Based on the inequality $d(\hat{M}_{n+m}^x, \hat{M}_n^x) \leq \exp(\sum_{k=1}^n \log L_k) d(F_{n+1} \circ ... \circ F_{n+m}(x), x)$ for all $n, m \geq 0$ and the observation that $(\sum_{k=1}^n \log L_k)_{n\geq 0}$ forms an ordinary random walk with negative drift, we will provide new estimates for $d(\hat{M}_\infty, \hat{M}_n^x)$ and $d(M_n^x, M_n^y)$, $x, y \in \mathbb{X}$, under polynomial as well as exponential moment conditions on $\log(1+L_1)$ and $\log(1+d(F_1(x_0), x_0))$. It will particularly be shown, that the decrease of the Prokhorov distance between $P^n(x,\cdot)$ and π to 0 is of polynomial, respectively exponential rate under these conditions where P^n denotes the n-step transition kernel of the Markov chain of forward iterations. The exponential rate was recently proved in [2] by different methods.

[1] Research partially supported by the National Science Council of the Republic of China.

AMS 1991 subject classifications. 60J05, 60J15, 60K05, 60G17.

Keywords and phrases. random function, Lipschitz map, forward iterations, backward iterations, stationary distribution, Prokhorov metric, level γ ladder epochs.

1. Introduction

The purpose of this article is to show how regenerative methods very effectively apply to establish known as well as new convergence results for iterations of i.i.d. mean contractive random Lipschitz functions. Somewhat surprisingly, such an approach has apparently not yet been used very much in the literature; two related articles by Babillot et al. [1] and Silvestrov and Stenflo [6] also draw on the idea of regeneration but in a different vein. In order to provide further information (and motivation) of the present work, we need a formal description of the underlying model including some necessary notation.

A sequence of the form

$$M_n \;=\; F(\theta_n, M_{n-1}), \quad n \ge 1, \tag{1.1}$$

is called an *iterated function system (IFS) of i.i.d. Lipschitz maps* providing

(1) $M_0, \theta_1, \theta_2, ...$ are independent random elements on a common probability space $(\Omega, \mathfrak{A}, \mathbb{P})$;
(2) $\theta_1, \theta_2, ...$ are identically distributed with common distribution Λ and take values in a measurable space $(\Theta, \mathcal{A})$;
(3) $M_0, M_1, ...$ take values in a complete separable metric space $(\mathbb{X}, d)$ with Borel-σ-field $\mathfrak{B}(\mathbb{X})$;
(4) $F : (\Theta \times \mathbb{X}, \mathcal{A} \otimes \mathfrak{B}(\mathbb{X})) \to (\mathbb{X}, \mathfrak{B}(\mathbb{X}))$ is jointly measurable and Lipschitz continuous in the second argument.

Clearly, $(M_n)_{n\ge 0}$ constitutes a temporally homogeneous Markov chain with state space $\mathbb{X}$ and transition kernel P, given by

$$P(x, B) \;=\; \Lambda(F(\cdot, x) \in B)$$

for $x \in \mathbb{X}$ and $B \in \mathfrak{B}(\mathbb{X})$. The n-step transition kernel is denoted P^n. For $x \in \mathbb{X}$, let $\mathbb{P}_x$ be the probability measure on the underlying measurable space under which $M_0 = x$ a.s. The associated expectation is denoted $\mathbb{E}_x$, as usual. For an arbitrary distribution ν on $\mathbb{X}$, we put $\mathbb{P}_\nu(\cdot) \stackrel{\text{def}}{=} \int \mathbb{P}_x(\cdot)\, \nu(dx)$ with associated expectation $\mathbb{E}_\nu$. We use $\mathbb{P}$ and $\mathbb{E}$ for probabilities and expectations, respectively, that do not depend on the initial distribution.

Let $\mathbb{X}_0$ be a countable dense subset of $\mathbb{X}$ and $\mathfrak{M}(\mathbb{X}_0, \mathbb{X})$ the space of all mappings $f : \mathbb{X}_0 \to \mathbb{X}$ endowed with product topology and product σ-field. Then the space $\mathfrak{C}_{Lip}(\mathbb{X}, \mathbb{X})$ of all Lipschitz continuous mappings $f : \mathbb{X} \to \mathbb{X}$ properly embedded forms a Borel subset of $\mathfrak{M}(\mathbb{X}_0, \mathbb{X})$ and the mappings

$$\mathfrak{C}_{Lip}(\mathbb{X}, \mathbb{X}) \times \mathbb{X} \;\ni\; (f, x) \;\mapsto\; f(x) \;\in\; \mathbb{X},$$
$$\mathfrak{C}_{Lip}(\mathbb{X}, \mathbb{X}) \;\ni\; f \;\mapsto\; l(f) \;\stackrel{\text{def}}{=}\; \sup_{x \ne y} \frac{d(f(x), f(y))}{d(x, y)}$$

are Borel, see Lemma 5.1 in [2] for details. Hence

$$L_n \;\stackrel{\text{def}}{=}\; l(F(\theta_n, \cdot)), \quad n \ge 1, \tag{1.2}$$

are also measurable and form a sequence of i.i.d. random variables.

In the following, we write $F_n(x)$ for $F(\theta_n, x)$. Let $F_{k:n} \stackrel{\text{def}}{=} F_k \circ ... \circ F_n$ and $F_{n:k} \stackrel{\text{def}}{=} F_n \circ ... \circ F_k$ for all $1 \le k \le n$. Put $F_{0:1}(x) = F_{1:0}(x) \stackrel{\text{def}}{=} x$. Hence

$$M_n \;=\; F_n(M_{n-1}) \;=\; F_{n:1}(M_0) \tag{1.3}$$

for all $n \ge 1$. Closely related to these *forward iterations*, and in fact a key tool to their analysis, is the following sequence of *backward iterations*

$$\hat{M}_n \;\stackrel{\text{def}}{=}\; F_{1:n}(M_0), \quad n \ge 1. \tag{1.4}$$

The connection is established by the identity

$$\mathbb{P}_x(M_n \in \cdot) \;=\; \mathbb{P}_x(\hat{M}_n \in \cdot)$$

for all $n \ge 0$. Put also $M_n^x \stackrel{\text{def}}{=} F_{n:1}(x)$ and $\hat{M}_n^x \stackrel{\text{def}}{=} F_{1:n}(x)$ for $x \in \mathbb{X}$ and note that

$$\mathbb{P}((M_n^x, \hat{M}_n^x)_{n\ge 0} \in \cdot) \;=\; \mathbb{P}_x((M_n, \hat{M}_n)_{n\ge 0} \in \cdot).$$

The reason for introducing these additonal sequences is that we will frequently do comparisons of $\hat{M}_n^x$ and $\hat{M}_n^y$, or M_n^x and M_n^y, for different x, y.

IFS have a wide range of applications including perfect simulation, the generation of fractal images, data compression, queuing theory and autoregressive processes; see Diaconis and Freedman [2] for an excellent recent survey including an extensive list of relevant literature.

A central question for an IFS $(M_n)_{n\ge 0}$ is under which conditions it stabilizes, that is, converges to a stationary distribution π. Elton [4] showed in the more general situation of a stationary sequence $(F_n)_{n\ge 1}$ that this holds true whenever $\mathbb{E}\log^+ l(F_1)$ and $\mathbb{E}\log^+ d(F_1(x_0), x_0)$ are both finite for some (and then all) $x_0 \in \mathbb{X}$ and the Liapunov exponent

$$l^* \;\stackrel{\text{def}}{=}\; \lim_{n\to\infty} n^{-1} \log l(F_{n:1})$$

which exists by Kingman's subadditive ergodic theorem, is a.s. negative. His results for i.i.d. $F_1, F_2, ...$ under the slightly stronger assumptions $\mathbb{E}\log^+ l(F_1) < 0$, $\mathbb{E}\log^+ d(F_1(x_0), x_0) < \infty$ for some $x_0 \in \mathbb{X}$ are restated in Theorem 2.1. The basic idea is to consider the backward iterations $\hat{M}_n^x = F_{1:n}(x)$ and to prove their a.s. convergence to a limit $\hat{M}_\infty$ which does not depend on x and which has distribution π. The obvious inequality

$$d(\hat{M}_{n+m}^x, \hat{M}_n^x) \;\le\; \left(\prod_{k=1}^n l(F_k)\right) d(F_{n+1:n+m}(x), x) \quad \text{a.s.}, \tag{1.5}$$

valid for all $n, m \ge 0$ and $x \in \mathbb{X}$, forms a key tool in the necessary analysis. The present article embarks on that same inequality together with the simple observation that

$$\log\left(\prod_{k=1}^n l(F_k)\right) \;=\; \sum_{k=1}^n \log l(F_k), \quad n \ge 0,$$

is an ordinary zero-delayed random walk and thus perfectly amenable to renewal theoretic (regenerative) arguments. Under the mean contraction assumption $\mathbb{E}\log^+ l(F_1) < 0$, it has negative drift whence, for arbitrary $\gamma \in (0,1)$, the level $\log\gamma$ ladder epochs $\sigma_0(\gamma) \stackrel{\text{def}}{=} 0$,

$$\sigma_n(\gamma) \stackrel{\text{def}}{=} \inf\left\{k > \sigma_{n-1}(\gamma) : \sum_{j=\sigma_{n-1}(\gamma)+1}^{k} \log l(F_j) \le \log\gamma\right\}, \quad n \ge 1, \tag{1.6}$$

are all a.s. finite and constituting an ordinary discrete renewal process. As a consequence, the subsequence $(M_{\sigma_n(\gamma)})_{n\ge 0}$ again forms an IFS of i.i.d. Lipschitz maps which further is *strictly contractive* because, by construction,

$$\max\{l(F_{1:\sigma_1(\gamma)}), l(F_{\sigma_1(\gamma):1})\} \le \gamma < 1.$$

For the associated backward iterations $\hat{M}^x_{\sigma_n(\gamma)} = F_{1:\sigma_n(\gamma)}(x)$, inequality (1.5) hence takes the very strong form

$$d(\hat{M}^x_{\sigma_{n+m}(\gamma)}, \hat{M}^x_{\sigma_n(\gamma)}) \le \gamma^n d(F_{\sigma_{n+1}(\gamma)+1:\sigma_{n+m}(\gamma)}(x), x) \tag{1.7}$$

for all $n, m \ge 0$ and $x \in \mathbb{X}$ and suggests the following procedure to prove convergence results for $(M_n)_{n\ge 0}$ and its associated sequence of backward iterations:

STEP 1. Given a set of conditions, find out what kind of results hold true for the strictly contractive sequences $(M_{\sigma_n(\gamma)})_{n\ge 0}$ or $(\hat{M}_{\sigma_n(\gamma)})_{n\ge 0}$ for any $\gamma \in (0,1)$.

STEP 2. Analyze the excursions of $(M_n)_{n\ge 0}$ or $(\hat{M}_n)_{n\ge 0}$ between two successive ladder epochs and adjust the results with respect to $(M_n)_{n\ge 0}$, respectively $(\hat{M}_n)_{n\ge 0}$ if necessary.

Our results in Section 2, some of which have been proved earlier in the literature by different methods, will show that this method is very powerful. They focus on estimates for $d(\hat{M}_\infty, \hat{M}_n)$ under $\mathbb{P}_x$, $x \in \mathbb{X}$, and $d(M^x_n, M^y_n)$ for $x, y \in \mathbb{X}$. The latter distance may be viewed as the coupling rate of the forward iterations at time n when started at different values x and y. The two sets of conditions we will consider are that, for some $p > 0$ and some $x_0 \in \mathbb{X}$, either

$$\mathbb{E}\log^{p+1}(1 + L_1) < \infty \quad \text{and} \quad \mathbb{E}\log^{p+1}(1 + d(F_1(x_0), x_0)) < \infty \tag{1.8}$$

or

$$\mathbb{E}L_1^p < \infty \quad \text{and} \quad \mathbb{E}d(F_1(x_0), x_0)^p < \infty \tag{1.9}$$

holds. Two major conclusions will concern the distance of $P^n(x,\cdot)$ for $x \in \mathbb{X}$ and π in the Prokhorov metric associated with d. Following [2], the latter is also denoted d and defined, for two probability measures λ_1, λ_2 on $\mathbb{X}$, as the infimum over all $\delta \ge 0$ such that

$$\lambda_1(B) < \lambda_2(B^\delta) + \delta \quad \text{and} \quad \lambda_2(B) < \lambda_1(B^\delta) + \delta$$

for all $B \in \mathfrak{B}(\mathbb{X})$, where $B^\delta \stackrel{\text{def}}{=} \{x \in \mathbb{X} : d(x,y) < \delta \text{ for some } y \in B\}$. We will prove that, for all $x \in \mathbb{X}$ and $n \geq 0$,

$$d(P^n(x,\cdot),\pi) \;\leq\; A_x(n+1)^{-p}, \tag{1.10}$$

if (1.8) holds, and

$$d(P^n(x,\cdot),\pi) \;\leq\; A_x r^n \tag{1.11}$$

for some $r \in (0,1)$ not depending on x and n, if (1.9) is true. (1.11) was also proved by different means in [2].

The further organization of the paper is as follows. The main results are presented in the next section. Section 3 collects some necessary lemmata for their proofs which in turn will be provided in Section 4.

2. Main Results

Let $\sigma_1(\gamma)$ be as defined in (1.6) for $\gamma \in (0,1)$, i.e.

$$\sigma_1(\gamma) \;\stackrel{\text{def}}{=}\; \inf\{n \geq 1 : L_{1:n} \leq \gamma\} \;=\; \inf\left\{n \geq 1 : \sum_{k=1}^n \log L_k \leq \log\gamma\right\}. \tag{2.1}$$

Providing $\mathbb{E}\log^+ L_1 < 0$, a condition which will always be in force throughout, $\sigma_1(\gamma)$ is an a.s. finite first passage time with finite mean $\mu(\gamma)$. It has also finite variance $\theta(\gamma)^2$, say, if $\mathbb{E}\log(1+L_1)^2 < \infty$, see [5, Theorem III.3.1]. Let further

$$\log\gamma^* \;\stackrel{\text{def}}{=}\; \inf_{\gamma\in(0,1)} \frac{\log\gamma}{\mu(\gamma)} \;=\; \inf_{\gamma\in(0,1)} \log\gamma^{1/\mu(\gamma)}. \tag{2.2}$$

If $\mathbb{E}|\log L_1| < \infty$, then it is well known from renewal theory, see [5, Section III.9], that

$$\frac{\log\gamma}{\mathbb{E}\log L_1} \;<\; \mu(\gamma) \;<\; \frac{\log\gamma}{\mathbb{E}\log L_1}(1+o(1)) \quad (\gamma \to 0). \tag{2.3}$$

It is then easily checked that in this case

$$\log\gamma^* \;=\; \lim_{\gamma\downarrow 0} \frac{\log\gamma}{\mu(\gamma)} \;=\; \mathbb{E}\log L_1. \tag{2.4}$$

Hence $\mathbb{E}\log(L_1/\gamma^*)$ either equals 0 or $-\infty$ which together with $\gamma^{1/\mu(\gamma)} > \gamma^*$ implies

$$\lim_{n\to\infty} \gamma^{-n/\mu(\gamma)} L_{1:n} \;=\; 0 \tag{2.5}$$

for all $\gamma > 0$ by the strong law of large numbers. The reason for introducing γ^* is that it constitutes a lower bound for the rate of exponential convergence in the results we are going to prove.

Theorem 2.1. *Given an IFS $(M_n)_{n\geq 0}$ of i.i.d. Lipschitz maps, suppose*

$$\mathbb{E}\log^+ L_1 < 0 \quad \textit{and} \quad \mathbb{E}\log^+ d(F_1(x_0),x_0) < \infty \tag{2.6}$$

for some $x_0 \in \mathbb{X}$. Then the following assertions hold:

(a) $\hat{M}_n$ *converges a.s. to a random element* $\hat{M}_\infty$ *with distribution* π *which does not depend on the initial distribution.*

(b) For each $\gamma \in (\gamma^*, 1)$, $\lim_{n\to\infty} \mathbb{P}_x(d(\hat{M}_\infty, \hat{M}_n) > \gamma^n) = 0$ *for all* $x \in \mathbb{X}$.

(c) M_n *converges in distribution to* π *under every* $\mathbb{P}_x$, $x \in \mathbb{X}$.

(d) π *is the unique stationary distribution of* $(M_n)_{n\geq 0}$ *and* $(\hat{M}_n)_{n\geq 0}$ *a stationary sequence under* P_π.

(e) $(M_n)_{n\geq 0}$ *is ergodic under* $\mathbb{P}_\pi$.

All parts of this theorem except for (b) were proved by Elton [4] for general stationary sequences $(F_n)_{n\geq 0}$ with a.s. negative Liapunov exponent. Part (b) will be proved at the beginning of Section 4.

THEOREM 2.2. *Given the situation of Theorem 2.1 and additionally condition (1.8) for some* $p > 0$, *the following assertions hold:*

(a) For each $\gamma \in (\gamma^*, 1)$,

$$\sum_{n\geq 1} n^{p-1}\mathbb{P}_x(d(\hat{M}_\infty, \hat{M}_n) > \gamma^n) \ \leq \ c_\gamma\Big(1 + \log^p(1 + d(x, x_0))\Big)$$

and

$$\lim_{n\to\infty} n^p\mathbb{P}_x(d(\hat{M}_\infty, \hat{M}_n) > \gamma^n) \ = \ 0$$

for all $x \in \mathbb{X}$ *and some* $c_\gamma \in (0, \infty)$.

(b) For each $\gamma \in (\gamma^*, 1)$,

$$\limsup_{n\to\infty} n^{\frac{p-1}{p}}\left(\frac{1}{n}\log d(\hat{M}_\infty, \hat{M}_n) - \log\gamma\right) \ \leq \ 0 \quad \mathbb{P}_x\text{-a.s.}$$

for all $x \in \mathbb{X}$. *In case* $0 < p \leq 1$ *this remains true for* $\gamma = \gamma^*$.

(c) If $p = 1$, *then* $\lim_{n\to\infty} \gamma^{-n} d(\hat{M}_\infty, \hat{M}_n) = 0$ $\mathbb{P}_x$*-a.s. for all* $x \in \mathbb{X}$ *and all* $\gamma \in (\gamma^*, 1)$.

(d) $d(P^n(x, \cdot), \pi) \leq A_x(n + 1)^{-p}$ *for all* $n \geq 0$, $x \in \mathbb{X}$ *and a positive constant* A_x *of the form* $\max\{A, 2d(x, x_0)\}$, *where* A *does neither depend on* x *nor on* n.

(e) $\int_{\mathbb{X}} \log^p(1 + d(x, x_0))\, \pi(dx) = \int_0^\infty pt^{p-1}\pi(x : \log(1 + d(x, x_0)) > t)\, dt < \infty$.

THEOREM 2.3. *Given the situation of Theorem 2.1 and additionally condition (1.9) for some* $p > 0$, *the following assertions hold:*

(a) For each $\gamma \in (\gamma^*, 1)$,

$$\lim_{n\to\infty} \alpha_\gamma^{-n}\mathbb{P}_x(d(\hat{M}_\infty, \hat{M}_n) > \gamma^n) \ = \ 0$$

for all $x \in \mathbb{X}$ *and some* $\alpha_\gamma \in (0, 1)$.

(b) There exists $\eta > 0$ *such that for each* $q \in (0, \eta)$,

$$\lim_{n\to\infty} \sup_{x\in\mathbb{X}} \alpha_q^{-n}(1 + d(x, x_0))^{-q}\mathbb{E}_x d(\hat{M}_\infty, \hat{M}_n)^q \ = \ 0$$

for some $\alpha_q \in (0,1)$. *The same holds true for* $q = \eta$ *with* $\alpha_q = 1$.

(c) $d(P^n(x,\cdot),\pi) \le A_x r^n$ *for all* $n \ge 0$, *some* $r \in (0,1)$ *and a constant* A_x *of the form* $\max\{A, d(x,x_0)\}$. *The constants* r *and* A *do not depend on* x *nor* n.

(d) $\int_{\mathbb{X}} d(x,x_0)^\eta \,\pi(dx) = \int_0^\infty \eta t^{\eta-1}\pi(x : d(x,x_0) > t)\,dt < \infty$ *for some* $\eta > 0$.

Let us mention that the constants $c_\gamma, \alpha_\gamma, \alpha_q, A_x$ and r in the previous theorems generally further depend on $p > 0$ of the supposed respective moment condition.

The assertions of the previous two theorems on $d(\hat{M}_\infty, \hat{M}_n)$ are easily translated into similar results on $d(M_n^x, M_n^y)$ for the forward iterations started at different values x and y. Essentially, this only takes the observation that (M_n^x, M_n^y) and $(\hat{M}_n^x, \hat{M}_n^y)$ are identically distributed for all $x, y \in \mathbb{X}$ and $n \ge 0$ and that

$$d(\hat{M}_n^x, \hat{M}_n^y) \;\le\; d(\hat{M}_\infty^{x_0}, \hat{M}_n^x) \;+\; d(\hat{M}_\infty^{x_0}, \hat{M}_n^y).$$

We summarize the results in the following two corollaries. The proofs are omitted.

Corollary 2.4. *Given the situation of Theorem 2.2, the following assertions hold:*

(a) *For each* $\gamma \in (\gamma^*, 1)$,

$$\sum_{n\ge 1} n^{p-1}\mathbb{P}(d(M_n^x, M_n^y) > \gamma^n) \;\le\; c_\gamma\Big(1 + \log^p(1 + d(x,x_0)) + \log^p(1 + d(y,x_0))\Big)$$

and

$$\lim_{n\to\infty} n^p\mathbb{P}(d(M_n^x, M_n^y) > \gamma^n) \;=\; 0$$

for all $x, y \in \mathbb{X}$ *and some* $c_\gamma \in (0,\infty)$.

(b) *For each* $\gamma \in (\gamma^*, 1)$,

$$\limsup_{n\to\infty} n^{\frac{p-1}{p}}\left(\frac{1}{n}\log d(M_n^x, M_n^y) - \log\gamma\right) \;\le\; 0 \quad a.s.$$

for all $x, y \in \mathbb{X}$. *In case* $0 < p \le 1$ *this remains true for* $\gamma = \gamma^*$.

(c) *If* $p = 1$, *then* $\lim_{n\to\infty}\gamma^{-n} d(M_n^x, M_n^y) = 0$ *a.s. for all* $x, y \in \mathbb{X}$ *and all* $\gamma \in (\gamma^*, 1)$.

Corollary 2.5. *Given the situation of Theorem 2.3, the following assertions hold:*

(a) *For each* $\gamma \in (\gamma^*, 1)$,

$$\lim_{n\to\infty} \alpha_\gamma^{-n}\mathbb{P}(d(M_n^x, M_n^y) > \gamma^n) \;=\; 0$$

for all $x, y \in \mathbb{X}$ *and some* $\alpha_\gamma \in (0,1)$.

(b) *There exists* $\eta > 0$ *such that for each* $q \in (0,\eta)$,

$$\lim_{n\to\infty}\sup_{x,y\in\mathbb{X}} \alpha_q^{-n}(1 + d(x,x_0) \vee d(y,x_0))^{-q}\mathbb{E}_x d(M_n^x, M_n^y)^q \;=\; 0$$

for some $\alpha_q \in (0,1)$. *The same holds true for* $q = \eta$ *with* $\alpha_q = 1$.

Notice that the previous results in combination with Theorem 2.2(e) and Theorem 2.3(d) provide also information on the distance of $M_n^x = F_{n:1}(x)$ for any $x \in \mathbb{X}$ to a stationary counterpart $M_n^\pi \stackrel{\text{def}}{=} F_{n:1}(M_0^\pi)$ where M_0^π has distribution π. We must only observe that

$$\mathbb{P}(d(M_n^x, M_n^\pi) \in \cdot) \;=\; \int_{\mathbb{X}} \mathbb{P}(d(M_n^x, M_n^y) \in \cdot)\;\pi(dy).$$

For instance, part (a) of Corollary 2.4 together with $\int_{\mathbb{X}} \log(1+d(y,x_0))\,\pi(dy) < \infty$ thus further gives for each $\gamma \in (\gamma^*, 1)$, that

$$\sum_{n\ge 1} n^{p-1}\mathbb{P}(d(M_n^x, M_n^\pi) > \gamma^n) \;\le\; c_\gamma\Big(1+\log^p(1+d(x,x_0))\Big)$$

and

$$\lim_{n\to\infty} n^p\mathbb{P}(d(M_n^x, M_n^\pi) > \gamma^n) \;=\; 0$$

for all $x \in \mathbb{X}$ and some $c_\gamma \in (0,\infty)$.

3. Auxiliary Lemmata

In the following the conditions of Theorem 2.1 shall always be assumed. Let us begin with the collection of some necessary notation and facts from renewal theory. Fix an arbitrary $\gamma \in (0,1)$ and consider the successive level γ ladder epochs $\sigma_n = \sigma_n(\gamma)$, $n \ge 0$, defined in (1.6). They constitute an aperiodic renewal sequence with finite mean $\mu = \mu(\gamma)$. Let us further mention, as being used in various places hereafter, that, for each $p > 0$, $\mathbb{E}\log^{p+1}(1+L_1) < \infty$ implies $\mathbb{E}\sigma_1^{p+1} < \infty$ and that $\mathbb{E}L_1^p < \infty$ implies $\mathbb{E}e^{t\sigma_1} < \infty$ for some $t > 0$, see [5, Section III.3].

Put $\tau(n) \stackrel{\text{def}}{=} \inf\{j \ge 0 : \sigma_j \ge n\}$ for $n \ge 0$. By the elementary renewal theorem

$$\frac{\tau(n)}{n} \;\to\; \frac{1}{\mu} \quad \text{a.s.} \tag{3.1}$$

Furnished by two subsequent lemmata, the proofs of our results are essentially based on an analysis of the sequences $(\gamma^{\tau(n)-1}C_{\tau(n)})_{n\ge 0}$ and $(\gamma^{\tau(n)}D_{\tau(n)})_{n\ge 0}$ with C_{n+1}, D_n defined through

$$C_{n+1} \;\stackrel{\text{def}}{=}\; \max\{d(F_{\sigma_n+1:\sigma_{n+1}}(x_0), x_0); d(F_{\sigma_n+1:\sigma_{n+1}}(x_0), F_{\sigma_n+1:k}(x_0)), \sigma_n < k < \sigma_{n+1}\}, \tag{3.2}$$

$$D_n \;\stackrel{\text{def}}{=}\; \sum_{j\ge 0} \gamma^j d(F_{\sigma_{n+j}+1:\sigma_{n+j+1}}(x_0), x_0). \tag{3.3}$$

for $n \ge 0$. The C_n are clearly i.i.d. and a standard renewal argument shows that $C_{\tau(n)}$ converges weakly to a limiting variable C_∞ with distribution function

$$\mathbb{P}(C_\infty \le t) \;=\; \frac{1}{\mathbb{E}\sigma_1}\sum_{n\ge 0}\mathbb{P}(\sigma_1 > n, C_1 \le t), \quad t \ge 0. \tag{3.4}$$

The $F_{\sigma_n+1:\sigma_{n+1}}$, $n \ge 0$, are also i.i.d. whence $(D_n)_{n\ge 0}$ forms a stationary sequence providing the D_n are a.s. finite. It will come out from the proofs of the two lemmata below that this is indeed guaranteed by $\mathbb{E}\log^+ d(F_1(x_0),x_0)<\infty$. $(D_n)_{n\ge 0}$ is further ergodic and autoregressive of order 1 for $D_n = d(F_{\sigma_n+1:\sigma_{n+1}}(x_0),x_0)+\gamma D_{n+1}$. Since $(F_{\sigma_{\tau(n)}+k})_{k\ge 1}\sim (F_k)_{k\ge 1}$ for each $n\ge 0$, where $\sim$ means identical distribution, we see that

$$D_{\tau(n)} \ \sim\ D_0 \tag{3.5}$$

for each $n\ge 0$. Moreover, $D_{\tau(n)}$ is independent of $(L_j,F_j)_{1\le j\le\sigma_{\tau(n)}}$ and of $\tau(n)$. Finally, since obviously $d(F_{\sigma_{n-1}+1:\sigma_n}(x_0),x_0)\le C_n$ for each $n\ge 1$, the inequality

$$D_n \ \le\ \sum_{j\ge 1}\gamma^{j-1}C_{n+j} \tag{3.6}$$

holds for all $n\ge 0$.

Notice that $\sigma_n\equiv n$, $C_n = d(F_n(x_0),x_0)$ and $D_n=\sum_{j\ge 1}\gamma^{j-1}d(F_{n+j}(x_0),x_0)$ in the strongly contractive case $L_1\le\gamma<1$ a.s. The reader should further always keep in mind that *all previous definitions and variables depend on the particularly chosen* $\gamma\in(0,1)$.

LEMMA 3.1. *Given the situation of Theorem 2.1 with* $L_1\le\gamma$ *a.s. for some* $\gamma\in(0,1)$, *let* $\hat M_\infty^{x_0}\overset{\text{def}}{=}\lim_{n\to\infty}\hat M_n^{x_0}$. *Then the* D_n *are a.s. finite and*

$$d(\hat M_\infty^{x_0},\hat M_n^{x})\ \le\ \gamma^n(D_n+d(x,x_0))\quad a.s. \tag{3.7}$$

for all $n\ge 0$ *and* $x\in\mathbb{X}$.

PROOF. If $L_1\le\gamma<1$, then

$$d(F_n(x),F_n(y))\ \le\ \gamma^n\, d(x,y)\quad \text{a.s.}$$

holds for all $x,y\in\mathbb{X}$ and $n\ge 1$. Consequently,

$$\begin{aligned} d(\hat M_{n+m}^{x_0},\hat M_n^{x_0}) &\le \gamma^n d(F_{n+1:n+m}(x_0),x_0)\\ &\le \gamma^n\left(d(F_{n+1}(x_0),x_0)+\sum_{i=2}^m d(F_{n+1:n+i}(x_0),F_{n+1:n+i-1}(x_0))\right)\\ &\le \gamma^n\sum_{i=1}^m\gamma^{i-1}d(F_{n+i}(x_0),x_0)\quad\text{a.s.}\end{aligned}$$

for all $m,n\ge 1$. As $m\to\infty$, the finite sum in the final inequality increases to D_n which is a.s. finite because (with $\beta\overset{\text{def}}{=}\log(1/\gamma)$ and a suitable constant $C>0$)

$$\begin{aligned}\sum_{i\ge 1}\mathbb{P}(\gamma^i d(F_{n+i}(x_0),x_0)>i^{-2}) &= \sum_{i\ge 1}\mathbb{P}(\log d(F_1(x_0),x_0)>i\beta-2\log i)\\ &\le C\,(\mathbb{E}\log^+ d(F_1(x_0),x_0)+1)\ <\ \infty.\end{aligned}$$

We have thus shown

$$d(\hat M_{n+m}^{x_0},\hat M_n^{x_0})\ \le\ \gamma^n D_n\quad\text{a.s.}$$

for all $m, n \geq 1$. The proof of the lemma is complete because $\hat{M}^{x_0}_{n+m} \to \hat{M}^{x_0}_\infty$ a.s. implies $d(\hat{M}^{x_0}_{n+m}, \hat{M}^{x_0}_n) \to d(\hat{M}^{x_0}_\infty, \hat{M}^{x_0}_n)$ a.s., as $m \to \infty$, and because

$$d(\hat{M}^{x_0}_\infty, \hat{M}^{x}_n) \;\leq\; d(\hat{M}^{x_0}_\infty, \hat{M}^{x_0}_n) + d(\hat{M}^{x}_n, \hat{M}^{x_0}_n) \;\leq\; d(\hat{M}^{x_0}_\infty, \hat{M}^{x_0}_n) + \gamma^n d(x, x_0) \quad \text{a.s.}$$

for all $x \in \mathbb{X}$ and $n \geq 0$. ♢

LEMMA 3.2. *Given the situation of Theorem 2.1 and an arbitrary $\gamma \in (0,1)$, the associated C_n in (3.2) have $\mathbb{E}\log^+ C_n < \infty$, the D_n in (3.3) are a.s. finite and*

$$d(\hat{M}^{x_0}_\infty, \hat{M}^{x}_n) \;\leq\; \gamma^{\tau(n)-1} C_{\tau(n)} + \gamma^{\tau(n)} D_{\tau(n)} + L_{1:n} d(x, x_0) \quad \textit{a.s.} \tag{3.8}$$

for all $n \geq 0$ and $x \in \mathbb{X}$, where $C_0 \stackrel{\text{def}}{=} 0$.

PROOF. Defining $F'_n \stackrel{\text{def}}{=} F_{\sigma_{n-1}+1:\sigma_n}$ and $L'_n \stackrel{\text{def}}{=} l(F'_n)$, we see that the (F'_n, L'_n) are i.i.d. and that

$$L'_n \;\leq\; \prod_{i=\sigma_{n-1}+1}^{\sigma_n} L_i \;\leq\; \gamma \quad \text{a.s.}$$

for all $n \geq 1$. So

$$F'_{n:1}(M_0), \quad n \geq 0,$$

is again an IFS of i.i.d. Lipschitz maps (generally $\neq (M_{\sigma_n})_{n\geq 1}$) with backward process

$$\hat{M}_{\sigma_n} \;=\; F'_{1:n}(M_0), \quad n \geq 0.$$

Notice that

$$d(\hat{M}^{x_0}_{\sigma_{\tau(n)}}, \hat{M}^{x_0}_n) \;\leq\; \gamma^{\tau(n)-1} d(\hat{M}^{x_0}_{\sigma_{\tau(n)-1}+1:\sigma_{\tau(n)}}, \hat{M}^{x_0}_{\sigma_{\tau(n)-1}+1:n}) \leq\; \gamma^{\tau(n)-1} C_{\tau(n)} \quad \text{a.s.} \tag{3.9}$$

for all $n \geq 1$ We now infer with the help of the previous lemma and (3.9)

$$\begin{aligned} d(\hat{M}^{x_0}_\infty, \hat{M}^{x}_n) \;&\leq\; d(\hat{M}^{x_0}_{\sigma_{\tau(n)}}, \hat{M}^{x_0}_n) + d(\hat{M}^{x_0}_\infty, \hat{M}^{x_0}_{\sigma_{\tau(n)}}) + d(\hat{M}^{x}_n, \hat{M}^{x_0}_n) \\ &\leq\; \gamma^{\tau(n)-1} C_{\tau(n)} + \gamma^{\tau(n)} D_{\tau(n)} + L_{1:n} d(x, x_0) \quad \text{a.s.} \end{aligned}$$

which is the asserted inequality (3.8).

It remains to show $\mathbb{E}\log^+ C_1 < \infty$ and $\mathbb{E}\log^+ d(F_{1:\sigma_1}(x_0), x_0) < \infty$, the latter to guarantee $D_n < \infty$ a.s. for all $n \geq 0$. Instead of $\log^+$ we will use $\log_* x \stackrel{\text{def}}{=} \log(1+x)$ which is subadditive and satisfies $\log_*(xy) \leq \log_* x + \log_* y$ for all $x, y \geq 0$. Note that $d(F_{1:\sigma_1}(x_0), x_0) \leq C_1$. Put

$$U_n \;\stackrel{\text{def}}{=}\; \max\left\{\prod_{j=1}^{k} L_j, 0 \leq k \leq n\right\}$$

for $n \geq 0$ (where $\prod_{k=1}^0 \stackrel{\text{def}}{=} 1$ as usual) and note that

$$U_n \;\leq\; \prod_{k=1}^n (1+L_k) \quad \text{and} \quad \log_* U_n \;\leq\; \sum_{k=1}^n \log_* L_k. \tag{3.10}$$

Now use

$$\begin{aligned} C_1 \;&\leq\; \sum_{j=1}^{\sigma_1} d(F_{1:j}(x_0), F_{1:j-1}(x_0)) \\ &\leq\; \sum_{j=1}^{\sigma_1} L_1 \cdot \ldots \cdot L_{j-1} d(F_j(x_0), x_0) \qquad (3.11) \\ &\leq\; U_{\sigma_1} \sum_{j=1}^{\sigma_1} d(F_j(x_0), x_0), \end{aligned}$$

(3.10) and Wald's first identity to infer

$$\begin{aligned} \mathbb{E} \log_* d(F_{1:\sigma_1}(x_0), x_0) \;&\leq\; \mathbb{E} \log_* C_1 \\ &\leq\; \mathbb{E}\left(\sum_{j=1}^{\sigma_1} \log_* L_j\right) \;+\; \mathbb{E}\left(\sum_{j=1}^{\sigma_1} \log_* d(F_j(x_0), x_0)\right) \\ &=\; \Big(\mathbb{E} \log_* L_1 \;+\; \mathbb{E} \log_* d(F_1(x_0), x_0)\Big)\mathbb{E}\sigma_1 \;<\; \infty. \end{aligned}$$

$\diamondsuit$

The next two lemmata will provide us with the necessary moment results to prove Theorems 2.2 and 2.3.

LEMMA 3.3. *Let $\gamma \in (0,1)$ and $p > 0$. If*

$$\mathbb{E} \log^{p+1}(1+L_1) < \infty \quad \textit{and} \quad \mathbb{E} \log^{p+1}(1 + d(F_1(x_0), x_0)) < \infty, \tag{3.12}$$

then the following assertions hold:

$$\mathbb{E} \log^{p+1}(1+C_1) \;<\; \infty, \tag{3.13}$$

$$\mathbb{E} \log^{p}(1+D_0) \;=\; \mathbb{E} \log^{p}(1+D_{\tau(n)}) \;<\; \infty \tag{3.14}$$

for all $n \geq 0$. The family $\{\log^p(1 + C_{\tau(n)}), n \geq 0\}$ is uniformly integrable and satisfies

$$\sup_{n \geq 0} \mathbb{E} \log^p(1 + C_{\tau(n)}) \;\leq\; \mathbb{E}\sigma_1 \log^p(1+C_1) \;<\; \infty. \tag{3.15}$$

Finally, the first condition of (3.12) also implies

$$\sum_{n \geq 1} n^{p-1} \mathbb{P}(L_{1:n} > \varepsilon \alpha^n) \;<\; \infty \tag{3.16}$$

as well as

$$\lim_{n\to\infty} n^p \mathbb{P}(L_{1:n} > \varepsilon \alpha^n) \;=\; 0 \tag{3.17}$$

for all $\alpha > \gamma^*$ and $\varepsilon > 0$.

Proof. We retain the notation of the proof of the previous lemma. If (3.12) holds, then $\mathbb{E}\sigma_1^{p+1} < \infty$. Using (3.10) and (3.11), we obtain

$$\mathbb{E}\log_*^{p+1} C_1 \ \le\ \mathbb{E}\left(\sum_{j=1}^{\sigma_1}\log_* L_j\right)^{p+1} + \mathbb{E}\left(\sum_{j=1}^{\sigma_1}\log_* d(F_j(x_0),x_0)\right)^{p+1}.$$

Now each of the terms on the right hand side is the $(p+1)$-st moment of a stopped sum of i.i.d. random variables. That they are finite follows from Theorem I.5.2 in [5].

Recalling (3.5), we must only show $\mathbb{E}\log_*^p D_0 < \infty$ for (3.14). To this end, let $a > 1$ be such that $a\gamma < 1$ and $b \stackrel{\text{def}}{=} \frac{1-a\gamma}{a\gamma} \le \frac{1}{\gamma}$. We then estimate with the help of (3.6)

$$\begin{aligned}\mathbb{P}(D_0 > e^t) \ &\le\ \mathbb{P}(\gamma^{j-1}C_j > b(a\gamma)^j e^t \text{ for some } j \ge 1)\\ &\le\ \sum_{j\ge1}\mathbb{P}(C_1 > a^j e^t)\end{aligned}$$

whence

$$\begin{aligned}\mathbb{E}((\log D_0)^+)^p \ &=\ \int_0^\infty pt^{p-1}\mathbb{P}(D_0 > e^t)\ dt\\ &\le\ \int_0^\infty pt^{p-1}\sum_{j\ge1}\mathbb{P}(C_1 > a^j e^t)\ dt\\ &=\ \sum_{j\ge1}\int_0^\infty pt^{p-1}\mathbb{P}(\log C_1 - j\log a > t)\ dt\\ &=\ \sum_{j\ge1}\mathbb{E}((\log C_1 - j\log a)^+)^p.\end{aligned}$$

A well-known result states that, given a nonnegative random variable X, $EX^{p+1} < \infty$ holds iff $\sum_{j\ge1}E((X-cj)^+)^p < \infty$ for some (and then all) $c > 0$. Since $\mathbb{E}((\log C_1)^+)^{p+1} \le \mathbb{E}\log_*^{p+1} C_1 < \infty$, we thus infer $\mathbb{E}((\log D_0)^+)^p < \infty$ which, of course, also gives $\mathbb{E}\log_*^p D_0 < \infty$.

To prove (3.15), let $H : [0,\infty) \to [0,\infty)$ be an arbitrary function and let U be the renewal measure of $(\sigma_n)_{n\ge0}$ which clearly satisfies $\sup_{n\ge0}U(\{n\}) \le 1$. A standard renewal argument gives the key inequality

$$\begin{aligned}\mathbb{P}(H(C_{\tau(n)}) > t) \ &=\ \sum_{j=1}^{n}U(\{n-j\})\mathbb{P}(\sigma_1 \ge j, H(C_1) > t)\\ &\le\ \sum_{j\ge0}\mathbb{P}(\sigma_1 > j, H(C_1) > t)\\ &=\ \sum_{j\ge1}j\,\mathbb{P}(\sigma_1 = j, H(C_1) > t)\\ &=\ \mathbb{E}\sigma_1\mathbf{1}_{\{H(C_1)>t\}}\end{aligned} \tag{3.18}$$

for all $t \ge 0$ and $n \ge 0$ whence

$$\sup_{n\ge0}\mathbb{E}H(C_{\tau(n)}) \ \le\ \mathbb{E}\sigma_1 H(C_1). \tag{3.19}$$

Furthermore, the family $\{H(C_{\tau(n)}), n \ge 0\}$ is uniformly integrable whenever $\mathbb{E}\sigma_1 H(C_1) < \infty$.

Choosing $H(t) = \log_*^p t$, we conclude with the help of Hölder's inequality

$$\sup_{n\ge 0} \mathbb{E} \log_*^p C_{\tau(n)} \;\le\; \mathbb{E}\sigma_1 \log_*^p C_1 \;\le\; (\mathbb{E}\sigma_1^{p+1})^{1/(p+1)} (\mathbb{E}\log_*^{p+1} C_1)^{p/(p+1)} \;<\; \infty.$$

In order to infer (3.16) and (3.17) from the first condition of (3.12), we note first that, possibly after a left truncation of the L_n, it is no loss of generality to assume $|\mathbb{E}\log L_1| < \infty$ and thus $\log\gamma^* = \mathbb{E}\log L_1$, see (2.4). In this case the i.i.d. $L_n^* \stackrel{\text{def}}{=} L_n/\gamma^*$, $n \ge 1$ satisfy $\mathbb{E}\log L_n^* = 0$. Since

$$\mathbb{P}(L_{1:n} > \alpha^n) \;=\; \mathbb{P}\left(\sum_{j=1}^n \log L_j^* > n\log(\alpha/\gamma^*)\right)$$

for all $n \ge 1$, the assertions are easily obtained from well-known one-sided tail estimates for centered random walks obtained by Chow and Lai [3]. Further details can be omitted. $\diamondsuit$

Lemma 3.4. *Let $\gamma \in (0,1)$ and $p > 0$. If*

$$\mathbb{E}L_1^p < \infty \quad \textit{and} \quad \mathbb{E}d(F_1(x_0), x_0)^p < \infty, \tag{3.20}$$

then the following assertions hold for some $\eta > 0$:

$$\mathbb{E}C_1^{2\eta} \;<\; \infty, \tag{3.21}$$

$$\mathbb{E}D_0^{2\eta} \;=\; \mathbb{E}D_{\tau(n)}^{2\eta} \;<\; \infty \tag{3.22}$$

for all $n \ge 0$. The family $\{C_{\tau(n)}^\eta, n \ge 0\}$ is uniformly integrable and satisfies

$$\sup_{n\ge 0} \mathbb{E}C_{\tau(n)}^\eta \;\le\; \mathbb{E}\sigma_1 C_1^\eta \;<\; \infty. \tag{3.23}$$

Moreover, the first condition of (3.20) implies

$$\mathbb{P}(L_{1:n} > \varepsilon\alpha^n) \;\le\; \varepsilon^{-1}\alpha^n \tag{3.24}$$

for all $n \ge 1$, $\varepsilon > 0$ and a suitable $\alpha \in (0,1)$.

It is to be noted that, by (3.5), the $D_{\tau(n)}$ are identically distributed whence (3.14) and (3.22) trivially imply the uniform integrability of $\{\log^p(1 + D_{\tau(n)}), n \ge 0\}$ and $\{D_{\tau(n)}^{2\eta}, n \ge 0\}$, respectively.

Proof. If (3.20) holds, which in particular means that $\log_* L_1$ has an exponential moment, then σ_1 has an exponential moment, too. Hence a standard argument shows that we can find a $\eta \le p/4$ suffciently small such that

$$\mathbb{E}\exp\left(4\eta\sum_{k=1}^{\sigma_1} \log_* L_k\right) \;<\; \infty. \tag{3.25}$$

It follows from (3.10) and (3.11) that

$$C_1^{2\eta} \;\le\; \exp\left(2\eta\sum_{j=1}^{\sigma_1}\log_* L_j\right)\left(\sum_{k=1}^{\sigma_1} d(F_k(x_0),x_0)\right)^{2\eta} \quad \text{a.s.}$$

and thus with Hölder's inequality

$$\mathbb{E}C_1^{2\eta} \;\le\; \left(\mathbb{E}\exp\left(4\eta\sum_{j=1}^{\sigma_1}\log_* L_j\right)\right)^{1/2}\left(\mathbb{E}\left(\sum_{k=1}^{\sigma_1} d(F_k(x_0),x_0)\right)^{4\eta}\right)^{1/2}.$$

The first expectation on the right hand side is finite by (3.25), while this holds for the second as being the expectation of a stopped sum of i.i.d. random variables with finite moments of order $4\eta \le p$ (see [5]).

(3.22) in case $2\eta \ge 1$ follows immediately by using (3.6) and the infinite version of Minkowski's inequality. They give

$$(\mathbb{E}D_0^{2\eta})^{1/2\eta} \;\le\; \sum_{j\ge 1}\gamma^{j-1}(\mathbb{E}C_1^{2\eta})^{1/2\eta} \;=\; \frac{(EC_1^{2\eta})^{1/2\eta}}{1-\gamma} \;<\; \infty.$$

If $0 < 2\eta < 1$, then $t \mapsto t^{2\eta}$ is subadditive and thus

$$\mathbb{E}D_0^{2\eta} \;\le\; \sum_{j\ge 1}\gamma^{2\eta(j-1)}\mathbb{E}C_1^{2\eta} \;\le\; \frac{\mathbb{E}C_1^{2\eta}}{1-\gamma^{2\eta}} \;<\; \infty.$$

As to the proof of (3.23), note first that $\mathbb{E}C_1^{2\eta} < \infty$ and $\mathbb{E}e^{s\sigma_1} < \infty$ for some $s > 0$ imply $\mathbb{E}C_1^{\eta}\sigma_1 < \infty$. The assertion now follows because (3.19) with $H(t) = t^{\eta}$ yields

$$\sup_{n\ge 0}\mathbb{E}C_{\tau(n)}^{\eta} \;\le\; \mathbb{E}\sigma_1 C_1^{\eta}.$$

In order to show (3.24) we note first that $\mathbb{E}L_1^r$ is a convex function of r on $[0,p)$ with negative derivative $\mathbb{E}\log L_1$ at 0. Hence there exists a $q \in (0, p\wedge 1)$ with $m_q \overset{\text{def}}{=} EL_1^q < 1$. Since $\mathbb{E}L_{1:n}^q = m_q^n$ for all $n \ge 1$, we infer

$$\mathbb{P}(L_{1:n} > \varepsilon\beta^n) \;\le\; \varepsilon^{-1}\left(\frac{m_q}{\beta^q}\right)^n$$

for all $\varepsilon, \beta > 0$ and $n \ge 1$. By choosing any $\beta \in (m_q^{1/q}, 1)$ and then $\alpha = \max\{\beta, m_q/\beta^q\}$, we arrive at the desired conclusion. This completes the proof of the lemma. ♢

Lemma 3.5. *Let $\gamma \in (0,1)$ and $p > 0$. Then $\mathbb{E}\log^{p+1}(1+L_1) < \infty$ implies*

$$\sum_{n\ge 0} n^{p-1}\mathbb{P}(\tau(n) \le an) \;<\; \infty \tag{3.26}$$

as well as

$$\lim_{n\to\infty} n^p\, \mathbb{P}(\tau(n) \le an) \;=\; 0 \qquad (3.27)$$

for all $0 < a < 1/\mu$, *while* $\mathbb{E}L_1^p < \infty$ *implies*

$$\lim_{n\to\infty} \alpha^{-n}\, \mathbb{P}(\tau(n) \le an) \;=\; 0 \qquad (3.28)$$

for all $0 < a < 1/\mu$ *and some* $\alpha \in (0,1)$ *depending on* a.

PROOF. Recall that $\mathbb{E}\log^{p+1}(1+L_1) < \infty$ ensures $\mathbb{E}\sigma_1^{p+1} < \infty$ which in turn yields

$$\mathbb{E}\left(\sup_{k\ge 0}(\sigma_k - k(\mu+\varepsilon))\right)^p \;<\; \infty$$

for all $\varepsilon > 0$ by a result of Chow and Lai [3]. Hence, with $\varepsilon > 0$ satisfying $a(\mu+\varepsilon) = 1-\varepsilon$, we infer

$$\begin{aligned} n^p\mathbb{P}(\tau(n) \le an) &= n^p\mathbb{P}(\sigma_{[an]} - a(\mu+\varepsilon)n > \varepsilon n) \\ &\le n^p\mathbb{P}(\sup_{k\ge an}(\sigma_k - k(\mu+\varepsilon)) > \varepsilon n) \\ &\le \varepsilon^{-p}\mathbb{E}\left(\sup_{k\ge an}(\sigma_k - k(\mu+\varepsilon))\right)^p \;\to\; 0, \end{aligned}$$

as $n\to\infty$ which proves (3.27).

For (3.28), recall that $\mathbb{E}L_1^p < \infty$ ensures $\phi(\beta) \stackrel{\text{def}}{=} \mathbb{E}e^{\beta\sigma_1} < \infty$ for all $\beta \le \beta^*$, $\beta^* > 0$. If $0 < a < 1/\mu$, we find $\beta_a \in (0,\beta^*]$ such that $\phi_a(\beta_a) \stackrel{\text{def}}{=} \mathbb{E}e^{\beta_a(\sigma_1 - 1/a)} = e^{-\beta_a/a}\phi(\beta_a) < 1$ because $\phi_a'(0) = \mathbb{E}(\sigma_1 - 1/a) = \mu - 1/a < 0$. A simple estimation gives

$$\mathbb{P}(\tau(n) \le an) \;\le\; \mathbb{P}(e^{\beta_a\sigma_{[an]}} > e^{\beta_a n}) \;\le\; e^{-\beta_a n}\phi(\beta_a)^{an} \;=\; \phi_a(\beta_a)^{an},$$

whence (3.28) holds for each $\alpha \in (1, \phi_a(\beta_a)^{-a})$. ◊

The final lemma of this section is Lemma 5.8 in [2] and provides a useful tool to estimate the Prokhorov distance of two probability measures.

LEMMA 3.6. *Let* X_1, X_2 *be two* $\mathbb{X}$*-valued random elements with distributions* λ_1, λ_2. *Then* $\mathbb{P}(d(X_1,X_2) \ge \delta) < \delta$ *implies* $d(\lambda_1,\lambda_2) \le \delta$. *(Here* d *denotes the metric on* $\mathbb{X}$ *in the first instance and the associated Prokhorov metric in the second.)*

PROOF. The assertion follows from the obvious inequality

$$\max\{\mathbb{P}(X_1 \in B, X_2 \notin B^\delta), \mathbb{P}(X_1 \notin B^\delta, X_2 \in B)\} \;\le\; \mathbb{P}(d(X_1,X_2) \ge \delta)$$

for all $B \in \mathfrak{B}(\mathbb{X})$. ◊

4. Proofs of Theorems 2.1–2.3

Proof of Theorem 2.1(b). Pick an arbitrary $\hat{\gamma} \in (\gamma^*, 1)$. By definition of γ^*, there is a $\gamma \in (0,1)$ such that $\log\gamma^* < \frac{\log\gamma}{\mu} < \log\hat{\gamma}$ whence $\hat{\gamma} = \gamma^b$ for some $0 < b < 1/\mu$. Recalling (3.1), we see that $\tau(n) - bn \to \infty$ a.s. and therefore

$$\gamma^{\tau(n)-1-bn} C_{\tau(n)} \ \stackrel{P}{\to}\ 0 \quad \text{and} \quad \gamma^{\tau(n)-1-bn} D_{\tau(n)} \ \stackrel{P}{\to}\ 0 \tag{4.1}$$

as $n \to \infty$, where $\stackrel{P}{\to}$ means convergence in probability. Moreover, $\mathbb{E}\log L_1 < -\log\gamma < -\log\hat{\gamma}$ (see (2.5) and the preceding discussion) implies $\sum_{i=1}^n (\log L_i + \log\hat{\gamma}) \to -\infty$ a.s. and thus

$$\hat{\gamma}^{-n} L_{1:n} \ \to\ 0 \quad \text{a.s.} \tag{4.2}$$

It follows from Lemma 3.2 that

$$\begin{aligned}\mathbb{P}_x(d(\hat{M}_\infty, \hat{M}_n) > \hat{\gamma}^n) \ &=\ \mathbb{P}(d(\hat{M}_\infty^{x_0}, \hat{M}_n^x) > \hat{\gamma}^n)\\ &\le\ \mathbb{P}(\gamma^{\tau(n)-1-bn}(C_{\tau(n)} + D_{\tau(n)}) + \hat{\gamma}^{-n} L_{1:n} d(x, x_0) > 1)\end{aligned}$$

for all $x \in \mathbb{X}$ and $n \ge 1$. Since the last expression converges to 0 by (4.1) and (4.2) we have shown the desired result. ♢

Proof of Theorem 2.2. (a) As in the proof of Theorem 2.1(b), pick any $\hat{\gamma} \in (\gamma^*, 1)$ and $\gamma \in (0,1)$ such that $\hat{\gamma} = \gamma^b$ for some $b \in (0, 1/\mu)$. Choose further any $a \in (b, 1/\mu)$ with $\gamma^a \in (\gamma^*, 1)$. Note that $\gamma^{an} d(x, x_0) \ge \hat{\gamma}^n/3$ iff $n \le n_0 \stackrel{\text{def}}{=} \frac{\log 3d(x,x_0)}{(a-b)\log(1/\gamma)}$. A simple estimation using (3.8) of Lemma 3.2 (with γ) leads to

$$\begin{aligned}\mathbb{P}(d(\hat{M}_\infty^{x_0}, \hat{M}_n^x) > \hat{\gamma}^n) \ &\le\ \mathbb{P}(\tau(n) \le an+1) \ +\ \mathbb{P}(\gamma^{an} C_{\tau(n)} > \gamma^{bn}/3)\\ &+\ \mathbb{P}(\gamma^{an} D_{\tau(n)}) > \gamma^{bn}/3) \ +\ \mathbb{P}(L_{1:n} > \gamma^{an}) \ +\ \mathbf{1}(\gamma^{an} d(x, x_0) > \gamma^{bn}/3)\end{aligned} \tag{4.3}$$

for all $x \in \mathbb{X}$ and $n \ge 0$. In the following $K \in (0, \infty)$ shall denote a generic constant which may differ from line to line but is always independent of x. All but the last term on the right hand side of (4.3) are independent of x. The first one multiplied with n^{p-1} is summable by Lemma 3.5(a). As to the second term, which is the critical one, we put $\varepsilon \stackrel{\text{def}}{=} (a-b)\log(1/\gamma)$ and estimate with the help of (3.18), Hölder's inequality and (3.13)

$$\begin{aligned}\sum_{n\ge1} n^{p-1}\mathbb{P}(\gamma^{an} C_{\tau(n)} > \gamma^{bn}/3) \ &=\ \sum_{n\ge1} n^{p-1}\mathbb{P}(\log_* 3C_{\tau(n)} > \varepsilon n)\\ &\le\ \sum_{n\ge1} n^{p-1}\mathbb{E}\sigma_1 \mathbf{1}_{\{\log_* 3C_1 > \varepsilon n\}}\\ &\le\ K\,\mathbb{E}\sigma_1 \log_*^p 3C_1\\ &\le\ K\,(\mathbb{E}\sigma_1^{p+1})^{1/(p+1)}(\mathbb{E}\log_*^{p+1} 3C_1)^{p/(p+1)} \ <\ \infty.\end{aligned}$$

Summability of the third and fourth term in (4.3) multiplied with n^{p-1} is guaranteed by

$$\sum_{n\ge1} n^{p-1}\mathbb{P}(\gamma^{an} D_{\tau(n)} > \gamma^{bn}/3) \ \le\ \sum_{n\ge1} n^{p-1}\mathbb{P}(\log_* 3D_0 > \varepsilon n) \ \le\ K\,\mathbb{E}\log_*^p D_0 \ <\ \infty$$

and by (3.16), respectively. Finally, we get for the last term in (4.3)

$$\begin{aligned}\sum_{n\ge 1} n^{p-1}\mathbf{1}(\gamma^{an}d(x,x_0)>\gamma^{bn}/3) &= \sum_{n=1}^{n_0} n^{p-1}\mathbf{1}(\gamma^{an}d(x,x_0)>\gamma^{bn}/3)\\ &\le K\,n_0^p \le c_\gamma \log_*^p d(x,x_0)\end{aligned}$$

for a suitable $c_\gamma\in(0,\infty)$.

The second assertion $n^p\mathbb{P}(d(\hat M_\infty^{x_0},\hat M_n^x)>\hat\gamma^n)\to 0$ is more easily inferred from (4.3) in combination with (3.17) and the uniform integrability of $\{\log_*^p(C_{\tau(n)}+D_{\tau(n)}),n\ge 0\}$ which holds by Lemma 3.3.

(b) Note first that $\frac{\tau(n)}{n}\to\frac{1}{\mu}$ a.s. together with (2.5) ensures for each $\gamma\in(0,1)$

$$\gamma^{-\tau(n)}L_{1:n} = \left(\gamma^{-\tau(n)/n}\right)^n L_{1:n} \to 0 \quad \text{a.s.}$$

In combination with the uniform integrability of $\{\log_*^p(\gamma^{-1}C_{\tau(n)}+D_{\tau(n)}),n\ge 0\}$, this implies

$$\lim_{n\to\infty} n^{-1/p}\log_*(\gamma^{-1}C_{\tau(n)}+D_{\tau(n)}+\gamma^{-\tau(n)}L_{1:n}d(x,x_0)) = 0 \quad \text{a.s.} \tag{4.4}$$

for each $p>0$. Another use of (3.8) in Lemma 3.2 gives for all $\gamma\in(0,1)$

$$\begin{aligned}\frac{1}{n}\log d(\hat M_\infty^{x_0},\hat M_n^{x_0}) &\le \left(\frac{\tau(n)}{n}\right)\log\gamma + \frac{1}{n}\log\left(\gamma^{-1}(C_{\tau(n)}+D_{\tau(n)})+\gamma^{-\tau(n)}L_{1:n}d(x,x_0)\right)\\ &\le \left(\frac{\tau(n)}{n}\right)\log\gamma + \frac{1}{n}\log_*\left(\gamma^{-1}(C_{\tau(n)}+D_{\tau(n)})+\gamma^{-\tau(n)}L_{1:n}d(x,x_0)\right) \quad \text{a.s.}\end{aligned}$$

where $\tau(n)\le n$ has been used for the second inequality. It follows for each $\hat\gamma\in(0,1)$

$$\begin{aligned}&n^{\frac{p-1}{p}}\left(\frac{1}{n}\log d(\hat M_\infty^{x_0},\hat M_n^{x_0})-\log\hat\gamma\right)\\ &\qquad = n^{\frac{p-1}{p}}\log\gamma\left(\left(\frac{\tau(n)}{n}-\frac{1}{\mu}\right)-c\right)\\ &\qquad + n^{-1/p}\log_*(\gamma^{-1}(C_{\tau(n)}+D_{\tau(n)})+\gamma^{-\tau(n)}L_{1:n}d(x,x_0)) \quad \text{a.s.,}\end{aligned} \tag{4.5}$$

where $c\overset{\text{def}}{=}\frac{\log\hat\gamma}{\log\gamma}-\frac{1}{\mu}$. As $n\to\infty$, the final expression in (4.5) converges a.s. to 0 by (4.4). Recall from (2.2) that $\log\gamma^*=\inf_{\gamma\in(0,1)}\log\gamma^{1/\mu}$. Now we infer upon choosing any $\hat\gamma\in(\gamma^{1/\mu},1)$, which implies $c<0$, and using $\frac{\tau(n)}{n}\to\frac{1}{\mu}$ a.s.

$$\limsup_{n\to\infty} n^{\frac{p-1}{p}}\log\gamma\left(\left(\frac{\tau(n)}{n}-\frac{1}{\mu}\right)-c\right) \le 0 \quad \mathbb{P}_x\text{-a.s.}$$

for all $x\in\mathbb{X}$. Consequently,

$$\limsup_{n\to\infty} n^{\frac{p-1}{p}}\left(\frac{1}{n}\log d(\hat M_\infty^{x_0},\hat M_n^{x_0})-\log\hat\gamma\right) \le 0 \quad \mathbb{P}_x\text{-a.s.} \tag{4.6}$$

for all $x \in \mathbb{X}$, all $\gamma \in (0,1)$ and all $\hat\gamma \in (\gamma^{1/\mu},1)$ which is the desired result. Since $n^{\frac{p-1}{p}} \to 0$ if $p \in (0,1)$, $n^{\frac{p-1}{p}} \equiv 1$ if $p=1$, we may obviously replace $\hat\gamma$ with γ^* in (4.6).

(c) This is an immediate consequence of (b) for $p=1$ which may also be stated as

$$d(\hat M_\infty, \hat M_n) \ \le\ (\gamma^* R_n)^n \quad \text{a.s.}$$

for all $n \ge 0$ where $(R_n)_{n\ge 0}$ is a suitable sequence of random variables satisfying $R_n \to 1$ $\mathbb{P}_x$-a.s. for all $x \in \mathbb{X}$. Hence we have for all $\gamma \in (\gamma^*,1)$ and $x \in \mathbb{X}$

$$\gamma^{-n} d(\hat M_\infty, \hat M_n) \ \le\ \left(\frac{\gamma^* R_n}{\gamma}\right)^n \ \to\ 0 \quad \mathbb{P}_x\text{-a.s.}$$

If $p>1$, the result may also be inferred from (a) which gives

$$\sum_{n\ge 1} \mathbb{P}_x(d(\hat M_\infty, \hat M_n) > \varepsilon\gamma^n) \ \le\ K(\varepsilon,x) \sum_{n\ge 1} n^{-p} \ <\ \infty$$

for all $\varepsilon > 0$, $\gamma \in (\gamma^*,1)$ and a suitable finite constant $K(\varepsilon,x)$. Hence $\mathbb{P}_x(d(\hat M_\infty, \hat M_n) > \varepsilon\gamma^n \text{ i.o.}) = 0$ by the Borel-Cantelli lemma.

(d) Pick first any $\gamma \in (0,1)$ with $\log\gamma < -p$, then any $a \in (0,1/\mu)$ and finally $A_x > 0$ so large that $\mathbb{P}(\tau(n) \le an+1) \le A_x(n+1)^{-p}/2$ for all $n \ge 0$ (Lemma 3.5) and

$$\begin{aligned}
&\mathbb{P}(\gamma^{an}(C_{\tau(n)} + D_{\tau(n)}) + L_{1:n} d(x,x_0) > A_x(n+1)^{-p}) \\
&\qquad \le\ \mathbb{P}(\log_*(C_{\tau(n)} + D_{\tau(n)}) > an\log(1/\gamma) + \log(A_x/2) - p\log(n+1)) \\
&\qquad +\ \mathbb{P}(L_{1:n} d(x,x_0) > A_x(n+1)^{-p}/2) \\
&\qquad \le\ A(n+1)^{-p}/2
\end{aligned}$$

for all $n \ge 0$ (Lemma 3.3). Now a similar estimation as in part (a) gives

$$\begin{aligned}
&\mathbb{P}(d(\hat M_\infty^{x_0}, \hat M_n^x) > A_x(n+1)^{-p}) \\
&\quad \le\ \mathbb{P}(\tau(n) \le an+1) \ +\ \mathbb{P}(\gamma^{an}(C_{\tau(n)} + D_{\tau(n)}) + L_{1:n} d(x,x_0) > A_x(n+1)^{-p}) \\
&\quad \le\ A_x(n+1)^{-p}
\end{aligned}$$

for all $n \ge 0$, which implies the assertion by an appeal to Lemma 3.6.

(e) The assertion follows immediately from $\tau(0) = C_0 = 0$,

$$\begin{aligned}
&\int_0^\infty t^{p-1}\pi(x : \log d(x,x_0) > t)\ dt \\
&\qquad = \int_0^\infty t^{p-1}\mathbb{P}(\log d(\hat M_\infty^{x_0}, x_0) > t)\ dt \\
&\qquad \le \int_0^\infty t^{p-1}\mathbb{P}(\log_* D_0 > t)\ dt
\end{aligned}$$

for all $\gamma \in (0,1)$ (use (3.8)) and the fact that (1.8) ensures $\mathbb{E}\log_*^p D_0 < \infty$ by Lemma 3.3. $\diamondsuit$

Proof of Theorem 2.3. (a) This is proved along the same path as part (a) of Theorem 2.2. The obvious necessary modifications are therefore omitted.

(b) Choose $\eta > 0$ such that Lemma 3.4 holds true and $m_q = \mathbb{E}L_1^q < 1$ for $q \in (0, \eta]$ (see at the end of the proof of Lemma 3.4). By (3.8) in Lemma 3.2, we get for all $\gamma \in (0,1)$ and $q \le \eta$

$$(1 + d(x, x_0))^{-q} d(\hat{M}_\infty^{x_0}, \hat{M}_n^x)^q \;\le\; \gamma^{q(\tau(n)-1)}(C_{\tau(n)} + D_{\tau(n)})^q + L_{1:n}^q \quad \text{a.s.}$$

The right hand side does not depend on $x \in \mathbb{X}$, converges a.s. to 0 and is uniformly integrable by Lemma 3.4. This implies the asserted result for $q = \eta$. For $q \in (0, \eta)$ use Hölder's inequality to obtain

$$\mathbb{E}\gamma^{q(\tau(n)-1)}(C_{\tau(n)} + D_{\tau(n)})^q \;\le\; \left(\mathbb{E}\gamma^{\eta q(\tau(n)-1)/(\eta-q)}\right)^{(\eta-q)/\eta} \left(\mathbb{E}(C_{\tau(n)} + D_{\tau(n)})^\eta\right)^{q/\eta}.$$

Note that $\mathbb{E}L_{1:n}^q = m_q^n$ and use Lemma 3.5 to see that $\lim_{n\to\infty} \alpha_q^{-n}(\mathbb{E}\gamma^{\eta q \tau(n)/(\eta-q)})^{(\eta-q)/\eta} = 0$ for some $\alpha_q \in (m_q, 1)$. The assertion now easily follows.

(c) Fix an arbitrary $\gamma \in (0, 1/2)$ and then $a \in (0, 1/\mu)$. Invoking Lemma 3.4, the moment assumption (1.9) guarantees $K_\eta \stackrel{\text{def}}{=} \sup_{n\ge0} \mathbb{E}(C_{\tau(n)} + D_{\tau(n)})^\eta < \infty$ for some $\eta > 0$. By Lemmata 3.4 and 3.5, $\mathbb{P}(L_{1:n} d(x, x_0) > \varepsilon\alpha^n/2) \le 2\varepsilon^{-1} d(x, x_0)\alpha^n$ and $\mathbb{P}(\tau(n) \le an + 1) \le c\alpha^n$ for all $n \ge 1$, $\varepsilon > 0$ and suitable $c > 0$ and $\alpha \in (0,1)$ (both depending on γ). Now

$$\begin{aligned}
&\mathbb{P}(\gamma^{\tau(n)-1}(C_{\tau(n)} + D_{\tau(n)}) > A(2\gamma)^{an}/2) \\
&\qquad\le\; \mathbb{P}(\tau(n) \le an + 1) \;+\; \mathbb{P}(\gamma^{an}(C_{\tau(n)} + D_{\tau(n)}) > A(2\gamma)^{an}/2) \\
&\qquad\le\; c\alpha^n \;+\; \frac{K_\eta}{2^{\eta a n} A} \\
&\qquad\le\; Ar^n/2
\end{aligned}$$

for all $n \ge 0$, all $A > 0$ sufficiently large and with $r \stackrel{\text{def}}{=} \max\{\alpha, 2^{-\eta a}, (2\gamma)^a\} \in (0,1)$. We conclude upon setting $A_x = \max\{A, 2d(x, x_0) + 1\}$ and by another use of (3.8) in Lemma 3.2

$$\mathbb{P}(d(\hat{M}_\infty^{x_0}, \hat{M}_n^x) > A_x r^n) \;\le\; A_x r^n$$

for all $n \ge 0$ and $x \in \mathbb{X}$. The assertion then follows by Lemma 3.6.

(d) Similar to part (e) of the previous theorem and with η as before, the assertion follows here from

$$\int_0^\infty t^{\eta-1}\pi(x : d(x, x_0) > t)\, dt \;\le\; \int_0^\infty t^{\eta-1}\mathbb{P}(D_0 > t)\, dt$$

for any $\gamma \in (0,1)$ in combination with $\mathbb{E}D_0^\eta < \infty$, see (3.22). ◊

Acknowledgement

The major part of this work was done during a visit of the first author at the Institute of Statistical Science, Academia Sinica to which he wishes to express his gratitude for providing financial support as well as a very kind and stimulating atmosphere.

References

[1] BABILLOT, M., BOUGEROL, P. and ELIE, L. (1997). The random difference equation $X_n = A_n X_{n-1} + B_n$ in the critical case. *Ann. Probab.* **25**, 478-493.

[2] DIACONIS, P. and FREEDMAN, D. (1999). Iterated random functions. *SIAM Review* **41**, 45-76.

[3] CHOW, Y.S. and LAI, T.L. (1975). Some one-sided theorems on the tail distribution of sample sums with applications to the last time and largest excess of boundary crossings. *Trans. Amer. Math. Soc.* **208**, 51-72.

[4] ELTON, J.H. (1990). A multiplicative ergodic theorem for Lipschitz maps. *Stoch. Proc. Appl.* **34**, 39-47.

[5] GUT, A. (1988). *Stopped Random Walks: Limit Theorems and Applications.* Springer, New York.

[6] SILVESTROV, D. and STENFLO, Ö. (1998). Ergodic theorems for iterated function systems controlled by regenerative sequences. *J. Theoret. Probab.* **11**, 589-608.