Ergodic Behaviour of Affine Recursions III

Positive Recurrence and Null Recurrence

Hans G. Kellerer

Mathematisches Institut der Universität München

Theresienstraße 39, D-8000 München 2, Germany

October 30, 1992

Summary. This paper is concerned with the discrete–time Markov process $(X_n)_{n\geq 0}$ solving the stochastic difference equation $X_n = Y_n X_{n-1} + Z_n$ for $n \in \mathbf{N}$, where $(Y_n, Z_n)_{n\in\mathbf{N}}$ is a sequence of i.i.d. random variables independent of the initial variable X_0 and, in accordance with most applications, the state space is restricted to $\mathbf{R}_+$. The study of the recurrent case is completed by distinguishing positive and null recurrence, corresponding to a finite and infinite invariant measure, respectively.

1991 Mathematics Subject Classification. Primary 47A35, 60J05; Secondary 54H20, 60G18, 60G30.

Introduction

This is the final part of work begun in ... and continued in .... It studies affine recursions on $\mathbf{R}_+$, i.e. sequences $(X_n)_{n\geq 0}$ defined by

$$X_n = Y_n X_{n-1} + Z_n \qquad \text{for } n \in \mathbf{N}.$$

Here, $(Y_n, Z_n)_{n\in\mathbf{N}}$ is a sequence of independent identically distributed $\mathbf{R}_+^2$–valued random variables which is independent of the initial variable $X_0 \geq 0$. Without loss of generality, the common distribution ν of $(X_n, Z_n), n \in \mathbf{N}$, will be assumed to belong to the class $\mathcal{N}$ defined in Section 0.

The object of Part II was to study the asymptotic behaviour of $(X_n)_{n\geq 0}$ in the recurrent case, i.e. under the assumption that the sequence does not diverge to infinity. A central result assigns an essentially unique invariant measure to (the transition kernel of) a recurrent sequence $(X_n)_{n\geq 0}$. As a consequence mean as well as pointwise ergodic theorems for ratios can be established. Part III strengthens this study by distinguishing positive recurrence and null recurrence; its contents are summarized below.

Section 8. It is only natural to classify the recurrent case further: the sequence $(X_n)_{n\geq 0}$ is called "positive recurrent" if the invariant measure is finite and "null recurrent" otherwise. This yields the connection between the present work and the existing literature, because a finite invariant measure can be normalized to a stationary distribution. An example for positive recurrence is the regenerative case (8.3) and a simple necessary condition requires divergence of the associated random walk $(S_n)_{n\geq 0}$ to $-\infty$ (8.4). If it exists, the stationary distribution equals the distribution of $\sum_{n\in\mathbf{N}} Y_1 \dots Y_{n-1} Z_n$ (8.2). Employing Kesten's extension of the strong law of large numbers a criterion for positive recurrence can be established (8.5), which can be shown to be best possible in a sense (8.6). When combined with earlier results, this leads to a very satisfasctory trichotomy: under a weak boundedness condition on Z_n, the affine recursion $(X_n)_{n\geq 0}$ is positive recurrent resp. null recurrent resp. transient, if the associated random walk $(S_n)_{n\geq 0}$ diverges to $-\infty$ resp. oscillates resp. diverges to $+\infty$ (8.7). The section closes with a stability result for the stationary distribution under a uniform boundedness condition for moments of Y_n and Z_n of some order (8.8).

Section 9. As in classical Markov chain theory positive recurrence can be characterized by weak convergence of the laws $\mathcal{L}(X_n)$ to the stationary distribution, while in the null recurrent case the total mass disappears to infinity (9.1). In the positive recurrent case the ergodic theorems from Section 7 can be strengthened considerably: Under the stationary distribution the sequence $(X_n)_{n\geq 0}$ is actually mixing (9.2), while under an arbitrary initial law almost sure convergence of successive averages holds more generally than for bounded continuous functions (9.3). The second half of this section treats mean passage times. While ascending ladder indices have a finite expectation in any case (9.4), the ergodic behaviour of the sequence is essential for descending ladder indices (9.5). The section closes with an extension of the main criterion for positive/null recurrence from classical Markov chain theory to the topological setting (9.6).

Section 10. Since the moments of a stationary distribution μ, as far as they exist, can be obtained recursively from the mixed moments of Y_n and Z_n (10.1), it is of interest whether μ is determined by all of the moments $\int x^k d\mu$, $k \in \mathbf{N}$. This leads in a natural way to the weakly contractive case (10.2). The strongly contractive case establishes the connection with self–similarity. It provides one of the rare situations allowing explicit results, for instance concerning uniform distributions (10.3). The second half of this section studies the multiplicative model with $Y_n \leq 1$, which is important in some applications. Here at least a partial answer concerning the singularity of the stationary distribution is possible (10.4). Finally, the exact order in which an upper limit ∞ is approached by $(X_n)_{n\geq 0}$ is derived (10.6).

In conclusion, here are some directions for further research:
— As already mentioned, the extension from $\mathbf{R}_+$ to $\mathbf{R}$ – for some results without difficulties – on the whole is a nontrivial problem.
— It is much easier to keep the state space $\mathbf{R}_+$ and to deal with other monotone recursions as for instance the process

$$X_n = Y_n X_{n-1} \vee Z_n \qquad \text{for } n \in \mathbf{N},$$

as considered by Goldie [16] and Rachev [34] (for a special case see also Alpuim [1]).
— New problems arise, if the total order in $\mathbf{R}$ is given up and $\mathbf{R}^k$ is considered instead; for some results in the additive model see Elton and Yan [12].
— A multidimensional generalization results also from autoregressive processes of higher order, treated in the context of random matrices for instance by Kesten [24].
— Even more general is a treatment in the framework of topological semigroups as begun by Mukherjea and Tserpes [31] and continued recently in the special case of nonnegative matrices by Mukherjea [30].
— There are, moreover, attempts to weaken the asumption on the sequence $(Y_n, Z_n)_{n\in\mathbf{N}}$ to stationarity and ergodicity; here Borovkov [4] and Brandt [5] have to be mentioned.
— Finally, it is only natural to ask for continuous–time analogues; for these Wolfe [40] and de Haan and Karandikar [19] may be consulted.

8. Positive recurrence and null recurrence

A further classification in the recurrent case is suggested by (5.6):

(8.1) Definition. *Let ν be recurrent with invariant measure μ. Then the distribution ν (or the kernel P or the process $(X_n)_{n\geq 0}$) is called*

(a) *"positive recurrent"* *if* $\mu(\mathbf{R}_+) < \infty$,

(b) *"null recurrent"* *if* $\mu(\mathbf{R}_+) = \infty$.

The simplest example for positive recurrence is provided by the case $\overline{x} < \infty$, in which the support of μ is bounded. A first general criterion, derived under moment conditions but not restricted to the case $Y, Z \geq 0$, was given by Vervaat [39]. In the present framework it is easily established:

(8.2) Theorem. *ν is positive recurrent if and only if*

$$W := \textstyle\sum_{n\in\mathbf{N}} Y_1 \dots Y_{n-1} Z_n < \infty \quad a.s.$$

In this case the stationary distribution is given by $\mathcal{L}(W)$.

Proof. 1. If the invariant measure μ is finite, $\mu \in \mathcal{M}_1(\mathbf{R}_+)$ may be assumed as well. Under the initial law μ the sequence $(X_n)_{n\geq 0}$ is stationary, and with the dual sequence $(W_n)_{n\geq 0}$, defined in Section 0, this implies by monotonicity

$$\begin{aligned}
\mu([0,t]) &= (\liminf_{n\to\infty}) \textstyle\int \mathbf{P}(X_n^x \leq t)\, \mu(dx) \\
&\leq \liminf_{n\to\infty} \mathbf{P}(X_n^0 \leq t) \\
&= \liminf_{n\to\infty} \mathbf{P}(W_n \leq t) \\
&= \mathbf{P}(W \leq t) \\
&\leq \mathbf{P}(W < \infty),
\end{aligned}$$

which for $t \to \infty$ yields the assertion.

2. To prove the converse, let (Y_0, Z_0) be independent of (Y_n, Z_n), $n \in \mathbf{N}$, with distribution ν. Then $Z_0 + Y_0 W$ is distributed as W, hence $\mu = \mathcal{L}(W)$ is a stationary distribution for P. Moreover, the recurrence of ν is a consequence of (2.2b), because the potential kernel $G := \sum_{n\geq 0} P^n$ satisfies by monotonicity

$$\begin{aligned}
G(0;[0,t]) &\geq \textstyle\int G(x;[0,t])\, \mu(dx) \\
&= \textstyle\sum_{n\geq 0} (\mu P^n)([0,t]) \\
&= \textstyle\sum_{n\geq 0} \mu([0,t]) \qquad \text{for all } t \geq 0. \quad \square
\end{aligned}$$

To derive the stationary distribution μ associated with a positive recurrent distribution ν, (8.2) is not of much use. For some explicit examples, obtained by an ad hoc principle in the framework of exponential families, see [7].

The next result strengthens (2.3):

(8.3) Proposition. *ν is positive recurrent whenever* $\mathbf{P}(Y = 0) > 0$.

Proof. Since $Y_n = 0$ for some $n \in \mathbf{N}$ with probability 1, the series in (8.2) is a finite sum almost surely. $\square$

In this context it is worthwhile to point out the following:

— The stationary distribution μ can be computed, at least in principle, from the probabilities $p := \mathbf{P}(Y=0)$ and $q := 1-p$ and the distributions ν_0 and ν' of (Y,Z) under the condition $Y=0$ and $Y \neq 0$, respectively (the assumption $q \neq 0$ is no real restriction). To this end let (Y_0, Z_0) and (Y'_n, Z'_n), $n \in \mathbf{N}$, be independent with distributions ν_0 and ν', respectively, and denote by W'_n the analogue of W_n with (Y'_m, Z'_m), $m < n$, replacing (Y_m, Z_m), $m < n$, and (Y_0, Z_0) replacing (Y_n, Z_n). Then it is easily seen that

$$\mu = \textstyle\sum_{n\in\mathbf{N}} p\, q^{n-1} \mu'_n \qquad \text{with } \mu'_n := \mathcal{L}(W'_n).$$

— While in the null recurrent case the hypotheses of (6.4) are satisfied and thus the invariant measure is nonatomic, in the positive recurrent case the mass $\mu(\{0\})$ of the stationary distribution can be computed from the equation

$$\begin{aligned} \mu(\{0\}) &= (\mu \otimes \nu)(\{(x; y, z) : yx + z = 0\}) \\ &= \mu(\{0\})\, \mathbf{P}(Z=0) + \mu(]0,\infty[)\, \mathbf{P}(Y = 0 = Z), \end{aligned}$$

resulting in

$$\mu(\{0\}) = \mathbf{P}(Y = 0 = Z) \,/\, (1 - \mathbf{P}(Y \neq 0 = Z)).$$

— If Y is restricted to the values 0 and 1, then the Laplace transform ψ of μ can be computed from (5.7). To this end let χ_i denote the Laplace transform of the distribution of Z under the condition $Y = i$ $(i = 0,\ 1)$. Then $\psi(0) = 1$ leads to

$$\psi(u) = \mathbf{P}(Y=0)\, \chi_0(u) \,/\, (1 - \mathbf{P}(Y=1)\, \chi_1(u))$$

(for a related result see [39]).

Next, the analogue of (2.4) can be stated, providing a necessary condition for positive recurrence instead of a sufficient condition for transience:

(8.4) Proposition. *If ν is positive recurrent, then $S_n \to -\infty$.*

Proof. Assuming $\sup_{n\geq 0} S_n = +\infty$, consider the random times $T_1 < T_2 < \dots$ when $(S_n)_{n\geq 0}$ hits $\mathbf{R}_+$ (being defined with probability 1). Since T_k, $k \in \mathbf{N}$, are stopping times with respect to $(Y_n, Z_n)_{n\in\mathbf{N}}$, the random variables Z_{T_k+1}, $k \in \mathbf{N}$, are independent and distributed as Z. Thus

$$W \geq \textstyle\sum_{k\in\mathbf{N}} Y_1 \dots Y_{T_k} Z_{T_k+1} \geq \sum_{k\in\mathbf{N}} Z_{T_k+1}$$

diverges almost surely and (8.2) yields a contradiction. □

That the converse of (8.4) does not hold in general is an immediate consequence of (2.5).

A first criterion for $\mathbf{P}(W < \infty) = 1$ appears in [27] under the assumption $\mathbf{E}(\log Y) < 0$, ensuring $S_n \to -\infty$. Later its sufficiency was shown by Vervaat [39] along the same lines but under a weaker moment assumption. Employing Kesten's extension of the strong law of large numbers, this assumption actually can be dropped completely:

(8.5) Theorem. *If $S_n \to -\infty$, then the condition*

$$\mathbf{E}(\log_+ Z) < \infty$$

(a) *is always sufficient for positive recurrence,*

(b) *is necessary for positive recurrence whenever $\mathbf{E}(|\log Y|) < \infty$.*

Proof. Applying Borel–Cantelli it follows easily that

$$\limsup_{n\to\infty} \frac{1}{n} \log_+ Z_n \begin{cases} = 0 & \text{a.s. if } \mathbf{E}(\log_+ Z) < \infty, \\ = +\infty & \text{a.s. if } \mathbf{E}(\log_+ Z) = \infty. \end{cases}$$

Applying [23], moreover, it follows from the hypothesis $S_n \to -\infty$ that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{1\le m<n} \log Y_m \begin{cases} \in]-\infty, 0[& \text{a.s. if } \mathbf{E}(|\log Y|) < \infty, \\ = -\infty & \text{a.s. if } \mathbf{E}(|\log Y|) = \infty, \end{cases}$$

because the trichotomy for $(S_n)_{n\ge0}$ carries over to $(S_n/n)_{n\in\mathbf{N}}$ in the case

$$\mathbf{E}(\log_- Y) = \infty = \mathbf{E}(\log_+ Y).$$

Therefore

(1) $\limsup_{n\to\infty} (Y_1 \ldots Y_{n-1} Z_n)^{\frac{1}{n}} < 1$ a.s. whenever $\mathbf{E}(\log_+ Z) < \infty$,

while

(2) $\limsup_{n\to\infty} (Y_1 \ldots Y_{n-1} Z_n)^{\frac{1}{n}} = \infty$ a.s. whenever $\mathbf{E}(\log_+ Z) = \infty$,

if in this case $\mathbf{E}(|\log Y|) < \infty$ holds in addition. Clearly, (1) and (2) imply almost sure convergence and divergence of W, respectively, and the assertion follows from (8.2). □

Specialized to $Y = \gamma \in]0,1[$ and combined with (3.1), this result shows that even in the additive model all three possibilities – transience, null recurrence, positive recurrence – really occur.

The statement of (8.5b) is best possible in some sense. However large Z may be, Y can be made small enough for positive recurrence without attaining the value 0, even if in addition independence of Y and Z is postulated. More precisely, (2.5) has the following counterpart:

(8.6) Proposition. *For any $\nu_z \in \mathcal{M}_1(\mathbf{R}_+)$ with $\nu_z(\{0\}) \neq 1$ there exists $\nu_y \in \mathcal{M}_1(\mathbf{R}_+)$ with $\nu_y(\{0\}) = 0$ and $\nu_y(\{1\}) \neq 1$ such that $\nu = \nu_y \otimes \nu_z$ is positive recurrent.*

Proof. Let Z_n, $n \in \mathbf{N}$, be independent with distribution ν_z. Then choose sequences $(z_n)_{n\in\mathbf{N}}, (y_n)_{n\in\mathbf{N}}, (p_n)_{n\in\mathbf{N}}$ satisfying

(1) $\sum_{n\in\mathbf{N}} q_n < \infty$ with $q_n := \mathbf{P}(Z > z_n)$,

(2) $1 = y_0 > y_1 > \ldots \to 0$ and $\sum_{n\in\mathbf{N}} y_n z_n < \infty$,

(3) $\quad 0 = p_0 < p_1 < \ldots \to 1 \quad$ and $\quad \sum_{n\in\mathbf{N}} p_n^{n-1} < \infty$

(e.g. $p_n = n^{-2/(n-1)}$ for $n > 1$). Independently of $Z_n, n\in\mathbf{N}$, let now $Y_n, n\in\mathbf{N}$, be independent with a distribution ν_y such that

$$\mathbf{P}(Y > y_n) = p_n \qquad \text{for } n \geq 0.$$

Then $0 < Y \leq 1$ and thus

$$\begin{aligned}\mathbf{P}(Y_1 \ldots Y_{n-1} Z_n > y_n z_n) \;&\leq\; \mathbf{P}(Y_1, \ldots, Y_{n-1} > y_n) \;+\; \mathbf{P}(Z_n > z_n)\\ &=\; p_n^{n-1} + q_n \qquad \text{for } n \in \mathbf{N}.\end{aligned}$$

Therefore, by the summability in (1) and (3), with probability 1

$$Y_1 \ldots Y_{n-1} Z_n \leq y_n z_n \qquad \text{for almost all } n \in \mathbf{N},$$

hence, by the summability in (2), the assertion follows from (8.2). $\square$

Now the close relationship between the behaviour of an affine recursion $(X_n)_{n\geq 0}$ and the associated random walk $(S_n)_{n\geq 0}$ can be summarized. Under a weak boundedness condition on Z, satisfied for instance in the multiplicative model, there is in fact a bijection:

(8.7) Theorem. *With arbitrary $\alpha > 0$ let one of the following two conditions be satisfied:*

(a) $\quad \mathbf{E}(Z^\alpha \mid Y) \leq \gamma \quad$ *for some* $\gamma < \infty$,

(b) $\quad \mathbf{P}(Y = 0) = 0 \quad$ *and* $\quad \mathbf{E}\Big(\big(\frac{Z}{Y}\big)^\alpha\Big) < \infty.$

Then the following trichotomy holds:

$$S_n \to +\infty \implies \nu \textit{ transient},$$

$$S_n \to \pm\infty \implies \nu \textit{ null recurrent},$$

$$S_n \to -\infty \implies \nu \textit{ positive recurrent}.$$

Proof. 1. Clearly, $\alpha \leq 1$ may be assumed and $\alpha = 1$ will be considered first. Then the case $S_n \to +\infty$ is settled by (2.4), while the case $S_n \to \pm\infty$ follows from (8.4). Thus it remains to consider the case $S_n \to -\infty$. Under condition (a) clearly $\mathbf{E}(\log_+ Z) < \infty$ and (8.5a) applies, while under condition (b) similarly $\mathbf{E}(\log_+(Z/Y)) < \infty$. In view of

$$W = \sum_{n\in\mathbf{N}} Y_1 \ldots Y_n \frac{Z_n}{Y_n}$$

the assertion follows from (8.2) as in the proof of (8.5).

2. For $\alpha < 1$ the transient case is again settled by (2.4), because the passage from $(Y', Z') = (Y^\alpha, Z^\alpha)$ to (Y, Z) does not affect the asymptotic behaviour of the associated random walk. The recurrent case follows from the inequalities

$$Z_1 Y_2 \dots Y_n + \dots + Z_n \leq (Z_1' Y_2' \dots Y_n' + \dots + Z_n')^{\frac{1}{\alpha}},$$

$$Z_1 + \dots + Y_1 \dots Y_{n-1} Z_n \leq (Z_1' + \dots + Y_1' \dots Y_{n-1}' Z_n')^{\frac{1}{\alpha}}.$$

Indeed, if $\underline{x}'$ and W' denote the analogues of $\underline{x}$ and W, this means

$$\underline{x} \leq (\underline{x}')^{\frac{1}{\alpha}} \qquad \text{and} \qquad W \leq (W')^{\frac{1}{\alpha}}. \quad \square$$

It should be mentioned that conditions (a) and (b) according to (3.4) can be simplified to

$$\mathbf{E}(Z^\alpha) < \infty \qquad \text{for some } \alpha > 0,$$

provided Y and Z are independent or Y is bounded away from 0.

This section will be concluded by a stability result. Related, but not comparable, conditions for continuous dependence of μ on ν can be found in [5] (see also [6], chapter 9.1). The following result strengthens (6.5):

(8.8) Proposition. *Let $\mathcal{N} \ni \nu_k \overset{\text{w}}{\to} \nu$ and ν_k be positive recurrent with stationary distributions μ_k such that*

$$\gamma := \sup_{k \in \mathbf{N}} \int y^\alpha \, d\nu_k < 1 \qquad \textit{and} \qquad \delta := \sup_{k \in \mathbf{N}} \int z^\alpha \, d\nu_k < \infty$$

for some $\alpha > 0$. Then

(a) *ν is positive recurrent,*

(b) $\mu_k \overset{\text{w}}{\to} \mu$,

where μ is the stationary distribution associated with ν.

Proof. (a) In view of the weak convergence

$$\int y^\alpha \, d\nu \leq \gamma \qquad \text{and} \qquad \int z^\alpha \, d\nu \leq \delta,$$

which by the hypotheses on γ and δ (and Jensen's inequality) yields

$$-\infty \leq \int \log y \, d\nu < 0 \qquad \text{and} \qquad \int \log_+ z \, d\nu < \infty.$$

Thus ν is positive recurrent according to (8.5a).

(b) Since the L_p–norms increase with the order, clearly $\alpha \leq 1$ may be assumed. In this case

$$\begin{aligned} \int x^\alpha \, d\mu_k &= \int \int (xy + z)^\alpha \, \mu_k(dx) \, \nu_k(dy, dz) \\ &\leq \int \int ((xy)^\alpha + z^\alpha) \, \mu_k(dx) \, \nu_k(dy, dz) \\ &\leq \left(\int x^\alpha \, d\mu_k\right) \gamma + \delta, \end{aligned}$$

which implies

$$\sup_{k\in\mathbf{N}} \int x^\alpha \, d\mu_k \le \frac{\delta}{1-\gamma} < \infty.$$

In view of $\alpha > 0$, therefore, the sequence $(\mu_k)_{k\in\mathbf{N}}$ is uniformly tight in $\mathcal{M}_1(\mathbf{R}_+)$ and each limit point μ is excessive with respect to ν according to (4.6), because weak convergence implies vague convergence. The assertion thus follows from (5.6). □

In this result the conditions on γ and δ both can be shown to be essential. The following counterexamples resemble those in [5], represent, however, even more extreme situations. In both cases the notation $\mathbf{P}_k$ refers to ν_k.

(1) $$\mathbf{P}_k(Z = \frac{k}{y^k}) = \frac{1}{\sqrt{k}}, \quad \mathbf{P}_k(Z = z) = 1 - \frac{1}{\sqrt{k}}$$

with $z > 0$ and $Y = y \in]0,1[$ yield positive recurrent ν_k, $k \in \mathbf{N}$, such that

$$\nu_k \overset{\mathrm{w}}{\to} \varepsilon_{(y,z)} \qquad \text{and} \qquad \mu_k \overset{\mathrm{w}}{\to} \varepsilon_\infty.$$

The first convergence is clear and the second one follows from

$$\begin{aligned} \mathbf{P}_k(W \le k) &\le \mathbf{P}_k(Z_i = z \ \text{ for } \ 1 \le i \le k) \\ &= (1 - \frac{1}{\sqrt{k}})^k \to 0. \end{aligned}$$

(2) $$\mathbf{P}_k(Y = 0) = \frac{1}{k^3}, \quad \mathbf{P}_k(Y = \frac{1}{k}) = 1 - \frac{1}{k}, \quad \mathbf{P}_k(Y = k^{(k^2)}) = \frac{1}{k} - \frac{1}{k^3}$$

and $Z = 1$ yield positive recurrent ν_k, $k \in \mathbf{N}$, such that

$$\nu_k \overset{\mathrm{w}}{\to} \varepsilon_{(0,1)} \qquad \text{and} \qquad \mu_k \overset{\mathrm{w}}{\to} \varepsilon_\infty.$$

Again the first convergence is clear, while the second one follows from

$$\begin{aligned} \mathbf{P}_k(W < k) &\le \mathbf{P}_k(Y_1 \dots Y_{k^2} < k) \\ &\le \mathbf{P}_k(\textstyle\bigcup_{1\le i\le k^2}\{Y_i = 0\}) + \mathbf{P}_k(\textstyle\bigcap_{1\le i\le k^2}\{Y_i = \frac{1}{k}\}) \\ &= 1 - (1 - \frac{1}{k^3})^{k^2} + (1 - \frac{1}{k})^{k^2} \to 0. \end{aligned}$$

9. Further ergodic theorems

The mean ergodic theorem (7.1) can be strengthened to weak convergence in the positive recurrent case, holding in an extended sense also in both the other cases. This is a consequence of more general results in [18], but for the state space $\mathbf{R}_+$ a direct proof is too simple to refer to other work. In the present framework the following holds, independently of the initial law:

(9.1) Theorem. *If* $\mu_n := \mathcal{L}(X_n)$ *for* $n \geq 0$, *then*

(a) $\mu_n \overset{\text{w}}{\to} \mu,$

whenever ν *is positive recurrent with stationary distribution* μ,

(b) $\mu_n \overset{\text{w}}{\to} \varepsilon_\infty,$ *i.e.* $\mu_n([0,t]) \to 0$ *for all* $t < \infty$,

whenever ν *is null recurrent or transient.*

Proof. (a) By bounded convergence (2.6) implies

$$\mu_n f - \mu f = \mathbf{E}(\textstyle\int (f(X_n) - f(X_n^x))\, \mu(dx)) \to 0 \qquad \text{for all } f \in \mathcal{K}(\mathbf{R}_+).$$

(b) It follows from the representation

$$W = Z_1 + \ldots + Y_1 \ldots Y_{n-1}(Z_n + Y_n Z_{n+1} + \ldots) \qquad \text{for } n \in \mathbf{N}$$

that the event $\{W = \infty\}$ is contained in (the completion of) the tail–field of $(Y_n, Z_n)_{n \in \mathbf{N}}$, because $\mathbf{P}(Y = 0) = 0$ by (8.3). Therefore $W = \infty$ a.s. by (8.2) and this implies by monotonicity

$$\begin{aligned} \mu_n([0,t]) &\leq \mathbf{P}(X_n^0 \leq t) \\ &= \mathbf{P}(W_n \leq t) \\ &\to \mathbf{P}(W \leq t) = 0 \qquad \text{for all } t < \infty. \quad \square \end{aligned}$$

Clearly, (9.1) strengthens (7.6) in the positive recurrent case. Moreover, the condition "X finite–valued" in (1.4b) can now be replaced by the condition $\mathbf{P}(X = \infty) \neq 1$, because convergence in probability implies weak convergence.

The pointwise ergodic theorem for the positive recurrent case can be derived from (7.4). To obtain it for as many functions as possible, it is, however, preferable to establish first ergodicity under stationarity. Actually a stronger result holds:

(9.2) Proposition. *If the process* $(X_n)_{n\geq 0}$ *is stationary, it is mixing.*

Proof. 1. Extending (Y_n, Z_n), $n \in \mathbf{N}$, let (Y_n, Z_n), $n \in \mathbf{Z}$, be independent with distribution ν. Then by (8.2)

$$X_n' := \textstyle\sum_{-\infty < m \leq n} Z_m Y_{m+1} \ldots Y_n < \infty \text{ a.s.}$$

and, moreover, $\mu = \mathcal{L}(X_n')$ is the stationary distribution. Since

$$X_n' = X_{n-1}' Y_n + Z_n \qquad \text{for } n \in \mathbf{N},$$

the processes $(X_n)_{n\geq 0}$ and $(X_n')_{n\geq 0}$ have the same distribution and it suffices to prove the assertion for $(X_n')_{n\geq 0}$.

2. Denote by σ and σ' the shift in $\prod_{n\in\mathbf{Z}} \mathbf{R}_+^2$ and in $\prod_{n\geq 0} \mathbf{R}_+$, respectively, and consider the mapping

$$\tau : (y_n, z_n)_{n\in\mathbf{Z}} \to (\textstyle\sum_{-\infty<m\leq n} z_m y_{m+1} \cdots y_n)_{n\geq 0}$$

from $\prod_{n\in\mathbf{Z}} \mathbf{R}_+^2$ to $\prod_{n\geq 0} \mathbf{R}_+$. It is compatible with σ and σ', i.e.

$$\tau \circ \sigma = \sigma' \circ \tau.$$

Therefore the mixing property of σ with respect to the measure $\bigotimes_{n\in\mathbf{Z}} \nu$ carries over to σ' with respect to its image by τ. Since this is the distribution of $(X'_n)_{n\geq 0}$, the assertion follows. □

In the case $\mathbf{P}(Y = 0) > 0$ it can be shown that (9.2) holds in the stronger sense of $(X_n)_{n\geq 0}$ having a trivial tail–field. This can fail extremely in the general case, as is seen by the following example:

$$Y = 1/2 \qquad \text{and} \qquad \mathbf{P}(Z = 0) = 1/2 = \mathbf{P}(Z = 1)$$

obviously represents the simplest nontrivial positive recurrent case. Here $(X_n)_{n\geq 0}$ is stationary, if the initial law is the uniform distribution on [0,2]. Since X_{n-1} can be reconstructed from X_n via

$$X_{n-1} = 2\,(X_n - 1_{\{X_n>1\}}) \text{ a.s.},$$

the tail–field of $(X_n)_{n\geq 0}$ coincides in this case with the full σ–algebra generated by $(X_n)_{n\geq 0}$ (modulo null sets).

Now the pointwise ergodic theorem is a simple consequence. Related results in the context of Lipschitz maps can be found in [3], [10], [11]. In the present framework the following is true, again independently of the initial law:

(9.3) Theorem. *Let ν be positive recurrent with stationary distribution μ and f be μ–integrable. Then the convergence*

$$\frac{1}{n} \textstyle\sum_{0\leq m<n} f(X_m) \to \mu f \quad a.s.$$

holds under each of the following conditions:

(a) $f \in \mathcal{C}_\mu(\mathbf{R}_+)$,

(b) *f uniformly continuous (with respect to Euclidean metric).*

Proof. 1. If $\mu_0 := \mathcal{L}(X_0)$ equals μ, the process $(X_n)_{n\geq 0}$ is ergodic by (9.2), hence the asserted convergence holds without any further condition on f. This extends to arbitrary μ_0 by (2.6), if f is restricted to $\mathcal{K}(\mathbf{R}_+)$.

2. If $f \in \mathcal{C}_\mu(\mathbf{R}_+)$ and for the moment $0 \leq f \leq 1$ is assumed, monotone approximation from $\mathcal{K}(\mathbf{R}_+)$ yields the inequality

$$\liminf_{n\to\infty} \frac{1}{n} \textstyle\sum_{0\leq m<n} f(X_m) \geq \mu f \text{ a.s.}$$

and its counterpart for $1-f$. Both inequalities together, in view of $\mu 1 = 1$, prove the assertion under condition (a), because the assumption $0 \le f \le 1$ is easily removed.

3. Under condition (b)

$$f(X_n^{x_1}) - f(X_n^{x_2}) \to 0 \text{ a.s.} \qquad \text{for all } x_i \in \mathbf{R}_+$$

follows from

$$X_n^{x_1} - X_n^{x_2} = (x_1 - x_2)\, e^{S_n} \to 0 \text{ a.s.},$$

which in turn follows from (8.4). Thus the initial law μ_0 is irrelevant for the asserted convergence. □

Applying (9.3) to a countable dense subset of the normed space $\mathcal{K}(\mathbf{R}_+)$ yields as usual

$$\frac{1}{n} \sum_{0 \le m < n} \varepsilon_{X_m} \xrightarrow{\text{w}} \mu \text{ a.s.},$$

i.e. the stationary distribution is the weak limit of the empirical distributions almost surely.

As an important special case (9.3b) yields the strong law of large numbers for the process $(X_n)_{n \ge 0}$, stating

$$\frac{1}{n} \sum_{0 \le m < n} X_m \to \int x \, d\mu \text{ a.s.}$$

and holding also, if the integral is infinite.

While in the case $\mathbf{P}(Y=0) > 0$ the conditions (a) and (b) can be shown to be superfluous, in the general case (9.3) can hardly be improved. Indeed convergence can fail, if f is assumed to be μ–integrable and continuous only, as is seen by the following example:

$$\nu = p\, \varepsilon_{(1/2\,,\,1)} + q\, \varepsilon_{(2\,,\,0)} \qquad \text{with } 0 < q < p < 1$$

is obviously positive recurrent with $\underline{x} = 2$ and $\overline{x} = \infty$. For $\mu_0 = \varepsilon_0$ consider the random times

$$T_k := \inf\{n \in \mathbf{N} : X_n \ge 2^k\} \qquad \text{for } k \in \mathbf{N}.$$

Since $X_n < 2^k$ implies $X_{n+1} < 2^{k+1}$, they satisfy $T_1 < T_2 < \dots$, both with probability 1. Choose $n_k \in \mathbf{N}$ such that

(1) $\quad \limsup_{k \to \infty} \mathbf{P}(T_k \le n_k) = 1$

and consider the sets

$$A_k := \{x : \mathbf{P}(T_k \le n_k, X_{T_k} = x) > 0\}.$$

Since A_k is a finite subset of

$$B_k := [2^k, 2^{k+1}[$$

and the stationary distribution μ is nonatomic by (6.4), a continuous function $f : \mathbf{R}_+ \to \mathbf{R}_+$ can be constructed such that

(2) $\quad f(x) \geq k\, n_k \qquad \text{for } x \in A_k,$

(3) $\quad \int_{B_k} f\, d\mu \leq 2^{-k} \qquad \text{for } k \in \mathbf{N}.$

Then by (2) the sequence

$$V_n := \frac{1}{n} \sum_{1 \leq m \leq n} f(X_m) \qquad \text{for } n \in \mathbf{N}$$

satisfies

$$\begin{aligned} \mathbf{P}(\limsup_{n\to\infty} V_n = \infty) &\geq \mathbf{P}(\limsup_{k\to\infty} \{V_{T_k} \geq k\}) \\ &\geq \mathbf{P}(\limsup_{k\to\infty} \{T_k \leq n_k\}), \end{aligned}$$

where the last inequality follows from

$$V_{T_k} \geq \frac{1}{T_k} f(X_{T_k}) \geq k \qquad \text{for } T_k \leq n_k.$$

In view of (1) this implies

$$\limsup_{n\to\infty} V_n = +\infty \ \text{ a.s.},$$

while by (3) on the other hand $\mu f \leq 1$.

The rest of this section treats mean passage times, where first the region above and below the initial level will be considered. Whether ν is recurrent or transient, the ascending ladder indices have a finite expectation:

(9.4) Proposition. *Let T_B denote the – possibly infinite – hitting time of a set $B \in \mathcal{B}(\mathbf{R}_+)$ by $(X_n)_{n\geq 0}$. Then, for arbitrary $x < \overline{x}$,*

$$\mathbf{E}^x(T_{[x,\infty[}) < \infty.$$

Proof. Due to $x < \overline{x}$ there exists $n \in \mathbf{N}$ such that

$$\vartheta := \mathbf{P}(X_n^0 \geq x) > 0.$$

Since by monotonicity

$$\begin{aligned} \{T_{[x,\infty[} > kn\} &\subset \{X_n < x, \ldots, X_{kn} < x\} \\ &\subset \{X_n^0 < x, \ldots, X_{kn}^0 < x\} \\ &\subset \bigcap_{0 \leq i < k} \{Z_{in+1} Y_{in+2} \ldots Y_{in+n} + \ldots + Z_{in+n} < x\}, \end{aligned}$$

independence implies

$$\mathbf{P}(T_{[x,\infty[} > kn) \leq (\mathbf{P}(X_n^0 < x))^k = (1-\vartheta)^k \qquad \text{for } k \geq 0.$$

Therefore the assertion holds for any initial distribution. □

The following counterpart of (9.4) distinguishes positive and null recurrence:

(9.5) Proposition. *Let T_B be defined as in* (9.4) *and ν be recurrent. Then, for arbitrary $x > \underline{x}$,*

(a) $\mathbf{E}^x(T_{[0,x]}) < \infty$ *whenever ν is positive recurrent,*

(b) $\mathbf{E}^x(T_{[0,x]}) = \infty$ *whenever ν is null recurrent.*

Proof. (a) Let μ be the stationary distribution and $\mathcal{L}(X_0) = \mu$. Then it follows from the recurrence theorem by Kac [21] that

$$\mathbf{E}(T_{[0,x]} \mid X_0 \leq x) = (\mu([0,x]))^{-1}.$$

Since by (5.2a) the hitting distribution on $[0,x]$ equals the distribution of X_0 under the condition $X_0 \leq x$, by induction this equation can be extended to the k–th hitting time T_k of $[0,x]$, i.e.

$$\mathbf{E}(T_k \mid X_0 \leq x) = k\,(\mu([0,x]))^{-1}.$$

This implies all that is needed in the sequel, namely

(1) $\int_{\{X_0 \leq x\}} T_k \, d\mathbf{P} < \infty$ for all $k \in \mathbf{N}$.

Now $x < \overline{x}$ may be assumed, because otherwise $\mathbf{E}^x(T_{[0,x]}) = 1$ by (1.2b). Under this assumption there exists $k \in \mathbf{N}$ such that

(2) $\mathbf{P}(X_0 \leq x < X_k) > 0.$

By the inequality

$$T := \inf\{n > k : X_n \leq x\} \leq T_{k+1},$$

with $\mu_{0,k}$ denoting the distribution of (X_0, X_k), and by monotonicity it follows that

$$\begin{aligned} \int_{\{X_0 \leq x\}} T_{k+1} \, d\mathbf{P} &\geq \int_{\{X_0 \leq x < X_k\}} T \, d\mathbf{P} \\ &= \int_{x_0 \leq x < x_k} (k + \mathbf{E}^{x_k}(T_{[0,x]}))\, \mu_{0,k}(dx_0, dx_k) \\ &\geq \mathbf{P}(X_0 \leq x < X_k)\, \mathbf{E}^x(T_{[0,x]}). \end{aligned}$$

This proves the assertion in view of (1) and (2).

(b) Let μ be the invariant measure and $x < t < \infty$. Then

$$\mu'(B) := (\mu([0,t]))^{-1}\mu(B) \qquad \text{for } B \in \mathcal{B}([0,t])$$

by (5.2a) defines a stationary distribution with respect to ${}^tP$. Now let X_0 be distributed according to (the trivial extension of) μ' and let T' denote the hitting time of $[0,x]$ by $({}^tX_n)_{n\geq 0}$. Then it is obvious that

$$T' \leq T_{[0,x]},$$

and it follows from the above–mentioned theorem by Kac that

$$\int_{\{X_0 \leq x\}} T' d\mathbf{P} = 1.$$

By monotonicity this implies

$$\begin{aligned} 1 &= \int_{x_0 \leq x} \mathbf{E}^{x_0}(T')\, \mu'(dx_0) \\ &\leq \int_{x_0 \leq x} \mathbf{E}^{x_0}(T_{[0,x]})\, \mu'(dx_0) \\ &\leq \mu'([0,x])\, \mathbf{E}^{x}(T_{[0,x]}), \end{aligned}$$

hence

$$\mu([0,t]) \leq \mu([0,x])\, \mathbf{E}^{x}(T_{[0,x]}).$$

In view of $\mu([0,x]) < \infty$ the assertion follows for $t \to \infty$. $\square$

Since $\mathbf{P}^x(T_{[0,x]} < \infty) = 1$ obviously implies $\underline{x} \leq x$, in the transient case the equation $\mathbf{E}^x(T_{[0,x]}) = \infty$ holds for all $x \in \mathbf{R}_+$.

The next result concerns mean recurrence times and makes again essential use of the monotonicity:

(9.6) Theorem. *Let T_B be defined as in* (9.4) *and ν be recurrent with invariant measure μ. Then, for arbitrary $x \in \operatorname{supp} \mu$, ν is positive recurrent if and only if*

$$\mathbf{E}^x(T_G) < \infty \qquad \textit{for all open subsets } G \textit{ of } \mathbf{R}_+ \textit{ containing } x.$$

Proof. 1. Let ν be positive recurrent and consider first the case $x = 0$. Then $G = [0, 2\varepsilon[$ may be assumed, which yields by (9.5a)

$$\mathbf{E}^0(T_G) \leq \mathbf{E}^\varepsilon(T_G) \leq \mathbf{E}^\varepsilon(T_{[0,\varepsilon]}) < \infty.$$

In the case $x > 0$ similarly $G =]x-\varepsilon, x+\varepsilon[\subset \mathbf{R}_+$ may be assumed. Choose first $G' =]x-\varepsilon', x+\varepsilon'[$ with arbitrary $\varepsilon' \in]0, \varepsilon[$. In view of $\mu(G') > 0$ it follows as in the proof of (9.5a) that

$$(1) \qquad \mathbf{E}^{x'}(T_{G'}) < \infty \qquad \text{for some } x' \in G'.$$

For fixed $\omega \in \Omega$ consider now $n \in \mathbf{N}$ with $X_n^{x'}(\omega) \in G'$. Then the inequality

$$\begin{aligned} |X_n^x(\omega) - X_n^{x'}(\omega)| &= |x - x'| \prod_{1 \leq m \leq n} Y_m(\omega) \\ &\leq \varepsilon' \frac{1}{x'} \big(x' \prod_{1 \leq m \leq n} Y_m(\omega)\big) \\ &< \varepsilon' \frac{1}{x - \varepsilon'} (x + \varepsilon') \end{aligned}$$

in view of $|X_n^{x'}(\omega) - x| < \varepsilon'$ implies that

$$|X_n^x(\omega) - x| < \varepsilon' \left(\frac{x + \varepsilon'}{x - \varepsilon'} + 1\right) = \frac{2\varepsilon' x}{x - \varepsilon'}.$$

Therefore by (1)

$$\mathbf{E}^x(T_G) \leq \mathbf{E}^{x'}(T_{G'}) < \infty \qquad \text{for } \frac{2\varepsilon' x}{x-\varepsilon'} \leq \varepsilon,$$

where the last condition is satisfied for sufficiently small ε'.

2. To prove the converse suppose only

$$\mathbf{E}^x(T_{[0,t]}) < \infty \qquad \text{for all } t > x.$$

Assuming now ν not to be positive recurrent implies

$$\text{(2)} \qquad \mathbf{P}(Yx + Z > t) = 0 \qquad \text{for all } t > x.$$

Indeed, with $\mu_1 := \mathcal{L}(X_1^x)$ monotonicity yields

$$\begin{aligned}\mathbf{E}^x(T_{[0,t]}) &\geq \int_{s>t} \mathbf{E}^s(1 + T_{[0,t]})\, \mu_1(ds) \\ &\geq \mathbf{P}(Yx + Z > t)\, \mathbf{E}^t(T_{[0,t]}),\end{aligned}$$

where $\mathbf{E}^t(T_{[0,t]}) = \infty$ by (9.5b). Letting $t \downarrow x$ in (2) leads to

$$\mathbf{P}(Yx + Z \leq x) = 1,$$

hence by (1.2b) to $\overline{x} \leq x < \infty$. This implies positive recurrence and thus a contradiction to the assumption. □

Together, (9.1) and (9.6) show that the two main characterizations of positive/null recurrence from classical Markov chain theory in essence carry over to affine recursions.

10. The contractive case

As outlined in Section 5 it is in general impossible to determine the invariant measure μ of a recurrent distribution ν explicitly. In spite of (8.2) this holds as well for the stationary distribution in the positive recurrent case (for an exception see (10.3)). This is compensated to some extent by the fact that, due to the algebraic form of the underlying recursion, the moments of μ can be easily computed from those of ν, as far as they exist. The following simple criterion for their existence extends a result by Vervaat [39]:

(10.1) Proposition. *Let* $(X_n)_{n\geq 0}$ *be stationary and* $0 < \alpha < \infty$. *Then*

$$\mathbf{E}(X_n^\alpha) < \infty$$

if and only if

$$\mathbf{E}(Y^\alpha) < 1 \qquad \textit{and} \qquad \mathbf{E}(Z^\alpha) < \infty.$$

Proof. With W as defined in (8.2) it follows from

$$\mathbf{E}(W^\alpha) = \sup_{n\in\mathbf{N}} \mathbf{E}((Z_1 + \ldots + Y_1 \ldots Y_{n-1} Z_n)^\alpha)$$

by elementary inequalities that

$$\mathbf{E}(W^\alpha) \left\{ \begin{matrix} \leq \\ \geq \end{matrix} \right\} \textstyle\sum_{n\in\mathbf{N}} (\mathbf{E}(Y^\alpha))^{n-1}\, \mathbf{E}(Z^\alpha) \qquad \text{for } \alpha \left\{ \begin{matrix} \leq \\ \geq \end{matrix} \right\} 1,$$

while on the other hand

$$W = \sup_{n\in\mathbf{N}} (Z_1 + \ldots + Y_1 \ldots Y_{n-1} Z_n)$$

by Minkowski's inequality implies

$$\|W\|_\alpha \left\{ \begin{matrix} \geq \\ \leq \end{matrix} \right\} \textstyle\sum_{n\in\mathbf{N}} \|Y\|_\alpha^{n-1} \, \|Z\|_\alpha \qquad \text{for } \alpha \left\{ \begin{matrix} \leq \\ \geq \end{matrix} \right\} 1.$$

This proves the assertion in view of $\mathbf{E}(Z^\alpha) \neq 0$ resp. $\|Z\|_\alpha \neq 0$. □

For $\alpha = 1$ the bounds in the proof coincide and yield the equation

$$\mathbf{E}(X_n) = \mathbf{E}(Z) \,/\, (1 - \mathbf{E}(Y)).$$

If finite, moments of any order $k \in \mathbf{N}$ can be computed from

$$\textstyle\int x^i \, d\mu = \int \int (xy + z)^i \, \mu(dx)\, \nu(dy, dz) \qquad \text{for } 1 \leq i \leq k,$$

a system of linear equations that can be solved recursively due to $\int y^i \, d\nu < 1$.

Two further comments are in order:

— The existence of moments of some order is clearly related to the tail–behaviour of μ. For relevant results see the papers by Kesten [24] and Goldie [16].

— (10.1) does not extend to the case $\alpha = \infty$. While $\|Y\|_\infty < 1$ and $\|Z\|_\infty < \infty$ obviously imply $\overline{x} < \infty$, conversely this condition yields only $\|Y\|_\infty \leq 1$ and $\|Z\|_\infty < \infty$ (consider for instance the case $N = \{(1,0), (0,1)\}$, where $\mu = \varepsilon_1$ according to (1.4)).

In view of (10.1) it is of interest, under which conditions μ has an exponential moment and is thus determined by its moments of natural order. Here contractivity enters:

(10.2) Proposition. *Let $(X_n)_{n\geq 0}$ be stationary. Then*

$$\mathbf{E}(e^{uX_n}) < \infty \quad \textit{for some } u > 0$$

if and only if

$$Y \leq 1 \qquad \textit{and} \qquad \mathbf{E}(e^{uZ}) < \infty \quad \textit{for some } u > 0.$$

Proof. 1. The conditions on Y and Z are necessary, because $\mathbf{E}(X_n^k) < \infty$ for all $k \in \mathbf{N}$ implies by (10.1)

$$\|Y\|_\infty = \lim_{k\to\infty} \|Y\|_k \leq 1,$$

while $\mathbf{E}(e^{uZ}) < \infty$ is immediate from $Z_1 \leq X_1$.

2. To prove sufficiency, consider first the special case $Y \leq \vartheta < 1$. With W as defined in (8.2) this yields the estimate

$$\begin{aligned} \mathbf{E}(e^{uW}) &\leq \mathbf{E}(\exp(u \textstyle\sum_{n\in\mathbf{N}} \vartheta^{n-1} Z_n)) \\ &= \textstyle\sum_{k\geq 0} \dfrac{u^k}{k!} \|\textstyle\sum_{n\in\mathbf{N}} \vartheta^{n-1} Z_n\|_k^k \\ &\leq \textstyle\sum_{k\geq 0} \dfrac{u^k}{k!} (\textstyle\sum_{n\in\mathbf{N}} \vartheta^{n-1} \|Z_n\|_k)^k \\ &= \textstyle\sum_{k\geq 0} \dfrac{1}{k!} (\dfrac{u}{1-\vartheta})^k \mathbf{E}(Z^k) \\ &= \mathbf{E}(e^{uZ/(1-\vartheta)}) \qquad \text{for } u \geq 0, \end{aligned}$$

i.e. if u is suited for Z, then $(1-\vartheta)u$ is suited for W.

3. To reduce the general case to this situation a partition into random blocks will be used, which is dual to that one used in the proof of (3.2). Applying $\mathbf{P}(Y = 1) < 1$ choose $\vartheta < 1$ satisfying $\mathbf{P}(Y > \vartheta) < 1$ and denote by $0 = T_0 < T_1 < \ldots$ the random times with $Y_n \leq \vartheta$ (being defined with probability 1). Now define

$$Y'_k := Y_{T_{k-1}+1} \ldots Y_{T_k},$$

$$Z'_k := Z_{T_{k-1}+1} + \ldots + Y_{T_{k-1}+1} \ldots Y_{T_k-1} Z_{T_k}.$$

Since $(T_k)_{k\geq 0}$ is a process with independent and identically distributed increments, the random variables (Y'_k, Z'_k), $k \in \mathbf{N}$, are independent with a distribution $\nu' \in \mathcal{N}$, where

(1) $\quad \sum_{k\in\mathbf{N}} Y'_1 \ldots Y'_{k-1} Z'_k = W.$

With $T := T_1$ they have in addition the properties

(2) $\quad Y' \leq \vartheta,$

$$\begin{aligned} \mathbf{E}(e^{uZ'}) &\leq \mathbf{E}(e^{u(Z_1+\ldots+Z_T)}) \\ &= \textstyle\sum_{n\in\mathbf{N}} \mathbf{E}(\prod_{1\leq m<n} 1_{\{Y_m>\vartheta\}} e^{uZ_m} 1_{\{Y_n\leq\vartheta\}} e^{uZ_n}) \\ &= \textstyle\sum_{n\in\mathbf{N}} (\psi_1(u))^{n-1} \psi_2(u), \end{aligned}$$

where

$$\psi_1(u) := \mathbf{E}(1_{\{Y>\vartheta\}} e^{uZ}) \to \mathbf{P}(Y > \vartheta) \qquad \text{for } u \to 0,$$

$$\psi_2(u) := \mathbf{E}(1_{\{Y\leq\vartheta\}} e^{uZ}) \leq \mathbf{E}(e^{uZ}) \qquad \text{for all } u \geq 0.$$

For $u_0 > 0$ with $\psi_1(u_0) < 1$ and $\psi_2(u_0) < \infty$ this yields

(3) $\quad \mathbf{E}(e^{u_0 Z'}) \leq \psi_2(u_0) \,/\, (1 - \psi_1(u_0)) < \infty.$

By (1) – (3) the reduction to the special case is settled. □

The simplest example for the situation of (10.2) is provided by the case $\overline{x} < \infty$. If in addition the support N is finite, the meanwhile classical field of self–similarity is entered. Because of the extensive literature on this subject, in particular in the context of fractals, only a question concerning uniform distributions will be considered here:

(10.3) Proposition. *Let ν be positive recurrent with $\underline{x} < \overline{x}$ and*

$$N = \{(y_i, z_i) : 0 \leq i \leq k\}$$

be such that $0 < y_i < 1$ and

$$]\underline{x}, \overline{x}] = \bigcup\nolimits_{0 \leq i \leq k} (y_i \,]\underline{x}, \overline{x}] + z_i)$$

is a partition. Then the stationary distribution μ satisfies

(a) $\quad \operatorname{supp} \mu = [\underline{x}, \overline{x}],$

(b) $\quad$ *μ is the uniform distribution on $[\underline{x}, \overline{x}]$ if and only if*

$$\nu(\{(y_i, z_i)\}) = y_i \qquad \text{for } 0 \leq i \leq k,$$

(c) $\quad$ *if ν' is any other distribution with support N, then its stationary distribution μ' is orthogonal to μ.*

Proof. (a) Since $M = \operatorname{supp} \mu$ and N are compact, by (4.1b)

$$M = \bigcup\nolimits_{0 \leq i \leq k} (y_i M + z_i).$$

By the hypothesis this equation holds as well, if M is replaced by $[\underline{x}, \overline{x}]$, and the assertion follows from well–known uniqueness results (see e.g. [20]).

(b) With $\gamma := \overline{x} - \underline{x}$ and $p_i := \nu(\{(y_i, z_i)\})$ the condition on μ translates via its distribution function into

$$\sum\nolimits_{0 \leq i \leq k} p_i \, \frac{1}{\gamma} \, \langle \, \underline{x} \, , \frac{t - z_i}{y_i} \, , \overline{x} \, \rangle = \frac{1}{\gamma} \, \langle \, \underline{x}, t, \overline{x} \, \rangle \qquad \text{for all } t \geq 0,$$

where $\langle a, b, c \rangle$ denotes the medium of $a, b, c \in \mathbf{R}$. The assertion follows from the partition properties.

(c) Consider the mapping

$$\tau : (y_{i_n}, z_{i_n})_{n \in \mathbf{N}} \to \sum\nolimits_{n \in \mathbf{N}} y_{i_1} \cdots y_{i_{n-1}} z_{i_n}$$

from $\prod_{n \in \mathbf{N}} N$ to $[\underline{x}, \overline{x}]$. Under the assumption $\underline{x} = z_0/(1 - y_0)$ it is easily seen that τ is bijective and bimeasurable, if it is restricted to sequences not terminating in the sense that $i_n = 0$ for almost all n and $\underline{x}$ is deleted from the range. Since the exceptional sequences form a null set with respect to the orthogonal measures $\bigotimes_{n\in\mathbf{N}}\nu$ and $\bigotimes_{n\in\mathbf{N}}\nu'$, orthogonality carries over to their images by τ, being μ and μ' by (8.2). $\square$

The simplest example for uniform distribution on [0,1] is provided by the additive model

$$Y = \frac{1}{k+1} \qquad \text{and} \qquad \mathbf{P}(Z = \frac{i}{k+1}) = \frac{1}{k+1} \qquad \text{for } 0 \le i \le k.$$

Since (10.3) extends easily to countable partitions, there are similar examples in the multiplicative model:

$$\mathbf{P}(Y = 2^{-k}) = 2^{-k} \text{ for } k \in \mathbf{N} \qquad \text{and} \qquad Z = 1$$

for instance yields the uniform distribution on [1,2].

It is obvious that the hypothesis in (10.3) implies $\sum_{0\le i\le k} y_i = 1$, and it is easily seen that in the case $\sum_{0\le i\le k} y_i < 1$ the support M is a Lebesgue null set and thus the measure μ is singular with respect to Lebesgue measure. More intricate is the case $\sum_{0\le i\le k} y_i > 1$, considered in detail by Garsia [15]. Even for

$$N = \{(0,1),(y,1)\} \qquad \text{with } \frac{1}{2} < y < 1$$

only partial results are known, going back essentially to Erdös [13].

The rest of this section is devoted to situations not excluding weak contractions, i.e. to the case $Y \le 1$. The first result is of interest mainly in the context of (10.6), though μ occurs as limit distribution also in learning theory (see e.g. [32]) or in a problem on random walks treated by Masimov [29]. The method from [13] yields the following:

(10.4) Proposition. *With $1 \ne l \in \mathbf{N}$ and $0 < p < 1$, $q = 1 - p$ suppose*

$$\mathbf{P}(Y = \frac{1}{l}) = p,\ \mathbf{P}(Y = 1) = q \qquad \textit{and} \qquad Z = 1.$$

Then ν is positive recurrent and the stationary distribution μ is singular with respect to Lebesgue measure.

Proof. Positive recurrence is immediate from (8.5a). The random times $0 = T_0 < T_1 < \ldots$ with $Y = 1/l$ (being defined with probability 1) have independent and geometrically distributed increments

$$U_k := T_k - T_{k-1} \qquad \text{for } k \in \mathbf{N}.$$

Since the random variable W from (8.2) has the representation

$$W = \textstyle\sum_{k\ge 0} l^{-k}\, U_{k+1},$$

the Fourier transform φ of μ is therefore given by

$$\varphi(u) = \textstyle\prod_{k\ge 0} p\, e^{i\,l^{-k}u} / (1 - q\, e^{i\,l^{-k}u}).$$

In view of

$$|\varphi(u)| = \prod_{k\geq 0} [1+\gamma\,(1-\cos\frac{u}{l^k})]^{-\frac{1}{2}} \qquad \text{with } \gamma := \frac{2q}{p^2}$$

it suffices to show

(1) $\liminf_{n\to\infty} \prod_{k\geq 0} [\ldots] < \infty,$

because then μ cannot be absolutely continuous with respect to Lebesgue measure by the Riemann–Lebesgue lemma and (6.2) applies. To verify (1) consider the values

$$u_m = 2\pi\, l^m \to \infty \quad \text{for} \quad m \to \infty,$$

which by shifting the index k to $k-m$ yield the constant value

$$\prod_{k\geq 0} [\ldots] = \prod_{k\in\mathbf{N}} (1+\gamma\,(1-\cos\frac{2\pi}{l^k})).$$

Since $1-\cos(2\pi/l^k)$ is bounded by $\frac{1}{2}(2\pi/l^k)^2$ for almost all k and thus is summable, the latter product converges and (1) is established. □

The final result, which originally motivated the present work, requires some preparation:

(10.5) Lemma. *With* $0\leq\gamma<1$ *and* $0<p<1$, $q=1-p$ *assume*

$$Y=\gamma \qquad \textit{and} \qquad \mathbf{P}(Z=k)=p\,q^{k-1} \quad \textit{for } k\in\mathbf{N}.$$

Then

$$\limsup_{n\to\infty} X_n/\log n = -1/\log q \quad \textit{a.s.}$$

Proof. 1. It is less cumbersome to prove instead

$$\limsup_{n\to\infty} X_n'/\log n = 1/\delta \quad \text{a.s.}$$

for a sequence $(X_n')_{n\geq 0}$ where $Y'=\gamma$, Z' has an exponential distribution with parameter δ and in addition $X_0'=0$. Then

$$Z = \sum_{k\in\mathbf{N}} k\, 1_{\{k-1\leq Z'<k\}}$$

has a geometric distribution with parameter $p=1-e^{-\delta}$, while

$$0\leq X_n - X_n' \leq X_0 + \sum_{m\geq 0}\gamma^m < \infty.$$

Moreover, replace Z' by $\delta Z'$ and thus X_n' by $\delta X_n'$ to see that $\delta=1$ means no loss of generality. Finally, the superscripts will be suppressed in the sequel.

2. To show that 1 is a lower bound for the upper limit, use

(1) $\mathbf{E}(e^{uZ}) = \dfrac{1}{1-u} \qquad \text{for } 0\leq u\leq 1$

to conclude that

$$\sum_{n>1} \mathbf{P}(Z_n / \log n > 1) = \sum_{n>1} \mathbf{P}(e^{Z_n} > n) = \infty,$$

which by Borel–Cantelli and $X_n \geq Z_n$ proves one half of the equation.

3. To prove the other half, choose an arbitrary $t > 1$. Then

$$\begin{aligned} \mathbf{P}(X_n / \log n > t) &= \mathbf{P}(Z_n > t \log n - (\gamma^{n-1} Z_1 + \ldots + \gamma^1 Z_{n-1})) \\ &\leq \mathbf{E}(\exp(\gamma^{n-1} Z_1 + \ldots + \gamma^1 Z_{n-1} - t \log n)) \end{aligned}$$

for $n > 1$, because $\mathbf{P}(Z > z) \leq e^{-z}$ holds for $z \leq 0$ as well and Fubini applies due to independence. By (1) therefore

$$\begin{aligned} \mathbf{P}(X_n / \log n > t) &\leq n^{-t} \prod_{1 \leq m < n} \frac{1}{1 - \gamma^m} \\ &\leq n^{-t} / \prod_{m \in \mathbf{N}} (1 - \gamma^m) \qquad \text{for } n > 1, \end{aligned}$$

where the infinite product is strictly positive by $\sum_{m \in \mathbf{N}} \gamma^m < \infty$. Thus

$$\sum_{n>1} \mathbf{P}(X_n / \log n > t) < \infty \qquad \text{for all } t > 1,$$

i.e. 1 is also an upper bound. □

In conclusion the order in which an upper limit $\overline{x} = \infty$ is approached will be studied for the weakly contractive multiplicative model, being of particular interest in applications. While the result provides just an upper bound in the case $\mathbf{P}(Y = 1) = 0$, it is exact otherwise and, somewhat surprisingly, depends only on this probability:

(10.6) Theorem. *If $Y \leq 1$ and $Z = 1$, then*

$$\limsup_{n \to \infty} X_n / \log n = -1 / \log \mathbf{P}(Y = 1) \quad a.s.$$

Proof. 1. Clearly, $X_0 = 1$ can be supposed in the sequel. In proving the right–hand side to be a lower bound for the upper limit, moreover, Y may be decreased to $1_{\{Y=1\}} Y$. Therefore in this part Y will be assumed to be $0, 1$–valued, where

$$p := \mathbf{P}(Y = 0) > 0.$$

Then the random times $0 = T_0 < T_1 < \ldots$ with $Y_n = 0$ are defined with probability 1 and have independent and geometrically distributed increments

$$U_k := T_k - T_{k-1} \qquad \text{for } k \in \mathbf{N}.$$

Thus by the strong law of large numbers

$$\frac{1}{k} T_k = \frac{1}{k} (U_1 + \ldots + U_k) \to \mathbf{E}(T_1) = \frac{1}{p} \text{ a.s.},$$

which implies

(1) $\log T_k / \log k \to 1$ a.s.

Therefore the equation

$$X_{T_k-1} = U_k \qquad \text{for } k \in \mathbf{N}$$

yields the estimate

$$\begin{aligned} \limsup_{n\to\infty} X_n / \log n &\geq \limsup_{k\to\infty} X_{T_k-1} / \log(T_k - 1) \\ &= \limsup_{k\to\infty} U_k / \log k. \end{aligned}$$

Now (10.5) applies with $\gamma = 0$, in which case X_n and Z_n agree.

2. In proving the inverse inequality Y may be increased to γ on $\{Y \leq \gamma\}$ and to 1 on $\{Y > \gamma\}$ for any $\gamma < 1$ satisfying $\mathbf{P}(Y \leq \gamma) > 0$, because a subsequent limiting procedure $\gamma \to 1$ yields the desired upper bound. Therefore restrict Y to the values γ and 1 and, with 0 replaced by γ, introduce the random variables T_k and U_k as in the first part of the proof. Now for $T_{k-1} \leq n < T_k$ clearly

$$X_n / \log n \leq X_{T_k-1} / \log T_{k-1},$$

where by (1)

$$\log T_{k-1} / \log(T_k - 1) \to 1 \text{ a.s.}$$

Again by (1) this implies

$$\begin{aligned} \limsup_{n\to\infty} X_n / \log n &\leq \limsup_{k\to\infty} X_{T_k-1} / \log(T_k - 1) \\ &= \limsup_{k\to\infty} X_{T_k-1} / \log k, \end{aligned}$$

where

$$X_{T_k-1} = \sum_{1\leq i\leq k} \gamma^{k-i} U_i \qquad \text{for } k \in \mathbf{N}.$$

Thus (10.5) applies again, with Z_n replaced by U_n. □

Finally, it should be mentioned that with probability 1 the limit points of the normalized sequence $(X_n / \log n)_{n>1}$ exhaust the interval

$$I := [0, -1/\log \mathbf{P}(Y = 1)].$$

Indeed, given $l \in \mathbf{N}$ and $\varepsilon > 0$, by $\underline{x} < \infty$ there exists a random time $T \geq l \vee e^{1/\varepsilon}$ such that $X_T / \log T < \varepsilon$, hence in view of

$$X_{n+1} / \log(n+1) \leq (X_n + 1) / \log n \leq X_n / \log n + \varepsilon \qquad \text{for } n \geq T$$

the values $X_n / \log n$, $n \geq l$, are ε–dense in I.

References

1. Alpuim, M.: An extremal Markovian sequence. J. Appl. Prob. **26**, 219–232 (1989)
3. Barnsley, M., Elton, J., Hardin, D.: Recurrent iterated function systems. Constr. Approx. **5**, 3–31 (1989)
4. Borovkov, A.: On the ergodicity and stability of the sequence $w_{n+1} = f(w_n, \xi_n)$. Theory Prob. Appl. **33**, 595–611 (1989)
5. Brandt, A.: The stochastic equation $Y_{n+1} = A_n Y_n + B_n$ with stationary coefficients. Adv. Appl. Prob. **18**, 211–220 (1986)
6. Brandt, A., Franken, P., Lisek, B.: Stationary stochastic models. Chichester: Wiley 1990
7. Chamayou, J., Letac, G.: Explicit stationary distributions for compositions of random functions and products of random matrices. J. Theor. Prob. **4**, 3–36 (1991)
10. Elton, J.: An ergodic theorem for iterated maps. Ergodic Theory Dyn. Syst. **7**, 481–488 (1987)
11. Elton, J.: A multiplicative ergodic theorem for Lipschitz maps. Stochastic Processes Appl. **34**, 39–47 (1990)
12. Elton, J., Yan, Z.: Approximation of measures by Markov processes and homogeneous affine iterated function systems. Constr. Approx. **5**, 69–87 (1989)
13. Erdös, P.: On a family of symmetric Bernoulli convolutions. Am. J. Math. **61**, 974–976 (1939)
15. Garsia, A.: Entropy and singularity of infinite convolutions. Pac. J. Math. **13**, 1159–1169 (1963)
16. Goldie, C.: Implicit renewal theory and tails of solutions of random equations. Ann. Appl. Prob. **1**, 126–166 (1991)
18. Grincevičius, A.: A random difference equation. Lith. Math. Trans. **21**, 302–306 (1982)
19. de Haan, L., Karandikar, R.: Embedding a stochastic difference equation into a continuous–time process. Stochastic Processes Appl. **32**, 225–235 (1989)
20. Hutchinson, J.: Fractals and self similarity. Indiana Univ. Math. J. **30**, 713–747 (1981)
21. Kac, M.: On the notion of recurrence in discrete stochastic processes. Bull. Am. Math. Soc. **53**, 1002–1010 (1947)
23. Kesten, H.: The limit points of a normalized random walk. Ann. Math. Stat. **41**, 1173–1205 (1970)
24. Kesten, H.: Random difference equations and renewal theory for products of random matrices. Acta Math. **131**, 207–248 (1973)
27. Lev, G.: Semi–Markov processes of multiplication with drift. Theory Prob. Appl. **17**, 159–164 (1972)
29. Masimov, V.: A generalized Bernoulli scheme and its limit distribution. Theory Prob. Appl. **18**, 521–530 (1973)
30. Mukherjea, A.: Recurrent random walk in nonnegative matrices: attractors of certain iterated function systems. Prob. Theory Related Fields **91**, 297–306 (1992)

31. Mukherjea, A., Tserpes, N.: Measures on topological semigroups: convolution products and random walks. Lect. Notes Math. **547**, Berlin–Heidelberg–New York: Springer 1976
32. Norman, M.: Limiting distributions for some random walks arising in learning models. Ann. Math. Stat. **37**, 393–405 (1966)
34. Rachev, S., Samorodnitsky, G.: Limit laws for a stochastic process and random recursion arising in probability modelling. Preprint (1992)
39. Vervaat, W.: On a stochastic difference equation and a representation of non–negative infinitely divisible random variables. Adv. Appl. Prob. **11**, 750–783 (1979)
40. Wolfe, S.: On a continuous analogue of the stochastic difference equation $X_n = \varrho X_{n-1} + B_n$. Stochastic Processes Appl. **12**, 301–312 (1982)