Ergodic Behaviour of Affine Recursions I

Criteria for Recurrence and Transience

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 $\mathbf{R}_+$. This results in quite natural notions of (topological) recurrence and transience and allows for rather explicit criteria to distinguish between both cases.

1991 Mathematics Subject Classification. Primary 60H25, 60J05; Secondary 54H20.

Introduction

A discrete–time Markov process $(X_n)_{n\geq 0}$ with a decent (say, Polish) state space E can be given the form of a "stochastically recursive sequence" (in the sense of Borovkov [4])

$$X_n = g_n(X_{n-1}) \qquad \text{for } n \in \mathbf{N},$$

where g_n, $n \in \mathbf{N}$, is a sequence of independent and identically distributed random mappings from E to E which is independent of the initial variable X_0.

A first systematic treatment under this aspect goes back to Dubins and Freedman [9], who, for $E = [0,1]$, considered in particular two cases:
(C) regarding the metric structure of E, they restricted the mappings g_n to the class of contractions;
(A) specializing further, they limited these mappings to the class of affinities.

Generalizations of (C) from [0,1] to a complete metric space E attracted much interest during the last decade, mainly in the context of fractals. Under the heading "iterated function systems" Barnsley, Elton and others considered questions concerning stationary distributions, weak convergence, recurrence, ergodicity etc. Crucial for their results is an "average contractivity" (see [2], [3], [10], [11], [12]).

Extensions of (A) from the compact interval [0,1] to $\mathbf{R}$ occur first in papers by Lev [27] and Masimov [29]. More comprehensive are independent treatments by Grincevičius [17] and Vervaat [39]. Explicitly, affine recursions on $\mathbf{R}$ have the form of a stochastic difference equation

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

(which is somewhat more general than $X_n = Y_n\,(X_{n-1} + Z_n)$). Here, X_0 is a real random variable, independent of a sequence $(Y_n, Z_n)_{n\in\mathbf{N}}$ of independent identically distributed $\mathbf{R}^2$-valued random variables.

These affine recursions include as degenerate cases random walks ($Y_n = 1$) and infinite products ($Z_n = 0$). Of particular interest are the "additive model", where

$$X_n = yX_{n-1} + Z_n \qquad \text{with constant } y \in \mathbf{R}$$

(the simplest case of an autoregressive process), and the "multiplicative model", where

$$X_n = Y_n X_{n-1} + z \qquad \text{with constant } z \in \mathbf{R}.$$

Both models can be subsumed under the case where Y_n and Z_n are independent.

Due to the particular role of 0 with respect to multiplication, the situation is especially simple for $\mathbf{P}(Y = 0) > 0$. This "regenerative" case can be studied in a more general context (see e.g. Nummelin [33]). Another particular situation is provided by the "contractive" case, where in the weak version $|Y_n| \leq 1$ and in the strong version $|Y_n| \leq \vartheta < 1$.

To conclude the historical remarks, two central results concerning the general case have to be mentioned:

— If an affine recursion has a stationary distribution, it is unique and the laws $\mathcal{L}(X_n)$ converge weakly to it, independently of the initial law $\mathcal{L}(X_0)$.

— Conditions on the existence of a stationary distribution can be formulated via logarithmic moments of Y_n and Z_n.

The present work originates in the observation that most applications in economics, biology, physics etc. (see the long list of references in [39]) in fact work in the state space $\mathbf{R}_+$, and thus the case $X_0 \geq 0$ and $Y_n, Z_n \geq 0$ is of special importance. From the mathematical point of view this restriction is supported by the fact that a state space $\mathbf{R}_+$ allows only one kind of divergence to infinity and the assumption $Y_n, Z_n \geq 0$ entails additional monotonicity properties of the associated transition kernel.

In the existing literature these aspects do not find much attention. Apart from an approach by Lamperti [25], [26], too general for affine recursions, there are only two exceptions: a recent paper by Mukherjea [30], coupling in the context of nonnegative matrices the sequence $(X_n)_{n\geq 0}$ with the partial products of $(Y_n)_{n\in\mathbf{N}}$, and a preprint by Rachev [34], concentrating on central limit theorems for suitably normalized variables X_n in the divergent case.

To summarize the main feature of the present work before going into details: affine recursions on $\mathbf{R}_+$ seem to provide one of the best suited models for extending classical Markov chain theory to an uncountable state space. Since the Harris theory (see e.g. [35]) is easily seen not to be adequate, the study has to be based on the topological structure. This, in general, leads to various notions of irreducibility and aperiodicity, of (positive or null) recurrence and transience (see the papers by Rosenblatt [37] and Tweedie [38]). In the present case, however, these notions merge into very natural definitions satisfying the classical criteria.

This allows for a rather complete theory, developed in the sections:

1. Lower and upper limit
2. Recurrence and transience
3. Recurrence criteria
4. Excessive and invariant measures
5. Existence and uniqueness of invariant measures
6. Main properties of the invariant measure
7. Ratio ergodic theorems
8. Positive and null recurrence
9. Further ergodic theorems
10. The contractive case

Thus the paper divides into three parts: Sections 1–3 classify affine recursions on $\mathbf{R}_+$ according to recurrence and transience, Sections 4–7 treat existence and uniqueness of invariant measures as well as ergodic theorems, Sections 8–10 continue to classify the recurrent case by introducing the notions of positive recurrence and null recurrence. Part II and III will appear in ... and ... ; the contents of Part I are summarized below.

Section 1. Since convergence of an affine recursion, even in probability, occurs only in an exceptional case (1.4), the lower and upper limit are of interest. It is a crucial consequence of restricting the state space to $\mathbf{R}_+$ that these limits are constants $\underline{x}$ and $\overline{x}$, independent of the initial law $\mathcal{L}(X_0)$ (1.1). Wherever starting, the sequence $(X_n)_{n\geq 0}$ approaches the interval $[\underline{x}, \overline{x}]$ monotonically (1.2). While the upper limit only depends on the support of the joint law $\mathcal{L}(Y_n, Z_n)$, this in general fails for the lower limit. In fact, it is not surprising that even in the simple example

$$\text{(E)} \qquad X_n = Y_n X_{n-1} + 1 \qquad \text{with} \qquad Y_n = 2^{-1} \quad \text{or} \quad 2^{+1}$$

the asymptotic behaviour depends essentially on the probabilities

$$p_- = \mathbf{P}(Y_n = 2^{-1}) \qquad \text{and} \qquad p_+ = \mathbf{P}(Y_n = 2^{+1}).$$

Provided, however, $\underline{x}$ and $\overline{x}$ are finite, a simple characterization by means of the support is available (1.3).

Section 2. Clearly, $(X_n)_{n\geq 0}$ has to be called "transient" in the case $\underline{x} = \infty$. It is less clear – and one of the central questions in the sequel –, whether the sequence may be called "recurrent" in the case $\underline{x} < \infty$. A first justification is supplied by an equivalent characterization through the associated potential kernel: the mean time spent in a bounded interval $[0, t]$ is finite in the transient case and infinite in the recurrent case, provided $t > \underline{x}$ (2.2). Examples for both situations are easily established: the affine recursion is certainly recurrent in the regenerative case $\mathbf{P}(Y_n = 0) > 0$ (2.3), and it is transient whenever the associated random walk $(S_n)_{n\geq 0}$ with increments $\log Y_n$ diverges to $+\infty$ (2.4). Though both conditions depend on the "primary" variable Y_n only, the "secondary" variable Z_n may be essential as well. In fact, however small Y_n is, a sufficiently large Z_n yields transience (2.5).

Section 3. In the additive model a nearly complete characterization of recurrence or transience by the asymptotic behaviour of $t\,\mathbf{P}(\log Z_n > t)$ for $t \to \infty$ can be derived, which extends to the case where Y_n is bounded away from 1 or 0 (3.1). The situation is less clear in the multiplicative model. It is not surprising that in example (E) above $p_- < p_+$ implies transience and $p_- > p_+$ implies recurrence. It is a nontrivial problem, however, to decide the balanced case $p_- = p_+$. Clearly the related multiplicative random walk, where

$$X_0 = 1 \qquad \text{and} \qquad X_n = Y_n X_{n-1} \ \text{ for } \ n \in \mathbf{N},$$

oscillates through the values $2^k, k \in \mathbf{Z}$. But it requires Spitzer's combinatorial identity to prove the drift term $+1$ not to change recurrence into transience. The relevant recurrence criterion settles the multiplicative model completely and extends in fact to the case where Z_n is only bounded (3.3). Assuming

Y_n and Z_n to be independent, even finiteness of the expectation of Z_n ensures recurrence (3.4).

0. Preliminaries

Throughout the paper $(X_n)_{n\geq 0}$ is a fixed affine recursion on $\mathbf{R}_+$, given by the stochastic difference equation $(*)$ of the introduction. Thus the distribution of $(X_n)_{n\geq 0}$ is completely determined by the laws $\mu_0 = \mathcal{L}(X_0)$ and $\nu = \mathcal{L}(Y, Z)$. Here, Y and Z is briefly written instead of Y_n and Z_n, as will be done whenever $n \in \mathbf{N}$ plays no role.

The initial law μ_0, as usual, is largely of only secondary significance. If in particular μ_0 is a unit measure ε_x, this will be expressed by the notation $(X_n^x)_{n\geq 0}$, i.e.

$$X_n^x := xY_1Y_2\ldots Y_n + Z_1Y_2\ldots Y_n + \ldots + Z_n \qquad \text{for } x \in \mathbf{R}_+ \text{ and } n \geq 0.$$

Thus conditional probability and expectation are simply given by

$$\mathbf{P}^x((X_n, n \geq 0) \in B) = \mathbf{P}((X_n^x, n \geq 0) \in B),$$

$$\mathbf{E}^x(g(X_n, n \geq 0)) = \mathbf{E}(g(X_n^x, n \geq 0)).$$

Roughly speaking, what follows is a theory of distributions ν on $\mathbf{R}_+^2$. Here an essential role will be played by their support, for which the notation N will be fixed. Since nothing new can be expected in the special cases $Y = 1$ resp. $Z = 0$, for simplification

$$N \cap \{(y, z) : y \neq 1\} \neq \emptyset \neq N \cap \{(y, z) : z \neq 0\}$$

is always assumed. The symbol $\mathcal{N}$ will throughout refer to the class of distributions ν on $\mathbf{R}_+^2$ that are admissible in this sense.

As is clear from the introduction, the ergodic behaviour of $(X_n)_{n\geq 0}$ is intimately related to the random walk

$$S_n := \textstyle\sum_{1\leq m\leq n} \log Y_m \qquad \text{for } n \geq 0,$$

generalized in the sense that it may attain the (absorbing) value $-\infty$. Due to $\mathbf{P}(Y = 1) < 1$ there are only three possibilities for the asymptotic behaviour of this random walk:

(1) $\quad S_n \to +\infty$ a.s.,

(2) $\quad S_n \to \pm\infty$ a.s.,

(3) $\quad S_n \to -\infty$ a.s.,

where the symbol in (2) serves as a short notation for

$$\mathbf{P}(\liminf_{n\to\infty} S_n = -\infty, \limsup_{n\to\infty} S_n = +\infty) = 1.$$

Another process closely related to $(X_n)_{n\geq 0}$ arises if, in the notation at the beginning of the introduction, the random variables $g_n \circ \ldots \circ g_1(X_0)$ are replaced by $g_1 \circ \ldots \circ g_n(X_0)$. Especially for $X_0 = 0$ this yields

$$W_n := Z_1 + Y_1 Z_2 + \ldots + Y_1 \ldots Y_{n-1} Z_n \qquad \text{for } n \geq 0.$$

The sequence $(W_n)_{n\geq 0}$ is no longer a Markov process, however, due to the exchangeability of $(Y_1, Z_1), \ldots, (Y_n, Z_n)$, satisfies

$$\mathcal{L}(W_n) = \mathcal{L}(X_n^0) \qquad \text{for } n \geq 0,$$

an equation, which will be important later on.

The transition kernel of the Markov process $(X_n)_{n\geq 0}$ will always be denoted by$P$. Thus the kernel P transforms a nonnegative function f into the function Pf given by

$$(Pf)(x) = \int f(yx + z)\, \nu(dy, dz)$$

and a measure μ on the Borel σ–algebra $\mathcal{B}(\mathbf{R}_+)$ into the measure μP given by

$$(\mu P)(B) = \int \nu(\{(y, z) : yx + z \in B\})\, \mu(dx).$$

If μ is σ–finite, the last equation amounts to

$$(\mu P)(B) = (\mu \otimes \nu)(\{(x; y, z) : yx + z \in B\}).$$

In accordance with the notations Pf and μP the μ–integral of a function f sometimes is simply denoted by μf.

The kernel P enjoys two important properties. First it is clearly a Feller kernel, transforming bounded continuous functions into the same type. Moreover, due to $Y \geq 0$, it is monotone in the sense that it transforms bounded increasing functions into the same type, too.

If E is a locally compact space with a countable base, the following concepts from topological measure theory will be used:

— $\mathcal{C}(E)$ denotes the space of bounded continuous functions $f : E \to \mathbf{R}$ and $\mathcal{K}(E)$ the subspace consisting of functions $f \in \mathcal{C}(E)$ with compact support.

— If μ is any measure on $\mathcal{B}(E)$, then $\mathcal{C}_\mu(E)$ denotes the space of bounded Borel–measurable functions $f : E \to \mathbf{R}$ that are μ–almost everywhere continuous and $\mathcal{K}_\mu(E)$ the corresponding subspace.

— The class $\mathcal{M}(E)$ of locally finite measures on E is endowed with the vague (weak*) topology, i.e. the initial topology with respect to the mappings

$$\mu \to \mu f,\ f \in \mathcal{K}(E).$$

In this topology convergence will be denoted by $\xrightarrow{\text{v}}$.

— The class $\mathcal{M}_1(E)$ of probability measures on E is endowed with the weak (narrow) topology, i.e. the initial topology with respect to the mappings

$$\mu \to \mu f,\ f \in \mathcal{C}(E).$$

In this topology convergence will be denoted by $\xrightarrow{\text{w}}$.

Finally it has to be emphasized that statements concerning random variables in case of doubt are always understood modulo $\mathbf{P}$–null sets. Thus the supplement "almost surely", as in the trichotomy concerning $(S_n)_{n\geq 0}$, will frequently be deleted.

1. Lower and upper limit

As outlined in the introduction, the following observation is of central importance throughout the paper:

(1.1) Theorem. *The random variables*

$$\underline{X} := \liminf_{n\to\infty} X_n \qquad \textit{and} \qquad \overline{X} := \limsup_{n\to\infty} X_n$$

are almost surely constant and independent of the initial value X_0, i.e. there exist constants $0 \leq \underline{x} \leq \overline{x} \leq \infty$, depending only on ν, such that

$$\mathbf{P}(\underline{X} = \underline{x}) = 1 \qquad \textit{and} \qquad \mathbf{P}(\overline{X} = \overline{x}) = 1.$$

Proof. 1. The following fact will be needed:

$$\text{(1)} \qquad \sup_{n\geq 0} S_n = +\infty \quad \text{implies} \quad \sup_{n\geq 0} X_n = \infty \quad \text{a.s.}$$

Indeed, since

$$\sup_{m\in\mathbf{N}} Z_m > 0 \text{ a.s.} \quad \text{and} \quad \sup_{n\geq m} e^{S_n - S_m} = \infty \text{ a.s.},$$

the assertion is a consequence of

$$\sup_{n\geq 0} X_n \geq \sup_{m\in\mathbf{N}} (Z_m \sup_{n\geq m} e^{S_n - S_m}).$$

2. Let $\underline{X}^x$ and $\overline{X}^x$ be defined in analogy to X_n^x. Then

$$\underline{X}^0 \geq \liminf_{n\to\infty} (Z_m Y_{m+1} \ldots Y_n + \ldots + Z_{m+n}) =: \underline{X}_m^0,$$

where $\mathcal{L}(\underline{X}^0)$ and $\mathcal{L}(\underline{X}_m^0)$ are identical. This implies

$$\underline{X}^0 = \underline{X}_m^0 \text{ a.s.} \qquad \text{for all } m \in \mathbf{N},$$

i.e. $\underline{X}^0$ is measurable with respect to (the completion of) the tail–field of $(Y_n, Z_n)_{n\in\mathbf{N}}$. Thus $\underline{X}^0$ – and similarly $\overline{X}^0$ – are almost surely constant. Therefore it suffices to show that $\underline{X}^x$ and $\overline{X}^x$ are in fact independent of $x \in \mathbf{R}_+$. Here $\sup_{n\geq 0} S_n = +\infty$ may be assumed, because otherwise

$$X_n^x - X_n^0 = x\, e^{S_n} \to 0 \text{ a.s.}$$

3. To treat the lower limit first, consider the random time

$$T := \inf\{n \geq 0 : X_n^0 \geq x\},$$

which is finite almost surely by (1). Then

$$\begin{aligned}\underline{X}^x &\geq \underline{X}^0 \\ &\geq \liminf_{n\to\infty} (X_T^0\, Y_{T+1} \ldots Y_{T+n} + \ldots + Z_{T+n}) \\ &\geq \liminf_{n\to\infty} (x\, Y_{T+1} \ldots Y_{T+n} + \ldots + Z_{T+n}) \\ &=: \underline{X}_T^x,\end{aligned}$$

where $\mathcal{L}(\underline{X}^x)$ and $\mathcal{L}(\underline{X}_T^x)$ are identical, because T is a stopping time with respect to $(Y_n, Z_n)_{n\in\mathbf{N}}$. Thus, indeed

$$\underline{X}^x = \underline{X}^0 \text{ a.s.} \qquad \text{for all } x \in \mathbf{R}_+.$$

The corresponding result for the upper limit is another consequence of (1), which yields

$$\overline{X}^x \geq \overline{X}^0 = \infty \text{ a.s.} \qquad \text{for all } x \in \mathbf{R}_+. \quad \square$$

The notation $\underline{x}$ and $\overline{x}$ will be used in the sequel without further reference. The interval defined by these limits attracts the sequence $(X_n)_{n\geq 0}$ in a strong sense:

(1.2) Proposition. *Whenever finite, the constants* $\underline{x}$ *and* $\overline{x}$ *are determined by the equivalences*

(a) $\quad x \leq \underline{x} \quad$ *if and only if* $\quad \mathbf{P}(Yx + Z \geq x) = 1,$

(b) $\quad x \geq \overline{x} \quad$ *if and only if* $\quad \mathbf{P}(Yx + Z \leq x) = 1;$

moreover the "if–part" holds without the finiteness assumption.

Proof. 1. The implication from right to left is an immediate consequence of (1.1), because e.g. the assumption

$$Yx + Z \geq x \text{ a.s.}$$

obviously implies

$$\{X_{n-1} \geq x\} \subset \{X_n \geq x\} \text{ a.s.},$$

which in turn yields

$$\underline{x} = \liminf_{n\to\infty} X_n^x \geq \inf_{n\geq 0} X_n^x \geq x.$$

2. To prove the converse, consider first assertion (a). Fix an arbitrary $x' \in\,]\underline{x}, \infty[$ and let $T_1 < T_2 \ldots$ denote the hitting times of $[0, x']$ by the sequence $(X_n^0)_{n\geq 0}$ (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

$$(Y_k', Z_k') := (Y_{T_k+1}, Z_{T_k+1}) \qquad \text{for } k \in \mathbf{N}$$

are again independent and distributed according to ν. Therefore

$$\begin{aligned}
\underline{x} &\leq \liminf_{k\to\infty} X^0_{T_k+1} \\
&= \liminf_{k\to\infty} (Y_{T_k+1} X^0_{T_k} + Z_{T_k+1}) \\
&\leq \liminf_{k\to\infty} (Y'_k x' + Z'_k) \\
&\leq yx' + z \qquad \text{for } (y,z) \in N,
\end{aligned}$$

because $(Y'_k, Z'_k)_{k\in\mathbf{N}}$ visits each neighborhood of (y,z) infinitely often with probability 1. Letting x' tend to $\underline{x}$ leads to

(1) $\qquad \underline{x} \leq y\underline{x} + z \qquad \text{for } (y,z) \in N.$

This inequality holds as well with 0 replacing $\underline{x}$, hence

(2) $\qquad x \leq yx + z \qquad \text{for } (y,z) \in N \text{ and } x \leq \underline{x},$

which is equivalent to the assertion

$$x \leq Yx + Z \text{ a.s.} \qquad \text{for } x \leq \underline{x}.$$

3. The corresponding proof for $\overline{x}$ requires only minor changes: choose now $x' \in [0, \overline{x}[$, which is possible due to

(3) $\qquad \overline{x} \geq \limsup_{n\to\infty} Z_n > 0,$

replace the interval $[0, x']$ by $[x', \infty[$, and use $\overline{x} > 0$ once more for the passage between (the counterparts of) (1) and (2). $\square$

From (1.2) it is easily deduced that for finite values $\underline{x}$ and $\overline{x}$

$$X_{n-1} \wedge \underline{x} \leq X_n \leq X_{n-1} \vee \overline{x} \text{ a.s.} \qquad \text{for all } n \in \mathbf{N}.$$

Now the constants $\underline{x}$ and $\overline{x}$ – except for a crucial ambiguity – can be evaluated explicitly, using a decomposition of the support N by its contractive and expansive part. With an asymmetry due to the different roles of 0 and ∞, the following holds:

(1.3) Proposition. *With the notations*

$$N_c := \{(y,z) \in N : y < 1\},$$
$$N_e := \{(y,z) \in N : y \geq 1 \text{ and } (y,z) \neq (1,0)\}$$

the constants $\underline{x}$ and $\overline{x}$ satisfy

(a) $\quad \underline{x} = \inf\{\frac{z}{1-y} : (y,z) \in N_c\}$

or $\quad \underline{x} = \infty$,

(b) $\quad \overline{x} = \sup\{\frac{z}{1-y} : (y,z) \in N_c\}$ $\quad$ *if* $N_e = \emptyset$,

and $\quad \overline{x} = \infty$ $\quad$ *otherwise.*

Proof. (a) If $\underline{x}$ is finite, according to (1.2a) it is the largest value x such that

$$x \le yx + z \qquad \text{for all } (y,z) \in N.$$

Since this inequality is trivially satisfied for $y \ge 1$, it amounts indeed to

$$x \le \frac{z}{1-y} \qquad \text{for } (y,z) \in N_c.$$

(b) If $\overline{x}$ is finite, according to (1.2b) it satisfies

$$\overline{x} \ge \overline{x}y + z \qquad \text{for all } (y,z) \in N.$$

In view of $\overline{x} > 0$ (see (3) in the proof of (1.2)) this yields $N_e = \emptyset$. Since $(y,z) = (1,0)$ can be disregarded, $\overline{x}$ is the smallest value x such that

$$x \ge xy + z \qquad \text{for } (y,z) \in N_c. \quad \square$$

It is a trivial consequence of this result that $X_n \to \infty$ a.s. whenever N_c is empty. Otherwise $\underline{x}$ – unlike $\overline{x}$ – may depend on the distribution ν not only through its support N. If, however, $\underline{x}$ resp. $\overline{x}$ is finite, it can be obtained graphically: $(0,\underline{x})$ resp. $(0,\overline{x})$ is that point where the lower resp. upper tangent from (1,0) to N_c intersects the z–axis. This implies in particular that the infimum or supremum in (1.3) need not be attained.

In contrast to the possibility $X_n \to \infty$ a.s., convergence of $(X_n)_{n\ge0}$ within $\mathbf{R}_+$ can occur only in a degenerate case, even if weakened to convergence in probability:

(1.4) Proposition. *The following assertions are equivalent:*

(a) $\quad \underline{x} = \gamma = \overline{x}$ *with* $0 < \gamma < \infty$,

(b) $\quad$ *the sequence* $(X_n)_{n\ge0}$ *converges in probability to a finite–valued random variable* X,

(c) $\quad Y \le 1$ $\quad$ *and* $\quad Z = \gamma(1-Y)$ *with* $0 < \gamma < \infty$.

Proof. 1. The implication (a) $\Rightarrow$ (b) is trivial.

2. Assume now (b) and let d be a bounded metric on $\mathbf{R}_+$. Then $(yX_n+z)_{n\ge0}$ converges in probability to $yX+z$, hence

(1) $\quad \mathbf{E}\,(d\,(yX_n+z, yX+z)) \to 0 \qquad$ for all $(y,z) \in \mathbf{R}_+^2$.

Since X_n and (Y_{n+1}, Z_{n+1}) are independent, moreover

$$\textstyle\int \mathbf{E}\,(d\,(yX_n+z, X_n))\,d\nu = \mathbf{E}\,(d\,(X_{n+1}, X_n)) \to 0.$$

Thus there is a subsequence $(X_{n_k})_{k\geq 0}$ such that

(2) $\quad \mathbf{E}\left(d\left(yX_{n_k}+z, X_{n_k}\right)\right) \to 0 \qquad$ for ν–almost all $(y,z) \in \mathbf{R}_+^2$.

When combined with the restriction of (1) to this subsequence, applied to (y,z) as well as to (1,0), (2) yields

$$\mathbf{E}\left(d\left(yX+z, X\right)\right) = 0 \qquad \text{for } \nu\text{–almost all } (y,z) \in \mathbf{R}_+^2.$$

With $\mu := \mathcal{L}(X)$ this implies

$$yx+z \stackrel{\mu}{=} x \qquad \text{for } \nu\text{–almost all } (y,z) \in \mathbf{R}_+^2;$$

therefore, by Fubini,

$$yx+z \stackrel{\nu}{=} x \qquad \text{for } \mu\text{–almost all } x \in \mathbf{R}_+.$$

Thus there exists indeed $\gamma \in \mathbf{R}_+$ such that

$$Z = \gamma(1-Y),$$

where $\gamma \neq 0$ (since otherwise $Z=0$) and $Y \leq 1$ (because of $Z \geq 0$).

3. The implication (c) $\Rightarrow$ (a) is immediate from (1.2). $\square$

Clearly, the constant γ is a common fixed point of the underlying affine maps almost surely.

2. Recurrence and transience

In view of $X_n \geq 0$ the first classification is very natural:

(2.1) Definition. *The distribution ν (or the kernel P or the process $(X_n)_{n\geq 0}$) is called*

(a) *"recurrent" if* $\underline{x} < \infty$,

(b) *"transient" if* $\underline{x} = \infty$.

Both cases can be distinguished as well by the associated potential kernel $G := \sum_{n\geq 0} P^n$, independently of the initial law:

(2.2) Theorem. *The following dichotomy holds:*

(a) *if ν is recurrent, then*

$$\sum_{n\geq 0} \mathbf{P}(X_n \leq t) = \infty \qquad \textit{for } t > \underline{x};$$

(b) *if ν is transient, then*

$$\sum_{n\geq 0} \mathbf{P}(X_n \leq t) < \infty \qquad \textit{for } t < \infty.$$

Proof. (a) This is an immediate consequence of

$$\mathbf{P}(X_n \leq t \text{ infinitely often}) = 1.$$

(b) Define recursively

$$T_0 := 0 \quad \text{and} \quad T_k := \inf\{n > T_{k-1} : X_n \le t\} \ \ (\le \infty).$$

Then in particular

$$0 = \mathbf{P}^0(X_n \le t \ \text{ infinitely often}) = \lim_{k\to\infty} \mathbf{P}^0(T_k < \infty),$$

hence there exists $l \in \mathbf{N}$ such that

$$\vartheta := \mathbf{P}^0(T_l < \infty) < 1.$$

With the decreasing function

$$g(x) := \mathbf{P}^x(T_l < \infty)$$

this implies

$$\begin{aligned} \mathbf{P}^0(T_{(k+1)\,l} < \infty) &= \int_{\{T_{kl}<\infty\}} g(X_{T_{kl}})\, d\mathbf{P}^0 \\ &\le \int_{\{T_{kl}<\infty\}} g(0)\, d\mathbf{P}^0 \\ &= \vartheta\, \mathbf{P}^0(T_{kl} < \infty). \end{aligned}$$

Therefore

$$\mathbf{P}^0(T_{kl} < \infty) \le \vartheta^k \qquad \text{for all } \ k \ge 0,$$

and again by monotonicity this yields

$$\begin{aligned} \textstyle\sum_{n\ge 0} \mathbf{P}(X_n \le t) &\le \textstyle\sum_{n\ge 0} \mathbf{P}^0(X_n \le t) \\ &= \mathbf{E}^0(|\{n \ge 0 : X_n \le t\}|) \\ &= \textstyle\sum_{i\ge 0} \mathbf{P}^0(T_i < \infty) \\ &\le l \textstyle\sum_{k\ge 0} \mathbf{P}^0(T_{kl} < \infty) \\ &\le l \textstyle\sum_{k\ge 0} \vartheta^k < \infty. \quad \square \end{aligned}$$

The next result is a simple consequence of (2.2) (and needed before a stronger version will be available):

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

Proof. For any $t < \infty$ satisfying

$$\mathbf{P}(Y = 0, Z \le t) > 0$$

the assertion follows from (2.2a) in view of

$$\textstyle\sum_{n\ge 0} \mathbf{P}(X_n \le t) \ge \sum_{n\in\mathbf{N}} \mathbf{P}(Y_n = 0, Z_n \le t). \quad \square$$

Equally simple is the following counterpart:

(2.4) Proposition. *ν is transient whenever $S_n \to +\infty$.*

Proof. Since

$$T := \inf\{n \in \mathbf{N} : Z_n > 0\}$$

defines an (almost surely finite) stopping time with respect to $(Z_n)_{n\in\mathbf{N}}$, by the hypothesis

$$S_n - S_T \to \infty \text{ a.s.}$$

Thus the assertion is a consequence of

$$X_n \geq Z_T\, e^{S_n - S_T} \qquad \text{for } n \geq T. \quad \square$$

That the converse of (2.4) does not hold in general can be demonstrated by a somewhat surprising result. However small the primary variable Y, in view of (2.3) only supposed to be strictly positive, may be, the secondary variable Z can be made large enough for transience, even if in addition independence of Y and Z is postulated. More precisely, in terms of distributions:

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

Proof. Let Y_n, $n \in \mathbf{N}$, be independent with distribution ν_y and assume without loss of generality $Y < 1$. Then choose a sequence $(c_n)_{n\geq 0}$ satisfying

(1) $\quad 0 < c_0 < c_1 < \ldots \to 1,$

(2) $\quad \sum_{n\geq 0} c_1 \ldots c_n < \infty$

(e.g. $c_n = (\frac{n}{n+1})^2$ for $n \in \mathbf{N}$) and a sequence $(\varepsilon_n)_{n\in\mathbf{N}}$ satisfying

(3) $\quad 1 > \varepsilon_1 > \varepsilon_2 > \ldots \to 0,$

(4) $\quad b_n := \mathbf{P}(Y_1 \ldots Y_{n-1} \leq \varepsilon_n) \;\leq\; 2^{-n}\, c_1 \ldots c_n \qquad \text{for } n \in \mathbf{N},$

where the assumption $\mathbf{P}(Y = 0) = 0$ enters. Independently of Y_n, $n \in \mathbf{N}$, let Z_n, $n \in \mathbf{N}$, be independent with a distribution ν_z such that

$$\mathbf{P}(Z = 0) = c_0 \qquad \text{and} \qquad \mathbf{P}(Z = 1/\varepsilon_n) = c_n - c_{n-1} \quad \text{for } n \in \mathbf{N},$$

hence in particular

$$\mathbf{P}(Z \leq 1/\varepsilon_n) = c_n \qquad \text{for } n \in \mathbf{N}.$$

With the dual sequence $(W_n)_{n\geq 0}$, defined in Section 0, consider now the events

$$\begin{aligned} A_n &:= \{W_n \leq 1\}, \\ B_n &:= \{Y_1 \ldots Y_{n-1} \leq \varepsilon_n\}, \\ C_n &:= \{Z_n \leq 1/\varepsilon_n\}, \end{aligned}$$

which obviously satisfy

$$A_n \subset A_{n-1} \cap (B_n \cup C_n) \subset (A_{n-1} \cap C_n) \cup B_n \qquad \text{for } n \in \mathbf{N}.$$

With the notation $a_n := \mathbf{P}(A_n)$ this yields by independence the inequality

$$a_n \leq a_{n-1}\, c_n + b_n \qquad \text{for } n \in \mathbf{N},$$

which by induction, using (4), leads to

$$a_n \leq (1 + \ldots + 2^{-n})\, c_1 \ldots c_n \qquad \text{for } n \geq 0.$$

By (2), therefore

$$\begin{aligned} \textstyle\sum_{n\geq 0} \mathbf{P}(X_n^0 \leq 1) &= \textstyle\sum_{n\geq 0} \mathbf{P}(W_n \leq 1) \\ &\leq 2 \textstyle\sum_{n\geq 0} c_1 \ldots c_n < \infty. \end{aligned}$$

According to (1.3a) the lower limit $\underline{x}$, due to $\mathbf{P}(Y < 1, Z = 0) > 0$, can only take the values 0 or ∞. But the first possibility is ruled out by (2.2a), and thus $\nu = \nu_y \otimes \nu_z$ is transient. $\square$

The final result of this section is a consequence of the monotonicity and will be crucial in Sections 5 and 7:

(2.6) Lemma. *If $f \in \mathcal{K}(\mathbf{R}_+)$, then*

$$f(X_n) - f(X_n^x) \to 0 \ \ a.s. \qquad \textit{for all } x \in \mathbf{R}_+.$$

Proof. 1. It suffices to prove the assertion under the hypothesis $X_0 = x_0 \in \mathbf{R}_+$, because its general validity then follows by integration. Comparing $f(X_n^{x_0})$ and $f(X_n^x)$ with $f(X_n^0)$ shows that in fact $x_0 = 0$ may be assumed. Since the assertion is obviously true in the transient case, moreover recurrence will be assumed in the sequel. According to (2.4) the random walk $(S_n)_{n\geq 0}$ then hits the interval $[-\infty, -\gamma]$ infinitely often with probability 1, where γ will be chosen at the end of the proof. If $T_1 < T_2 < \ldots$ are the corresponding hitting times, the random variables

$$(Y_k', Z_k') := (Y_{T_k+1}, Z_{T_k+1}) \qquad \text{for } k \in \mathbf{N}$$

are again independent and distributed according to ν.

2. Next, in view of

$$0 < \mathbf{P}(Z > 0) = \lim_{m\to\infty} \mathbf{P}(Y < mZ)$$

there exists $l \in \mathbf{N}$ such that

$$\vartheta := \mathbf{P}(Y < lZ) > 0.$$

Therefore, by part 1 of the proof, the random time

$$T := \inf\{T_k : Y_{T_k+1} < l\, Z_{T_k+1}\}$$

may be assumed to be finite for all $\omega \in \Omega$.

3. Now for fixed $\omega \in \Omega$ and $n > T(\omega)$ obviously

$$X_n^x(\omega) - X_n^0(\omega) = x \prod{}_{0<m\le T(\omega)} Y_m(\omega) \prod{}_{T(\omega)<m\le n} Y_m(\omega),$$

where the two products $\prod_1$ and $\prod_2$ can be estimated by

$$\prod{}_1 = e^{S_{T(\omega)}(\omega)} \le e^{-\gamma},$$

$$\prod{}_2 \le l\, Z_{T(\omega)+1}(\omega) \prod{}_{T(\omega)+1<m\le n} Y_m(\omega) \le l\, X_n^0(\omega).$$

Together this yields

(1) $$0 \le X_n^x(\omega) - X_n^0(\omega) \le x\, e^{-\gamma}\, l\, t \qquad \text{whenever } X_n^0(\omega) \le t.$$

4. Finally, fix t satisfying the requirement

$$\text{supp } f \subset [0,t]$$

and, given an arbitrary $\varepsilon > 0$, choose $\delta > 0$ such that

$$|f(t_1) - f(t_2)| < \varepsilon \qquad \text{for } |t_1 - t_2| < \delta.$$

This yields

$$f(X_n^x(\omega)) - f(X_n^0(\omega)) = 0 \qquad \text{for } X_n^0(\omega) > t,$$

in view of $X_n^x(\omega) \ge X_n^0(\omega)$, and

$$|f(X_n^x(\omega)) - f(X_n^0(\omega))| < \varepsilon \qquad \text{for } X_n^0(\omega) \le t,$$

if in addition

$$|X_n^x(\omega) - X_n^0(\omega)| < \delta.$$

By (1) this condition is satisfied for $x\, e^{-\gamma}\, l\, t < \delta$, i.e. if γ is chosen sufficiently large. □

3. Recurrence criteria

The sufficient conditions for recurrence and transience, given in (2.3) and (2.4), apply only to extreme cases. To deal with concrete situations stronger criteria are necessary. The first relevant result concerns essentially the additive model. A slight generalization yields:

(3.1) Theorem. *For $0 < \gamma < 1$ the following dichotomy holds:*

(a) *ν is transient, if*

$$Y \ge \gamma \qquad \textit{and} \qquad \liminf_{t\to\infty} t\, \mathbf{P}(\log Z > t) > \log \frac{1}{\gamma},$$

(b) *ν is recurrent, if*

$$Y \le \gamma \qquad \textit{and} \qquad \limsup_{t\to\infty} t\,\mathbf{P}(\log Z > t) < \log \frac{1}{\gamma}.$$

Proof. (a) Clearly, $Y = \gamma$ may be assumed. By the hypothesis

$$\liminf_{t\to\infty} t\,\mathbf{P}(\log Z > s + t) > \log \frac{1}{\gamma} \qquad \text{for all } s \ge 0,$$

which for fixed s implies the existence of $\alpha > 1$ and $l \in \mathbf{N}$ such that

$$m \log \frac{1}{\gamma}\,\mathbf{P}(\log Z > s + m \log \frac{1}{\gamma}) \ge \alpha \log \frac{1}{\gamma} \qquad \text{for } m \ge l.$$

Therefore, by independence,

$$\begin{aligned}
\mathbf{P}(X_n \le e^s) &\le \mathbf{P}(\textstyle\bigcap_{l\le m<n} \{\gamma^m Z_{n-m} \le e^s\}) \\
&= \textstyle\prod_{l\le m<n} \mathbf{P}(\log Z \le s + m \log \frac{1}{\gamma}) \\
&\le \textstyle\prod_{l\le m<n} (1 - \frac{\alpha}{m}) \\
&\le \textstyle\prod_{l\le m<n} (1 - \frac{1}{m})^\alpha \\
&= \Big(\frac{l-1}{n-1}\Big)^\alpha \qquad \text{for } n > l,
\end{aligned}$$

where the last inequality makes use of $\alpha \ge 1$, implying

$$(1-x)^\alpha \ge 1 - \alpha x \qquad \text{for } 0 \le x \le 1.$$

Summation over n yields

$$\textstyle\sum_{n\ge 0} \mathbf{P}(X_n \le e^s) \le (l+1) + (l-1)^\alpha \sum_{n\ge l} \frac{1}{n^\alpha} < \infty,$$

and the transience follows from (2.2a), because s is arbitrary.

(b) Again, $Y = \gamma$ may be assumed. In addition Z may be replaced by 0 on the set $\{Z \le z\}$ for fixed $z < \infty$, because this is irrelevant for the hypothesis and changes the values of X_n by $z/(1-\gamma)$ at most. By an appropriate choice of $\alpha < 1$ and $z < \infty$, therefore, the following conditions can be satisfied (in the given order):

(1) $\quad t\,\mathbf{P}(\log Z > t) \le \alpha \log \frac{1}{\gamma} \qquad \text{for all } t \ge 0,$

(2) $\quad \mathbf{P}(Z = 0) \ge \delta := \gamma^{1-\alpha}.$

With the abbreviation

$$m_n := \log n / \log \frac{1}{\gamma} + 2 \qquad \text{for } n \in \mathbf{N}$$

this yields by independence

$$\begin{aligned}
\mathbf{P}(X_n^0 \le 1) &\ge \mathbf{P}(\textstyle\bigcap_{\,0\le m<n} \{\gamma^m Z_{n-m} \le \frac{1}{n}\}) \\
&\ge \textstyle\prod_{\,0\le m\le m_n} \mathbf{P}(Z=0) \;\prod_{\,m_n<m<n} \mathbf{P}(Z \le \frac{1}{n}\gamma^{-m}).
\end{aligned}$$

In view of (2) and (1) the two products $\prod_1$ and $\prod_2$ can be estimated by

$$\begin{aligned}
\textstyle\prod_1 &\ge \delta^{m_n+1} = \delta^3\, n^{\alpha-1}, \\
\textstyle\prod_2 &= \textstyle\prod_{\,m_n<m<n} (1-\mathbf{P}(\log Z > [m\log\frac{1}{\gamma} - \log n])) \\
&\ge \textstyle\prod_{\,m_n<m<n} (1-\alpha\log\frac{1}{\gamma} \,/\, [m\log\frac{1}{\gamma} - \log n]),
\end{aligned}$$

because for $m > m_n$ the difference $[\ldots]$ satisfies the conditions

$$[\ldots] > 0 \quad \text{and} \quad \alpha\log\frac{1}{\gamma}\,/\,[\ldots] \le 1.$$

The estimation of $\prod_2$ can be continued by

$$\begin{aligned}
\textstyle\prod_2 &\ge \textstyle\prod_{\,m_n<m<n} (1-\alpha\,/\,(m-(m_n-2))) \\
&\ge \textstyle\prod_{\,0<l<n} (1-\frac{\alpha}{l+1}) \\
&\ge \textstyle\prod_{\,0<l<n} (1-\frac{1}{l+1})^{\alpha} \\
&= n^{-\alpha},
\end{aligned}$$

where the last inequality makes use of $0 \le \alpha \le 1$, implying

$$(1-x)^\alpha \le 1-\alpha x \qquad \text{for } \; 0 \le x \le 1.$$

Together the bounds for $\prod_1$ and $\prod_2$ yield

$$\textstyle\sum_{\,n\ge 0} \mathbf{P}(X_n^0 \le 1) \ge \sum_{\,n\in\mathbf{N}} \delta^3\, n^{\alpha-1} n^{-\alpha} = \infty,$$

and the recurrence follows from (2.2b). □

Two comments on this result are in order:

— The sufficient condition for transience is strictly stronger than the condition $\mathbf{E}(\log_+ Z) = \infty$, while the sufficient condition for recurrence is strictly stronger than the condition $\mathbf{E}((\log_+ Z)^{1-\varepsilon}) < \infty$ for all $\varepsilon > 0$.

— If $t\,\mathbf{P}(\log Z > t) \to \infty$, then the sequence $(X_n)_{n\ge 0}$ is transient whenever $Y = \gamma > 0$; if $t\,\mathbf{P}(\log Z > t) \to 0$, then it is recurrent whenever $Y = \gamma < 1$.

The rest of this section is devoted to recurrence criteria for the case that the underlying affine maps are not necessarily contractions. Here the possibility $S_n \to +\infty$ is ruled out by (2.4), i.e. $\inf_{n\ge 0} S_n = -\infty$ has to be assumed.

Then the following auxiliary result can be obtained by a partition into random blocks:

(3.2) Lemma. *If* $\inf_{n\geq 0} S_n = -\infty$ *and*

$$\mathbf{E}(X_T^0) < \infty \quad \textit{for} \quad T := \inf\{n \in \mathbf{N} : S_n < 0\},$$

then ν *is recurrent.*

Proof. 1. In view of (2.3) the hypothesis $\mathbf{P}(Y = 0) = 0$ can be used in the sequel. Set $T_0 := 0$ and let $T = T_1 < T_2 < \ldots$ denote the strictly descending ladder indices of the random walk $(S_n)_{n\geq 0}$ (being defined with probability 1). Moreover, define

$$\begin{aligned} Y'_k &:= Y_{T_{k-1}+1} \ldots Y_{T_k}, \\ Z'_k &:= Z_{T_{k-1}+1} Y_{T_{k-1}+2} \ldots Y_{T_k} + \ldots + Z_{T_k}. \end{aligned}$$

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 $\nu' \in \mathcal{N}$ such that

$$\begin{aligned} \gamma &:= \mathbf{E}(Y'_k) = \mathbf{E}(e^{S_T}) < 1, \\ \delta &:= \mathbf{E}(Z'_k) = \mathbf{E}(X_T^0) < \infty. \end{aligned}$$

2. Let now $(X'_k)_{k\geq 0}$ denote the sequence associated with $(Y'_k, Z'_k)_{k\in\mathbf{N}}$, and the initial value $X'_0 = 0$. Then, clearly

$$X'_k = X^0_{T_k} \qquad \text{for } k \geq 0,$$

and this implies by Fatou

$$\begin{aligned} \underline{x} &= \mathbf{E}(\liminf_{n\to\infty} X_n^0) \\ &\leq \mathbf{E}(\liminf_{k\to\infty} X'_k) \\ &\leq \liminf_{k\to\infty} \mathbf{E}(X'_k) \\ &= \liminf_{k\to\infty} (\delta\gamma^{k-1} + \ldots + \delta) \\ &= \delta \,/\, (1-\gamma) < \infty. \quad \square \end{aligned}$$

Now the two main recurrence criteria can be derived simultaneously:

(3.3) Theorem. *If* $\inf_{n\geq 0} S_n = -\infty$, *then* ν *is recurrent in each of the following cases:*

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

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

Proof. 1. In view of (2.3) the hypothesis $\mathbf{P}(Y=0)=0$ can be used in both cases. Moreover, the assumptions $\gamma \le 1$ in (a) and $\mathbf{E}(Z/Y) \le 1$ in (b) are admissible simplifications, because a scalar multiplication has the same effect on X_n^0, $n \ge 0$, as on Z_n, $n \in \mathbf{N}$.

2. Consider now first case (a). Then, with T as in (3.2),

$$\begin{aligned} \mathbf{E}(1_{\{T=n\}} Z_m Y_{m+1} \dots Y_n) &= \mathbf{E}(\mathbf{E}(\dots \mid Y_1, \dots, Y_n)) \\ &= \mathbf{E}(1_{\{T=n\}} Y_{m+1} \dots Y_n \, \mathbf{E}(Z_m \mid Y_m)) \\ &\le \mathbf{E}(1_{\{T=n\}} Y_{m+1} \dots Y_n) \qquad \text{for } 1 \le m \le n, \end{aligned}$$

where the last equation uses that $\{T=n\}$ is measurable with respect to $Y_1, \dots, Y_n$ and (Y_m, Z_m) is independent of Y_l, $l \ne m$. In view of $Y_1 \dots Y_n \le 1$ on $\{T=n\}$ this yields

$$\begin{aligned} \mathbf{E}(X_T^0) &= \textstyle\sum_{n\in\mathbf{N}} \sum_{1\le m\le n} \mathbf{E}(1_{\{T=n\}} Z_m Y_{m+1} \dots Y_n) \\ &\le \textstyle\sum_{m\in\mathbf{N}} \sum_{n>m} \mathbf{E}(1_{\{T=n\}} Y_{m+1} \dots Y_n) + 1 \\ &\le \textstyle\sum_{m\in\mathbf{N}} \sum_{n>m} \mathbf{E}(1_{\{T=n\}} \frac{1}{Y_1} \cdots \frac{1}{Y_m}) + 1 \\ &= \textstyle\sum_{m\ge 0} \mathbf{E}(1_{\{T>m\}} \frac{1}{Y_1} \cdots \frac{1}{Y_m}), \end{aligned}$$

where the summand 1 stands first for $\sum_{n\in\mathbf{N}} \mathbf{E}(1_{\{T=n\}})$ and then for $\mathbf{E}(1_{\{T>0\}})$. The final result can be rewritten as

$$(1) \qquad \mathbf{E}(X_T^0) \le \mathbf{E}(\textstyle\sum_{0\le m<T} e^{-S_m}).$$

3. This inequality extends to case (b). Indeed, here

$$\begin{aligned} \mathbf{E}(X_T^0) &= \textstyle\sum_{m\in\mathbf{N}} \sum_{n\ge m} \mathbf{E}(1_{\{T=n\}} \frac{Z_m}{Y_m} Y_m \dots Y_n) \\ &\le \textstyle\sum_{m\in\mathbf{N}} \sum_{n\ge m} \mathbf{E}(1_{\{T=n\}} \frac{1}{Y_1} \cdots \frac{1}{Y_{m-1}} \frac{Z_m}{Y_m}) \\ &= \textstyle\sum_{m\in\mathbf{N}} \mathbf{E}(1_{\{T>m-1\}} \frac{1}{Y_1} \cdots \frac{1}{Y_{m-1}}) \, \mathbf{E}\Big(\frac{Z_m}{Y_m}\Big) \\ &\le \textstyle\sum_{m\ge 0} \mathbf{E}(1_{\{T>m\}} \frac{1}{Y_1} \cdots \frac{1}{Y_m}), \end{aligned}$$

where the last equation uses that $\{T > m-1\}$ is measurable with respect to $Y_1, \dots, Y_{m-1}$ and Z_m/Y_m is independent of $Y_1, \dots, Y_{m-1}$.

4. Now the proof can be jointly completed. The essential tool is Spitzer's identity, which will be applied in its real form (see e.g. [28], p.395). Letting there $t \uparrow 1$ leads to

$$(2) \qquad \mathbf{E}(\textstyle\sum_{0\le m<T} e^{-S_m}) = \exp\big(\sum_{n\in\mathbf{N}} \frac{1}{n} \mathbf{E}(1_{\{S_n \ge 0\}} e^{-S_n})\big).$$

Moreover, a result by Rosén [36] provides $\delta < \infty$ such that

$$(3) \qquad \mathbf{P}(x \le S_n < x+1) \le \delta \frac{1}{\sqrt{n}} \qquad \text{for all } n \in \mathbf{N} \text{ and } x \in \mathbf{R}.$$

Combination of (1) – (3) yields

$$\begin{aligned} \log \mathbf{E}(X_T^0) \;&\le\; \textstyle\sum_{n\in\mathbf{N}} \frac{1}{n} \sum_{k\ge 0} \mathbf{P}(k \le S_n < k+1)\, e^{-k} \\ &\le\; \delta \textstyle\sum_{n\in\mathbf{N}} n^{-3/2} \sum_{k\ge 0} e^{-k} < \infty \end{aligned}$$

and, in view of (3.2), establishes the assertion. □

The main application of (3.3a) concerns the multiplicative model. More generally: if only $(S_n)_{n\ge 0}$ does not converge to $+\infty$, boundedness of Z ensures recurrence. Version (3.3b) is better adapted to a model where the role of multiplication and addition is interchanged, i.e. to the recursion

$$X'_n = Y'_n(X'_{n-1} + Z'_n) \qquad \text{for } n \in \mathbf{N};$$

in obvious notation: if only $(S'_n)_{n\ge 0}$ does not converge to $+\infty$, the existence of $\mathbf{E}(Z')$ already ensures recurrence (the additional hypothesis $\mathbf{P}(Y' = 0) = 0$ is no real restriction because of (2.3)).

It is an open problem, however, whether the conditional expectation in (3.3a) may be replaced simply by $\mathbf{E}(Z)$. Only a partial answer is possible:

(3.4) Proposition. *If* $\inf_{n\ge 0} S_n = -\infty$ *and* $\mathbf{E}(Z) < \infty$, *then* ν *is recurrent in each of the following cases:*

(a) *Y and Z independent,*

(b) *$Y \ge \gamma$ for some $\gamma > 0$.*

Proof. Both assertions are immediate consequences of their counterparts in (3.3). □

References

1. Alpuim, M.: An extremal Markovian sequence. J. Appl. Prob. **26**, 219–232 (1989)
2. Barnsley, M., Elton, J.: A new class of Markov processes for image encoding. Adv. Appl. Prob. **20**, 14–32 (1988)
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)
9. Dubins, L., Freedman, D.: Invariant probabilities for certain Markov processes. Ann. Math. Stat. **37**, 837–848 (1966)

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)
17. Grincevičius, A.: On the continuity of the distribution of a sum of dependent variables connected with independent walks on lines. Theory Prob. Appl. **19**, 163–168 (1974)
25. Lamperti, J.: Criteria for the recurrence or transience of stochastic processes I. J. Math. Anal. Appl. **1**, 314–330 (1960)
26. Lamperti, J.: Criteria for stochastic processes II: passage–time moments. J. Math. Anal. Appl. **7**, 127–145 (1963)
27. Lev, G.: Semi–Markov processes of multiplication with drift. Theory Prob. Appl. **17**, 159–164 (1972)
28. Loève, M.: Probability theory I, 4th edition. New York–Heidelberg–Berlin: Springer 1977
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)
33. Nummelin, E.: General irreducible Markov chains and non–negative operators. Cambridge: University Press 1984
34. Rachev, S., Samorodnitsky, G.: Limit laws for a stochastic process and random recursion arising in probability modelling. Preprint (1992)
35. Revuz, D.: Markov chains, 2nd edition. Amsterdam–New York: North–Holland 1984
36. Rosén, B.: On the asymptotic distribution of sums of independent identically distributed random variables. Ark. Mat. **4**, 323–332 (1961)
37. Rosenblatt, M.: Recurrent points and transition functions acting on continuous functions. Z. Wahrscheinlichkeitstheorie Verw. Geb. **30**, 173–183 (1974)
38. Tweedie, R.: Criteria for classifying general Markov chains. Adv. Appl. Prob. **8**, 737–771 (1976)
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)