ELSEVIER

Statistics & Probability Letters 40 (1998) 139–148

STATISTICS & PROBABILITY LETTERS

Verifying irreducibility and continuity of a nonlinear time series

Daren B.H. Cline *, Huay-min H. Pu

Department of Statistics, Texas A&M University, College Station, TX 77843, USA

Received September 1997

Abstract

When considering the stability of a nonlinear time series, verifying aperiodicity, irreducibility and smoothness of the transitions for the corresponding Markov chain is often the first step. Here, we provide reasonably general conditions applicable to nonlinear autoregressive time series, including many with nonadditive errors. © 1998 Elsevier Science B.V. All rights reserved

Keywords: ψ-Irreducibility; Aperiodicity; T-chain; Feller chain

1. Introduction

Suppose $\mathbb{Y} \subset \mathbb{R}^m$ and $\{\xi_t\}$ is a $\mathbb{Y}$-valued nonlinear autoregressive time series of order p, defined by

$$\xi_t = \zeta(e_t; \xi_{t-1}, \ldots, \xi_{t-p}), \quad t \geqslant 1, \tag{1.1}$$

where ζ is Borel measurable and $\{e_t\}$ is an i.i.d. sequence of random variables in some space $\mathbb{E}$, with distribution F and independent of the initial state $X_0 = (\xi_0, \ldots, \xi_{1-p})$. Thus $\{\xi_t\}$ is embedded in a Markov chain $\{X_t\}$ on $\mathbb{X} = \mathbb{Y}^p$ with

$$X_t = (\xi_t, \ldots, \xi_{t-p+1}), \quad t \geqslant 0. \tag{1.2}$$

Having embedded the time series into a Markov chain, we are interested in whether the chain is irreducible, or aperiodic or has smooth transitions (as does a Feller chain). Knowing these properties hold may make it possible or easier to establish other properties, such as the stability of the time series. For example, papers on the stability of a nonlinear time series typically start by assuming aperiodicity and irreducibility.

For time series with additive errors, establishing aperiodicity and irreducibility is often as simple as looking at the error distribution. Having a continuous density, positive on $\mathbb{R}^m$, for example, suffices and the chain will also be strong Feller in that case. However, in a nonlinear time series the error often is not additive. Its variance could depend on the current state as in conditionally heteroscedastic (ARCH) models and bilinear models, or it could be involved in a very nonlinear way as in the mixture transition distribution (MTD)

* Corresponding author. Tel.: +1 409-845-3181; fax: +1 409-845-3144; e-mail:dcline@stat.tamu.edu.

0167-7152/98/$ – see front matter
PII S0167-7152(98)00081-9

models of Le et al. (1996). With the increasing popularity of nonlinear time series modeling, there is also increased interest in fitting general, even nonparametric, models and in moving away from such simplifying assumptions as additivity.

In the present paper, we avoid assuming the model has additive errors. Specifically, we provide relatively general, but simple to check, conditions for aperiodicity, ψ-irreducibility and T-continuity (the last of which we define precisely below). Results and examples are in Section 2 and proofs are in Section 3.

We follow standard notation and terminology for a time homogeneous Markov chain $\{X_t\}$ on a topological space $\mathbb{X}$ and with transition kernel $P(x,A) = P_x(A)$. The Borel class of sets for $\mathbb{X}$ is denoted $\mathscr{B}(\mathbb{X})$ and a kernel T is any function on $\mathbb{X} \times \mathscr{B}(\mathbb{X})$ such that $T(\cdot,A)$ is measurable for all $A \in \mathscr{B}(\mathbb{X})$ and $T(x,\cdot)$ is a measure for all $x \in \mathbb{X}$. When conditioning on the initial state we will indicate expectations so:

$$E_x(\cdot) = \boldsymbol{E}(\cdot|X_0 = x).$$

We assume the reader is familiar with the definitions of small sets, aperiodicity and ψ-irreducibility. If not, the topics are thoroughly covered by Meyn and Tweedie (1993) whose notation and definitions we adopt. The notions of T-continuity and petite sets are perhaps less well known so we provide their definitions here.

Definition 1.1. Assume $\{X_t\}$ is a Markov chain on a locally compact, separable and metrizable space $\mathbb{X}$. The chain is a T-chain (or is T-continuous) if there exists a kernel T and a probability distribution $\{a_n\}$ on the nonnegative integers such that

(i) $T(x,\mathbb{X}) > 0$ for all $x \in \mathbb{X}$,
(ii) $T(\cdot,A)$ is lower semicontinuous for all $A \in \mathscr{B}(\mathbb{X})$ and
(iii) $\sum_{n=0}^{\infty} a_n P^n(x,A) \geqslant T(x,A)$ for all $x \in \mathbb{X}$, $A \in \mathscr{B}(\mathbb{X})$.

Note that, like irreducibility and aperiodicity, T-continuity may be determined by a sufficiently smooth *component* of the transition kernel rather than of the kernel itself. Thus, T-continuity generalizes the notion of a strong Feller chain but is more useful than the weak Feller property. In particular, it relates the topology of $\mathbb{X}$ to the behavior of the chain. T-continuity also appears to be just the right assumption for studying stability results (cf. Meyn and Tweedie, 1993).

A key concept for T-chains is that of the so-called petite sets, defined next.

Definition 1.2. A set $A \in \mathscr{B}(\mathbb{X})$ is petite if there exists a nontrivial measure v on $\mathscr{B}(\mathbb{X})$ and a probability distribution $\{a_n\}$ on the nonnegative integers such that

$$\sum_{n=0}^{\infty} a_n P^n(x,B) \geqslant v(B) \text{ for all } x \in A, B \in \mathscr{B}(\mathbb{X}).$$

If each compact set in $\mathbb{X}$ is petite then $\{X_t\}$ is a T-chain and if $\{X_t\}$ is a ψ-irreducible T-chain then all relatively compact sets are petite (Meyn and Tweedie, 1993, Theorem 6.2.5). Showing T-continuity typically involves, in essence, showing that all compact sets are petite. However, some important and possibly non-compact sets are also petite, as we will show. For ψ-irreducible aperiodic chains, petite sets are the same as small sets (cf. Meyn and Tweedie, 1993, Theorem 5.5.7). Identifying particular petite sets or small sets is a critical step in establishing stability (e.g., ergodicity) of the Markov chain.

2. Results and examples

Our own stability results (Cline and Pu, 1997a,b, 1998) assume that $\{X_t\}$ is an aperiodic, ψ-irreducible T-chain. Chan (1993) discusses this assumption for models with additive errors. We will provide simple ways for determining this without assuming the model has additive errors. Meyn and Tweedie (1993, Ch. 7) present a general control model approach, but their conditions require a high degree of smoothness and are therefore somewhat more stringent and difficult to check than are ours. Included with the proofs in Section 3 are lemmas for proving ψ-irreducibility and T-continuity for general Markov chains.

First, we require some notation and assumptions, essentially to ensure the transition kernel has a Lebesgue component.

Assumption 2.1. (i) $\mathbb{Y}$ *and* $\mathbb{E}$ *are open subsets of* $\mathbb{R}^m$, *and* $\mathbb{X} = \mathbb{Y}^p$. *Let* $\mu_{\mathbb{E}}$ *and* $\mu_{\mathbb{X}}$ *be Lebesgue measure on* $\mathbb{E}$ *and* $\mathbb{X}$, *respectively.*

(ii) *F has a nontrivial Lebesgue component F_a with density f.*

(iii) *For fixed $x \in \mathbb{Y}^p$, $\zeta(\cdot;x) : \mathbb{E} \to \mathbb{Y}$ is 1-1, onto and continuous. Its inverse, denoted by $\zeta^{\leftarrow}(\cdot;x)$, is differentiable and has Jacobian denoted by $J(\cdot;x)$.*

Remark. For consistency, the "current" state x comes first in kernels such as $P(x,A)$ or $T(x,A)$, but in other functions we have placed it second, behind a semicolon as in $\zeta(\cdot;x)$ and $J(\cdot;x)$.

Given the situation in Assumption 2.1, there are two simple but separate conditions each of which will give the result we want. The first assumes more about ζ (continuity in x) while the second assumes more about f. We state these in Theorem 2.2. A more general version of the second condition is given in Theorem 2.3 after the examples.

Theorem 2.2. *Assume $\{X_t\}$ is defined by* (1.1) *and* (1.2) *and Assumption* 2.1 *holds. Each of the following conditions implies $\{X_t\}$ is aperiodic, ψ-irreducible and T-continuous.*
(i) *$\zeta(u;\cdot)$ is continuous on $\mathbb{X}$ a.e.$(\mu_{\mathbb{E}})$, f is positive a.e.$(\mu_{\mathbb{E}})$ and J is locally bounded on $\mathbb{Y} \times \mathbb{X}$.*
(ii) *$f \circ \zeta^{\leftarrow}$ and J are each bounded away from 0 on compact sets of $\mathbb{Y} \times \mathbb{X}$.*

The second part of Theorem 2.2 probably gives the most useful, easily stated assumption for nonparametrically defined models, since it requires little of ζ. Chan (1993) proves a special case of this for models with additive noise. On the other hand, the first case in Theorem 2.2 will apply for some parametric models such as the amplitude dependent exponential autoregressive models (EXPAR) (Jones, 1976; Ozaki and Oda, 1978; cf. also Tong, 1990) and the Gaussian mixture transition distribution models of Le et al. (1996).

Also, note that since all three properties (aperiodicity, ψ-irreducibility and T-continuity) may be determined from a component of the transition kernel one does not necessarily have to know the density f precisely in order to check the conditions of the theorems. An appropriate lower bound on that density will suffice.

Example 2.1. Let $\mathbb{Y} = \mathbb{E} = \mathbb{R}^m$. Consider the nonlinear autoregressive model

$$\xi_t = a(\xi_{t-1},\dots,\xi_{t-p}) + b(\xi_{t-1},\dots,\xi_{t-p})e_t.$$

Both conditionally heteroscedastic (ARCH) model (Engle, 1982; Guégan and Diebolt, 1994) and bilinear models (Granger and Andersen, 1978) are of this form. (See also Tong, 1990.) Assume $a : \mathbb{R}^{pm} \to \mathbb{R}^m$ is locally bounded, $b(x)$ is an invertible $m\times m$ matrix for each $x \in \mathbb{R}^{pm}$ such that both $b(x)$ and $b^{-1}(x)$ are locally bounded and e_t has density f which is positive everywhere on $\mathbb{R}^m$. Note that $\zeta^{\leftarrow}(\cdot;x) = b^{-1}(x)(\cdot - a(x))$ and $J(\cdot;x) = \det(b^{-1}(x))$. If a and b are continuous then $\{(\xi_t,\dots,\xi_{t-p+1})\}$ is an aperiodic ψ-irreducible T-chain by Theorem 2.2(i). The same conclusion holds, by Theorem 2.2(ii), if f is locally bounded away from 0.

Example 2.2. In Example 2.1, only the variance of the noise term depends on the current state. In a self-exciting threshold (SETAR) model, however, the distribution of the noise term may depend on which of several regions the current state is in. A model which incorporates this is

$$\xi_t = a(\xi_{t-1},\dots,\xi_{t-p}) + b(\xi_{t-1},\dots,\xi_{t-p})F_j^{\leftarrow}(e_t) \quad \text{if } \xi_{t-1} \in R_j, \tag{2.1}$$

where a and b are as in Example 2.1, e_t has uniform (0,1) distribution, $\{R_1,\dots,R_k\}$ partitions $\mathbb{R}^m$, and $F_1,\dots,F_k$ are probability distributions with positive densities $f_1,\dots,f_k$. Clearly, if the conditions referred to in Example 2.1 hold with $f = f_j$ for each j then the Markov process associated with (2.1) is aperiodic, ψ-irreducible and T-continuous.

Example 2.3. A mixture version of (1.1) would be

$$\xi_t = \zeta_i(e_t;\xi_{t-1},\dots,\xi_{t-p}) \quad \text{w.p. } p_i,$$

where $p_i > 0$, $i = 1,\dots,k$ and $p_1 + \cdots + p_k = 1$. This is a general, and possibly nonparametric, version of the MTD models (cf. Le et al., 1996). If F and some ζ_i, $1 \leqslant i \leqslant k$, satisfy the conditions of either Theorem 2.2(i) or 2.2(ii) then the conclusion of that theorem will hold. This is so since only a component of the transition distribution needs to be sufficiently smooth. Such ζ_i could be as in Example 2.1 or 2.2 above.

The next theorem is a generalization of the second part of Theorem 2.2.

Theorem 2.3. *Assume* $\{X_t\}$ *is defined by* (1.1)–(1.2) *and Assumption* 2.1 *holds. For* $x = (x_1,\dots,x_p)$ *and* $x^* = (x_{p+1},\dots,x_{2p})$, *define*

$$g(x;x^*) = \prod_{k=1}^{p} J(x_k;x_{k+1},\dots,x_{k+p}) f(\zeta^{\leftarrow}(x_k;x_{k+1},\dots,x_{k+p})). \tag{2.2}$$

If there exists a nonempty open set $G \subset \mathbb{X}$ *and for each* $y \in \mathbb{X}$ *there exists an open* $B_y \subset \mathbb{X}$ *such that* $y \in B_y$ *and*

$$\inf_{x^* \in B_y} g(x;x^*) > 0 \quad \textit{for almost all } x \in G$$

then $\{X_t\}$ *is aperiodic,* ψ*-irreducible and* T*-continuous.*

The function g in (2.2) is essentially the transition density (or a component of it) for X_p, given $X_0 = x^*$. The behavior of g is critical since the dimension of the error space $\mathbb{E}$ is typically less than that of the state space $\mathbb{X}$ and studying only the one-step transition $P(x,\cdot)$ usually does not suffice.

To show only that $\{X_t\}$ is a T-chain, one may weaken the conditions slightly. For part (i) of Theorem 2.2, the condition that f is positive a.e. is not needed for T-continuity. For Theorem 2.3 (which includes Theorem 2.2(ii)), the open set G may depend on y.

If f is not positive almost everywhere (e.g., has compact support), it may still be possible to verify ψ-irreducibility. Let T be the kernel defined by

$$T(x^*,A) = \int_A g(x;x^*)\,\mathrm{d}x,$$

define $K = \sum_{i=1}^{\infty} 2^{-i}T^i$ and let $k(x;x^*)$ be the corresponding kernel density. To prove ψ-irreducibility (as well as T-continuity), the assumption f is positive almost everywhere in Theorem 2.2(i) may be replaced with $k(\cdot;x)$ is positive almost everywhere for each x. Likewise, if k replaces g in Theorem 2.3, both ψ-irreducibility and T-continuity are assured.

In case $m = p = 1$ and the process is real-valued with additive noise, specific conditions for T-continuity may be formulated depending on the discontinuities of the autoregression function. The next result illustrates how these conditions can range from one extreme to another.

Corollary 2.4. *Assume* $\mathbb{X} = \mathbb{E} = \mathbb{R}$ *and*

$$\xi_t = a(\xi_{t-1}) + e_t$$

where e_t *has density* f. *Define*

$$\omega(x) = \lim_{\delta \downarrow 0} \sup_{\substack{|x_1 - x| < \delta \\ |x_2 - x| < \delta}} |a(x_1) - a(x_2)|.$$

Any one of the following is sufficient for $\{\xi_t\}$ *to be a* T-*chain.*

(i) *a is continuous on* $\mathbb{R}$.

(ii) *For each* $x \in \mathbb{R}$, $\omega(x) < \infty$ *and there exists an interval* I_x *of length greater than* $\omega(x)$ *on which* f *is bounded away from 0.*

(iii) *For each* $x \in \mathbb{R}$, *the set of limit points of* a *at* x *is finite,* $\omega(x) < \infty$ *and there exists an open interval* I_x *of length greater than* $\omega(x)$ *such that every open interval in* I_x *contains a nonempty subinterval on which* f *is bounded away from zero.*

(iv) *The set of discontinuities of* a *is locally finite, the set of limit points of* a *at each* $x \in \mathbb{R}$ *is finite and* f *is positive everywhere.*

(v) *a is locally bounded and* f *is locally bounded away from 0.*

With the additional condition, of course, that f is positive almost everywhere the model in Corollary 2.4 is aperiodic and ψ-irreducible.

We observed earlier that for ψ-irreducible T-chains, compact sets are petite. Petiteness of a set, however, is a characteristic of the continuity properties of the chain and as such is defined by the chain. In particular, certain noncompact sets associated with the chain can be petite. This is described in our final result.

Theorem 2.5. *Suppose* $\{X_t\}$ *is a* ψ-*irreducible* T-*chain on* $\mathbb{X}$, *a closed subset of* $\mathbb{R}^{pm}$. *Suppose also*

$$X_t = \alpha(X_{t-1}) + \beta(e_t; X_{t-1}),$$

where $\alpha(x)$ *is locally bounded and, for some* $r > 0$,

$$\sup_{\|\alpha(x)\| \leqslant M} E_x(\|\beta(e_t;x)\|^r) < \infty \quad \text{for all } M < \infty. \tag{2.3}$$

Then $\{x : \|\alpha(x)\| \leqslant M\}$ *is petite for every finite* M.

The observation that $\{x : \|\alpha(x)\| \leqslant M\}$ is petite makes it possible to construct better criteria for ergodicity or recurrence of the process (Cline and Pu, 1997a).

3. Proofs

Recall from Definition 1.1 that showing T-continuity requires finding a nontrivial kernel T for which $T(\cdot, A)$ is lower semicontinuous for all $A \in \mathscr{B}(\mathbb{X})$. The lemma given next is useful for identifying a lower semicontinuous kernel.

Lemma 3.1. *Assume $\mathbb{X}$ is a locally compact, separable metrizable space, $T : \mathbb{X} \times \mathscr{B}(\mathbb{X}) \to [0,1]$ is a kernel and μ is a measure on $\mathbb{X}$ which is bounded on compact sets. If*

(i) *for each $\varepsilon > 0$ and compact K_1, K_2 there is $\delta > 0$ such that if $A \subset K_2$ and $\mu(A) < \delta$ then $\sup_{y\in K_1} T(y,A) < \varepsilon$, and*

(ii) *$T(\cdot, O)$ is lower semicontinuous for all (relatively compact) open sets O, then $T(\cdot,A)$ is lower semicontinuous for all $A \in \mathscr{B}(\mathbb{X})$.*

Proof. Note that $\mathbb{X}$ is also topologically complete (cf. Ash, 1972, Theorem A.9.12). Let C be a compact subset of $\mathbb{X}$. For some open A and compact K_2, we have $C \subset A \subset K_2$. Fix $x \in \mathbb{X}$ and choose K_1 to be a compact set containing a neighborhood of x. Given $\varepsilon > 0$, let δ be as in (i). Also there exists an open set O such that $C \subset O \subset A$ and $\mu(OC^c) < \delta$. Hence,

$$\begin{aligned}\liminf_{y\to x} T(y,C) &\geqslant \liminf_{y\to x} T(y,O) - \varepsilon \\ &\geqslant T(x,O) - \varepsilon \\ &\geqslant T(x,C) - \varepsilon.\end{aligned}$$

It follows that $T(\cdot, C)$ is lower semicontinuous for every compact C.

By standard arguments (cf. Ash, 1972, Corollary 4.3.7 and Theorem 4.3.8), $T(\cdot,A)$ is lower semicontinuous for all $A \in \mathscr{B}(\mathbb{X})$. □

Remark. The conditions in Lemma 3.1 are not any stronger than we need them to be. They are obviously satisfied when $T(x,A) = s(x)\mu(A)$, where s is a positive continuous function on $\mathbb{X}$ and μ is a nontrivial measure defined on $\mathscr{B}(\mathbb{X})$. But in fact if $\{X_t\}$ is a ψ-irreducible T-chain then it is always possible to choose such a kernel T and to have it also satisfy Definition 1.1(iii) (Meyn and Tweedie, 1993, Proposition 6.2.6).

Next is the lemma we use to verify the chain is aperiodic and ψ-irreducible.

Lemma 3.2. *Suppose $\{X_t\}$ is a T-chain on a locally compact, separable metrizable space $\mathbb{X}$. Suppose there exists a positive integer k and a kernel T_* such that*

(i) *$T_*(x,A) \leqslant P^k(x,A)$ for all $x \in \mathbb{X}$ and $A \in \mathscr{B}(\mathbb{X})$,*

(ii) *$T_*(\cdot,A)$ is lower semicontinuous for all $A \in \mathscr{B}(\mathbb{X})$, and*

(iii) *there exists $x^* \in \mathbb{X}$ such that $T_*(x,O) > 0$ for any $x \in \mathbb{X}$ and open set O containing x^*.*

Then $\{X_t\}$ is aperiodic and ψ-irreducible.

Proof. To show irreducibility we will follow the argument in Meyn and Tweedie (1993, Proposition 6.2.1) and then extend it to show aperiodicity. Suppose $A \in \mathscr{B}(\mathbb{X})$ is such that $T_*(x^*,A) > 0$. By (ii), there exists an open set O_1 such that $x^* \in O_1$ and

$$T_*(y,A) > \tfrac{1}{2}T_*(x^*,A) \quad \text{for all } y \in O_1.$$

Since $T_*(y, O_1) > 0$ by (iii), for any $y \in \mathbb{X}$, we have

$$\begin{aligned}P^{2k}(y,A) &\geqslant \int_{O_1} T_*(x,A)T_*(y,dx) \\ &\geqslant \tfrac{1}{2}T_*(x^*,A)T_*(y,O_1) > 0.\end{aligned} \tag{3.1}$$

This shows $\{X_t\}$ is ψ-irreducible with irreducibility measure $T_*(x^*,\cdot)$ (cf. Meyn and Tweedie, 1993, Ch. 4).

In addition, $T_*(x^*,O_1) > 0$ and $T_*(\cdot,O_1)$ is lower semicontinuous so there exists an open set O_2 which contains x^* and for which

$$T_*(y,O_1) > \tfrac{1}{2}T_*(x^*,O_1) \qquad \text{for all } y \in O_2.$$

Using (3.1), we obtain

$$P^{2k}(y,A) \geqslant \tfrac{1}{4}T_*(x^*,O_1)T_*(x^*,A) \quad \text{for all } y \in O_2,\ A \in \mathscr{B}(\mathbb{X}), \tag{3.2}$$

and therefore O_2 is a small set (cf. Meyn and Tweedie, 1993, Ch. 5).

Likewise, $PT_*(\cdot,O_1)$ is lower semicontinuous and

$$PT_*(x^*,O_1) = \int T_*(y,O_1)P(x^*,\mathrm{d}y) > 0$$

so there exists an $O_3 \subset O_2$ such that $x^* \in O_3$ and

$$PT_*(y,O_1) > \tfrac{1}{2}PT_*(x^*,O_1) \quad \text{for all } y \in O_3.$$

Furthermore, by repeating the argument for (3.2),

$$P^{2k+1}(y,A) \geqslant \tfrac{1}{4}PT_*(x^*,O_1)T_*(x^*,A) \quad \text{for all } y \in O_3,\ A \in \mathscr{B}(\mathbb{X}). \tag{3.3}$$

From (3.2) and (3.3) it follows that O_3 is small and that $\{X_t\}$ is aperiodic (Meyn and Tweedie, 1993, Theorem 5.4.4). □

Proof of Theorem 2.2. (i) We use the notation as in (2.1): $x = (x_1,\dots,x_p)$ and $x^* = (x_{p+1},\dots,x_{2p})$, where $x_i \in \mathbb{Y}$. With this notation and given $X_0 = x^*$, we have

$$X_1 = (\zeta(e_1;x^*),x_{p+1},\dots,x_{2p-1}).$$

Define

$$T(x^*,A) = \int_{\mathbb{E}} 1_A((\zeta(u;x^*),x_{p+1},\dots,x_{2p-1}))f(u)\,\mathrm{d}u \quad \text{for } x^* \in \mathbb{X},\ A \in \mathscr{B}(\mathbb{X}), \tag{3.4}$$

where $1_A(\cdot)$ is the indicator function over A. By Assumption 2.1(ii), $P(x^*,A) \geqslant T(x^*,A)$, and thus

$$P^k(x^*,A) \geqslant T^k(x^*,A) \quad \text{for all } k,x^*,A.$$

Let $\mu_{\mathbb{E}^p}$ be Lebesgue measure on $\mathbb{E}^p$. For $w = (w_1,\dots,w_p) \in \mathbb{E}^p$, recursively define

$$\begin{aligned}
\zeta_p(w;x^*) &= \zeta(w_p;x_{p+1},\dots,x_{2p}),\\
\zeta_{p-1}(w;x^*) &= \zeta(w_{p-1};\zeta_p(w;x^*),x_{p+1},\dots,x_{2p-1}),\\
&\ \vdots\\
\zeta_1(w;x^*) &= \zeta(w_1;\zeta_2(w;x^*),\dots,\zeta_p(w;x^*),x_{p+1}).
\end{aligned} \tag{3.5}$$

Now fix $\varepsilon > 0$, $x \in \mathbb{X}$ and choose $\delta > 0$ so that $\mu_{\mathbb{E}^p}(B) < \delta$ implies

$$\int_B \prod_{k=1}^{p} f(w_k)\,\mathrm{d}w < \varepsilon.$$

Let K_1, K_2 be compact subsets of $\mathbb{X}$. For $x^* \in K_1, x \in K_2$ there exists $M < \infty$ such that

$$\prod_{k=1}^{p} J(x_k; x_{k+1}, \ldots, x_{k+p}) \leqslant M.$$

Let A be a relatively compact set in $\mathscr{B}(\mathbb{X})$ such that $A \subset K_2$ and $\mu_{\mathbb{X}}(A) < \delta/M$. Define

$$B = \{w \in \mathbb{E}^p : \ (\zeta_1(w; x^*), \ldots, \zeta_p(w; x^*)) \in A\}. \tag{3.6}$$

Thus, if $x^* \in K_1$ then

$$\begin{aligned}\mu_{\mathbb{E}^p}(B) &= \int_A \prod_{k=1}^{p} J(x_k; x_{k+1}, \ldots, x_{k+p}) \, \mathrm{d}x \\ &\leqslant M \mu_{\mathbb{X}}(A) < \delta\end{aligned}$$

and therefore,

$$\begin{aligned}T^p(x^*, A) &= \int_{\mathbb{E}^p} 1_A(\zeta_1(w; x^*), \ldots, \zeta_p(w; x^*)) \, \mathrm{d}w \\ &= \int_B \prod_{k=1}^{p} f(w_k) \, \mathrm{d}w < \varepsilon \quad \text{for all } x^* \in K_1.\end{aligned}$$

This verifies condition 3.1(i) for the kernel T^p and the measure $\mu_{\mathbb{X}}$.

Let A be open in $\mathbb{X}$. By assumption (i), $\{y \in \mathbb{X} : \ \zeta(u; y) \in A\}$ is open a.e.(F_a) and therefore $1_A(\zeta(u; \cdot))$ is lower semicontinuous a.e.(F_a). It follows, by Fatou's lemma, that $T(\cdot, A)$ is lower semicontinuous and by similar reasoning that $T^p(\cdot, A)$ is lower semicontinuous. Since A is an arbitrary open set, we have verified condition 3.1(ii) for the kernel T^p and it follows from Lemma 3.1 that $T^p(\cdot, A)$ is lower semicontinuous for all $A \in \mathscr{B}(\mathbb{X})$. Thus $\{X_t\}$ is a T-chain.

By Assumption 2.1(iii), the vector $(\zeta_p(\cdot; x^*), \ldots, \zeta_1(\cdot; x^*))$ defined in (3.5) is a 1–1, onto and continuous mapping of $\mathbb{E}^p$ to $\mathbb{X}$. Therefore, if A is an open and nonempty subset of $\mathbb{X}$ then B defined in (3.6) is also open and nonempty. Since f is positive a.e.,

$$T^p(x^*, A) = \int_B \prod_{k=1}^{p} f(w_k) \, \mathrm{d}w > 0.$$

By Lemma 3.2, $\{X_t\}$ is aperiodic and ψ-irreducible.

(ii) This follows from Theorem 2.3 as a special case. □

Proof of Theorem 2.3. From the definition of T in (3.4) we have

$$T^p(x^*, A) = \int_A g(x; x^*) \, \mathrm{d}x.$$

Define

$$T_y(\cdot, A) = 1_{B_y}(\cdot) \int_A \inf_{x^* \in B_y} g(x; x^*) \, \mathrm{d}x. \tag{3.7}$$

For fixed A the integral is a constant and B_y is open and so $T_y(\cdot, A)$ is lower semicontinuous. Borrowing an argument from Meyn and Tweedie (1993, Proposition 6.2.4), let $\{y_i\}$ be a countable set of points in $\mathbb{X}$ such that $\bigcup_{i=1}^{\infty} B_{y_i} = \mathbb{X}$ and let $T_* = \sum_{i=1}^{\infty} 2^{-i} T_{y_i}$. Then $T_*(\cdot, A)$ is lower semicontinuous for each $A \in \mathscr{B}(\mathbb{X})$ and

$$P^p(y,A) \geqslant T^p(y,A) \geqslant T_*(y,A) \quad \text{for all } y \in \mathbb{X},\ A \in \mathscr{B}(\mathbb{X}).$$

Therefore, $\{X_t\}$ is a T-chain.

Furthermore, for any open set O intersecting G and any $y \in \mathbb{X}$,

$$T_*(y,O) = \sum_{i=1}^{\infty} 2^{-i} 1_{B_{y_i}}(y) \int_O \inf_{x^* \in B_{y_i}} g(x;x^*)\,\mathrm{d}x > 0.$$

So by Lemma 3.2, $\{X_t\}$ is aperiodic and ψ-irreducible. □

Proof of Corollary 2.4. Cases (i), (ii) and (v) are direct consequences of the T-chain part of Theorems 2.2 and 2.3. Case (iv) is proved in a manner analogous to Theorem 2.3. See Pu (1995) for proofs. Case (iii) also follows from Theorem 2.3 and we provide its proof here:

Fix y and let $a_1 < a_2 < \cdots < a_k$ be the limit points of a at y (including $a(y)$ itself). Since I_y has length greater than $\omega(y)$, there exist q and c so that the intervals $H_i = (q - a_i - c, q - a_i + c)$ are all in I_y and are disjoint. Choose open $J_1 \subset H_1$ so that f is bounded away from 0 on J_1. Now, recursively for $j = 2, \ldots, k$ choose open $J_j \subset J_{j-1} + a_{j-1} - a_j$ (since $J_{j-1} + a_{j-1} - a_j \subset H_j$) so that f is bounded away from 0 on J_j. By this construction, we have $J_k + a_k - a_j \subset J_j$ for each j.

Let G_y be the middle third of $J_k + a_k$ and let ε be the length of G_y. Choose δ so that $|x^* - y| < \delta$ implies $|a(x^*) - a_j| < \varepsilon$ for some j and let $B_y = (y - \delta, y + \delta)$. Therefore, $x \in G_y$ and $x^* \in B_y$ implies $x - a(x^*) \in J_j$ for some j. Hence, if $x \in G_y$,

$$\begin{aligned} \inf_{x^* \in B_y} g(x;x^*) &= \inf_{x^* \in B_y} f(x - a(x^*)) \\ &\geqslant \min_j \inf_{u \in J_j} f(u) > 0. \end{aligned}$$

As remarked following Theorem 2.3, this suffices to show T-continuity. □

Proof of Theorem 2.5. Given $M < \infty$ choose $K_M > \sup_{\|\alpha(x)\| \leqslant M} E_x(\|X_1\|^r)$, which is finite by (2.3). Thus, by Markov's inequality

$$\inf_{\|\alpha(x)\| \leqslant M} P_x(\|X_1\| \leqslant K_M) > 0.$$

Since $C = \{x : \|x\| \leqslant K_M\}$ is compact, it is petite (Meyn and Tweedie, 1993, Theorem 6.2.5). Therefore, $D = \{x : \|\alpha(x)\| \leqslant M\}$ is petite (Meyn and Tweedie, 1993, Proposition 5.5.4(i)). □

References

Ash, R.B., 1972. Real Analysis and Probability. Academic Press, Orlando, FL.

Chan, K.S. 1993. A review of some limit theorems of Markov chains and their applications. In: H. Tong (Ed.), Dimensions, Estimation and Models. World Scientific, Singapore, pp. 108–135.

Cline, D.B.H., Pu, H.H., 1997a. Stability of nonlinear AR(1) times series with delay. Statistics Dept., Texas A&M University, College Station, TX.

Cline, D.B.H., Pu, H.H., 1997b. Geometric ergodicity of nonlinear time series. Statistics Dept., Texas A&M University, College Station, TX.

Cline, D.B.H., Pu, H.H., 1998. Geometric transience of nonlinear time series. Statistics Dept., Texas A&M University, College Station, TX.

Engle, R.F., 1982. Autoregressive conditional heteroscedasticity with estimates of the U.K. inflation. Econometrica 50, 987–1008.

Granger, C.W.J., Andersen, A.P., 1978. An Introduction to Bilinear Time Series Models. Vandenhoek and Ruprecht, Gottingen.

Guégan, D., Diebolt, J., 1994. Probabilistic properties of the β-ARCH model. Statist. Sinica 4, 71–87.

Jones, D.A., 1976. Nonlinear autoregressive processes. Proc. Roy. Soc. A 360, 71–95.

Le, N.D, Martin, R.D., Raftery, A.E., 1996. Modeling flat stretches, bursts, and outliers in time series using mixture transition distribution models. J. Amer. Statist. Assoc. 91, 1504–1515.

Meyn, S.P., Tweedie, R.L., 1993. Markov Chains and Stochastic Stability. Springer, London.

Ozaki, T., Oda, H., 1978. Non-linear time series model identification by Akaike's information criterion. Proc. IFAC Workshop on Information and Systems. Compeign, France, October 1977.

Pu, H.H., 1995. Recurrence and transience for autoregressive nonlinear time series. Dissertation, Texas A&M University, College Station, TX.

Tong, H., 1990. Non-linear Time Series Analysis: A Dynamical System Approach. Oxford University Press, London.