Chapter 17
Stochastic Recursive Sequences

Abstract The chapter begins with introducing the concept of stochastic random sequences in Sect. 17.1. The idea of renovating events together with the key results on ergodicity of stochastic random sequences and the boundedness thereof is presented in Sect. 17.2, whereas the Loynes ergodic theorem for the case of monotone functions specifying the recursion is proved in Sect. 17.3. Section 17.4 establishes ergodicity conditions for contracting in mean Lipschitz transformations.

17.1 Basic Concepts

Consider two measurable state spaces $\langle \mathcal{X}, \mathfrak{B}_{\mathcal{X}} \rangle$ and $\langle \mathcal{Y}, \mathfrak{B}_{\mathcal{Y}} \rangle$, and let $\{\xi_n\}$ be a sequence of random elements taking values in $\mathcal{Y}$. If $\langle \Omega, \mathfrak{F}, \mathbf{P} \rangle$ is the underlying probability space, then $\{\omega : \xi_k \in B\} \in \mathfrak{F}$ for any $B \in \mathfrak{B}_{\mathcal{Y}}$. Assume, moreover, that a measurable function $f : \mathcal{X} \times \mathcal{Y} \to \mathcal{X}$ is given on the measurable space $\langle \mathcal{X} \times \mathcal{Y}, \mathfrak{B}_{\mathcal{X}} \times \mathfrak{B}_{\mathcal{Y}} \rangle$, where $\mathfrak{B}_{\mathcal{X}} \times \mathfrak{B}_{\mathcal{Y}}$ denotes the σ-algebra generated by sets $A \times B$ with $A \in \mathfrak{B}_{\mathcal{X}}$ and $B \in \mathfrak{B}_{\mathcal{Y}}$.

For simplicity's sake, by $\mathcal{X}$ and $\mathcal{Y}$ we can understand the real line $\mathbb{R}$, and by $\mathfrak{B}_{\mathcal{X}}$, $\mathfrak{B}_{\mathcal{Y}}$ the σ-algebras of Borel sets.

Definition 17.1.1 A sequence $\{X_n\}$, $n = 0, 1, \ldots$, taking values in $\langle \mathcal{X}, \mathfrak{B}_{\mathcal{X}} \rangle$ is said to be a *stochastic recursive sequence* (s.r.s.) driven by the sequence $\{\xi_n\}$ if X_n satisfies the relation

$$X_{n+1} = f(X_n, \xi_n) \tag{17.1.1}$$

for all $n \geq 0$. For simplicity's sake we will assume that the initial state X_0 is independent of $\{\xi_n\}$.

The distribution of the sequence $\{X_n, \xi_n\}$ on $\langle (\mathcal{X} \times \mathcal{Y})^\infty, (B_{\mathcal{X}} \times B_{\mathcal{Y}})^\infty \rangle$ can be constructed in an obvious way from finite-dimensional distributions similarly to the manner in which we constructed on $\langle \mathcal{X}^\infty, \mathfrak{B}_{\mathcal{X}}^\infty \rangle$ the distribution of a Markov chain X with values in $\langle \mathcal{X}, \mathfrak{B}_{\mathcal{X}} \rangle$ from its transition function $P(x, B) = \mathbf{P}(X_1(x) \in B)$. The finite-dimensional distributions of $\{(X_0, \xi_0), \ldots, (X_k, \xi_k)\}$ for the s.r.s. are given by the relations

$\mathbf{P}(X_l \in A_l, \, \xi_l \in B_l; \, l = 0, \dots, k)$

$$= \int_{B_0} \cdots \int_{B_k} \mathbf{P}(\xi_l \in dy_l, \, l = 0, \dots, k) \prod_{l=1}^{k} \mathrm{I}\big(f_l(X_0, y_0, \dots, y_l) \in A_l\big),$$

where $f_1(x, y_0) := f(x, y_0)$, $f_l(x, y_0, \dots, y_l) := f(f_{l-1}(x, y_0, \dots, y_{l-1}), y_l)$.

Without loss of generality, the sequence $\{\xi_n\}$ can be assumed to be given for all $-\infty < n < \infty$ (as we noted in Sect. 16.1, for a stationary sequence, the required extension to $n < 0$ can always be achieved with the help of Kolmogorov's theorem).

A *stochastic recursive sequence is a more general object than a Markov chain*. It is evident that if ξ_k are independent, then the X_n form a Markov chain. A stronger assertion is true as well: under broad assumptions about the space $\langle \mathcal{X}, \mathfrak{B}_{\mathcal{X}} \rangle$, for any Markov chain $\{X_n\}$ in $\langle \mathfrak{X}, \mathfrak{B}_{\mathfrak{X}} \rangle$ one can construct a function f and a sequence of independent identically distributed random variables $\{\xi_n\}$ such that (17.1.1) holds. We will elucidate this statement in the simplest case when both $\mathcal{X}$ and $\mathcal{Y}$ coincide with the real line $\mathbb{R}$. Let $P(x, B)$, $B \in \mathfrak{B}$, be the transition function of the chain $\{X_n\}$, and $F_x(t) = P(x, (-\infty, t))$ the distribution function of $X_1(x)$ $(X_0 = x)$. Then if $F_x^{-1}(t)$ is the function inverse (in t) to $F_x(t)$ and $\alpha \Subset \mathbf{U}_{0,1}$ is a random variable, then, as we saw before (see e.g. Sect. 6.2), the random variable $F_x^{-1}(\alpha)$ will have the distribution function $F_x(t)$. Therefore, if $\{\alpha_n\}$ is a sequence of independent random variables uniformly distributed over $[0, 1]$, then the sequence $X_{n+1} = F_{X_n}^{-1}(\alpha_n)$ will have the same distribution as the original chain $\{X_n\}$. Thus the Markov chain is an s.r.s. with the function $f(x, y) = F_x^{-1}(y)$ and driving sequence $\{\alpha_n\}$, $\alpha_n \Subset \mathbf{U}_{0,1}$.

For more general state spaces $\mathcal{X}$, a similar construction is possible if the σ-algebra $\mathfrak{B}_{\mathcal{X}}$ is countably-generated (i.e. is generated by a countable collection of sets from $\mathcal{X}$). This is always the case for Borel σ-algebras in $\mathcal{X} = \mathbb{R}^d$, $d \geq 1$ (see [22]).

One can always consider $f(\cdot, \xi_n)$ as a sequence of *random mappings* of the space $\mathcal{X}$ into itself. The principal problem we will be interested in is again (as in Chap. 13) that of the existence of the limiting distribution of X_n as $n \to \infty$.

In the following sections we will consider three basic approaches to this problem.

17.2 Ergodicity and Renovating Events. Boundedness Conditions

17.2.1 Ergodicity of Stochastic Recursive Sequences

We introduce the σ-algebras

$$\mathfrak{F}_{l,n}^{\xi} := \sigma\{\xi_k; l \leq k \leq n\},$$

$$\mathfrak{F}_n^{\xi} := \sigma\{\xi_k; k \leq n\} = \mathfrak{F}_{-\infty,n}^{\xi},$$

$$\mathfrak{F}^\xi := \sigma\{\xi_k; -\infty < k < \infty\} = \mathfrak{F}^\xi_{-\infty,\infty}.$$

In the sequel, for the sake of definiteness and simplicity, we will assume the initial value X_0 to be constant unless otherwise stated.

Definition 17.2.1 An event $A \in \mathfrak{F}^\xi_{n+m}$, $m \geq 0$, is said to be *renovating* for the s.r.s. $\{X_n\}$ on the segment $[n, n+m]$ if there exists a measurable function $g : \mathcal{Y}^{m+1} \to \mathcal{X}$ such that, on the set A (i.e. for $\omega \in A$),

$$X_{n+m+1} = g(\xi_n, \ldots, \xi_{n+m}). \tag{17.2.1}$$

It is evident that, for $\omega \in A$, relations of the form $X_{n+m+k+1} = g_k(\xi_n, \ldots, \xi_{n+m+k})$ will hold for all $k \geq 0$, where g_k is a function depending on its arguments only and determined by the event A.

The *sequence of events* $\{A_n\}$, $A_n \in \mathfrak{F}^\xi_{n+m}$, where the integer m is fixed, is said to be *renovating* for the s.r.s. $\{X_n\}$ if there exists an integer $n_0 \geq 0$ such that, for $n \geq n_0$, one has relation (17.2.1) for $\omega \in A_n$, the function g being common for all n.

We will be mainly interested in "positive" renovating events, i.e. renovating events having positive probabilities $\mathbf{P}(A_n) > 0$.

The simplest example of a renovating event is the hitting by the sequence X_n of a fixed point $x_0 : A_n = \{X_n = x_0\}$ (here $m = 0$), although such an event could be of zero probability. Below we will consider a more interesting example.

The motivation behind the introduction of renovating events is as follows. After the trajectory $\{X_k, \xi_k\}$, $k \leq n+m$, has entered a renovating set $A \in \mathfrak{F}^\xi_{n+m}$, the future evolution of the process will not depend on the values $\{X_k\}$, $k \leq n+m$, but will be determined by the values of $\xi_k, \xi_{k+1}, \ldots$ only. It is not a complete "regeneration" of the process which we dealt with in Chap. 13 while studying Markov chains (first of all, because the ξ_k are now, generally speaking, dependent), but it still enables us to establish ergodicity of the sequence X_n (in approximately the same sense as in Chap. 13).

Note that, generally speaking, the event A and hence the function g may depend on the initial value X_0. If X_0 is random then a renovating event is to be taken from the σ-algebra $\mathfrak{F}^\xi_{n+m} \times \sigma(X_0)$.

In what follows it will be assumed that the sequence $\{\xi_n\}$ *is stationary.* The symbol U will denote the measure preserving shift transformation of $\mathfrak{F}^\xi$-measurable random variables generated by $\{\xi_n\}$, so that $U\xi_n = \xi_{n+1}$, and the symbol T will denote the shift transformation of sets (events) from the σ-algebra $\mathfrak{F}^\xi : \xi_{n+1}(\omega) = \xi_n(T\omega)$. The symbols U^n and T^n, $n \geq 0$, will denote the powers (iterations) of these transformations respectively (so that $U^1 = U$, $T^1 = T$; U^0 and T^0 are identity transformations), while U^{-n} and T^{-n} are transformations inverse to U^n and T^n, respectively.

A sequence of events $\{A_k\}$ is said to be *stationary* if $A_k = T^k A_0$ for all k.

Example 17.2.1 Consider a real-valued sequence

$$X_{n+1} = (X_n + \xi_n)^+, \quad X_0 = \text{const} \geq 0, \quad n \geq 0, \tag{17.2.2}$$

where $x^+ = \max(0, x)$ and $\{\xi_n\}$ is a stationary metric transitive sequence. As we already know from Sect. 12.4, the sequence $\{X_n\}$ describes the dynamics of waiting times for customers in a single-channel service system. The difference is that in Sect. 12.4 the initial value has subscript 1 rather than 0, and that now the sequence $\{\xi_n\}$ has a more general nature. Furthermore, it was established in Sect. 12.4 that Eq. (17.2.2) has the solution

$$X_{n+1} = \max(\overline{S}_{n,n}, X_0 + S_n), \tag{17.2.3}$$

where

$$S_n := \sum_{k=0}^{n} \xi_k, \qquad \overline{S}_{n,k} := \max_{-1 \le j \le k} S_{n,j}, \qquad S_{n,j} := \sum_{k=n-j}^{n} \xi_k, \qquad S_{n,-1} := 0 \tag{17.2.4}$$

(certain changes in the subscripts in comparison to (17.2.4) are caused by different indexing of the initial values). From representation (17.2.3) one can see that the event

$$B_n := \{X_0 + S_n \le 0, \overline{S}_{n,n} = 0\} \in \mathfrak{F}_n^{\xi}$$

implies the event $\{X_{n+1} = 0\}$ and so is renovating for $m = 0$, $g(y) \equiv 0$. If $X_{n+1} = 0$ then

$$X_{n+2} = g_1(\xi_n, \xi_{n+1}) := \xi_{n+1}^+, \quad X_{n+3} = g_2(\xi_n, \xi_{n+1}, \xi_{n+2}) := \left(\xi_{n+1}^+ + \xi_{n+2}\right)^+,$$

and so on do not depend on X_0.

Now consider, for some $n_0 > 1$ and any $n \ge n_0$, the narrower event

$$A_n := \left\{ X_0 + \sup_{j \ge n_0} S_{n,j} \le 0, \ \overline{S}_{n,\infty} := \sup_{j \ge -1} S_{n,j} = 0 \right\}$$

(we assume that the sequence $\{\xi_n\}$ is defined on the whole axis). Clearly, $A_n \subset B_n \subset \{X_{n+1} = 0\}$, so A_n is a renovating event as well. But, unlike B_n, the renovating event A_n is *stationary*: $A_n = T^n A_0$.

We assume now that $\mathbf{E}\xi_0 < 0$ and show that in this case $\mathbf{P}(A_0) > 0$ for sufficiently large n_0. In order to do this, we first establish that $\mathbf{P}(\overline{S}_{0,\infty} = 0) > 0$. Since, by the ergodic theorem, $S_{0,j} \xrightarrow{a.s.} -\infty$ as $j \to \infty$, we see that $\overline{S}_{0,\infty}$ is a proper random variable and there exists a v such that $\mathbf{P}(\overline{S}_{0,\infty} < v) > 0$. By the total probability formula,

$$0 < \mathbf{P}(\overline{S}_{0,\infty} < v) = \sum_{j=0}^{\infty} \mathbf{P}\left(\overline{S}_{0,j-1} < S_{0,j} < v, \ \sup_{k \ge j}(S_{0,k} - S_{0,j}) = 0\right).$$

Therefore there exists a j such that

$$\mathbf{P}\left(\sup_{k \ge j}(S_{0,k} - S_{0,j}) = 0\right) > 0.$$

But the supremum in the last expression has the same distribution as $\overline{S}_{0,\infty}$. This proves that $p := \mathbf{P}(\overline{S}_{0,\infty} = 0) > 0$. Next, since $S_{0,j} \xrightarrow{a.s.} -\infty$, one also has $\sup_{j \geq k} S_{0,j} \xrightarrow{a.s.} -\infty$ as $k \to \infty$. Therefore, $\mathbf{P}(\sup_{j \geq k} S_{0,j} < -X_0) \to 1$ as $k \to \infty$, and hence there exists an n_0 such that

$$\mathbf{P}\left(\sup_{j \geq n_0} S_{0,j} < -X_0 \right) > 1 - \frac{p}{2}.$$

Since $\mathbf{P}(AB) \geq \mathbf{P}(A) + \mathbf{P}(B) - 1$ for any events A and B, the aforesaid implies that $\mathbf{P}(A_0) \geq p/2 > 0$.

In the assertions below, we will use the existence of stationary renovating events A_n with $\mathbf{P}(A_0) > 0$ as a condition insuring convergence of the s.r.s. X_n to a stationary sequence. However, in the last example such convergence can be established directly. Let $\mathbf{E}\xi_0 < 0$. Then by (17.2.3), for any fixed v,

$$\mathbf{P}(X_{n+1} > v) = \mathbf{P}(\overline{S}_{n,n} > v) + \mathbf{P}(\overline{S}_{n,n} \leq v, X_0 + S_n > v),$$

where evidently

$$\mathbf{P}(X_0 + S_n > v) \to 0, \quad \mathbf{P}(\overline{S}_{n,n} > v) \uparrow \mathbf{P}(\overline{S}_{0,\infty} > v)$$

as $n \to \infty$. Hence the following limit exists

$$\lim_{n \to \infty} \mathbf{P}(X_n > v) = \mathbf{P}(\overline{S}_{0,\infty} > v). \tag{17.2.5}$$

Recall that in the above example the sequence of events A_n becomes renovating for $n \geq n_0$. But we can define other renovating events C_n along with a number m and function $g : \mathbb{R}^{m+1} \to \mathbb{R}$ as follows:

$$m := n_0, \qquad C_n := T^m A_n, \qquad g(y_0, \dots, y_m) := 0.$$

The events $C_n \in \mathfrak{F}^{\xi}_{n+m}$ are renovating for $\{X_n\}$ on the segment $[n, n+m]$ *for all* $n \geq 0$, so in this case the n_0 in the definition of a renovating sequence will be equal to 0.

A similar argument can also be used in the general case for arbitrary renovating events. Therefore we will assume in the sequel that the number n_0 from the definition of renovating events is equal to zero.

In the general case, the following assertion is valid.

Theorem 17.2.1 *Let $\{\xi_n\}$ be an arbitrary stationary sequence and for the s.r.s. $\{X_n\}$ there exists a sequence of renovating events $\{A_n\}$ such that*

$$\mathbf{P}\left(\bigcup_{j=1}^{n} A_j T^{-s} A_{j+s} \right) \to 1 \quad as\ n \to \infty \tag{17.2.6}$$

uniformly in $s \geq 1$. Then one can define, on a common probability space with $\{X_n\}$, a stationary sequence $\{X^n := U^n X^0\}$ satisfying the equations $X^{n+1} = f(X^n, \xi_n)$ and such that

$$\mathbf{P}\{X_k = X^k \text{ for all } k \geq n\} \to 1 \quad \text{as } n \to \infty. \tag{17.2.7}$$

If the sequence $\{\xi_n\}$ is metric transitive and the events A_n are stationary, then the relations $\mathbf{P}(A_0) > 0$ and $\mathbf{P}(\bigcup_{n=0}^{\infty} A_n) = 1$ are equivalent and imply (17.2.6) and (17.2.7).

Note also that if we introduce the measure $\pi(B) = \mathbf{P}(X^0 \in B)$ (as we did in Chap. 13), then (17.2.7) will imply convergence in total variation:

$$\sup_{B \in \mathcal{B}_{\mathcal{X}}} \left| \mathbf{P}(X_n \in B) - \pi(B) \right| \to 0 \quad \text{as } n \to \infty.$$

Proof of Theorem 17.2.1 First we show that (17.2.6) implies that

$$\mathbf{P}\left(\bigcap_{k=0}^{\infty} \{X_{n+k} \neq U^{-s} X_{n+k+s}\} \right) \to 0 \quad \text{as } n \to \infty \tag{17.2.8}$$

uniformly in $s \geq 0$. For a fixed $s \geq 1$, consider the sequence $X_j^s = U^{-s} X_{s+j}$. It is defined for $j \geq -s$, and

$$X_{-s}^s = X_0, \qquad X_{-s+1}^s = f(X_{-s}^s, \xi_{-s}) = f(X_0, \xi_{-s})$$

and so on. It is clear that the event

$$\{X_j = X_j^s \text{ for some } j \in [0, n]\}$$

implies the event

$$\{X + n + k = X_{n+k}^s \text{ for all } k \geq 0\}.$$

We show that

$$\mathbf{P}\left(\bigcup_{j=1}^{n} \{X_j = X_j^s\} \right) \to 1 \quad \text{as } n \to \infty.$$

For simplicity's sake put $m = 0$. Then, for the event $X_{j+1} = X_{j+1}^s$ to occur, it suffices that the events A_j and $T^{-s} A_{j+s}$ occur simultaneously. In other words,

$$\bigcup_{j=0}^{n-1} A_j T^{-s} A_{j+s} \subset \bigcup_{j=1}^{n} \{X_j = X_j^s\} \subset \bigcap_{k=0}^{\infty} \{X_{n+k} = X_{n+k}^s\}.$$

Therefore (17.2.6) implies (17.2.8) and convergence

$$\mathbf{P}\left(X_k^n \neq X_k^{n+s} \right) \to 0 \quad \text{as } n \to \infty$$

uniformly in $k \geq 0$ and $s \geq 0$. If we introduce the metric ρ putting $\rho(x, y) := 1$ for $x \neq y$, $\rho(x, x) = 0$, then the aforesaid means that, for any $\delta > 0$, there exists an N such that

$$\mathbf{P}\left(\rho\left(X_k^n, X_k^{n+s}\right) > \delta\right) = \mathbf{P}\left(\rho\left(X_k^n, X_k^{n+s}\right) \neq 0\right) < \delta$$

for $n \geq N$ and any $k \geq 0$, $s \geq 0$, i.e. X_k^n is a Cauchy sequence with respect to convergence in probability for each k. Because any space $\mathcal{X}$ is complete with such a metric, there exists a random variable X^k such that $X_k^n \xrightarrow{p} X^k$ as $n \to \infty$ (see Lemma 4.2). Due to the specific nature of the metric ρ this means that

$$\mathbf{P}\left(X_k^n \neq X^k\right) \to 0 \quad \text{as } n \to \infty. \tag{17.2.9}$$

The sequence X^k is stationary. Indeed, as $n \to \infty$,

$$\mathbf{P}\left(X^{k+1} \neq U X^k\right) = \mathbf{P}\left(X_{k+1}^n \neq U X_k^n\right) + o(1) = \mathbf{P}\left(X_{k+1}^n \neq X_{k+1}^{n-1}\right) + o(1) = o(1).$$

Since the probability $\mathbf{P}(X^{k+1} \neq U X^k)$ does not depend on n, $X^{k+1} = U X^k$ a.s.

Further, $X_{n+k+1} = f(X_{n+k}, \xi_{n+k})$, and therefore

$$X_{k+1}^n = U^{-n} f(X_{n+k}, \xi_{n+k}) = f\left(X_k^n, \xi_k\right). \tag{17.2.10}$$

The left and right-hand sides here converge in probability to X^{k+1} and $f(X^k, \xi_k)$, respectively. This means that $X^{k+1} = f(X^k, \xi_k)$.

To prove convergence (17.2.7) it suffices to note that, by virtue of (17.2.10), the values X_k^n and X^k, after having become equal for some k, will never be different for greater values of k. Therefore, as well as (17.2.9) one has the relation

$$\mathbf{P}\left(\bigcup_{k \geq 0}\{X_k^n \neq X^k\}\right) = \mathbf{P}\left(\bigcup_{k \geq 0}\{X_{k+n} \neq X^{k+n}\}\right) \to 0 \quad \text{as } n \to \infty,$$

which is equivalent to (17.2.7).

The last assertion of the theorem follows from Theorem 16.2.5. The theorem is proved. $\qquad \square$

Remark 17.2.1 It turns out that condition (17.2.6) is also a necessary one for convergence (17.2.7) (see [6]). For more details on convergence of stochastic recursive sequences and their generalisations, and also on the relationship between (17.2.6) and conditions (I) and (II) from Chap. 13, see [6].

In Example 17.2.1 the sequence X^k was actually found in an explicit form (see (17.2.3) and (17.2.5)):

$$X^k = \overline{S}_{k,-\infty} = \sup_{j \geq 0} S_{k-1}^j. \tag{17.2.11}$$

These random variables are proper by Corollary 16.3.1. It is not hard to also see that, for $X_0 = 0$, one has (see (17.2.3))

$$U^{-1}X_{n+k} \uparrow X^k. \tag{17.2.12}$$

17.2.2 Boundedness of Random Sequences

Consider now conditions of boundedness of an s.r.s. in spaces $\mathcal{X} = [0, \infty)$ and $\mathcal{X} = (-\infty, \infty)$. Assertions about boundedness will be stated in terms of existence of stationary majorants, i.e. stationary sequences M_n such that

$$X_n \le M_n \quad \text{for all } n.$$

Results of this kind will be useful for constructing stationary renovating sequences.

Majorants will be constructed for a class of random sequences more general than stochastic recursive sequences. Namely, we will consider the class of random sequences satisfying the inequalities

$$X_{n+1} \le \big(X_n + h(X_n, \xi_n)\big)^+, \tag{17.2.13}$$

where the measurable function h will in turn be bounded by rather simple functions of X_n and ξ_n. The sequence $\{\xi_n\}$ will be assumed given on the whole axis.

Theorem 17.2.2 *Assume that there exist a number $N > 0$ and a measurable function g_1 with $\mathbf{E}g_1(\xi_n) < 0$ such that (17.2.13) holds with*

$$h(x, y) \le \begin{cases} g_1(y) & \text{for } x > N, \\ g_1(y) + N - x & \text{for } x \le N. \end{cases} \tag{17.2.14}$$

If $X_0 \le M < \infty$, then the stationary sequence

$$M_n = \max(M, N) + \sup_{j \ge -1} S_{n-1, j}, \tag{17.2.15}$$

where $S_{n,-1} = 0$ and $S_{k,j} = g_1(\xi_k) + \cdots + g_1(\xi_{k-j})$ for $j \ge 0$, is a majorant for X_n.

Proof For brevity's sake, put $\zeta_i := g_1(\xi_i)$, $Z := \max(M, N)$, and $Z_n := X_n - Z$. Then Z_n will satisfy the following inequalities:

$$Z_{n+1} \le \begin{cases} (Z_n + Z + \zeta_n)^+ - Z \le (Z_n + \zeta_n)^+ & \text{for } Z_n > N - Z, \\ (N + \zeta_n)^+ - Z \le \zeta_n^+ & \text{for } Z_n \le N - Z. \end{cases}$$

Consider now a sequence $\{Y_n\}$ defined by the relations $Y_0 = 0$ and

$$Y_{n+1} = (Y_n + \zeta_n)^+.$$

Assume that $Z_n \leq Y_n$. If $Z_n > N - Z$ then

$$Z_{n+1} \leq (Z_n + \zeta_n)^+ \leq (Y_n + \zeta_n)^+ = Y_{n+1}.$$

If $Z_n \leq N - Z$ then

$$Z_{n+1} \leq \zeta_n^+ \leq (Y_n + \zeta_n)^+ = Y_{n+1}.$$

Because $Z_0 \leq 0 = Y_0$, it is evident that $Z_n \leq Y_n$ for all n. But we know the solution of the equation for Y_n and, by virtue of (17.2.11) and (17.2.13),

$$X_n - Z \leq \sup_{j \geq -1} S_{n-1, j}.$$

The theorem is proved. $\qquad\qquad\square$

Theorem 17.2.2A *Assume that there exist a number $N > 0$ and measurable functions g_1 and g_2 such that*

$$\mathbf{E}g_1(\xi_n) < 0, \qquad \mathbf{E}g_2(\xi_n) < 0 \qquad\qquad (17.2.16)$$

and

$$h(x, y) \leq \begin{cases} g_1(y) & \text{for } x > N, \\ g_1(y) + g_2(y) & \text{for } x \leq N. \end{cases} \qquad (17.2.17)$$

If $Z_0 \leq M < \infty$, then the conditions of Theorem 17.2.2 are satisfied (possibly for other N and g_1) and for X_n there exists a stationary majorant of the form (17.2.15).

Proof We set $g := -\mathbf{E}g_1(\xi_n) > 0$ and find $L > 0$ such that $\mathbf{E}(g_2(\xi_n); \, g_2(\xi_n) > L) \leq g/2$. Introduce the function

$$g_1^*(y) := g_1(y) + g_2(y)\mathrm{I}\big(g_2(y) > L\big).$$

Then $\mathbf{E}g_1^*(\xi_n) \leq -g/2 < 0$ and

$$\begin{aligned} h(x, y) &\leq g_1(y) + g_2(y)\mathrm{I}(x \leq N) \\ &\leq g_1^*(y) + g_2(y)\mathrm{I}(x \leq N) - g_2(y)\mathrm{I}\big(g_2(y) > L\big) \\ &\leq g_1^*(y) + L\mathrm{I}(x \leq N) \leq g_1^*(y) + (L + N - x)\mathrm{I}(x \leq N) \\ &\leq g_1^*(y) + (L + N - x)\mathrm{I}(x \leq L + N). \end{aligned}$$

This means that inequalities (17.2.14) hold with N replaced with $N^* = N + L$. The theorem is proved. $\qquad\qquad\square$

Note again that in Theorems 17.2.2 and 17.2.2A we did not assume that $\{X_n\}$ is an s.r.s.

The reader will notice the similarity of the conditions of Theorems 17.2.2 and 17.2.2A to the boundedness condition in Sect. 15.5, Theorem 13.7.3 and Corollary 13.7.1.

The form of the assertions of Theorems 17.2.2 and 17.2.2A enables one to construct stationary renovating events for a rather wide class of nonnegative stochastic recursive sequences (so that $\mathcal{X} = [0, \infty)$) having, say, a "positive atom" at 0. It is convenient to write such sequences in the form

$$X_{n+1} = \big(X_n + h(X_n, \xi_n)\big)^+. \tag{17.2.18}$$

Example 17.2.2 Let an s.r.s. (see (17.1.1)) be described by Eq. (17.2.18) and satisfy conditions (17.2.14) or (17.2.17), where the function h is sufficiently "regular" to ensure that

$$B_{n,T} = \bigcap_{t \le T} \big\{ h(t, \xi_n) \le -t \big\}$$

is an event for any T. (For instance, it is enough to require $h(t, v)$ to have at most a countable set of discontinuity points t. Then the set $B_{n,T}$ can be expressed as the intersection of countably many events $\bigcap_k \{ h(t_k, \xi_n) \le -t_k \}$, where $\{t_k\}$ form a countable set dense on $[0, T]$.) Furthermore, let there exist an $L > 0$ such that

$$\mathbf{P}(M_n < L, B_{n,L}) > 0 \tag{17.2.19}$$

(M_n was defined in (17.2.15)). Then the event $A_n = \{M_n < L\} B_{n,L}$ is clearly a positive stationary renovating event with the function $g(y) = (h(0, y))^+$, $m = 0$. (On the set $A_n \in \mathfrak{F}_n^\xi$ we have $X_{n+1} = 0$, $X_{n+2} = h(0, \xi_{n+1})^+$ and so on.) Therefore, an s.r.s. satisfying (17.2.18) satisfies the conditions of Theorem 17.2.1 and is ergodic in the sense of assertion (17.2.7).

It can happen that, from a point $t \le L$, it would be impossible to reach the point 0 in one step, but it could be done in $m > 1$ steps. If B is the set of sequences $(\xi_n, \dots, \xi_{n+m})$ that effect such a transition, and $\mathbf{P}(M_n < L)$, then $A_n = \{M_n < L\} B$ will also be stationary renovating events.

17.3 Ergodicity Conditions Related to the Monotonicity of f

Now we consider ergodicity conditions for stochastic recursive sequences that are related to the analytic properties of the function f from (17.1.1). As we already noted, the sequence $f(x, \xi_k)$, $k = 1, 2, \dots$, may be considered as a sequence of random transformations of the space $\mathcal{X}$. Relation (17.1.1) shows that X_{n+1} is the result of the application of $n + 1$ random transformations $f(\cdot, \xi_k)$, $k = 0, 1, \dots, n$, to the initial value $X_0 = x \in \mathcal{X}$. Denoting by ξ_n^{n+k} the vector $\xi_n^{n+k} = (\xi_n, \dots, \xi_{n+k})$

and by f_k the k-th iteration of the function f: $f_1(x, y_1) = f(x, y_1)$, $f_2(x, y_1, y_2) = f(f(x, y_1), y_2)$ and so on, we can re-write (17.1.1) for $X_0 = x$ in the form

$$X_{n+1} = X_{n+1}(x) = f_{n+1}(x, \xi_0^n),$$

so that the "forward" and "backward" equations hold true:

$$f_{n+1}(x, \xi_0^n) = f(f_n(x, \xi_0^{n-1}), \xi_n) = f_n(f(x, \xi_0), \xi_1^n).   \qquad (17.3.1)$$

In the present section we will be studying stochastic recursive sequences for which the function f from representation (17.1.1) is monotone in the first argument. To this end, we need to assume that a partial order relation "$\geq$" is defined in the space $\mathcal{X}$. In the space $\mathcal{X} = \mathbb{R}^d$ of vectors $x = (x(1), \ldots, x(d))$ (or its subspaces) the order relation can be introduced in a natural way by putting $x_1 \geq x_2$ if $x_1(k) \geq x_2(k)$ for all k.

Furthermore, we will assume that, for each non-decreasing sequence $x_1 \leq x_2 \leq \cdots \leq x_n \leq \ldots$, there exists a limit $x \in \mathcal{X}$, i.e. the smallest element $x \in \mathcal{X}$ for which $x_k \leq x$ for all k. In that case we will write $x_k \uparrow x$ or $\lim_{k \to \infty} x_k = x$. In $\mathcal{X} = \mathbb{R}^d$ such convergence will mean conventional convergence. To facilitate this, we will need to complete the space $\mathbb{R}^d$ by adding points with infinite components.

Theorem 17.3.1 (Loynes) *Suppose that the transformation $f = f(x, y)$ and space $\mathcal{X}$ satisfy the following conditions*:

(1) *there exists an $x_0 \in \mathcal{X}$ such that $f(x_0, y) \geq x_0$ for all $y \in \mathcal{Y}$;*
(2) *the function f is monotone in the first argument: $f(x_1, y) \geq f(x_2, y)$ if $x_1 \geq x_2$;*
(3) *the function f is continuous in the first argument with respect to the above convergence: $f(x_n, y) \uparrow f(x, y)$ if $x_n \uparrow x$.*

Then there exists a stationary random sequence $\{X_n\}$ satisfying Eq. (17.1.1): $X^{n+1} = U X^n = f(X^n, \xi_n)$, such that

$$U^{-n} X_{n+s}(x)) \uparrow X^s \quad as \ n \to \infty, \qquad (17.3.2)$$

where convergence takes place for all elementary outcomes.

Since the distributions of X_n and $U^{-n} X_n$ coincide, in the case where convergence of random variables $\eta_n \uparrow \eta$ means convergence (in a certain sense) of their distributions (as is the case when $\mathcal{X} = \mathbb{R}^d$), Theorem 17.2.1 also implies convergence of the distributions of X_n to that of X^0 as $n \to \infty$.

Remark 17.3.1 A substantial drawback of this theorem is that it holds only for a single initial value $X_0 = x_0$. This drawback disappears if the point x_0 is accessible with probability 1 from any $x \in \mathcal{X}$, and ξ_k are independent. In that case x_0 is likely to be a positive atom, and Theorem 13.6.1 for Markov chains is also applicable.

The limiting sequence X^s in (17.3.2) can be "improper" (in spaces $\mathcal{X} = \mathbb{R}^d$ it may assume infinite values). The sequence X^s will be proper if the s.r.s. X_n satisfies, say, the conditions of the theorems of Sect. 15.5 or the conditions of Theorem 17.2.2.

Proof of Theorem 17.3.1 Put

$$v_s^{-k} := f_{k+s}\big(x_0, \xi_{-k}^{s-1}\big) = U^{-k} f_{k+s}\big(x_0, \xi_0^{s+k-1}\big) = U^{-k} X_{k+s}(x_0).$$

Here the superscript $-k$ indicates the number of the element of the driving sequence $\{\xi_n\}_{n=-\infty}^{\infty}$ such that the elements of this sequence starting from that number are used for constructing the s.r.s. The subscript s is the "time epoch" at which we observe the value of the s.r.s. From the "backward" equation in (17.3.1) we get that

$$v_s^{-k-1} = f_{k+s}\big(f(x_0, \xi_{-k-1}), \xi_{-k}^{s-1}\big) \ge f_{k+s}\big(x_0, \xi_{-k}^{s-1}\big) = v_s^{-k}.$$

This means that the sequence v_s^{-k} increases as k grows, and therefore there exists a random variable $X^s \in \mathcal{X}$ such that

$$v_s^{-k} = U^{-k} X_{k+s}(x_0) \uparrow X^s \quad \text{as } k \to \infty.$$

Further, v_s^{-k} is a function of ξ_{-k}^{s-1}. Therefore, X^s is a function of $\xi_{-\infty}^{s-1}$:

$$X^s = G\big(\xi_{-\infty}^{s-1}\big).$$

Hence

$$U X^s = U G\big(\xi_{-\infty}^{s-1}\big) = G\big(\xi_{-\infty}^{s}\big) = X^{s+1},$$

which means that $\{X^s\}$ is stationary. Using the "forward" equation from (17.3.1), we obtain that

$$v_s^{-k-1} = f\big(f_{k+s}\big(x_0, \xi_{-k-1}^{s-2}\big), \xi_{s-1}\big) = f\big(v_{s-1}^{-k-1}, \xi_{s-1}\big).$$

Passing to the limit as $k \to \infty$ gives, since f is continuous, that

$$X^s = f\big(X^{s-1}, \xi_{s-1}\big).$$

The theorem is proved.   □

Example 17.2.1 clearly satisfies all the conditions of Theorem 17.3.1 with $\mathcal{X} = [0, \infty)$, $x_0 = 0$, and $f(x, y) = (x + y)^+$.

17.4 Ergodicity Conditions for Contracting in Mean Lipschitz Transformations

In this section we will assume that $\mathcal{X}$ is a complete separable metric space with metric ρ. Consider the following conditions on the iterations $X_k(x) = f_k(x, \xi_0^{k-1})$.

Condition (B) (boundedness). *For some $x_0 \in \mathcal{X}$ and any $\delta > 0$, there exists an $N = N_\delta$ such that, for all $n \geq 1$,*

$$\mathbf{P}\big(\rho(x_0, X_n(x_0) > N)\big) = \mathbf{P}\big(\rho(x_0, f_n(x_0, \xi_0^{n-1})) > N\big) < \delta.$$

It is not hard to see that condition **(B)** holds (possibly with a different N) as soon as we can establish that, for some $m \geq 1$, the above inequality holds for all $n \geq m$.

Condition **(B)** is clearly met for stochastic random sequences satisfying the conditions of Theorems 17.2.2 and 17.2.2A or the theorems of Sect. 15.5.

Condition (C) (contraction in mean). *The function f is continuous in the first argument and there exist $m \geq 1$, $\beta > 0$ and a measurable function $q : \mathbb{R}^m \to \mathbb{R}$ such that, for any x_1 and x_2,*

$$\rho\big(f_m(x_1, \xi_0^{m-1}), f_m(x_2, \xi_0^{m-1})\big) \leq q\big(\xi_0^{m-1}\big)\rho(x_1, x_2),$$
$$m^{-1}\mathbf{E}\ln q\big(\xi_0^{m-1}\big) \leq -\beta < 0.$$

Observe that conditions **(B)** and **(C)** are, generally speaking, not related to each other. Let, for instance, $\mathcal{X} = \mathbb{R}$, $X_0 \geq 0$, $\xi_n \geq 0$, $\rho(x, y) = |x - y|$, and $f(x, y) = bx + y$, so that

$$X_{n+1} = bX_n + \xi_n.$$

Then condition **(C)** is clearly satisfied for $0 < b < 1$, since

$$\big|f(x_1, y) - f(x_2, y)\big| = b|x_1 - x_2|.$$

At the same time, condition **(B)** will be satisfied if and only if $\mathbf{E}\ln \xi_0 < \infty$. Indeed, if $\mathbf{E}\ln \xi_0 = \infty$, then the event $\{\ln \xi_k > -2k \ln b\}$ occurs infinitely often a.s. But X_{n+1} has the same distribution as

$$b^{n+1}X_0 + \sum_{k=0}^{n} b^k \xi_k = b^{n+1}X_0 + \sum_{k=0}^{n} \exp\{k \ln b + \ln \xi_k\},$$

where, in the sum on the right-hand side, the number of terms exceeding $\exp\{-k \ln b\}$ increases unboundedly as n grows. This means that $X(n+1) \xrightarrow{p} \infty$ as $n \to \infty$. The case $\mathbf{E}\ln \xi_0 < \infty$ is treated in a similar way. The fact that **(B)**, generally speaking, does not imply **(C)** is obvious.

As before, we will assume that the "driving" stationary sequence $\{\xi_n\}_{n=-\infty}^{\infty}$ is given on the whole axis. Denote by U the respective distribution preserving shift operator.

Convergence in probability and a.s. of a sequence of $\mathcal{X}$-valued random variables $\eta_n \in \mathcal{X}$ ($\eta_n \xrightarrow{p} \eta$, $\eta_n \xrightarrow{a.s.} \eta$) is defined in the natural way by the relations $\mathbf{P}(\rho(\eta_n, \eta) > \delta) \to 0$ as $n \to \infty$ and $\mathbf{P}(\rho(\eta_k, \eta) > \delta$ for some $k \geq n) \to 0$ as $n \to \infty$ for any $\delta > 0$, respectively.

Theorem 17.4.1 *Assume that conditions* **(B)** *and* **(C)** *are met. Then there exists a stationary sequence* $\{X_n\}$ *satisfying* (17.1.1):

$$X^{n+1} = U X^n = f\left(X^n, \xi_n\right)$$

such that, for any fixed x,

$$U^{-n} X_{n+s}(x) \xrightarrow{a.s.} X^s \quad as \quad n \to \infty. \tag{17.4.1}$$

This convergence is uniform in x *over any bounded subset of* $\mathcal{X}$.

Theorem 17.2.2 implies the weak convergence, as $n \to \infty$, of the distributions of $X_n(x)$ to that of X^0. Condition **(B)** is clearly necessary for ergodicity. As the example of a generalised autoregressive process below shows, condition **(C)** is also necessary in some cases.

Set $Y_n := U^n X_n(x_0)$, where x_0 is from condition **(B)**. We will need the following auxiliary result.

Lemma 17.4.1 *Assume that conditions* **(B)** *and* **(C)** *are met and the stationary sequence* $\{q(\xi_{km}^{km+m-1})\}_{k=-\infty}^{\infty}$ *is ergodic. Then, for any* $\delta > 0$, *there exists an* n_δ *such that, for all* $k \geq 0$,

$$\sup_{k \geq 0} \mathbf{P}\left(\rho(Y_{n+k}, Y_n) < \delta \text{ for all } n \geq n_\delta\right) \geq 1 - \delta. \tag{17.4.2}$$

For ergodicity of $\{q(\xi^{km+m-1})\}_{k=-\infty}^{\infty}$ *it suffices that the transformation* T^m *is metric transitive.*

The lemma means that, with probability 1, the distance $\rho(Y_{n+k}, Y_n)$ tends to zero uniformly in k as $n \to \infty$. Relation (17.4.2) can also be written as $\mathbf{P}(A_\delta) \leq \delta$, where

$$A_\delta := \bigcup_{n \geq n_\delta} \left\{\rho(Y_{n+k}, Y_n) \geq \delta\right\}.$$

Proof of Lemma 17.4.1 By virtue of condition **(B)**, there exists an $N = N_\delta$ such that, for all $k \geq 1$,

$$\mathbf{P}\left(\rho(x_0, X_k(x_0)) > N\right) \leq \frac{\delta}{4}.$$

Hence

$$\mathbf{P}(A_\delta) \leq \delta/3 + \mathbf{P}\left(A_\delta; \rho(x_0, \theta_{n,k}) \leq N\right).$$

The random variable $\theta_{n,k} := U^{-n-k} X_k(x_0)$ has the same distribution as $X_k(x_0)$. Next, by virtue of **(C)**,

$$\rho(Y_{n+k}, Y_n) \leq \rho\left(f_{n+k}(x_0, \xi_{-n-k}^{-1}), f_n(x_0, \xi_{-n}^{-1})\right)$$

$$\leq q\big(\xi_{-m}^{-1}\big)\rho\big(f_{n+k-m}\big(x_0,\xi_{-n-k}^{-m-1}\big),\,f_{n-m}\big(x_0,\xi_{-n}^{-m-1}\big)\big)$$

$$= q\big(\xi_{-m}^{-1}\big)\rho\big(U^{-n-k}X_{n+k-m}(x_0),\,U^{-n}X_{n-m}(x_0)\big). \qquad (17.4.3)$$

Denote by B_s the set of numbers n of the form $n = lm + s$, $l = 0, 1, 2, \ldots$, $0 \leq s < m$, and put

$$\lambda_j := \ln q\big(\xi_{-jm}^{-jm+m-1}\big), \quad j = 1, 2, \ldots.$$

Then, for $n \in B_s$, we obtain from (17.4.3) and similar relations that

$$\rho(Y_{n+k}, Y_n) \leq \exp\left\{\sum_{j=1}^{l}\lambda_j\right\}\rho\big(U^{-n-k}X_{k+s}(x_0),\,U^{-n}X_s(x_0)\big), \qquad (17.4.4)$$

where the last factor (denote it just by ρ) is bounded from above:

$$\rho \leq \rho\big(x_0, U^{-n-k}X_{k+s}(x_0)\big) + \rho\big(x_0, U^{-n}X_s(x_0)\big).$$

The random variables $U^{-n}X_j(x_0)$ have the same distribution as $X_j(x_0)$. By virtue of (**B**), there exists an $N = N_\delta$ such that, for all $j \geq 1$,

$$\mathbf{P}\big(\rho\big(x_0, X_j(x_0)\big) > N\big) \leq \frac{\delta}{4m}.$$

Hence, for all n, k and s, we have $\mathbf{P}(\rho > 2N) < \delta/(2m)$, and the right-hand side of (17.4.4) does not exceed $2N \exp\{\sum_{j=1}^{l}\lambda_j\}$ on the complement set $\{\rho \leq 2N\}$.

Because $\mathbf{E}\lambda_j \leq -m\beta < 0$ and the sequence $\{\lambda_j\}$ is metric transitive, by the ergodic Theorem 16.3.1 we have

$$\sum_{j=1}^{l}\lambda_j < -m\beta l/2$$

for all $l \geq l(\omega)$, where $l(\omega)$ is a proper random variable. Choose l_1 and l_2 so that the inequalities

$$-\frac{m\beta l_1}{2} < \ln\delta - \ln 2N, \qquad \mathbf{P}\big(l(\omega) > l_2\big) < \frac{\delta}{2}$$

hold. Then, putting

$$l_\delta := \max(l_1, l_2), \qquad n_\delta := ml_\delta, \qquad A_\delta^s := \bigcup_{n \geq n_\delta,\,n \in B_s}\{\rho(Y_{n+k}, Y_n) \geq \delta\},$$

we obtain that

$$\mathbf{P}\big(A_\delta^s\big) \leq \mathbf{P}(\rho > 2N) + \mathbf{P}\big(A_\delta^s; \rho \leq N\big) \leq \frac{\delta}{2m} + \mathbf{P}\left(\bigcup_{l \geq l_\delta}\left\{2N \exp\left\{-\sum_{j=0}^{l}\lambda_j\right\} \geq \delta\right\}\right).$$

But the intersection of the events from the term with $\{l_\delta \geq l(\omega)\}$ is empty. Therefore, the former event is a subset of the event $\{l(\omega) > l_\delta\}$, and

$$\mathbf{P}\big(A_\delta^s\big) \leq \frac{\delta}{m}, \quad \mathbf{P}(A_\delta) \leq \sum_{s=0}^{m-1} \mathbf{P}\big(A_\delta^s\big) \leq \delta.$$

The lemma is proved.   □

Lemma 17.4.2 (Completeness of $\mathcal{X}$ with respect to convergence in probability) *Let $\mathcal{X}$ be a complete metric space. If a sequence of $\mathcal{X}$-valued random elements η_n is such that, for any $\delta > 0$,*

$$P_n := \sup_{k \geq 0} \mathbf{P}\big(\rho(\eta_{n+k}, \eta_n) > \delta\big) \to 0$$

as $n \to \infty$, then there exists a random element $\eta \in \mathcal{X}$ such that $\eta \overset{p}{\to} \eta$ (that is, $\mathbf{P}(\rho(\eta_n, \eta) > \delta) \to 0$ as $n \to \infty$).

Proof For given ε and δ choose n_k, $k = 0, 1, \ldots$, such that

$$\sup_s \mathbf{P}\big(\rho(\eta_{n_k+s}, \eta_{n_k}) > 2^{-k}\delta\big) < \varepsilon 2^{-k},$$

and, for the sake of brevity, put $\zeta_k := \eta_{n_k}$. Consider the set

$$D := \bigcap_{k=0}^{\infty} D_k, \quad D_k := \big\{\omega \, \rho(\zeta_{k+1}, \zeta_k) \leq 2^{-k}\delta\big\}.$$

Then $\mathbf{P}(D) > 1 - 2\varepsilon$ and, for any $\omega \in D$, one has $\rho(\zeta_{k+s}(\omega), \zeta_k(\omega)) < \delta 2^{k-1}$ for all $s \geq 1$. Hence $\zeta_k(\omega)$ is a Cauchy sequence in $\mathcal{X}$ and there exists an $\eta = \eta(\omega) \in \mathcal{X}$ such that $\zeta_k(\omega) \to \eta(\omega)$. Since ε is arbitrary, this means that $\zeta_k \xrightarrow{a.s.} \eta$ as $k \to \infty$, and

$$\mathbf{P}\big(\rho(\zeta_0, \eta) > 2\delta\big) \leq \mathbf{P}\left(\bigcup_{k=0}^{\infty} \rho(\zeta_{k+1}, \zeta_k) > 2^{-k}\delta\right)$$

$$\leq \sum_{k=0}^{\infty} \mathbf{P}\big(\rho(\zeta_{k+1}, \zeta_k) > 2^{-k}\delta\big) \leq 2\varepsilon.$$

Therefore, for any $n \geq n_0$,

$$\mathbf{P}\big(\rho(\eta_n, \eta) > 3\delta\big) \leq \mathbf{P}\big(\rho(\eta_n, \eta_{n_0}) > \delta\big) + \mathbf{P}\big(\rho(\zeta_0, \eta) > 2\delta\big) \leq 3\varepsilon.$$

Since ε and δ are arbitrary, the lemma is proved.   □

Proof of Theorem 17.4.1 From Lemma 17.4.1 it follows that

$$\sup_k \mathbf{P}\big(\rho(Y_{n+k}, Y_n) > \delta\big) \to 0 \quad \text{as } n \to \infty.$$

This means that Y_n is a Cauchy sequence with respect to convergence in probability, and by Lemma 17.4.2 there exists a random variable X^0 such that

$$Y_n \xrightarrow{P} X^0,$$

$$U^{-n} X_{n+s}(x_0) = U^s\big(U^{-n-s} X_{n+s}(x_0)\big) = U^s Y_{n+s} \to U^s X^0 \equiv X^s. \tag{17.4.5}$$

By continuity of f,

$$U^{-n} X_{n+s+1}(x_0) = U^{-n} f\big(X_{n+s}(x_0), \xi_{n+s}\big)$$

$$= f\big(U^{-n} X_{n+s}(x_0), \xi_s\big) \xrightarrow{P} f\big(X^s, \xi_s\big) = X^{s+1}.$$

We proved the required convergence for a fixed initial value x_0. For an arbitrary $x \in C_n = \{z : \rho(x_0, z) \le N\}$, one has

$$\rho\big(U^{-n} X_n(x), X^0\big) \le \rho\big(U^{-n} X_n(x), U^{-n} X_n(x_0)\big) + \rho\big(U^{-n} X_n(x_0), X^0\big), \tag{17.4.6}$$

where the first term on the right-hand side converges in probability to 0 uniformly in $x \in C_N$. For $n = lm$ this follows from the inequality (see condition **(C)**)

$$\rho\big(U^{-n} X_n(x), U^{-n} X_n(x_0)\big) \le N \exp\left\{\sum_{j=1}^{l} \lambda_j\right\} \tag{17.4.7}$$

and the above argument. Similar relations hold for $n = lm + s$, $m > s > 0$. This, together with (17.4.5) and (17.4.6), implies that

$$U^{-n} X_{n+s}(x) \xrightarrow{P} X^s = U^s X^0$$

uniformly in $x \in C_N$. This proves the assertion of the theorem in regard to convergence in probability.

We now prove convergence with probability 1. To this end, one should repeat the argument proving Lemma 17.4.1, but bounding $\rho(X^0, U^{-n} X_n(x))$ rather than $\rho(Y_{n+k}, Y_n)$. Assuming for simplicity's sake that $s = 0$ (n is a multiple of m), we get (similarly to (17.4.4)) that, for any x,

$$\rho\big(X^0, U^{-n} X_n(x)\big) \le \rho\big(x, U^{-n} X^0\big) \exp\left\{\sum_{j=1}^{l} \lambda_j\right\}. \tag{17.4.8}$$

The rest of the argument of Lemma 17.4.1 remains unchanged. This implies that, for any $\delta > 0$ and sufficiently large n_δ,

$$\mathbf{P}\left(\bigcup_{n \ge n_\delta} \{\rho\big(X^0, U^{-n} X_n(x)\big) > \delta\}\right) < \delta.$$

Theorem 17.4.1 is proved.   □

Example 17.4.1 (Generalised autoregression) Let $\mathcal{X} = \mathbb{R}$. A generalised autoregression process is defined by the relations

$$X_{n+1} = G\big(\zeta_n F(X_n) + \eta_n\big), \tag{17.4.9}$$

where F and G are functions mapping $\mathbb{R} \mapsto \mathbb{R}$ and $\xi_n = (\zeta_n, \eta_n)$ is a stationary ergodic driving sequence, so that $\{X_n\}$ is an s.r.s. with the function

$$f(x, y) = G\big(y_1, F(x) + y_2\big), \quad y = (y_1, y_2) \in \mathcal{Y} = \mathbb{R}^2.$$

If the functions F and G are nondecreasing and left continuous, $G(x) \geq 0$ for all $x \in \mathbb{R}$, and the elements ζ_n are nonnegative, then the process (17.4.9) satisfies the condition of Theorem 17.3.1, and therefore $U^{-n+s} X_n(0) \uparrow X^s$ with probability 1 (as $n \to \infty$). To establish convergence to a proper stationary sequence X^s, one has to prove uniform boundedness in probability (in n) of the sequence $X_n(0)$ (see below).

Now we will establish under what conditions the sequence (17.4.9) will satisfy the conditions of Theorem 17.4.1. Suppose that the functions F and G satisfy the Lipschitz condition:

$$\big|G(x_1) - G(x_2)\big| \leq c_G|x_1 - x_2|, \qquad \big|F(x_1) - F(x_2)\big| \leq c_F|x_1 - x_2|.$$

Then

$$\big|f(x_1, \xi_0) - f(x_2, \xi_0)\big| \leq c_G\big|\zeta_0\big(F(x_1) - F(x_2)\big)\big| \leq c_F c_G|\zeta_0||x_1 - x_2|. \tag{17.4.10}$$

Theorem 17.4.2 *Under the above assumptions, the sequence* (17.4.9) *will satisfy condition* (**C**) *if*

$$\ln c_G c_F + \mathbf{E}\ln|\zeta_0| < 0. \tag{17.4.11}$$

The sequence (17.4.9) *will satisfy condition* (**B**) *if* (17.4.11) *holds and, moreover,*

$$\mathbf{E}\big(\ln|\eta_0|\big)^+ < \infty. \tag{17.4.12}$$

When (17.4.11) *and* (17.4.12) *hold, the sequence* (17.4.9) *has a stationary majorant, i.e. there exists a stationary sequence* M_n *(depending on X_0) such that* $|X_n| \leq M_n$ *for all n.*

Proof That condition (**C**) for $\rho(x_1, x_2) = |x_1 - x_2|$ follows from (17.4.10) is obvious. We prove (**B**). To do this, we will construct a stationary majorant for $|X_n|$. One could do this using Theorems 17.2.2 and 17.2.2A. In our case, it is simpler to prove it directly, making use of the inequalities

$$\big|G(x)\big| \leq \big|G(0)\big| + c_G|x|, \qquad \big|F(x)\big| \leq \big|F(0)\big| + c_F|x|,$$

where we assume, for simplicity's sake, that $G(0)$ and $F(0)$ are finite. Then

$$|X_{n+1}| \le |G(0)| + c_G|\zeta_n| \cdot |F(X_n)| + c_G|\eta_n|$$
$$\le |G(0)| + c_G c_F|\zeta_n| \cdot |X_n| + c_G|\zeta_n| \cdot |F(0)| + c_G|\eta_n| = \beta_n|X(n)| + \gamma_n,$$

where

$$\beta_n := c_G c_F|\zeta_n| \ge 0, \qquad \gamma_n := |G(0)| + c_G|\zeta_n| \cdot |F(0)| + c_G|\eta_n|$$

$$\mathbf{E}\ln\beta_n < 0, \quad \mathbf{E}(\ln\gamma_n)^+ < \infty.$$

From this we get that, for $X_0 = x$,

$$|X_{n+1}| \le |x| \prod_{j=0}^{n} \beta_j + \sum_{l=0}^{n-1} \left(\prod_{j=n-l}^{n} \beta_j \right) \gamma_{n-l-1} + \gamma_n,$$

$$U^{-n}|X_{n+1}| \le |x| \prod_{j=-n}^{0} \beta_j + \sum_{l=0}^{\infty} \left(\prod_{j=-l}^{0} \beta_j \right) \gamma_{-l-1} + \gamma_0. \tag{17.4.13}$$

Put

$$\alpha_i := \ln\beta_j, \qquad S_l := \sum_{j=-l}^{0} \alpha_j.$$

By the strong law of large numbers, there are only finitely many positive values $S_l - al$, where $2a = \mathbf{E}\alpha_j < 0$. Therefore, for all l except for those with $S_l - al > 0$,

$$\prod_{j=-l}^{0} \beta_j < e^{al}.$$

On the other hand, γ_{-l-1} exceeds the level l only finitely often. This means that the series in (17.4.13) (denote it by R) converges with probability 1. Moreover,

$$S = \sup_{k \ge 0} S_k \ge S_n$$

is a proper random variable. As result, we obtain that, for all n,

$$U^{-n}|X_{n+1}| \le |x|e^S + R + \gamma_0,$$

where all the terms on the right-hand side are proper random variables. The required majorant

$$M_n := U^{n-1}\left(|x|e^S + R + \gamma_0 \right)$$

is constructed. This implies that **(B)** is met. The theorem is proved.   □

The assertion of Theorem 17.4.2 can be extended to the multivariable case $\mathcal{X} = \mathbb{R}^d$, $d > 1$, as well (see [6]).

Note that conditions (17.4.11) and (17.4.12) are, in a certain sense, necessary not only for convergence $U^{-n+s} X_n(x) \to X^s$, but also for the boundedness of $X_n(x)$ (or of X^0) only. This fact can be best illustrated in the case when $F(t) \equiv G(t) \equiv t$. In that case, $U^{-n} X_{n+s+1}(x)$ and X^{s+1} admit explicit representations

$$U^{-n} X_{n+s+1}(x) = x \prod_{j=-n}^{s} \zeta_j + \sum_{l=0}^{n+s} \prod_{j=s-l}^{s} \zeta_j \eta_{s-l-1} + \eta_s,$$

$$X^{s+1} = \sum_{l=0}^{\infty} \prod_{j=s-l}^{s} \zeta_j \eta_{s-l-1} + \eta_s.$$

Assume that $\mathbf{E} \ln \zeta \geq 0$, $\eta \equiv 1$, and put

$$s := 0, \qquad z_j := \ln \zeta_j, \qquad Z_l := \sum_{j=-l}^{0} z_j.$$

Then

$$X^1 = 1 + \sum_{l=0}^{\infty} e^{Z_l}, \quad \text{where } \sum_{l=0}^{\infty} \mathbf{I}(Z_l \geq 0) = \infty$$

with probability 1, and consequently $X^1 = \infty$ and $X_n \to \infty$ with probability 1.

If $\mathbf{E}[\ln \eta]^+ = \infty$ and $\zeta = b < 1$ then

$$X^1 = \eta_0 + b \sum_{l=0}^{\infty} \exp\{y_{-l-1} + l \ln b\},$$

where $y_j = \ln \eta_j$; the event $\{y_{-l-1} > -l \ln b\}$ occurs infinitely often with probability 1. This means that $X^1 = \infty$ and $X_n \to \infty$ with probability 1.