Chapter Eleven

Stochastic Max-Plus Systems

This chapter is devoted to the study of sequences $\{x(k) : k \in \mathbb{N}\}$ satisfying the recurrence relation

$$x(k+1) = A(k) \otimes x(k), \qquad k \geq 0, \tag{11.1}$$

where $x(0) = x_0 \in \mathbb{R}^n_{\max}$ is the initial value and $\{A(k) : k \in \mathbb{N}\}$ is a sequence of $n \times n$ matrices over $\mathbb{R}_{\max}$. In order to develop a meaningful mathematical theory, we need some additional assumptions on $\{A(k) : k \in \mathbb{N}\}$. The approach presented in this chapter assumes that $\{A(k) : k \in \mathbb{N}\}$ is a sequence of random matrices in $\mathbb{R}^{n \times n}_{\max}$, defined on a common probability space. Specifically, we address the case where $\{A(k) : k \in \mathbb{N}\}$ consists of independent identically distributed (i.i.d.) random matrices. The theory is also available for the more general case of $\{A(k) : k \in \mathbb{N}\}$ being an ergodic sequence. However, for ease of exposition, we restrict our presentation to the i.i.d. case.

We focus on the asymptotic growth rate of $x(k)$. Note that $x(k)$ and thus $x(k)/k$ are random variables. We have to be careful about how to interpret the asymptotic growth rate. The key result of this chapter will be that under appropriate conditions the asymptotic growth rate of $x(k)$ defined in (11.1) is, with probability one, a constant.

The stochastic max-plus theory is dissimilar to the deterministic theory developed in this book so far, not only with respect to the applied techniques but also with respect to the obtained results. In the deterministic theory, proofs are usually constructive, and a rich variety of numerical procedures for computing eigenvalues and eigenvectors, for example, can be provided. In the stochastic theory, proofs are usually proofs of existence, and no efficient numerical algorithms for computing, say, the asymptotic growth rate for large-scale models, are available. In highlighting this difference one could say that while deterministic theory comes up with efficient algorithms for computing the asymptotic growth rate, the stochastic theory has to be content with showing that the asymptotic growth rate exists (with probability one) and that it equals some finite constant with probability one. The reader is referred to the notes section for some recently developed numerical approaches.

The stochastic limit theory will be discussed for three different cases of max-plus systems. First, we will study recurrence relations with the properties that (i) the arc set of the communication graph of $A(k)$ is nonrandom and (ii) the communication graph of $A(k)$ is strongly connected with probability one (this is the stochastic equivalent to the study of irreducible matrices). Second, as in deterministic theory, we will drop condition (ii) and study recurrence relations satisfying only condition (i) (this is the stochastic equivalent to the study of reducible matrices). Finally, we will examine recurrence relations not satisfying condition (i).

The chapter is organized as follows. In Section 11.1 basic concepts are introduced for stochastic max-plus recurrence relations (concepts familiar from deterministic theory, such as irreducibility, have to be redefined in a stochastic context). Moreover, examples of stochastic max-plus systems are given. Section 11.2 is devoted to subadditive ergodic theory for stochastic sequences. The limit theory for matrices with property (i) is provided in Section 11.3. Possible relaxations of the rather restrictive conditions needed for the analysis in the latter section are provided in Section 11.4. An overview of the stochastic theory not covered in this book is given in the notes section.

11.1 BASIC DEFINITIONS AND EXAMPLES

For a sequence of square matrices $\{A(k) : k \in \mathbb{N}\}$, we set

$$\bigotimes_{k=l}^{m} A(k) \stackrel{\text{def}}{=} A(m) \otimes A(m-1) \otimes \cdots \otimes A(l+1) \otimes A(l),$$

where $m \geq l$ and $\bigotimes_{k=l}^{m} A(k) \stackrel{\text{def}}{=} E$ otherwise.

A few words on the fundamentals of the stochastic setup are in order here. Let X be a random element in $\mathbb{R}_{\max}$ defined on a probability space $(\Omega, \mathcal{F}, P)$ modeling the underlying randomness.[1] When defining the expected value of X, denoted by $\mathbb{E}[X]$, one has to take care of the fact that X may take value ε $(= -\infty)$ with positive probability. This is reflected in the following extension to $\mathbb{R}_{\max}$ of the usual definition of integrability of a random variable on $\mathbb{R}$. We call $X \in \mathbb{R}_{\max}$ *integrable* if $X \oplus e = \max(X, 0)$ and $X \oplus' e = \min(X, 0)$ are integrable and if $\mathbb{E}[X \oplus e]$ is finite. The expected value of X is then given by $\mathbb{E}[X] = \mathbb{E}[X \oplus e] + \mathbb{E}[X \oplus' e]$. This definition implies that $\mathbb{E}[X] = -\infty$ if $P(X = \varepsilon) > 0$. A random matrix A in $\mathbb{R}_{\max}^{n \times m}$ is called integrable if its elements a_{ij} are integrable for $i \in \underline{n}, j \in \underline{m}$. The expected value of A is given by the matrix $\mathbb{E}[A]$ with elements $\left[\mathbb{E}[A]\right]_{ij} = \mathbb{E}[a_{ij}]$.

In order to define irreducibility for random matrices, we introduce the concept of a *fixed support* of a matrix.

DEFINITION 11.1 *We say that $\{A(k) : k \in \mathbb{N}\}$ has fixed support if the set of arcs of the communication graph of $A(k)$ is nonrandom and does not depend on k, or, more formally, for all $i, j \in \underline{n}$,*

$$\Big(\forall k \geq 0 : \; P\big([A(k)]_{ij} = \varepsilon\big) = 0\Big) \vee \Big(\forall k \geq 0 : \; P\big([A(k)]_{ij} = \varepsilon\big) = 1\Big).$$

With the definition of fixed support at hand, we say that a random matrix A is *irreducible* if it has fixed support and any sample of A is irreducible with probability one. Hence, for random matrices, irreducibility presupposes fixed support.

Stochasticity occurs quite naturally in real-life railway networks. For example, travel times become stochastic due to, for example, weather conditions or the individual behavior of the driver. Another source of randomness is the time durations

[1] It is a assumed that the reader is familiar with basic probability theory.

for boarding or alighting of passengers. Also, the lack of information about the future specification of a railway system, such as the type of rolling stock, the capacity of certain tracks, and so forth, can be modeled by randomness.

Example 11.1.1 *Consider the railway network described in Example 7.2.1 and assume that the travel times are random. More specifically, denote the kth travel time from station S_i to S_{i+1} by $a_{i+1,i}(k)$, for $i \in \underline{2}$ and the kth travel time from station S_3 to S_1 by $a_{1,3}(k)$. It is assumed that the travel times are stochastically independent and that the travel times for a certain track have the same distribution. If we follow the reasoning put forward in Example 7.2.1, together with exercise 5 in Chapter 7, then this system can be modeled through $x(k) = (x_1(k), x_2(k))^\top$, which satisfies*

$$x(k+1) = \begin{pmatrix} a_{21}(k) \oplus a_{13}(k+1) & a_{13}(k+1) \otimes a_{32}(k) \\ a_{21}(k) & a_{32}(k) \end{pmatrix} \otimes x(k),$$

where $x_1(k)$ denotes the kth departure time from station S_1 and $x_2(k)$ denotes the kth departure time from station S_2. Notice that the matrix on the right-hand side of the above equation has fixed support and is irreducible.

Example 11.1.2 *Consider the railway network described in Example 7.3.1, and assume, as in the previous example, that the travel times (and the interarrival times) are stochastically independent and that the travel times for a certain track as well as the interarrival times are identically distributed. Following the reasoning put forward in Example 7.3.1, this system can be modeled through $x(k) = (x_0(k), x_1(k), x_2(k))^\top$, which satisfies*

$$x(k+1) = A(k) \otimes x(k),$$

where the matrix $A(k)$ looks like

$$\begin{pmatrix} a_0(k) & \varepsilon & \varepsilon \\ a_0(k) \otimes a_{10}(k) & e & \varepsilon \\ a_0(k) \otimes a_{10}(k) \otimes a_{21}(k) & a_{21}(k) & e \end{pmatrix},$$

for $k \geq 0$. Observe that $A(k)$ has fixed support but fails to be irreducible.

Example 11.1.3 *Consider a simple railway network consisting of two stations with deterministic travel times between the stations. Specifically, the travel time from Station 2 to Station 1 equals σ', and the dwell time at Station 1 equals d, whereas the travel time from Station 1 to Station 2 equals σ and the dwell time at Station 2 equals d'. At Station 1 there is one platform at which trains can stop, whereas at Station 2 there are two platforms. Three trains circulate in the network. Initially, one train is present at Station 1, one train at Station 2, and the third train is just about to enter Station 2. The time evolution of this network is described by a max-plus linear sequence of vectors $x(k) = (x_1(k), \ldots, x_4(k))^\top$, where $x_1(k)$ is the kth arrival time of a train at Station 1 and $x_2(k)$ is the kth departure time of a train from the Station 1, $x_3(k)$ is the kth arrival time of a train at Station 2, and $x_4(k)$ is the kth departure time of a train from Station 2. Figure 11.1 on the following page shows the Petri net model of this system. The sample-path dynamics*

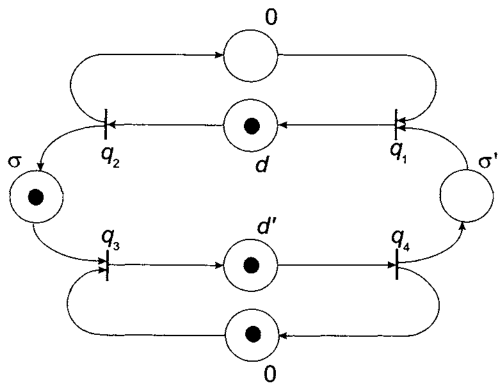

Figure 11.1: The initial state of the railway system with two platforms at Station 2.

of the network with two platforms at Station 2 is given by

$$\begin{aligned} x_1(k+1) &= x_2(k+1) \oplus \big(x_4(k+1) \otimes \sigma'\big), \\ x_2(k+1) &= x_1(k) \otimes d, \\ x_3(k+1) &= \big(x_2(k) \otimes \sigma\big) \oplus x_4(k), \\ x_4(k+1) &= x_3(k) \otimes d', \end{aligned}$$

for $k \geq 0$. Replacing $x_2(k+1)$ and $x_4(k+1)$ in the first equation by the expression on the right-hand side of the second and fourth equations above, respectively, yields

$$x_1(k+1) = \big(x_1(k) \otimes d\big) \oplus \big(x_3(k) \otimes d' \otimes \sigma'\big).$$

Hence, for $k \geq 0$,

$$\begin{aligned} x_1(k+1) &= \big(x_1(k) \otimes d\big) \oplus \big(x_3(k) \otimes d' \otimes \sigma'\big), \\ x_2(k+1) &= x_1(k) \otimes d, \\ x_3(k+1) &= \big(x_2(k) \otimes \sigma\big) \oplus x_4(k), \\ x_4(k+1) &= x_3(k) \otimes d', \end{aligned}$$

which reads in vector-matrix notation

$$x(k+1) = D_2 \otimes x(k),$$

where

$$D_2 = \begin{pmatrix} d & \varepsilon & d' \otimes \sigma' & \varepsilon \\ d & \varepsilon & \varepsilon & \varepsilon \\ \varepsilon & \sigma & \varepsilon & e \\ \varepsilon & \varepsilon & d' & \varepsilon \end{pmatrix}.$$

Notice that D_2 is irreducible.

Consider the railway network again, but one of the platforms at Station 2 is not available. The initial condition is as in the previous example. Figure 11.2 on the next page shows the Petri net of the system with one blocked platform at Station 2.

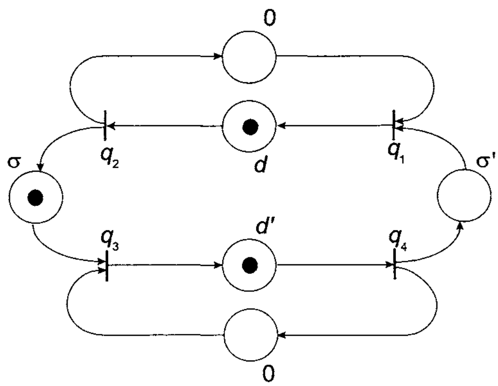

Figure 11.2: The initial state of the railway system with one blocked platform.

Note that the blocking is modeled by the absence of the token in the bottom place, yielding that $x_3(k+1) = (x_2(k) \otimes \sigma) \oplus x_4(k+1)$. *Following the line of argument put forward for the network with two platforms at Station 2, one arrives at*

$$x(k+1) = D_1 \otimes x(k),$$

where

$$D_1 = \begin{pmatrix} d & \varepsilon & d' \otimes \sigma' & \varepsilon \\ d & \varepsilon & \varepsilon & \varepsilon \\ \varepsilon & \sigma & d' & \varepsilon \\ \varepsilon & \varepsilon & d' & \varepsilon \end{pmatrix}.$$

Notice that D_1 *fails to be irreducible.*

Assume that whenever a train arrives at Station 2, one platform is blocked with probability p, *with* $0 < p < 1$. *This is modeled by introducing* $A(k)$ *with distribution*

$$P(\, A(k) = D_1\,) = p$$

and

$$P(\, A(k) = D_2\,) = 1 - p.$$

Then

$$x(k+1) = A(k) \otimes x(k)$$

describes the time evolution of the system with resource restrictions. Notice that $A(k)$ *fails to have a fixed support (and that* $A(k)$ *is thus not irreducible).*

11.2 THE SUBADDITIVE ERGODIC THEOREM

Subadditive ergodic theory is based on Kingman's subadditive ergodic theorem and its application to generalized products of random matrices. Kingman's result [56] is formulated in terms of *subadditive processes*. These are double-indexed processes $X = \{X_{m\,l} : m, l \in \mathbb{N}\}$ satisfying the following conditions:

(S1) For $i, j, k \in \mathbb{N}$, such that $i < j < k$, the inequality $X_{ik} \leq X_{ij} + X_{jk}$ holds with probability one.

(S2) All joint distributions of the process $\{X_{m+1,l+1} : l, m \in \mathbb{N}, l > m\}$ are the same as those of $\{X_{ml} : l, m \in \mathbb{N}, l > m\}$.

(S3) The expected value $g_l = \mathbb{E}[X_{0l}]$ exists and satisfies $g_l \geq -c \times l$ for some finite constant $c > 0$ and all $l \in \mathbb{N}$.

Kingman's celebrated ergodic theorem can now be stated as follows.

THEOREM 11.2 *(Kingman's subadditive ergodic theorem) If $X = \{X_{ml} : m, l \in \mathbb{N}\}$ is a subadditive process, then a finite number ξ exists such that*

$$\xi = \lim_{k \to \infty} \frac{X_{0k}}{k}$$

with probability one and

$$\xi = \lim_{k \to \infty} \frac{\mathbb{E}[X_{0k}]}{k}.$$

The surprising part of Kingman's ergodic theorem is that the random variables X_{0k}/k converge, with probability one, towards the same finite value, which is the limit of $\mathbb{E}[X_{0k}]/k$.

We will apply Kingman's subadditive ergodic theorem to the maximal (resp., minimal) finite element of a matrix. The basic concepts are defined in the following. For $A \in \mathbb{R}_{\max}^{n \times m}$, the minimal finite entry of A, denoted by $||A||_{\min}$, is given by

$$||A||_{\min} = \min\{a_{ij} \mid (i,j) \in \mathcal{D}(A)\},$$

where $||A||_{\min} = \varepsilon' \ (= +\infty)$ if $\mathcal{D}(A) = \emptyset$. (Recall that $\mathcal{D}(A)$ denotes the set of arcs in the communication graph of A.) In the same vein, we denote the maximal finite entry of $A \in \mathbb{R}_{\max}^{n \times m}$ by $||A||_{\max}$, which implies

$$||A||_{\max} = \max\{a_{ij} \mid (i,j) \in \mathcal{D}(A)\},$$

where $||A||_{\max} = \varepsilon$ if $\mathcal{D}(A) = \emptyset$. A direct consequence of the above definitions is that for any regular $A \in \mathbb{R}_{\max}^{n \times m}$

$$||A||_{\min} \leq ||A||_{\max}.$$

Notice that $||A||_{\min}$ and $||A||_{\max}$ can have negative values. It is easily checked (see exercise 4) that for regular $A \in \mathbb{R}_{\max}^{n \times m}$ and regular $B \in \mathbb{R}_{\max}^{m \times l}$

$$||A \otimes B||_{\max} \leq ||A||_{\max} \otimes ||B||_{\max} \tag{11.2}$$

and

$$||A \otimes B||_{\min} \geq ||A||_{\min} \otimes ||B||_{\min}. \tag{11.3}$$

We now revisit our basic max-plus recurrence relation

$$x(k+1) = A(k) \otimes x(k),$$

for $k \geq 0$, with $x(0) = x_0$. To indicate the initial value of the sequence, we sometimes use the notation

$$x(k; x_0) = \bigotimes_{l=0}^{k-1} A(l) \otimes x_0, \qquad k \in \mathbb{N}. \tag{11.4}$$

To abbreviate the notation, we set for $m \geq l \geq 0$

$$A[m, l] \stackrel{\text{def}}{=} \bigotimes_{k=l}^{m-1} A(k).$$

With this definition (11.4) can be written as

$$x(k; x_0) = A[k, 0] \otimes x_0,$$

for $k \geq 0$. Notice that for $0 \leq l \leq p \leq m$

$$A[m, l] = A[m, p] \otimes A[p, l]. \tag{11.5}$$

LEMMA 11.3 *Let $\{A(k) : k \in \mathbb{N}\}$ be an i.i.d. sequence of integrable matrices such that $A(k)$ is regular with probability one. Then $\{-||A[m, l]||_{\min} : m > l \geq 0\}$ and $\{||A[m, l]||_{\max} : m > l \geq 0\}$ are subadditive ergodic processes.*

Proof. For $2 \leq m$ and $0 \leq l < p < m$, we obtain

$$\begin{aligned} ||A[m, l]||_{\max} &\stackrel{(11.5)}{=} ||A[m, p] \otimes A[p, l]||_{\max} \\ &\stackrel{(11.2)}{\leq} ||A[m, p]||_{\max} + ||A[p, l]||_{\max}, \end{aligned}$$

which establishes (S1) for $||A[m, l]||_{\max}$. The proof that (S1) also holds for $-||A[m, l]||_{\min}$ follows from the same line of argument, where (11.3) is used for establishing the inequality, and the proof is therefore omitted.

The stationarity condition (S2) follows immediately from the i.i.d. assumption for $\{A(k) : k \in \mathbb{N}\}$.

We now turn to condition (S3). The fact that $\{A(k) : k \in \mathbb{N}\}$ is an i.i.d. sequence implies

$$\begin{aligned} \mathbb{E}\big[||A[k, 0]||_{\max}\big] &\geq \mathbb{E}\big[||A[k, 0]||_{\min}\big] \\ &\stackrel{(11.3)}{\geq} k \times \mathbb{E}\big[||A(0)||_{\min}\big] \\ &\geq k \times \big(- \big| \mathbb{E}\big[||A(0)||_{\min}\big] \big| \big). \end{aligned}$$

Integrability of $A(0)$ together with regularity implies that $\mathbb{E}[||A(0)||_{\min}]$ is finite (for a proof use the fact that $\min(X, Y) \leq |X| + |Y|$). This establishes condition (S3) for $||A[m, l]||_{\max}$. For the proof that $-||A[m, l]||_{\min}$ satisfies (S3), notice that (11.2) implies

$$-\mathbb{E}\big[||A[k, 0]||_{\min}\big] \geq -\mathbb{E}\big[||A[k, 0]||_{\max}\big] \geq -k \times \mathbb{E}\left[|| \, A(0) \, ||_{\max}\right].$$

Since integrability of $A(0)$ together with regularity implies that $\mathbb{E}\left[|| \, A(0) \, ||_{\max}\right]$ is finite, we have proved the claim. □

The above lemma shows that Kingman's subadditive ergodic theorem can be applied to $||A[k, 0]||_{\min}$ and $||A[k, 0]||_{\max}$. The precise statement is given in the following theorem.

THEOREM 11.4 *Let $\{A(k) : k \in \mathbb{N}\}$ be an i.i.d. sequence of integrable matrices such that $A(k)$ is regular with probability one. Then, finite constants λ^{top} and λ^{bot} exist such that with probability one*

$$\lambda^{\mathrm{bot}} \stackrel{\mathrm{def}}{=} \lim_{k\to\infty} \frac{1}{k} \, ||A[k,0]||_{\min} \leq \lambda^{\mathrm{top}} \stackrel{\mathrm{def}}{=} \lim_{k\to\infty} \frac{1}{k} \, ||A[k,0]||_{\max}$$

and

$$\lambda^{\mathrm{bot}} = \lim_{k\to\infty} \frac{1}{k}\mathbb{E}\Big[\, ||A[k,0]||_{\min} \Big], \qquad \lambda^{\mathrm{top}} = \lim_{k\to\infty} \frac{1}{k}\mathbb{E}\Big[\, ||A[k,0]||_{\max} \Big].$$

The constant λ^{top} is called the *top* or *maximal Lyapunov exponent* of $\{A(k) : k \in \mathbb{N}\}$, and λ^{bot} is called the *bottom* or *minimal Lyapunov exponent* of $\{A(k) : k \in \mathbb{N}\}$. The top and bottom Lyapunov exponents of $A(k)$ are related to the asymptotic growth rate of $x(k)$ defined in (11.1) as follows. The top Lyapunov exponent equals the asymptotic growth rate of the maximal entry of $x(k)$, and the bottom Lyapunov exponent equals the asymptotic growth rate of the minimal entry of $x(k)$. The precise statement is given in the following corollary.

COROLLARY 11.5 *Let $\{A(k) : k \in \mathbb{N}\}$ be an i.i.d. sequence of integrable matrices such that $A(k)$ is regular with probability one. Then, for any finite and integrable initial condition x_0, it holds with probability one that*

$$\lambda^{\mathrm{bot}} = \lim_{k\to\infty} \frac{||x(k;x_0)||_{\min}}{k} \leq \lambda^{\mathrm{top}} = \lim_{k\to\infty} \frac{||x(k;x_0)||_{\max}}{k}$$

and

$$\lambda^{\mathrm{bot}} = \lim_{k\to\infty} \frac{1}{k}\mathbb{E}\Big[||x(k;x_0)||_{\min} \Big], \qquad \lambda^{\mathrm{top}} = \lim_{k\to\infty} \frac{1}{k}\mathbb{E}\Big[||x(k;x_0)||_{\max} \Big].$$

Proof. Note that $x(k;x_0) = A[k,0] \otimes x_0$ for any $k \in \mathbb{N}$. Provided that x_0 is finite, it is easily checked (see exercise 4) that

$$||A[k,0]||_{\min} \otimes ||x_0||_{\min} \leq ||x(k;x_0)||_{\min} \leq ||A[k,0]||_{\min} \otimes ||x_0||_{\max}.$$

Dividing the above row of inequalities by k and letting k tend to ∞ yields

$$\lim_{k\to\infty} \frac{1}{k} ||x(k;x_0)||_{\min} = \lambda^{\mathrm{bot}}$$

with probability one. The proof for the other limit follows from the same line of argument.

The arguments used for the proof of the first part of the corollary remain valid when expected values are applied (we omit the details). This concludes the proof of the corollary. □

A sufficient condition for $A(k)$ to be regular with probability one is the irreducibility of $A(k)$. Therefore, in the literature, Theorem 11.4 and Corollary 11.5 are often stated with irreducibility (instead of regularity) as a condition.

11.3 MATRICES WITH FIXED SUPPORT

11.3.1 Irreducible matrices

In this section, we consider i.i.d. sequences $\{A(k) : k \in \mathbb{N}\}$ of integrable and irreducible matrices such that with probability one finite entries are bounded from below by a finite constant. As we will show in the following theorem, the setting of this section implies that $\lambda^{\mathrm{top}} = \lambda^{\mathrm{bot}}$, which in particular implies convergence of $x_i(k)/k$ as k tends to ∞, for $i \in \underline{n}$. The main technical result is provided in the following lemma.

LEMMA 11.6 *Let $D \in \mathbb{R}_{\max}^{n\times n}$ be a nonrandom irreducible matrix such that its communication graph has cyclicity σ. If $A(k) \geq D$ with probability one, for any k, then integers L and N exist such that for any $k \geq N$*

$$\|x(k)\|_{\min} \geq \|x(k-L)\|_{\max} + (\|D^{\otimes\sigma}\|_{\min})^{\otimes L}.$$

Proof. Denote the communication graph of D by $\mathcal{G} = (\mathcal{N}, \mathcal{D})$, and let $\mathcal{G}$ be of cyclicity one. Denote the number of elementary circuits in $\mathcal{G}$ by q, and let β_i denote the length of circuit ξ_i, for $i \in \underline{q}$. Then the greatest common divisor of $\{\beta_1, \ldots, \beta_q\}$ is equal to one. According to Theorem 3.2 a natural number N exists such that for all $\kappa \geq N$ there are integers $n_1, \ldots, n_q \geq 0$ such that $\kappa = n_1\beta_1 + \cdots + n_q\beta_q$.

Let l_{ij} denote the minimal length of a path from j to i containing *all* nodes of $\mathcal{G}$. Such paths exist because D is irreducible (and, hence, $\mathcal{G}$ is strongly connected). Let the maximal length of all these paths be denoted by l, i.e., $l = \max_{i,j\in\underline{n}} l_{ij}$.

Next, choose an L with $L \geq N+l$. Then for any $i, j \in \underline{n}$, there is a path from j to i of length L. Indeed, take any $i, j \in \underline{n}$ and choose a path, as mentioned above, from j to i containing *all* nodes of $\mathcal{G}$ and having minimal length l_{ij}. Clearly, the path has at least one node in common with each of the q circuits in $\mathcal{G}$. As $L - l_{ij} \geq N$, there are integers $n_1, \ldots, n_q \geq 0$ such that $L - l_{ij} = n_1\beta_1 + \cdots + n_q\beta_q$. Hence, by adding n_1 copies of circuit ξ_1, and so on, up to n_q copies of circuit ξ_q to the chosen path from i to j of length l_{ij}, a new path from j to i is created of length L.

In graph-theoretical terms, the element $[A(k, k-L)]_{ij}$ denotes the maximal weight of a path of length L from node j to node i on the "interval" $[k-L, k)$. Since $A[k, k-L] \geq D^{\otimes L}$ by assumption, it follows that for all $k \geq N$ and all $i \in n$

$$\begin{aligned}
x_i(k) &= \bigoplus_{j=1}^{n} [A(k, k-L)]_{ij} \otimes x_j(k-L) \\
&\geq \bigoplus_{j=1}^{n} \left[D^{\otimes L}\right]_{ij} \otimes x_j(k-L) \\
&\geq \bigoplus_{j=1}^{n} (\|D\|_{\min})^{\otimes L} \otimes x_j(k-L) \\
&\geq (\|D\|_{\min})^{\otimes L} \otimes \bigoplus_{j=1}^{n} x_j(k-L),
\end{aligned}$$

implying that

$$||x(k)||_{\min} \geq ||x(k-L)||_{\max} + (||D||_{\min})^{\otimes L}, \qquad \forall k \geq N.$$

Similarly as in Lemma 3.3 it can be shown that if $\mathcal{G}(D)$ has cyclicity σ, then $\mathcal{G}(D^{\otimes \sigma})$ has cyclicity one. Applying the arguments put forward above to $D^{\otimes \sigma}$ extends the result to the case of matrices with cyclicity greater than one. □

The condition that $A(k) \geq D$ with probability one for any $k \in \mathbb{N}$ and with D being irreducible will be referred to as condition (H_1).

(H_1) *There exists a nonrandom irreducible matrix D such that $A(k) \geq D$ for any $k \in \mathbb{N}$, with probability one.*

Notice that Example 11.1.1 satisfies (H_1), whereas Example 11.1.2 and Example 11.1.3 fail to satisfy (H_1). Lemma 11.6 provides the main technical means for establishing sufficient conditions for equality of maximal, minimal, and individual growth rates. The precise statement is provided in the following theorem.

THEOREM 11.7 *Let $\{A(k) : k \in \mathbb{N}\}$ be a random sequence of integrable matrices satisfying (H_1). For $x(k)$ defined in (11.1) it holds, with probability one, that*

$$\lim_{k\to\infty} \frac{1}{k} ||x(k;x_0)||_{\min} = \lim_{k\to\infty} \frac{1}{k} x_i(k;x_0) = \lim_{k\to\infty} \frac{1}{k} ||x(k;x_0)||_{\max}$$

for any $i \in \underline{n}$ and any finite initial state x_0.

Proof. Let D be given as in (H_1); then D satisfies the condition put forward in Lemma 11.6, and finite positive numbers L and N exist such that for $k \geq N$

$$||x(k;x_0)||_{\min} \geq ||x(k-L;x_0)||_{\max} + (||D^{\otimes \sigma}||_{\min})^{\otimes L},$$

where σ denotes the cyclicity of the communication graph of D. Dividing both sides of the above inequality by k and letting k tend to ∞ yields

$$\lim_{k\to\infty} \frac{1}{k} ||x(k;x_0)||_{\min} \geq \lim_{k\to\infty} \frac{1}{k} ||x(k;x_0)||_{\max}, \tag{11.6}$$

for any finite initial vector x_0. The existence of the above limits is guaranteed by Corollary 11.5, where we use the fact that (H_1) implies that $A(k)$ is regular with probability one. Following the line of argument in the proof of Corollary 11.5, the limits in (11.6) are independent of the initial state.

Combining (11.6) with the obvious fact that $||x(k;x_0)||_{\max} \geq x_j(k;x_0) \geq ||x(k;x_0)||_{\min}$, for $j \in \underline{n}$, proves the claim. □

By Theorem 11.7, integrability of $A(k)$ together with (H_1) is a sufficient condition for the top and bottom Lyapunov exponent to coincide. Moreover, a random matrix $A(k)$ satisfies condition (H_1) if $A(k)$ is irreducible and if, with probability one, all finite elements are bounded from below by a finite number. Combining this with Theorem 11.4 and Corollary 11.5, we arrive at the following limit theorem for i.i.d. sequences of irreducible matrices.

THEOREM 11.8 *Let $\{A(k) : k \in \mathbb{N}\}$ be an i.i.d. sequence of integrable and irreducible matrices such that with probability one all finite elements are bounded* *from below by a finite number. Then, it holds that* $\lambda \stackrel{\text{def}}{=} \lambda^{\text{top}} = \lambda^{\text{bot}}$*, and with probability one for all* $i, j \in \underline{n}$

$$\lim_{k\to\infty} \frac{1}{k}\big[A[k,0]\big]_{ij} = \lim_{k\to\infty}\frac{1}{k}\mathbb{E}\Big[\big[A[k,0]\big]_{ij}\Big] = \lambda.$$

Moreover, for any finite integrable initial condition x_0 *it holds with probability one that*

$$\lim_{k\to\infty} \frac{x_j(k;x_0)}{k} = \lim_{k\to\infty}\frac{1}{k}\mathbb{E}\Big[x_j(k;x_0)\Big] = \lambda, \qquad j \in \underline{n}.$$

The constant λ, defined in Theorem 11.8, is referred to as the *max-plus Lyapunov exponent* of the sequence of random matrices $\{A(k) : k \in \mathbb{N}\}$. There is no ambiguity in denoting the Lyapunov exponent of $\{A(k) : k \in \mathbb{N}\}$ and the eigenvalue of a matrix A by the same symbol, since the Lyapunov exponent of $\{A(k) : k \in \mathbb{N}\}$ is just the eigenvalue of A whenever $A(k) = A$ for all $k \in \mathbb{N}$. To see this, compare Theorem 11.8 with Lemma 3.12.

The system in Example 11.1.1 satisfies the conditions in Theorem 11.8, and the existence of the Lyapunov exponent is thus guaranteed. Notice that the systems in Examples 11.1.2 and 11.1.3 cannot be analyzed by Theorem 11.8.

11.3.2 Reducible matrices

Now suppose that $A(k)$ has a fixed support and drop the assumption that it is irreducible. To deal with reducible matrices $A(k)$, we decompose $A(k)$ into its irreducible parts. The limit theorem, to be presented shortly, then states that the Lyapunov exponent of the overall matrix equals the maximum of the Lyapunov exponent of its irreducible components. This result presents the stochastic version of Theorem 3.17.

Let $\{A(k) : k \in \mathbb{N}\}$ be a sequence of matrices in $\mathbb{R}_{\max}^{n\times n}$ with fixed support, and consider the associated communication graph of $A(k)$ (with nonrandom arc set). For $i \in \underline{n}$, $[i]$ denotes the set of nodes of the m.s.c.s. that contains node i, and denote by $\lambda_{[i]}$ the Lyapunov exponent associated to the matrix obtained by restricting $A(k)$ to the nodes in $[i]$. We state the theorem without proof. A proof can, for example, be found in [5].

THEOREM 11.9 *Let* $\{A(k) : k \in \mathbb{N}\}$ *be an i.i.d. sequence of integrable matrices in* $\mathbb{R}_{\max}^{n\times n}$ *with fixed support such that with probability one all finite elements are bounded from below by a finite number. For any finite integrable initial value* x_0*, it holds with probability one that*

$$\lim_{k\to\infty} \frac{x_j(k;x_0)}{k} = \lim_{k\to\infty}\frac{1}{k}\mathbb{E}\Big[x_j(k;x_0)\Big] = \lambda_j,$$

with

$$\lambda_j = \bigoplus_{i\in\pi^*(j)} \lambda_{[i]}, \qquad j \in \underline{n}.$$

The system in Example 11.1.2 satisfies the conditions in Theorem 11.9, and the existence of the Lyapunov exponent is thus guaranteed. Notice that the system in Example 11.1.3 cannot be analyzed by Theorem 11.9 because its support is not fixed.

11.4 BEYOND FIXED SUPPORT

In this section we discuss possible relaxations of the conditions put forward in Theorem 11.8. The main technical condition is the following.

(H_2) *There exists a nonrandom irreducible matrix D such that*

$$P(\,A(k) \geq D) \geq p, \quad k \in \mathbb{N},$$

for some $p \in (0, 1]$.

Condition (H_2) suffices to guarantee that the top and bottom Lyapunov exponent coincide. The precise statement is given in the following lemma.

LEMMA 11.10 *Let $\{A(k) : k \in \mathbb{N}\}$ be an i.i.d. sequence of integrable matrices such that $A(0)$ is regular with probability one. If condition (H_2) holds, then the top and bottom Lyapunov exponents of $\{A(k) : k \in \mathbb{N}\}$ coincide.*

Proof. By Lemma 11.6, there exists an integer L such that there is a path of length L from any node j to any node i in the graph of D with weight at least $(||D||_{\min})^{\otimes L}$, where we assume, for ease of exposition, that the communication graph of D is of cyclicity one. Consider the event that for some k it holds that

$$\forall l \in \underline{L}: \quad A(k-l) \geq D. \tag{11.7}$$

On this event,

$$\bigotimes_{l=1}^{L} A(k-l) \geq D^{\otimes L},$$

and in accordance with Lemma 11.6 it follows that

$$||x(k)||_{\min} \geq ||x(k-L)||_{\max} + (||D||_{\min})^{\otimes L}. \tag{11.8}$$

Notice that by assumption (H_2) the event characterized in (11.7) occurs at least with probability $p^L > 0$. Let $\{\tau_m\}$ be the sequence of times k when the event characterized in (11.7) occurs. The i.i.d. assumption implies that $\tau_m < \infty$ for $m \in \mathbb{N}$ and that $\lim_{m\to\infty} \tau_m = \infty$. By inequality (11.8),

$$||x(\tau_m)||_{\min} \geq ||x(\tau_m - L)||_{\max} + (||D||_{\min})^{\otimes L},$$

and dividing both sides of the above inequality by τ_m and letting m tend to ∞ yields with probability one

$$\lim_{m\to\infty} \frac{1}{\tau_m} ||x(\tau_m)||_{\min} \geq \lim_{m\to\infty} \frac{1}{\tau_m} ||x(\tau_m)||_{\max}.$$

The existence of the top and the bottom Lyapunov exponents is guaranteed by Corollary 11.5, and the above inequality for a subsequence of $x(k)$ is sufficient to establish equality of the top and bottom Lyapunov exponents. □

Lemma 11.10 allows us to extend Theorem 11.8 to matrices that fail to have a fixed support. More precisely, the fixed support condition can be replaced by the assumption that $A(k)$ is, with positive probability, bounded from below by an irreducible nonrandom matrix. Notice that D_2 in Example 11.1.3 is irreducible, and $\{A(k) : k \in \mathbb{N}\}$ in Example 11.1.3 thus satisfied condition (H_2) (take $D = D_2$). The extended version of Theorem 11.8 thus applies to this example.

11.5 EXERCISES

1. Show that if $A \in \mathbb{R}_{\max}^{n\times m}$ and $B \in \mathbb{R}_{\max}^{m\times l}$ are integrable, then $A \otimes B$ is integrable.
2. Show that if $A \in \mathbb{R}_{\max}^{n\times m}$ and $B \in \mathbb{R}_{\max}^{m\times l}$ are regular with probability one, then $A \otimes B$ is regular with probability one.
3. Show that if A is regular with probability one, then $||A||_{\min}$ and $||A||_{\max}$ are finite with probability one.
4. Let $A \in \mathbb{R}_{\max}^{n\times m}$ and $B \in \mathbb{R}_{\max}^{m\times l}$ be regular. Show that

$$||A||_{\min} \otimes ||B||_{\min} \leq ||A \otimes B||_{\min}, \qquad ||A \otimes B||_{\max} \leq ||A||_{\max} \otimes ||B||_{\max},$$

   and

$$||A||_{\min} \otimes ||B||_{\min} \leq ||A \otimes B||_{\min} \leq ||A||_{\min} \otimes ||B||_{\max}.$$

5. Suppose that for $\{x(k) : k \in \mathbb{N}\}$ defined in (11.1) it holds that $\mathbb{E}[x(k+1) - x(k)]$ converges to $\mathbf{u}[\lambda]$ as k tends to ∞ for some finite constant λ. Show that this implies that λ is the Lyapunov exponent of $\{A(k) : k \in \mathbb{N}\}$. (Hint: Use a Cesaro averaging argument.)
6. Show that condition (H_2) can be relaxed as follows. There exists a finite number M and nonrandom matrices $D_i \in \mathbb{R}_{\max}^{n\times n}$, for $i \in \underline{M}$, such that $D_M \otimes \cdots \otimes D_2 \otimes D_1$ is irreducible and $P(A(k) \geq D_i) > 0$, for $i \in \underline{M}$.
7. Consider the system $x(k+1) = A(k) \otimes x(k)$, with $A(k) = D_1$ with probability 0.5 and $A(k) = D_2$, also with probability 0.5. The matrices D_1 and D_2 are taken from Example 11.1.3 into which the numerical values $\sigma = \sigma' = d = 1$ and $d' = 2$ are substituted. The elements in the sequence $A(k)$, $k \in \mathbb{N}$, are assumed to be independent.

   - If one starts with an arbitrary initial state, say, $x(0) = (0,0,0,0)^\top$, then one considers the evolution of the state $\overline{x(k)}$ in the projective space (see Section 1.4). For $\overline{x(1)}$ one gets two possibilities according to whether D_1 or D_2 was the transition matrix. Each of these possibilities leads to two possible $\overline{x(2)}$ states and so on. Show that this projective space consists of ten elements and that the set of absorbing states consists of $\overline{x}^{(1)} \stackrel{\text{def}}{=} \overline{(0,0,-1,-1)^\top}$, $\overline{x}^{(2)} \stackrel{\text{def}}{=} \overline{(0,-1,-1,-1)^\top}$, and $\overline{x}^{(3)} \stackrel{\text{def}}{=} \overline{(0,-1,-2,-1)^\top}$.
   - A Markov chain can be constructed with these three states, as indicated in Figure 11.3, left.

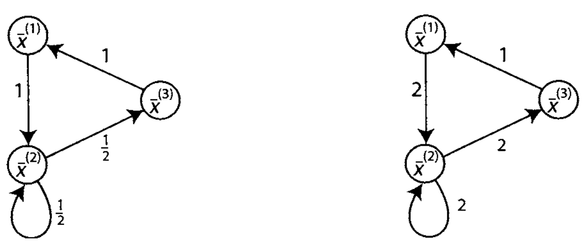

Figure 11.3: Markov chain with transition probabilities (left) and with time durations (right).

   Show that the stationary distribution for this Markov chain is $p_1 = p_3 = 0.25$ and $p_2 = 0.5$, where p_i corresponds to $\overline{x}^{(i)}$.

- The Lyapunov exponent can be calculated as

$$\lambda = p_1 t_{21} + p_2 \left(\frac{1}{2} t_{22} + \frac{1}{2} t_{32} \right) + p_3 t_{13} = \frac{7}{4},$$

  where the t_{ij}'s are the time durations as indicated in Figure 11.3, right.

- Note that $\lambda(D_1) = 2$ and $\lambda(D_2) = \frac{5}{3}$ and that $\frac{1}{2}(\lambda(D_1) + \lambda(D_2)) \neq \frac{7}{4}$.

8. Show that condition (H_2) in Lemma 11.10 can be replaced by the following (weaker) condition:

   (H_3) A nonrandom irreducible matrix D and a fixed number N exist such that

$$P\left(\bigotimes_{i=k+1}^{k+N} a(i) \geq D \right) \geq p$$

   for some $p \in (0, 1]$.

11.6 NOTES

Example 11.1.3 is an adaptation of an example by Baccelli and Hong [6]. A different approach to stability theory elaborating on the projective space can be found in [63].

A discussion of max-plus linearity in terms of queueing systems can be found in [51]. A max-plus-based analysis of a train network with stochastic travel times can be found in [58]. In [52], a control-theoretic approach to train networks with stochastic travel times based on a max-plus model can be found.

Computing the Lyapunov exponent exactly is a long-standing problem. Upper and lower bounds can be found in [7] and [8]. Approaches that use parallel simulation to estimate the growth rate $x_j(k)/k$ for large k are described in [4]. A classical reference on Lyapunov exponents of products of random matrices is [14], and a more recent one, dedicated to nonnegative matrices, is [53].

Based on a limit theorem for Markov chains, strong limit theorems for max-plus systems providing results on Lyapunov exponents have been developed; see [70], [74] , [79], and [84]. Exercise 7 is an example of this approach, where the Lyapunov exponent can actually be computed. Unfortunately, apart from simple problems, computing the Lyapunov exponent in this manner becomes extremely difficult.

The lack of numerical approaches for stochastic max-plus systems has lead to an increased interest in Taylor series approximations of performance characteristics of max-plus systems. The pioneering paper of Baccelli and Schmidt [10] has initiated an ongoing search for better and more efficient algorithms for approximately computing characteristics of stochastic max-plus systems. Recent results in this area are [3], [6], and [39].

One of the celebrated results in the field of stochastic max-plus theory is the extension of Loyne's result on the stability of waiting times in the G/G/1 queue [60] to max-plus linear queueing systems. Readers interested in the max-plus theory of waiting times are referred to [5], [9], and [63].

In [13] the stochastic approach of this chapter is combined with the so-called model predictive control problem, which is well known in system theory.