Chapter 2

Ergodic Theory

Ergodic theory for stochastic max-plus linear systems studies the asymptotic behavior of the sequence

$$x(k+1) = A(k) \otimes x(k), \quad k \geq 0,$$

where $\{A(k)\}$ is a sequence of regular matrices in $\mathbb{R}_{\max}^{J \times J}$ and $x(0) = x_0 \in \mathbb{R}_{\max}^J$. One distinguishes between two types of asymptotic results:

(Type I) *first-order* limits

$$\lim_{k \to \infty} \frac{x(k)}{k},$$

(Type II) *second-order* limits of type

$$(a) \qquad \lim_{k \to \infty} \Big(x_i(k) - x_j(k) \Big) \quad \text{and} \quad (b) \qquad \lim_{k \to \infty} \Big(x_j(k+1) - x_j(k) \Big).$$

A first-order limit of departure times is an inverse throughput in a queuing network. For example, the throughput of the tandem queuing network in Example 1.5.2 can be obtained from

$$\lim_{k \to \infty} \frac{k}{x_J(k)},$$

provided that the limit exists.

Second-order limits are related to steady-state waiting times and cycle times. Consider the closed tandem network in Example 1.5.1. There are J customers circulating through the system. Thus, the k^{th} and the $(k+J)^{th}$ departure from queue j refers to the same (physical) customer and the cycle time of this customer equals

$$x_j(k+J) - x_j(k).$$

Hence, the existence of the second-order limit $x_j(k+1) - x_j(k)$ implies limit results on steady-state cycle times of customers. For more examples of the modeling of performance characteristics of queuing systems via first-order and second-order expressions we refer to [10, 77, 84].

The chapter is organized as follows. Section 2.1 and Section 2.2 are devoted to limits of type I. Section 2.1 presents background material from the theory of deterministic max-plus systems. In Section 2.2 we present Kingman's celebrated subadditive ergodic theorem. We will show that max-plus recurrence relations constitute in a quite natural way subadditive sequences and we will apply the subadditive ergodic theorem in order to obtain a first ergodic theorem for max-plus linear systems. Limits of type IIa will be addressed in Section 2.3, where the stability theorem for waiting times in max-plus linear networks is addressed. In Section 2.4, limits of type I and type IIa will be discussed. This section is devoted to the study of max-plus linear systems $\{x(k)\}$ such that the relative difference between the components of $x(k)$ constitutes a Harris recurrent Markov chain. Section 2.5 and Section 2.6 are devoted to limits of type IIb and type I. In Section 2.5, we study ergodic theorems in the so called projective space. In Section 2.6, we show how the type I limit can be represented as a second-order limit.

2.1 Deterministic Limit Theory (Type I)

This section provides results from the theory of deterministic max-plus linear systems that will be needed for ergodic theory of max-plus linear stochastic systems. This monograph is devoted to stochastic systems and we state the results presented in this section without proof. To begin with, we state the celebrated cyclicity theorem for deterministic matrices, which is of key importance for our analysis.

Let $A \in \mathbb{R}_{\max}^{J \times J}$, if $x \in \mathbb{R}_{\max}^{J}$ with at least one finite element and $\lambda \in \mathbb{R}_{\max}$ satisfy

$$\lambda \otimes x = A \otimes x ,$$

then we call λ an *eigenvalue of* A and x an *eigenvector associated with* λ. Note that the set of all eigenvectors associated with an eigenvalue is a vector space. We denote the set of eigenvectors of A by $V(A)$. The following theorem states a key result from the theory of deterministic max-plus linear systems, namely, that any irreducible square matrix in the max-plus semiring possesses a unique eigenvalue. Recall that $x^{\otimes n}$ denotes the n^{th} power of $x \in \mathbb{R}_{\max}$, see equation (1.5).

Theorem 2.1.1 *(Cohen et al. [33, 34] and Heidergott et al. [65]) For any irreducible matrix $A \in \mathbb{R}_{\max}^{J \times J}$, uniquely defined integers $c(A)$, $\sigma(A)$ and a uniquely defined real number $\lambda = \lambda(A)$ exist such that for all $n \geq c(A)$:*

$$A^{n+\sigma(A)} = \lambda^{\otimes \sigma(A)} \otimes A^{n} .$$

In the above equation, $\lambda(A)$ is the eigenvalue of A; the number $c(A)$ is called the coupling time of A and $\sigma(A)$ is called the cyclicity of A.

Moreover, for any finite initial vector $x(0)$ the sequence $x(k+1) = A \otimes x(k)$, $k \geq 0$, satisfies

$$\lim_{k \to \infty} \frac{x_j(k)}{k} = \lambda , \quad 1 \leq j \leq J .$$

The above theorem can be seen as the max-plus analog of the Perron-Frobenius theorem in conventional linear algebra and it is for this reason that it is sometimes referred to as 'max-plus Perron-Frobenius theorem.' We illustrate the above definition with a numerical example.

Example 2.1.1 *Matrix*

$$A = \begin{pmatrix} 1 & \varepsilon & 2 & \varepsilon \\ 1 & \varepsilon & \varepsilon & \varepsilon \\ \varepsilon & e & \varepsilon & e \\ \varepsilon & \varepsilon & 2 & \varepsilon \end{pmatrix}$$

has eigenvalue $\lambda(A) = 1$ *and coupling time* $c(A) = 4$. *The critical graph of* A *consists of the circuits* $(1,1)$ *and* $((1,2),(2,3),(3,1))$, *and* A *is thus of cyclicity* $\sigma(A) = 1$. *In accordance with Theorem 2.1.1,* $A^{n+1} = 1 \otimes A^n$, *for* $n \geq 4$ *and*

$$\lim_{k\to\infty} \frac{(A^k \otimes x_0)_j}{k} = 1, \quad 1 \leq j \leq 4,$$

for any finite initial condition x_0. *For matrix*

$$B = \begin{pmatrix} 1 & \varepsilon & 2 & \varepsilon \\ 1 & \varepsilon & \varepsilon & \varepsilon \\ \varepsilon & e & 2 & e \\ \varepsilon & \varepsilon & 2 & \varepsilon \end{pmatrix}$$

we obtain $\lambda(B) = 2$, *coupling time* $c(B) = 4$. *The critical graph of* B *consists of the selfloop* $(3,3)$, *which implies that* $\sigma(B) = 1$. *Theorem 2.1.1 yields* $B^{n+1} = 2 \otimes B^n$, *for* $n \geq 4$ *and*

$$\lim_{k\to\infty} \frac{(B^k \otimes x_0)_j}{k} = 2, \quad 1 \leq j \leq 4,$$

for any finite initial condition x_0. *Matrix*

$$C = \begin{pmatrix} \varepsilon & \varepsilon & 7 & \varepsilon \\ 3 & \varepsilon & \varepsilon & \varepsilon \\ \varepsilon & e & \varepsilon & e \\ \varepsilon & \varepsilon & 7 & \varepsilon \end{pmatrix}$$

has eigenvalue $\lambda(C) = 3.5$, *coupling time* $c(C) = 4$. *The critical graph of* C *consists of the circuit* $((3,4),(4,3))$, *which implies that* $\sigma(C) = 2$. *Theorem 2.1.1 yields* $C^{n+2} = 3.5^{\otimes 2} \otimes C^n = 7 \otimes C^n$, *for* $n \geq 4$ *and*

$$\lim_{k\to\infty} \frac{(C^k \otimes x_0)_j}{k} = 3.5, \quad 1 \leq j \leq 4,$$

for any finite initial condition x_0.

Let $A \in \mathbb{R}_{\max}^{J \times J}$ and recall that the communication graph of A is denoted by $\mathcal{G}(A)$. For each circuit $\xi = ((i = i_1, i_2), (i_2, i_3), \ldots, (i_n, i_{n+1} = i))$, with arcs (i_m, i_{m+1}) in $\mathcal{G}(A)$ for $1 \leq m \leq n$, we define the average weight of ξ by

$$w(\xi) = \frac{1}{n} \bigotimes_{m=1}^{n} A_{i_{m+1} i_m} = \frac{1}{n} \sum_{m=1}^{n} A_{i_{m+1} i_m} .$$

Let $\mathcal{C}(A)$ denote the set of all circuits in $\mathcal{G}(A)$. One of the main results of deterministic max-plus theory is that for any irreducible square matrix A its eigenvalue can be obtained from

$$\lambda = \max_{\xi \in \mathcal{C}(A)} w(\xi) .$$

In words, the eigenvalue is equal to the maximal average circuit weight in $\mathcal{G}(A)$.

A circuit ξ in $\mathcal{G}(A)$ is called *critical* if its average weight is maximal, that is, if $w(\xi) = \lambda$. The critical graph of A, denoted by $\mathcal{G}^c(A)$, is the graph consisting of those nodes and arcs that belong to a critical circuit in $\mathcal{G}(A)$. Eigenvectors of A are characterized through the critical graph. However, before we are able to present the precise statement we have to introduce the necessary concepts from graph theory.

Let (E, V) denote a graph with set of nodes E and edges V. A graph is called *strongly connected* if for any two different nodes $i \in E$ and $j \in E$ there exists a path from i to j. For $i, j \in E$, we say that $i\mathcal{R}j$ if either $i = j$ or there exists a path from i to j and from j to i. We split (E, V) up into equivalence classes $(E_1, V_1), \ldots, (E_q, V_q)$ with respect to the relation $\mathcal{R}$. Any equivalence class (E_i, V_i), $1 \leq i \leq q$, constitutes a strongly connected graph. Moreover, (E_i, V_i) is maximal in the sense that we cannot add a node from (E, V) to (E_i, V_i) such that the resulting graph would still be strongly connected. For this reason we call $(E_1, V_1), \ldots, (E_q, V_q)$ *maximal strongly connected subgraphs* (m.s.c.s.) of (E, V). Note that this definition implies that an isolated node or a node with just incoming or outgoing arcs constitutes a m.s.c.s. with an empty arc set. We define the reduced graph, denoted by $(\tilde{E}, \tilde{V})$, by $\tilde{E} = \{1, \ldots, q\}$ and $(i, j) \in \tilde{V}$ if there exists $(k, l) \in V$ with $k \in E_i$ and $l \in E_j$. The *cyclicity* of a strongly connected graph is the greatest common divisor of the lengths of all circuits, whereas the cyclicity of a graph is the least common multiple of the cyclicities of the maximal strongly connected sub-graphs. As shown in [10], the cyclicity of a square matrix A (that is, $\sigma(A)$ in Theorem 2.1.1) is given by the cyclicity of the critical graph of A. A class of matrices that is of importance in applications are irreducible square matrices whose critical graph has a single m.s.c.s. of cyclicity one. Following [65], we call such matrices *primitive*. In the literature, primitive matrices are also referred to as *scs1-cyc1* matrices. For example, matrices A and B in Example 2.1.1 are primitive whereas matrix C in Example 2.1.1 is not.

Example 2.1.2 *We revisit the open tandem queuing system with initially one customer present at each server. The max-plus model for this system is given in*

Example 1.5.12. Suppose that the service times are deterministic, that is, $\sigma_j = \sigma_j(k)$ for $k \in \mathbb{N}$ and $0 \leq j \leq J$. The communication graph of $A = A_1(k)$ consists of the circuit $((0,1),(1,2),\ldots,(J,0))$ and the recycling loops $(0,0)$, $(1,1)$ to (J,J). Set

$$L = \{j : \sigma_j = \max\{\sigma_i : 0 \leq i \leq J\}\}.$$

We distinguish between three cases.

- *If $1 = |L|$, then the critical graph of A consists of the node $j \in L$ and the arc (j,j). The critical graph has thus a single m.s.c.s. of cyclicity one, A is therefore primitive.*

- *If $1 < |L| < J$, then the critical graph of A consists of the nodes $j \in L$ and the arcs (j,j), $j \in L$. The critical graph has thus $|L|$ m.s.c.s. each of which has cyclicity one and A fails to be primitive.*

- *If $|L| = J$, then the critical graph and the communication graph coincide and A. The critical graph has a single m.s.c.s. of cyclicity one, and A is primitive.*

Let $A \in \mathbb{R}_{\max}^{J \times J}$ be irreducible. Denote by A_λ the normalized matrix, that is, the matrix which is obtained by subtracting (in conventional algebra) the eigenvalue of A from all components, in formula: $(A_\lambda)_{ij} = A_{ij} - \lambda$, for $1 \leq i,j \leq J$. The eigenvalue of a normalized matrix is e. For a normalized matrix of dimension $J \times J$ we set

$$A^+ \stackrel{\text{def}}{=} \bigoplus_{k \geq 1} (A_\lambda)^k . \tag{2.1}$$

It can be shown that $A^+ = A_\lambda \oplus (A_\lambda)^2 \oplus \cdots \oplus (A_\lambda)^J$. See, for example, Lemma 2.2 in [65]. The eigenspaces of A and A_λ are equal. To see this, let $\mathbf{e}$ denote the vector with all components equal to e; for $x \in V(A)$, it then holds that

$$\lambda \otimes x = A \otimes x \Leftrightarrow x = A \otimes x - \lambda \otimes \mathbf{e} \Leftrightarrow e \otimes x = A_\lambda \otimes x .$$

The following theorem is an adaptation of Theorem 3.101 in [10] which characterizes the eigenspace of A_λ. We write $A_{\cdot i}$ to indicate the i^{th} column of A.

Theorem 2.1.2 *(Baccelli et al. [10]) Let A be irreducible and let A^+ be defined as in (2.1).*

(i) If i belongs to the critical graph, then $A^+_{\cdot i}$ is an eigenvector of A.

(ii) For i,j belonging to the critical graph, there exists $a \in \mathbb{R}$ such that

$$a \otimes A^+_{\cdot i} = A^+_{\cdot j}$$

if and only if i,j belong to the same m.s.c.s.

(iii) *Every eigenvector of A can be written as a linear combination of critical columns, that is, for every $x \in V(A)$ it holds that*

$$x = \bigoplus_{i \in G^c(A)} a_i \otimes A^+_{\cdot i} ,$$

where $G^c(A)$ denotes the set of nodes belonging to the critical graph, and $a_i \in \mathbb{R}^J_{\max}$ such that

$$\bigoplus_{i \in G^c(A)} a_i \neq \varepsilon .$$

Example 2.1.3 *Consider the matrix*

$$A = \begin{pmatrix} 0 & -2 \\ 1 & 0 \end{pmatrix} .$$

A is irreducible with eigenvalue 0 and the critical graph of A consists of the nodes $\{1, 2\}$ and recycling loops $(1, 1)$ and $(2, 2)$. The critical graph has thus two m.s.c.s., namely, the recycling loops $(1, 1)$ and $(2, 2)$, and $\sigma(A) = 1$. For A it holds that

$$A = A^n = A_\lambda = A^+ , \quad n \in \mathbb{N} .$$

Theorem 2.1.2 yields the following representation of the eigenspace of A: A vector $x \in \mathbb{R}^2_{\max}$ belongs to $V(A)$ if and only if numbers $a_1, a_2 \in \mathbb{R}_{\max}$ exist with $a_1 \oplus a_2 \neq \varepsilon$ (in words: at least one of two numbers is finite) such that

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = a_1 \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} \oplus a_2 \otimes \begin{pmatrix} -2 \\ 0 \end{pmatrix} ,$$

see (1.3) for the definition of scalar multiplication of vectors.

Let $A \in \mathbb{R}^{J \times J}_{\max}$ be irreducible with cyclicity one. Recall that we call $v, w \in \mathbb{R}^J_{\max}$ linear dependent if an $\alpha \in \mathbb{R}$ exists such that $v = \alpha \otimes w$. We say that *the eigenvector of A is unique* if any two eigenvectors of A are linear dependent, or, equivalently, if there exists $v \in \mathbb{R}^J$ such that

$$V(A) = \{\alpha \otimes v : \alpha \in \mathbb{R}\} .$$

This can conveniently be expressed by saying that the eigenspace of A reduces to a single point in $\mathbb{R}^J_{\max}$.

An important consequence of Theorem 2.1.2 is that eigenvectors of primitive matrices are unique. Primitive matrices enjoy the additional properties that, for sufficiently large k, $A^k \otimes x$ becomes an eigenvector of A for any finite vector x. These properties of primitive matrices will be of use in Section 2.5 and Section 2.6. The precise statement is as follows.

Corollary 2.1.1 *If $A \in \mathbb{R}^{J \times J}_{\max}$ is a primitive matrix, then the eigenvector of A is unique.*

Let $x(k+1) = A \otimes x(k)$, for $k \geq 0$, and let $x(0)$ be a finite vector. Then, it holds that $x(k) \in V(A)$ for $k \geq c(A)$. Specifically, it holds that

$$x(k+1) = \lambda \otimes x(k), \quad k \geq c(A),$$

where λ denotes the eigenvalue of A, and consequently, for $k \geq c(A)$, it holds that $||x(k)||_{\mathbb{P}} = a$ for some finite constant a.

Proof: Because A is primitive, the critical graph has only one m.s.c.s. Thus, by Theorem 2.1.2 (ii), there exists i_0 in the critical graph such that

$$A^+_{\cdot i} = \alpha_i \otimes A_{\cdot i_0}, \quad i \in G^c(A).$$

Hence, by Theorem 2.1.2 (iii), any eigenvector v of A can be written

$$\begin{aligned} v &= \bigoplus_{i \in G^c(A)} a_i \otimes A^+_{\cdot i} \\ &= \bigoplus_{i \in G^c(A)} a_i \otimes \left(\alpha_i \otimes A^+_{\cdot i_0}\right) \\ &= \left(\bigoplus_{i \in G^c(A)} a_i \otimes \alpha_i\right) \otimes A^+_{\cdot i_0} \\ &= \gamma \otimes A^+_{\cdot i_0}, \end{aligned}$$

where

$$\gamma = \bigoplus_{i \in G^c(A)} a_i \otimes \alpha_i \in \mathbb{R}_{\max},$$

which establishes uniqueness of the eigenvector.

We now turn to the proof of the second part of the corollary. Since A is primitive, $\sigma(A)$ in Theorem 2.1.1 is equal to one. This yields for $k \geq c(A)$: $A^{k+1} = \lambda \otimes A^k$ for any $k \geq c(A)$. Multiplying both sides of the above equation with the initial vector x_0 concludes the proof. □

Eigenvalues and eigenvectors of matrices over the max-plus semi-ring can be computed in an iterative way. A classical reference is [73]. For more methods for computing max-plus eigenvalues and eigenvectors we refer to [10, 65]. A recent alternative method based on policy iteration is given in [32], see also [65] for a detailed discussion. A general approach for computing cycle times (gives eigenvalues only) for so-called min-max-plus systems (an extension of max-plus linear systems) is established in [57, 56, 49]. Algorithms for computing eigenvalues and eigenvectors of both max-plus and min-max-plus systems can be found in [98, 101]. In particular, the algorithm given in [98] yields an upper bound for the cyclicity of a matrix in the max-plus semiring. Computing the eigenvalue of a matrix A can be achieved in polynomial time. In contrast to this, computing the coupling time is NP-hard (in the number of circuits of the critical graph), see [25]. Feasible upper bounds for the coupling time can be found in [60] and [25].

2.2 Subadditive Ergodic Theory (Type I)

Subadditive ergodic theory is based on Kingman's subadditive ergodic theorem [74, 75] and its application to generalized products of random matrices. We start with an elementary result which appears as an exercise in [91]. A sequence $a = \{a_n : n \in \mathbb{N}\}$ of real numbers is called subadditive if

$$a_{m+n} \leq a_n + a_m \,, \quad \text{for } n, m \geq 1 \,.$$

If a is subadditive, then a_n/n has a limit as $n \to \infty$, which may be $-\infty$. To see this, note that for given m, any n can be written as $n = k_n m + l_n$, where $l_n < m$ and k_n is a multiplier that depends on n. The subadditivity of a implies

$$a_n = a_{k_n m + l_n} \leq k_n a_m + a_{l_n} \,.$$

Dividing both sides by n yields

$$\frac{a_n}{n} = \frac{k_n}{n} a_m + \frac{1}{n} a_{l_n} \,.$$

Noticing that $k_n/n \leq 1/m$ and $k_n/n \to 1/m$, we have

$$\limsup_n \frac{a_n}{n} \leq \frac{a_m}{m} \,.$$

Since m is arbitrary, we may take the infimum w.r.t. m over the right-hand side and get

$$\limsup_n \frac{a_n}{n} \leq \liminf_m \frac{a_m}{m} \,.$$

Therefore, the limit a_n/n exists (and is equal to $\liminf_n a_n/n$).

Kingman's [75] result is formulated in terms of *subadditive processes.* These are double indexed processes $X = \{X_{mn} : m, n \in \mathbb{N}\}$ satisfying the following conditions:

(S1) If $i < j < k$, then $X_{ik} \leq X_{ij} + X_{jk}$ a.s.

(S2) For $m \geq 0$, the joint distributions of the process $\{X_{m+1n+1} : m < n\}$ are the same as those of $\{X_{mn} : m < n\}$.

(S3) The expected value $g_n = \mathbb{E}[X_{0n}]$ exists and satisfies $g_n \geq -cn$ for some finite constant $c > 0$ and all $n \geq 1$.

A consequence of **(S1)**, **(S3)** and the elementary result given above is that

$$\lambda = \lim_{n \to \infty} \frac{g_n}{n}$$

exists and is finite. We can now state Kingman's subadditive ergodic theorem: if X is a subadditive process (that is, **(S1)**, **(S2)** and **(S3)** hold), then the limit

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

exists almost surely, and $\mathbb{E}[\xi] = \lambda$. Condition **(S2)**, on the shift $\{X_{mn}\} \to \{X_{m+1n+1}\}$, is a stationarity condition. If all events defined in terms of X that are invariant under this shift have probability zero or one, then X is *ergodic*. In this case, as discussed in Kingman [75], the limiting random variable ξ is almost surely constant and equal to λ. Note that the limit also holds when expected values are considered.

We now turn to homogeneous equations, that is, to max-plus linear systems whose dynamic can be described via

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

for $k \geq 0$, with $x(0) = x_0$ given. In particular, we write

$$x(n+1, x_0) = \bigotimes_{k=0}^{n} A(k) \otimes x_0 \,, \quad n \geq 0 \,, \tag{2.2}$$

to indicate the initial value of the sequence. Recall that $\mathbf{e}$ denotes the vector with all components equal to e. We set

$$x_{nm} = \bigotimes_{k=n}^{m-1} A(k) \otimes \mathbf{e}$$

From this we recover $x(k+1, \mathbf{e})$ through $x_{0k+1} = x(k+1, \mathbf{e})$.

Lemma 2.2.1 *Let $\{A(k)\}$ be a stationary sequence of a.s. regular and integrable matrices in $\mathbb{R}_{\max}^{J\times J}$. Then $\{-||x_{nm}||_{\min} : m > n \geq 0\}$ and $\{||x_{nm}||_{\max} : m > n \geq 0\}$ are subadditive ergodic processes.*

Proof: For $x, y \in \mathbb{R}_{\max}^{J}$, let $x \leq y$ denote the component-wise order. Note that $x \leq y$ implies $||x||_{\max} \leq ||y||_{\max}$; in particular, $x \leq ||x||_{\max} \otimes \mathbf{e}$, where $\mathbf{e}$ denotes the vector whose components are equal to e (we refer to (1.3) for a definition of the $\otimes$-product of a scalar and a vector). Furthermore, for any $A \in \mathbb{R}_{\max}^{J\times J}$ it holds that $x \leq y$ implies $A \otimes x \leq A \otimes y$. Combining these statements it follows for $x \in \mathbb{R}_{\max}^{J}$ and $A \in \mathbb{R}_{\max}^{J\times J}$:

$$||A \otimes x||_{\max} \leq ||A \otimes (||x||_{\max} \otimes \mathbf{e})||_{\max} \,. \tag{2.3}$$

In the same vein, for $x \in \mathbb{R}_{\max}^{J}$ and $A \in \mathbb{R}_{\max}^{J\times J}$:

$$||A \otimes x||_{\min} \geq ||A \otimes (||x||_{\min} \otimes \mathbf{e})||_{\min} \,. \tag{2.4}$$

We now show the subadditive property of $||x_{nm}||_{\max}$. For $0 \leq n < p < m$, we obtain

$$
\begin{aligned}
||x_{nm}||_{\max} &= \left|\left| \bigotimes_{i=n}^{m-1} A(i) \otimes \mathbf{e} \right|\right|_{\max} \\
&= \left|\left| \bigotimes_{i=p}^{m-1} A(i) \otimes x_{np} \right|\right|_{\max} \\
&\overset{(2.3)}{\leq} \left|\left| \bigotimes_{i=p}^{m-1} A(i) \otimes \Big(||x_{np}||_{\max} \otimes \mathbf{e} \Big) \right|\right|_{\max} \\
&= \left|\left| ||x_{np}||_{\max} \otimes \left(\bigotimes_{i=p}^{m-1} A(i) \otimes \mathbf{e} \right) \right|\right|_{\max} \\
&= ||x_{np}||_{\max} + \left|\left| \bigotimes_{i=p}^{m-1} A(i) \otimes \mathbf{e} \right|\right|_{\max} \\
&= ||x_{np}||_{\max} + ||x_{pm}||_{\max} ,
\end{aligned}
$$

which establishes (**S1**) for $||x_{nm}||_{\max}$. The proof that (**S1**) holds for $-||x_{nm}||_{\min}$ as well follows the same line of argument: for $0 \leq n < p < m$,

$$
\begin{aligned}
||x_{nm}||_{\min} &= \left|\left| \bigotimes_{i=n}^{m-1} A(i) \otimes \mathbf{e} \right|\right|_{\min} \\
&= \left|\left| \bigotimes_{i=p}^{m-1} A(i) \otimes x_{np} \right|\right|_{\min} \\
&\overset{(2.4)}{\geq} \left|\left| \bigotimes_{i=p}^{m-1} A(i) \otimes \Big(||x_{np}||_{\min} \otimes \mathbf{e} \Big) \right|\right|_{\min} \\
&= \left|\left| ||x_{np}||_{\min} \otimes \left(\bigotimes_{i=p}^{m-1} A(i) \otimes \mathbf{e} \right) \right|\right|_{\min} \\
&= ||x_{np}||_{\min} + \left|\left| \bigotimes_{i=p}^{m-1} A(i) \otimes \mathbf{e} \right|\right|_{\min} \\
&= ||x_{np}||_{\min} + ||x_{pm}||_{\min} ,
\end{aligned}
$$

which establishes (**S1**) for $-||x_{nm}||_{\min}$.

The stationarity condition (**S2**) follows immediately from the stationarity of $\{A(k)\}$.

We now turn to condition (**S3**). We have assumed that each row of $A(k)$ contains at least one non-ε element, which implies $x(k, \mathbf{e}) \in \mathbb{R}^J$ for any k. We may now prove by induction that $x(k, \mathbf{e})$ is absolutely integrable where we use the fact that (i) $|\min(a,b)|, |\max(a,b)| \leq |a| + |b|$, (ii) $A(k)$ is integrable, and that (iii) the initial condition $\mathbf{e}$ of $x(k, \mathbf{e})$ is integrable. From

$$\mathbb{E}[\,||x_{0k}||_{\max}\,] = \mathbb{E}[\,||x(k,\mathbf{e})||_{\max}\,] \tag{2.5}$$

it follows that x_{0k} is integrable for any k. Let $|||A|||$ denote the smallest non-ε element of A (note that (i) and (ii) above imply that $\mathbb{E}[|||A(k)|||]$ is finite). With this definition it is immediate that

$$\sum_{j=0}^{k-1} \mathbb{E}[\,|||A(j)|||] \leq \mathbb{E}[\,||x(k,\mathbf{e})||_{\max}\,]\,. \tag{2.6}$$

Stationarity of $\{A(k)\}$ implies that $\mathbb{E}[\,|||A(k)|||] = c$ for any k. Integrability of $A(k)$ together with the fact that there are at least J finite elements in $A(k)$ yields $c > -\infty$. We obtain from (2.6):

$$\begin{aligned} -k\,|c| &\leq \mathbb{E}[\,||x(k,\mathbf{e})||_{\max}\,] \\ &\overset{(2.5)}{=} \mathbb{E}[\,||x_{0k}||_{\max}\,]\,, \end{aligned}$$

which establishes (**S3**) for $\{||x_{nm}||_{\max} : m \geq 1; m > n \geq 0\}$.

We now turn to $\{-||x_{nm}||_{\min} : m \geq 1; m > n \geq 0\}$. Following the above line of argument it holds that, for $k \in \mathbb{N}$,

$$\mathbb{E}[\,||x_{0k}||_{\min}\,] = \mathbb{E}[\,||x(k,\mathbf{e})||_{\min}\,] < \infty$$

and

$$\sum_{j=0}^{k-1} \mathbb{E}[\,||A(j)||_{\max}] \geq \mathbb{E}[\,||x_{0k}||_{\min}\,]\,.$$

Hence,

$$\sum_{j=0}^{k-1} -\mathbb{E}[\,||A(j)||_{\max}] \leq \mathbb{E}[\,-||x_{0k}||_{\min}\,]\,,$$

for $k \in \mathbb{N}$, and for $\tilde{c} = \mathbb{E}[\,||A(1)||_{\max}]$, we obtain

$$-|\tilde{c}|\,k \leq \mathbb{E}[\,-||x_{0k}||_{\min}]\,,$$

which concludes the proof of the lemma. □

The above lemma provides the means of applying Kingman's subadditive ergodic theorem to $||x(k)||_{\min}$ and $||x(k)||_{\max}$, respectively. The precise statement is given in the following theorem.

Theorem 2.2.1 *Let $\{A(k)\}$ be a stationary sequence of a.s. regular, integrable square matrices. Then, finite constants λ^{top} and λ^{bot} exist, so that for all (non-random) finite initial conditions x_0:*

$$\lambda^{\mathrm{bot}} \stackrel{\mathrm{def}}{=} \lim_{k\to\infty} \frac{||x(k)||_{\min}}{k} \leq \lambda^{\mathrm{top}} \stackrel{\mathrm{def}}{=} \lim_{k\to\infty} \frac{||x(k)||_{\max}}{k} \quad a.s.$$

and

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

The above limits also hold for random initial conditions provided that the initial condition is a.s. finite and integrable.

Proof: Lemma 2.2.1 applies and subadditivity of $||x(k,\mathrm{e})||_{\min}$ and $||x(k,\mathrm{e})||_{\max}$, respectively, follows. Therefore, Kingman's subadditive ergodic theorem applies and the proof with respect to the limit of $||x(k,\mathrm{e})||_{\min}$ as k tends to ∞ and the limit of $||x(k,\mathrm{e})||_{\max}$ as k tends to ∞ follows.

It remains to be shown that the limit exists for any finite initial condition. To see this note that for any finite initial condition y it holds that:

$$\begin{aligned} ||y||_{\min} + ||x(k,\mathrm{e})||_{\max} &= ||x(k, ||y||_{\min} \otimes \mathrm{e})||_{\max} \\ &\leq ||x(k,y)||_{\max} \\ &\leq ||x(k, ||y||_{\max} \otimes \mathrm{e})||_{\max} \\ &= ||y||_{\max} + ||x(k,\mathrm{e})||_{\max} \end{aligned}$$

(for a proof use the fact that $x \leq y$ implies $A \otimes x \leq A \otimes y$). Thus,

$$||y||_{\min} + ||x(k,\mathrm{e})||_{\max} \leq ||x(k,y)||_{\max} \leq ||y||_{\max} + ||x(k,\mathrm{e})||_{\max}$$

and, by similar arguments,

$$||y||_{\min} + ||x(k,\mathrm{e})||_{\min} \leq ||x(k,y)||_{\min} \leq ||x(k,\mathrm{e})||_{\min} + ||y||_{\max} .$$

Therefore, for $k > 0$,

$$\frac{1}{k}||y||_{\min} + \frac{1}{k}||x(k,\mathrm{e})||_{\max} \leq \frac{1}{k}||x(k,y)||_{\max} \leq \frac{1}{k}||x(k,\mathrm{e})||_{\max} + \frac{1}{k}||y||_{\max} \tag{2.7}$$

and

$$\frac{1}{k}||y||_{\min} + \frac{1}{k}||x(k,\mathrm{e})||_{\min} \leq \frac{1}{k}||x(k,y)||_{\min} \leq \frac{1}{k}||x(k,\mathrm{e})||_{\min} + \frac{1}{k}||y||_{\max} . \tag{2.8}$$

Letting k tend to infinity, it follows from (2.7) that the limits of $||x(k,\mathrm{e})||_{\max}/k$ and $||x(k,y)||_{\max}/k$ coincide. In the same vein, (2.8) implies that the limits of $||x(k,\mathrm{e})||_{\min}/k$ and $||x(k,y)||_{\min}/k$ coincide. If, in addition, x_0 is integrable, we first prove by induction that $x(k,x_0)$ is integrable for any $k > 0$. Then, we take expected values in (2.7) and (2.8). Using the fact that, by Kingman's subadditive ergodic theorem, the limits of $\mathbb{E}[||x(k,\mathrm{e})||_{\max}]/k$ and $\mathbb{E}[||x(k,\mathrm{e})||_{\min}]/k$ as k tends to ∞ exist, the proof follows from letting k tend to ∞. □

The constant λ^{top} is called the *top* or *maximal Lyapunov exponent* of $\{A(k)\}$ and λ^{bot} is called the *bottom Lyapunov exponent.*

Remark 2.2.1 *Irreducibility is a sufficient condition for $A(k)$ to be a.s. regular, see Remark 1.4.1. Therefore, in the literature, Theorem 2.2.1 is often stated with irreducibility as a condition.*

Remark 2.2.2 *Note that integrablity of $\{A(k)\}$ is a necessary condition for applying Kingman's subadditive ergodic theorem in the proof of the path-wise statement in Theorem 2.2.1.*

Remark 2.2.3 *Provided that (i) any finite element of $A(k)$ is positive, (ii) $A(k)$ is a.s. regular, and (iii) the initial state x_0 is positive, the statement in Theorem 2.2.1 holds for $||\cdot||_\oplus$ as well. This stems from the fact that under conditions (i) to (iii) it holds that $||A(k)||_{\max} = ||A(k)||_\oplus$. In particular, following the line of argument in the proof of Lemma 2.2.1, one can show that under the conditions of the lemma the sequence $||x_{nm}||_\oplus$ constitutes a subadditive process.*

2.2.1 The Irreducible Case

In this section, we consider stationary sequences $\{A(k)\}$ of integrable and irreducible matrices in $\mathbb{R}_{\max}^{J\times J}$ with the additional property that all finite elements are non-negative and that all diagonal elements are non-negative. We consider $x(k+1) = A(k) \otimes x(k)$, $k \geq 0$, and recall that $x(k)$ may model an autonomous system (for example, a closed queuing network). See Section 1.4.3. Indeed, $A(k)$ for the closed tandem queuing system in Example 1.5.1 is irreducible. As we will show in the following theorem, the setting of this section implies that $\lambda^{\text{top}} = \lambda^{\text{bot}}$, which in particular implies convergence of $x_i(k)/k$, $1 \leq i \leq J$. The condition that all finite elements of $A(k)$ are non-negative is not very restrictive when working with queuing networks. Here the non-ε elements of $A(k)$ represent sums of service times at the stations, which are by definition non-negative. In contrast, the assumption that all diagonal elements are non-negative (and thus different from ε) is indeed a restriction as illustrated by Example 1.5.5. The following theorem goes back to Cohen [35] and Baccelli et al. [10].

Theorem 2.2.2 *Let $\{A(k)\}$ be a stationary sequence of integrable and irreducible matrices in $\mathbb{R}_{\max}^{J\times J}$ such that all finite elements are non-negative and all diagonal elements are different from ε. Then, a finite constant λ exists, so that for any non-random finite initial condition x_0:*

$$\lim_{k\to\infty} \frac{x_j(k)}{k} = \lim_{k\to\infty} \frac{||x(k)||_{\min}}{k} = \lim_{k\to\infty} \frac{||x(k)||_{\max}}{k} = \lambda \quad a.s. \tag{2.9}$$

and

$$\lim_{k\to\infty} \frac{1}{k}\mathbb{E}[x_j(k)] = \lim_{k\to\infty} \frac{1}{k}\mathbb{E}[||x(k)||_{\min}] = \lim_{k\to\infty} \frac{1}{k}\mathbb{E}[||x(k)||_{\max}] = \lambda\,,$$

for $1 \leq j \leq J$. The above limits also hold true for random initial conditions provided that the initial condition is a.s. finite and integrable.

Proof: The existence of the limits (except that for $x_j(k)/k$) is guaranteed by Theorem 2.2.1 and in order to prove the theorem we have to show that the component-wise limits (that is, the limit of $x_j(k)/k$ as k tends to ∞, for $1 \leq j \leq J$) equal the limits of $||\cdot||_{\min}$ and $||\cdot||_{\max}$.

Irreducibility of $A(k)$ implies that $A(k)$ has fixed support and the communication graph of $A(k)$ is thus non-random. We have assumed that all elements different from ε are non-negative and all diagonal elements are non-negative. Hence, Lemma 1.4.1 applies and

$$G(k) = \bigotimes_{j=k-J}^{k-1} A(j), \quad k \geq J,$$

has all elements larger than or equal to zero for all k. This implies for any component j

$$\begin{aligned} x_j(k, \mathrm{e}) &= \bigoplus_{i=1}^{J} (G(k))_{ji} \otimes x_i(k-J, \mathrm{e}) \\ &\geq \bigoplus_{i=1}^{J} 0 \otimes x_i(k-J, \mathrm{e}) \\ &= ||x(k-J, \mathrm{e})||_{\max}, \end{aligned}$$

for $k \geq J$, which yields

$$||x(k, \mathrm{e})||_{\min} \geq ||x(k-J, \mathrm{e})||_{\max}. \tag{2.10}$$

By (2.10),

$$\frac{1}{k}||x(k, \mathrm{e})||_{\min} \geq \frac{1}{k}||x(k-J, \mathrm{e})||_{\max},$$

which implies

$$\lambda^{\mathrm{bot}} = \lim_{k\to\infty} \frac{||x(k, \mathrm{e})||_{\min}}{k} \geq \lim_{k\to\infty} \frac{||x(k, \mathrm{e})||_{\max}}{k} = \lambda^{\mathrm{top}} \quad \text{a.s.}$$

By Theorem 2.2.1, it holds that $\lambda^{\mathrm{bot}} \leq \lambda^{\mathrm{top}}$ and we have thus shown $\lambda^{\mathrm{bot}} = \lambda^{\mathrm{top}}$. In other words, setting $\lambda \stackrel{\mathrm{def}}{=} \lambda^{\mathrm{bot}} = \lambda^{\mathrm{top}}$ we have shown

$$\lim_{k\to\infty} \frac{||x(k, \mathrm{e})||_{\min}}{k} = \lim_{k\to\infty} \frac{||x(k, \mathrm{e})||_{\max}}{k} = \lambda \quad \text{a.s.} \tag{2.11}$$

and from

$$||x(k, \mathrm{e})||_{\max} \geq x_j(k, \mathrm{e}) \geq ||x(k, \mathrm{e})||_{\min}, \quad 1 \leq j \leq J,$$

follows:

$$\lim_{k\to\infty} \frac{x_j(k, \mathrm{e})}{k} = \lambda \quad \text{a.s.} \tag{2.12}$$

for $1 \leq j \leq J$.

Like for the proof of Theorem 2.2.1, we show that the limits in (2.11) and (2.12) are independent of the initial condition. This concludes the proof of the first part of the theorem.

We now turn to the proof of the second part of the theorem. Let λ, as defined in the first part of Theorem 2.2.2, exist. Then Theorem 2.2.1 yields,

$$\lambda = \lim_{k\to\infty} \frac{1}{k}\mathbb{E}[||x(k)||_{\max}] = \lim_{k\to\infty} \frac{1}{k}\mathbb{E}[||x(k)||_{\min}]$$

and

$$\frac{1}{k}\mathbb{E}[||x(k)||_{\min}] \leq \frac{1}{k}\mathbb{E}[x_j(k)] \leq \frac{1}{k}\mathbb{E}[||x(k)||_{\max}]$$

implies

$$\lambda = \lim_{k\to\infty} \frac{1}{k}\mathbb{E}[x_j(k)] ,$$

for $1 \leq j \leq J$. □

The constant λ, as defined in (2.9) in Theorem 2.2.2, is called *max-plus Lyapunov exponent* of the sequence of random matrices $\{A(k)\}$. There is no ambiguity in denoting the Lyapunov exponent of $\{A(k)\}$ and the eigenvalue of a matrix A by the same symbol, since for $A(k) = A$, for all k, the Lyapunov exponent of $\{A(k)\}$ is just the eigenvalue of A.

Remark 2.2.4 *Depending on the sequence $\{A(k)\}$, it is sometimes possible to replace an element of x_0 that is equal to ε by a finite element without changing the value of $x(k)$, for $k \geq 1$. In these cases, Theorem 2.2.2 applies even though not all elements of x_0 are finite.*

Remark 2.2.5 *We say that $A, B \in \mathbb{R}_{\max}^{J\times J}$ have the same structure if any element (ij) is either finite in A and B, or, is equal to ε (that is, the arc sets of communication graph of A and B coincide). The irreducibility condition in the above theorem can be replaced by the following weaker condition. There exists a.s. a sequence $\{m_n\}$ with $\lim_{n\to\infty} m_n = \infty$, such that $A(k + m_n)$, $1 \leq k \leq J$, have the same structure and are irreducible.*

Remark 2.2.6 *If the initial condition x_0 is positive, then the statement in Theorem 2.2.2 holds for $||\cdot||_\oplus$ as well. See Remark 2.2.3 for details.*

Computing exactly, or approximating the Lyapunov exponent of products of matrices over the max-plus semiring is a long standing problem [35, 96, 93, 10, 36, 46, 11, 50, 21, 8, 7, 42]. Only for special cases exact formulae are known. Upper and lower bounds can be found in [14, 18, 53, 28, 29]. In [12] approaches are described which use parallel simulation to estimate the ratio $x_j(k)/k$ for large k. When it comes to discrete event systems, Lyapunov exponents measure the cycle time, i.e., the average time between two events. A classical reference on Lyapunov exponents of products of random matrices is [24] and a more recent one, dedicated to non-negative matrices, is [66].

Consider the system in Example 1.5.1. If we assume that (i) the service times $\sigma_j(k)$ are i.i.d. with finite mean for each j and (ii) the sequences $\{\sigma_j(k)\}$ $(1 \leq j \leq J)$ are mutually independent, then Theorem 2.2.2 applies (indeed, $\{A(k)\}$ is an i.i.d. sequence of irreducible matrices with fixed support).

Comparing the conditions in Theorem 2.2.2 with those in Theorem 2.2.1, Theorem 2.2.2 imposes the additional conditions that (i) the matrices are irreducible (and have thus fixed support), (ii) all elements different from ε are non-negative and that (iii) all diagonal elements are non-negative. However, conditions (i)-(iii) are only needed to establish the pathwise statement in Theorem 2.2.2. Hence, the second part of Theorem 2.2.2 is valid under weaker conditions. The exact statement is as follows:

Corollary 2.2.1 *Let $\{A(k)\}$ be a stationary sequence of a.s. regular and integrable matrices in $\mathbb{R}_{\max}^{J\times J}$. If*

- $\lambda^{\mathrm{bot}} \geq \lambda^{\mathrm{top}}$, *and*
- *the initial condition is integrable,*

then

$$\lim_{k\to\infty} \frac{1}{k}\mathbb{E}[x_j(k)] = \lambda ,$$

for all components $1 \leq j \leq J$ of $x(k)$.

Proof: By assumption,

$$\lim_{k\to\infty} \frac{||x(k)||_{\min}}{k} = \lim_{k\to\infty} \frac{||x(k)||_{\max}}{k} = \lambda ,$$

with $\lambda = \lambda^{\mathrm{bot}} = \lambda^{\mathrm{top}}$, and Theorem 2.2.1 yields

$$\lim_{k\to\infty} \frac{||x(k)||_{\min}}{k} = \lim_{k\to\infty} \frac{1}{k}\mathbb{E}[||x(k)||_{\min}] = \lambda ,$$

$$\lim_{k\to\infty} \frac{||x(k)||_{\max}}{k} = \lim_{k\to\infty} \frac{1}{k}\mathbb{E}[||x(k)||_{\max}] = \lambda .$$

For any $k \in \mathbb{N}$ and $1 \leq j \leq J$,

$$\frac{1}{k}\mathbb{E}[||x(k)||_{\min}] \leq \frac{1}{k}\mathbb{E}[x_j(k)] \leq \frac{1}{k}\mathbb{E}[||x(k)||_{\max}]$$

and taking limits yields

$$\lambda = \lim_{k\to\infty} \frac{1}{k}\mathbb{E}[x_j(k)] ,$$

which concludes the proof. □

2.2.2 The Reducible Case

The setup is as in the previous section except that we now suppose that $A(k)$ has fixed support and drop the assumption that it is irreducible. An example of a model that has fixed support but fails to be irreducible is the open tandem queuing system in Example 1.5.2. We study the homogeneous equation

$$x(k+1) = A(k) \otimes x(k), \quad k \geq 0.$$

Notice that this setup comprises inhomogeneous equations, such as the standard autonomous equation as well, see Section 1.4.3 for details.

To deal with reducible matrices $A(k)$, we decompose $A(k)$ into its 'irreducible' components. The ergodic theorem, to be proved presently, then states that the Lyapunov exponent of the overall matrix is given by the maximal top Lyapunov exponent of its irreducible components. However, before we are able to present the ergodic theorem and give the proof, we need to introduce some concepts from graph theory. For the basic definitions we refer to Section 2.1.

Let $\{A(k)\}$ be a sequence of matrices in $\mathbb{R}_{\max}^{J \times J}$ with fixed support. If we replace any element of $A(k)$ that is different from ε by e, then the resulting communication graph of $A(k)$, denoted by $\mathcal{G}_e(A)$, is independent of k (and thus non-random). Let $\mathcal{G}_e^r(A)$ denote the reduced graph of $\mathcal{G}_e(A)$. We denote by $[i] \stackrel{\text{def}}{=} \{j \in \{1, \ldots, J\} : i \mathcal{R} j\}$ the set of nodes of the m.s.c.s. that contains i. The set of all nodes j such that there exists a path from j to i in $\mathcal{G}_e(A)$ is denoted by $\pi^+(i)$. Furthermore, we set $\pi^*(i) = \{i\} \cup \pi^+(i)$; and we define predecessor sets

$$[\leq i] = \bigcup_{j \in \pi^*(i)} [j]$$

and $[< i] = [\leq i] \setminus [i]$. We denote by $\lambda_{[i]}^{\text{top}}$ the top Lyapunov exponent associated with the matrix obtained by restricting $A(k)$ to the nodes in $[i]$. In case i is an isolated node or node with only incoming or outgoing arcs, we set $\lambda_{[i]}^{\text{top}} = \varepsilon$. The following theorem goes back to [6].

Theorem 2.2.3 *Let $\{A(k)\}$ be a stationary sequence of integrable matrices in $\mathbb{R}_{\max}^{J \times J}$ with fixed support such that with probability one all finite elements are non-negative and the diagonal elements are different from ε. For any (non-random) finite initial value x_0 it holds true that*

$$\lim_{k \to \infty} \frac{x_j(k)}{k} = \lambda_j \qquad a.s. ,$$

with

$$\lambda_j = \bigoplus_{i \in \pi^*(j)} \lambda_{[i]}^{\text{top}} ,$$

and

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

for $1 \leq j \leq J$. The above limits also hold for random initial conditions provided that the initial condition is a.s. finite and integrable.

Proof: Under the conditions of the theorem, it a.s. holds, for any k, that $||x(k)||_{\max} = ||x(k)||_{\oplus}$, see Remark 2.2.3. In the following proof we will only work with upper bounds on the growth rate $||x(k)||_{\max}/k$ and thus adopt the notation $||\cdot||_{\oplus}$ for the maximal element of a vector/matrix.

Let $A_{[i]\,[i]}(k)$ denote the matrix that is obtained from $A(k)$ by restricting $A(k)$ to the nodes in $[i]$ and write $x_{[i]}(k)$ for $x(k)$ restricted to the nodes in $[i]$. To understand the difficulty that arises when proving the theorem, it is worth noting that in general

$$\lim_{k\to\infty} \frac{1}{k} ||x_{[i]}(k)||_{\oplus} \neq \lambda^{\text{top}}_{[i]} \text{ a.s.}$$

This stems from the fact that $\lambda^{\text{top}}_{[i]}$ is the top Lyapunov exponent of the matrix restricted to the nodes in $[i]$, whereas $x_{[i]}(k)$ is also influenced by nodes others than those in $[i]$ namely those in $[\leq i] \setminus [i]$.

We now turn to the proof. In the same way as we have defined $A_{[i]\,[i]}(k)$ and $x_{[i]}(k)$, we write $A_{[\leq i]\,[\leq i]}(k)$ for the restriction of $A(k)$ to the nodes in $[\leq i]$ and $x_{[\leq i]}(k)$ for $x(k)$ restricted to the nodes in $[\leq i]$. By Theorem 2.2.1, the maximal Lyapunov exponent of $A_{[\leq i]\,[\leq i]}(k)$, given by $\lambda^{\text{top}}_{[\leq i]}$, exists (indeed, Theorem 2.2.1 applies to reducible matrices). Note that

$$\frac{1}{k} ||x_{[i]}(k)||_{\oplus} \leq \frac{1}{k} ||x_{[\leq i]}(k)||_{\oplus}$$

and thus

$$\begin{aligned} \limsup_{k\to\infty} \frac{1}{k} ||x_{[i]}(k)||_{\oplus} &\leq \limsup_{k\to\infty} \frac{1}{k} ||x_{[\leq i]}(k)||_{\oplus} \\ &= \lambda^{\text{top}}_{[\leq i]} \,. \end{aligned} \tag{2.13}$$

Fixed support of $A(k)$ implies that $\mathcal{G}_e(A)$ is non-random. Node i can be reached from any node $h \in \pi^*(i)$ and since $A(k)$ is of dimension $J \times J$ such a path is at most of length J. We have assumed that the diagonal elements of $A(k)$ are all different from ε. Hence, if there is a path of length l from h to i, then there is for any $p \geq l$ a path of length p from h to i (just add sufficiently many loops of length one at h). Any finite element of $A(k)$ is positive and paths have therefore positive weights. We thus obtain for any $j \in [i]$

$$\begin{aligned} x_j(k) &\geq \bigoplus_{h\in\pi^*(i)} x_h(k-J) \\ &= ||x_{[\leq i]}(k-J)||_{\oplus} \,, \end{aligned} \tag{2.14}$$

for $k \geq J$. Therefore,

$$||x_{[i]}(k)||_{\oplus} \geq ||x_{[\leq i]}(k-J)||_{\oplus} \,,$$

for $k \geq J$, which implies that

$$\begin{aligned} \liminf_{k\to\infty} \frac{1}{k} ||x_{[i]}(k)||_{\oplus} &\geq \liminf_{k\to\infty} \frac{1}{k} ||x_{[\leq i]}(k)||_{\oplus} \\ &= \lambda^{\text{top}}_{[\leq i]} \quad \text{a.s.} \end{aligned}$$

Together with (2.13) we obtain

$$\lim_{k\to\infty}\frac{1}{k}||x_{[i]}(k)||_\oplus = \lambda^{\mathrm{top}}_{[\le i]} \quad \text{a.s.} \tag{2.15}$$

By (2.14), it holds a.s. for any $j \in [i]$ that

$$\frac{1}{k}||x_{[i]}(k)||_\oplus \ge \frac{1}{k}x_j(k) \ge \frac{1}{k}||x_{[\le i]}(k-J)||_\oplus , \tag{2.16}$$

and by (2.15) it follows that

$$\lim_{k\to\infty}\frac{1}{k}||x_{[i]}(k)||_\oplus = \lim_{k\to\infty}\frac{1}{k}||x_{[\le i]}(k-J)||_\oplus = \lambda^{\mathrm{top}}_{[\le i]} , \tag{2.17}$$

which yields

$$\lim_{k\to\infty}\frac{1}{k}x_j(k) = \lambda^{\mathrm{top}}_{[\le i]} \quad \text{a.s.}, \quad j \in [i] .$$

In the integrable case, (2.16) implies

$$\frac{1}{k}\mathbb{E}[||x_{[\le i]}(k)||_\oplus] \ge \frac{1}{k}\mathbb{E}[x_j(k)] \ge \frac{1}{k}\mathbb{E}[||x_{[\le i]}(k-J)||_\oplus] .$$

By Theorem 2.2.1, the expected values on the right-hand side and on the left-hand side in the above inequality converge to $\lambda^{\mathrm{top}}_{[\le i]}$ as k tends to ∞. Hence,

$$\lim_{k\to\infty}\frac{1}{k}\mathbb{E}[x_j(k)] = \lambda^{\mathrm{top}}_{[\le i]} , \quad j \in [i] .$$

It remains to be shown that

$$\lambda^{\mathrm{top}}_{[\le i]} = \bigoplus_{j\in\pi^*(i)} \lambda_{[j]} . \tag{2.18}$$

The reduced graph $\mathcal{G}^r_e(A)$ is acyclic and we obtain

$$x_{[i]}(k+1) = A_{[i]\,[i]}(k) \otimes x_{[i]}(k) \oplus s(i, k+1) , \tag{2.19}$$

where

$$s(i,k+1) \stackrel{\mathrm{def}}{=} A_{[i]\,[<i]}(k) \otimes x_{[<i]}(k)$$

and $A_{[i]\,[<i]}(k)$ is defined in the obvious way. By definition,

$$\begin{aligned} ||s(i,k+1)||_\oplus &\le ||A_{[i]\,[<i]}(k)||_\oplus \otimes ||x_{[<i]}(k)||_\oplus \\ &\le ||A(k)||_\oplus \otimes ||x_{[<i]}(k)||_\oplus . \end{aligned} \tag{2.20}$$

Note that

$$\lim_{k\to\infty}\frac{1}{k}||A(k)||_\oplus = 0 \quad \text{a.s.} \tag{2.21}$$

Indeed, integrability of $\{A(k)\}$ together with stationarity and ergodicity implies that

$$\lim_{k\to\infty}\frac{1}{k}\sum_{n=1}^{k}||A(k)||_\oplus = \mathbb{E}[||A(1)||_\oplus] < \infty \quad \text{a.s.}$$

(integrability of $||A(1)||_\oplus$ is guaranteed by integrability of $A(1)$), which gives

$$\begin{aligned}\mathbb{E}[||A(1)||_\oplus] &= \lim_{k\to\infty} \frac{1}{k}\sum_{n=1}^{k} ||A(n)||_\oplus \\ &= \lim_{k\to\infty} \frac{k-1}{k}\frac{1}{k-1}\sum_{n=1}^{k-1} ||A(n)||_\oplus + \lim_{k\to\infty}\frac{1}{k}||A(k)||_\oplus \\ &= \mathbb{E}[||A(1)||_\oplus] + \lim_{k\to\infty}\frac{1}{k}||A(k)||_\oplus \quad \text{a.s.}\end{aligned}$$

and thus establishes (2.21).

We obtain from (2.20) together with (2.21)

$$\limsup_{k\to\infty}\frac{1}{k}||s(i,k+1)||_\oplus \le \lambda^{\text{top}}_{[<i]} \quad \text{a.s.}$$

At the same time, following the line of argument that has lead to (2.14), we obtain

$$||s(i,k+1)||_\oplus \ge ||x_{[<i]}(k-J)||_\oplus \quad \text{a.s.},$$

which implies

$$\liminf_{k\to\infty}\frac{1}{k}||s(i,k+1)||_\oplus \ge \lambda^{\text{top}}_{[<i]} \quad \text{a.s.}$$

and thus

$$\lim_{k\to\infty}\frac{1}{k}||s(i,k+1)||_\oplus = \lambda^{\text{top}}_{[<i]} \quad \text{a.s.}$$

It is clear from the definition of $s(i,k)$ that

$$||x_{[\le i]}(k)||_\oplus \ge ||s(i,k)||_\oplus ,$$

so that

$$\frac{1}{k}||x_{[\le i]}(k)||_\oplus \ge \frac{1}{k}||s(i,k)||_\oplus ,$$

which in turn implies

$$\lambda^{\text{top}}_{[\le i]} \ge \lambda^{\text{top}}_{[<i]} . \tag{2.22}$$

Now suppose that $\lambda^{\text{top}}_{[\le i]} > \lambda^{\text{top}}_{[<i]}$. The existence of the individual limits implies that for sufficiently large $K \in \mathbb{N}$ it holds that

$$A_{[i]\,[i]}(k) \otimes x_{[i]}(k) \ge s(i,k+1) , \quad k \ge K .$$

Accordingly, equation (2.19) reads

$$x_{[i]}(k+1) = A_{[i]\,[i]}(k) \otimes x_{[i]}(k) \ge s(i,k+1) , \quad k \ge K ,$$

which, by Theorem 2.2.1, yields

$$\lim_{k\to\infty}\frac{1}{k}||x_{[i]}(k)||_\oplus = \lambda^{\text{top}}_{[i]} \quad \text{a.s.}$$

and, by (2.17), this implies

$$\lambda^{\text{top}}_{[\leq i]} = \lambda^{\text{top}}_{[i]} .$$

We have thus shown that

$$\lambda^{\text{top}}_{[\leq i]} > \lambda^{\text{top}}_{[<i]} \quad \Rightarrow \quad \lambda^{\text{top}}_{[\leq i]} = \lambda^{\text{top}}_{[i]} . \tag{2.23}$$

Combining (2.22) and (2.23) we reach at:

$$\lambda^{\text{top}}_{[\leq i]} = \lambda^{\text{top}}_{[<i]} \oplus \lambda^{\text{top}}_{[i]} .$$

Any node $i \in \mathcal{G}_e(A)$ belongs to a m.s.c.s. that is represented in $\mathcal{G}^r_e(A)$ by the single node $[i]$. Let $\pi([i])$ denote the set of direct predecessors of $[i]$ in $\mathcal{G}^r_e(A)$ and set $\pi([i]) = \emptyset$ if there is no predecessor. Each element of $\pi([i])$ represents a m.s.c.s. in $\mathcal{G}_e(A)$ and we denote by $\tau(i)$ the set of nodes in $\mathcal{G}_e(A)$ that belong to the m.s.c.s. corresponding to the elements of $\pi([i])$. If $\pi([i]) = \emptyset$, we set $\tau(i) = \emptyset$. Then

$$\lambda^{\text{top}}_{[<i]} = \bigoplus_{j \in \tau(i)} \lambda^{\text{top}}_{[\leq j]}$$

and inserting this into the above equation yields

$$\lambda^{\text{top}}_{[\leq i]} = \lambda^{\text{top}}_{[i]} \oplus \bigoplus_{j \in \tau(i)} \lambda^{\text{top}}_{[\leq j]} .$$

We now repeat the argument until applying τ yields no more nodes. In particular, going from $\tau(i)$ to $\{\tau(j) : j \in \tau(i)\}$ and so forth, we will eventually cover the set $\pi^*(i)$. This concludes the proof of (2.18). □

Remark 2.2.7 *Suppose that the conditions in Theorem 2.2.3 are satisfied. Continuity of the operators* max *and* min *yields that it holds with probability one that*

$$\lambda^{\text{bot}} = \min(\lambda_j : 1 \leq j \leq J)$$

and

$$\lambda^{\text{top}} = \max(\lambda_j : 1 \leq j \leq J) .$$

The vector $\vec{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_J)$, with λ_j defined in Theorem 2.2.3, is called the *Lyapunov vector* of $\{A(k)\}$. In the light of Theorem 2.2.2 we can state that irreducibility of $\{A(k)\}$ is a sufficient condition for the components of $\vec{\lambda}$ to be equal.

Recalling that $\lim_{k\to\infty} x_j(k)/k$ is the (asymptotic) speed with which transition j operates, the above theorem matches our intuition that the (asymptotic) speed with which the system operates is determined by the slowest component of the system. In terms of queuing networks, the throughput of a system is determined by the smallest throughput of one of its components. Moreover, if the queuing network is irreducible in the max-plus sense, then the throughput is the same at any station.

The key conditions on $A(k)$ are that any element of $A(k)$ is either equal to ε or non-negative, that the elements on the diagonal are non-negative and that it has fixed support. As we have already explained, the condition that any element different from ε has only non-negative values is a natural condition for queuing systems, and all examples presented in this monograph enjoy this property. The fixed support condition is satisfied by the queuing systems in Example 1.5.1 and Example 1.5.2. An example of a system that fails to have fixed support is given in Example 1.5.5. Such a system cannot be analyzed via the subadditive ergodic theory developed so far.

2.2.3 Variations and Extensions

One of the marvels of max-plus theory is that the existence of the top and bottom Lyapunov exponent follows so easily from Kingman's subadditive ergodic theorem. See the proof of Theorem 2.2.1. However, the conditions in Theorem 2.2.1 are too weak to guarantee that the top and bottom Lyapunov exponents are equal, or, in other words, that the individual growth rates (that is, $\lim_{k\to\infty} x_i(k)/k$, $1 \leq j \leq J$) have the same limit. In this section, we discuss approaches to establish equality of the top and bottom Lyapunov exponent without imposing conditions on the elements of $A(k)$.

2.2.3.1 The 'Up-Crossing' Property

In order to show that the individual growth rates coincide we had to impose the assumption that (i) any non-ε element of $A(k)$ is non-negative, that (ii) all diagonal elements are non-negative, and that (iii) $A(k)$ has fixed support. The 'non-negativity' condition on the finite elements causes no restriction for queuing systems. Therefore, we focus in this section on a relaxation of the 'fixed support' and the 'diagonal' condition.

Inspecting the proof of Theorem 2.2.2 one sees that what is actually needed is the following 'up-crossing' property: a subsequence $\{x(k_n)\}$ and a constant M exist, such that for any $n \geq 1$

$$||x(k_n + M)||_{\min} \geq a_n + b_n ||x(k_n)||_{\max} \quad \text{a.s.},$$

with

$$\lim_{n\to\infty} \frac{a_n}{n} = 0 \quad \text{and} \quad \lim_{n\to\infty} b_n = 1,$$

see (2.10) on page 72 in the proof of Theorem 2.2.2, where $a_n = 0$ and $b_n = 1$ for all n. Indeed, Vincent uses in [102] this type of condition to show that the top and bottom Lyapunov exponent coincide. Provided that finite elements of $A(k)$ are positive, the diagonal condition together with fixed support are sufficient for the above 'up-crossing' property to hold, see Lemma 1.4.1.

2.2.3.2 The 'Memory Loss' Property

In this section we present an alternative approach to finding sufficient conditions for $\lambda^{\text{top}} = \lambda^{\text{bot}}$. This approach goes back to [48, 84] and applies to sequences with countable state-space.

The key observation for this approach is the following. Let $A \in \mathbb{R}_{\max}^{J \times J}$ be such that any two columns of A are linear dependent. Then, a finite number a exists such that

$$||A \otimes x||_{\max} - ||A \otimes x||_{\min} = a\,, \quad x \in \mathbb{R}^J \tag{2.24}$$

(for a proof use the argument put forward in the proof of Corollary 2.1.1). A matrix with the property that any two columns are linear dependent is said to be of *rank 1*. While the notation of rank 1 is undisputed, there are several notions of rank in the literature, see [37] and [103].

Definition 2.2.1 *A sequence $\{A(k)\}$ of square matrices is said to have memory loss property (MLP) if there exists an N such that $A(N-1) \otimes A(N-2) \otimes \cdots \otimes A(0)$ with positive probability has only mutually linear dependent columns, i.e., is of rank 1.*

Let A be a matrix with mutually linear dependent columns and assume that $\{A(k)\}$ has MLP with respect to A and N, that is, assume that a finite number N exists such that $P(A(N-1) \otimes A(N-2) \otimes \cdots \otimes A(0) = A) > 0$ and A is of rank 1. Let

$$T_0 \;=\; \inf\{k \geq N-1 : A(k) \otimes A(k-1) \otimes \cdots \otimes A(k-N+1) \;=\; A\}$$

denote the first time a partial product of the series of matrix generates A. This gives

$$x(T_0) \;=\; A \otimes \bigotimes_{k=0}^{T_0 - N} A(k) \otimes x_0\,,$$

where we set the product to E for $T_0 = N-1$ and we assume that $x_0 \in \mathbb{R}^J$. By (2.24) a finite number a exists such that

$$||x(T_0)||_{\max} - ||x(T_0)||_{\min} \;=\; a\,,$$

for any finite initial value x_0. For $n \geq 0$, introduce the time of the $(n+1)^{st}$ occurrence of the event that a partial product of $\{A(k)\}$ generates A by

$$T_{n+1} = \inf\{k \geq N + T_n : A(k) \otimes A(k-1) \otimes \cdots \otimes A(k-N+1) = A\} \tag{2.25}$$

and we obtain

$$||x(T_k)||_{\max} - ||x(T_k)||_{\min} \;=\; a\,, \quad k \geq 0\,. \tag{2.26}$$

If $\{A(k)\}$ is stationary and ergodic, then $\lim_{n \to \infty} T_n = \infty$ and $T_n < \infty$ with probability one; for details see Section E.3 in the Appendix. Specifically, by equation (2.26),

$$\lim_{k \to \infty} \frac{1}{T_k} ||x(T_k)||_{\max} - \frac{1}{T_k} ||x(T_k)||_{\min} \;=\; 0 \quad \text{a.s.} \tag{2.27}$$

If, in addition, $\{A(k)\}$ is a sequence of a.s. regular and integrable matrices, Theorem 2.2.1 yields

$$\lim_{k\to\infty} \frac{1}{T_k}||x(T_k)||_{\max} = \lambda^{\rm top} \text{ and } \lim_{k\to\infty} \frac{1}{T_k}||x(T_k)||_{\min} = \lambda^{\rm bot}$$

with probability one, and equality of $\lambda^{\rm top}$ and $\lambda^{\rm bot}$ follows from (2.27). We summarize our analysis in the following theorem:

Theorem 2.2.4 *Let $\{A(k)\}$ be a stationary and ergodic sequence of integrable and a.s. regular matrices in $\mathbb{R}_{\max}^{J\times J}$. If $\{A(k)\}$ has MLP, then a finite constant λ exists such that, for any (non-random) finite initial conditions x_0:*

$$\lim_{k\to\infty} \frac{x_j(k)}{k} = \lim_{k\to\infty} \frac{||x(k)||_{\min}}{k} = \lim_{k\to\infty} \frac{||x(k)||_{\max}}{k} = \lambda \quad a.s.$$

and

$$\lim_{k\to\infty} \frac{1}{k}\mathbb{E}\big[x_j(k)\big] = \lim_{k\to\infty} \frac{1}{k}\mathbb{E}\big[||x(k)||_{\min}\big] = \lim_{k\to\infty} \frac{1}{k}\mathbb{E}\big[||x(k)||_{\max}\big] = \lambda\,,$$

for $1 \le j \le J$. The above limits also hold for random initial conditions provided that the initial condition is a.s. finite and integrable.

It is worth noting that, in contrast to Theorem 2.2.3, the Lyapunov exponent is unique, or, in other words, the components of the Lyapunov vector are equal. In view of Theorem 2.2.2 the above theorem can be phrased as follows: Theorem 2.2.2 remains valid in the presence of reducible matrices if MLP is satisfied.

MLP is a technical condition and typically impossible to verify directly. A sufficient condition for $\{A(k)\}$ to have MLP is the following:

(C) There exists a primitive matrix C and $N \in \mathbb{N}$ such that

$$P\Big(A(N-1) \otimes A(N-2) \otimes \cdots \otimes A(0) = C \Big) > 0\,.$$

The following lemma illustrates the close relationship between primitive matrices and matrices of rank 1.

Lemma 2.2.2 *If A is primitive with coupling time c, then A^c has only finite entries and is of rank 1. Moreover, for any matrix A that has only finite entries it holds that A is of rank 1 if and only if the projective image of A is a single point in the projective space.*

Proof: We first prove the second part of the lemma. '$\Rightarrow$': Let $A \in \mathbb{R}_{\max}^{I\times J}$ be such that all elements are finite and that it is of rank 1. Denote the j^{th} column of A by $A_{\cdot j}$. Since A is of rank 1, there exits finite numbers α_j, with $2 \le j \le J$, such that $A_{\cdot 1} = \alpha_j \otimes A_{\cdot j}$ for $2 \le j \le J$. Hence, for $x \in \mathbb{R}^J$ it holds that

$$A \otimes x = \bigotimes_{j=1}^{J} \alpha_j \otimes x_j \otimes A_{\cdot 1}, \tag{2.28}$$

with $\alpha_1 = 0$. Let $\gamma_x \stackrel{\text{def}}{=} \bigotimes_{j=1}^{J} \alpha_j \otimes x_j$. Let $y \in \mathbb{R}^J$, with $x \neq y$. By (2.28), $A \otimes x = \gamma_x \otimes A_{\cdot 1}$ and $A \otimes y = \gamma_y \otimes A_{\cdot 1}$, which implies that $A \otimes x$ and $A \otimes y$ are linear dependent. Hence, the projective image of A contains only the single point $\overline{A_{\cdot 1}}$.

'$\Leftarrow$': We give a proof by contradiction. Suppose that A is not of rank 1, then there exist at least two columns $A_{\cdot j}$ and $A_{\cdot i}$ of A such that $A_{\cdot j}$ and $A_{\cdot i}$ are linear independent. Then $x^i, x^j \in \mathbb{R}^J$ can be chosen such that $A \otimes x^i = \beta^i \otimes A_{\cdot i}$ and $A \otimes x^j = \beta^j \otimes A_{\cdot j}$ for finite constants β^i, β^j. Since $A_{\cdot j}$ and $A_{\cdot i}$ are linear independent, the projective image of A contains at least the two distinct points $\overline{A_{\cdot i}}$ and $\overline{A_{\cdot j}}$.

We now turn to the proof of the first part of the lemma. For $1 \leq j \leq J$, let $\mathbf{e}_j$ be the vector with ε entries except for element j which is equal to e. Hence, $A^c \otimes \mathbf{e}_j = A^c_{\cdot j}$, where $A^c_{\cdot j}$ denotes the j^{th} column of A^c. By Theorem 2.1.1,

$$A \otimes A^c_{\cdot j} = A \otimes A^c \otimes \mathbf{e}_j = \lambda \otimes A^c \otimes \mathbf{e}_j = \lambda \otimes A^c_{\cdot j},$$

with λ the unique eigenvector of A, and the columns of A^c are thus eigenvectors of A. Using the fact that eigenvectors of irreducible matrices have only finite entries (see, for example, Lemma 2.8 in [65]), it follows that A^c has only finite elements. On the one hand, by Corollary 2.1.1, the eigenvector of A is unique. On the other hand, by Theorem 2.1.1, $A^c \otimes x$ is an eigenvector of A for any x. Hence, the projective image of A is a single point (in formula: $\exists v \in \mathbb{PR}^J \; \forall x \in \mathbb{R}^J : \overline{A^c \otimes x} = v$). Applying the second part of the lemma then proves the claim. □

We present a version of Theorem 2.2.4 with a condition that can be directly verified.

Lemma 2.2.3 *Let $\{A(k)\}$ be an i.i.d. sequence of a.s. regular integrable matrices in $\mathbb{R}^{J\times J}_{\max}$ with countable state space. If condition* (**C**) *holds, then the statement put forward in Theorem 2.2.4 holds.*

Proof: Let C be as given as in (**C**) and denote the coupling time of C by c. Because $\{A(k)\}$ is i.i.d. with countable state-space,

$$P\Big(A(N-1) = A(N-2) = \cdots = A(0) = C \Big) > 0\,,$$

implies

$$P\Big(A(cN-1) \otimes A(cN-2) \otimes \cdots \otimes A(0) = C^c \Big) > 0.$$

Since C is primitive, Lemma 2.2.2 implies that C^c is of rank 1 and $\{A(k)\}$ has thus MLP. Hence, Theorem 2.2.4 applies. □

Example 2.2.1 *Consider Example 1.5.5. Matrix D_2 is primitive. Hence, applying Lemma 2.2.3 shows that the Lyapunov exponent of the system exists.*

Remark 2.2.8 *In principle, MLP and condition* (**C**) *restrict the class of sequences $\{A(k)\}$ that can be analyzed to those with countable state-space. A* *possible generalization is the following. Suppose that the distribution of $A(k)$ is a mixture of a discrete distribution on a countable state-space, say $\mathcal{A}^c$, and a general distribution on an arbitrary state-space, say $\mathcal{A}^g$. If we require $P(A(N-1) \otimes A(N-2) \otimes \cdots \otimes A(0) \in \mathcal{A}^c) > 0$ in Definition 2.2.1 and $C \in \mathcal{A}^c$ in condition* (**C**)*, respectively, then the results in this section hold for $\{A(k)\}$ with state space $\mathcal{A}^c \cup \mathcal{A}^g$ as well.*

We conclude this section by presenting a generalization of Theorem 2.2.4. As Baccelli and Mairesse show in [11], using the arguments put forward in this section, a limit result can be obtained under a slightly weaker condition than MLP.

Theorem 2.2.5 *Let $\{A(k)\}$ be a stationary and ergodic sequence of integrable and a.s. regular square matrices in $\mathbb{R}_{\max}^{J\times J}$. If there exists $N \in \mathbb{N}$ such that with positive probability $A(N-1) \otimes A(N-2) \otimes \cdots \otimes A(0)$ has a bounded projective image, then the statement put forward in Theorem 2.2.4 holds.*

Proof: By assumption, there exist finite numbers $a, b \in \mathbb{R}$ such that

$$\forall x \in \mathbb{R}^J : \quad ||A(N-1) \otimes A(N-2) \otimes \cdots \otimes A(0) \otimes x||_{\mathbb{P}} \in [a, b].$$

In analogy to (2.25), let T_k denote the time index such that for the k^{th} time a product $A(T_k) \otimes A(T_k - 1) \otimes \cdots \otimes A(T_k - N + 1)$ has been observed whose projective image lies within the interval $[a, b]$; in formula:

$$a \leq ||x(T_k)||_{\max} - ||x(T_k)||_{\min} \leq b$$

for all k. We have assumed that $\{A(k)\}$ is stationary and ergodic, which implies $\lim_{n\to\infty} T_k = \infty$ and $T_k < \infty$ with probability one; for details see Section E.3 in the Appendix. Since $[a, b]$ is compact, the Bolzano-Weierstrass Theorem yields the existence of a subsequence $\{T_{k_n}\}$ of $\{T_k\}$ such that

$$\lim_{n\to\infty} ||x(T_{k_n})||_{\max} - ||x(T_{k_n})||_{\min} = c,$$

for some finite constant c, which implies

$$\lim_{n\to\infty} \frac{1}{k_n} ||x(T_{k_n})||_{\max} = \lim_{n\to\infty} \frac{1}{k_n} ||x(T_{k_n})||_{\min}. \tag{2.29}$$

By Theorem 2.2.1, convergence of the sequences $||x(k)||_{\max}/k$ and $||x(k)||_{\max}/k$ as k tends to infinity is guaranteed. Hence,

$$\begin{aligned} \lim_{k\to\infty} \frac{1}{k} ||x(k)||_{\max} &= \lim_{n\to\infty} \frac{1}{k_n} ||x(T_{k_n})||_{\max} \\ &\overset{(2.29)}{=} \lim_{n\to\infty} \frac{1}{k_n} ||x(T_{k_n})||_{\min} \\ &= \lim_{k\to\infty} \frac{1}{k} ||x(k)||_{\min}, \end{aligned}$$

which proves the claim. □

2.2.3.3 Weak Irreducibility

An approach relaxing the concept of fixed support can be found in [69, 70]. This approach is based on an interpretation of the concept of 'irreducibility' for random matrices which we will explain in the following.

Irreducibility of a matrix A is defined via the communication graph of A, denoted by $\mathcal{G}(A)$. Specifically, A is called irreducible if for any two nodes in $\mathcal{G}(A)$ a path from i to j exists in $\mathcal{G}(A)$. Let $\{A(k)\}$ be a random sequence of $J \times J$ dimensional matrices. The communication graph of a random sequence is itself a random variable and we extend the definition of a path to the sequence $\mathcal{G}(A(k))$ as follows. For any two nodes i, j, a sequence of arcs $\rho = ((i_n, j_n) : 1 \leq n \leq m)$, with $i = i_1$, $j = j_m$ and $j_n = i_{n+1}$ for $1 \leq n < m$, is called a *path* of *length* m from i to j in $\{A(k)\}$ if (i_n, j_n) is an arc in $\mathcal{G}(A(k+n-1))$ for $1 \leq n \leq m$, for some $k \in \mathbb{N}$. We say that ρ is a path in $\mathcal{G}(A(k+n-1) : 1 \leq n \leq m)$.

The *weight of a path* in $\mathcal{G}(A)$ is defined by the sum of the weights of all arcs constituting the path; more formally: let $\rho = ((i_n, j_n) : 1 \leq n \leq m)$ be a path from i to j of length m, then the weight of ρ, denoted by $|\rho|_w$, is given by

$$|\rho|_w = \bigotimes_{n=1}^{m} (A(k+n-1))_{j_n i_n},$$

with $i = i_1$ and $j = j_m$, for some k.

We now are able to introduce the concept of weak irreducibility: A sequence $\{A(k)\}$ of square matrices is said to be *weakly irreducible* if for any pair of nodes $i, j \in \{1, \ldots, J\}$ a finite number m_{ij} exists such that there is with positive probability a path of length m_{ij} from i to j; in formula: for any i, j, with $1 \leq i, j \leq J$, $m_{ij} \in \mathbb{N}$ exists such that

$$P\left(\left(\bigotimes_{k=0}^{m_{ij}-1} A(k) \right)_{ji} > \varepsilon \right) > 0.$$

Theorem 2.2.6 *Let $\{A(k)\}$ be an i.i.d. sequence of regular, integrable matrices in $\mathbb{R}_{\max}^{J \times J}$ with countable state-space. Assume that $\{A(k)\}$ is weakly irreducible. If there exists at least one node j such that j lies with positive probability on a circuit of length one, then the Lyapunov exponent of $\{A(k)\}$ exists.*

Proof: Consider the collection of numbers m_{ij} for $1 \leq i, j \leq J$. We have assumed that there exits at least one node j^* such that $m_{j^* j^*} = 1$ and the greatest common divisor of the collection of numbers m_{ij}, with $1 \leq i, j \leq J$, is thus equal to one. This implies that a finite number N exists such that each $m \geq N$ can be written as a linear combination of m_{ij}'s, see [26]. Weak analyticity thus implies that for any $m \geq N$ there exists with positive probability a path from any node to any other node; in formula: for any $m \geq N$

$$\forall i, j \in \{1, \ldots, J\} : \quad P\left(\left(\bigotimes_{k=0}^{m-1} A(k) \right)_{ji} > \varepsilon \right) > 0.$$

Let $h \geq N$. Since J is finite, we can choose $j, i \in \{1, \ldots, J\}$ such that a sequence $\{m_n\}$, with $\lim_{n\to\infty} m_n = \infty$, exists for which it holds that

$$||x(m_n + h)||_{\min} = x_i(m_n + h) \quad \text{and} \quad ||x(m_n)||_{\max} = x_j(m_n) ,$$

for $n \in \mathbb{N}$. By assumption, $\{A(k)\}$ is a weakly irreducible i.i.d. sequence. Hence, we may select a subsequence $\{m_{n_l}\}$ of $\{m_n\}$ such that there is (at least) a fixed path ρ from j to i of length h in $\mathcal{G}(A(m_{n_l+k}) : 0 \leq k < h)$ for any l and

$$w_{ij} \stackrel{\text{def}}{=} \left(\bigotimes_{m=m_{n_l}}^{m_{n_l}+h-1} A(m) \right)_{ij}$$

is finite. With slight abuse of notation we will identify $\{m_n\}$ and $\{m_{n_l}\}$. This yields

$$\begin{aligned} ||x(m_n + h)||_{\min} &= x_i(m_n + h) \\ &= \bigoplus_{k=1}^{J} \left(\bigotimes_{m=m_n}^{m_n+h-1} A(m) \right)_{ik} \otimes x_k(m_n) \\ &\geq -|w_{ij}| \otimes x_j(m_n) \\ &= -|w_{ij}| + ||x(m_n)||_{\max} , \end{aligned}$$

which establishes the up-crossing property with $M = h$. □

Theorem 2.2.6 provides a sufficient condition for the existence of the Lyapunov exponent completely avoiding the concept of fixed support. The following example illustrates this. Consider $A_1, A_2 \in \mathcal{A}$, with

$$A_1 = \begin{pmatrix} Y_1 & \varepsilon \\ \varepsilon & Y_4 \end{pmatrix} \quad \text{and} \quad A_2 = \begin{pmatrix} \varepsilon & Y_2 \\ Y_3 & \varepsilon \end{pmatrix} ,$$

for some finite integrable random variables Y_i, $1 \leq i \leq 4$. Let $\{A(k)\}$ be an i.i.d. sequence such that $P(A(k) = A_1) = p > 0$ and $P(A(k) = A_2) = 1 - p > 0$, for $k \geq 0$. Then $\{A(k)\}$ satisfies the condition put forward in Theorem 2.2.6. However, neither does $\{A(k)\}$ have fixed support nor does it satisfy the diagonal condition. Note that the situation in Example 1.5.5 is covered by Theorem 2.2.6, which follows from the fact that D_2 is irreducible and contains one finite element on its diagonal.

As Hong shows in [69, 70], the condition that there is at least one node that lies with positive probability on a circuit of length one is not necessary for Theorem 2.2.6 to hold. Without this simplifying assumption the proof of the theorem becomes however rather technical and the interested reader is referred to [69, 70] for details.

2.3 Stability Analysis of Waiting Times (Type IIa)

A classical result in queuing theory states that if in a G/G/1 queue the expected interarrival time is larger than the expected service time, then the sequence of waiting times converges, independent of the initial condition, to a unique stationary regime. The proof of this result goes back to [81]. In this section, we generalize the classical result on stability of waiting times in the GI/G/1 queue to that of stability of waiting times in open max-plus linear networks. It is worth noting that by virtue of the max-plus formalism we can almost literally copy the proof of the classical result in [81].

We consider the following situation. An open queuing network with J stations is given such that the vector of departure times from the stations, denoted by $x(k)$, follows the recurrence relation

$$x(k+1) = A(k) \otimes x(k) \oplus \tau(k+1) \otimes B(k) , \tag{2.30}$$

with $x(0) = \mathbf{e}$, where $\tau(k)$ denotes the time of the k^{th} arrival to the system. See, equation (1.15) in Section 1.4.2.2 and equation (1.27) in Example 1.5.2, respectively. As usually, we denote by $\sigma_0(k)$ the k^{th} interarrival time, so that the k^{th} arrival of a customer at the network happens at time

$$\tau(k) = \sum_{i=1}^{k} \sigma_0(i) , \quad k \geq 1 ,$$

with $\tau(0) = 0$. Then, $W_j(k) = x_j(k) - \tau(k)$ denotes the time the k^{th} customer arriving to the system spends in the system until completion of service at server j. The vector of k^{th} sojourn times, denoted by $W(k) = (W_1(k), \ldots, W_J(k))$, follows the recurrence relation

$$W(k+1) = A(k) \otimes C(\sigma_0(k+1)) \otimes W(k) \oplus B(k) , \quad k \geq 0 ,$$

with $W(0) = \mathbf{e}$, where $C(h)$ denotes a diagonal matrix with $-h$ on the diagonal and ε elsewhere. See Section 1.4.4 for details. Alternatively, $x_j(k)$ in (2.30) may model the times of the k^{th} beginning of service at station j. With this interpretation of $x(k)$, $W_j(k)$ defined above represents the time spent by the k^{th} customer arriving to the system until beginning of her/his service at j. For example, in the G/G/1 queue $W(k)$ models the waiting time.

In the following we will establish sufficient conditions for $W(k)$ to converge to a unique stationary regime. The main technical assumptions are:

(W1) For $k \in \mathbb{Z}$, let $A(k) \in \mathbb{R}_{\max}^{J \times J}$ be a.s. regular and assume that the maximal Lyapunov exponent of $\{A(k)\}$ exists.

(W2) There exists a fixed number α, with $1 \leq \alpha \leq J$, such that the vector $B^{\alpha}(k) = (B_j(k) : 1 \leq j \leq \alpha)$ has finite elements for any k, and $B_j(k) = \varepsilon$, for $\alpha < j \leq J$ and any k.

(W3) The sequence $\{(A(k), B^{\alpha}(k))\}$ is stationary and ergodic, and independent of $\{\tau(k)\}$, where $\tau(k)$ is given by

$$\tau(k) = \sum_{i=1}^{k} \sigma(i) , \quad k \geq 1 ,$$

with $\tau(0) = 0$ and $\{\sigma(k) : k \in \mathbb{Z}\}$ a stationary and ergodic sequence of positive random variables with mean $\nu \in (0, \infty)$.

In what follows, we establish sufficient conditions for $\{W(k)\}$, with

$$W(k+1) = A(k) \otimes C(\sigma(k+1)) \otimes W(k) \oplus B(k) , \quad k \geq 0 , \tag{2.31}$$

to have a unique stationary solution.

Provided that $\{A(k)\}$ is a.s. regular and stationary, integrability of $A(k)$ is a sufficient condition for (**W1**), see Theorem 2.2.1. In terms of queuing networks, the main restriction imposed by these conditions stems from the non-negativity of the diagonal of $A(k)$, see Section 2.2 for a detailed discussion and possible relaxations. The part of condition (**W3**) that concerns the arrival stream of the network is, for example, satisfied for Poisson arrival streams.

The proof goes back to [19] and has three main steps. First, we introduce Loynes' scheme for sojourn times. In a second step we show that the Loynes variable converges a.s. to a finite limit. Finally, we show that this limit is the unique stationary solution of equations of type (2.31).

Step 1 (the Loynes's scheme): Let $M(k)$ denote the vector of sojourn times at time zero provided that the sequence of waiting time vectors was started at time $-k$ in $B(-(k+1))$. For $k > 0$, we set

$$\tau(-k) = -\sum_{i=0}^{k-1} \sigma(-i) .$$

By recurrence relation (2.31),

$$M(1) = A(-1) \otimes C(\sigma(0)) \otimes B(-2) \oplus B(-1) .$$

For $M(2)$ we have to replace $B(-2)$ by

$$A(-2) \otimes C(\sigma(-1)) \otimes B(-3) \oplus B(-2) , \tag{2.32}$$

which yields

$$\begin{aligned} M(2) = {} & A(-1) \otimes C(\sigma(0)) \otimes A(-2) \otimes C(\sigma(-1)) \otimes B(-3) \\ & \oplus A(-1) \otimes C(\sigma(0)) \otimes B(-2) \oplus B(-1) . \end{aligned} \tag{2.33}$$

By finite induction, we obtain for $M(k)$

$$M(k) = \bigoplus_{j=0}^{k} \bigotimes_{i=1}^{j} A(-i) \otimes C(\sigma(-i+1)) \otimes B(-(j+1)) , \tag{2.34}$$

where we set the product

$$\bigotimes_{i=1}^{j} A(-i) \otimes C(\sigma(-i+1))$$

to E for $j = 0$.

The sequence $\{M(k)\}$ is called *Loynes sequence.* The above construction implies that $\{M(k)\}$ is monotone increasing in k. To see this, denote for $x, y \in \mathbb{R}^J_{\max}$ the component-wise ordering of x and y by $x \leq y$. By calculation,

$$\begin{aligned} M(k) &= \bigoplus_{j=0}^{k} \bigotimes_{i=1}^{j} A(-i) \otimes C(\sigma(-i+1)) \otimes B(-(j+1)) \\ &\leq \bigotimes_{i=1}^{k+1} A(-i) \otimes C(\sigma(-i+1)) \otimes B(-(k+1)) \\ &\quad \oplus \bigoplus_{j=0}^{k} \bigotimes_{i=1}^{j} A(-i) \otimes C(\sigma(-i+1)) \otimes B(-(j+1)) \\ &= \bigoplus_{j=0}^{k+1} \bigotimes_{i=1}^{j} A(-i) \otimes C(\sigma(-i+1)) \otimes B(-(j+1)) \\ &= M(k+1) \,, \end{aligned}$$

for $k \geq 0$, which proves that $M(k)$ is monotone increasing in k.

The matrix $C(\cdot)$ has the following properties. For any $y \in \mathbb{R}$, $C(y)$ commutes with any matrix $A \in \mathbb{R}^{J \times J}_{\max}$:

$$C(y) \otimes A = A \otimes C(y) \,.$$

Furthermore, for $y, z \in \mathbb{R}$, it holds that

$$C(y) \otimes C(z) = C(z) \otimes C(y) = C(y+z) \,.$$

Specifically,

$$\bigotimes_{i=1}^{j} C(\sigma(-i+1)) = C\left(\bigotimes_{i=1}^{j} \sigma(-i+1) \right) = C(-\tau(-j)) \,.$$

Elaborating on these rules of computation, we obtain

$$\bigotimes_{i=1}^{j} A(-i) \otimes C(\sigma(-i)) \otimes B(-(j+1)) = C(-\tau(-j)) \otimes \bigotimes_{i=1}^{j} A(-i) \otimes B(-(j+1)) \,.$$

Set

$$D(k) = \bigotimes_{i=1}^{k} A(-i) \otimes B(-(k+1)) \,, \quad k \geq 1 \,,$$

and, for $k = 0$, set $D(0) = B(-1)$. Note that $\tau(0) = 0$ implies that $C(-\tau(0)) = E$. Equation (2.34) now reads

$$M(k) = \bigoplus_{j=0}^{k} C(-\tau(-j)) \otimes D(j) \,.$$

Step 2 (pathwise limit): We now show that the limit of $M(k)$ as k tends to ∞ exists and establish a sufficient condition for the limit to be a.s. finite.

Because $M(k)$ is monotone increasing in k, the random variable M, defined by

$$\lim_{k\to\infty} M(k) = \bigoplus_{j\geq 0} C(-\tau(-j)) \otimes D(j)$$
$$\stackrel{\text{def}}{=} M ,$$

is either equal to ∞ or finite. The variable M is called *Loynes variable.* In what follows we will derive a sufficient condition for M to be a.s. finite. As a first step towards this result, we study three individual limits.

(i) Under condition (**W1**), a number $\mathbf{a} \in \mathbb{R}$ exists such that, for any $x \in \mathbb{R}^J$,

$$\lim_{k\to\infty} \frac{1}{k} \left|\left| \bigotimes_{j=1}^{k} A(-j) \otimes x \right|\right|_{\max} = \mathbf{a} \qquad \text{a.s.}$$

(ii) Under condition (**W3**), the strong law of large numbers (which is a special case of Theorem 2.2.3) implies

$$\begin{aligned}\lim_{k\to\infty} \frac{1}{k} ||C(-\tau(-k))||_{\max} &= \lim_{k\to\infty} \frac{1}{k}\tau(-k) \\ &= -\lim_{k\to\infty} \frac{1}{k} \sum_{i=-k+1}^{0} \sigma(i) \\ &= -\nu \qquad \text{a.s.}\end{aligned}$$

(iii) Ergodicity of $\{B^\alpha(k)\}$ (condition (**W3**)) implies that, for $1 \leq j \leq \alpha$, a $b_j \in \mathbb{R}$ exists such that

$$\lim_{k\to\infty} \frac{1}{k} \sum_{i=1}^{k} B_j(-i) = b_j \qquad \text{a.s.} ,$$

which implies that it holds with probability one that

$$\begin{aligned}b_j &= \lim_{k\to\infty} \frac{1}{k} \sum_{i=1}^{k} B_j(-i) \\ &= \lim_{k\to\infty} \frac{1}{k} B_j(-k) + \lim_{k\to\infty} \frac{k-1}{k} \frac{1}{k-1} \sum_{i=1}^{k-1} B_j(-i) \\ &= \lim_{k\to\infty} \frac{1}{k} B_j(-k) + b_j ,\end{aligned}$$

and thus

$$\lim_{k\to\infty} \frac{1}{k} B_j(-k) = \lim_{k\to\infty} \frac{1}{k} B_j(-(k+1)) = 0 \qquad \text{a.s.} ,$$

for $j \leq \alpha$. From the above we conclude that

$$\lim_{k\to\infty} \frac{1}{k} ||B(-k)||_{\max} = 0 \qquad \text{a.s.}$$

From Lemma 1.6.2 it follows that

$$\begin{aligned} ||C(-\tau(-k)) \otimes D(k)||_{\max} &= \left|\left| C(-\tau(-k)) \otimes \bigotimes_{i=1}^{k} A(-i) \otimes B(-(k+1)) \right|\right|_{\max} \\ &\leq ||C(-\tau(-k))||_{\max} + \left|\left| \bigotimes_{i=1}^{k} A(-i) \otimes \mathbf{e} \right|\right|_{\max} \\ &\quad + ||B(-(k+1))||_{\max} . \end{aligned}$$

Combining the individual limits (i)-(iii), we obtain

$$\lim_{k\to\infty} \frac{1}{k} ||C(-\tau(-k)) \otimes D(k)||_{\max} \leq \mathbf{a} - \nu \qquad \text{a.s.}$$

and $\nu > \mathbf{a}$ implies

$$\lim_{k\to\infty} ||C(-\tau(-k)) \otimes D(k)||_{\max} = -\infty \qquad \text{a.s.} \tag{2.35}$$

Hence, for k sufficiently large, the vector $C(-\tau(-k)) \otimes D(k)$ has only negative elements and thus doesn't contribute to $M(k)$ (note that $M(k) \geq 0$ by definition). Consequently, $M(k)$ is dominated by the maximum over finitely many vectors $C(-\tau(-k)) \otimes D(k)$ whose elements are all finite. We have thus shown that $\nu > \mathbf{a}$ implies that M is an a.s. finite random variable. In the same vein, one can show that $\nu < \mathbf{a}$ implies $M = \infty$ with probability one.

Step 3 (stationarity and uniqueness): We revisit the construction of $\{M(k)\}$. Under assumption (**W3**), let θ denote an ergodic shift operator such that $A(k) = A \circ \theta^k$, $B(k) = B \circ \theta^k$ and $\sigma(k) = \sigma \circ \theta^k$, for appropriately defined random variables A, B, σ, see Section E.3 in the Appendix. Equation (2.33) thus reads

$$M(2) = A \circ \theta^{-1} \otimes C(\sigma) \otimes M(1) \circ \theta^{-1} \oplus B \circ \theta^{-1}$$

(to see this, note that the expression in (2.32) is equivalent to $M(1) \circ \theta^{-1}$). By finite induction,

$$M(k+1) = A \circ \theta^{-1} \otimes C(\sigma) \otimes M(k) \circ \theta^{-1} \oplus B \circ \theta^{-1}$$

and letting k tend to ∞ in the above equation shows that

$$M = A \circ \theta^{-1} \otimes C(\sigma) \otimes M \oplus B \circ \theta^{-1} .$$

In other words, M is the stationary solution of (2.31).

It remains to be shown that M is the unique limit. Let $M(k,w)$ denote the vector of sojourn times at time 0 provided that the sequence is started at time $-k$ with initial vector $w \in \mathbb{R}^J$, or, more formally, set

$$M(k,w) = \bigotimes_{i=1}^{k} A(-i) \otimes C(\sigma(-i+1)) \otimes w \oplus \bigoplus_{j=0}^{k-1} C(-\tau(-j)) \otimes D(j) .$$

Because w has only finite elements, we have $||w||_{\max} < \infty$. Following the line of argument in step 2 above, it readily follows that

$$\lim_{k\to\infty} \left\| \bigotimes_{i=1}^{k} A(-i) \otimes C(\sigma(-i+1)) \otimes w \right\|_{\max} = -\infty \quad \text{a.s.} ,$$

for $\nu > \mathbf{a}$, and

$$\lim_{k\to\infty} \bigotimes_{i=1}^{k} A(-i) \otimes C(\sigma(-i+1)) \otimes w \oplus \bigoplus_{j=0}^{k-1} C(-\tau(-j)) \otimes D(j) = M \quad \text{a.s.}$$

Hence, for any finite initial value w, $M(k,w)$ has the same limit as $M(k)$, which establishes uniqueness. We have thus shown that $M(k,w)$ converges a.s. to a unique stationary limit M, independent of the initial value w.

For $w \in \mathbb{R}^J$, write $W(k,w)$ for the vector of k^{th} system times, initiated at 0 to w. Assumption (**W3**) implies that $M(k,w)$ and $W(k,w)$ are equal in distribution. Hence, M is the unique weak limit of $\{W(k,w)\}$ for arbitrary $w \in \mathbb{R}^J$. We summarize our analysis in the following theorem.

Theorem 2.3.1 *Assume that assumptions* (**W1**), (**W2**) *and* (**W3**) *are satisfied and denote the maximal Lyapunov exponent of* $\{A(k)\}$ *by* $\mathbf{a}$. *If* $\nu > \mathbf{a}$, *then the sequence*

$$W(k+1) = A(k) \otimes C(\sigma(k+1)) \otimes W(k) \oplus B(k)$$

converges with strong coupling to an unique stationary regime W, *with*

$$W = D(0) \oplus \bigoplus_{j\geq 1} C(-\tau(-j)) \otimes D(j) ,$$

where $D(0) = B \circ \theta^{-1}$ *and*

$$D(j) = \bigotimes_{i=1}^{j} A(-i) \otimes B(-(j+1)) , \quad j \geq 1 .$$

Proof: It remains to be shown that the convergence of $\{W(k)\}$ towards W happens with strong coupling. For $w \in \mathbb{R}^J$, let $W(k,w)$ denote the vector of k^{th}

sojourn times, initiated at 0 to w. From the forward construction, see (1.22) on page 20, we obtain

$$W(k+1,w) = \bigotimes_{i=0}^{k} A(i) \otimes C(\sigma(i+1)) \otimes w \oplus \bigoplus_{i=0}^{k} \bigotimes_{j=i+1}^{k} A(j) \otimes C(\sigma(j+1)) \otimes B(i)\,.$$

From the arguments provided in step 2 of the above analysis it follows that

$$\lim_{k\to\infty} \left|\left| \bigotimes_{i=0}^{k} A(i) \otimes C(\sigma(i+1)) \otimes w \right|\right|_{\max} = -\infty \quad \text{a.s.}\,,$$

provided that $\nu > \mathbf{a}$. Hence, there exists an a.s. finite random variable $\beta(w)$, such that

$$\forall k \geq \beta(w) : \quad \left|\left| \bigotimes_{i=0}^{k} A(i) \otimes C(\sigma(i+1)) \otimes w \right|\right|_{\max} < 0\,.$$

In words, after $\beta(w)$ transitions the influence of the initial vector w dies out. We now compare two versions of $\{W(k)\}$. One version is initialized to W, the stationary regime, and the other version is initialized to an arbitrary finite vector w. We obtain that

$$\forall k \geq \max(\beta(w), \beta(W)) : \quad W(k,w) = W(k,W)\,.$$

Hence, $\{W(k,w)\}$ couples after a.s. finitely many transitions with the stationary version $\{W(k,W)\}$. □

It is worth noting that $\beta(w)$, defined in the proof of Theorem 2.3.1, fails to be a stopping time adapted to the natural filtration of $\{(A(k),B(k)) : k \geq 0\}$. More precisely, $\beta(w)$ is measurable with respect to the σ-field $\sigma((A(k),B(k)) : k \geq 0)$ but, in general, $\{\beta(w) = m\} \notin \sigma((A(k),B(k)) : m \geq k \geq 0)$, for $m \in \mathbb{N}$.

Due to the max-plus formalism, the proof of Theorem 2.3.1 is a rather straightforward extension of the proof of the classical result for the G/G/1 queue. To fully appreciate the conceptual advantage offered by the max-plus formalism, we refer to [6, 13] where the above theorem is shown without using max-plus formalism.

2.4 Harris Recurrent Max-Plus Linear Systems (Type I and Type IIa)

The Markov chain approach to stability analysis of max-plus linear systems presented in this section goes back to [93, 41]. Consider the recurrence relation $x(k+1) = A(k) \otimes x(k)$, $k \geq 0$, and let

$$Z_{j-1}(k) = x_j(k) - x_1(k)\,, \quad j \geq 2\,, \tag{2.36}$$

denote the discrepancy between component $x_1(k)$ and $x_j(k)$ in $x(k)$. The sequence $\{Z(k)\}$ constitutes a Markov chain, as the following theorem shows.

Theorem 2.4.1 *The process $\{Z(k) : k \geq 0\}$ is a Markov chain. Suppose $Z(k) = (z_1, \ldots, z_{J-1})$ for fixed $z_j \in \mathbb{R}$. Then the conditional distribution of $Z_j(k+1)$ given $Z(k) = (z_1, \ldots, z_{J-1})$, is equal to the distribution of the random variable*

$$A_{j+1\,1}(k) \oplus \bigoplus_{i=2}^{J} A_{j+1\,i}(k) \otimes z_{i-1} \; - \; A_{11}(k) \oplus \bigoplus_{i=2}^{J} A_{1i}(k) \otimes z_{i-1} \,,$$

for $1 \leq j \leq J-1$.

Proof: Note that

$$\begin{aligned} a \otimes x \oplus b \otimes y - x &= \max(a + x, b + y) - x \\ &= \max(a, b + (y - x)) \\ &= a \oplus b \otimes (y - x) \,. \end{aligned}$$

Using the above equality, we obtain for $2 \leq j \leq J$:

$$\begin{aligned} Z_{j-1}(k+1) &= x_j(k+1) - x_1(k+1) \\ &= (A(k) \otimes x(k))_j - (A(k) \otimes x(k))_1 \\ &= A_{j1}(k) \otimes x_1(k) \oplus A_{j2}(k) \otimes x_2(k) \oplus \cdots \oplus A_{jJ}(k) \otimes x_J(k) \; - \\ &\quad A_{11}(k) \otimes x_1(k) \oplus A_{12}(k) \otimes x_2(k) \oplus \cdots \oplus A_{1J}(k) \otimes x_J(k) \\ &= A_{j1}(k) \otimes x_1(k) \oplus A_{j2}(k) \otimes x_2(k) \oplus \cdots \oplus A_{jJ}(k) \otimes x_J(k) - x_1(k) \; - \\ &\quad (A_{11}(k) \otimes x_1(k) \oplus A_{12}(k) \otimes x_2(k) \oplus \cdots \oplus A_{1J}(k) \otimes x_J(k) - x_1(k)) \\ &= A_{j1}(k) \oplus A_{j2}(k) \otimes (x_2(k) - x_1(k)) \oplus \cdots \oplus A_{jJ}(k) \otimes (x_J(k) - x_1(k)) \; - \\ &\quad A_{11}(k) \oplus A_{12}(k) \otimes (x_2(k) - x_1(k)) \oplus \cdots \oplus A_{1J}(k) \otimes (x_J(k) - x_1(k)) \\ &= A_{j1}(k) \oplus A_{j2}(k) \otimes Z_1(k) \oplus \cdots \oplus A_{jJ}(k) \otimes Z_{J-1}(k) \; - \\ &\quad A_{11}(k) \oplus A_{12}(k) \otimes Z_1(k) \oplus \cdots \oplus A_{1J}(k) \otimes Z_{J-1}(k) \,. \end{aligned}$$

From this expression it follows that the conditional distribution of $Z(k+1)$ given $Z(0), \ldots, Z(k)$ equals the conditional distribution of $Z(k+1)$ given $Z(k)$ and hence the process $\{Z(k) : k \geq 0\}$ is a Markov chain. □

Now define

$$D(k) = x_1(k) - x_1(k-1) \,, \quad k \geq 1 \,.$$

Then, we have

$$x_1(k) = x_1(0) + \sum_{n=1}^{k} D(n) \,, \quad k \geq 1 \,, \tag{2.37}$$

and

$$x_j(k) = x_j(0) + (Z_{j-1}(k) - Z_{j-1}(0)) + \sum_{n=1}^{k} D(n) \,, \quad k \geq 1 \,, j \geq 2 \,. \tag{2.38}$$

Theorem 2.4.2 *For any $k \geq 0$, the distribution of $(D(k+1), Z(k+1))$ given $(Z(0), D(1), Z(1), \dots, D(k), Z(k))$ depends only on $Z(k)$. If $Z(k) = (z_2, \dots, z_J)$, then the conditional distribution of $D(k+1)$ given $(Z(0), D(1), Z(1), \dots, D(k), Z(k))$ is equal to the distribution of the random variable*

$$A_{11}(k) \oplus \bigoplus_{j=2}^{J} A_{1j}(k) \otimes z_{j-1} .$$

Proof: We have

$$\begin{aligned} D(k+1) &= x_1(k+1) - x_1(k) \\ &= A_{11}(k) \otimes x_1(k) \oplus A_{12}(k) \otimes x_2(k) \oplus \cdots \oplus A_{1J}(k) \otimes x_J(k) \; - x_1(k) \\ &= A_{11}(k) \oplus A_{12}(k) \otimes (x_2(k) - x_1(k)) \oplus \cdots \oplus A_{1J}(k) \otimes (x_J(k) - x_1(k)) \\ &= A_{11}(k) \oplus A_{12}(k) \otimes Z_1(k) \oplus \cdots \oplus A_{1J}(k) \otimes Z_{J-1}(k) , \end{aligned}$$

which, together with the previous theorem, yields the desired result. □

If $\{Z(k)\}$ is uniformly ϕ-recurrent and aperiodic (for a definition we refer to the Appendix), then it is ergodic and, as will be shown in the following theorem, a type IIa limit holds. Elaborating on a result from Markov theory for so-called *chain dependent processes*, ergodicity of $\{Z(k)\}$ yields existence of the type I limit and thus of the Lyapunov exponent.

Theorem 2.4.3 *Suppose that the Markov chain $\{Z(k) : k \geq 1\}$ is aperiodic and uniformly ϕ-recurrent, and denote its unique invariant probability measure by π. Then the following holds:*

(i) For $1 \leq i, j \leq J$, $x_i(k) - x_j(k)$ converges weakly to the unique stationary regime π.

(ii) If the elements of $A(k)$ have finite first moments, then a finite number λ exists such that

$$\lim_{k \to \infty} \frac{x_j(k)}{k} = \lambda , \quad j = 1, \dots, J ,$$

almost surely for any finite initial value, and

$$\lambda = \mathbb{E}_\pi[D(1)] ,$$

where $\mathbb{E}_\pi$ indicates that the expected value is taken with $Z(0)$ distributed according to π.

Proof: For the proof of the first part of the theorem note that

$$x_i(k) - x_j(k) = Z_{i-1}(k) - Z_{j-1}(k) ,$$

for $2 \leq i, j \leq J$, and

$$x_i(k) - x_1(k) = Z_{i-1}(k) , \quad x_1(k) - x_i(k) = -Z_{i-1}(k) ,$$

for $2 \leq i \leq J$. Hence, weak convergence of $Z(k)$ to a unique stationary regime implies weak convergence of $x_i(k) - x_j(k)$ to a unique stationary regime. Weak convergence of $Z(k)$ to a unique stationary regime follows from uniform ϕ-recurrence and aperiodicity of $Z(k)$, see Section F in the Appendix, and we have thus shown the first part of the theorem.

We now turn to the second part. The process $\{D(k) : k \geq 1\}$ is a so-called chain dependent process and the limit theorem of Griggorescu and Oprisan [55] implies

$$\lim_{k\to\infty} \frac{1}{k} \sum_{n=1}^{k} D(n) = \lambda = \mathbb{E}_\pi[D(1)] \quad \text{a.s.} ,$$

for all initial values x_0. This yields for the limit of $x_1(k)/k$ as k tends to ∞:

$$\begin{aligned} \lim_{k\to\infty} \frac{1}{k} x_1(k) &\overset{(2.37)}{=} \lim_{k\to\infty} \left(\frac{1}{k} x_1(0) + \frac{1}{k} \sum_{n=1}^{k} D(n) \right) \\ &= \lim_{k\to\infty} \frac{1}{k} \sum_{n=1}^{k} D(n) \\ &= \lambda \quad \text{a.s.} \end{aligned}$$

It remains to be shown that, for $j \geq 2$, the limit of $x_j(k)/k$ as k tends to ∞ equals λ. Suppose that for $j \geq 2$:

$$\lim_{k\to\infty} \frac{1}{k} Z_{j-1}(k) = \lim_{k\to\infty} \frac{1}{k} \Big(Z_{j-1}(k) - Z_{j-1}(0) \Big) = 0 \quad \text{a.s.} \tag{2.39}$$

With (2.39) it follows from (2.38) that

$$\begin{aligned} \lim_{k\to\infty} \frac{1}{k} x_j(k) &= \lim_{k\to\infty} \frac{1}{k} (Z_{j-1}(k) - Z_{j-1}(0)) + \lambda \\ &= \lambda \quad \text{a.s.} , \end{aligned}$$

for $j \geq 2$. In what follows we show that (2.39) indeed holds under the conditions of the theorem.

Uniform ϕ-recurrence and aperiodicity of the Markov chain $\{Z(k) : k \geq 1\}$ implies Harris ergodicity. Hence, for $J - 1 \geq j \geq 1$, finite constants c_j exists, such that

$$\lim_{k\to\infty} \frac{1}{k} \sum_{n=1}^{k} Z_j(n) = c_j \quad \text{a.s.}$$

This implies

$$\begin{aligned}
c_j &= \lim_{k\to\infty} \frac{1}{k}\sum_{n=1}^{k} Z_j(n) \\
&= \lim_{k\to\infty}\left(\frac{1}{k}Z_j(k) + \frac{k-1}{k}\frac{1}{k-1}\sum_{n=1}^{k-1} Z_j(n)\right) \\
&= \lim_{k\to\infty}\frac{1}{k}Z_j(k) + \lim_{k\to\infty}\frac{k-1}{k}\lim_{k\to\infty}\frac{1}{k-1}\sum_{n=1}^{k-1} Z_j(n) \\
&= \lim_{k\to\infty}\frac{1}{k}Z_j(k) + c_j ,
\end{aligned}$$

which yields, for $J-1 \geq j \geq 1$,

$$\lim_{k\to\infty}\frac{1}{k}Z_j(k) = \lim_{k\to\infty}\frac{1}{k}\Big(Z_j(k) - Z_j(0)\Big) = 0 \quad \text{a.s.}$$

□

Remark 2.4.1 *Let the conditions in Theorem 2.4.3 be satisfied. If, in addition, the elements of $A(k)$ and the initial vector have finite second moments, then*

$$0 \leq \sigma^2 \stackrel{\text{def}}{=} \sum_{n=1}^{\infty} \mathbb{E}_\pi[(D(1)-\lambda)(D(n)-\lambda)] < \infty ,$$

and if $\sigma^2 > 0$, the sequence

$$\frac{(x_1(k),\ldots,x_J(k)) - (k\lambda,\ldots,k\lambda)}{\sigma\sqrt{k}}, \quad k \geq 1 ,$$

converges in distribution to the random vector $(\mathcal{N},\ldots,\mathcal{N})$, where $\mathcal{N}$ is a standard normal distributed random variable. For details and proof we refer to [93].

Remark 2.4.2 *If the state space of $Z(k)$ is finite, then the convergence in part (i) of Theorem 2.4.3 happens in strong coupling.*

The computational formula for λ put forward in Theorem 2.4.3 is also known as 'Furstenberg's cocycle representation of the Lyapunov exponent;' see [45].

Example 2.4.1 *Consider $x(k)$ as defined in Example 1.5.5, and let $\sigma = 1$ and $\sigma' = 2$. Matrix D_2 is primitive and has (unique) eigenvector $(1,1,0,1)^\top$. Let $z(1) = z((1,1,0,1)^\top) = (0,-1,0)^\top$. It is easily checked that $\{Z(k)\}$ is a Markov chain on state space $\{z(i) : 1 \leq i \leq 5\}$, with*

$$z(2) = \begin{pmatrix}0\\0\\0\end{pmatrix}, \; z(3) = \begin{pmatrix}-1\\0\\0\end{pmatrix}, \; z(4) = \begin{pmatrix}-1\\-2\\0\end{pmatrix} \;\; \textit{and} \;\; z(5) = \begin{pmatrix}0\\-1\\-1\end{pmatrix}.$$

Denoting the transition probability of $Z(k)$ from state $z(i)$ to state $z(j)$, for $1 \le i,j \le 5$, one obtains the following transition matrix

$$\mathbf{P} = \begin{pmatrix} 1-\theta & \theta & 0 & 0 & 0 \\ 0 & 0 & \theta & 1-\theta & 0 \\ 0 & 0 & \theta & 1-\theta & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1-\theta & \theta & 0 & 0 & 0 \end{pmatrix}.$$

The chain is aperiodic and since the state space is finite it is uniformly ϕ-recurrent. Moreover, the unique stationary distribution of $Z(k)$ is this given by

$$\pi_{z(1)} = \frac{(1-\theta)^2}{1+\theta(1-\theta)},\ \pi_{z(2)} = \frac{\theta(1-\theta)}{1+\theta(1-\theta)},\ \pi_{z(3)} = \frac{\theta^2}{1+\theta(1-\theta)},$$
$$\pi_{z(4)} = \frac{\theta(1-\theta)}{1+\theta(1-\theta)}\ \text{ and }\ \pi_{z(5)} = \frac{\theta(1-\theta)}{1+\theta(1-\theta)}.$$

Applying Theorem 2.4.3, yields $\lambda = \mathbb{E}_\pi[D(1)]$. Evoking Theorem 2.4.2, this expected value can thus be computed as follows:

$$\begin{aligned}\lambda &= \sum_{i=1}^{5} \pi_{z(i)}(1 \oplus 2 \otimes z_2(i)) \\ &= \pi_{z(1)} + 2\pi_{z(2)} + 2\pi_{z(3)} + \pi_{z(4)} + \pi_{z(5)} \\ &= 1 + \frac{\theta}{1+\theta-\theta^2},\end{aligned}$$

for any $\theta \in [0,1]$. For a different example of this kind, see [65].

Example 2.4.2 *Let $\{A(k)\}$, with $A(k) \in \{0,1\}^{2\times 2}$, be an i.i.d. sequence following the distribution $P(A_{ij}(k) = 0) = 1/2 = P(A_{ij}(k) = 1)$ for $1 \le i,j \le J$. We turn to the Markov process $\{Z(k)\}$ as defined in (2.36). This process has state space $\{-1,0,1\}$. By Theorem 2.4.1, the transition probability of $Z(k)$ is given by*

$$\begin{aligned}&P(Z(k+1) = m \,|\, Z(k) = z) \\ &= P((A_{21}(k+1) \oplus (A_{22}(k+1) \otimes z)) - A_{11}(k+1) \oplus (A_{12}(k+1) \otimes z) = m),\end{aligned}$$

for $m, z \in \{-1,0,1\}$, and the transition matrix on $\{Z(k)\}$ can be computed as

$$\mathbf{P} = \begin{pmatrix} \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{3}{16} & \frac{5}{8} & \frac{3}{16} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \end{pmatrix}.$$

The Markov chain $\{Z(k)\}$ is aperiodic (all elements of $\mathbf{P}$ are positive), uniformly ϕ-recurrent (the state space is finite) and has unique stationary distribution

$$\pi_{-1} = \frac{3}{14},\ \pi_0 = \frac{8}{14},\ \pi_1 = \frac{3}{14}.$$

From Theorem 2.4.3 together with Theorem 2.4.2 follows that

$$\begin{aligned}\lambda &= \sum_{z\in\{-1,0,1\}} \mathbb{E}[\,A_{11}(1) \oplus (A_{12}(1) \otimes z)\,]\pi_z \\ &= \frac{3}{14} \times \frac{1}{2} + \frac{8}{14} \times \frac{3}{4} + \frac{3}{14} \times \frac{3}{2} \\ &= \frac{6}{7}\,.\end{aligned}$$

As shown in [93], for this example σ (as defined in Remark 2.4.1) is equal to 33/343.

The above examples are deceitfully simple in the sense that (i) the transition probability (in this case a matrix) of $\{Z(k)\}$ can be calculated easily and (ii) we can deduce that $\{Z(k)\}$ is aperiodic and uniformly ϕ-recurrent from inspecting the transition matrix of $\{Z(k)\}$. In [93], examples with countable state space are discussed. For one example, the elements of $A(k)$ are exponentially distributed with the same parameter; for another example, the elements are assumed to be uniformly distributed over the unit interval. Unfortunately, even when the elements of $A(k)$ are governed by these ostensibly simple distributions, the analysis leads to cumbersome computations. It is mainly for this reason that the Markov chain approach, as presented in this section, will be of avail only in special cases.

2.5 Limits in the Projective Space (Type IIb)

In the previous section, we studied the limit of *differences within* $x(k)$, that is, $x_j(k) - x_1(k)$, for $2 \le j \le J$. In what follows, we take a slightly different point of view and consider differences *between* $x(k)$ *and* $x(k-1)$, that is, $x_j(k) - x_j(k-1)$, for $1 \le j \le J$. The basic recurrence relation we study is given by

$$x(k+1) \;=\; A(k) \otimes x(k)\,, \quad k \ge 0\,, \tag{2.40}$$

with $x(0) = x_0 \in \mathbb{R}^J_{\max}$ and $A(k) \in \mathbb{R}^{J\times J}_{\max}$, for $k \ge 0$.

For the following we use a definition of $Z(k)$ that slightly differs from the definition in Section 2.4. We now let

$$Z(k) = x(k) - x(k-1)\,, \quad k \ge 1\,, \tag{2.41}$$

denote the component-wise increase of $x(k)$. In particular, the components of $Z(k)$ are given by

$$\begin{aligned}Z_j(k) &= F_j(A(k-1), x(k-1)) \\ &\stackrel{\text{def}}{=} \left(\bigoplus_{i=1}^{J} A_{ji}(k-1) \otimes x_i(k-1) \right) - x_j(k-1) \\ &= A_{jj}(k-1) \oplus \bigoplus_{\substack{i=1 \\ i\neq j}}^{J} A_{ji}(k-1) \otimes (x_i(k-1) - x_j(k-1))\,, \quad j \ge 1\,.\end{aligned}$$

For $\overline{x(k-1)} \in \mathbb{PR}^J_{\max}$, the value of F_j doesn't depend on the representative, that is, for all $X \in \overline{x(k-1)}$ we have $Z_j(k) = F_j(A(k-1), X)$, for $1 \leq j \leq J$, and we write $Z_j(k) = F_j(A(k-1), \overline{x(k-1)})$ to express this fact. For the definition of the modes of convergence used in the following lemma we refer to Section E.4 in the Appendix.

Lemma 2.5.1 *Consider the situation in (2.40) and let $\{A(k) : k \geq 0\}$ be stationary. If $\overline{x(k)} \in \mathbb{PR}^J$ converges weakly to a unique invariant distribution, uniformly over all initial conditions, then $Z(k)$ converges weakly to a unique invariant distribution, uniformly over all initial conditions.*

Proof: Consider the sequence $y(k) = (A(k), \overline{x(k)})$, $k \geq 0$. The sequence $A(k)$ is stationary by assumption with stationary distribution π_A. Let A be distributed according to π_A. If $\overline{x(k)}$ converges weakly to $\overline{x}$, then $y(k)$ converges weakly to $(A, \overline{x})$. Because $F = (F_1, \ldots, F_J)$ defined above is continuous, we obtain from the continuous mapping theorem (see Appendix, Section E.4) the weak convergence of $F(A(k), \overline{x(k)})$. □

In what follows we establish sufficient conditions for weak convergence of $\overline{x(k)}$. By Lemma 2.5.1, this already implies weak convergence of $Z(k)$ which in turn yields type IIb second-order ergodic theorems. As we will show in the following, in many situations, the convergence of $Z(k)$ occurs even in strong coupling. In Section 2.5.1, we will study systems with countable state space and, in Section 2.5.2, we will address the general situation. In Section 2.5.3 we revisit the deterministic setup. Finally, we present a representation of type IIb limits via a renewal type approach in Section 2.5.4.

2.5.1 Countable Models

In this section, we study models with countable state space. Let $\mathcal{A}$ be a finite or countable collection of $J \times J$-dimensional irreducible matrices. We think of $\mathcal{A}$ as the state space of the random sequence $\{A(k)\}$ following a discrete law.

Definition 2.5.1 *Let $\{A(k)\}$, with $A(k) \in \mathcal{A}$, be a random sequence. A matrix $\tilde{A} \in \mathcal{A}$ is called a pattern of $\{A(k)\}$ if a sequence $\tilde{a} = (a_1, \ldots, a_N) \in \mathcal{A}^N$ exists such that*

$$(a) \qquad \tilde{A} = a_N \otimes a_{N-1} \otimes \cdots \otimes a_1$$

$$(b) \qquad P\big(A(N+k) = a_N, \ldots, A(1+k) = a_1\big) > 0 \,, \quad k \in \mathbb{N} \,.$$

We call $\tilde{a}$ the sequence constituting $\tilde{A}$.

Note that if $\{A(k)\}$ is i.i.d., then the second condition in the above definition is satisfied if we let $\mathcal{A}$ contain only those possible outcomes of $A(k)$ that have a positive probability. In other words, in the i.i.d. case, the second condition is satisfied if we restrict $\mathcal{A}$ to the support of $A(k)$. Existence of a pattern essentially implies that $\mathcal{A}$ is at most countable, see Remark 2.2.8.

The main technical assumptions we need are the following:

(C1) The sequence $\{A(k)\}$ is i.i.d. with countable state space $\mathcal{A}$.

(C2) Each $A \in \mathcal{A}$ is regular.

(C3) There is a primitive matrix C that is a pattern of $\{A(k)\}$.

Observe that we have already encountered the concept of a pattern - as expressed in condition **(C3)** - in condition **(C)** on page 82, although we haven't coined the name 'pattern' for it at that stage.

The following theorem provides a sufficient condition for $\{\overline{x(k)}\}$ to converge in strong coupling.

Theorem 2.5.1 *Let* **(C1)** - **(C3)** *be satisfied, then* $\{\overline{x(k)}\}$ *converges with strong coupling to a unique stationary regime for all initial conditions in* $\mathbb{R}^J$. *In particular,* $\overline{x(k)}$ *converges in total variation.*

Proof: Let C be defined as in **(C3)** and denote the coupling time of C by c. For the sake of simplicity, assume that $C \in \mathcal{A}$, which implies $N = 1$. Set $\tau_0 = 0$ and

$$\tau_{k+1} = \inf\{m \geq \tau_k + c : A(m-i) = C : 0 \leq i \leq c-1\}\,, \quad k \geq 0\,.$$

In words, at time τ_k we have observed for the k^{th} time a sequence of c consecutive occurrences of C. The i.i.d. assumption **(C1)** implies that $\tau_k < \tau_{k+1} < \infty$ for all k and that $\lim_{k\to\infty} \tau_k = \infty$ with probability one. Let p denote the probability of observing C, then we observe C^c with probability p^c. By construction, the probability of the event $\{\tau_1 = m\}$ is less than or equal to the probability of the event $A(k) \neq C$, $0 \leq k \leq m-c$, and $A(k) = C$, for $k = m-c+1, \ldots, m$. In other words, for $m \geq c$, it holds that $P(\tau_1 = m) \geq (1-p)^{m-c}p^c$. Hence,

$$\begin{aligned}
\mathbb{E}[\tau_1] &\leq \sum_{m=c}^{\infty} m\,(1-p)^{m-c}p^c \\
&= \sum_{m=0}^{\infty} (m+c)\,(1-p)^m p^c \\
&= c\,p^c \sum_{m=0}^{\infty} (1-p)^m \; + p^c \sum_{m=0}^{\infty} m\,(1-p)^m \\
&= \frac{c\,p^c}{p} + \frac{p^c(1-p)}{p^2} \\
&< \infty\,,
\end{aligned}$$

which implies that $\mathbb{E}[\tau_{k+1} - \tau_k] < \infty$, for $k \in \mathbb{N}$.

At τ_k, $x(\tau_k) \in V(C)$, see Theorem 2.1.1. By condition **(C3)**, C is primitive and, by Corollary 2.1.1, the eigenspace of C is a single point in the projective space (that is, the eigenvector of C is unique). In other words, $\{A(k)\}$ has MLP, see Lemma 2.2.2. By **(C2)**, $\overline{x(k)} \in \mathbb{R}^J$, for any k, and from the above line argument it follows that $\{\overline{x(k)}\}$ is a Harris ergodic Markov chain and regenerates whenever the chain hits the single point $\overline{V(C)}$. This implies that $\{\overline{x(k)}\}$ converges with strong coupling to a unique stationary regime. See Section F in the Appendix. □

What happens if we consider in Theorem 2.5.1 a stationary and ergodic sequence instead of an i.i.d. sequence? The key argument in the proof of Theorem 2.5.1 is that $\{A(k)\}$ has MLP. This is guaranteed by the fact that we observe with positive probability a sequence of occurrences of $A(k)$ such that the partial product over that sequence equals C^c, for some primitive matrix C, where c denotes the coupling time of C, see Lemma 2.2.2. If the coupling time of C is larger than 1, then, under i.i.d. regime, the event that C occurs c times in a row has positive probability. However, this reasoning doesn't apply in the stationary case. To see this, consider the following example. Let $\Omega = \{\omega_1, \omega_2\}$ and $P(\omega_i) = 1/2$, for $i = 1, 2$. Define the shift operator θ by $\theta(\omega_1) = \omega_2$ and $\theta(\omega_2) = \omega_1$. Then θ is stationary and ergodic. Consider the matrices A, B as defined in Example 2.1.1 and let

$$\{A(k,\omega_1)\} = A, B, A, B, \ldots \qquad \{A(k,\omega_2)\} = B, A, B, A, \ldots$$

The sequence $\{A(k)\}$ is thus stationary and ergodic. Furthermore, A, B are primitive matrices whose coupling time is 4 each. But with probability one we never observe a sequence of 4 occurrences in a row of either A or B. Since neither A or B is of rank 1, we cannot conclude that $\{A(k)\}$ has MLP and, consequently, that $\overline{x(k)}$, with $x(k+1) = \bigotimes_{i=0}^{k} A(i) \otimes x_0$, is regenerative. However, if we replace, for example, A by A^m, for $m \geq 4$ (i.e., a matrix of rank 1), then the argument would apply again. For this reason, we require for the stationary and ergodic setup that a matrix of rank 1 exists that is a pattern, so that $\overline{x(k)}$ becomes a regenerative process. Note that the condition 'there exits a pattern of rank 1' is equivalent to the condition '$\{A(k)\}$ has MLP.' The precise statement is given in the following theorem. For a proof we refer to [84].

Theorem 2.5.2 *Let $\{A(k)\}$ be a stationary and ergodic sequence of a.s. regular square matrices. If $\{A(k)\}$ has MLP, then $\{\overline{x(k)}\}$ converges with strong coupling to a unique stationary regime for all initial conditions in $\mathbb{R}^J$. In particular, $\{\overline{x(k)}\}$ converges in total variation.*

2.5.2 General Models

In this section, we consider matrices $A(k)$ the elements of which may follow a distribution that is either discrete or absolutely continuous with respect to the Lebesgue measure, or a mixture of both. For general state-space, the event $\{A(N+k) \otimes \cdots \otimes A(2+k) \otimes A(1) = \tilde{A}\}$ in Definition 2.5.1 typically has probability zero. For this reason we introduce the following extension of the definition of a pattern. Let $M \in \mathbb{R}_{\max}^{J \times J}$ be a deterministic matrix and $\eta > 0$. We denote by $B(M, \eta)$ the open ball with center M and radius η in the supremum norm on $\mathbb{R}^{J \times J}$. More precisely, $A \in B(M, \eta)$ if for all i, j, with $1 \leq i, j \leq J$, it holds that

$$A_{ij} \begin{cases} \in (M_{ij} - \eta, M_{ij} + \eta) & \text{for } M_{ij} \neq \varepsilon, \\ = \varepsilon & \text{for } M_{ij} = \varepsilon \,. \end{cases}$$

With this notation, we can state the fact that a matrix A belongs to the support of a random matrix $\tilde{A}$ by

$$\forall \eta > 0 \qquad P(\tilde{A} \in B(A, \eta)) > 0 \,.$$

This includes the case where A is a boundary point of the support. We now state the definition of a pattern for non-countable state-space.

Definition 2.5.2 *Let $\{A(k)\}$ be a random sequence of matrices over $\mathbb{R}_{\max}^{J\times J}$ and let $\tilde{A} \in \mathbb{R}_{\max}^{J\times J}$ be a deterministic matrix. We call $\tilde{A}$ a pattern of $\{A(k)\}$ if a deterministic number N exists such that for any $\eta > 0$ it holds that*

$$P\Big(A(N-1) \otimes A(N-2) \otimes \cdots \otimes A(0) \in B(\tilde{A}, \eta) \Big) > 0 \,.$$

Definition 2.5.2 can be phrased as follows: Matrix $\tilde{A}$ is a pattern of $\{A(k)\}$ if $N \in \mathbb{N}$ exists such that $\tilde{A}$ lies in the support of the random matrix $A(N-1) \otimes A(N-2) \otimes \cdots \otimes A(0)$. The key condition for general state space is the following:

(C4) There exists a (measurable) set of matrices $\mathcal{C}$ such that for any $C \in \mathcal{C}$ it holds that C is a pattern of $\{A(k)\}$ and C is of rank 1. Moreover, a finite number N exists such that

$$P\Big(A(N-1) \otimes A(N-2) \otimes \cdots \otimes A(0) \in \mathcal{C} \Big) > 0.$$

Under condition **(C4)**, the following counterpart of Theorem 2.5.2 for models with general state space can be established; for a proof we refer to [84].

Theorem 2.5.3 *Let $\{A(k)\}$ be a stationary and ergodic sequence of a.s. regular matrices in $\mathbb{R}_{\max}^{J\times J}$. If condition* **(C4)** *is satisfied, then $\{\overline{x(k)}\}$ converges with strong coupling to a unique stationary regime. In particular, $\{\overline{x(k)}\}$ converges in total variation to a unique stationary regime.*

In Definition 2.5.2, we required that after a fixed number of transitions the pattern lies in the support of the matrix product. The following, somewhat weaker, definition requires that an arbitrarily small η-neighborhood of the pattern can be reached in a finite number of transitions where the number of transitions is deterministic and may depend on η.

Definition 2.5.3 *Let $\{A(k)\}$ be a random sequence of matrices over $\mathbb{R}_{\max}^{J\times J}$ and let $\tilde{A} \in \mathbb{R}_{\max}^{J\times J}$ be a deterministic matrix. We call $\tilde{A}$ an asymptotic pattern of $\{A(k)\}$ if for any $\eta > 0$ a deterministic number N_η exists, such that*

$$P\Big(\overline{A(N_\eta - 1) \otimes A(N_\eta - 2) \otimes \cdots \otimes A(0)} \in \overline{B(\tilde{A}, \eta)} \Big) > 0 \,.$$

Accordingly, we obtain a variant of condition (**C4**).

(C4)' There exists a matrix C such that C is an asymptotic pattern of $\{A(k)\}$ and C is of rank 1.

Under condition **(C4)'** only weak convergence of $\{\overline{x(k)}\}$ can be established, whereas **(C4)** even yields total variation convergence. The precise statement is given in the following theorem.

Theorem 2.5.4 *Let $\{A(k)\}$ be a stationary and ergodic sequence of a.s. regular matrices in $\mathbb{R}_{\max}^{J\times J}$. If condition* $(\mathbf{C4})'$ *is satisfied, then $\{\overline{x(k)}\}$ converges with δ-coupling to a unique stationary regime. In particular, $\{\overline{x(k)}\}$ converges weakly to a unique stationary regime.*

Proof: We only give a sketch of the proof, for a detailed proof see [84]. Suppose that a stationary version $\overline{x \circ \theta^k}$ of $\overline{x(k)}$ exists, where θ denotes a stationary and ergodic shift. We will show that $\overline{x(k)}$ converges with δ-coupling to $\overline{x \circ \theta^k}$. Fix $\eta > 0$ and let τ denote the time of the first occurrence of the pattern. Condition $(\mathbf{C4})'$ implies that at time τ the projective distance of the two versions is at most η, in formula:

$$d_{\mathbb{P}}(\overline{x}(\tau), \overline{x \circ \theta^{\tau}}) \leq \eta\,. \tag{2.42}$$

As Mairesse shows in [84], the projective distance of two sequences driven by the same sequence $\{A(k)\}$ is non-expansive which means that (2.42) already implies

$$\forall k \geq \tau : \quad d_{\mathbb{P}}(\overline{x(k)}, \overline{x \circ \theta^k}) \leq \eta\,.$$

Hence, for any $\eta > 0$,

$$P(\, d_{\mathbb{P}}(\overline{x(k)}, \overline{x \circ \theta^k}) \leq \eta\,,\; k \geq \tau\,) = 1\,.$$

Stationarity of $\{A(k)\}$ implies $\tau < \infty$ a.s. and the above formula can be written

$$\lim_{k\to\infty} P(\, d_{\mathbb{P}}(\overline{x(k)}, \overline{x \circ \theta^k}) \leq \eta\,) = 1\,.$$

Hence, $\overline{x(k)}$ converges with δ-coupling to a stationary regime. See the Appendix. Uniqueness of the limit follows from the same line of argument. □

We conclude this presentation of convergence results by stating the most general result, namely, that existence of an asymptotic pattern is a necessary and sufficient condition for weak convergence of $\{\overline{x(k)}\}$.

Theorem 2.5.5 *(Theorem 7.4 in [84]) Let $\{A(k)\}$ be a stationary and ergodic sequence on $\mathbb{R}_{\max}^{J\times J}$. A necessary and sufficient condition for $\{\overline{x(k)}\}$ to converge in δ-coupling (respectively, weakly) to a unique stationary regime is that* $(\mathbf{C4})'$ *is satisfied.*

2.5.3 Periodic Regimes of Deterministic Max-Plus DES

Consider the deterministic max-plus linear system

$$x(k+1) = A \otimes x(k), \quad k \geq 0,$$

with $x(0) = x_0 \in \mathbb{R}^J$ and $A \in \mathbb{R}_{\max}^{J \times J}$ a regular matrix. A periodic regime of period d is a set of vectors $x^1, \dots, x^d \in \mathbb{R}^J$ such that (i) $\overline{x^i} \neq \overline{x^j}$, for $1 \leq i \neq j \leq d$, and (ii) a finite number μ exists which satisfies

$$x^{i+1} = A \otimes x^i, \quad 1 \leq i \leq d,$$

and $\mu \otimes x^1 = A \otimes x^d$. A consequence of the above definition is that $x^1, \dots, x^d$ are eigenvectors of A^d and μ is an eigenvalue of A^d. If A is irreducible with cyclicity $\sigma(A)$, then A will possess periodic regimes of period $\sigma(A)$, see Theorem 2.1.1, and $A^{\sigma(A)}$ will have $\sigma(A)$ mutually linear independent eigenvectors.

From a system theoretic point of view, one is interested in the limiting behavior of $x(k)$. More precisely, one is interested in the behaviour of $\overline{x(k)}$ for k large. If A is primitive, $\overline{x(k)}$ converges in a finite number of steps to $\overline{x}$, where x denotes the unique eigenvector of A. In the general situation, however, there are two sources for non-uniqueness of the limiting behavior of $\overline{x(k)}$. First, if A has cyclicity $\sigma(A) > 1$, then $\{\overline{x(k)}\}$ may eventually reach a periodic regime of period $\sigma(A)$. Secondly, even if A has cyclicity one, if the communication graph of A possesses m strongly connected subgraphs, with $m > 1$, then the eigenspace of A is a m-dimensional vector space. See Theorem 2.1.2.

Example 2.5.1 *Consider matrix*

$$A = \begin{pmatrix} 3 & 6 \\ 4 & 4 \end{pmatrix}.$$

A is irreducible with eigenvalue 5 and the critical graph of A consists of the circuit $((1,2),(2,1))$. The critical graph has thus one m.s.c.s. and $\sigma(A) = 2$. It is easily checked that the eigenspace of A is given by

$$V(A) = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}_{\max}^2 \,\middle|\, \exists a \in \mathbb{R} : \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = a \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}.$$

Starting in $x(0) \notin V(A)$, will lead to a periodic regime of period 2. For example, taking $x(0) = (0,0)$, yields

$$x(1) = \begin{pmatrix} 6 \\ 4 \end{pmatrix}, \quad x(2) = \begin{pmatrix} 10 \\ 10 \end{pmatrix}, \quad x(3) = \begin{pmatrix} 16 \\ 14 \end{pmatrix}, \quad x(4) = \begin{pmatrix} 20 \\ 20 \end{pmatrix} \quad \cdots$$

In other words, A^2 has eigenvalue 10 and two linear independent eigenvectors, namely

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 2 \\ 0 \end{pmatrix}.$$

We call the set of all initial conditions x_0 such that $\overline{A^k \otimes x_0}$ eventually reaches $\overline{x}$, for some eigenvector x, (resp. periodic regime $\overline{x^1}, \ldots, \overline{x^d}$) the domain of attraction of x (resp. $x^1, \ldots, x^d$). For example, for the matrix given in Example 2.5.1 above, the vector $x = (0,0)$ lies in the domain of attraction of the periodic regime $(6,4), (10,10)$.

For $J = 3$, Mairesse provides a graphical representation of the domain of attraction in the projective space, see [83] and the extended version [82]. In particular, the eigenvector (resp. periodic regime) in whose domain of attraction an initial value x_0 lies can be deduced from a graphical representation of the eigenspace of A in the projective space.

2.5.4 The Cycle Formula

We revisit the situation in Section 2.2.3.2 and use the notation as introduced therein. Specifically, we assume that $\{A(k)\}$ has MLP. Elaborating on the projective space, (2.26) reads

$$\overline{x(T_k)} = \overline{x}\,, \quad k \geq 0\,,$$

for some fixed $x \in \mathbb{R}^J$. This constitutes a regenerative property of $\{\overline{x(k)}\}$. Specifically, the cycles $\{\overline{x(k)} : T_k < n \leq T_{k+1}\}$ constitute an i.i.d. sequence. Moreover, $\{T_k\}$ is a sequence of renewal times for the process $\{x(k)-x(k-1)\}$ as well. Stationarity and ergodicity of $\{A(k)\}$ imply that $\overline{x(k)}$ hits $\overline{x}$ a.s. infinitely often. Hence, $\{x(k) - x(k-1)\}$ is a regenerative process with renewal times $\{T_k\}$, see Section E.9 in the Appendix. Note that

$$\mathbb{E}\left[\sum_{k=T_0+1}^{T_1} \big(x(k) - x(k-1)\big)\right] = \mathbb{E}\left[x(T_1) - x(T_0)\right].$$

Let $\overline{X}$ denote the unique stationary regime of $\{\overline{x(k)}\}$. Provided that $\mathbb{E}[x(T_1) - x(T_0)] < \infty$ and $\mathbb{E}[T_1 - T_0] < \infty$, the limit theorem for regenerative processes yields

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} \big(x(k) - x(k-1)\big) = \frac{1}{\mathbb{E}[T_1 - T_0]}\, \mathbb{E}\left[\sum_{k=T_0+1}^{T_1} \big(x(k) - x(k-1)\big)\right] \quad \text{a.s.}$$

Moreover, ergodicity of $\{A(k)\}$ yields

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} \big(x(k) - x(k-1)\big) = \mathbb{E}[X \circ \theta - X] \quad \text{a.s.}\,,$$

for $X \in \overline{X}$. In particular, for $X \in \overline{X}$, it holds that

$$\mathbb{E}[\overline{X} \circ \theta - \overline{X}] = \mathbb{E}[X \circ \theta - X]\,.$$

We summarize the above analysis in the following lemma.

Lemma 2.5.2 (Cycle Formula) *Let $\{A(k)\}$ be a stationary and ergodic sequence in $\mathbb{R}_{\max}^{J\times J}$ that has MLP. If $\mathbb{E}[x(T_1) - x(T_0)] < \infty$ and $\mathbb{E}[T_1 - T_0] < \infty$, then*

$$\mathbb{E}[\overline{X} \circ \theta - \overline{X}] = \frac{\mathbb{E}\,[\,x(T_1) \,-\, x(T_0)\,]}{\mathbb{E}[T_1 - T_0]} \,,$$

where $\overline{X}$ denotes the unique stationary regime of $\{\overline{x(k)}\}$.

Remark 2.5.1 *Note that*

$$\begin{aligned}\overline{X \circ \theta} &= \{Y \,|\, \exists \alpha \,:\, Y = \alpha \otimes (X \circ \theta)\,\} \\ &= \{Y \,|\, \exists \alpha \,:\, Y = (\alpha \otimes X) \circ \theta\,\} \\ &= \overline{X} \circ \theta\end{aligned}$$

and the cycle formula can alternatively be phrased

$$\mathbb{E}[\overline{X \circ \theta} - \overline{X}] = \frac{\mathbb{E}\,[\,x(T_1) \,-\, x(T_0)\,]}{\mathbb{E}[T_1 - T_0]} \,.$$

Remark 2.5.2 *If $\{A(k)\}$ is i.i.d., then in the above theorem the condition that $\{A(k)\}$ has MLP can be replaced by condition* (**C**)*, see Lemma 2.2.2. Moreover, a simple geometrical trial argument, like the one used in the proof of Theorem 2.5.1, shows that $\mathbb{E}[T_1 - T_0] < \infty$. If, in addition, $A(k)$ is integrable, one can show that $\mathbb{E}[x(T_1) - x(T_0)] < \infty$ holds as well.*

In the following section we will establish sufficient conditions for $\mathbb{E}[\overline{X} \circ \theta - \overline{X}]$ to be equal to the Lyapunov exponent.

2.6 Lyapunov Exponents via Second Order Limits (Type IIb)

The Lyapunov exponent can be defined as a first-order limit, as explained in Section 2.2. However, as we will show in this section, under suitable conditions, the Lyapunov exponent can be obtained by a second-order limit as well. In Section 2.6.1 we establish the general result, whereas in Section 2.6.2 we provide a direct analysis via backward coupling. It is this result that will prove valuable for the analysis provided in Part II. The basic recurrence relation we study is given by

$$x(k+1) = A(k) \otimes x(k) \,, \quad k \geq 0 \,, \tag{2.43}$$

with $x(0) = x_0 \in \mathbb{R}^J$ and $\{A(k)\}$ a stationary sequence of a.s. regular matrices on $\mathbb{R}_{\max}^{J\times J}$.

2.6.1 The Projective Space

Suppose that $\overline{x(k)}$ converges in total variation and let $\overline{X}$ denote the limiting random variable. Goldstein's maximal coupling implies the existence of a random variable N so that for all $k \geq N$

$$\overline{x(k)} = \overline{X \circ \theta^k} \quad \text{a.s.} ,$$

where, for notational convenience, we have identified the versions of the random variables on the underlying common probability space with the original ones. Let x_0 denote the initial value of the recurrence relation, then we may rephrase the above equation as

$$\overline{A(k) \otimes \cdots \otimes A(0) \otimes x_0} = \overline{X \circ \theta^k} , \quad k \geq N ,$$

or, equivalently,

$$\overline{A(0) \otimes A(-1) \otimes \cdots \otimes A(-k) \otimes x_0} = \overline{X} , \quad k \geq N ,$$

where $\{A(k) : k = \ldots 1, 0, -1, \ldots\}$ denotes the continuation of the stationary sequence $\{A(k)\}$ to the negative numbers. Hence, for $X \in \overline{X}$ there exists $a \in \mathbb{R}$ so that

$$A(0) \otimes A(-1) \otimes \cdots \otimes A(-k) \otimes x_0 = a \otimes X , \quad k \geq N .$$

This implies, for $k \geq N$,

$$\begin{aligned} A(1) \otimes & A(0) \otimes A(-1) \otimes \cdots \otimes A(-k) \otimes x_0 - A(0) \otimes \cdots \otimes A(-k) \otimes x_0 \\ &= A(1) \otimes a \otimes X - a \otimes X \\ &= A(1) \otimes X - X , \end{aligned}$$

where a.s. regularity of $\{A(k)\}$ and our assumption that $x_0 \in \mathbb{R}^J$ implies that the above differences are well-defined. Taking the limit,

$$\begin{aligned} \lim_{k \to \infty} & A(1) \otimes A(0) \otimes A(-1) \otimes \cdots \otimes A(-k) \otimes x_0 - A(0) \otimes \cdots \otimes A(-k) \otimes x_0 \\ &= A(1) \otimes A(0) \otimes A(-1) \otimes \cdots \otimes A(-N) \otimes x_0 - A(0) \otimes \cdots \otimes A(-N) \otimes x_0 \\ &= A(1) \otimes X - X , \end{aligned}$$

for all $X \in \overline{X}$. We introduce the following condition:

(D) *A random variable $Z \in [0, \infty)^J$ exists such that with probability one*

$$\sup_k |A(1) \otimes A(0) \otimes A(-1) \otimes \cdots \otimes A(-k) \otimes x_0 - A(0) \otimes \cdots \otimes A(-k) \otimes x_0| \leq Z$$

and $\mathbb{E}[Z]$ is finite.

In the next section we will provide sufficient conditions for **(D)**.

Suppose that condition (**D**) is satisfied, applying the dominated convergence theorem then yields

$$\begin{aligned}
&\lim_{k\to\infty} \mathbb{E}[x(k+1) - x(k)] \\
&= \lim_{k\to\infty} \mathbb{E}\Big[A(1)\otimes A(0)\otimes A(-1)\otimes\cdots\otimes A(-k)\otimes x_0 \;-\; A(0)\otimes\cdots\otimes A(-k)\otimes x_0\Big] \\
&= \mathbb{E}\left[\lim_{k\to\infty}\Big(A(1)\otimes A(0)\otimes A(-1)\otimes\cdots\otimes A(-k)\otimes x_0 \;-\; A(0)\otimes\cdots\otimes A(-k)\otimes x_0\Big)\right] \\
&= \mathbb{E}[A(1)\otimes X - X] < \infty \,.
\end{aligned}$$

Convergence of $\mathbb{E}[x(k+1) - x(k)]$ implies convergence of the Cesàro-sums (see Section G.1 in the Appendix) and we obtain

$$\begin{aligned}
\lim_{k\to\infty} \mathbb{E}[x(k+1) - x(k)] &= \lim_{k\to\infty} \frac{1}{k+1}\sum_{i=1}^{k+1} \mathbb{E}[x(i) - x(i-1)] \\
&= \lim_{k\to\infty} \mathbb{E}\left[\frac{1}{k+1}\sum_{i=1}^{k+1}(x(i) - x(i-1))\right] \\
&= \lim_{n\to\infty} \mathbb{E}\left[\frac{1}{k+1}x(k+1) - \frac{1}{k+1}x(0)\right] \\
&= \lim_{k\to\infty} \mathbb{E}\left[\frac{1}{k}x(k)\right] \,.
\end{aligned}$$

We summarize our analysis in the following theorem:

Theorem 2.6.1 *Consider the situation in (2.43). If*

- $\{\overline{x(k)} : k \geq 1\}$ *converges in total variation to* $\overline{x}$,
- $\{A(k)\}$ *is a.s. regular and stationary,*
- *condition* (**D**) *is satisfied,*

then there is an a.s. finite random variable N *so that*

$$\lim_{k\to\infty} \mathbb{E}\left[\frac{x(k)}{k}\right] = \mathbb{E}\left[A(1)\otimes \bigotimes_{i=-N}^{0} A(i)\otimes x_0 \;-\; \bigotimes_{i=-N}^{0} A(i)\otimes x_0\right] ,$$

for any finite initial value $x_0 \in \mathbb{R}^J$.

Under the conditions in Theorem 2.2.3, $\mathbb{E}[x_j(k)]/k$, $1 \leq j \leq J$, tends to the Lyapunov vector of $\{A(k)\}$ as k tends to ∞. This yields the following representation for the Lyapunov vector:

Lemma 2.6.1 *Consider the situation in (2.43). If*

(i) $\{\overline{x}(k) : k \geq 1\}$ *converges in total variation to* $\overline{x}$,

(ii) *condition* (**D**) *is satisfied,*

(iii) $\{A(k)\}$ *is an a.s. regular and stationary sequence of integrable matrices such that*

- $\{A(k)\}$ *has fixed support,*
- *any finite element is a.s. non-negative, and*
- *the elements on the diagonal are a.s. different from* ε,

then there is an a.s. finite random variable N *such that*

$$\mathbb{E}\left[A(1) \otimes \bigotimes_{i=-N}^{0} A(i) \otimes x_0 \; - \; \bigotimes_{i=-N}^{0} A(i) \otimes x_0 \right] = \vec{\lambda} \,,$$

for any integrable initial value $x_0 \in \mathbb{R}^J$, *where* $\vec{\lambda}$ *denotes the Lyapunov vector of* $\{A(k)\}$.

Lemma 2.6.1 can be stated in various forms. For example, if we replace condition (iii) by the condition that $\{A(k)\}$ has MLP, then we obtain that the components of the Lyapunov vector are equal, see Theorem 2.2.4.

Recall that we have introduced $\mathbf{e}$ as the vector with all elements equal to e. For $x \in \mathbb{R}$, the vector with all elements equal to x is then given by $x \otimes \mathbf{e}$. For sequences $\{A(k)\}$ with countable state-space, Lemma 2.6.1 can be phrased as follows:

Lemma 2.6.2 *Consider the situation in (2.43). If*

- (**C1**) – (**C3**) *are satisfied, and*
- *condition* (**D**) *is satisfied,*

then there is an a.s. finite random variable N *so that*

$$\mathbb{E}\left[A(1) \otimes \bigotimes_{i=-N}^{0} A(i) \otimes x_0 \; - \; \bigotimes_{i=-N}^{0} A(i) \otimes x_0 \right] = \lambda \otimes \mathbf{e} \,,$$

for any integrable initial value x_0, *where* λ *denotes the Lyapunov exponent of* $\{A(k)\}$.

Proof: Conditions (**C1**) – (**C3**) imply convergence of $\{\overline{x(k)} : k \geq 1\}$ in total variation, see Theorem 2.5.1. By condition (**C3**), a primitive matrix, say, C exists that is a pattern of $\{A(k)\}$, and we assume, for the sake of simplicity, that $C \in \mathcal{A}$, which implies $N = 1$, Let c denote the coupling time of C. From the i.i.d. assumption it follows that the event $\{A(c-1) = A(c-2) = \cdots = A(0) = C\}$ has positive probability and matrix C therefore satisfies condition (**C**). By Theorem 2.2.4 we obtain $\lim_{k\to\infty} \mathbb{E}[x_j(k)]/k = \lambda$, for $1 \leq j \leq J$. Hence, the proof of the lemma follows directly from Theorem 2.6.1. □

We conclude this section with a remark on the cycle formula in Section 2.5.4. Under the conditions put forward in the above Lemma it holds that

$$\mathbb{E}[\overline{A(1) \otimes X} - \overline{X}] = \lambda \otimes \mathbf{e} . \tag{2.44}$$

The cycle formula can therefore be rephrased as follows: let the conditions in Lemma 2.6.2 be satisfied and let $\{T_k\}$ denote the time of the k^{th} occurrence of the c-fold concatenation of C, see Section 2.5.4 for a formal definition. A simple geometrical trial argument, like the one used in the proof of Theorem 2.5.1, shows that

$$\mathbb{E}[T_1 - T_0] < \infty. \tag{2.45}$$

Elaborating on the limit theorem for regenerative processes (see Section 2.5.4 for details), (2.45) together with (2.44) implies $\mathbb{E}[x(T_1) - x(T_0)] < \infty$, and the cycle formula reads

$$\lambda \otimes \mathbf{e} = \frac{\mathbb{E}[x(T_1) - x(T_0)]}{\mathbb{E}[T_1 - T_0]} .$$

2.6.2 Backward Coupling

In the previous section, the existence of a coupling time N was shown. In this section, we will provide an explicit construction of N via *backward coupling.* In Markov chain theory, backward coupling, or, coupling from the past, is an approach that allows sampling from the stationary distribution of a finite-state Markov chain. Suppose that we consider a family of Markov chains X^s on a finite state space S, each with the same transition probabilities and with common unique stationary distribution π, but with version X^s starting in state $s \in S$. If we can find a time T in the past such that *all* versions X^s starting, not at time 0, but at time $-T$, have the *same value* at time 0, then this common value is a sample from π, see Theorem 1 in [92]. Intuitively, it is clear why this result holds with such a random time T. Consider a chain starting at $-\infty$ with π. This chain must at time $-T$ pick some value s, and from then on it follows the trajectory from that value. By definition of T, this trajectory reaches at time 0 the same state s' that is reached by X^s no matter what choice of s. Therefore, s' is a sample from π. Propp and Wilson coin the name 'coupling-from-the-past' for this algorithm since in essence $-T$ is a coupling time with the stationary version started at $-\infty$. Based on the same principles, Borovkov and Foss developed in [23, 22] the so-called 'renovating events' approach to stability analysis of stochastically recursive sequences. In particular, the approach to stability analysis via patterns (see Section 2.5) was originally inspired by backward coupling via 'renovating events.'

Elaborating on backward coupling, we combine our results for second-order limits with results for first-order limits in order to represent the Lyapunov exponent (a first-order limit) by the difference of two finite horizon experiments. We follow the line of argument in [7]. The key assumption for our analysis is that $\{A(k)\}$ possesses a pattern $\tilde{A}$ such that $\tilde{A}$ is primitive. The fact that $\{A(k)\}$ admits a pattern resembles a sort of memory loss property of max-plus linear systems. To see this, let $x(k+1) = A(k) \otimes x(k)$ be a stochastic sequence defined via $\{A(k)\}$ and assume that $\{A(k)\}$ has a pattern with associated matrix $\tilde{A}$ and that $\{A(k)\}$ is a.s. regular. For vectors $x, y \in \mathbb{R}^J$, let $x - y$ denote the component-wise difference, that is, $(x-y)_j = x_j - y_j$. In what follows we consider the limit of $x(k+1) - x(k)$ as k tends to ∞, where the limit has to be understood component-wise. In order to prove the existence of this limit we will work with a backward coupling argument. For this reason it is more convenient to let the index k run backwards. More precisely, we set

$$A^0_{-m} \stackrel{\text{def}}{=} A(0) \otimes A(-1) \otimes \cdots \otimes A(-m) \stackrel{\text{def}}{=} \bigotimes_{k=-m}^{0} A(k)$$

and

$$x^0_{-m} \stackrel{\text{def}}{=} A^0_{-m} \otimes x_0 = \bigotimes_{k=-m}^{0} A(k) \otimes x_0 ,$$

with $x^0_0 = x_0 \in \mathbb{R}^J$, that is, x^0_{-m} is the state of the sequence $\{x(k)\}$, started at time $-m$ in x_0, at time 0. The sequence $\{x^0_{-m} : m \geq 0\}$ evolves backwards in time according to

$$x^0_{-(m+1)} = A^0_{-m} \otimes A(-(m+1)) \otimes x_0 .$$

Note that $x(k+1)$ and x^0_{-k} are equal in distribution. With this notation the second-order limit reads

$$\lim_{k\to\infty} A(1) \otimes x^0_{-k} - x^0_{-k} = \lim_{k\to\infty} \left(A(1) \otimes \bigotimes_{m=-k}^{0} A(m) \otimes x_0 - \bigotimes_{m=-k}^{0} A(m) \otimes x_0 \right) .$$

Note that the above differences are well-defined due to the a.s. regularity of $\{A(k)\}$ and our assumption that $x_0 \in \mathbb{R}^J$.

Let condition (**C3**) be satisfied. Suppose that, after going η steps backwards in time, we observe for the first time the $c(\tilde{A})$-fold concatenation of the sequence constituting $\tilde{A}$, the pattern of $\{A(k)\}$. More precisely, let $(a_N, a_{N-1}, \ldots, a_1)$ denote the sequence constituting $\tilde{A}$, that is, $\tilde{A} = a_N \otimes \cdots \otimes a_1$, and let $\tilde{a}$ denote the $c(\tilde{A})$-fold concatenation of the string $(a_N, a_{N-1}, \ldots, a_1)$, which implies that $\tilde{a}$ has $M = c(\tilde{A}) \cdot N$ components. Then,

$$\tilde{A}^{c(\tilde{A})} = \bigotimes_{k=1}^{M} a_k$$

and η is defined by

$$\eta = \inf\{k \geq 0 \,|\, A(-k) = a_1, A(-k+1) = a_2, \ldots, A(-k+(M-1)) = a_M\} . \tag{2.46}$$

In accordance with Theorem 2.1.1, we obtain that the random variable

$$\bigotimes_{k=-(\eta+n)}^{0} A(k) \otimes x_0\,, \quad n \geq 0\,,$$

is an eigenvector of $\tilde{A}$, in formula:

$$\bigotimes_{k=-(\eta+n)}^{0} A(k) \otimes x_0 \in V(\tilde{A})\,, \quad n \geq 0\,.$$

Remark 2.6.1 *The random variable η denotes the index of the matrix that completes the first occurrence of $\tilde{a}$. Since we start counting the elements of the series of matrices from zero, the total number of transitions until this happens is $\eta + 1$.*

Recall that multiplication of a vector $v \in \mathbb{R}^J_{\max}$ with a scalar $\gamma \in \mathbb{R}_{\max}$ is defined by component-wise multiplication: $(\gamma \otimes u)_j = \gamma \otimes u_j$. It can be easily checked that

$$\forall \gamma \in \mathbb{R}_{\max}\,,\, v \in \mathbb{R}^J_{\max} : \quad B \otimes v - C \otimes v = B \otimes (\gamma \otimes v) - C \otimes (\gamma \otimes v)\,, \tag{2.47}$$

for all $B, C \in \mathbb{R}^{I \times J}_{\max}$. We now use the fact that the eigenvector of a primitive matrix is unique (up to scalar multiplication): if $u, v \in V(A)$, then a $\gamma \in \mathbb{R}_{\max}$ exists such that $v = \gamma \otimes u$, see Corollary 2.1.1. Hence, (2.47) implies

$$\forall v, u \in V(A) : \quad B \otimes v - C \otimes v = B \otimes u - C \otimes u\,, \tag{2.48}$$

for matrices $A, B, C \in \mathbb{R}^{J \times J}_{\max}$. Combining the above arguments, we obtain

$$\begin{aligned}
&\lim_{k\to\infty} A(1) \otimes x^0_{-k} - x^0_{-k} \\
&= \lim_{k\to\infty} \left(A(1) \otimes \bigotimes_{m=-k}^{0} A(m) \otimes x_0 - \bigotimes_{m=-k}^{0} A(m) \otimes x_0 \right) \\
&= A(1) \otimes \bigotimes_{m=-\eta+M}^{0} A(m) \otimes \underbrace{\tilde{A}^{c(\tilde{A})} \otimes \bigotimes_{m=-\infty}^{-\eta-1} A(m) \otimes x_0}_{\in V(\tilde{A})} \\
&\quad - \bigotimes_{m=-\eta+M}^{0} A(m) \otimes \underbrace{\tilde{A}^{c(\tilde{A})} \otimes \bigotimes_{m=-\infty}^{-\eta-1} A(m) \otimes x_0}_{\in V(\tilde{A})}
\end{aligned}$$

$$
\begin{aligned}
&\overset{(2.48)}{=} A(1) \otimes \bigotimes_{m=-\eta+M}^{0} A(m) \otimes \underbrace{\tilde{A}^{c(\tilde{A})} \otimes x_0}_{\in V(\tilde{A})} \\
&\qquad - \bigotimes_{m=-\eta+M}^{0} A(m) \otimes \underbrace{\tilde{A}^{c(\tilde{A})} \otimes x_0}_{\in V(\tilde{A})} \\
&= A(1) \otimes \bigotimes_{m=-\eta}^{0} A(m) \otimes x_0 - \bigotimes_{m=-\eta}^{0} A(m) \otimes x_0 \\
&= A(1) \otimes A_{-\eta}^{0} \otimes x_0 - A_{-\eta}^{0} \otimes x_0 < \infty .
\end{aligned}
$$

Hence, the second-order limit can be represented by a random horizon experiment.

Next, we will show that the above limit representation also holds if we consider expected values. We have assumed that $x_0 \in \mathbb{R}^J$. This together with a.s. regularity of $A(k)$ yields that $x(k) \in \mathbb{R}^J$ a.s. for all k. Let $(\cdot)_j$ denote the projection on the j^{th} component. Applying Lemma 1.6.1 yields

$$
\begin{aligned}
&\left| \left(A(1) \otimes \bigotimes_{k=-m}^{0} A(k) \otimes x_0 \right)_j - \left(\bigotimes_{k=-m}^{0} A(k) \otimes x_0 \right)_j \right| \\
&\qquad \leq \left\| A(1) \otimes \bigotimes_{k=-m}^{0} A(k) \otimes x_0 \right\|_{\oplus} + \left\| \bigotimes_{k=-m}^{0} A(k) \otimes x_0 \right\|_{\oplus} \\
&\qquad \leq 2 \sum_{k=-m}^{1} ||A(k)||_{\oplus} + 2||x_0||_{\oplus} .
\end{aligned}
$$

From the preceding analysis follows that, for any m,

$$
\begin{aligned}
&\left| \left(A(1) \otimes \bigotimes_{k=-m}^{0} A(k) \otimes x_0 \right)_j - \left(\bigotimes_{k=-m}^{0} A(k) \otimes x_0 \right)_j \right| \\
&\qquad \leq 2 \sum_{k=-\eta}^{1} ||A(k)||_{\oplus} + 2||x_0||_{\oplus} .
\end{aligned}
$$

Let $A(1)$ be integrable, then $\mathbb{E}[||A(1)||_{\oplus}] < \infty$, and assume that $\mathbb{E}[\eta] < \infty$. By construction, for $m \geq 0$, the event $\{\eta = m\}$ is independent of $\{A(-k) : k > m\}$. Provided that $\{A(k)\}$ is i.i.d., Wald's equality (see Section E.8 in the Appendix) yields

$$
\mathbb{E}\left[\sum_{k=-\eta}^{1} ||A(n)||_{\oplus} \right] = \mathbb{E}[\eta + 1]\, \mathbb{E}[||A(1)||_{\oplus}] < \infty .
$$

Hence, provided that $\mathbb{E}[\eta] < \infty$, we may apply the dominated convergence theorem to the second-order limit and obtain

$$
\begin{aligned}
&\lim_{k\to\infty} \mathbb{E}[x(k+1) - x(k)] \\
&\quad = \lim_{k\to\infty} \mathbb{E}\left[A(1) \otimes x^0_{-k} - x^0_{-k}\right] \\
&\quad = \mathbb{E}\left[\lim_{k\to\infty} \left(A(1) \otimes x^0_{-k} - x^0_{-k}\right)\right] \\
&\quad = \mathbb{E}\left[\bigotimes_{k=-\eta}^{1} A(k) \otimes x_0 \; - \; \bigotimes_{k=-\eta}^{0} A(k) \otimes x_0\right] < \infty \, .
\end{aligned} \tag{2.49}
$$

In particular, the above analysis shows that if $\mathbb{E}[\eta] < \infty$, then **(C1)** – **(C3)** already imply **(D)**, and Lemma 2.6.2 can be phrased as follows:

Theorem 2.6.2 *Let $\{A(k)\}$ be a sequence of integrable matrices. If* **(C1)**–**(C3)** *are satisfied, then the Lyapunov exponent of $\{A(k)\}$, denoted by λ, exists and it holds for any initial vector $x_0 \in \mathbb{R}^J$:*

$$
\begin{aligned}
\lambda \otimes \mathbf{e} &= \mathbb{E}\left[\bigotimes_{k=-\eta}^{1} A(k) \otimes x_0 \; - \; \bigotimes_{k=-\eta}^{0} A(k) \otimes x_0\right] \\
&= \lim_{k\to\infty} \frac{1}{k} \mathbb{E}\left[\bigotimes_{i=0}^{k-1} A(i) \otimes x_0\right] \\
&= \lim_{k\to\infty} \frac{1}{k} \bigotimes_{i=0}^{k-1} A(i) \otimes x_0 \qquad a.s.
\end{aligned}
$$

Proof: We show that $\mathbb{E}[\eta]$ is finite. By assumption **(C3)**, a primitive matrix, say, C exists that is a pattern, and we assume, for the sake of simplicity, that $C \in \mathcal{A}$, which implies $N = 1$. Let c denote the coupling time of C. Because the state space is discrete and the sequence is i.i.d., the probability of observing C, denoted by p, is larger than 0. If $p = 1$, then $\mathbb{E}[\eta] = c$. In case $0 < p < 1$, we argue as follows. By construction, the probability of the event $\{\eta = m\}$ is less than or equal to the probability of the event that $A(k) \neq C$, $0 \geq k \geq -m + c$, and $A(k) = C$, for $k = -m + c - 1, \ldots, -m$. In other words, for $m \geq c$, it holds that $P(\eta = m) \geq (1-p)^{m-c} p^c$. This implies

$$
\begin{aligned}
\mathbb{E}[\eta] &\leq \sum_{m=c}^{\infty} m \, (1-p)^{m-c} p^c \\
&= \sum_{m=0}^{\infty} (m + c) \, (1-p)^m p^c \\
&= c\, p^c \sum_{m=0}^{\infty} (1-p)^m \; + \; p^c \sum_{m=0}^{\infty} m \, (1-p)^m
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{c\,p^c}{p} + \frac{p^c(1-p)}{p^2} \\
&< \infty \,,
\end{aligned}
$$

which concludes the proof. □

We conclude this section with revisiting the cycle formula in Lemma 2.5.2.

Corollary 2.6.1 *Let* **(C1)** – **(C3)** *be satisfied. If* C *is a pattern of* $\{A(k)\}$*, then, for any finite initial vector* $x_0 \in V(C)$*,*

$$
\lambda \otimes \mathbf{e} = \frac{\mathbb{E}[x(\eta) - x_0]}{\mathbb{E}[\eta]} \,,
$$

where λ *denotes the Lyapunov exponent of* $\{A(k)\}$*.*