

# Max-Plus Algebra

Kasie G. Farlow

Master's Thesis submitted to the Faculty of the  
Virginia Polytechnic Institute and State University  
in partial fulfillment of the requirements for the degree of

Masters  
in  
Mathematics

Dr. Martin Day, Chair  
Dr. Peter Haskell,  
Dr. Robert Wheeler

April 27, 2009  
Blacksburg, Virginia

Keywords: Max-Plus, Linear Algebra, Markov Chains

Copyright 2009, Kasie G. Farlow

# Max-Plus Algebra

Kasie G. Farlow

(Abstract)

In max-plus algebra we work with the max-plus semi-ring which is the set  $\mathbb{R}_{\max} = \{-\infty\} \cup \mathbb{R}$  together with operations  $a \oplus b = \max(a, b)$  and  $a \otimes b = a + b$ . The additive and multiplicative identities are taken to be  $\epsilon = -\infty$  and  $e=0$  respectively. Max-plus algebra is one of many idempotent semi-rings which have been considered in various fields of mathematics. Max-plus algebra is becoming more popular not only because its operations are associative, commutative and distributive as in conventional algebra but because it takes systems that are non-linear in conventional algebra and makes them linear. Max-plus algebra also arises as the algebra of asymptotic growth rates of functions in conventional algebra which will play a significant role in several aspects of this thesis. This thesis is a survey of max-plus algebra that will concentrate on max-plus linear algebra results. We will then consider from a max-plus perspective several results by Wentzell and Freidlin for finite state Markov chains with an asymptotic dependence.

# Contents

**Abstract** **ii**

**Table Of Contents** **iii**

**List of Figures** **v**

**1 Introduction** **1**

    1.1 Brief History . . . . . 1

    1.2 Definitions and Basic Algebraic Properties . . . . . 8

    1.3 Matrices and Vectors in Max Plus Algebra . . . . . 11

        1.3.1 Matrices . . . . . 11

        1.3.2 Vectors . . . . . 13

    1.4 Max-Plus and Graph Theory . . . . . 13

**2 Max-Plus Linear Algebra** **18**

    2.1 Inverse Matrices . . . . . 18

    2.2 Determinants . . . . . 23

|          |  |           |
|----------|--|-----------|
| 2.3      | Linear Systems . . . . .   | 26        |
| 2.3.1    | Principal Sub-Solution . . . . .                                   | 26        |
| 2.3.2    | Cramer's Rule . . . . .  | 28        |
| 2.3.3    | Solving $x = (A \otimes x) \oplus b$ . . . . .                     | 31        |
| 2.4      | Eigenvalue and Eigenvectors . . . . .                              | 32        |
| 2.5      | Cayley-Hamilton and the Max-Plus Characteristic Equation . . . . . | 45        |
| 2.6      | Linear Dependence and Independence . . . . .                       | 51        |
| 2.6.1    | Bases . . . . .  | 52        |
| 2.7      | Asymptotic and Limiting Behavior . . . . .                         | 53        |
| <b>3</b> | <b>Markov Chains</b>   | <b>61</b> |
| 3.0.1    | Problem 1 . . . . .  | 63        |
| 3.0.2    | Problem 2 . . . . .  | 72        |
| 3.0.3    | Problem 3 . . . . .  | 77        |
| <b>4</b> | <b>Concluding Remarks</b>  | <b>80</b> |
|          | <b>Bibliography</b>  | <b>81</b> |
|          | <b>List of Symbols</b>   | <b>85</b> |

# List of Figures

|     |                         |    |
|-----|-------------------------|----|
| 1.1 | Train problem . . . . . | 3  |
| 3.1 | $G(U_1)$ . . . . .      | 66 |
| 3.2 | $G(U_2)$ . . . . .      | 68 |
| 3.3 | $G(U_3)$ . . . . .      | 75 |

# Chapter 1

## Introduction

### 1.1 Brief History

In max-plus algebra we work with the max-plus semi-ring which is the set  $\mathbb{R}_{\max} = \{-\infty\} \cup \mathbb{R}$  together with operations  $a \oplus b = \max(a, b)$  and  $a \otimes b = a + b$ . The additive and multiplicative identities are taken to be  $\epsilon = -\infty$  and  $e = 0$  respectively. Its operations are associative, commutative and distributive as in conventional algebra.

Max-plus algebra is one of many idempotent semi-rings which have been considered in various fields of mathematics. One other is min-plus algebra. There  $\oplus$  means minimum and the additive identity is  $\infty$ . See [26] for other examples. We will only be concerned with max-plus algebra. It first appeared in 1956 in Kleene's paper on nerve sets and automata [18],[15]. It has found applications in many areas such as combinatorics, optimization, mathematical physics and algebraic geometry [15]. It is also used in control theory, machine scheduling, discrete event processes, manufacturing systems, telecommunication networks, parallel processing systems and traffic control, see [7], [12] and [9]. Many equations that are used to describe the behavior of these applications are nonlinear in conventional algebra but become linear in max-plus algebra. This is a primary reason for its utility in various areas [7].

Many of the theorems and techniques we use in classical linear algebra have analogues in the max-plus semi-ring. Cunningham-Green, Gaubert, Gondran and Minoux are among the researchers who have devoted a lot of time creating much of the max-plus linear algebra theory we have today. Many of Cunningham-Greens' results are found in [9]. They have studied concepts such as solving systems of linear equations, the eigenvalue problem, and linear independence in the max-plus sense. In Chapter 2 we will see the extent to which max-plus algebra is an analogue of classical linear algebra and look at many max-plus counterparts of conventional results.

To illustrate the usefulness of max-plus algebra in a simple example, let's look at a railroad network between two cities. A similar example can be found in [16]. This is an example of how max-plus algebra can be applied to a discrete event system. Assume we have two cities such that  $S_1$  is the station in the first city, and  $S_2$  is the station in the second city. This system contains 4 trains. The time it takes a train to go from  $S_1$  to  $S_2$  is 3 hours where the train travels along track 1. It takes 2 hours to go from  $S_2$  to  $S_1$  where the train travels along track 2. These tracks can be referred to as the long distance tracks. There are two more tracks in this network, one which runs through city 1 and one which runs through city 2. We can refer to these as the inner city tracks. Call them tracks 3 and 4 respectively. We can picture track 3 as a loop beginning and ending at  $S_1$ . Similarly, track 4 starts and ends at  $S_2$ . The time it takes to traverse the loop on track 3 is 2 hours. The time it takes to travel from  $S_2$  to  $S_2$  on track 4 is 4 hours. Track 3 and track 4 each contain a train. There are two trains that circulate along the two long distance tracks. In this network we also have the following criteria:

1. The travel times along each track indicated above are fixed
2. The frequency of the trains must be the same on all four tracks
3. Two trains must leave a station simultaneously in order to wait for the changeover of passengers

4. The two  $(k + 1)^{st}$  trains leaving  $S_i$  can not leave until the  $k^{th}$  train that left the other station arrives at  $S_i$ .

$x_i(k - 1)$  will denote the the  $k^{th}$  departure time for the two trains from station  $i$ . Therefore  $x_1(k)$  denotes the departure time of the pair of  $k + 1$  trains from  $S_1$  and  $x_2(k)$  is the departure time of the  $k + 1$  trains from  $S_2$ .  $x(0)$  is a vector denoting the departure times of the first trains from  $S_1$  and  $S_2$ . So  $x_1(0)$  denotes the departure time of the first pair of trains from station 1 and likewise  $x_2(0)$  denotes the departure time of the first pair of trains from station 2. See figure 1.1.

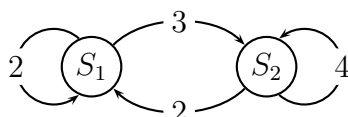


Figure 1.1: Train problem

Let's say we want to determine the departure time of the  $k^{th}$  trains from station 1. We can see that

$$\begin{aligned} x_1(k + 1) &\geq x_1(k) + a_{11} + \delta \quad \text{and} \\ x_1(k + 1) &\geq x_2(k) + a_{12} + \delta \end{aligned}$$

where  $a_{ij}$  denotes the travel time from station  $j$  to station  $i$  and  $\delta$  is the time allowed for the passengers to get on and off the train. So in our situation we have  $a_{12} = 2$ ,  $a_{11} = 2$ ,  $a_{22} = 4$  and  $a_{21} = 3$ . We will assume  $\delta = 0$  in this example. So it follows that

$$x_1(k + 1) = \max\{x_1(k) + a_{11}, x_2(k) + a_{12}\}.$$

Similarly we can see that

$$x_2(k + 1) = \max\{x_1(k) + a_{21}, x_2(k) + a_{22}\}.$$

In conventional algebra we would determine successive departure times by iterating the nonlinear system



$$x_i(k+1) = \max_{j=1,2,\dots,n} \{a_{ij} + x_j(k)\}.$$

In max-plus we would express this as

$$x_i(k+1) = \bigoplus_{j=1}^n (a_{ij} \otimes x_j(k)), \quad i = 1, 2, \dots, n$$

where  $\bigoplus_{j=1}^n (a_{ij} \otimes x_j) = (a_{i1} \otimes x_1) \oplus (a_{i2} \otimes x_2) \oplus \dots \oplus (a_{in} \otimes x_n)$  for  $i = 1, 2, \dots, n$ .

In our example we have  $x_1(1) = 0 \oplus 2 = 2$  and  $x_2(1) = 1 \oplus 4 = 4$  provided we are given  $x_1(0) = -2$  and  $x_2(0) = 0$ .

We can create a matrix  $A$  using the values  $a_{ij}$  such that  $A = \begin{pmatrix} 2 & 2 \\ 3 & 4 \end{pmatrix}$  and  $x(0) = \begin{pmatrix} -2 \\ 0 \end{pmatrix}$ .

We can also express this system using matrices and vectors such that  $x(k) = A \otimes x(k-1)$ :

$$x(1) = A \otimes x(0), \quad x(2) = A \otimes x(1) = A \otimes A \otimes x(0) = A^{\otimes 2} \otimes x(0).$$

(See section 1.2 for max-plus matrix operations.) Continuing in this fashion we see that  $x(k) = A^{\otimes k} \otimes x(0)$ . This gives us a simple example of how a system of equations which is not linear in conventional algebra is linear in max-plus algebra.

The questions of eigenvectors and eigenvalues are very prevalent in the max-plus algebra. In the previous problem  $x(0)$  happens to be an eigenvector of  $A$  with eigenvalue  $\lambda = 4$ . (We will see how to compute this later.) We will look at the asymptotics of and limiting behavior of  $x(k)$  as  $k \rightarrow \infty$  and show how  $A^{\otimes k} \otimes x(0)$  is related to the eigenvalue problem. We will see that this leads to a simple relationship between  $x(k)$  and  $\lambda$ .

Max-plus algebra also arises as the algebra of asymptotic growth rates of functions in conventional algebra. The max-plus operations  $\oplus$  and  $\otimes$  are operations on exponents induced by conventional algebra applied to exponential functions  $e^{sa}$ :

$$e^{sa} e^{sb} \asymp e^{s(a \otimes b)}$$

$$e^{sa} + e^{sb} \asymp e^{s(a \oplus b)}$$

The following definition and lemma makes this precise. (We follow the conventions  $\ln(0) = -\infty$  and  $e^{-\infty} = 0$ .)

**Definition 1.** If  $p : (0, \infty) \rightarrow (0, \infty)$  and  $u \in [-\infty, \infty)$ , we define  $p \asymp e^{su}$  to mean  $\lim_{s \rightarrow \infty} s^{-1} \ln(p) = u$ .

**Lemma 1.** If  $f \asymp e^{sa}$  and  $g \asymp e^{sb}$ , then

$$f + g \asymp e^{s(a \oplus b)} \text{ and } fg \asymp e^{s(a \otimes b)}.$$

*Proof.* First we see that

$$\lim_{s \rightarrow \infty} s^{-1} \ln(fg) = \lim_{s \rightarrow \infty} s^{-1} \ln(f) + \lim_{s \rightarrow \infty} s^{-1} \ln(g) = a + b = a \otimes b$$

Thus  $fg \asymp e^{s(a \otimes b)}$ . Now notice that  $\max(f, g) \leq f + g \leq 2 \max(f, g)$ . Then by applying Definition 1 we have

$$\lim_{s \rightarrow \infty} s^{-1} \ln(\max(e^{sa}, e^{sb})) \leq \lim_{s \rightarrow \infty} s^{-1} \ln(e^{sa} + e^{sb}) \leq \lim_{s \rightarrow \infty} s^{-1} [\ln(\max(e^{sa}, e^{sb})) + \ln(2)].$$

Since  $s^{-1} \ln(\max(e^{sa}, e^{sb})) = \max(a, b)$ , by applying the squeeze theorem we see that

$$\lim_{s \rightarrow \infty} s^{-1} \ln(f + g) = \max(a, b) = a \oplus b.$$

□

This connection with conventional algebra has been used to prove many properties in max-plus-algebra [16]. It is used to generalize Cramer's Rule and the Cayley-Hamilton Theorem as we will see in Chapter 2. We will also exploit it in Chapter 3 where we will see how some results of Wentzell and Freidlin about finite Markov chains relate to max-plus linear algebra.

Applications of max-plus in infinite dimensional settings is an emerging area of research. Although this thesis is limited to finite dimensional settings, we want to indicate what

some of the infinite dimensional issues are. Instead of vectors  $v \in \mathbb{R}_{\max}$  this will involve problems for  $\mathbb{R}_{\max}$ -valued functions  $\phi : X \rightarrow \mathbb{R}_{\max}$  defined on some domain. We might call this *max-plus functional analysis*. Just as many finite dimensional optimization problems become linear from the max-plus perspective, the nonlinear equations of continuous state optimization problems (such as optimal control) likewise become linear in a max-plus context.

Imagine that we are considering optimization problems for a controlled system of ordinary differential equations,

$$\dot{x}(t) = f(x(t), u(t)),$$

where the control function  $u(t)$  takes values in some prescribed control set  $U$ . Typical optimization problems involve some sort of running cost function  $L(x, u)$ . For instance a finite horizon problem would have a specified terminal time  $T$  and terminal cost function  $\Phi(\cdot)$ . For a given initial state  $x(t) = x$ ,  $t < T$  the goal would be to maximize

$$J(x, t, u(\cdot)) = \int_t^T L(x(s), u(s)) ds + \Phi(x(T))$$

for a given initial condition  $x = x(t)$  over all allowed control functions  $u(\cdot)$ . In other words we would want

$$V(x) = S_T[\Phi](x),$$

where

$$S_T[\phi](x) = \sup_{u(\cdot)} \left\{ \int_0^T L(x(s), u(s)) ds + \phi(x(T)) \right\}.$$

So  $S_T$  is the solution operator. In other problems, like the nonlinear  $H_\infty$  problem of [21], the desired solution turns out to be a fixed point:  $W = S_T[W]$ .

In the conventional sense  $S_T$  is a nonlinear operator. (In fact  $\{S_T : T \geq 0\}$  forms a nonlinear semigroup; see [11].) However it is linear in the max-plus sense:

$$S_T[c \otimes \phi] = S_T[c + \phi] = c + S_T[\phi] = c \otimes S_T[\phi]$$

and

$$S_T[\phi \oplus \psi] = S_T[\max(\phi, \psi)] = \max(S_T[\phi], S_T[\psi]) = S_T[\phi] \oplus S_T[\psi].$$

With this observation one naturally asks if it is possible to develop max-plus analogues of eigenfunction expansions and something like the method of separation of variables in the context of these nonlinear problems. The idea would be to make an appropriate choice of basis functions  $\psi_i : X \rightarrow \mathbb{R}_{\max}$ , use an approximation

$$\psi(x) \approx \bigoplus_1^N a_i \otimes \psi_i(x),$$

and then take advantage of the max-plus linearity to write

$$S_T[\psi](x) \approx \bigoplus_1^N a_i \otimes S_T[\psi_i](x).$$

If the  $\psi_i$  are chosen so that the expansion  $S_T[\psi_i] \approx \bigoplus_j b_{ij} \psi_j$  can be worked out, then an approximation to the finite time optimization problems  $S_T[\Phi] \asymp \bigoplus_1^N c_i \otimes \psi_i$  where  $\Phi = \bigoplus_j a_j \psi_j$  would be given by a max-plus matrix product:

$$[c_i] = B \otimes [a_j].$$

For problems that reduce to a fixed point of  $S_T$ , we might hope to obtain approximations by solving a (finite-dimensional) max-plus eigenvector problem:

$$[c_i] = B \otimes [c_i].$$

To do all this carefully, one must choose the appropriate function spaces in which to work, and carry out some sort of error analysis. This has in fact been done by W. M. McEneaney for the  $H_\infty$  problem referred to above; see [21], [22]. Moreover methods of this type offer the prospect of avoiding the so-called ‘‘curse of dimensionality;’’ see [23]. But there are many questions about how to do this effectively in general cases. For instance, what basis functions should one use? At present relatively little research has been done in this direction, aside from the papers of McEneaney cited. McEneaney’s book [20] provides an introduction to many of these ideas.

## 1.2 Definitions and Basic Algebraic Properties

In this section we look more carefully at the algebraic properties of max-plus algebra. Many references discuss this material, such as [3], [6], [16], [20] and [25]. Recall that in max-plus algebra for  $a, b \in \mathbb{R}_{\max} = \{-\infty\} \cup \{\mathbb{R}\}$ , we define two operations  $\oplus$  and  $\otimes$  by

$$a \oplus b = \max(a, b) \quad a \otimes b = a+b.$$

For example,

$$\begin{aligned} 6 \oplus 2 &= \max(6, 2) = 6 = \max(2, 6) = 2 \oplus 6 \\ 7 \otimes 5 &= 7+5=12=5+7=5 \otimes 7. \end{aligned}$$

In max-plus algebra  $e = 0$  is the multiplicative identity:

$$a \otimes e = e \otimes a = a + 0 = a \text{ for all } a \in \mathbb{R}_{\max}.$$

The additive identity is  $\epsilon = -\infty$ :

$$a \oplus \epsilon = \epsilon \oplus a = \max(a, -\infty) = a \text{ for } a \in \mathbb{R}_{\max}.$$

Clearly  $\oplus$  and  $\otimes$  are commutative, and obey many other properties similar to  $+$  and  $\times$  in conventional algebra. For instance we can see that  $\otimes$  distributes over  $\oplus$  as follows:

for  $a, b, c \in \mathbb{R}_{\max}$  we have

$$a \otimes (b \oplus c) = a + \max(b, c) = \max(a + b, a + c) = (a \otimes b) \oplus (a \otimes c).$$

This and other basic algebraic properties are listed in the following lemma. The proofs are elementary and are not included.

**Lemma 2.** *For all  $x, y, z \in \mathbb{R}_{\max}$*

1. Associativity  $x \oplus (y \oplus z) = (x \oplus y) \oplus z$  and  $x \otimes (y \otimes z) = (x \otimes y) \otimes z$
2. Commutativity  $x \oplus y = y \oplus x$  and  $x \otimes y = y \otimes x$
3. Distributivity  $x \otimes (y \oplus z) = (x \otimes y) \oplus (x \otimes z)$
4. Zero Element  $x \oplus \epsilon = \epsilon \oplus x = x$
5. Unit Element  $x \otimes e = e \otimes x = x$
6. Multiplicative Inverse if  $x \neq \epsilon$  then there exists a unique  $y$  with  $x \otimes y = e$
7. Absorbing Element  $x \otimes \epsilon = \epsilon \otimes x = \epsilon$
8. Idempotency of Addition  $x \oplus x = x$

**Definition 2.** For  $x \in \mathbb{R}_{\max}$  and  $n \in \mathbb{N}$

$$x^{\otimes n} = \underbrace{x \otimes x \otimes \dots \otimes x}_{n \text{ times}}.$$

In the max-plus algebra exponentiation reduces to conventional multiplication  $x^{\otimes n} = nx$ .

It would be natural to extend max-plus exponentiation to more general exponents as follows.

- if  $x \neq \epsilon$ ,  $x^{\otimes 0} = e = 0$ .
- if  $\alpha \in \mathbb{R}$ ,  $x^{\otimes \alpha} = \alpha x$
- if  $k > 0$  then  $\epsilon^{\otimes k} = \epsilon$  ( $k \leq 0$  is not defined)

However none of these are needed below.

Here are the laws of exponents in max-plus.

**Lemma 3.** For  $m, n \in \mathbb{N}$ ,  $x \in \mathbb{R}_{\max}$

1.  $x^{\otimes m} \otimes x^{\otimes n} = mx + nx = (m + n)x = x^{\otimes (m \oplus n)}$

$$2. (x^{\otimes m})^{\otimes n} = (mx)^{\otimes n} = nm x = x^{\otimes (m^{\otimes n})}$$

$$3. x^{\otimes 1} = 1x = x$$

$$4. x^{\otimes m} \otimes y^{\otimes m} = (x \otimes y)^{\otimes m}$$

Using  $\oplus$  we can define the existence of order in the max- plus semi-ring.

**Definition 3.** *We say that:*

$$a \leq b \text{ if } a \oplus b = b$$

The following definitions which are used to describe max-plus algebra and can be found in [17]:

**Definition 4.** *A binary operation  $*$  is called idempotent on a set  $R$  if for all  $x \in R$   $x*x = x$*

**Definition 5.** *A monoid is a closed set under an associative binary operations which has a multiplicative identity.*

**Definition 6.** *A semi-ring is a commutative monoid which has no additive identity.*

Two important aspects of max-plus algebra are that it does not have additive inverses and it is idempotent. This is why max-plus algebra is considered a semi-ring and not a ring. The following lemma from [16] generalizes to all idempotent semi-rings.

**Lemma 4.** *The idempotency of  $\oplus$  in the max-plus semi-ring implies that for every  $a \in \mathbb{R}_{\max} \setminus \{\epsilon\}$ ,  $a$  does not have an additive inverse.*

*Proof.* Suppose  $a \in \mathbb{R}_{\max}$  such that  $a \neq \epsilon$  has an inverse with respect to  $\oplus$ . Let  $b$  be the inverse of  $a$ . Then we would have

$$a \oplus b = \epsilon.$$

By adding  $a$  to the left of both sides of the equation we get

$$a \oplus (a \oplus b) = a \oplus \epsilon = a.$$

Using the associativity property and the idempotency property of  $\oplus$

$$a = a \oplus (a \oplus b) = (a \oplus a) \oplus b = a \oplus b = \epsilon$$

which is a contradiction since we assumed  $a \neq \epsilon$ . □

### 1.3 Matrices and Vectors in Max Plus Algebra

We have pointed out that many nonlinear optimization problems become linear in  $\mathbb{R}_{\max}$ . In this section we define vectors and matrices in  $\mathbb{R}_{\max}$ . In the next chapter we will develop max-plus versions of the standard results of linear algebra. Much of the material discussed in this section can be found in [3], [12], [16] and [4].

#### 1.3.1 Matrices

Here we will begin a discussion on matrices over  $\mathbb{R}_{\max}$ . The set of  $n \times m$  matrices for  $n, m \in \mathbb{N}$  over  $\mathbb{R}_{\max}$  is denoted by  $\mathbb{R}_{\max}^{n \times m}$ . The number of rows in such a matrix is  $n$  and  $m$  denotes the number of columns. As in conventional algebra we write a matrix  $A \in \mathbb{R}_{\max}^{n \times m}$  as follows:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix}$$



The entry in the  $i^{\text{th}}$  row and  $j^{\text{th}}$  column of  $A$  is denoted by  $a_{ij}$  or sometimes as  $[A]_{ij}$ .

Sums and products of max-plus vectors and matrices are defined in the usual way, replacing  $+$  and  $\times$  by  $\oplus$  and  $\otimes$ .

**Definition 7.** 1. For  $A, B \in \mathbb{R}_{\max}^{n \times n}$  define their sum,  $A \oplus B$  by

$$[A \oplus B]_{ij} = a_{ij} \oplus b_{ij} = \max(a_{ij}, b_{ij})$$

2. For  $A \in \mathbb{R}_{\max}^{n \times k}$  and  $B \in \mathbb{R}_{\max}^{k \times m}$ , define their product,  $A \otimes B$  by

$$[A \otimes B]_{il} = \bigoplus_{j=1}^k (a_{ij} \otimes b_{jl}) = \max_{j \in \{1, 2, \dots, k\}} (a_{ij} + b_{jl})$$

3. The transpose of a matrix is denoted by  $A^T$  and is defined as in conventional algebra

$$[A^T]_{ij} = [A]_{ji}.$$

4. The  $n \times n$  max-plus identity matrix,  $E_n$ , is defined as

$$[E_n]_{ij} = \begin{cases} e & \text{if } i = j \\ \epsilon & \text{if } i \neq j \end{cases}.$$

We will use  $E$  when the dimension is clear.

5. For a square matrix and positive integer  $k$ , the  $k^{\text{th}}$  power of  $A$  is denoted by  $A^{\otimes k}$  is defined by

$$A^{\otimes k} = \underbrace{A \otimes A \otimes \dots \otimes A}_{k \text{ times}}.$$

For  $k = 0$ ,  $A^{\otimes 0} = E_n$

6. For any matrix  $A \in \mathbb{R}_{\max}^{n \times m}$  and any scalar  $\alpha \in \mathbb{R}_{\max}$ ,  $\alpha \otimes A$  is defined by:

$$[\alpha \otimes A]_{ij} = \alpha \otimes [A]_{ij}$$

We have the following examples and elementary observations.

**Example:** Let  $A = \begin{pmatrix} 2 & 3 \\ e & 4 \end{pmatrix}$  and  $B = \begin{pmatrix} 3 & 5 \\ -1 & 4 \end{pmatrix}$ , then  $A \oplus B = B \oplus A = \begin{pmatrix} 3 & 5 \\ e & 4 \end{pmatrix}$ ,  
 $A \otimes B = \begin{pmatrix} 5 & 7 \\ 3 & 8 \end{pmatrix}$ , and  $B \otimes A = \begin{pmatrix} 5 & 9 \\ 4 & 8 \end{pmatrix}$

As usual  $\oplus$  is commutative for matrices, but  $\otimes$  is not.

The identity matrix is an identity with respect to  $\otimes$ .

$$A \otimes E_n = A \text{ for all } A \in \mathbb{R}_{\max}^{m \times n}, \text{ and}$$

$$E_m \otimes A = A \text{ for all } A \in \mathbb{R}_{\max}^{n \times m}$$

Observe that as before  $\otimes$  distributes over  $\oplus$  for matrices. Also  $\oplus$  is idempotent in  $\mathbb{R}_{\max}^{n \times n}$  since we have  $A \oplus A = A$ . So  $\mathbb{R}_{\max}^{n \times n}$  is another idempotent semi-ring. Note however that it is an idempotent semi-ring in which  $\otimes$  is noncommutative.

### 1.3.2 Vectors

The elements  $x \in \mathbb{R}_{\max}^n$  are called vectors (or max-plus vectors). The  $j^{\text{th}}$  component of a vector  $x$  is denoted by  $x_j$  or  $[x]_j$ . The  $j^{\text{th}}$  column of the identity matrix  $E_n$  is known as the  $j^{\text{th}}$  basis vector of  $\mathbb{R}_{\max}^n$ . This vector is denoted by  $e_j = (\epsilon, \epsilon, \dots, \epsilon, e, \epsilon, \epsilon, \dots, \epsilon)$ , in other words  $e$  is in the  $j^{\text{th}}$  entry of the vector. The concept of bases will be discussed further in Chapter 2.

## 1.4 Max-Plus and Graph Theory

Many results in max-plus algebra can be interpreted graphically. In particular, graph theory plays an important role in the eigenvalue and eigenvector problem, as we will discuss in the next chapter. We collect some basic definitions and results from [16] in this section.

**Definition 8.** A directed graph  $G$  is a pair  $(V, E)$  where  $V$  is the set of vertices of the graph  $G$  and  $E \subset V \times V$  is the set of edges of  $G$ .

A typical edge  $(i, j) \in E$  where  $i, j \in V$  is thought of as an arrow directed from  $i$  to  $j$ .

**Definition 9.** For  $A \in \mathbb{R}_{\max}^{n \times n}$  the communication graph of  $A$  is the graph  $G(A)$  with vertices  $V = \{1, 2, \dots, n\}$  and edges  $E = \{(i, j) : a_{ji} \neq \epsilon\}$ . For  $(i, j) \in E$   $a_{ji}$  is the weight of the edge.

The “edge reversal” is common in max-plus literature where weight of edge  $(i, j)$  is denoted by  $a_{ji}$ .

**Definition 10.** A path  $p$  from  $i$  to  $j$  in a graph is a sequence of edges  $p = (i_1, i_2, \dots, i_{s+1})$  with  $i_1 = i, i_{s+1} = j$  such that each  $(i_k, i_{k+1})$  is an edge of the graph. We say this path has length  $s$  and denote the length of the path by  $\|p\|_l = s$ . The set of paths from  $i$  to  $j$  of length  $k$  will be denoted by  $P(i, j, k)$ .

**Definition 11.** To say that a vertex  $j$  is reachable from a vertex  $i$  means that there exists a path from  $i$  to  $j$ . A strongly connected graph is a graph such that every vertex is reachable from every other vertex. We say a matrix  $A \in \mathbb{R}_{\max}^{n \times m}$  is irreducible if  $G(A)$  is strongly connected.

**Definition 12.** A circuit of length  $s$  is a closed path i.e. a path such that  $i_1 = i_{s+1}$ . We call a circuit consisting of one edge a loop. An elementary circuit is one in which  $i_1, i_2, \dots, i_s$  are distinct.

The paths and circuits of  $G(A)$  have weights. These weights are determined by the entries in the matrix  $A$ .

**Definition 13.** The weight of a path  $p$  from vertex  $i$  to  $j$  of length  $s$  is denoted by  $\|p\|_w = \bigotimes_{k=1}^s a_{i_{k+1}, i_k}$  where  $i = i_1$  and  $j = i_{s+1}$ .

For circuits we will be especially concerned with the average weight of a circuit. This plays an important role in the eigenvalue problem.

**Definition 14.** *The average weight of a circuit  $c$  is given by  $\frac{\|c\|_w}{\|c\|_1}$  (calculated in conventional algebra).*

The  $ij^{\text{th}}$  entry in  $A^{\otimes k}$  is the maximum weight of all paths of length  $k$  from vertex  $j$  to  $i$  in  $G(A)$ . The entry is  $\epsilon$  if no such paths exist. This is stated as the following theorem.

**Theorem 1.** *Let  $A \in \mathbb{R}_{\max}^{n \times n}$ . For all  $k \geq 1$*

$$\begin{aligned} [A^{\otimes k}]_{ji} &= \max\{\|p\|_w : p \in P(i, j, k)\}, \\ [A^{\otimes k}]_{ji} &= \epsilon \text{ if } P(i, j, k) = \emptyset \end{aligned}$$

*Proof.* This proof is by induction. Choose  $i$  and  $j$  arbitrarily.

Let  $k = 1$ . The only path in  $P(i, j, 1)$  is the edge  $(i, j)$  which has weight  $[A]_{ji}$ . Note that if  $[A]_{ji} = \epsilon$  then  $P(i, j, 1) = \emptyset$ . Now assume the theorem holds true for  $k - 1$  and consider a path  $p \in P(i, j, k)$ . We know that  $p$  is made up of a path  $\tilde{p}$  of length  $k - 1$  from  $i$  to some vertex  $\ell$  followed by an edge from  $\ell$  to  $j$ . So

$$p = ((i_1, i_2), (i_2, i_3), \dots, (i_{k-1}, i_k)) \text{ with } \tilde{p} = ((i_1, i_2), \dots, (i_{k-2}, i_{k-1})), \ell = i_{k-1}.$$

With this decomposition we can now obtain the maximal weight of the paths in  $P(i, j, k)$  as

$$\max_{\ell} ([A]_{j\ell} + \max\{\|\tilde{p}\|_w : \tilde{p} \in P(i, \ell, k - 1)\}).$$

By the induction hypothesis we know that

$$\max_{\ell} \{\|\tilde{p}\|_w : \tilde{p} \in P(i, \ell, k - 1)\} = [A^{\otimes k-1}]_{\ell i},$$

is the maximum path of length  $k - 1$  from  $i$  to  $\ell$ . Therefore the maximum weight of paths of length  $k$  from  $i$  to  $j$  is

$$\max_{\ell}(a_{j\ell} + [A^{\otimes k-1}]_{\ell i}) = \bigoplus_{\ell=1}^n (a_{j\ell} \otimes [A^{\otimes k}]_{\ell i}) = [A \otimes A^{\otimes k-1}]_{ji} = [A^{\otimes k}]_{ji}.$$

We need to mention the case where  $P(i, j : k) = \emptyset$ . This means that for any vertex  $\ell$  either there exists no path of length  $k - 1$  from  $i$  to  $\ell$  or there exists no edge from  $\ell$  to  $j$  or both. So  $P(i, j, k) = \emptyset$  and  $[A^{\otimes k}]_{ji} = \epsilon$ .  $\square$

We will see the significance of the following definitions and lemma in Chapter 2 when dealing with the eigenvalue problem, as well as in Chapter 3 with Markov chains.

**Definition 15.** For  $A \in \mathbb{R}_{\max}^{n \times n}$  let  $A^+ = \bigoplus_{k=1}^{\infty} A^{\otimes k}$  and  $A^* = E \oplus A^+ = \bigoplus_{k \geq 0} A^{\otimes k}$

This says that  $[A^+]_{ij}$  is the maximal weight of any path from vertex  $j$  to  $i$ .  $[A^+]_{ij} = \epsilon$  if no such path exists. It is also possible that  $[A^+]_{ij} = \infty$ . In this case we would say that  $[A^+]_{ij}$  is undefined.

**Lemma 5.** Suppose that  $A \in \mathbb{R}_{\max}^{n \times n}$  is such that the maximum average weight of any circuit in  $G(A)$  is less than or equal to  $e$  then  $A^+$  exists and is given by

$$A^+ = A \oplus A^{\otimes 2} \oplus \dots \oplus A^{\otimes n} = \bigoplus_{k=1}^n A^{\otimes k}.$$

*Proof.* By definition of  $A^+$  we have  $[A^+]_{ji} \geq \max\{[A^{\otimes k}]_{ji} : 1 \leq k \leq n\}$ . Since  $A$  is a  $n \times n$  matrix, all paths in  $G(A)$  from  $i$  to  $j$  of length greater than  $n$  must be made up of at least one circuit and a path from  $i$  to  $j$  of length at most  $n$ . By assumption the circuits in  $G(A)$  have weights less than or equal to zero. Therefore  $[A^+]_{ji} \leq \max\{[A^{\otimes k}]_{ji} : 1 \leq k \leq n\}$  which gives us  $A^+ = A \oplus A^{\otimes 2} \oplus \dots \oplus A^{\otimes n} = \bigoplus_{k=1}^n A^{\otimes k}$ .  $\square$

**Lemma 6.** If the circuits of  $G(A)$  all have negative weight then for all  $x \in \mathbb{R}_{\max}^n$

$$\lim_{k \rightarrow \infty} A^{\otimes k} \otimes x = \epsilon$$

where  $\epsilon$  is the matrix with all entries equal to  $\epsilon$ .

This is obvious since the entries  $[A^{\otimes k}]_{ij}$  tend to  $\epsilon$  as  $k \rightarrow \infty$ .

Next we define the cyclicity of a graph. This definition is the same as in conventional algebra and is equivalent to the period of a Markov chain.

**Definition 16.** *The cyclicity of a graph  $G$ , denoted by  $\sigma_G$ , is defined as follows:*

- *If  $G$  is strongly connected, then  $\sigma_G$  is the greatest common divisor of the lengths of all the elementary circuits in  $G$ .*
- *If  $G$  consists of just one node without a loop, then  $\sigma_G$  is one.*
- *If  $G$  is not strongly connected, then  $\sigma_G$  is the least common multiple of all the maximal strongly connected subgraphs of  $G$ .*

# Chapter 2

## Max-Plus Linear Algebra

This chapter is devoted to max-plus linear algebra. We will see that many concepts of conventional linear algebra have a max-plus version. Cuningham-Green [9], Gaubert [12], Gondran and Minoux [13] are all contributors to the development of max-plus linear algebra. Specifically we will consider matrix inverses, generalization of the determinant of a matrix, the solvability of linear systems such as  $A \otimes x = b$  and linear independence and dependence. We will also study the eigenvalue and eigenvector problem. The main question is whether these conventional linear algebra concepts have max-plus versions and if so, how they are similar and or different from the conventional algebra results.

### 2.1 Inverse Matrices

In conventional algebra we know that not all matrices have inverses. We will see that in max-plus algebra the invertible matrices are even more limited. First we need a few definitions.

**Definition 17.** *A matrix  $A \in \mathbb{R}_{\max}^{n \times n}$  is called invertible in the max-plus sense if there exists a matrix  $B$  such that  $A \otimes B = E$ , and we write  $A^{\otimes -1} = B$ .*

To be precise we should say “right inverse” in the above definition. But Theorem 3 below

will show that a right inverse is also a left inverse. Our immediate task is to identify the invertible matrices

**Definition 18.** A permutation matrix is a matrix in which each row and each column contains exactly one entry equal to  $e$  and all other entries are equal to  $\epsilon$ . If  $\sigma : \{1, 2, \dots, n\} \rightarrow \{1, 2, \dots, n\}$  is a permutation we define the max-plus permutation matrix  $P_\sigma = [p_{ij}]$  where

$$p_{ij} = \begin{cases} e & : i = \sigma(j) \\ \epsilon & : i \neq \sigma(j) \end{cases} .$$

So that the  $j^{th}$  column of  $P_\sigma$  has  $e$  in the  $\sigma(j)^{th}$  row.

Left multiplication by  $P_\sigma$  permutes the rows of a matrix, so that the  $i^{th}$  row of  $A$  appears as the  $\sigma(i)^{th}$  row of  $P_\sigma \otimes A$ .

**Definition 19.** If  $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{R}_{max}$ ,  $\lambda_i \neq \epsilon$  we define the diagonal matrix:

$$D(\lambda_i) = \begin{pmatrix} \lambda_1 & \epsilon & \dots & \epsilon \\ \epsilon & \lambda_2 & \epsilon \dots & \epsilon \\ \dots & \dots & \dots & \dots \\ \epsilon & \epsilon \dots & \epsilon & \lambda_n \end{pmatrix}$$

**Theorem 2.**  $A \in \mathbb{R}_{max}^{n \times n}$  has a right inverse if and only if there is a permutation  $\sigma$  and values  $\lambda_i > \epsilon$ ,  $i \in \{1, 2, \dots, n\}$  such that  $A = P_\sigma \otimes D(\lambda_i)$ .

*Proof.* Suppose there exists  $B$  such that  $A \otimes B = E$ . This implies that

- (1)  $\max_k(a_{ik} + b_{ki}) = e = 0$  for each  $i$
- (2)  $\max_k(a_{ik} + b_{kj}) = \epsilon = -\infty$  for all  $i \neq j$

By (1) for each  $i$  there exists a  $k$  so the  $a_{ik} + b_{ki} = e$ . Therefore we have a function  $k = \theta(i)$  with  $a_{i\theta(i)} > \epsilon$  and  $b_{\theta(i)i} > \epsilon$ . From (2) we find that



(3)  $a_{i\theta(j)} = \epsilon$  for all  $i \neq j$ .

Since  $a_{i\theta(i)} > \epsilon = a_{i\theta(j)}$  for  $i \neq j$ , it follows that  $\theta$  is an injection and therefore a permutation. (3) also tells us that  $a_{i\theta(i)}$  is the only entry of the  $\theta(i)^{th}$  column of  $A$  that is not  $\epsilon$ . Now let  $\tilde{A} = P_\theta \otimes A$ . The  $\theta(i)^{th}$  row of  $\tilde{A}$  is the  $i^{th}$  row of  $A$ , which has an entry greater than  $\epsilon$  in the  $\theta(i)^{th}$  column. Thus we have that all the diagonal entries of  $\tilde{A}$  are greater than  $\epsilon$ . We also have that  $A$  has only one non- $\epsilon$  entry in each column. It follows that this is also true of  $\tilde{A}$ . Thus

$$P_\theta \otimes A = \tilde{A} = D(\lambda_i) \text{ with } \lambda_i = a_{\theta^{-1}(i)i} > \epsilon.$$

Let  $\sigma = \theta^{-1}$ . Since  $P_\sigma \otimes P_\theta = E$ , it follows that

$$A = P_\sigma \otimes D(\lambda_i).$$

Now for the converse we assume that  $A = P_\sigma \otimes D(\lambda_i)$  with  $\lambda_i \in \mathbb{R}_{\max}$  and  $\lambda_i > \epsilon$ . If this is true then we let  $B = D(-\lambda_i) \otimes P_{\sigma^{-1}}$ . Note that  $-\lambda_i = \lambda_i^{\otimes -1}$ . So it follows that:

$$A \otimes B = P_\sigma \otimes D(\lambda_i) \otimes D(-\lambda_i) \otimes P_{\sigma^{-1}} = P_\sigma \otimes P_{\sigma^{-1}} = E$$

So  $A \otimes B = E$  and  $B$  is the right inverse of  $A$ . □

The previous theorem gives us a simple characterization of invertible matrices in the max-plus algebra. We now know that an invertible matrix is a permuted diagonal matrix.

**Theorem 3.** For  $A, B \in \mathbb{R}_{\max}^{n \times n}$  if  $A \otimes B = E$  then  $B \otimes A = E$ , and  $B$  is uniquely determined by  $A$ .

*Proof.* By Theorem 2 we know that  $A = P_\sigma \otimes D(\lambda_i)$  for some values  $\lambda_i > \epsilon$  and permutation  $\sigma$ . Observe that  $\tilde{B} = D(-\lambda_i) \otimes P_{\sigma^{-1}}$  is a left inverse of  $A$ . If  $A \otimes B = E$ , then  $\tilde{B} = \tilde{B} \otimes (A \otimes B) = (\tilde{B} \otimes A) \otimes B = E \otimes B = B$ , showing that  $B$  is uniquely determined and also a left inverse. □

**Lemma 7.** *If  $A \in \mathbb{R}_{\max}^{n \times n}$  and  $B \in \mathbb{R}_{\max}^{n \times n}$  are invertible then  $A \otimes B$  is invertible.*

*Proof.* By the previous theorems we can write

$$A = P_{\sigma_a} \otimes D(\lambda_i^a) \text{ and } B = D(\lambda_i^b) \otimes P_{\sigma_b}.$$

Then  $A \otimes B = P_{\sigma_a} \otimes D(\lambda_i^a) \otimes D(\lambda_i^b) \otimes P_{\sigma_b}$ .

Note that the product of two diagonal matrices is a diagonal matrix so we have that

$$A \otimes B = P_{\sigma_a} \otimes D(\lambda_i^a \otimes \lambda_i^b) \otimes P_{\sigma_b}.$$

So it follows that  $A \otimes B$  is a permuted diagonal matrix. Therefore  $A \otimes B$  is invertible.  $\square$

**Theorem 4.** *Suppose  $A \in \mathbb{R}_{\max}^{n \times n}$  and let  $L_A : \mathbb{R}_{\max}^n \rightarrow \mathbb{R}_{\max}^n$  be the linear map  $L_A(x) = A \otimes x$ . Then the following are equivalent.*

1.  $A = P_{\sigma} \otimes D(\lambda_i)$  for some permutation and  $\lambda_i > \epsilon$
2.  $L_A$  is surjective
3.  $A$  has a right inverse:  $A \otimes B = E$
4.  $A$  has a left inverse:  $B \otimes A = E$
5.  $L_A$  is injective

*Proof.* The proof that  $2 \Rightarrow 3$ ,  $1 \Rightarrow 2$ , and  $1 \Rightarrow 4 \Rightarrow 5$  are all elementary. We have already proven  $3 \iff 1$ . All we have left to prove is  $5 \Rightarrow 1$ .

So suppose that  $L_A$  is injective. For each  $i$  we define the sets

$$F_i = \{j : a_{ji} > \epsilon\} \text{ and } G_i = \{j : a_{jk} > \epsilon \text{ for some } k \neq i\}.$$

We claim that  $F_i \not\subseteq G_i$ . By contradiction suppose that  $F_i \subset G_i$ . We will show that this contradicts the fact that  $L_A$  is injective. Define  $x = [x_k]$  where  $x_k = \begin{cases} e & : k \neq i \\ \epsilon & : k = i \end{cases}$ . Let  $b = A \otimes x = \bigoplus_{k \neq i} a_{*k}$ , where  $a_{*k}$  denotes the  $k^{\text{th}}$  column of  $A$ .

Now suppose  $j \in F_i$ , then  $j \in G_i$ . This means that there exists  $k \neq i$  for which  $a_{jk} > \epsilon$ . Therefore we have  $b_j \geq a_{jk} > \epsilon$ . Since  $a_{ji} > \epsilon$ , then we can find  $\beta_j > \epsilon$  such that  $\beta_j \otimes a_{ji} \leq b_j$ . If  $j \notin F_i$  then  $a_{ji} = \epsilon$ . Therefore  $\beta \otimes a_{ji} \leq b_j$  for all  $j$ . Now let  $\beta = \min_{j \in F_i} \beta_j$ . Then  $\beta > 0$  and  $\beta \otimes a_{ji} \leq b_j$  for all  $j$ . This says that  $\beta \otimes a_{*i} \leq b$ . Then we have

$$\begin{aligned} A \otimes [x \oplus \beta \otimes e_i] &= [A \otimes x] \oplus [\beta \otimes A \otimes e_i] \\ &= b \oplus \beta \otimes a_{*i} = b. \end{aligned}$$

So for  $\tilde{x} = x \oplus \beta \otimes e_i$ ,  $L_A(\tilde{x}) = L_A(x)$ . But  $x_i = \epsilon < \tilde{x}_i = \beta$ , contradicting the injectivity of  $L_A$ . Thus our claim has been proved.

The claim means that for each  $i$  there is an index  $j = \sigma(i)$  with the property that  $a_{ji} > \epsilon$  but  $a_{jk} = \epsilon$  for all  $k \neq i$ . In other words  $a_{ji}$  is the only entry not equal to  $\epsilon$  in the  $j = \sigma(i)$  row. But then if  $j = \sigma(i')$  it follows that  $i = i'$ . In other words  $\sigma$  is injective. This means that  $\sigma$  is a permutation. Therefore for each row  $j$  there is a unique column  $i$  ( $j = \sigma(i)$ ) so that  $a_{ji}$  is the only entry not equal to  $\epsilon$ . For each column  $i$  and any row  $k$  with  $k \neq \sigma(i)$  we know that  $k = \sigma(i')$  for some  $i' \neq i$ . This means  $a_{ki}$  is not the unique non- $\epsilon$  entry in the  $k^{\text{th}}$  row, so  $a_{ki} = \epsilon$ . Therefore  $a_{\sigma(i)i}$  is the only non- $\epsilon$  entry in column  $i$ . Thus  $A$  is a permuted diagonal matrix,

$$A = P_\sigma \otimes D(\lambda_i), \lambda_i = a_{\sigma(i)i} > \epsilon$$

□

## 2.2 Determinants

In conventional algebra we know that for  $A \in \mathbb{R}^{n \times n}$   $\det(A) = \sum_{\sigma \in P_n} \text{sgn}(\sigma) \prod_{i=1}^n a_{i\sigma(i)}$ , where  $P_n$  denotes the set of all permutations of  $\{1, 2, \dots, n\}$  and  $\text{sgn}(\sigma)$  is the sign of the permutation  $\sigma$ . In max-plus algebra the determinant has no direct analogue because of the absence of additive inverses. Two related quantities, the permanent of  $A$  and the dominant of  $A$ , which are defined below, partially take over the role of the determinant. The permanent of  $A$  is defined similarly to the determinant but with the  $\text{sgn}(\sigma)$  simply omitted [25]. The following definitions of the permanent and dominant are from [25].

**Definition 20.** For a matrix  $A \in \mathbb{R}_{\max}^{n \times n}$  the permanent of  $A$  is defined to be  $\text{perm}(A) = \bigoplus_{\sigma \in P_n} \bigotimes_{i=1}^n (a_{i\sigma(i)})$ , with  $\sigma$  and  $P_n$  is the set of all permutations of  $\{1, 2, \dots, n\}$ .

**Lemma 8.** If  $A \in \mathbb{R}_{\max}^{n \times n}$  is invertible, then  $\text{perm}(A) \neq \epsilon$ .

*Proof.* In max-plus algebra an invertible matrix is a diagonal matrix times a permutation matrix. So if  $A$  is invertible the  $\text{perm}(A)$  is just the max-plus product of the diagonal entries of the diagonal matrix. Therefore if  $A$  is invertible  $\text{perm}(A) \neq \epsilon$ .  $\square$

Note that the  $\text{perm}(A) \neq \epsilon$  is not sufficient for  $A$  to be invertible. For example the matrix

$$A = \begin{pmatrix} 2 & 3 \\ 4 & 0 \end{pmatrix}$$

has  $\text{perm}(A) = \max\{2+0, 3+4\} = 7 \neq \epsilon$  but  $A$  is not invertible because it is not a permuted diagonal matrix.

The dominant is a refinement of the permanent and leads to results such as the max-plus version of Cramer's Rule [25]. These and other references use the matrix  $z^A$  to describe the dominant. Given  $A \in \mathbb{R}_{\max}^{n \times n}$   $z^A$  is the  $n \times n$  matrix  $z^A$  with entries  $z^{a_{ij}}$ , where  $z$  is a variable. The typical definition is

$$\text{dom}(A) = \begin{cases} \text{highest exponent in } \det(z^A) & \text{if } \det(z^A) \neq 0 \\ \epsilon & \text{if } \det(z^A) \equiv 0 \end{cases}.$$

In light of Lemma 1, we replace  $z$  by  $e^s$ , leading to the following two definition.

**Definition 21.** Given a matrix  $A \in \mathbb{R}_{\max}^{n \times n}$  the matrix  $e^{sA}$  has entries  $e^{sa_{ij}}$  where  $a_{ij} \in \mathbb{R}_{\max}$  are the entries in  $A$ .

Note that this does not refer to the matrix exponential. There is no analogue of the matrix exponential in this thesis. Here  $[e^{sA}]_{ij} = e^{sa_{ij}}$ .

**Definition 22.**  $\text{dom}(A) = \begin{cases} \lim_{s \rightarrow \infty} \frac{1}{s} \ln |\det(e^{sA})| & \text{if } \det(e^{sA}) \not\equiv 0 \\ \epsilon & \text{if } \det(e^{sA}) \equiv 0 \end{cases}$

In terms of Definition 1, this says that  $|\det(e^{sA})| \asymp e^{s \text{dom}(A)}$ . This asymptotic connection with the conventional determinants provides the basic approach to generalizing Cramer's Rule and the Cayley- Hamilton Theorem. (Note that a similar connection holds for permanent :  $\text{perm}(e^{sA}) \asymp e^{s \text{perm}(A)}$ ).

Since  $\text{perm}(A)$  is the maximum diagonal value for all permutations of the columns of  $A$ , we have the following lemma [25].

**Lemma 9.**  $\text{dom}(A) \leq \text{perm}(A)$

By the diagonal value we mean  $\bigotimes_i^n a_{i\sigma(i)}$  for some  $\sigma \in P_n$ . This is true when calculating the dominant we can have cancellations which will not occur when calculating the permanent. Note that due to the cancellations the  $\text{dom}(A)$  can be  $\epsilon$ . In order for it to be possible to have  $\text{perm}(A) = \epsilon$  each column of  $A$  must have at least one entry equal to  $\epsilon$ . For example take the matrix

$$A = \begin{pmatrix} 7 & 4 \\ 5 & 2 \end{pmatrix}.$$

We can see that  $\text{dom}(A) = \epsilon$  since  $\det(e^{sA}) = e^{s9} - e^{s9} = 0$ . But  $\text{perm}(A) = 9 \oplus 9 = 9$ .

**Lemma 10.** If  $A \in \mathbb{R}_{\max}^{n \times n}$  is invertible, then  $\text{dom}(A) \neq \epsilon$

*Proof.* Since  $A$  is invertible  $A$  is a permuted diagonal matrix. Therefore the  $\text{dom}(A)$  is equal to the max-plus product of the diagonal entries of the diagonal matrix. Therefore  $\text{dom}(A) \neq \epsilon$ .  $\square$

**Example 1.**  $A = \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix}$

We see that  $\text{dom}(A) = 5 > \epsilon$  but  $A$  is not invertible. In conventional algebra we know that  $A$  is invertible if and only if  $\det(A) \neq 0$ . So the max-plus version of determinants and invertible matrices is not completely analogous to the conventional case. With regards to invertibility this is not too disappointing because Theorem 2 provides such a simple characterization of invertible max-plus matrices.

Using the definitions of  $\text{dom}(A)$ ,  $\text{perm}(A)$  and invertibility we can also see that.

**Lemma 11.** *If  $A \in \mathbb{R}_{\max}^{n \times n}$  is invertible, then  $\text{dom}(A) = \text{perm}(A)$*

*Proof.* The proof follows from the proof of Lemma 8 and Lemma 10.  $\square$

In conventional algebra it is known that for  $A, B \in \mathbb{R}_{\max}^{n \times n}$   $\det(AB) = \det(A) \det(B)$ . However it is easy to exhibit examples for which  $\text{dom}(A \otimes B) \neq \text{dom}(A) \otimes \text{dom}(B)$ . This again is due to the fact that there are cancellations when computing the dominant. Examples that show  $\text{perm}(A \otimes B) \neq \text{perm}(A) \otimes \text{perm}(B)$  are also possible. For instance

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 3 & 1 & 2 \\ 2 & 1 & 1 \end{pmatrix}$$

and

$$B = \begin{pmatrix} 2 & 2 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

We can see that  $\text{perm}(A) = 6$  and  $\text{perm}(B) = 6$  but  $\text{perm}(A \otimes B) = 13$ . Also,  $\text{dom}(A) = 5$ ,  $\text{dom}(B) = 6$ , and  $\text{dom}(A \otimes B) = \epsilon$ .

However we can state the following lemma.

**Lemma 12.** *If  $A, B \in \mathbb{R}_{\max}^{n \times n}$  is invertible, then  $\text{dom}(A \otimes B) = \text{dom}(A) \otimes \text{dom}(B)$  and  $\text{perm}(A \otimes B) = \text{perm}(A) \otimes \text{perm}(B)$ .*

*Proof.* This follows from lemma 7 and 11 □

Gondran and Minoux also define what is called the bideterminant of  $A$ : [13], [14]. This is another approach to handling the absence of subtraction.

**Definition 23.** *For  $A \in \mathbb{R}_{\max}^{n \times n}$  let  $w_A(\sigma) = a_{1,\sigma(1)} \otimes a_{2,\sigma(2)} \otimes \dots \otimes a_{n,\sigma(n)}$ ,  $P_n^e$  be the set of even permutations and  $P_n^o$  be the set of odd permutations of  $\{1, 2, \dots, n\}$ . The bideterminant of  $A$  is  $\begin{pmatrix} \Delta_1(A) \\ \Delta_2(A) \end{pmatrix}$  is where  $\Delta_1(A) = \bigoplus_{\sigma \in P_n^e} w_A(\sigma)$  and  $\Delta_2(A) = \bigoplus_{\sigma \in P_n^o} w_A(\sigma)$*

Using the definition of the bideterminant we can write the definition of the permanent in a different form. Note that  $\text{perm}(A) = \Delta_1(A) \oplus \Delta_2(A)$ .

## 2.3 Linear Systems

In max-plus algebra the lack of additive inverses also causes difficulty when solving linear systems of equations such as  $A \otimes x = b$ . As in conventional algebra the solution to  $A \otimes x = b$  does not always exist in max-plus algebra and if it does it is not necessarily unique. We will explore other linear systems in max-plus algebra as well.

### 2.3.1 Principal Sub-Solution

First let us consider the case where  $A$  is a matrix which is not necessarily a square matrix. For the matrix  $A$  we can always find what is known as the principal or greatest sub-solution

to  $A \otimes x = b$  The principal sub-solution is the vector largest vector  $x$  such that  $A \otimes x \leq b$ . This sub-solution will be denoted by  $x^*(A, b)$ . The principle sub-solution is not necessarily a solution of  $A \otimes x = b$ . The following theorem is found in [9],[16] and [3].

**Theorem 5.** *Let  $A \in \mathbb{R}_{\max}^{m \times n}$  is an irreducible matrix and  $b \in \mathbb{R}_{\max}^m$ . Then  $[x^*(A, b)]_j = \min\{b_i - a_{ij} : i \in \{1, 2, \dots, m\} \text{ and } a_{ij} > \epsilon\}$ .*

*Proof.* First observe that  $A \otimes x \leq b$  is equivalent to each of the following:

1. for all  $i$  and  $j$ ,  $a_{ij} + x_j \leq b_i$
2. for all  $i$  and  $j$ ,  $x_j \leq b_i - a_{ij}$  or  $a_{ij} = \epsilon$
3. for all  $j$ ,  $x_j \leq \min\{b_i - a_{ij} : i \in \{1, 2, \dots, m\} \text{ and } a_{ij} > \epsilon\}$ .

This tells us that  $x$  is a solution to  $A \otimes x \leq b$  if and only if  $x_j \leq \min\{b_i - a_{ij} : i \in \{1, 2, \dots, m\}\}$  for all  $j$ . So  $x^*(A, B)_j = \min\{b_i - a_{ij} : i \in \{1, 2, \dots, m\} \text{ and } a_{ij} > \epsilon\}$  for all  $j$  is the maximum solution of  $A \otimes x \leq b$ . □

**Lemma 13.** *If a solution of  $A \otimes x = b$  exists, then the principle sub-solution is a solution.*

*Proof.* let  $x'$  be the maximum solution of  $A \otimes x = b$  then  $x$  satisfies the equation  $A \otimes x \leq b$ . So  $x'$  must be the principle sub-solution. Like wise the principle sub-solution  $x^*(A, b)$  is the maximum solution of  $A \otimes x \leq b$ . Which means  $x^*(A, b)$  must be the maximum solution of  $A \otimes x = b$  since we know a solution exists. Therefore the principle sub-solution is a solution. □

The following examples illustrate the above lemma.



1. Let  $A = \begin{pmatrix} 2 & 0 \\ 1 & 3 \end{pmatrix}$ ,  $b = \begin{pmatrix} 5 \\ 4 \end{pmatrix}$ . Using the previous theorem we see that the principle solution is  $x = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$ . We find that it is in fact a solution to  $A \otimes x = b$ .
2. Let  $A = \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix}$ ,  $b = \begin{pmatrix} 20 \\ 4 \end{pmatrix}$ . The principle solution is  $x = \begin{pmatrix} 3 \\ 0 \end{pmatrix}$ . In this case  $A \otimes x \neq b$  so there are no solutions.

Going back to our train example in the introduction, suppose that  $A$  is the matrix of travel times in the train network between stations and suppose that  $b$  is the vector containing the planned departure times of the trains for each station. Then  $x^*(A, b)$  gives the latest departure times of the trains from the previous station such that the times contained in  $b$  still can be met [16].

### 2.3.2 Cramer's Rule

In conventional matrix algebra, when  $A$  is a non-singular matrix, Cramer's Rule yields a solution to the linear equation  $Ax = b$ . The solution is given by:

$$x_i = \frac{\det(a_{*1}, \dots, a_{*i-1}, b, a_{*i+1}, \dots, a_{*n})}{\det(A)}, \quad i = 1, 2, \dots, n,$$

where  $a_{*j}$  denotes the  $j^{\text{th}}$  column of  $A$  and  $1 \leq j \leq n$ .

The max-plus analog to this formula developed in [25] is :

$$x_i \otimes \text{dom}(A) = \text{dom}(a_{*1}, \dots, a_{*i-1}, b, a_{*i+1}, \dots, a_{*n}).$$

Note that  $(a_{*1}, \dots, a_{*i-1}, b, a_{*i+1}, \dots, a_{*n})$  is the matrix  $A$  with its  $i^{\text{th}}$  column replaced by the vector  $b$ . Unlike the conventional case, however,  $\text{dom}(A) > \epsilon$  is not sufficient for this formula to produce a solution. An additional condition is needed. The extra condition is that

$\text{sign}(a_{*1}, \dots, a_{*i-1}, b, a_{*i+1}, \dots, a_{*n}) = \text{sign}(A)$  for all  $1 \leq i \leq n$ . Intuitively  $\text{sign}(A)$  is the sign of the coefficient in  $\det(e^{sA})$  that contributes to the  $\text{dom}(A)$ . To define  $\text{sign}(A)$  more precisely let  $P_n$  be the set of permutations  $\sigma : \{1, 2, \dots, n\} \rightarrow \{1, 2, \dots, n\}$  and let  $t_1, t_2, \dots, t_L$  be all possible values such that  $t_i = \bigotimes_1^n (a_{i\sigma(i)})$  for some  $\sigma \in P_n$ .

**Definition 24.** *Let*

$$\begin{aligned} S_i &= \{\sigma \in P_n : t_i = \bigotimes_1^n (a_{i\sigma(i)}) \text{ for some } \sigma \in P_n\} \\ S_{ie} &= \{\sigma \in S_i : \sigma \in P_n^e\} \\ S_{io} &= \{\sigma \in S_i : \sigma \in P_n^o\} \\ k_{ie} &= |S_{ie}| \text{ and } k_{io} = |S_{io}| \end{aligned}$$

Then we say  $\text{sign}(A) = 1$  if  $k_{ie} - k_{io} \geq 0$  and  $\text{sign}(A) = -1$  if  $k_{ie} - k_{io} \leq 0$ . If  $\text{dom}(A) = \epsilon$  then  $\text{sign}(A) = \epsilon$ .

Using the above definition we can write,  $\det(e^{sA}) = \sum_{i=1}^L (k_{ie} - k_{io})e^{st_i}$ . Observe that if  $\text{sign}(A) \neq \epsilon$ , then  $\text{sign}(A) \det(e^{sA}) > 0$  for all sufficiently large  $s$ .

**Theorem 6.** *If  $\text{sign}(a_{*1}, \dots, a_{*i-1}, b, a_{*i+1}, \dots, a_{*n}) = \text{sign}(A)$  for all  $i$  and  $\text{dom}(A) > \epsilon$  then*

$$x_i \otimes \text{dom}(A) = \text{dom}(a_{*1}, \dots, a_{*i-1}, b, a_{*i+1}, \dots, a_{*n}) \quad (2.1)$$

*yields a solution to  $A \otimes x = b$ .*

*Proof.* Assume that  $\text{dom}(A) > \epsilon$  and  $\text{sign}(a_{*1}, \dots, a_{*i-1}, b, a_{*i+1}, \dots, a_{*n}) = \text{sign}(A)$ . We consider the equation

$$e^{sA} \zeta = e^{sb}.$$

Since  $\text{dom}(A) > \epsilon$  then  $\det(e^{sA}) \neq 0$  so we can use Cramer's Rule to solve the equation above. Cramer's Rule yields the following:

$$\zeta_i(s) = \frac{\det(e^{sa_{*1}}, \dots, e^{sa_{*i-1}}, e^{sb}, e^{sa_{*i+1}}, \dots, e^{sa_{*n}})}{\det(e^{sA})}, \quad 1 \leq i \leq n.$$

Because of the hypothesis on the signs,  $\zeta_i(s) > 0$  for all sufficiently large  $s$ . Using Lemma 1 we see that

$$\lim_{s \rightarrow \infty} \frac{1}{s} \ln \zeta_i = (d_i - \text{dom}(A)), \quad 1 \leq i \leq n$$

where  $d_i = \text{dom}(a_{*1}, \dots, a_{*i-1}, b, a_{*i+1}, \dots, a_{*n})$ .

The value  $x_i = d_i - \text{dom}(A)$  are the unique solutions of (2.1). So what we have shown is that  $\zeta_i(s) \asymp e^{sx_i}$ . By substitution into  $e^{sA}\zeta(s) = e^{sb}$  we have

$$e^{sb} = e^{sA}\zeta(s) \asymp e^{sA}e^{sx} \asymp e^{s(A \otimes x)}.$$

By applying Lemma 1 we have shown that  $A \otimes x = b$ . □

Even if the assumptions of this theorem are not met, Cramer's Rule can sometimes still produce a solution. Consider the following example from [25]:

$$A = \begin{pmatrix} 1 & 3 & 5 \\ 2 & e & 1 \\ e & 4 & 2 \end{pmatrix}, \quad b = \begin{pmatrix} 8 \\ 4 \\ 6 \end{pmatrix}.$$

First observe that  $\text{dom}(A) = 11$ . This is because  $\det(e^{sA}) = e^{s(1+0+2)} + e^{s(3+1+0)} + e^{s(5+2+4)} - e^{s(0+0+5)} - e^{s(4+1+1)} - e^{s(2+2+3)}$ . This gives us that 11 is the highest exponent in  $\det(e^{sA})$ . We

can see that 11 is the max-plus product of the entries  $a_{13}, a_{32}$  and  $a_{21}$  of  $A$ . This is associated with the permutation  $(13)(32)(21)=(132)$  which is an even permutation. So  $\text{sign}(A) = 1$ .

Also note that  $e^{s(5+2+4)} = e^{s11}$  which has a positive coefficient in the  $\det(e^{sA})$ . Similarly we can find the  $\text{sign}(A_i)$  and  $\text{dom}(A_i)$  where  $A_i = (a_1, \dots, a_{i-1}, b, a_{i+1}, \dots, a_n)$ ,  $i = \{1, 2, 3\}$ . It turns

out that  $\text{dom}(A_1) = 10$ ,  $\text{dom}(A_2) = 13$ , and  $\text{dom}(A_3) = 14$ . The vector described by (2.1)

is  $x = \begin{pmatrix} -1 \\ 2 \\ 3 \end{pmatrix}$  which is in fact a solution. However note that in this case  $\text{sign}(A_1) = -1$

and  $\text{sign}(A_i) = 1$ ,  $i \in \{2, 3\}$ . So even though  $A$  and  $b$  do not meet the requirements of the previous theorem Cramer's Rule (2.1) still did give us a solution.

### 2.3.3 Solving $x = (A \otimes x) \oplus b$

Now we want to consider solving the linear equation  $x = (A \otimes x) \oplus b$ . We will see this type of equation in Chapter 3. Under certain constraints we can solve this linear system. The following is found in [16] and [3]. ( $A^*$  below is the same as defined in Definition 15 on pg 16.)

**Theorem 7.** *Let  $A \in \mathbb{R}_{\max}^{n \times n}$  and  $b \in \mathbb{R}_{\max}^n$ . If  $G(A)$  has maximal average circuit weight less than or equal to  $e$ , then the vector  $x = A^* \otimes b$  solves the equation  $x = (A \otimes x) \oplus b$ . If the circuit weights of  $G(A)$  are negative then the solution is unique.*

*Proof.* We will show that  $A^* \otimes b = A \otimes (A^* \otimes b) \oplus b$ . Using Definition 15 of  $A^*$ , and Lemma 5 on page 16 we know that since the maximal average circuit in  $G(A)$  is less than or equal to  $e$  then  $A^*$  exists. First observe that  $A^* = A \otimes A^* \oplus E$ . Therefore we have:

$$\begin{aligned} A^* \otimes b &= \bigoplus_{k \geq 0} A^{\otimes k} \otimes b \\ &= (\bigoplus_{k \geq 1} A^{\otimes k} \otimes b) \oplus (E \otimes b) \\ &= A \otimes (\bigoplus_{k \geq 0} A^{\otimes k} \otimes b) \oplus (E \otimes b) \\ &= A \otimes (A^* \otimes b) \oplus b. \end{aligned}$$

This shows that  $x = A^* \otimes b$  is a solution. To show uniqueness we assume that all circuits in  $G(A)$  are negative. Suppose that  $x$  is a solution of  $x = (A \otimes x) \oplus b$ . Now substitute this expression for  $x$  in for  $x$  on the right side. It follows that

$$x = b \oplus A \otimes [(A \otimes x) \oplus b] = b \oplus (A \otimes b) \oplus (A^{\otimes 2} \otimes x).$$

By repeating this argument we see that:

$$\begin{aligned} x &= b \oplus (A \otimes b) \oplus (A^{\otimes 2} \otimes b) \oplus (A^{\otimes 3} \otimes x) \\ &= b \oplus (A \otimes b) \oplus \dots \oplus (A^{\otimes (k-1)} \otimes b) \oplus (A^{\otimes k} \otimes x) \\ &= \bigoplus_{l=0}^{k-1} (A^{\otimes l} \otimes b) \oplus (A^{\otimes k} \otimes x). \end{aligned}$$

Using Lemma 6, Definition 15 and letting  $k \rightarrow \infty$  we get that:

$$\begin{aligned}
& \lim_{k \rightarrow \infty} \bigoplus_{l=0}^{k-1} (A^{\otimes l} \otimes b) \oplus (A^{\otimes k} \otimes x) \\
&= \lim_{k \rightarrow \infty} \bigoplus_{l=0}^{k-1} (A^{\otimes l} \otimes b) \oplus \lim_{k \rightarrow \infty} (A^{\otimes k} \otimes x) \\
&= \lim_{k \rightarrow \infty} (\bigoplus_{l=0}^{k-1} A^{\otimes l}) \otimes b \oplus \lim_{k \rightarrow \infty} (A^{\otimes k} \otimes x) \\
&= A^* \otimes b \oplus \varepsilon
\end{aligned}$$

So  $x = A^* \otimes b$  is the unique solution. □

## 2.4 Eigenvalue and Eigenvectors

Here we will study max-plus eigenvalues and eigenvectors. The max-plus eigenvalue and eigenvectors have a graph theoretical interpretation. The relevance of many of the theorems, lemmas and definitions from the graph theory section will be seen here. The significance of max-plus eigenvalues and eigenvectors will also be seen in Chapter 3. The results in this section are from [16] and [3].

**Definition 25.** Let  $A \in \mathbb{R}_{\max}^{n \times n}$  be a square matrix. If  $\mu \in \mathbb{R}_{\max}$  is a scalar,  $v \in \mathbb{R}_{\max}^n$  is a vector that contains at least one element not equal to  $\epsilon$ , and

$$A \otimes v = \mu \otimes v,$$

then  $\mu$  is called an eigenvalue of  $A$  and  $v$  is an associated eigenvector of  $A$ .

**Definition 26.** An eigenvector is called a finite eigenvector if it has all finite entries.

In general there can be more than one eigenvalue. The definition above also allows the eigenvalue to be equal to  $\epsilon$  and the eigenvector to have entries equal to  $\epsilon$ . Consider the following lemma.

**Lemma 14.**  $\epsilon$  is an eigenvalue of  $A$  if and only if  $A$  has a column of all  $\epsilon$  entries.

*Proof.* Let  $u$  be an eigenvector of  $A$  associated with eigenvalue  $\lambda = \epsilon$ . Define the set  $I = \{i : u_i > \epsilon\}$  then  $J = \{i : u_i = \epsilon\}$ . Since  $\lambda$  is an eigenvalue the following is true.

- For each,  $j \in J \bigoplus_{i \in I} a_{ji} \otimes u_i = \epsilon$ . Then  $a_{ji} = \epsilon$  for all  $j$ .
- For each,  $i \in I \bigoplus_{i' \in I} a_{ii'} \otimes u_{i'} = \epsilon$  Therefore  $a_{ii'} = \epsilon$  for all  $i$ .

Therefore  $A$  has a column of all  $\epsilon$ . □

As a result of the previous lemma we have the following corollary.

**Corollary 1.**  $\epsilon$  is not an eigenvalue of  $A$  if  $A$  is irreducible.

*Proof.* For each  $i$  there exists  $j$  such that  $a_{ji} > \epsilon$ . So the  $i^{\text{th}}$  column is not all  $\epsilon$ . □

**Example 2.**  $A = \begin{pmatrix} \epsilon & \epsilon & \epsilon \\ 1 & \epsilon & \epsilon \\ 2 & 3 & \epsilon \end{pmatrix}$ .

Then  $\begin{pmatrix} \epsilon & \epsilon & \epsilon \\ 1 & \epsilon & \epsilon \\ 2 & 3 & \epsilon \end{pmatrix} \otimes \begin{pmatrix} \epsilon \\ \epsilon \\ \alpha \end{pmatrix} = \lambda \begin{pmatrix} \epsilon \\ \epsilon \\ \alpha \end{pmatrix}$  where  $\alpha$  is a finite element of  $\mathbb{R}_{\max}$  and  $\lambda = \epsilon$ . In

this example  $\lambda$  is an eigenvalue and there are an infinite number of eigenvectors associated with this eigenvalue.

We will soon see an important case in which there is a unique eigenvalue, see Theorem 8. Eigenvectors certainly are not unique. Note that if  $v$  is an eigenvector of  $A$  then for any finite number  $\alpha$ ,  $\alpha \otimes v$  is also an eigenvector of  $A$ . The eigenvectors associated with a specific eigenvalue form a max-plus vector space called the eigenspace. The eigenspace will be described explicitly once we have some more background on eigenvalues and their associated eigenvectors.

In the max-plus algebra the following lemma states that any finite eigenvalue of a matrix is always the average weight of an elementary circuit in  $G(A)$ .

**Lemma 15.** Any finite eigenvalue  $\mu$  of a square matrix  $A$  is the average value of some elementary circuit in  $G(A)$ .

*Proof.* By definition an associated eigenvector  $v$  of  $\mu$  has at least one finite element. This means that there exists  $\nu_1$  such that  $v_{\nu_1} \neq \epsilon$ . Therefore  $[A \otimes v]_{\nu_1} = \mu \otimes v_{\nu_1} \neq \epsilon$ . Hence we can find a vertex  $\nu_2$  with  $a_{\nu_1 \nu_2} \otimes v_{\nu_2} = \mu \otimes v_{\nu_1}$ , implying that  $a_{\nu_1 \nu_2} \neq \epsilon$ ,  $v_{\nu_2} \neq \epsilon$  and  $(\nu_2, \nu_1) \in E(A)$ . Similarly there exists  $\nu_3$  such that  $a_{\nu_2 \nu_3} \otimes v_{\nu_3} = \mu \otimes v_{\nu_2}$  with  $(\nu_3, \nu_2) \in E(A)$ . Continuing in this fashion there exists a vertex  $\nu_k$  that we will encounter twice for the first time,  $\nu_k = \nu_{k+\ell}$ , since the number of vertices is finite. So we have found an elementary circuit,

$$c = ((\nu_k, \nu_{k+\ell-1}), (\nu_{k+\ell-1}, \nu_{k+\ell-2}), \dots, (\nu_{k+1}, \nu_k)).$$

This has length  $\|c\|_l = \ell$  and weight  $\|c\|_w = \bigotimes_{j=0}^{\ell-1} a_{\nu_{k+j} \nu_{k+j+1}}$ , where  $\nu_k = \nu_{k+\ell}$ . By construction we have

$$\bigotimes_{j=0}^{\ell-1} (a_{\nu_{k+j} \nu_{k+j+1}} \otimes v_{\nu_{k+j+1}}) = \mu^{\otimes \ell} \otimes \bigotimes_{j=0}^{\ell-1} v_{\nu_{k+j}}.$$

Since  $\otimes$  converts to  $+$  in conventional algebra the equation above can be written as

$$\sum_{j=0}^{\ell-1} (a_{\nu_{k+j} \nu_{k+j+1}} + v_{\nu_{k+j+1}}) = \ell \mu + \sum_{j=0}^{\ell-1} v_{\nu_{k+j}}.$$

We also have that

$$\sum_{j=0}^{\ell-1} v_{\nu_{k+j+1}} = \sum_{j=0}^{\ell-1} v_{\nu_{k+j}}$$

since  $\nu_k = \nu_{k+\ell}$ . Using this fact we can subtract  $\sum_{j=0}^{\ell-1} v_{\nu_{k+j}}$  from both sides giving us,

$$\bigotimes_{j=0}^{\ell-1} a_{\nu_{k+j} \nu_{k+j+1}} = \ell \times \mu.$$

This means that  $\|c\|_w = \ell \times \mu = \mu^{\otimes \ell}$ . Then the average value of the circuit  $c$  is

$$\frac{\|c\|_w}{\|c\|_l} = \frac{\mu^{\otimes \ell}}{\ell} = \mu.$$

□

**Lemma 16.** *Let  $C(A)$  denote the set of elementary circuits in  $G(A)$ . Then*

$\max_{p \in C(A)} \frac{\|p\|_w}{\|p\|_l}$  is the maximal average circuit.

The proof follows from the fact that every circuit is made up of elementary circuits. Recall Definition 12 of an elementary circuit.

**Definition 27.** A circuit  $p \in C(A)$  is called critical if its average weight is maximal.

**Definition 28.** The critical graph of  $A$  is denoted by  $G^c(A)$  and contains the vertices and edges belonging to the critical circuits of  $G(A)$ . The set of vertices in  $G^c(A)$  are denoted by  $V^c(A)$ .

The vertices in  $G^c(A)$  are often called the critical vertices. We now have the following definition from [1].

**Definition 29.** The critical classes of  $A$  are the strongly connected components of  $G^c(A)$

**Lemma 17.** If  $G(A)$  contains at least one circuit then any circuit in  $G^c(A)$  is critical.

*Proof.* Let  $\lambda$  be the maximal average circuit of  $A$ . Without loss of generality we can assume that  $\lambda = 0$ . Now suppose by contradiction that  $G^c(A)$  contains a circuit  $\rho$  with average value not equal to zero. Since  $\rho$  is a circuit in  $G^c(A)$  then  $\rho$  is a circuit in  $G(A)$ . If the average weight of  $\rho$  is larger than zero, then the maximal average circuit weight of  $A$  is larger than zero which contradicts the fact that  $\lambda = 0$ . Now consider the case when  $\rho$  is less than zero. Note that  $\rho$  is a path made up of sub-paths  $\rho_i$ , a sub-path of some critical circuit  $c_i$ ,  $i \in \{1, 2, \dots, k\}$ . Therefore there exists sub-paths  $\zeta_i$  such that  $c_i$  is made up of the two sub-paths  $\zeta_i$  and  $\rho_i$ . Since all circuits  $c_i$  have weight zero then the circuit made of the sub-paths  $\zeta_i$  is a circuit with weight greater than zero. This again contradicts the fact that  $\lambda = 0$ . Therefore every circuit in  $G^c(A)$  is a critical circuit.  $\square$

**Definition 30.** Let  $\lambda$  be a finite eigenvalue of  $A$ . The matrix  $A_\lambda$  is defined by  $[A_\lambda]_{ij} = a_{ij} - \lambda$



This matrix is called the normalized matrix and will be important in the proof of Lemma 19. Note that the eigenspace of  $A$  and  $A_\lambda$  coincide and  $e$  is an eigenvalue of  $A_\lambda$ . This is easy to see since

$$[\lambda \otimes v]_j = [A \otimes v]_j \text{ if and only if } v_j = [A \otimes v]_j - \lambda \text{ if and only if } e \otimes v_j = [A_\lambda \otimes v]_j.$$

Similarly we have the following lemma.

**Lemma 18.** *Let  $A \in \mathbb{R}_{\max}^{n \times n}$  be an irreducible matrix with finite eigenvalue  $\lambda$ , and let  $v$  be an associated eigenvector. Then,  $A_\lambda^*$  has eigenvalue  $e$  and  $v$  is an associated eigenvector.*

*Proof.* It can be shown that

$$(E \oplus A_\lambda) \otimes v = v.$$

It follows from Lemma 5 that

$$A_\lambda^* = (E \oplus A_\lambda)^{\otimes(n-1)}.$$

Therefore,

$$A_\lambda^* \otimes v = (E \oplus A_\lambda)^{\otimes(n-1)} \otimes v = v.$$

□

**Lemma 19.** *Let the communication graph  $G(A)$  of  $A \in \mathbb{R}_{\max}^{n \times n}$  have a finite maximal average circuit weight  $\lambda$ . Then  $\lambda$  is an eigenvalue of  $A$  and for any  $\nu \in V^c(A)$ , the column  $[A_\lambda^*]_\nu$  is an eigenvector of  $A$  associated with  $\lambda$ .*

*Proof.* Let  $\lambda$  be the maximal average circuit of  $G(A)$ . It is clear that the maximal average circuit of  $G(A_\lambda)$  is  $e = 0$ . Therefore, by Lemma 5 on pg 16,  $A_\lambda^+$  is well defined. As we saw

above the critical circuits of  $A$  and  $A_\lambda$  coincide. Likewise except for the weights of the edges, the graphs  $G^c(A)$  and  $G^c(A_\lambda)$  coincide. Therefore

$$\text{for any } \nu \in V^c(A), [A_\lambda^+]_{\nu\nu} = e, \quad (2.2)$$

since any critical circuit from  $\nu$  to  $\nu$  in  $G(A)$  has average weight  $\lambda$ .

Recalling Definition 15 and Lemma 5 on page 16, we see that

$$A_\lambda^* = E \oplus A^+.$$

It follows that

$$[A_\lambda^*]_{i\nu} = [E \oplus A_\lambda^+]_{i\nu} = \begin{cases} \varepsilon \oplus [A_\lambda^+]_{i\nu} & : \text{ for } i \neq \nu \\ e \oplus [A_\lambda^+]_{i\nu} & : \text{ for } i = \nu \end{cases}.$$

Using (2.2) this implies  $[A_\lambda^*]_\nu = [A_\lambda^+]_\nu$ . It is easy to see that  $A_\lambda^+ = A_\lambda \otimes A_\lambda^*$ . So we have that, for  $\nu \in V^c(A)$ ;

$$\begin{aligned} & A_\lambda \otimes [A_\lambda^*]_\nu \\ &= [A_\lambda \oplus A_\lambda^*]_\nu \\ &= [A_\lambda^+]_\nu \\ &= [A_\lambda^*]_\nu. \end{aligned}$$

$$\text{Therefore } A \otimes [A_\lambda^*]_\nu = \lambda \otimes A_\lambda \otimes [A_\lambda^*]_\nu = \lambda \otimes [A_\lambda^*]_\nu.$$

Thus  $\lambda$  is an eigenvalue of  $A$  and the  $\nu^{\text{th}}$  column of  $A_\lambda^*$  is an eigenvector of  $A$  for any  $\nu \in V^c(A)$ .  $\square$

**Definition 31.** *The columns of  $[A_\lambda^*]_\nu$  for  $\nu \in V^c(A)$  are called the critical columns of  $A$ .*

We have mentioned that eigenvalues and eigenvectors are not unique. However if  $A \in \mathbb{R}_{\max}^{n \times n}$  is an irreducible matrix then we have the following theorem

**Theorem 8.** *If  $A \in \mathbb{R}_{\max}^{n \times n}$  is an irreducible matrix the maximal average circuit weight is the unique eigenvalue.*

*Proof.* Assume  $A$  is irreducible and  $\lambda$  is the maximal average circuit in  $G(A)$ . If  $A$  is irreducible then  $G(A)$  contains at least one circuit so  $\lambda$  must be finite where  $\lambda$  is a eigenvalue by Lemma 15. Now we will show that the eigenvalue of  $A$  is unique if  $A$  is irreducible. First pick an elementary circuit  $c = ((i_1, i_2), (i_2, i_3), \dots, (i_\ell, i_{\ell+1}))$  in  $G(A)$  of length  $\|c\|_l = \ell$  and  $i_1 = i_{\ell+1}$ . So  $a_{i_{k+1}i_k} \neq \epsilon$  for  $k = \{1, 2, \dots, \ell\}$ . Let  $\mu$  be the eigenvalue of  $A$  and let  $v$  be an associated eigenvector. Since  $A$  is irreducible the  $\mu$  must be finite. By assumption we have  $\mu \otimes v = A \otimes v$  so it follows that

$$a_{i_{k+1}i_k} \otimes v_{i_k} \leq \mu \otimes v_{i_{k+1}} \text{ for } k \in \{1, 2, \dots, \ell\}.$$

Using the same argument as lemma 15 we have that

$$\frac{\|c\|_w}{\|c\|_l} \leq \frac{\mu^{\otimes \ell}}{\ell} = \mu.$$

This holds for any circuit in  $C(A)$ . So any finite eigenvalue  $\mu$  has to be larger or equal to the maximal average circuit  $\lambda$ . By lemma 15 we know that any finite eigenvalue is the average value of a circuit in  $G(A)$ . Therefore  $\lambda$  is a finite eigenvalue of  $A$  which is uniquely determined.

□

**Example 3.** *let*

$$A = \begin{pmatrix} 2 & 5 \\ 3 & 2 \end{pmatrix}.$$

*Then it is easy to see that  $\lambda = \frac{3+5}{2} = 4$  with critical circuit  $((1, 2), (2, 1))$ . It follows that*

$$A_\lambda^* = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

So  $\begin{pmatrix} 0 \\ -1 \end{pmatrix}$  and  $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$  are eigenvectors of  $A$  which are the critical columns of  $A$ .

Now consider the following counterexample to the uniqueness of the eigenvalue and eigenvector of a matrix. Note that the matrix  $A$  below is not irreducible.

**Example 4.**

$$A = \begin{pmatrix} 3 & \epsilon \\ \epsilon & 2 \end{pmatrix}$$

Then  $\begin{pmatrix} 3 & \epsilon \\ \epsilon & 2 \end{pmatrix} \otimes \begin{pmatrix} e \\ \epsilon \end{pmatrix} = 3 \otimes \begin{pmatrix} e \\ \epsilon \end{pmatrix}$  and  $\begin{pmatrix} 3 & \epsilon \\ \epsilon & 2 \end{pmatrix} \otimes \begin{pmatrix} \epsilon \\ e \end{pmatrix} = 2 \otimes \begin{pmatrix} \epsilon \\ e \end{pmatrix}$ .

**Lemma 20.** *If  $A$  is irreducible and  $v$  is an eigenvector, then  $v_i > \epsilon$  for all  $i$ .*

*Proof.* We know the eigenvalue  $\lambda$  of  $A$  is finite, since  $A$  is irreducible. By the definition of an eigenvector we know that  $v_j > \epsilon$  for some  $j$ . Since  $A$  is irreducible  $[A^{\otimes n}]_{ij} > \epsilon$  for some  $n$ . Since  $A^{\otimes n} \otimes v = \lambda^{\otimes n} \otimes v$  we know

$$[A^{\otimes n}]_{ij} \otimes v_j \leq v_i.$$

(See Lemma 24 on pg 54.) We know that all terms except  $v_i$  are finite, it follows that  $v_i$  must also be finite.  $\square$

The saturation graph of  $A$  is related to the critical graph of  $A$  and plays a role in the description of the eigenvectors of  $A$ . The following is from [16] and [1].

**Definition 32.** *Suppose  $\lambda$  is a finite eigenvalue of  $A$  and  $v$  is an associated eigenvector. The saturation graph denoted by  $S_{\lambda,v}(A)$  consists of the edges  $(j, i)$  such that  $a_{ij} \otimes v_j = \lambda \otimes v_i$  with  $v_i \neq \epsilon$  and  $v_j \neq \epsilon$ .*

Recall that a finite vector has all finite entries. As a result of the previous definition we have the following lemma.

**Lemma 21.** *Suppose  $A \in \mathbb{R}_{\max}^{n \times n}$  with finite eigenvalue  $\lambda$  and associated finite eigenvector  $v$*

- (1) *For each vertex  $i \in S_{\lambda,v}(A)$  there exists a circuit in  $S_{\lambda,v}(A)$  from which  $i$  can be reached in a finite number of steps.*
- (2) *All circuits of  $S_{\lambda,v}(A)$  have average circuit weight  $\lambda$ .*
- (3) *If  $A$  is irreducible, then circuits of  $S_{\lambda,v}(A)$  are critical.*
- (4) *If  $A$  is irreducible,  $G^c(A) \subseteq S_{\lambda,v}(A)$ .*

*Proof.* For (1) let  $i$  be a vertex in  $S_{\lambda,v}(A)$ . Then there exists a vertex  $j \in S_{\lambda,v}(A)$  such that  $\lambda \otimes v_i = a_{ij} \otimes v_j$  with  $v_i, v_j \neq \epsilon$ . Repeating this we can find a vertex  $k$  such that  $\lambda \otimes v_j = a_{jk} \otimes v_k$  with  $v_j, v_k \neq \epsilon$ . We can repeat this an arbitrary number of times. Assume we continue this  $m$  times, then we get a path of length  $m$  in  $S_{\lambda,v}(A)$ . If  $n < m$  then the path we constructed contains an elementary circuit. Thus our proof of (1) is complete.

For (2) Consider the elementary circuit  $c \in S_{\lambda,v}(A)$  where  $c = ((i_1, i_2), (i_2, i_3), \dots, (i_\ell, i_{\ell+1} = i_1))$ . By the definition of a saturation graph we have

$$\lambda \otimes v_{i_{k+1}} = a_{i_{k+1}i_k} \otimes v_{i_k}, \quad 1 \leq k \leq \ell.$$

This implies that

$$\lambda^{\otimes \ell} \otimes v_{i_1} = \bigotimes_{k=1}^{\ell} a_{i_{k+1}i_k} \otimes v_{i_1}.$$

Therefore we have  $\lambda^{\otimes \ell} = \bigotimes_{k=1}^{\ell} a_{i_{k+1}i_k}$ . The expression on the right-hand side is the weight of circuit  $c$ . Therefore  $c$  has average weight  $\lambda$ .

For (3) we recall that since  $A$  is irreducible it has a unique eigenvalue  $\lambda$  which is the maximum average circuit weight. By part (2) we know that any circuit in the saturation graph has

average circuit weight  $\lambda$ . Therefore the average circuit weight of any circuit in the saturation graph is the maximal average circuit weight of  $G(A)$ , so it belongs to the critical graph of  $A$ .

For (4) since  $A$  is irreducible, then  $A$  has a unique finite eigenvalue  $\lambda$ . The irreducibility of  $A$  also implies that  $v_i > \epsilon$  for all the entries of the eigenvector  $v$ . (See the previous lemma.)

Consider any edge  $(i, j) \in G^c(A)$ . We want to show that

$$a_{ji} \otimes v_i = \lambda \otimes v_j, \quad (2.3)$$

which would imply that  $(i, j) \in S_{\lambda, v}(A)$ . Since  $A \otimes v = \lambda \otimes v$  we know  $a_{ji} \otimes v_i \leq \lambda \otimes v_j$ . Suppose that the inequality was strict,

$$a_{ji} \otimes v_i < \lambda \otimes v_j. \quad (2.4)$$

Since  $(i, j) \in G^c(A)$  there is a critical circuit  $c : ((i_1, i_2), (i_2, i_3), \dots, (i_\ell, i_{\ell+1} = i_1))$  with  $i_1 = i$  and  $i_2 = j$ . Then for each  $k = 1, 2, \dots, \ell$  we have

$$a_{i_{k+1}i_k} \otimes v_k \leq \lambda \otimes v_{k+1}$$

and for the first ( $k = 1$ ) the inequality is strict. Summing around the circuit gives us,

$$\bigotimes_{k=1}^{\ell} a_{i_{k+1}i_k} \otimes \bigotimes_{k=1}^{\ell} v_{i_k} < \lambda^{\otimes \ell} \otimes \bigotimes_{k=1}^{\ell} v_{i_{k+1}} = \lambda^{\otimes \ell} \otimes \bigotimes_{k=1}^{\ell} v_{i_k}.$$

The inequality is strict because of (2.4). Since  $v_i > \epsilon$  it follows that

$$\bigotimes_{k=1}^{\ell} a_{i_{k+1}i_k} < \lambda^{\otimes \ell}.$$

But this means the average weight of out circuit  $c$  is less than  $\lambda$ , which is contrary to the assumption that  $c$  is critical. Therefore  $a_{ji} \otimes v_i = \lambda \otimes v_j$  so  $(i, j) \in S_{\lambda, v}(A)$ . Thus  $G^c(A) \subseteq S_{\lambda, v}(A)$ .

□

As previously mentioned the eigenvectors of  $A$  form a max-plus vector space called the eigenspace of  $A$  which is the set of all eigenvectors of  $A$ . Let  $A$  have finite eigenvalue  $\lambda$ . We will denote the eigenspace of  $A$  by  $V_\lambda(A)$ .

**Theorem 9.** *If  $A$  is an irreducible matrix then the eigenspace of  $A$  is given by:*

$$V_\lambda(A) = \{v \in \mathbb{R}_{\max}^n : v = \bigoplus_{i \in V^c(A)} a_i \otimes [A_\lambda^*]_i \text{ for } a_i \in \mathbb{R}_{\max}\}$$

*Proof.* First we know from Lemma 19 that  $[A_\lambda^*]_i$  for all  $i \in V^c(A)$  are eigenvectors of  $A$ . Since any max-plus linear combination of eigenvectors is an eigenvector, it follows that

$$\bigoplus_{i \in V^c(A)} a_i \otimes [A_\lambda^*]_i$$

is an eigenvector where  $a_i \in \mathbb{R}_{\max}$  and at least one  $a_i \neq \epsilon$ . For the converse we will show that any eigenvector can be written as a linear combination of the critical columns of  $A$ . Since  $A$  has eigenvalue  $\lambda$  with associated eigenvector  $v$  then  $A_\lambda^*$  has eigenvalue  $e$  and eigenvector  $v$ . Now consider vertices  $i$  and  $j$  in the saturation graph  $S_{e,v}(A_\lambda)$  such that there exists a path from  $i$  to  $j$ . By Lemma 21, for each  $j$  there exists  $i$  that belongs to a critical circuit. Now let  $((i_1, i_2), (i_2, i_3), \dots, (i_\ell, i_{\ell+1}))$ , with  $i = i_1$  and  $j = i_{\ell+1}$ . By the definition of the saturation graph we have

$$[A_\lambda]_{i_{k+1}i_k} \otimes v_{i_k} = v_{i_{k+1}}, \quad 1 \leq k \leq \ell.$$

Hence

$$v_j = \alpha \otimes v_i,$$

$$\text{where } \alpha = \bigotimes_{k=1}^{\ell} [A_\lambda]_{i_{k+1}i_k} \leq [A_\lambda^{\otimes \ell}]_{ji} \leq [A_\lambda^*]_{ji}. \quad (2.5)$$

Using the fact that  $v_j = \alpha \otimes v_i$  for any vertex  $\eta$  we have

$$\begin{aligned}
[A_\lambda^*]_{\eta j} \otimes v_j &= [A_\lambda^*]_{\eta j} \otimes \alpha \otimes v_i \\
&\leq [A_\lambda^*]_{\eta j} \otimes [A_\lambda^*]_{j i} \otimes v_i \text{ ( by (2.5))} \\
&\leq [A_\lambda^*]_{\eta i} \otimes v_i
\end{aligned} \tag{2.6}$$

The last inequality follows from the fact that  $A_\lambda^* \otimes A_\lambda^* = A_\lambda^*$ . (2.6) gives us that

$$\bigoplus_{j \in S_{e,v}(A_\lambda)} [A_\lambda^*]_{\eta j} \otimes v_j \leq \bigoplus_{i \in V^c(A_\lambda)} [A_\lambda^*]_{\eta i} \otimes v_i. \tag{2.7}$$

Since  $v$  is the eigenvector of  $A_\lambda$  associated with eigenvalue  $e$ , we know  $v = A_\lambda^* \otimes v$ . This means that  $v_\eta = [A_\lambda^*]_{\eta j} \otimes v_j$  for some  $j$  in the saturation graph. We don't know which  $j$  but since  $j$  is in the saturation graph it holds that

$$v_\eta = \bigoplus_{j \in S_{e,v}(A_\lambda)} [A_\lambda^*]_{\eta j} \otimes v_j \leq \bigoplus_{i \in V^c(A_\lambda)} [A_\lambda^*]_{\eta i} \otimes v_i \text{ (by (2.7)).}$$

Therefore

$$v_\eta \leq \bigotimes_{i \in V^c(A_\lambda)} [A_\lambda^*]_{\eta i} \otimes v_i.$$

Conversely since  $v$  is an eigenvector of  $A_\lambda^*$  for eigenvalue  $e$ ,

$$v_\eta = [A_\lambda^* \otimes v]_\eta = \bigoplus_{j=1}^n [A_\lambda^*]_{\eta j} \otimes v_j \geq \bigoplus_{i \in V^c(A_\lambda)} [A_\lambda^*]_{\eta i} \otimes v_i.$$

Hence  $v_\eta = \bigoplus_{i \in V^c(A_\lambda)} [A_\lambda^*]_{\eta i} \otimes v_i$ .

□

**Lemma 22.** *Let  $A$  be an irreducible matrix. For vertices  $i$  and  $j$  which belong to the critical graph of  $A$ , there exists  $a \in \mathbb{R}$  such that*

$$a \otimes [A_\lambda^*]_{*i} = [A_\lambda^*]_{*j}$$

*if and only if  $i$  and  $j$  belong to the same critical class.*



In other words all columns from the same critical class are scalar multiples of each other.

*Proof.* First assume that  $i$  and  $j$  belong to the same critical class of  $A_\lambda$  (recall that the critical classes of  $A$  and  $A_\lambda$  coincide). It follows that  $[A_\lambda^*]_{ji} \otimes [A_\lambda^*]_{ij} = e$ . Therefore for all  $\ell$ ,

$$\begin{aligned} [A_\lambda^*]_{\ell i} \otimes [A_\lambda^*]_{ij} &\leq [A_\lambda^*]_{\ell j} \\ &= [A_\lambda^*]_{\ell j} \otimes [A_\lambda^*]_{ji} \otimes [A_\lambda^*]_{ij} \quad (\text{since } [A_\lambda^*]_{ji} \otimes [A_\lambda^*]_{ij} = e) \\ &\leq [A_\lambda^*]_{\ell i} \otimes [A_\lambda^*]_{ij}. \end{aligned}$$

Which tells us that

$$[A_\lambda^*]_{\ell i} \otimes [A_\lambda^*]_{ij} \leq [A_\lambda^*]_{\ell j} \leq [A_\lambda^*]_{\ell i} \otimes [A_\lambda^*]_{ij},$$

so  $[A_\lambda^*]_{\ell i} \otimes [A_\lambda^*]_{ij} = [A_\lambda^*]_{\ell j}$ . By letting  $a = [A_\lambda^*]_{ij}$  we have shown that  $a \otimes [A_\lambda^*]_{*i} = [A_\lambda^*]_{*j}$ .

Now we want to show that if  $i$  and  $j$  do not belong to the same critical class then  $a \otimes [A_\lambda^*]_{*i} = [A_\lambda^*]_{*j}$  does not hold. So suppose that  $i$  and  $j$  do not belong to the same critical class but that  $a \otimes [A_\lambda^*]_{*i} = [A_\lambda^*]_{*j}$ . It follows that

$$a \otimes e = [A_\lambda^*]_{ij} \quad \text{and} \quad a \otimes [A_\lambda^*]_{ji} = e$$

so that  $[A_\lambda^*]_{ji} \otimes [A_\lambda^*]_{ij} = e$ . Therefore the circuits formed by the edges  $(i, j)$  and  $(j, i)$  have average weight  $e$  where  $i$  and  $j$  belong to the critical graph of  $A$ . So they belong to the same critical class which is a contradiction.  $\square$

In general the eigenvectors of  $A \in \mathbb{R}_{\max}^{n \times n}$  are not unique. However we will consider a situation when the eigenvectors are unique (up to scalar multiples).

**Definition 33.** For  $A \in \mathbb{R}_{\max}^{n \times n}$  the eigenvector of  $A$  is unique if for any two eigenvectors  $v, w \in \mathbb{R}_{\max}^n$   $v = \alpha \otimes w$  for some  $\alpha \in \mathbb{R}$

As result of the previous lemma we have:

**Lemma 23.** *If  $A \in \mathbb{R}_{\max}^{n \times n}$  is irreducible the  $A$  has a unique critical class if and only if the eigenvector of  $A$  is unique.*

*Proof.* First assume  $A$  has a unique critical class. By Lemma 22 we know if  $i$  and  $j$  belong to the same critical class then  $a \otimes [A_\lambda^*]_{*i} = [A_\lambda^*]_{*j}$ . So since  $A$  has only one critical class then the eigenvector is unique. For the other direction if  $v$  is a unique eigenvector then  $a \otimes [A_\lambda^*]_{*i} = [A_\lambda^*]_{*j}$  for all critical vertices. By the previous theorem these vertices belong to the same critical class. Hence  $A$  has a unique critical class.  $\square$

There are several algorithms for computing the eigenvalue of an irreducible matrix such as Karp's algorithm and the power algorithm; see [16] and [5]. These algorithms will not be described here.

For reducible matrices the eigenvalue and eigenvectors of  $A$  are not necessarily finite and the eigenvalue may not be unique. Thus the eigenspace is much more complicated than in the irreducible case. We will not discuss the case when  $A$  is reducible here. See [16] for more discussion on this matter.

## 2.5 Cayley-Hamilton and the Max-Plus Characteristic Equation

Recall that in conventional linear algebra the Cayley-Hamilton Theorem states that every square matrix  $A$  satisfies its own characteristic equation. The characteristic equation in conventional algebra is used to solve for the eigenvalue of a matrix. We begin with the following theorem which is the conventional version of the Cayley-Hamilton Theorem. Let  $C_n^k$  be the set of all subsets of  $k$  elements of the set  $\{1, 2, \dots, n\}$ . If  $A$  is a  $n \times n$  matrix and  $\varphi \subset \{1, 2, \dots, n\}$ , the submatrix obtained by removing all rows and columns of  $A$  except those denoted by  $\varphi$  is denoted by  $A_{\varphi\varphi}$ .

**Theorem 10.** Suppose  $A \in \mathbb{R}^{n \times n}$ , if

$$\begin{aligned} \det(\lambda I - A) &= \lambda^n + c_1 \lambda^{n-1} + \dots + c_{n-1} \lambda + c_n \text{ then} \\ A^n + c_1 A^{n-1} + \dots + c_{n-1} A + c_n I &= 0 \end{aligned}$$

where the  $c_k$  are given by  $c_k = (-1)^k \sum_{\sigma \in C_n^k} \det(A_{\sigma\sigma})$ .

We will show how the Cayley- Hamilton Theorem can be translated into max-plus algebra. These results are found in [25] and [6]. Again we consider the matrix  $e^{sA}$ . We also need a few more definitions. First we have that the characteristic polynomial of the matrix- valued function  $e^{sA}$  is given by:

$$\det(\lambda I - e^{sA}) = \lambda^n + \gamma_1(s) \lambda^{n-1} + \dots + \gamma_{n-1}(s) \lambda + \gamma_n(s) \quad (2.8)$$

with coefficients

$$\gamma_k(s) = (-1)^k \sum_{\varphi \in C_n^k} \det(e^{sA_{\varphi\varphi}}).$$

Therefore

$$(e^{sA})^n + \gamma_1(s)(e^{sA})^{n-1} + \dots + \gamma_{n-1}(s)e^{sA} + \gamma_n I = 0. \quad (2.9)$$

This is just the result of applying Theorem 10 to the matrix  $e^{sA}$ . Next recall that  $P_n$  represents the set of permutations  $\sigma : \{1, 2, \dots, n\} \rightarrow \{1, 2, \dots, n\}$  where  $P_n^e$  represents the even permutations and  $P_n^o$  represents the odd permutations.

Now we define  $\Gamma_k = \{\zeta : \exists \{i_1, i_2, \dots, i_k\} \in C_n^k, \exists \rho \in P_k \text{ such that } \zeta = \sum_{r=1}^k a_{i_r i_{\rho(r)}}\}$  for  $k = 1, 2, \dots, n$ . These values are the values  $\zeta$  for which  $e^{s\zeta}$  occurs in  $\gamma_k(s)$ . To describe the coefficients of  $e^{s\zeta}$  in  $\gamma_k(s)$  we define the following. For every  $k \in \{1, 2, \dots, n\}$  and for every  $\zeta \in \Gamma_k$  we define the values:

$I_k^e(\zeta)$  is the number of  $\rho \in P_k^e$  such that there exists  $\{i_1, i_2, \dots, i_k\} \in C_n^k$  and  $\zeta = \sum_r a_{i_r i_{\rho(r)}}$   
 $I_k^o(\zeta)$  is the number of  $\rho \in P_k^o$  such that there exists  $\{i_1, i_2, \dots, i_k\} \in C_n^k$  and  $\zeta = \sum_r a_{i_r i_{\rho(r)}}$ .

Then  $I_k(\zeta) = I_k^e(\zeta) - I_k^o(\zeta)$ .

Now we can write

$$\gamma_k(s) = (-1)^k \sum_{\zeta \in \Gamma_k} I_k(s) e^{s\zeta}. \quad (2.10)$$

The dominant term in  $\gamma_k(s)$  is given by  $e^{d_k}$  where

$$d_k = \max\{\zeta \in \Gamma_k : I_k(\zeta) \neq 0\}.$$

Applying Lemma 1 to (2.10) we have

$$|\gamma_k(s)| \asymp e^{sd_k}.$$

Now define the leading coefficients of  $\gamma_k(s)$ :  $\tilde{\gamma}_k = (-1)^k I(d_k)$  for  $k = 1, 2, \dots, n$ . In brief  $\gamma_k(s) = \tilde{\gamma}_k e^{sd_k} + (\text{lower order terms})$ . We need to separate the terms ( $k$ ) with positive  $\tilde{\gamma}_k$  from those with negative  $\tilde{\gamma}_k$ . Let  $\ell = \{k : \tilde{\gamma}_k > 0\}$  and  $j = \{k : \tilde{\gamma}_k < 0\}$ . For every  $\zeta \in \Gamma_1 = \{a_{ii} : i = 1, 2, \dots, n\}$ , we have that  $I_1^o(\zeta) = 0$  and  $I_1^e(\zeta) > 0$ . This implies that  $I_1(d_1) > 0$  and that  $\tilde{\gamma}_1 < 0$ . Therefore we always have  $1 \in j$ .

The max-plus characteristic equation of  $A$  is also found by rearranging (2.8) and moving the  $\gamma_k(s)\lambda^{n-k}$  for which  $\gamma_k(s)$  has a negative leading coefficient to the right side, replace  $\lambda$  with  $e^{s\lambda}$ , and then applying Lemma 1. The reason for doing this is that subtraction is not defined in the max-plus algebra. We state this in the following definition.

**Definition 34.** *The max-plus characteristic equation of  $A \in \mathbb{R}_{\max}^{n \times n}$  is defined as*

$$\lambda^{\otimes n} \oplus \bigoplus_{k \in \ell} d_k \otimes \lambda^{\otimes n-k} = \bigoplus_{k \in j} d_k \otimes \lambda^{\otimes n-k} \quad (2.11)$$

Now we can state and prove the following theorem

**Theorem 11.**  *$A \in \mathbb{R}_{\max}^{n \times n}$ . The  $A$  satisfies its own characteristic equation:*

$$A^{\otimes n} \oplus \bigoplus_{k \in \ell} d_k \otimes A^{\otimes n-k} = \bigoplus_{k \in j} d_k \otimes A^{\otimes n-k}. \quad (2.12)$$

*Proof.* We continue with the application of the conventional Cayley-Hamilton Theorem to the matrix  $e^{sA}$ . If  $A \in \mathbb{R}_{\max}^{n \times n}$  then it is easy to see that

$$(e^{sA})^k \asymp e^{sA^{\otimes k}}.$$

Now since  $\gamma_k(s) \asymp \tilde{\gamma}_k e^{sd_k}$  so after rearrangement (2.9) becomes

$$e^{sA^{\otimes n}} + \sum_{k \in \ell} \tilde{\gamma}_k(s) e^{sd_k} e^{sA^{\otimes n-k}} \asymp - \sum_{k \in j} \tilde{\gamma}_k(s) e^{sd_k} e^{sA^{\otimes n-k}}. \quad (2.13)$$

Using Lemma 1 we get the following expression in  $\mathbb{R}_{\max}$ ,

$$A^{\otimes n} \oplus \bigoplus_{k \in \ell} d_k \otimes A^{\otimes n-k} = \bigoplus_{k \in j} d_k \otimes A^{\otimes n-k}. \quad (2.14)$$

□

Let us now consider an example from [6].

**Example 5.** *Let*

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

*Using the definitions above we find that  $\Gamma_1 = \{2, 0, -2\}$ ,  $\Gamma_2 = \{2, 1, 0, -2, \epsilon\}$  and  $\Gamma_3 = \{4, 0, -1, \epsilon\}$ . Now we see that*

$$\begin{aligned} I_1(2) &= 1, & I_1(0) &= 1, & I_1(-2) &= 1 \\ I_2(2) &= 0, & I_2(1) &= -1, & I_2(0) &= 1, & I_2(-2) &= 1, & I_2(\epsilon) &= -1 \\ I_3(4) &= -1, & I_3(0) &= 1, & I_3(-1) &= -1, & I_3(\epsilon) &= 1. \end{aligned}$$

*Hence  $d_1 = 2$ ,  $d_2 = 1$  and  $d_3 = 4$ . Note that the maximum value in  $\Gamma_2$  is 2 however  $I_2(2) = 0$  So an even and an odd permutation gives us the diagonal value 2 which means the*

two permutations cancel each other out so  $d_2 = 1$ .  $\tilde{\gamma}_1 = -1$ ,  $\tilde{\gamma}_2 = -1$  and  $\tilde{\gamma}_3 = 1$  so  $\ell = 3$  and  $j = \{1, 2\}$ . Therefore the max-plus characteristic equation of  $A$  is

$$\lambda^{\otimes 3} \oplus 4 = 2 \otimes \lambda^{\otimes 2} \oplus 1 \otimes \lambda \quad (2.15)$$

and  $A$  satisfies its max plus characteristic equation with

$$A^{\otimes 3} \oplus 4 \otimes E = 2 \otimes A^{\otimes 2} \oplus 1 \otimes A = \begin{pmatrix} 4 & 3 & 4 \\ 3 & 4 & 5 \\ 3 & 4 & 6 \end{pmatrix}. \quad (2.16)$$

In [25] Olsder and Roos described the max-plus characteristic equation, but had a subtle flaw. De Schutter and De Moor [6] found a counterexample, and were able to fix Olsder and Roos's mistake. The correction will be summarized here. In [25] Olsder and Roos claimed that the highest exponential factor in  $\gamma_k(s)$  is equal to the  $\max\{\text{dom}(A_{\sigma\sigma}) : \sigma \in C_n^k\}$  rather than  $d_k$  and that

$$\gamma_k(s) = (-1)^k \tilde{\gamma}_k e^{s \max\{\text{dom}(A_{\sigma\sigma}) : \sigma \in C_n^k\}}.$$

However this is not necessarily true because the even and odd permutations may cancel, eliminating that exponential factor. Our example  $A$  above is the counterexample cited by De Schutter and De Moor. Consider  $\gamma_k(s)$ . First we have that

$$\begin{aligned} \det(e^{sA_{\{1,2\}\{1,2\}}}) &= e^{s(-2)} - e^{s2} \\ \det(e^{sA_{\{1,3\}\{1,3\}}}) &= e^{s0} - e^{s(-\infty)} = 1 \\ \det(e^{sA_{\{2,3\}\{2,3\}}}) &= e^{s2} - e^{s1}. \end{aligned}$$

$\text{dom}(A_{\{1,2\}\{1,2\}}) = 2$ ,  $\text{dom}(A_{\{1,3\}\{1,3\}}) = 0$  and  $\text{dom}(A_{\{2,3\}\{2,3\}}) = 2$ , so  $\max\{\text{dom}(A_{\sigma\sigma}) : \sigma \in C_3^2\} = 2$ . However,

$$\begin{aligned} \gamma_2(s) &= (-1)^2 (\det(e^{sA_{\{1,2\}\{1,2\}}}) + \det(e^{sA_{\{1,3\}\{1,3\}}}) + \det(e^{sA_{\{2,3\}\{2,3\}}})) \\ &= e^{s(-2)} - e^{s2} + e^{s0} - e^{s(-\infty)} + e^{s2} - e^{s1} \\ &= e^{s(-2)} + e^{s0} - e^{s(-\infty)} - e^{s1} \end{aligned}$$

and the highest degree in  $\gamma_2(s)$  is 1.

Now consider the following theorem.

**Theorem 12.** *The eigenvalues of  $A \in \mathbb{R}_{\max}^{n \times n}$  satisfy the characteristic equation of  $A$*

*Proof.* Suppose  $v$  is an eigenvalue associated with the eigenvalue  $\lambda$  of  $A$ . Now consider (2.14).

We can multiply  $v$  on both sides of that equation giving us

$$\lambda^{\otimes n} \otimes v \oplus \bigoplus_{k \in \ell} d_k \otimes \lambda^{\otimes n-k} \otimes v = \bigoplus_{k \in j} d_k \otimes \lambda^{\otimes n-k} \otimes v.$$

(See Lemma 24 on pg 54.) Now we can subtract (in the conventional sense)  $v$  from both sides of the equation which gives us

$$\lambda^{\otimes n} \oplus \bigoplus_{k \in \ell} d_k \otimes \lambda^{\otimes n-k} = \bigoplus_{k \in j} d_k \otimes \lambda^{\otimes n-k}.$$

Thus our proof is complete. □

If  $\rho$  is a root of the characteristic equation of  $A$  then  $\rho$  is not necessarily an eigenvalue of  $A$ .

To see this consider the matrix

$$A = \begin{pmatrix} 2 & 7 \\ -3 & 4 \end{pmatrix}.$$

Using the formulas we have defined for the max-plus characteristic equation of  $A$  we see that:

$$\Gamma_1 = \{2, 4\} \text{ and } \Gamma_2 = \{6, 4\} \text{ with } I_1(2) = 1, I_1(4) = 1, I_2(6) = 1 \text{ and } I_2(4) = -1.$$

Therefore  $d_1 = 4$ ,  $d_2 = 6$ ,  $\tilde{\gamma}_1 = -1$  and  $\tilde{\gamma}_2 = 1$  with  $\ell = \{2\}$  and  $j = \{1\}$ . So the characteristic equation is  $\lambda^{\otimes 2} \oplus 6 = 4 \otimes \lambda$ . We observe that both 2 and 4 are roots of the characteristic equation but since  $A$  is irreducible it has a unique eigenvalue given by the maximal average circuit weight, which turns out to be 4. Thus  $\rho = 2$  satisfies the characteristic equation of  $A$  but is not an eigenvalue of  $A$ .

## 2.6 Linear Dependence and Independence

As in conventional linear algebra we can define the linear dependence and independence of vectors in the max-plus sense. However the definitions are a little more complex. The following can be found in [12], [13], [3] and [2].

Recall that the max-plus algebra is an idempotent semi-ring. In order to define linear dependence, independence and bases we need the definition of a semi-module. A semi-module is essentially a linear space over a semi-ring. Semimodules and subsemimodules are analogous to modules and submodules over rings [2].

**Definition 35.** *A set  $V \subset \mathbb{R}_{\max}^n$  is a commutative idempotent semi-module over  $\mathbb{R}_{\max}$  if it is closed under  $\oplus$  and scalar multiplication;  $\alpha \otimes v \in V$  and  $u \oplus v \in V$  for all  $\alpha \in \mathbb{R}_{\max}$  and  $u, v \in \mathbb{R}_{\max}^n$ .*

**Definition 36.** *A finitely generated semi-module  $V \subset \mathbb{R}_{\max}^n$  is the set of all linear combinations of a finite set  $\{u_1, u_2, \dots, u_r\}$  of vectors in  $\mathbb{R}_{\max}^n$ :*

$$V = \left\{ \bigoplus_{i=1}^r \alpha_i \otimes u_i \mid \alpha_1, \alpha_2, \dots, \alpha_r \in \mathbb{R}_{\max} \right\}.$$

**Definition 37.** *A element  $x$  can be written as a finite linear combination of elements of  $F \subseteq V$  if  $x = \bigoplus_{f \in F} \lambda_f \otimes f$ , for some  $\lambda_f \in \mathbb{R}_{\max}$  such that  $\lambda_f = \epsilon$  for all but finitely many  $f \in F$ .*

As noted above linear independence and dependence in the max-plus sense are not completely analogous to the conventional definition. There are different interpretations of linear independence, dependence and bases. We will consider the definitions of linear dependence and linear independence due to Gondran and Minoux [13].

**Definition 38.** *A set of  $p$  vectors  $\{v_1, v_2, \dots, v_p\} \in \mathbb{R}_{\max}^n$  is linearly dependent if the set  $\{1, 2, \dots, p\}$  can be partitioned into disjoint subsets  $I$  and  $K$  such that for  $j \in I \cup K$  there exist  $\alpha_j \in \mathbb{R}_{\max}$ , not all equal to  $\epsilon$  and*



$$\bigoplus_{i \in I} \alpha_i v_i = \bigoplus_{k \in K} \alpha_k v_k.$$

**Definition 39.** A set of  $p$  vectors  $\{v_1, v_2, \dots, v_p\} \in \mathbb{R}_{\max}^n$  is linearly independent if for all disjoint subsets  $I$  and  $K$  of  $\{1, 2, \dots, p\}$   $j \in I \cup K$  and all  $\alpha_j \in \mathbb{R}_{\max}$  we have

$$\bigoplus_{i \in I} \alpha_i v_i \neq \bigoplus_{k \in K} \alpha_k v_k$$

unless  $\alpha_j = \epsilon$  for all  $j \in I \cup K$

In other words linearly independent simply means not linearly dependent.

**Definition 40.** A subset  $F$  of a semi-module  $M$  over  $\mathbb{R}_{\max}$  spans  $M$  or is a spanning family of  $M$  if every element  $x \in M$  can be written as a finite linear combination of the elements of  $F$

**Definition 41.** A family of vectors  $\{u_i\}_{i=1}^p$  is a weak basis of a semi-module  $V$  if it is a minimal spanning family.

**Definition 42.** A semi-module is finitely generated if it has a finite spanning family.

### 2.6.1 Bases

The following basis theorem and definitions are found in [13] and [12]. The proof is not stated but can be derived from the results found in [26].

**Theorem 13.** A finitely generated semi-module  $V \subset \mathbb{R}_{\max}^n$  has a weak bases. Any two weak bases have the same number of generators. For any two weak bases  $\{u_1, u_2, \dots, u_r\}$  and  $\{v_1, v_2, \dots, v_r\}$  there exist invertible scalars in  $\mathbb{R}_{\max}$  and a permutation  $\sigma$  of  $\{1, 2, \dots, r\}$  such that  $u_i = \alpha_i \otimes v_{\sigma(i)}$ .

**Definition 43.** The cardinality of a weak basis is called the weak rank of the semi-module.

We can also identify the column space and row space of a matrix. For the column space of a matrix  $A$  consider the  $n \times p$  matrix  $A = [c_1, c_2, \dots, c_p]$  where  $c_i$  are the  $p$  columns of  $A$ . The column space of  $A$  is the semi-module  $V = \{A \otimes x : x \in \mathbb{R}_{\max}^p\}$ . The row space of  $A$  is the column space of  $A^T$ .

**Definition 44.** *The Gondran-Minoux row rank of  $A \in \mathbb{R}_{\max}^{n \times n}$  is the maximum  $r \geq 1$  such that  $A$  has  $r$  linearly independent rows in the Gondran-Minoux sense. The column rank is the maximum  $r \geq 1$  such that  $A$  has  $r$  linearly independent columns in the Gondran-Minoux sense.*

Using these definitions we can find a weak basis which is linear dependent. Take  $X = \{(e, \epsilon), (\epsilon, e), (e, e)\}$ . It is not hard to see that  $X$  generates  $\mathbb{R}_{\max}^2$ . However we can write  $(e, e) = (e, \epsilon) \oplus (\epsilon, e)$  so  $X$  is not linear independent [26]. Therefore we see that independence is not needed to form a weak basis.

Note that in Chapter 1 we defined standard bases in the max-plus sense. (See Section 1.3.2.)

## 2.7 Asymptotic and Limiting Behavior

We now want to consider the sequences  $\{x(k) : k \in \mathbb{N}\}$  which are generated by  $x(k+1) = A \otimes x(k)$  with  $A \in \mathbb{R}_{\max}^{n \times n}$  and  $x(0) \in \mathbb{R}_{\max}^n$  as the initial condition. The sequences  $x(k)$  can also be described by

$$x(k) = A \otimes x(k-1) = A \otimes A \otimes x(k-2) = A^{\otimes 2} \otimes x(k-2).$$

Continuing in the fashion we see that  $x(k) = A^{\otimes k} \otimes x(0)$ .

We want to look at the quantitative asymptotic behavior of  $x(k)$ . This involves the cycle time vector and the asymptotic growth rate of  $x(k)$  which are defined below. In particular if  $A$  is irreducible the asymptotic growth rate of any  $x_j(k)$   $j \in \{1, 2, \dots, n\}$  is equal to the

unique eigenvalue of  $A$ . We will also consider the limiting behavior of  $x(k)$  and see when  $x(k)$  ultimately behaves like an eigenvector of  $A$ . The results here are from [16].

**Definition 45.** Let  $\{x(k) : k \in \mathbb{N}\}$  be a sequence in  $\mathbb{R}_{\max}^n$  such that for all  $j \in 1, 2, \dots, n$

$$\tau_j = \lim_{k \rightarrow \infty} \frac{x_j(k)}{k}$$

exists. Then vector  $\tau = (\tau_1, \tau_2, \dots, \tau_n)$  is called the cycle time vector of  $x(k)$ . If all  $\tau_i$  are equal then this value is called the asymptotic growth rate of the sequence  $x(k)$ .

An interesting observation is that for an irreducible matrix  $A \in \mathbb{R}_{\max}^{n \times n}$  with unique eigenvalue  $\lambda$  and associated eigenvector  $v$ , an eigenvalue of  $A^{\otimes k}$  is  $\lambda^{\otimes k}$  with the same corresponding eigenvector  $v$ . This is stated in the following lemma.

**Lemma 24.** If  $v$  is an eigenvector of  $A$  with eigenvalue  $\lambda$ , then

$$A^{\otimes k} \otimes v = \lambda^{\otimes k} \otimes v \text{ for all } k \geq 0$$

*Proof.* Note that since  $A \otimes (\lambda^{\otimes k} \otimes v) = \lambda^{\otimes k} \otimes (A \otimes v) = \lambda^k \otimes (\lambda \otimes v) = \lambda^{\otimes(k+1)} \otimes v$ , the proof follows by induction.

□

This tells us that if  $v$  is an eigenvector then for any  $j \in \{1, 2, \dots, n\}$

$$\lim_{k \rightarrow \infty} \frac{x_j(k)}{k} = \lambda.$$

So we see that for  $x(0) = v$ , where  $v$  is an eigenvector, the asymptotic growth rate of  $x(k)$  coincides with the associated eigenvalue of  $A$ . The question we want to consider is what happens when the sequence  $x(k)$  is initialized with a vector other than an eigenvector of  $A$ . We will soon see conditions under which the cycle time vector will always exist.

Before we go any further we need the following lemma which uses the  $l^\infty$  norm for a vector  $v \in \mathbb{R}^n$  and a definition. The  $l^\infty$  norm is defined by  $\|v\|_\infty = \max_{i \in \{1, 2, \dots, n\}} |v_i|$  for every  $v \in \mathbb{R}^n$ .

**Definition 46.** A matrix is called regular if it contains at least one element different from  $\epsilon$  in each row.

**Lemma 25.** Let  $A \in \mathbb{R}_{\max}^{m \times n}$  be a regular (not necessarily square) matrix and  $v, u \in \mathbb{R}^n$ . Then

$$\|(A \otimes u) - (A \otimes v)\|_\infty \leq \|u - v\|_\infty$$

*Proof.* First we note that  $A \otimes u$  and  $A \otimes v$  are finite vectors in  $\mathbb{R}^m$ . Define  $\beta = \|(A \otimes u) - (A \otimes v)\|_\infty$ . This means that there exists  $i_o \in \{1, 2, \dots, m\}$  such that  $\beta = |[(A \otimes u) - (A \otimes v)]_{i_o}|$ . So the  $i_o$  is the index of the entry with the maximum absolute value in  $(A \otimes u) - (A \otimes v)$ . Without loss of generality assume that that  $\beta = [(A \otimes u) - (A \otimes v)]_{i_o} \geq 0$ . Then  $\beta = \max_{j \in \{1, 2, \dots, n\}} (a_{i_o j} + u_j) - \max_{l \in \{1, 2, \dots, n\}} (a_{i_o l} + v_l)$  by the definition of max-plus matrix multiplication. Therefore there exists a  $j_o \in \{1, 2, \dots, n\}$  such that  $\beta = (a_{i_o j_o} + u_{j_o}) - \max_{l \in \{1, 2, \dots, n\}} (a_{i_o l} + v_l) \leq (a_{i_o j_o} + u_{j_o}) - (a_{i_o j_o} + v_{j_o}) = u_{j_o} - v_{j_o}$ . This implies that  $\beta \leq u_{j_o} - v_{j_o} \leq \max_{j \in \{1, 2, \dots, n\}} (u_j - v_j) \leq \max_{j \in \{1, 2, \dots, n\}} |u_j - v_j| = \|u - v\|_\infty$ . So  $\beta \leq \|u - v\|_\infty$ . Hence our proof is complete.

□

In the next theorem we consider what happens when  $x(0)$  is not necessarily an eigenvector of  $A$  and  $A$  is not necessarily irreducible. We use the notation  $x(k, x(0))$  to denote the vector  $x(k)$  initiated by  $x(0)$ . Also note that for  $x(0)$  to be a vector in  $\mathbb{R}^n$  means that  $x(0)$  has no  $\epsilon$  entries.

**Theorem 14.** Consider the recurrence relation  $x(k+1) = A \otimes x(k)$  for  $k \geq 0$  and  $A \in \mathbb{R}_{\max}^{n \times n}$  and initial condition  $x(0)$  where  $A \in \mathbb{R}_{\max}^{n \times n}$  is regular. If  $x(0) \in \mathbb{R}^n$  is a initial condition such that  $\lim_{k \rightarrow \infty} \frac{x(k, x(0))}{k}$  exists, then this limit has the same value for any initial condition  $y(0) \in \mathbb{R}^n$ .

*Proof.* Assume that  $x(0) \in \mathbb{R}^n$  and that  $\lim_{k \rightarrow \infty} \frac{x(k, x(0))}{k} = \tau$  with  $\tau \in \mathbb{R}^n$ . Then for any  $y(0) \in \mathbb{R}^n$  we have,

$$0 \leq \left\| \frac{x(k, y(0))}{k} - \frac{x(k, x(0))}{k} \right\|_{\infty} \leq \frac{1}{k} \|(A^{\otimes k} \otimes y(0)) - (A^{\otimes k} \otimes x(0))\|_{\infty} \leq \frac{1}{k} \|y(0) - x(0)\|_{\infty}.$$

Taking the limit as  $k \rightarrow \infty$  the last inequality goes to 0. Therefore we have that

$$\left\| \frac{x(k, y(0))}{k} - \frac{x(k, x(0))}{k} \right\|_{\infty} \rightarrow 0 \text{ as } k \rightarrow \infty,$$

and consequently  $\lim_{k \rightarrow \infty} \frac{x(k, y(0))}{k} = \tau$

□

So by the previous theorem we know that if the cycle time vector exists it is independent of the finite initial condition for any regular matrix. For an irreducible matrix A, the next lemma gives us the existence of the cycle time vector. The next lemma follows from Theorem 14 and Lemma 24. It tells us that we can initiate the recurrence relation  $x(k)$  with any initial vector and the cycle time vector is still equal to  $\lambda$  as we saw in Lemma 24.

**Lemma 26.** *Consider the recurrence relation  $x(k+1) = A \otimes x(k)$  for  $k \geq 0$ , with  $A \in \mathbb{R}_{\max}^{n \times n}$  irreducible such that  $v$  is an eigenvector associated with eigenvalue  $\lambda \in \mathbb{R}$ . Then for all  $j \in \{1, 2, \dots, n\}$*

$$\lim_{k \rightarrow \infty} \frac{x_j(k, x(0))}{k} = \lambda$$

for any initial condition  $x(0) \in \mathbb{R}_{\max}^n$ .

*Proof.* Let  $v$  be an eigenvector of A. Initializing the above recurrence relation with  $x(0) = v \in \mathbb{R}^n$  gives

$$\lim_{k \rightarrow \infty} \frac{x_j(k, x(0))}{k} = \lambda$$

for all  $j$  as we saw above. Note that since  $A$  is irreducible then  $A$  is regular and  $v$  is finite. By the previous theorem we know that once the cycle time vector (i.e. asymptotic growth rate) exists it is independent of  $x(0)$  therefore the result follows.  $\square$

So we have looked at the asymptotic growth rate of the sequence  $x(k)$ . Now we want to consider the limiting behavior of  $x(k)$ . For an irreducible matrix  $A$  the qualitative limiting behavior of  $x(k)$  has two possibilities. Either  $x(k)$  enters in to the eigenspace of  $A$  or  $x(k)$  enters a periodic regime. In order to see this we need to examine the cyclicity of  $A$ .

**Definition 47.** Let  $A \in \mathbb{R}_{\max}^{n \times n}$  be a matrix whose communication graph contains at least one circuit. The cyclicity of  $A$ ,  $\sigma(A)$ , is the cyclicity of the critical graph of  $A$ .

Here we are defining the cyclicity of a matrix. Previously we defined the cyclicity of a graph. So the cyclicity of a matrix  $A \in \mathbb{R}_{\max}^{n \times n}$  is the cyclicity of  $G^c(A)$ .

The following theorem gives us some insight about the cyclicity of  $A$ . We will omit the proof, as we stated above it can be found in [16]

**Theorem 15.** Let  $A \in \mathbb{R}_{\max}^{n \times n}$  be an irreducible matrix with eigenvalue  $\lambda$  and cyclicity  $\sigma = \sigma(A)$ . Then there is  $N$  such that  $A^{\otimes(k+\sigma)} = \lambda^{\otimes\sigma} \otimes A^{\otimes k}$  for all  $k \geq N$

Because of the previous theorem there exists a smallest number  $t(A)$  called the transient time.

**Definition 48.** Suppose  $A \in \mathbb{R}_{\max}^{n \times n}$  is an irreducible matrix with eigenvalue  $\lambda$  and cyclicity  $\sigma = \sigma(A)$ . The smallest number  $t(A)$  such that  $A^{\otimes(k+\sigma)} = \lambda^{\otimes\sigma} \otimes A^{\otimes k}$  for all  $k \geq t(A)$ .  $t(A)$  is called the transient time of  $A$ .

So for any initial condition  $x(0)$  we have

$$x(k + \sigma) = A^{\otimes(k+\sigma)} \otimes x(0) = \lambda^{\otimes\sigma} \otimes A^{\otimes k} \otimes x(0) = \lambda^{\otimes\sigma} \otimes x(k) \quad (2.17)$$

for all  $k \geq t(A)$ . This says that the sequence  $x(k)$  will show a periodic behavior after at most  $t(A)$  transition. So the periodic behavior of  $x(k)$  is characterized by the eigenvalue and the cyclicity of  $A$ .

**Lemma 27.** *If  $A \in \mathbb{R}_{\max}^{n \times n}$  is irreducible and has cyclicity one then  $x(k+1) = A \otimes x(k) = \lambda \otimes x(k)$  for  $k \geq t(A)$ .*

*Proof.* This lemma is a consequence of the previous theorem and follows from the definition of the transient time and (2.17).  $\square$

So if  $A$  has cyclicity one then for any initial vector  $x(0)$ ,  $x(k)$  becomes an eigenvector of  $A$  after  $t(A)$  steps. So the effect of the initial vector dies out after  $t(A)$  steps. Now what about the case where  $A$  does not have cyclicity one? To consider this case we need to define a periodic regime.

**Definition 49.** *Let  $A \in \mathbb{R}_{\max}^{n \times n}$ . A periodic regime is a set of vectors  $\{x^1, x^2, \dots, x^d\} \in \mathbb{R}_{\max}^n$ , each  $x_i$  contains at least one finite entry, for some  $d \geq 1$  such that there exists a finite number  $\rho$  that satisfies*

$$\rho \otimes x^1 = A \otimes x^d \text{ and } x^{i+1} = A \otimes x^i, i \in \{1, 2, \dots, d-1\}.$$

$$\{x^1, x^2, \dots, x^d\} \text{ is said to be of period } d.$$

Its not hard to see that  $x^1, x^2, \dots, x^d$  are all eigenvectors of  $A^{\otimes d}$  with eigenvalue  $\rho$ . If  $A$  is irreducible with cyclicity  $\sigma(A)$  then  $A$  will possess a periodic regime of period  $\sigma(A)$  or less.

**Theorem 16.** *Let  $\{x^1, \dots, x^d\}$  be a periodic regime for the matrix  $A$  with  $\rho \otimes x^1 = A \otimes x^d$ . Then  $A$  has an eigenvalue  $\lambda$  that satisfies  $\rho = \lambda^{\otimes d}$  and a corresponding eigenvector  $v$  given by  $v = \bigoplus_{j=1}^d \lambda^{\otimes(d-j)} \otimes x^j$*

*Proof.* Let  $v$  be defined as above in the theorem. We must show that  $A \otimes v = \lambda \otimes v$

$$A \otimes v = A \otimes \left( \bigoplus_{j=1}^d \lambda^{\otimes(d-j)} \otimes x^j \right) = \bigoplus_{j=1}^d A \otimes \lambda^{\otimes(d-j)} \otimes x^j$$

By the definition of a periodic regime we have  $A \otimes x^j = x^{j+1}$  and  $A \otimes x^d = \lambda^{\otimes d} \otimes x^1$  This gives us:

$$\begin{aligned} \bigoplus_{j=1}^d A \otimes \lambda^{\otimes(d-j)} \otimes x^j &= \lambda^{\otimes d} \otimes x^1 \oplus \bigoplus_{j=1}^{d-1} \lambda^{\otimes(d-j)} \otimes x^{j+1} = \lambda^{\otimes d} \otimes x^1 \oplus \bigoplus_{l=2}^d \lambda^{\otimes(d-l+1)} \otimes x^l \\ &= \bigoplus_{l=1}^d \lambda^{\otimes(d-l+1)} \otimes x^l = \lambda \otimes \left( \bigoplus_{l=1}^d \lambda^{\otimes(d-l)} \otimes x^l \right) = \lambda \otimes v \end{aligned}$$

So  $A \otimes v = \lambda \otimes v$  □

To illustrate the concept of periodic regime consider the following matrix.

$$A = \begin{pmatrix} 2 & 6 \\ 4 & 4 \end{pmatrix}$$

Let  $x(0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ . This generates the sequence:

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

This is a periodic regime with period 2 and  $\rho = 10$ . It is easy to see that the eigenvalue of  $A$  is 5. Note that  $\sigma(A) = 2$ . Using the previous theorem an eigenvector of  $A$  is given by:

$$\lambda^{\otimes 1} \otimes x^1 \oplus \lambda^{\otimes 0} \otimes x^2 = 5 \otimes \begin{pmatrix} 6 \\ 4 \end{pmatrix} \oplus 0 \otimes \begin{pmatrix} 10 \\ 10 \end{pmatrix} = \begin{pmatrix} 11 \\ 10 \end{pmatrix}$$

Now by computing  $A_\lambda^*$  since the critical vertices are 1 and 2 we find that  $\begin{pmatrix} 0 \\ -1 \end{pmatrix}$  and  $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$  are critical columns of  $A_\lambda^*$  and thus eigenvectors of  $A$ . So  $v = \begin{pmatrix} 11 \\ 10 \end{pmatrix}$  satisfies the equation

$$A \otimes v = \lambda \otimes v \text{ and } \begin{pmatrix} 11 \\ 10 \end{pmatrix} = 11 \otimes \begin{pmatrix} 0 \\ -1 \end{pmatrix} \oplus \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Therefore  $v$  must be an eigenvector of  $A$ .

So we see that when  $A$  is irreducible for any initial vector  $x(0) \in \mathbb{R}^n$ ,  $x(k)$  enters after at most  $t(A)$  iterations into a periodic regime with period  $\sigma(A)$  or smaller. If  $A$  has cyclicity



one then  $x(k)$  enters after at most  $t(A)$  iterations the eigenspace of  $A$  and  $x(k)$  behaves like an eigenvector of  $A$ . If  $x(k)$  enters a periodic regime the effect of the initial condition doesn't fade away and the sequence  $x(k)$  does not ultimately behave like an eigenvector of  $A$ . So for an irreducible matrix  $A$  we have determined the limiting behavior of  $x(k)$ .

# Chapter 3

## Markov Chains

We indicated above that max-plus algebra arises as the “algebra of exponents” for quantities having asymptotic expressions of the form  $e^{su}$  as  $s \rightarrow \infty$ ; see Lemma 1. Such asymptotic expressions arise in many places. One in particular is large deviations analysis in probability. Large deviations analysis is concerned with the asymptotic decay of some probability  $p^\epsilon$  which depends on a vanishing parameter  $\epsilon \downarrow 0$ . The *large deviations rate*, if it exists, is defined by

$$V = -\lim_{\epsilon \downarrow 0} \epsilon^2 \ln(p^\epsilon).$$

Using our notation this is expressed as  $p^\epsilon \asymp e^{-\frac{V}{\epsilon^2}}$ , where  $V \geq 0$ . For the connection with standard max-plus algebra we replace  $-V$  by  $u$  (with  $u \leq 0$ ). (Otherwise we would need to work in the min-plus algebra — we prefer to stay with the standard max-plus algebra throughout this thesis.) We also replace  $\epsilon^{-2}$  ( $\epsilon \downarrow 0$ ) by  $s \rightarrow \infty$ . This is because we want to reserve  $\epsilon$  for the max-plus additive identity  $\epsilon = -\infty$ , as is customary.

The book [27] of Wentzell and Freidlin is concerned with large deviations analysis of continuous processes (diffusions) on Euclidean spaces, and connected manifolds in general. In Chapter 6 they show that for certain problems involving large time intervals, the analysis can be reduced to the study of transitions between a finite number of compact subsets  $K_i$ . The continuous process moves between the  $K_i$  in a way which can be approximated by a

Markov chain on a finite set of states, one for each  $K_i$ , having transition probabilities of the form

$$p_{ij} \asymp e^{su_{ij}}, \text{ as } s \rightarrow \infty. \quad (3.1)$$

(This is after changing from their  $V, \epsilon$  notation to our  $u, s$  notation as above.) Wentzell and Freidlin establish several results for finite state Markov chains with this kind of asymptotic dependence. Our purpose in this chapter is to consider their Markov chain results as applications of the max-plus linear algebra results described in the previous chapter.

We begin with the definition of a Markov chain. A random process  $X = \{X_0, X_1, \dots\}$  is a sequence of random variables which take values in some countable set  $L$ , called the state space, is a Markov chain if it satisfies the Markov condition:  $P(X_n = s | X_0 = x_0, X_1 = x_1, \dots, X_{n-1} = x_{n-1}) = P(X_n = s | X_{n-1} = x_{n-1})$  for all  $n \geq 1$ . The chain is described by its transition probabilities  $P(X_{n+1} = j | X_n = i) = p_{ij}$ . We want to consider a Markov chain where the  $p_{ij}$  depend on a parameter  $s$  converging to infinity, in such a way that  $\lim_{s \rightarrow \infty} s^{-1} \ln(p_{ij}) = u_{ij}$ , in other words  $p_{ij} \asymp e^{su_{ij}}$ . To be transition probabilities of a Markov chain we require  $p_{ij} \geq 0$  and  $\sum_j p_{ij} = 1$  for each  $i$ . This implies that  $u_{ij} \leq e$  and  $\bigoplus_j u_{ij} = e$ . Note that in this chapter  $u_{ij}$  represents the weight of an edge from  $(i, j)$ , not  $(j, i)$  as is the customary max-plus interpretation.

For a Markov chain with such an asymptotic description the following three problems are considered by Wentzell and Freidlin in [27]. We will describe the max-plus interpretation of their results.

1. Determine the asymptotics of the stationary distribution.
2. For  $B \subset L$ , determine the asymptotic probability that the first contact with  $B$  occurs at  $k \in B$ .
3. Determine the asymptotics of the expected number of steps to reach  $B \subset L$ .

In this chapter we consider these results from a max-plus perspective assuming that  $L$  is finite.

### 3.0.1 Problem 1

In conventional algebra the stationary distribution is a vector  $v$  such that  $v = vP$ , with  $v_i \geq 0$  and  $\sum v_i = 1$ . It is unique if the probability matrix  $P$  is irreducible. It will follow from Theorem 17 there exist  $w_i$  for which we have  $v_i \asymp e^{sw_i}$ . Our goal is to characterize  $w = [w_i]$  in max-plus terms. Now let  $U$  be the matrix of entries  $u_{ij}$  taken from the asymptotic description (3.1) of  $P$ . From Lemma 1 it follows that  $w = U^T \otimes w$ . It is assumed in [27] that  $P$  is irreducible (for each  $s \geq 0$ ). That does not necessarily imply that  $U$  is irreducible. However if  $U$  is irreducible then  $P$  is irreducible. The fact that  $v_i \geq 0$  and  $\sum_i v_i = 1$  tells us that  $\bigoplus w_i = e$  and  $\epsilon \leq w_i \leq e$ .

We see that the exponents  $w_i$  for the stationary distribution form a max-plus eigenvector for  $U^T$  with eigenvalue  $e = 0$ . Since  $u_{ij} \leq e$  we know that all circuits have weight less than or equal to  $e$ . Lemma 19 and the existence of an eigenvector tells us that there is a circuit of average weight  $e$ . Thus  $U^T$  has a maximal average circuit weight  $e$ , even though it may fail to be irreducible. In the conventional sense we know that the stationary distribution  $v$  is unique when the probability matrix  $P$  is irreducible. However even for an irreducible matrix the max-plus eigenvector is not unique in general, see Example 1 below. Recall that if  $U$  is irreducible the critical columns of  $U^T$  are eigenvectors associated with eigenvalue  $\lambda = e$  and every other eigenvector is a max-plus linear combination of the critical columns of  $U^T$ . So for an irreducible matrix we would like to know which linear combination of columns gives us the correct eigenvector for the asymptotic description  $v$ ? In general we would like to have a concise max-plus characterization of the eigenvector  $w$  for the asymptotic description of  $v$ . Although we do not know of a simple criteria which uniquely identifies  $w$  out of all max-plus eigenvectors, we will be able to identify it constructively; see Theorem 17. In [1] Akian, Bapat and Gaubert have some additional insight into this question. We will discuss their results shortly.

Freidlin and Wentzell describe a graph theoretical approach to finding the stationary distri-

bution. In this approach they use  $B$ -graphs, which we will now define.

**Definition 50.** *Suppose we have a subset  $B$  of a finite set  $L$ . A  $B$ -graph is a graph with vertices  $L$  which consists of edges  $(m, n)$  with  $m \in L \setminus B, n \in L$  and  $n \neq m$  and which satisfies the following conditions:*

1. every point  $m \in L \setminus B$  is an initial point of exactly one edge.
2. for any point  $m \in L \setminus B$  there is a sequence of edges leading from it to some point  $n \in B$

The set of  $B$ -graphs is denoted by  $G(B)$ . For  $g \in G(B)$  we define  $\pi(g) = \prod_{(m,n) \in g} p_{mn}$  where  $p_{ij}$  are the transition probabilities for the Markov chain.

The following is Lemma 3.1 of Chapter 6 from [27], describing the stationary distribution in terms of  $B$ -graphs.

**Lemma 28.** *Consider an irreducible Markov chain with finite set of states  $L$  and transition probabilities  $p_{ij}$ . Then the stationary distribution of the chain is  $v_i = \{(\sum_{i \in L} Q_i)^{-1} Q_i\}$  where,*

$$Q_i = \sum_{g \in G\{i\}} \pi(g).$$

The  $B$ -graphs translate directly into the max-plus setting. Since  $p_{ij} \asymp e^{su_{ij}}$  we have  $\pi(g) \asymp e^{s\rho(g)}$  where  $\rho(g) = \bigotimes_{(i,j) \in g} u_{ij}$ . Thus  $\rho(g)$  is the max-plus product of the weights of the edges in  $g$ . Now we can state the max-plus version of the previous lemma.

**Theorem 17.** *Consider an irreducible Markov chain with the set of finite states  $L$  and transition probabilities  $p_{ij} \asymp e^{su_{ij}}$ . There exists a max-plus eigenvector  $w$  of  $U^T$  whose  $k^{\text{th}}$  component is  $w_k = \tilde{w}_k - \bigoplus_{k \in L} \tilde{w}_k$ , where*

$$\tilde{w}_k = \bigoplus_{g \in G(\{k\})} \rho(g).$$

Moreover  $v_k \asymp e^{sw_k}$  where  $v_k$  is the  $k^{\text{th}}$  component of the unique stationary distribution of  $P$ .

*Proof.* Applying Lemma 1 to the  $Q_k$  of Lemma 28 we have that  $Q_k \asymp e^{s\tilde{w}_k}$ . In other words  $\tilde{w}_k = \lim_{s \rightarrow \infty} s^{-1} \ln(Q_k) = \bigoplus_{g \in G(\{k\})} \rho(g)$ . Since  $v_k = \frac{Q_k}{\sum_j Q_j}$  it follows that  $v_k \asymp e^{sw_k}$ . Since  $v = vP \asymp e^{s(w \otimes U)}$  (coordinate wise) the equation  $vP = v$  implies  $U^T \otimes w = w$ .

□

This theorem tells us that the max-plus version of Freidlin and Wentzell's graph theoretical construction of the stationary distribution produces the max-plus eigenvector  $w$  associated with the asymptotic description of  $P$ . Note that  $w$  is the normalized vector of  $\tilde{w}$  in the max-plus sense. So using this construction we can identify the right max-plus eigenvector  $w$  that corresponds to the asymptotic description of the stationary distribution even though max-plus eigenvectors are not unique. From Chapter 2, if  $U$  is irreducible, then  $w$  is a linear combination of the critical columns of  $U^T$ . However we do not have a procedure that tells us which linear combination gives us the eigenvector  $w$  we are concerned with.

Below we emphasize the important results of Problem 1.

- Assuming the transition matrix  $P$  is an irreducible probability matrix, the stationary distribution (normalized left Perron eigenvector)  $v$  *does* have exponential asymptotics  $v_i \asymp e^{sw_i}$  where  $w$  is a max-plus eigenvector of the max-plus matrix  $U^T$  associated with  $P$ .
- Although the eigenspace for the Perron eigenvalue of  $P$  ( $p_{ij} \asymp e^{su_{ij}}$ ) is one-dimensional (for each  $s$ ), the max-plus eigenspace is not "one-dimensional" in general (i.e. there can be other eigenvectors beyond  $c \otimes w$ ,  $c \in \mathbb{R}_{\max}$ ) except in the case that  $U$  has a unique critical class ( see Lemma 23 on page 45).
- The correct max-plus eigenvector  $w$  for the asymptotic description of the stationary distribution *is* determined by the max-plus matrix  $U$ . The max-plus version of the

$B$ -graph gives us a graph theoretical procedure for finding  $w$ . What we don't know is a nice characterizing property that picks  $w$  out of all the possible max-plus eigenvectors of  $U^T$ . As far as we know this is an unsolved problem except for the case that  $U$  has a unique critical class.

To get a better idea of what a  $B$ -graph is in max-plus, let's consider the example illustrated in Figure 3.1. In addition to the matrix  $U$  the figure shows the graph with the  $u_{ij}$  indicated as edge weights.

$$U_1 = \begin{pmatrix} 0 & 0 & -\infty & -\infty \\ -1 & 0 & -9 & -9 \\ -\infty & -6 & 0 & -6 \\ -\infty & 0 & 0 & 0 \end{pmatrix}$$

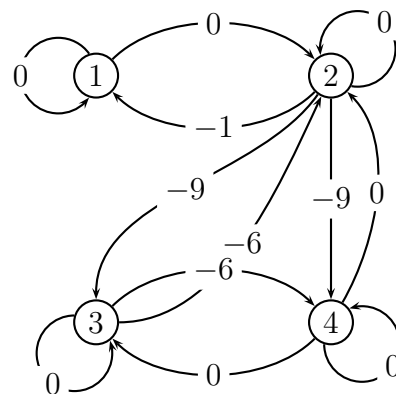


Figure 3.1:  $G(U_1)$

Let us calculate the set of graphs in  $G(\{1\})$  of the Markov chain with transition probabilities  $u_{ij}$  below. Here  $L = \{1, 2, 3, 4\}$  and  $B = \{1\}$ . For each vertex in  $L \setminus B = \{2, 3, 4\}$  we pick one edge with that vertex as its initial point and so that there are no circuits for our selected edges. (If we limit ourselves to  $g$  with  $\rho(g) > \epsilon$ , then we can ignore edges with  $u_{ij} = \epsilon$ .) Each such selection of edges produces a  $g \in G(\{1\})$ . For example  $g_1 = \{(2, 1), (3, 2), (4, 2)\}$  make up one graph in  $G(\{1\})$ . We also have  $g_2 = \{(2, 1), (3, 2), (4, 3)\}$  and  $g_3 = \{(3, 4), (4, 2), (2, 1)\}$  in  $G(\{1\})$ .

To demonstrate Theorem 17 we will consider a few examples. Given a matrix  $U_1^T$  obtained from the asymptotics of  $P$  we will use the previous theorem to find the max-plus eigenvector  $w$  associated with  $P$ . We will also show that  $w$  satisfies the max-plus linear algebra results in Chapter 2 about eigenvectors. As you will see the eigenspace of  $U_1^T$  is rather complicated and there is no obvious way to determine which linear combination of critical columns determine

the eigenvector  $w$ .

**Example 6.** *Continuing with the example of Figure 3.1, we compute the max-plus eigenvector  $w$  for  $U_1^T$ .*

First we will compute  $w_1$  using Theorem 17. For each  $g_j \in G(\{1\})$  we need to compute  $\rho(g_j)$ . Note that this is just the max-plus product of the edges in  $g_j$  for each  $i \in L$ .

$$\rho(g_1) = -1 \otimes -6 \otimes 0 = -7$$

$$\rho(g_2) = -1 \otimes -6 \otimes 0 = -7$$

$$\rho(g_3) = -6 \otimes 0 \otimes -1 = -7.$$

Therefore  $\tilde{w}_1 = \oplus_{g_j \in G(\{1\})} \rho(g_j) = -7 \oplus -7 \oplus -7 = -7$ . Similarly we find that

$$\tilde{w}_2 = -6, \tilde{w}_3 = -9, \tilde{w}_4 = -15.$$

Normalizing  $\tilde{w}$  in the max-plus sense we see that  $w = \begin{pmatrix} -1 \\ 0 \\ -3 \\ -9 \end{pmatrix}$ .

Now since  $w$  is an eigenvector of  $U_1^T$  we should be able to write  $w$  as a max-plus linear combination of the critical columns of  $U_1^T$ . By Chapter 2 we find that

$$(U_1^T)^* = \begin{pmatrix} 0 & -1 & -7 & -1 \\ 0 & 0 & -6 & 0 \\ -9 & -9 & 0 & 0 \\ -9 & -9 & -6 & 0 \end{pmatrix}.$$

Notice that since the graph of  $U_1$  contains a loop of weight  $e$  on each vertex then  $V^c(U_1^T) = \{1, 2, 3, 4\}$  and



$$\begin{pmatrix} -1 \\ 0 \\ -3 \\ -9 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ -9 \\ -9 \end{pmatrix} \oplus -3 \otimes \begin{pmatrix} -7 \\ -6 \\ 0 \\ -6 \end{pmatrix}.$$
 This shows us that  $w$  can be written as a linear combination of the second and third critical columns of  $U_1^T$ .

**Example 7.** : Consider the Markov chain described by matrix and graph in Figure 3.2.

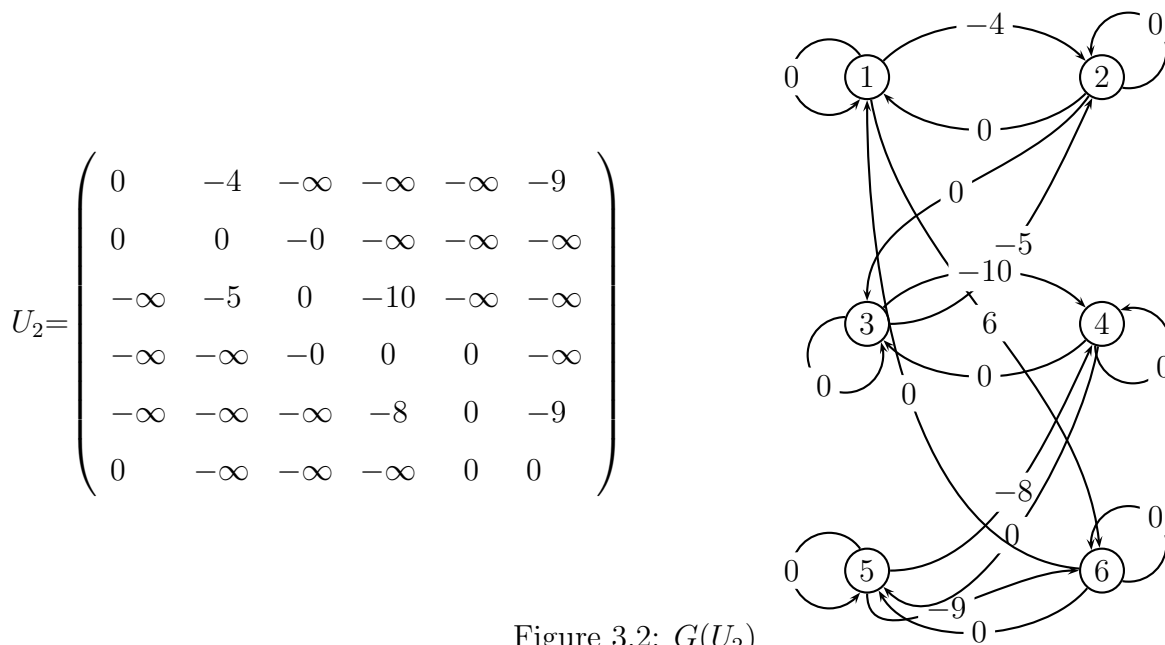


Figure 3.2:  $G(U_2)$

We can calculate the stationary distribution using Theorem 17. We find that

$$\tilde{w}_1 = \bigoplus_{g \in G(\{1\})} \rho(g) = -13$$

$\{(2, 1), (3, 2), (4, 3), (5, 4), (6, 1)\}$  is an example of a graph from  $G(\{1\})$  with  $\rho(g) = -13$ .

Similarly we find that  $\tilde{w}_2 = -17, \tilde{w}_3 = -17, \tilde{w}_4 = -19, \tilde{w}_5 = -14, \tilde{w}_6 = -19$ .

Therefore  $w = \begin{pmatrix} 0 \\ -4 \\ -4 \\ -6 \\ -1 \\ -6 \end{pmatrix}$  since  $\bigoplus_i \tilde{w}_i = -13$ . So  $w$  is the correct stationary distribution

corresponding to the asymptotics of  $P$  represented in Figure 3.2 and  $U_2^T \otimes w = w$ . As in the previous example we can write  $w$  as a linear combination of the critical columns of  $U_2^T$ . First we calculate  $(U_2^T)^*$ :

$$(U_2^T)^* = \begin{pmatrix} 0 & 0 & -5 & -5 & -9 & 0 \\ -4 & 0 & -5 & -5 & -13 & -4 \\ -4 & 0 & 0 & 0 & -8 & -4 \\ -14 & -10 & -10 & 0 & -8 & -8 \\ -6 & -6 & -10 & 0 & 0 & 0 \\ -6 & -6 & -11 & -9 & -9 & 0 \end{pmatrix}.$$

As in the previous example  $V^c(U_2^T) = \{1, 2, 3, 4, 5, 6\}$  and

$$w = \begin{pmatrix} 0 \\ -4 \\ -4 \\ -6 \\ -1 \\ -6 \end{pmatrix} = \begin{pmatrix} 0 \\ -4 \\ -4 \\ -14 \\ -6 \\ -6 \end{pmatrix} \oplus -6 \otimes \begin{pmatrix} -5 \\ -5 \\ 0 \\ 0 \\ 0 \\ -9 \end{pmatrix} \oplus -1 \otimes \begin{pmatrix} -9 \\ -13 \\ -8 \\ -8 \\ 0 \\ -9 \end{pmatrix}.$$

This confirms that  $w$  can be written as a linear combination of the first, fourth and fifth critical column of  $U_2^T$ .

Before moving on to Problem 2, we want to discuss the work of Akian, Bapat and Gaubert [1] and its implication for identifying the max-plus eigenvector  $w$  of Theorem 17. The stationary distribution  $v$  is the unique (up to normalization) left eigenvector guaranteed by the classical Perron-Frobenius Theorem. Akian, Bapat and Gaubert consider the asymp-

otics of the Perron eigenvalue and eigenvector of irreducible non-negative matrices. They consider the matrix  $\mathcal{A} = \lim_{s \rightarrow \infty} \mathcal{A}(s)^{1/s}$  which going back to Lemma 1 is the same as our  $U = \lim_{s \rightarrow \infty} s^{-1} \ln(\mathcal{A}(s))$  with  $U = \ln(\mathcal{A})$ . The natural logarithm maps their max-times algebra on  $[0, \infty)$  to the usual max-plus algebra on  $[-\infty, \infty)$ . They consider matrices  $\mathcal{A}(s)$  which are nonnegative, but not necessarily probability matrices. That is very significant. They offer the following example:

$$\mathcal{A}(s) = \begin{pmatrix} 1 + \cos(s)e^{-s} & e^{-2s} \\ e^{-2s} & 1 \end{pmatrix}.$$

The (right) Perron eigenvector turns out to be  $\mathcal{U}(s) = (e^s \cos(s) + \sqrt{4 + e^{2s} \cos^2(s)}, 2)$ . We find that  $s^{-1} \ln$  of the first coordinate is oscillatory as  $s \rightarrow \infty$ . This shows that the Perron eigenvector  $\mathcal{U}(s)$  can be such that  $\lim_{s \rightarrow \infty} \mathcal{U}_i(s)^{1/s}$  fails to exist even though the  $\lim_{s \rightarrow \infty} \mathcal{A}_{ij}(s)^{1/s}$  do exist. However their matrix  $\mathcal{A}$  is not a probability matrix; neither the rows nor the columns sum to 1. For irreducible probability matrices we know that the (left) Perron eigenvector (which is the stationary distribution, when normalized) *always does* have exponential asymptotics, because of applying Theorem 17. We can transpose their  $\mathcal{A}(s)$  (so that what was a left eigenvector becomes a right eigenvector) and then multiply the rows by constants so that they sum to 1. The left Perron eigenvector now works out to be  $\mathcal{U}(s) = \left( \frac{1 + e^{2s} + e^s \cos(s)}{1 + e^{2s}}, 1 \right)$ . The point is that  $\lim_{s \rightarrow \infty} s^{-1} \ln$  does exist for this. Thus the asymptotic behavior of the Perron eigenvector is much better for probability matrices than for nonnegative matrices in general. In Problem 1 we considered only probability matrices.

In the general case of [1], there need not exist a particular max-plus eigenvector  $\mathcal{U}$  associated with the Perron eigenvector of  $\mathcal{A}(s)$ ; there could be multiple limit points of  $\mathcal{U}(s)_i^{1/s}$  as  $s \rightarrow \infty$ . In [1] they use  $\mathcal{U}$  to represent the max-plus eigenvector which is not the same as our max-plus probability matrix  $U$ . In [1] they do not consider the problem of identifying the correct eigenvector if the eigenspace is multidimensional. In general there need not be one. They simply focus on sufficient conditions for the max-plus eigenspace to be “one-dimensional.” They do this first with just max-plus ideas in their Theorem 2 which is stated

below.

**Theorem 18.** *If  $\mathcal{A} = \lim_{s \rightarrow \infty} \mathcal{A}(s)^{1/s}$ ,  $\mathcal{A}$  is irreducible and  $\mathcal{A}$  has a unique critical class, then for all vertices  $j$  in the critical class,*

$$\lim_{s \rightarrow \infty} \mathcal{U}_i(s)^{1/s} = \frac{(\tilde{\mathcal{A}}^*)_{ij}}{\bigoplus_k (\mathcal{A}^*)_{kj}} \text{ for } 1 \leq i \leq n.$$

In [1] they state that this theorem is a consequence of max-plus results from Cunnigham-Green, [3], Gondran and Minoux and [12]. We can see that it follows from our Lemma 23 and Lemma 1.

Then they consider the same problem using the algebra of max-jets in place of max-plus. The max-plus idea is to just identify the asymptotic exponent  $B$  in expressions like  $f(s) \sim B^s$ . (We would write  $f(s) \asymp e^{s \ln(B)}$ .) The idea of max-jets is to include a pre-exponential coefficient along with the exponent:

$$f(p) = b \cdot B^p + o(B^p).$$

The pair  $(b, B)$  is a typical element of the max-jet algebra  $\mathbb{J}_{\max}$ . The definitions of  $\oplus$  and  $\otimes$  in (6) on page 930 are just the operations on pairs  $(b, B)$  and  $(c, C)$  that correspond to  $f(p) + g(p)$  and  $f(p) \cdot g(p)$ , if  $f$  is described by  $(b, B)$  and  $g$  is described by  $(c, C)$ . A max-jet asymptotic description of the Perron eigenvector would be an eigenvector in the max-jet algebra, and their Theorem 4 is a description of when that eigenspace is “one-dimensional.” Their Theorem 5 tells us when they can deduce that there is a unique max-jet vector describing the asymptotics of the Perron eigenvector of  $\mathcal{A} \in \mathbb{J}_{\max}^{n \times n}$ . One could ask about how the max-jet analysis works out if we insist that the matrix is a probability matrix. We might expect a max-jet version of the  $B$ -graph representation to hold.

They also claim that if  $\mathcal{A}(s)$  has a Dirichlet series representation, then the Perron eigenvector is uniquely determined (in the general nonnegative case, not just the probability matrix case) and is described by a Dirichlet series representation. We might view this as another step in the same direction. The more structure you assume about  $\mathcal{A}(s)$ , the more structure you can

expect the eigenvector to have. We have already seen this if the structure is a probability matrix. We should expect the same thing if the structure is in terms of a more specific type of asymptotic description or series representation.

### 3.0.2 Problem 2

As in Problem 1 we will consider a Markov chain on a state space  $L$  with  $B \subset L$ . Let  $D = L \setminus B$  and pick  $k \in B$ . We want to determine the probability of starting at state  $i \in D$  and hitting  $B$  first at state  $k$ . In conventional algebra for  $i \in D$  we let  $h_i$  be the probability that the chain starting at  $i \in D$  will first reach  $B$  at  $k \in B$ . Denote the size of  $D$  by  $d$ . The  $h_i$   $i \in D$  are characterized by the equation  $h_i = \sum_{j \in D} p_{ij} h_j + p_{ik}$ ,  $i \in D$ . So  $h$  is a  $d \times 1$  vector. We define  $P^D = [P_{ij}]$  where  $i, j \in D$  and  $b^{(k)} = [p_{ik}]$  with  $i \in D$ . Then we have the matrix equation  $h = P^D h + b^k$ . According to Lemma 3.3 from Chapter 6 of [27],  $h_i \asymp e^{sa_i}$  for some  $a_i \in \mathbb{R}_{\max}$ . Continuing to assume  $P$  is an irreducible matrix with  $p_{ij} \asymp e^{su_{ij}}$  we apply Lemma 1 to see that for each  $i \in D$ ,  $a_i = [\bigoplus_{j \in D} u_{ij} \otimes a_j] \oplus u_{ik}$ . As a max-plus matrix equation this is  $a = [U^D \otimes a] \oplus c^k$  where  $c^k = [u_{ik}]$ ,  $i \in D$ .

We know that the average circuit of  $U^D$  must be less than or equal to  $e = 0$ , since  $-\infty \leq u_{ij} \leq e$ . By Theorem 7 on page 31 the equation  $a = [U^D \otimes a] \oplus c^k$  has at least one solution. But in general this solution is not unique unless the maximal average circuit in  $U^D$  is negative. So the question is how to identify the correct solution for the asymptotic description of  $h_i$ .

As in problem 1, Freidlin and Wentzell have a graph theoretical technique for determining the probability  $h_i$  described above. Again we will consider the set of  $B$ -graphs. We need to consider a subset of  $G(B)$  denoted by  $G_{ik}(B)$  where  $i \in D$  and  $k \in B$ .

**Definition 51.** *The set of  $G_{ik}(B)$  graphs are a subset of the  $B$ -graphs such that each graph is  $G_{ik}(B)$  has the following properties:*

- *They do not contain circuits.*

- For each vertex in  $D$  there is a sequence of edges going into  $B$ .
- There exists a sequence of edges from  $i$  to  $k$ .
- The remaining vertices  $j \neq i$  in  $D$  may be included in this sequence of edges or may be contained in a sequence of edges ending in  $B$ .

The following is a restatement of Lemma 3.3 found in Chapter 6 of [27].

**Lemma 29.** *Given an irreducible Markov chain on a finite state space  $L$ , the probability that chain starting at  $i \in D = L \setminus B$  hits  $B \subset L$  first at  $k \in B$  is given by:*

$$h_i = \frac{\sum_{g \in G_{ik}(B)} \pi(g)}{\sum_{g \in G(B)} \pi(g)}.$$

Now we state the max-plus version of the previous lemma.

**Theorem 19.** *Given an irreducible Markov chain on a finite state space  $L$  and transition probabilities  $p_{ij} \asymp e^{su_{ij}}$  where  $U$  is irreducible, let  $h$  be the probability of first hitting  $B$  at  $k \in B$ . Then  $h_i \asymp e^{sa_i}$  where  $a_i = \bigoplus_{g \in G_{ik}(B)} \rho(g) - \bigoplus_{g \in G(B)} \rho(g)$  for each  $i \in L \setminus B$ . This is the solution to  $a = [U^D \otimes a] \oplus c^k$  which provides the asymptotic description of  $h$ .*

We have added the extra hypothesis that  $U$  is irreducible to guarantee that  $\bigoplus_{g \in G(B)} \rho(g) \neq \epsilon$ .

Note that if  $U$  is irreducible,  $\bigoplus_{g \in G(B)} \rho(g) > \epsilon$ .

*Proof.* Using Lemma 1 and the same notation from problem 1, we see that the description of  $h_i$  in Lemma 29 implies that  $h_i \asymp e^{sa_i}$  where  $a_i = \bigoplus_{g \in G_{ik}(B)} \rho(g) - \bigoplus_{g \in G(B)} \rho(g)$ .  $\square$

We have already observed that  $a = [a_i]$ ,  $i \in D$  is a solution of  $a = [U^D \otimes a] \oplus c^k$ . Theorem 19 provides an explicit construction of the correct solution for the asymptotic description of  $h$ .

To illustrate the procedure of finding the probability of hitting  $B \subset L$  first at  $k \in B$  we will consider the examples of Figure 3.1 and Figure 3.2 again.

**Example 8.** Consider the example of Figure 3.1. Let  $B = \{1, 3\}$  and  $D = \{2, 4\}$  and  $k = 3$ .

We see that

$$\begin{aligned}\bigoplus_{g \in G_{23}(B)} \rho(g) &= u_{24} + u_{43} = -9, \\ \bigoplus_{g \in G_{43}(B)} \rho(g) &= u_{43} + u_{21} = -1 \\ \bigoplus_{g \in G(B)} \rho(g) &= u_{21} + u_{43} = -1.\end{aligned}$$

Therefore

$$a = \begin{pmatrix} -8 \\ 0 \end{pmatrix}.$$

We can check directly that  $a = (U_1^D \otimes a) \oplus c^k: \begin{pmatrix} 0 & -9 \\ 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} -8 \\ 0 \end{pmatrix} \oplus \begin{pmatrix} -9 \\ 0 \end{pmatrix} = \begin{pmatrix} -8 \\ 0 \end{pmatrix}.$

**Example 9.** Consider the example of Figure 3.2. Let  $B = \{4, 5, 6\}$  and  $k = 4$ .

We see that

$$\begin{aligned}\bigoplus_{g \in G_{16}(B)} \rho(g) &= u_{12} + u_{23} + u_{34} = -14, \\ \bigoplus_{g \in G_{26}(B)} \rho(g) &= u_{12} + u_{23} + u_{34} = -14, \\ \bigoplus_{g \in G_{36}(B)} \rho(g) &= u_{12} + u_{23} + u_{34} = -14 \\ \bigoplus_{g \in G(B)} \rho(g) &= u_{32} + u_{21} + u_{16} = -11.\end{aligned}$$

Therefore

$$a = \begin{pmatrix} -3 \\ -3 \\ -3 \end{pmatrix}.$$

This is the solution that corresponds to the asymptotics of  $h$ . Again we check that  $a = [U_2^D \otimes a] \oplus c^k$  directly.

$$\left[ \begin{pmatrix} 0 & -4 & -\infty \\ 0 & 0 & 0 \\ -\infty & -5 & 0 \end{pmatrix} \otimes \begin{pmatrix} -3 \\ -3 \\ -3 \end{pmatrix} \right] \oplus \begin{pmatrix} -\infty \\ -\infty \\ -10 \end{pmatrix} = \begin{pmatrix} -3 \\ -3 \\ -3 \end{pmatrix}.$$

**Example 10.** Now we consider a slightly more difficult example (see figure 3.3).

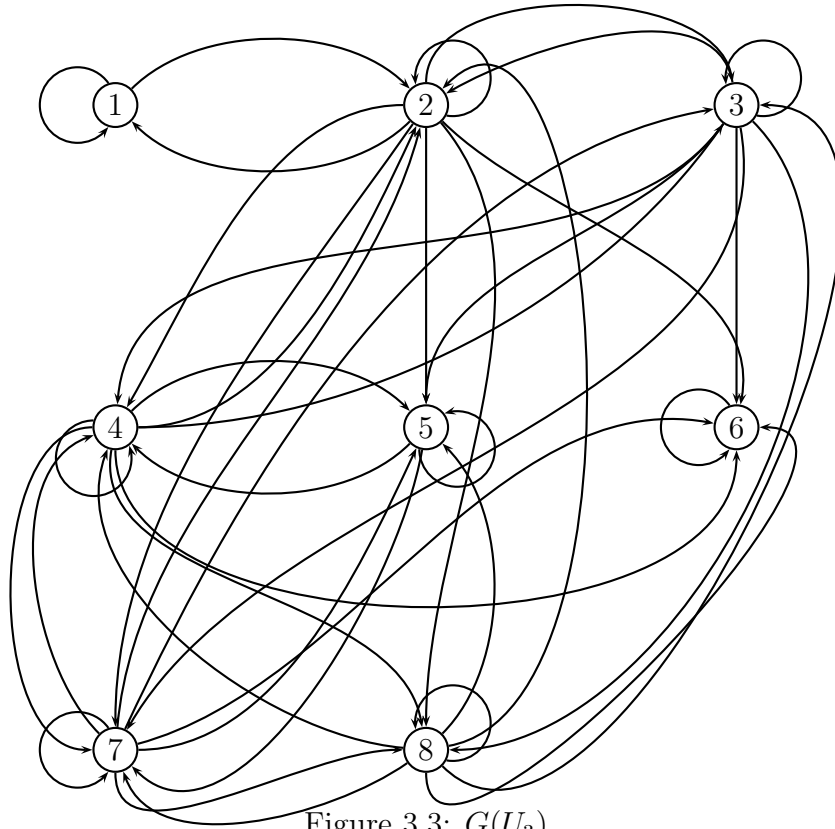


Figure 3.3:  $G(U_3)$



$$U_3 = \begin{pmatrix} 0 & -0 & -\infty & -\infty & -\infty & -\infty & -\infty & -\infty \\ -1 & 0 & -9 & -9 & -10 & -9 & -12 & -8 \\ -\infty & -6 & 0 & -6 & -7 & -6 & -7 & -7 \\ -\infty & 0 & 0 & 0 & -5 & -4 & -7 & -5 \\ -\infty & 0 & 0 & 0 & 0 & 0 & -5 & 0 \\ -\infty & -3 & 0 & -3 & -7 & 0 & -5 & -5 \\ -\infty & 0 & 0 & 0 & -5 & 0 & 0 & 0 \\ -\infty & 0 & -8 & -8 & -10 & -9 & -11 & 0 \end{pmatrix}$$

Let  $B = \{5, 6, 7, 8\}$  and  $k = 5$ . We see that

$$\bigoplus_{g \in G_{15}(B)} \rho(g) = -16,$$

$$\bigoplus_{g \in G_{25}(B)} \rho(g) = -16,$$

$$\bigoplus_{g \in G_{35}(B)} \rho(g) = -15,$$

$$\bigoplus_{g \in G_{45}(B)} \rho(g) = -15$$

$$\bigoplus_{g \in G(B)} \rho(g) = -14$$

$$a = \begin{pmatrix} -2 \\ -2 \\ -1 \\ -1 \end{pmatrix}.$$

Again we check that  $a$  also satisfies the equation  $a = [U_3^D \otimes a] \oplus c^k$ :

$$\left[ \begin{pmatrix} 0 & -0 & -\infty & -\infty \\ -1 & 0 & -9 & -9 \\ -\infty & -6 & 0 & -6 \\ -\infty & 0 & 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} -2 \\ -2 \\ -1 \\ -1 \end{pmatrix} \right] \oplus \begin{pmatrix} -\infty \\ -10 \\ -7 \\ -5 \end{pmatrix} = \begin{pmatrix} -2 \\ -2 \\ -1 \\ -1 \end{pmatrix}.$$

### 3.0.3 Problem 3

As in Problem 1 and 2 we have a Markov chain on a state space  $L$  with  $p_{ij} \asymp e^{su_{ij}}$ . In this problem we will compute the expected number of steps to reach  $B \subset L$  from an initial state  $i \in D = L \setminus B$ . Let  $n_B$  be the smallest  $n$  for which  $X_n \in B$  and  $m_B(i) = E_i[n_B]$ . In conventional algebra we have for  $i \in L \setminus B$ ,  $m_B(i) = \mathbf{1} + \sum_{j \notin B} p_{ij} m_B(j)$  where  $\mathbf{1} = [1]_{i \in D}$ . If  $P^D = [p_{ij}]_{i,j \in D}$  and  $m_B = [m_B(i)]_{i \in D}$  then  $m_B = \mathbf{1} + P^D m_B$ . Using Lemma 1 as in Problem 1 and 2 the conventional equation converts to a max-plus equation, provided  $m_B(i) \asymp e^{sz_i}$  and  $p_{ij} = e^{su_{ij}}$ . So we get  $z = \mathbf{e} \oplus [U^D \otimes z]$ , where  $\mathbf{e}$  is the max-plus vector of all zeros. By Theorem 7 on page 31 since  $U$  has circuits of average weight less than or equal to  $\mathbf{e}$  then one solution is  $z = (U^D)^* \otimes \mathbf{e}$ . But, as in Problem 1 and Problem 2, the solution  $z$  found using Theorem 7 is not unique unless the maximal average circuit in  $U^D$  is negative. Therefore that solution may not be the correct solution corresponding to the asymptotic description of  $m_B$ . We will consider Freidlin and Wentzell's graph theoretical approach to solving this problem (in the conventional sense) and find the max-plus analogue of their results. Here is the result from [27](Lemma 3.4 of Chapter 6).

**Lemma 30.** *Given an irreducible Markov chain on a finite state space  $L$  the expected number of steps until the first entrance into  $B \subset L$  starting at  $i \in D$  is given by*

$$m_B(i) = \frac{\sum_{g \in G(B \cup \{i\})} \pi(g) + \sum_{i \in D, j \neq i} \sum_{g \in G_{ij}(B \cup \{j\})} \pi(g)}{\sum_{g \in G(B)} \pi(g)}$$

Note that the numerator in the equation for  $m_B(i)$  contains graphs without cycles and  $d-1$  edges  $(m, n)$ ,  $m \in, n \in L, m \neq n$  such that the sequence of edges does not contain a path from  $i$  into  $B$ . From this lemma we now derive the following max-plus theorem that gives the correct solution to  $z = \mathbf{e} \oplus [U^D \otimes z]$  that corresponds to  $m_B(i)$  above.

**Theorem 20.** *Under the same hypotheses as Theorem 19 with  $m_B(i) \asymp e^{sz_i}$ ,*

$$z_i = [\bigoplus_{g \in G(B \cup \{i\})} \rho(g)] \oplus [\bigoplus_{i \in D, j \neq i} \bigoplus_{g \in G_{ij}(B \cup \{j\})} \rho(g)] - \bigoplus_{g \in G(B)} \rho(g)$$

is the expected number of steps until the first entrance into  $B \subset L$  starting at  $i \in D$  where  $z = [z_i]_{i \in D}$  satisfies  $z = \mathbf{e} \oplus [U^D \otimes z]$ .

*Proof.* As in the Theorem 19 we just apply Lemma 1 to Lemma 30. □

Now let's look at a few examples that demonstrate the previous result for Problem 3.

**Example 11.** Consider again the example of Figure 3.1 with  $B = \{3, 4\}$ .

First we see that

$$\bigoplus_{g \in G(B \cup \{1\})} \rho(g) = u_{21} = -1 \text{ and } \bigoplus_{g \in G(B \cup \{2\})} \rho(g) = u_{12} = 0.$$

Next we consider the  $G_{ij}(B \cup \{j\})$ :

$$\bigoplus_{g \in G_{12}(B \cup \{2\})} \rho(g) = u_{12} = 0 \text{ and } \bigoplus_{g \in G_{21}(B \cup \{1\})} \rho(g) = u_{21} = -1.$$

Since  $\bigoplus_{g \in G(B)} \rho(g) = -9$  we conclude that

$$z = \left[ \begin{pmatrix} -1 \\ 0 \end{pmatrix} \oplus \begin{pmatrix} 0 \\ -1 \end{pmatrix} \right] - \begin{pmatrix} -9 \\ -9 \end{pmatrix} = \begin{pmatrix} 9 \\ 9 \end{pmatrix}.$$

We check that  $z$  satisfies the equation  $z = \mathbf{e} \oplus [U_1^D \otimes z]$ :

$$z = \begin{pmatrix} 9 \\ 9 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \oplus \left[ \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix} \otimes \begin{pmatrix} 9 \\ 9 \end{pmatrix} \right].$$

**Example 12.** Consider the example of figure 3.2 with  $B = \{4, 5, 6\}$ .

First we see that:

$$\begin{aligned}\bigoplus_{g \in G(B \cup \{1\})} \rho(g) &= u_{21} + u_{32} = -4 \\ \bigoplus_{g \in G(B \cup \{2\})} \rho(g) &= u_{12} + u_{32} = -9 \\ \bigoplus_{g \in G(B \cup \{3\})} \rho(g) &= u_{12} + u_{23} = -4.\end{aligned}$$

Next we consider the  $G_{ij}(B \cup \{j\})$ .

For  $i = 1$  we have

$$\bigoplus_{g \in G_{12}(B \cup \{2\})} \rho(g) = u_{12} + u_{34} = -9 \text{ and } \bigoplus_{g \in G_{13}(B \cup \{3\})} \rho(g) = u_{21} + u_{23} = -4.$$

For  $i = 2$  we have

$$\bigoplus_{g \in G_{21}(B \cup \{1\})} \rho(g) = u_{21} + u_{32} = -5 \text{ and } \bigoplus_{g \in G_{23}(B \cup \{3\})} \rho(g) = u_{21} + u_{23} = -4.$$

For  $i = 3$  we have

$$\bigoplus_{g \in G_{31}(B \cup \{1\})} \rho(g) = u_{32} + u_{21} = -5 \text{ and } \bigoplus_{g \in G_{32}(B \cup \{2\})} \rho(g) = u_{32} + u_{12} = -9.$$

Finally we see that  $\bigoplus_{g \in G(B)} \rho(g) = -11$ . Therefore we conclude that

$$z = \left[ \begin{pmatrix} -5 \\ -9 \\ -4 \end{pmatrix} \oplus \begin{pmatrix} -4 \\ -4 \\ -5 \end{pmatrix} \right] - \begin{pmatrix} -11 \\ -11 \\ -11 \end{pmatrix} = \begin{pmatrix} 7 \\ 7 \\ 7 \end{pmatrix}.$$

We find that  $z$  does indeed satisfy the required max-plus equation:

$$z = \begin{pmatrix} 7 \\ 7 \\ 7 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \oplus \left[ \begin{pmatrix} 0 & -4 & -\infty \\ -0 & 0 & 0 \\ -\infty & -5 & 0 \end{pmatrix} \otimes \begin{pmatrix} 7 \\ 7 \\ 7 \end{pmatrix} \right].$$

# Chapter 4

## Concluding Remarks

This thesis has given the reader a survey of max-plus algebra which focuses on the linear algebra aspects of max-plus algebra. We also described how we could apply these results to several results by Wentzell and Freidlin for finite state Markov chains with an asymptotic dependence. We would like to consider directions for future investigation of max-plus. Below we describe a few possibilities.

1. One obvious direction is the issue we came across in Problem 1 of Chapter 3. We would like to try and develop a nice max-plus characterization of the stationary distribution produced in Theorem 17.
2. Professor Marianne Akian, who gave us some insight into our issue in Problem 1, has developed a theory of max-plus probabilities motivated by dynamical programming and large deviations [12]. We might want to explore her research and continue in this direction with our work in max-plus algebra.
3. The authors in [3] briefly mentioned the max-plus analogues of Brownian motion and probability laws such as the central limit theorem. This is an intriguing avenue to pursue further.

4. As previously mentioned this thesis focused on the finite dimensional linear algebra results. Another path to take is to explore the infinite dimensional max-plus algebra. Above we briefly discussed this and what we called *max-plus function analysis*. We may consider proceeding in this direction and the path Professor McEaney has taken with max-plus algebra.

# Bibliography

- [1] M Akian R. Bapat, and S. Gaubert, *Asymptotics of the perron eigenvalue and eigenvector using max-algebra* C. R. Acad. Sci. Paris, t.327, Seire 1,(1998) pp. 927–932 .
- [2] M.Akian, R. Bapat, S. Gaubert, *Max-plus algebras*, Chapter 25 in the HANDBOOK OF LINEAR ALGEBRA, L. Hogben, R. Brualdi, A. Greenbaum, and R. Mathias (editors), Discrete Mathematics and Its Applications, **39**, Chapman and Hall, 2007.
- [3] F. Baccelli, G. Cohen, G.L. Olsder and J.P Quadrat, *SYNCRONIZATION AND LINEARITY*, Wiley, New York, 1992.
- [4] R.B. Bapat, David P. Standford, P. Van Den Driessche, *Pattern, properties and spectral inequalities in max algebra*, SIAM. J Matrix Anal Appl. **16** No 3 (1985), pp. 964-976.
- [5] J.G. Braker and G.J. Olsder, *The power algorithm in the max algebra*, Linear Algebra and its Applications. **182** (1993) pp. 67-89.
- [6] Bart De Schutter, Bart De Moor *A note on the Characteristic Equation in the Max-Plus Algebra*, Linear Albebra and its Applications. **261** (1997), pp. 237-250.
- [7] Bart De Shutter, *On the ultimate behavior of the sequence of consecutive powers of a matrix in the max-plus algebra*, Linear Algebra and it Applications. **30** (2000), pp. 103-117.
- [8] R.A. Cuninghame-Green and P Butkovic, *Bases in max-algebra*, Linear Algebra and its Applications, **389** (2004), pp. 104-120.

- [9] R.A. Cuninghame-Green, MINIMAX ALGEBRA LECTURE NOTES IN ECONOMICS AND MATHEMATICAL SYSTEMS. **166** Springer, 1979.
- [10] Wendell H. Fleming and William M. McEneaney, *A max-plus based algorithm for a Hamilton-Jacobi-Bellman equation of nonlinear filtering*, SIAM J. Control Optim. **38** (2000), pp. 683–710.
- [11] Wendell H. Fleming and H. Mete Soner, CONTROLLED MARKOV PROCESSES AND VISCOSITY SOLUTIONS, Springer-Verlag, New York, 1993.
- [12] Gaubert,S, *Methods and applications of (max,+) linear algebra*, STACS 1997, Lecture Notes in Computer Science **500**, Springer-Verlag, Berlin, (1997), pp. 261-282.
- [13] M. Gondran, and M.Minoux, *Linear algebra in dioids: A survey of recent results*, Annals of Discrete Mathematics, **19** (1984), pp. 147-164.
- [14] M.Gondran and M.Minoux, *L'independance lineaire dans les dioides*, Bulletin De La Direction Des Etudes et Recherches Series C. **1**, (1978), pp. 67-90.
- [15] R.G. Halburd, N.J. Southall, *Tropical nevanlinna theory and ultra -discrete equations*, Loughborough University, 2007.
- [16] B. Heidergot, G.J.Olsder, J.Woude, MAX PLUS AT WORK, Princeton Universty Press, New Jersey, 2006.
- [17] T.W. Hungerford, ALGEBRA: GRADUATE TEXTS IN MATHEMATICS, Springer-Verlag, New York, 1974.
- [18] S.C. Kleene. *Representation of events in nerve sets and finite automata*, Automata Studies, Annals of Mathematical Studies. No.34 (1956), pp. 3-41.
- [19] G.L. Litvinov, V.P. Maslov and G.B. Shipz, *Idempotent functional analysis: an algebraic approach*, Mathematical Notes. **69(5)**: (2001), pp. 696-729.



- [20] William M. McEneaney, MAX-PLUS METHODS FOR NONLINEAR CONTROL AND ESTIMATION, Birkhäuser, Boston, 2006.
- [21] William M. McEneaney, *Max-plus eigenvector representations for solution of nonlinear  $H_\infty$  problems: basic concepts*, IEEE Trans. Automat. Control **48** (2003), pp. 1150–1163.
- [22] William M. McEneaney, *Max-plus eigenvector methods for nonlinear  $H$  problems: error analysis*, SIAM J. Control Optim. **43** (2004), pp. 379–412.
- [23] William M. McEneaney, *A curse-of-dimensionality-free numerical method for solution of certain HJB pdes*, SIAM J. Control Optim. **46** (2007), pp. 1239–1276.
- [24] G.J. Olsder, *On the Characteristic Equation and Minimal Realizations for Discrete-event Dynamic Systems*, Lecture Notes in Control and Inform. Sci. **83** (1986), pp.189-2001.
- [25] G.J. Olsder and C. Roos, *Cramer and Cayley-Hamilton in the max-algebra*, Linear Algebra and its Applications. **101** (1988), pp. 87-108.
- [26] E. Wagneur, *Moduloids and pseudomodules.1.dimesion theory*, Discrete Math. **98**(1991), pp. 57-73.
- [27] M. I. Freidlin and A. D. Wentzell, RANDOM PERTURBATIONS OF DYNAMICAL SYSTEMS, Springer-Verlag, New York, 1984.

**List of Figures:**

- $\oplus, \otimes$ : pg 8
- $E$ : pg 12
- $G(A)$ : pg 14 item  $\|p\|_l$ : pg 14
- $\|p\|_w$ : pg 14
- $\frac{\|c\|_w}{\|c\|_l}$ : pg 15
- $A^{\otimes k}$ : pg 15
- $P(i, j, k)$ : pg 15
- $A^*$ : pg 16
- $A^+$ : pg 16
- $\text{perm}(A)$ : pg 23
- $\text{dom}(A)$ : pg 23
- $\begin{pmatrix} \Delta_1(A) \\ \Delta_2(A) \end{pmatrix}$ : pg 26
- $C(A)$ : pg 35
- $G^c(A)$ : pg 35
- $A_\lambda$ : pg 35
- $V^c(a)$ : pg 37
- $S_{\lambda v}(A)$ : pg 39
- $V_\lambda(A)$ : pg 42

- $A_{\sigma\sigma}$  : pg 45
- $G(B)$  : pg 64
- $\pi(g)$  : pg 64
- $\rho(g)$  : pg 64
- $G_{ij}(B)$  : pg 73