Université de Liège
Faculté des Sciences
Département de Mathématique

Relations associées aux matrices de Parikh

Mémoire de fin d'études réalisé par Delphine Wintgens en vue de l'obtention du diplôme de Master en Sciences Mathématiques

Promoteur : Michel Rigo
Année académique 2016-2017

Remerciements

Ce mémoire de fin d'études marque la fin de mon master en sciences mathématiques à finalité didactique et n'aurait pas pu aboutir sans la conjugaison des efforts de plusieurs personnes. Je souhaite adresser mes remerciements aux personnes qui m'ont apporté leur soutien ainsi que leur aide et qui ont donc contribué à l'élaboration de mon mémoire.

Je voudrais d'abord remercier mon promoteur, le Professeur Michel Rigo, et son assistante, doctorante à l'Université de Liège, Madame Manon Stipulanti, pour leur aide précieuse et pour le temps qu'ils ont bien voulu me consacrer tout au long de la réalisation de mon mémoire.

Je tiens également à remercier Monsieur Gaëtan Brouwers qui m'a encouragé ainsi que donné de précieux conseils et Monsieur Benoît Balthazard qui a pris le temps de relire l'entièreté de mon mémoire.

Table des matières

Introduction

D'une manière générale, la combinatoire des mots s'intéresse aux propriétés des suites de symboles appartenant à un alphabet. Par exemple, nous listons fréquemment les facteurs d'un mot afin d'obtenir certaines informations sur la structure de ce mot. En outre, nous pouvons étendre la notion de facteurs en considérant les sous-mots c'est-à-dire les sous-suites d'un mot donné. Ainsi, l'étude des sous-mots d'un mot donné fournit certains renseignements sur ce dernier. Par exemple, un problème classique de reconstruction qui a été largement étudié consiste à construire un mot à partir d'un certain nombre de ces sous-mots.

Ce mémoire mêle la combinatoire des mots et l'algèbre linéaire. En effet, à tout mot fini sur un alphabet ordonné constitué de k lettres, nous associons une matrice de dimension $(k+1)\times(k+1)$ à coefficients entiers appelée matrice de Parikh. Par exemple, sur l'alphabet ordonné $\Sigma = \{a, b, c\}$, la matrice de Parikh associée au mot w se révèle être :

$$\begin{pmatrix} 1 & |w|_a & |w|_{ab} & |w|_{abc} \\ 0 & 1 & |w|_b & |w|_{bc} \\ 0 & 0 & 1 & |w|_c \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

où $|w|_u$ (u un mot sur Σ) compte le nombre de sous-mots de w égaux à u. Les entiers $|w|_a, |w|_b$ et $|w|_c$ forment ce que nous nommons classiquement le vecteur de Parikh de w sur Σ.

Ce travail s'articule essentiellement autour de l'article d'Atanasiu, *Binary amiable words* [1], de l'article de Salomaa, *Criteria for the matrix equivalence of words* [11], et de l'article de Şerbănuţă, *Injectivity of the Parikh matrix mappings revisited* [13]. De plus, il se découpe en sept chapitres décrits brièvements ci-dessous.

Dans le premier chapitre, nous donnons quelques définitions et notions de base indispensables à la compréhension de ce travail.

Dans le deuxième chapitre, nous définissons le nombre d'occurrences et nous montrons que ce nombre est une généralisation du coefficient binomial d'entiers. Une fois cette généralisation établie, nous pouvons définir le vecteur de Parikh ainsi que la matrice de Parikh dont la forme est étudiée en détails. Ce chapitre se termine avec des propriétés relatives au nombre d'occurrences, dont l'inégalité de Cauchy sur les mots : $|w|_{xyz}|w|_y \leq |w|_{xy}|w|_{yz}$ avec w, x, y, z des mots sur un alphabet ordonné quelconque.

Deux mots ayant le même vecteur de Parikh sont bien évidemment des anagrammes (équivalence abélienne). Ainsi, forts de cette constatation, nous nous intéressons dans le troisième chapitre à un raffinement de l'équivalence abélienne et nous disons que deux mots sont équivalents s'ils possèdent la même matrice de Parikh. Plusieurs méthodes sont introduites afin de trouver rapidement des mots équivalents.

Dans le quatrième chapitre, nous introduisons des matrices particulières appelées matrices à signes alternés et matrices retournées. Nous montrons le lien existant entre ces matrices étant donné que ce lien nous sera utile dans le chapitre suivant.

Dans le cinquième chapitre, nous affirmons qu'un mot w est ambigu s'il existe un mot distinct de w ayant la même matrice de Parikh que w. Premièrement, si nous travaillons sur un alphabet ordonné constitué de deux lettres, nous montrons qu'il est facile de déterminer si un mot est ambigu ou non ambigu. Deuxièmement, nous insistons sur le fait qu'il s'avère plus difficile (nous ne disposons pas, à ce jour, de critères) de déterminer si un mot est ambigu ou non ambigu en travaillant sur un alphabet ordonné constitué de plus de deux lettres. Néanmoins, nous donnons plusieurs méthodes afin de trouver des mots ambigus ou non ambigus.

Dans le sixième chapitre, nous introduissons des règles de réécriture qui permettent d'obtenir un raffinement de l'équivalence de Parikh à savoir l'équivalence élémentaire de Parikh. Lorsque nous travaillons sur un alphabet ordonné constitué de deux lettres, nous obtenons une information supplémentaire étant donné que deux mots sont équivalents si et seulement si ces mots sont élémentairement équivalents.

Dans le septième chapitre, nous nous demandons sous quelles conditions une matrice de Parikh nous permet de déterminer uniquement un mot. Lorsque nous travaillons sur un alphabet ordonné constitué de deux lettres, nous pouvons remarquer qu'il nous suffit de posséder le nombre d'occurrences de certains mots particuliers et le vecteur de Parikh afin de déterminer de manière unique un mot. Lorsque nous travaillons sur un alphabet ordonné quelconque, nous utilisons les projections et la notion d'ambiguïté pour déterminer un mot unique.

Chapitre 1

Définitions et notions de base

Il est primordial de donner quelques définitions et notions de base afin de pouvoir comprendre les différents concepts qui vont être analysés au cours de ce travail. Ce chapitre est consacré à l'introducion de ces définitions et notions qui vont être réparties en quatre sections intitulées "Coefficient binomial", "Mot et longueur de mot", "Concaténation" et "Restriction d'un mot". Ces quatres sections ont été respectivement écrites sur base des articles [4], [5], [5] et [7].

1.1 Coefficient binomial

Pour commencer, nous allons nous intéresser aux coefficients binomiaux. Plus précisément, nous allons définir ces coefficients ainsi qu'énoncer la formule de Pascal [1] liant certains de ces coefficients.

Définition 1.1.1. Si n et k sont des entiers tels que $0 \leq k \leq n$, le *coefficient binomial (d'entiers)*, noté comme suit :

$$\binom{n}{k},$$

compte le nombre de choix possibles de k éléments parmi n éléments sans tenir compte de l'ordre de ces élements et est défini comme suit [2] :

$$\frac{n!}{k!(n-k)!}.$$

Nous allons voir qu'il existe un lien entre les différents coefficients binomiaux. Ce lien est représenté par la formule de Pascal que nous allons énoncer ci-desous.

1. Blaise Pascal est un mathématicien, physicien et écrivain français né en 1623 et mort en 1662.
2. La factorielle d'un naturel n, notée $n!$, est le produit des nombres naturels inférieurs à n. Par exemple, la factorielle du naturel 4 est $4 \cdot 3 \cdot 2 \cdot 1 = 24$.

Théorème 1.1.2. *Si n et k sont des entiers tels que $1 \leq k \leq n$, nous pouvons écrire la relation suivante connue sous le nom de "Formule de Pascal" :*

$$\binom{n}{k-1} + \binom{n}{k} = \binom{n+1}{k}.$$

1.2 Alphabet, mot et longueur d'un mot

Tout au long de ce mémoire, nous allons travailler avec des mots. Ainsi, il s'avère indispensable de définir la notion de mots ainsi que la notion d'alphabets sans laquelle nous ne pourrions pas définir la première notion. De plus, chaque mot possède une caractéristique bien particulière à savoir sa longueur que nous allons également définir.

Définition 1.2.1. Un *alphabet* est un ensemble fini de symboles dont les éléments sont appelés lettres. On représente un alphabet par la lettre grecque majuscule Σ.

Définition 1.2.2. Un *mot* est une suite, finie ou infinie, de lettres. On note Σ^* l'ensemble des mots finis sur l'alphabet Σ et $\Sigma^{\mathbb{N}}$ l'ensemble des mots infinis sur l'alphabet Σ.

Définition 1.2.3. La *longueur* d'un mot fini w, notée $|w|$, est le nombre de lettres qui le constituent.

Exemple 1.2.4. Si nous prenons l'alphabet ordonné $\Sigma = \{a, b, c\}$, $aabca$ est un mot sur Σ et la longueur de ce mot vaut 5.

Remarque 1.2.5. Une lettre est un mot de longueur 1.

Remarque 1.2.6. Le mot vide, noté ε, est l'unique mot de longueur 0.

1.3 Concaténation

Il est intéressant de remarquer que nous pouvons juxtaposer plusieurs mots afin de construire un mot de longueur supérieure aux autres [3] d'où la définition suivante.

Définition 1.3.1. Soient u et v deux mots finis sur l'alphabet Σ. La *concaténation* des mots u et v, notée $u \cdot v = uv$, est la juxtaposition de ces deux mots. Formellement, si $u = a_1 \cdots a_l$ et $v = b_1 \cdots b_k$ avec $a_i, b_j \in \Sigma \ \forall \ i, j \in \mathbb{N}$ et $l, k \in \mathbb{N}$, nous obtenons $w = uv = c_1 \cdots c_{l+k}$ avec :

$$\begin{cases} c_i = a_i & \forall i \in \{1, \ldots, l\}, \\ c_i = b_{i-l} & \forall i \in \{l+1, \ldots, l+k\}. \end{cases}$$

Exemple 1.3.2. Si $\Sigma = \{a, b\}$, $u = aaba$ et $v = bba$, alors nous obtenons le mot $w = aababba$ en effectuant la concaténation des mots u et v.

Remarque 1.3.3. Le mot vide est le neutre pour la concaténation.

3. Cette définition est évidente vu la Définition 1.2.2 et la Remarque 1.2.5.

1.4 Restriction d'un mot

Vu la Définition 1.3.1, nous savons que nous pouvons juxtaposer plusieurs mots afin d'en construire un nouveau de longueur supérieure. Mais, il peut être intéressant de construire un mot de longueur inférieure à un mot initial plutôt que de longueur supérieure. C'est pourquoi, nous allons définir la notion de sous-mots et la notion de facteurs.

Définition 1.4.1. Si w est un mot sur l'alphabet Σ, *un sous-mot* u de w est un mot obtenu en supprimant certaines lettres de w.

Exemple 1.4.2. Si $\Sigma = \{a, b, c\}$, le mot abc est un sous-mot de $abbbacbba$.

Définition 1.4.3. Si w, u, x et y sont des mots sur l'alphabet Σ tels que $w = xuy$, alors u est un *facteur* de w. En d'autres termes, un *facteur* est un mot qui se trouve dans le mot initial ou encore, un sous-mot formé de lettres qui s'avèrent être consécutives dans le mot initialement choisi.

Exemple 1.4.4. Si $\Sigma = \{a, b\}$ et $w = abbaa$, le mot $u = ba$ est un facteur de w.

Remarque 1.4.5. Vu la Définition 1.4.1 et la Définition 1.4.3, tout facteur est un sous-mot. Cependant, la réciproque est fausse. En effet, si nous prenons $\Sigma = \{a, b\}$ et $w = abbaa$, le mot $u = aba$ s'avère être un sous-mot de w, mais pas un facteur de w étant donné que les lettres comprises dans le mot u ne sont pas des lettres apparaissant dans le mot w de manière consécutives.

Chapitre 2

Le nombre d'occurrences

Dans ce chapitre, basé essentiellement sur l'article [7], nous allons nous intéresser au nombre d'occurrences sur les mots. Plus précisément, nous allons d'abord définir ce nombre ainsi que le lien existant entre celui-ci et le coefficient binomial d'entiers. Ensuite, nous introduirons le vecteur de Parikh et la matrice de Parikh tous deux définis à l'aide du nombre d'occurrences. Enfin, nous finirons ce chapitre en citant quelques propriétes liées au nombre d'occurrences, dont l'inégalité de Cauchy (sur les mots).

Définition 2.0.1. Soient k un entier strictement positif et $\Sigma = \{e_1, \ldots, e_k\}$ un alphabet. L'alphabet Σ est un *alphabet ordonné* si ces lettres sont telles que $e_1 < \ldots < e_k$.

Définition 2.0.2. Soient $\Sigma = \{e_1, \ldots, e_k\}$ un alphabet ordonné et u, w des mots sur l'alphabet Σ. Si w s'écrit $w_1 \cdots w_{|w|}$ avec $w_1, \ldots, w_{|w|} \in \Sigma$, alors une *occurrence* de u dans w est une suite strictement croissante d'entiers $n_1 < \ldots < n_{|u|}$ compris entre 1 et $|w|$ tel que $w_{n_1} \cdots w_{n_{|u|}} = u$.

Remarque 2.0.3. Soient $\Sigma = \{e_1, \ldots, e_k\}$ un alphabet ordonné et u, w des mots sur l'alphabet Σ. Les occurrences du mot u dans le mot w peuvent être représentées sous forme vectorielle. En effet, nous pouvons représenter une occurrence du mot u dans le mot w par un vecteur de la forme $(i_1, \ldots, i_{|u|})$ où l'entier j est tel que $1 \leq j \leq |u|$, où l'entier i_j est tel que $1 \leq i_j \leq |w|$ et où la j^{e} lettre de u se révèle être la i_j^{e} lettre de w.

Exemple 2.0.4. Si $\Sigma = \{a, b\}$, $u = ba$ et $w = abbaa$, alors les quatre occurences du mot u dans le mot w sont $(2, 4), (2, 5), (3, 4)$ et $(3, 5)$.

Définition 2.0.5. Soient $\Sigma = \{e_1, \ldots, e_k\}$ un alphabet ordonné et u, w des mots sur l'alphabet Σ. Le *nombre d'occurences* du mot u dans le mot w, noté $|w|_u$, est le nombre de fois que le mot u apparaît dans le mot w en tant que sous-mot de w. Par convention, nous notons $|w|_\varepsilon = 1$.

Exemple 2.0.6. Si $\Sigma = \{a, b\}$, $u = ba$ et $w = abbaa$, alors le nombre d'occurrences du mot u dans le mot w vaut 4.

Proposition 2.0.7. *Soient $\Sigma = \{e_1, \dots, e_k\}$ un alphabet ordonné et u, w des mots sur l'alphabet Σ. Si $|w| < |u|$, le nombre d'occurrences du mot u dans le mot w est nul et, si $|w| = |u|$, nous pouvons exprimer le nombre d'occurrences comme suit :*

$$|w|_u = \begin{cases} 0 & si \quad w \neq u, \\ 1 & si \quad w = u. \end{cases}$$

Démonstration. Premièrement, si $|w| < |u|$, le mot u ne peut pas apparaître dans le mot w. Par conséquent, le nombre d'occurrences de u dans w est nul. Deuxièmement, si $|w| = |u|$, les mots u et w sont soit égaux soit différents. S'ils sont égaux, nous en déduisons que $|w|_u = 1$ et, s'ils sont différents, nous en déduisons que $|w|_u = 0$.

□

2.1 Lien avec les coefficients binomiaux

Nous allons montrer qu'il existe un lien (expliqué dans l'article [10]) entre le coefficient binomial et le nombre d'occurrences. Plus précisément, nous allons montrer que le nombre d'occurrences est une généralisation du coefficient binomial.

Remarque 2.1.1. Soit $\Sigma = \{e_1\}$ un alphabet constitué d'une lettre. Les mots sur l'alphabet Σ sont le mot vide et les mots e_1^i, $i \in \mathbb{N}_0$. Par convention, nous notons ε comme étant égal à e_1^0.

Proposition 2.1.2. *Soit $\Sigma = \{e_1\}$ un alphabet constitué d'une lettre et u, w des mots sur l'alphabet Σ. Si $|w| \geq |u|$, nous obtenons ce qui suit :*

$$|w|_u = \binom{|w|}{|u|}.$$

Démonstration. Étant donné que nous travaillons sur l'alphabet constitué uniquement de la lettre e_1, nous pouvons supposer que le mot w est égal à e_1^n et que le mot u est égal à e_1^m avec m et n deux entiers positifs. Par conséquent, nous savons que la longueur de w vaut n, que la longueur de u vaut m et que n est supérieur [1] à m. De plus, vu la Définition 2.0.5, le nombre d'occurrences du mot u dans le mot w s'avère être le nombre de choix possibles de m éléments parmi n éléments ce qui nous permet de conclure vu la Définition 1.1.1.

□

Vu la proposition précédente, nous pouvons généraliser le coefficient binomial à l'aide du nombre d'occurrences. Pour cette raison, le nombre d'occurrences peut aussi s'appeler le coefficient binomial de mots.

1. Par hypothèse, la longueur de w est supérieure à la longueur de n.

2.2 Vecteur de Parikh

Grâce à la Définition 2.0.5 et la Remarque 1.2.5, nous sommes en mesure de dire que le nombre d'occurences d'une lettre a dans un mot w est notée $|w|_a$. À l'aide de cette notation, nous pouvons définir le vecteur de Parikh.

Définition 2.2.1. Soit $\Sigma = \{e_1, \ldots, e_k\}$ un alphabet ordonné. La *fonction de Parikh* est la fonction $\psi : \Sigma^* \to \mathbb{N}^k$ qui à un mot w associe son vecteur de Parikh. Le *vecteur de Parikh* de w, noté $\psi(w)$, est le vecteur suivant :

$$\psi(w) = \begin{pmatrix} |w|_{e_1} \\ \vdots \\ |w|_{e_k} \end{pmatrix}.$$

Exemple 2.2.2. Soient $\Sigma = \{a, b, c\}$ un alphabet ordonné et $w = abcab$ un mot sur l'alphabet Σ. Le vecteur de Parikh du mot w est le suivant :

$$\psi(w) = \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}.$$

Comme nous allons le voir à l'aide du théorème suivant, certains mots peuvent posséder le même vecteur de Parikh si nous travaillons sur un alphabet ordonné constitué de plus de deux lettres.

Théorème 2.2.3. *Soient k un entier tel que $k \geq 2$ et $\Sigma = \{e_1, \ldots, e_k\}$ un alphabet ordonné. La fonction de Parikh n'est pas injective.*

Démonstration. Nous procédons par l'absurde en supposant que la fonction de Parikh est injective. Nous pouvons facilement trouver un contre-exemple à cette affirmation. Effectivement, si nous prenons l'alphabet ordonné $\Sigma = \{a, b\}$ ainsi que les mots *abbaba* et *baabba*, nous remarquons que ces mots sont différents et qu'ils possèdent le même vecteur de Parikh à savoir :

$$\begin{pmatrix} 3 \\ 3 \end{pmatrix}.$$

Ainsi, la fonction de Parikh qui à un mot de Σ^* associe son vecteur de Parikh n'est pas injective.

□

2.3 Matrice de Parikh

Le vecteur de Parikh nous renseigne uniquement sur le nombre d'occurences des lettres dans un mot. Cependant, nous pouvons aussi nous intéresser au nombre d'occurrences de certains mots dans un mot. C'est pourquoi nous allons introduire la matrice de Parikh qui est une généralisation du vecteur de Parikh.

Définition 2.3.1. Soient $\Sigma = \{e_1, \dots, e_k\}$ un alphabet ordonné, $\mathbb{N}_{|u|+1}^{|u|+1}$ l'ensemble des matrices carrées de dimension $(|u|+1)$ à coefficients dans $\mathbb{N}$ et $u = b_1 \cdots b_{|u|}$ un mot sur l'alphabet Σ ($b_1, \dots, b_{|u|} \in \Sigma$). L'application ψ_u est le morphisme de monoïdes [2] :

$$\psi_u : (\Sigma^*, \cdot, \varepsilon) \to (\mathbb{N}_{|u|+1}^{|u|+1}, \cdot, I_{|u|+1}).$$

qui à un mot w associe sa *matrice de Parikh* induite par le mot u sur l'alphabet Σ et qui est défini par la condition suivante :

$$\text{Si } a \in \Sigma \text{ et } \psi_u(a) = (m_{i,j})_{1 \leq i,j \leq (|u|+1)}, \text{ alors } m_{i,j} = \begin{cases} 1 \text{ si } j = i, \\ \delta_{b_i,a} \text{ si } j = i+1, \\ 0 \text{ sinon.} \end{cases}$$

Remarque 2.3.2. L'application ψ_u est définie partout sur Σ^*. En effet, vu la Définition 2.3.1, nous savons que ψ_u est un morphisme et nous connaissons les valeurs prises par ce morphisme en les lettres de Σ.

Remarque 2.3.3. Si nous remplaçons u par $e_1 \cdots e_k$ dans la Définition 2.3.1, nous notons ψ_k au lieu de ψ_u. De plus, le morphisme de monoïdes $\psi_k : \Sigma^* \to N_{k+1}^{k+1}$ qui à un mot w associe sa matrice de Parikh [3], est défini par la condition suivante :

$$\text{Si } q \in \{1, \dots, k\} \text{ et } \psi_k(e_q) = (m_{i,j})_{1 \leq i,j \leq (k+1)}, \text{ alors on a :}$$

$$m_{i,j} = \begin{cases} 1 \text{ si } i = j, \\ 1 \text{ si } i = q \text{ et } j = i+1, \\ 0 \text{ sinon.} \end{cases}$$

2. Il est évident que $(\mathbb{N}_{|u|+1}^{|u|+1}, \cdot, I_{|u|+1})$ et $(\Sigma^*, \cdot, \varepsilon)$ sont des monoïdes. Effectivement, leurs opérations sont binaires, internes, partout définies ainsi qu'associatives et les ensembles Σ^* et $\mathbb{N}_{|u|+1}^{|u|+1}$ admettent tout deux un neutre à savoir ε pour Σ^* et la matrice identité pour $\mathbb{N}_{|u|+1}^{|u|+1}$. De plus, l'opération $\cdot$ dans le monoïde $(\mathbb{N}_{|u|+1}^{|u|+1}, \cdot, I_{|u|+1})$ est le produit matriciel, tandis que l'opération $\cdot$ dans le monoïde $(\Sigma^*, \cdot, \varepsilon)$ est la concaténation.

3. Lorsque le mot u est égal au mot $e_1 \cdots e_k$, nous ne précisons pas que la matrice de Parikh est induite par u.

Forme des matrices de Parikh

Vu la Définition 2.3.1, nous savons déterminer la forme des matrices $\psi_u(e_l)$ avec u un mot quelconque sur l'alphabet ordonné $\Sigma = \{e_1, \ldots, e_k\}$ et e_l une lettre de Σ. Cependant, cette définition ne nous apprend rien concernant la forme des matrices $\psi_u(w)$ avec w un mot quelconque sur l'alphabet Σ. Ainsi, pour déterminer la forme de telles matrices, il est nécessaire d'introduire une propriété concernant le nombre d'occurrences à savoir la proposition suivante.

Proposition 2.3.4. *Soit $\Sigma = \{e_1, \ldots, e_k\}$ un alphabet ordonné. Pour tous mots $u, w \in \Sigma^*$ et pour toutes lettres $a, b \in \Sigma$, nous pouvons écrire l'identité suivante :*

$$|wb|_{ua} = |w|_{ua} + \delta_{a,b}|w|_u.$$

Démonstration. Si les lettres a et b sont égales, nous avons l'égalité entre $|wb|_{ua}$ et $|wa|_{ua}$. Ainsi, la dernière lettre du mot ua est soit la dernière lettre du mot wa soit une lettre apparaissant dans le mot w. Si le dernière lettre de ua et la dernière lettre de wa coïncident, le mot u est un sous-mot de w et, si la dernière lettre de ua apparaît dans w, le mot ua est un sous-mot de w. Par contre, si les lettres a et b diffèrent, le mot ua ne peut être qu'un sous-mot de w. En regroupant ces deux configurations à savoir le cas où les lettres a et b sont égales et le cas où ces lettres diffèrent, nous obtenons le résultat attendu. □

Théorème 2.3.5. *Soient $\Sigma = \{e_1, \ldots, e_k\}$ un alphabet ordonné, $u = b_1 \cdots b_{|u|}$ un mot sur l'alphabet Σ ($b_1, \ldots, b_{|u|} \in \Sigma$) et w un mot sur l'alphabet Σ. La matrice de Parikh du mot w induite par le mot u sur Σ, notée $\psi_u(w) = (m_{i,j})_{1 \leq i,j \leq (|u|+1)}$, possède les propriétés suivantes[4] :*

- *$m_{i,i} = 1$ pour tout entier i tel que $1 \leq i \leq (|u| + 1)$,*
- *$m_{i,j} = 0$ pour tous les entiers i, j tels que $1 \leq j < i \leq (|u| + 1)$,*
- *$m_{i,j+1} = |w|_{u_{i,j}}$ pour tous les entiers i, j tels que $1 \leq i \leq j \leq |u|$.*

En d'autres mots, la matrice de Parikh du mot w induite par le mot u sur Σ est une matrice carrée triangulaire supérieure de dimension $(|u| + 1)$ étant représentée comme suit :

$$\psi_u(w) = \begin{pmatrix} 1 & |w|_{b_1} & |w|_{b_1 b_2} & \cdots & \cdots & |w|_{b_1 \cdots b_{|u|}} \\ 0 & \ddots & |w|_{b_2} & \cdots & \cdots & |w|_{b_2 \cdots b_{|u|}} \\ \vdots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & |w|_{b_{|u|}} \\ 0 & \cdots & \cdots & \cdots & 0 & 1 \end{pmatrix}.$$

4. Les mots $u_{i,j}$, $1 \leq i \leq j \leq (|u|+1)$, sont égaux aux mots $b_i \cdots b_j$ et les mots $u_{i,j}$, $1 \leq j < i \leq (|u|+1)$, sont tous égaux au mot vide.

Démonstration. Nous notons $w = a_1 \cdots a_n$ avec $a_1, \ldots, a_n \in \Sigma$ et $n \in \mathbb{N}$. Pour démontrer ce théorème, nous allons procéder par récurrence sur la longueur du mot w, à savoir n. Si n vaut 0, nous avons l'égalité entre $\psi_u(w)$ et $I_{|u|+1}$ étant donné que le mot w est égal au mot vide [5]. Le théorème est donc satisfait pour $n = 0$, ce qui signifie que le cas de base est vérifié.

Nous supposons que le théorème est vrai pour $n < m$ ($m \in \mathbb{N}$) et nous allons vérifier qu'il reste vrai pour $n = m$.

Par hypothèse, nous avons l'égalité entre u et $b_1 \cdots b_{|u|}$. De plus, si nous notons $w' = a_1 \cdots a_{m-1}$ et $a_m = e_l$ avec e_l une lettre de Σ, nous obtenons l'égalité entre $\psi_u(w)$ et $\psi_u(w')\psi_u(e_l)$ vu la définition d'un morphisme. Ainsi, vu la définition du produit matriciel, les éléments de la matrice $\psi_u(w)$ sont donnés par [6]

$$m_{i,j} = \begin{cases} \sum_{m=1}^{|u|+1} (\psi_u(w'))_{i,m}(\psi_u(e_l))_{m,j} \text{ si } j > i, \\ 1 \text{ si } i = j, \\ 0 \text{ sinon.} \end{cases}$$

Grâce à la forme des éléments $m_{i,j}$ donnée ci-dessus, nous observons directement l'égalité entre $m_{i,i}$ et 1 pour tout entier i tel que $1 \leq i \leq (|u| + 1)$ ainsi que l'égalité entre $m_{i,j}$ et 0 pour tous les entiers i, j tels que $1 \leq j < i \leq (|u| + 1)$. Par conséquent, il nous reste à montrer l'égalité entre $m_{i,j+1}$ et $|w|_{u_{i,j}}$ pour tous les entiers i, j tels que $1 \leq i \leq j \leq |u|$. Pour montrer cette dernière égalité, nous allons réexprimer les éléments $m_{i,j}$, $1 \leq i < j \leq (|u| + 1)$, en utilisant l'hypothèse de récurrence [7] sur w' :

$$(\psi_u(w'))_{i,m} = \begin{cases} 0 \text{ si } i > m, \\ 1 \text{ si } i = m, \\ |w'|_{u_{i,m-1}} \text{ sinon,} \end{cases}$$

ainsi que la Définition 2.3.1 :

$$(\psi_u(e_l))_{m,j} = \begin{cases} 1 \text{ si } m = j, \\ \delta_{b_m, e_l} \text{ si } j = m + 1, \\ 0 \text{ sinon.} \end{cases}$$

5. Voir la Définition 2.3.1

6. Vu que ψ_u est un morphisme et vu que les matrices $\psi_u(a_j)$, $1 \leq j \leq n$, sont des matrices triangulaires supérieures, nous obtenons :

$$(\psi_u(w))_{i,i} = \prod_{j=1}^{n} (\psi_u(a_j))_{i,i}.$$

Ainsi, vu la Définition 2.3.1, nous obtenons directement l'égalité entre $m_{i,i}$ et 1 pour tout entier i compris entre 1 et $|u| + 1$. De plus, vu que le produit de matrices triangulaires supérieures est lui-même une matrice triangulaire supérieure, nous obtenons l'égalité entre $m_{i,j}$ et 0 pour tous les entiers i et j tels que $1 \leq j < i \leq (|u| + 1)$.

7. Le mot w' est de longueur inférieure au mot w. Ainsi, nous pouvons appliquer l'hypothèse de récurrence sur w'.

Vu la définition du produit matriciel, nous pouvons facilement en déduire ce qui suit :

$$\sum_{m=1}^{|u|+1}(\psi_u(w'))_{i,m}(\psi_u(e_l))_{m,j} = \delta_{b_{j-1},e_l}\delta_{i,j-1} + |w'|_{u_{i,j-1}} + |w'|_{u_{i,j-2}}\delta_{b_{j-1},e_l}\chi_{i<j-1}.$$

Grâce à l'égalité ci-dessus, nous remarquons que les éléments $m_{i,j+1}$, $1 \leq i \leq j \leq (|u|+1)$, sont égaux [8] aux éléments $|w'|_{u_{i,j-1}}\delta_{b_j,e_l} + |w'|_{u_{i,j}}$ et vu la Proposition [9] 2.3.4, nous pouvons conclure que les éléments $m_{i,j+1}$ et $|w|_{u_{i,j}}$ sont égaux pour tous les entiers i, j tels que $1 \leq i \leq j \leq |u|$.

□

Si le mot u est égal au mot $e_1 \cdots e_k$ dans le Théorème 2.3.5, nous obtenons un cas particulier de ce théorème à savoir le Corollaire 2.3.6.

Corollaire 2.3.6. *Soient $\Sigma = \{e_1, \cdots, e_k\}$ un alphabet ordonné et w un mot sur l'alphabet Σ. La matrice de Parikh du mot w, notée $\psi_k(w) = (m_{i,j})_{1\leq i,j\leq (k+1)}$, possède les propriétés suivantes [10] :*

- $m_{i,j} = 0$ *pour tous les entiers i, j tels que $1 \leq j < i \leq (k+1)$,*
- $m_{i,i} = 1$ *pour tout entier i tel que $1 \leq i \leq (k+1)$,*
- $m_{i,j+1} = |w|_{e_{i,j}}$ *pour tous les entiers i, j tels que $1 \leq i \leq j \leq k$.*

Démonstration. Nous procédons de manière analogue à la démonstration du Théorème 2.3.5 en remplacant le morphisme de monoïdes ψ_u par le morphisme de monoïdes ψ_k.

□

Remarque 2.3.7. Vu le Corollaire 2.3.6 et la Définition 2.2.1, il est évident que nous pouvons obtenir le vecteur de Parikh d'un mot à partir de la matrice de Parikh de ce mot. En effet, le i^{e} élément du vecteur de Parikh est l'élément se trouvant à la i^{e} ligne et à la $(i+1)^{\text{e}}$ colonne de la matrice de Parikh.

Injectivité de la matrice de Parikh

Comme nous allons le voir à l'aide de la proposition suivante, certains mots peuvent posséder la même matrice de Parikh induite par un mot quelconque.

Proposition 2.3.8. *Soient k un entier tel que $k \geq 2$, $\Sigma = \{e_1, \ldots, e_k\}$ un alphabet ordonné et u un mot sur l'alphabet Σ. La fonction ψ_u n'est pas injective.*

8. Si i est égal à j, nous avons l'égalité entre $\delta_{b_j,e_l}\delta_{i,j}$ et δ_{b_j,e_l} ainsi que l'égalité entre $|w'|_{u_{i,j-1}}\delta_{b_j,e_l}\chi_{i<j}$ et $\delta_{b_j,e_l}\chi_{i<j}$. Ainsi, nous avons également l'égalité entre $m_{i,j+1}$ et $|w'|_{u_{i,j-1}}\delta_{b_j,e_l} + |w'|_{u_{i,j}}$ si $1 \leq i \leq j \leq (|u|+1)$.

9. Vu la définition des mots $u_{i,j}$,$1 \leq i \leq j \leq (|u|+1)$, et vu l'égalité entre w et $w'e_l$, nous pouvons facilement appliquer la Proposition 2.3.4.

10. Les mots $e_{i,j}$, $1 \leq i \leq j \leq (k+1)$, sont égaux aux mots $e_ie_{i+1}\cdots e_j$ et les mots $e_{i,j}$, $1 \leq j < i \leq (k+1)$, sont tous égaux au mot vide.

Démonstration. Nous procédons par l'absurde en supposant que l'application ψ_u est injective. Nous pouvons facilement trouver un contre-exemple à cette affirmation. Effectivement, si nous travaillons sur l'alphabet ordonné $\Sigma = \{a, b\}$ et si nous considérons le mot $u = ab$, les mots $w_1 = abbaba$ et $w_2 = baabba$ diffèrent. Cependant, les matrices de Parikh de w_1 et w_2 induites par $u = ab$ sur Σ sont identiques et égales à la matrice suivante :

$$\begin{pmatrix} 1 & 3 & 4 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix}.$$

Ainsi, la fonction ψ_u qui à un mot de Σ^* associe sa matrice de Parikh n'est pas injective ce qui conclut la preuve.

□

2.4 Inégalité de Cauchy

Dans cette section, nous allons nous intéresser à une propriété bien particulière du nombre d'occurrences à savoir l'inégalité de Cauchy (sur les mots). Cependant, avant de vérifier cette inégalité, il est nécessaire d'introduire le Théorème 2.4.1.

Théorème 2.4.1. *Soient $\Sigma = \{e_1, \dots, e_k\}$ un alphabet ordonné, w un mot sur l'alphabet Σ et u un mot sur l'alphabet Σ tel que $u = b_1 \cdots b_{|u|}$ $(b_1, \dots, b_{|u|} \in \Sigma)$. Chaque mineur*[11] *d'ordre 2 de la matrice de Parikh de w induite par u est un entier positif.*

Démonstration. Pour démontrer ce théorème, nous allons procéder par récurrence sur la longueur de w. Si la longueur du mot w vaut 0, ce mot est égal au mot vide et la matrice de Parikh $\psi_u(w)$ est égale[12] à la matrice $I_{|u|+1}$. Par conséquent, tous les mineurs d'ordre 2 de la matrice $\psi_u(w)$ valent 0 ou 1. Le théorème est donc satisfait pour $|w| = 0$ ce qui signifie que le cas de base est vérifié.

Nous supposons que le théorème est vrai pour $|w| < m$ $(m \in \mathbb{N}_0)$ et nous allons vérifier qu'il reste vrai pour $|w| = m$.

Dans la suite de cette démonstration, nous notons $w = w'e_i$ avec w' un mot de longueur $m - 1$ sur l'alphabet Σ et e_i une lettre de Σ. Étant donné que ψ_u est un morphisme (de monoïdes), nous avons l'égalité entre $\psi_u(w)$ et $\psi_u(w')\psi_u(e_i)$. Ainsi, vu le Théorème 2.3.5 et vu la définition du produit matriciel, les éléments $[\psi_u(w)]_{p,q}$, $1 \leq p \leq |u| + 1$ et $1 \leq q \leq |u| + 1$, s'écrivent comme suit[13] :

11. Soient $A = (a_{i,j})_{1 \leq i,j \leq n}$ une matrice carrée de dimension n et r un entier tel que $1 \leq r \leq n$. Un mineur d'ordre r de A est le déterminant d'une sous-matrice carrée de A de dimension r.

12. Cette égalité de matrices provient de la Définition 2.3.1

13. Vu le Théorème 2.3.5, nous savons que $[\psi_u(w')]_{p,q}$ est égal à 1 si p est égal à q et que $[\psi_u(w')]_{p,q}$ est égal à 0 si p est strictement supérieur à q.

$$\begin{aligned}
[\psi_u(w)]_{p,q} &= \sum_{n=1}^{|u|+1} [\psi_u(w')]_{p,n}[\psi_u(e_i)]_{n,q},\\
&= \begin{cases} [\psi_u(w')]_{p,q-1}\delta_{b_{q-1},e_i} + [\psi_u(w')]_{p,q} \text{ si } p < q,\\ [\psi_u(w')]_{p,q} \text{ si } p \geq q,\end{cases}
\end{aligned}$$

ce qui nous permet de dire que les matrices $\psi_u(w)$ et $\psi_u(w')$ sont toutes les deux des matrices triangulaires supérieures dont les éléments diagonaux valent 1. Afin de conclure cette preuve, nous notons D une sous-matrice carrée de dimension 2 obtenue à partir de $\psi_u(w)$ et nous envisageons plusieurs cas.

Tout d'abord, si la matrice D ne contient pas d'éléments de la matrice $\psi_u(w)$ se trouvant au dessus de la diagonale principale de $\psi_u(w)$, la matrice D est une sous-matrice de $\psi_u(w')$. Comme le mot w' est de longueur inférieure au mot w, nous pouvons appliquer l'hypothèse de récurrence à la matrice de Parikh de w' induite par u ce qui nous permet de conclure [14].

Ensuite, si la matrice D contient un seul élément de la matrice $\psi_u(w)$ se trouvant au dessus de la diagonale principale de $\psi_u(w)$, la matrice D se représente comme suit :

$$\begin{pmatrix} [\psi_u(w')_{p,p} & [\psi_u(w')]_{p,q} + [\psi_u(w')]_{p,q-1}\delta_{b_{q-1},e_i} \\ [\psi_u(w')]_{q,p} & [\psi_u(w')]_{q,q} \end{pmatrix}$$

où p et q sont des entiers tels que $1 \leq p < q \leq |u| + 1$. Ainsi, vu le Théorème 2.3.5, le déterminant de la matrice D vaut 1 ce qui permet de conclure.

En outre, si la matrice D contient trois éléments de la matrice $\psi_u(w)$ se trouvant au dessus de la diagonale principale de $\psi_u(w)$, la matrice D se représente comme suit :

$$\begin{pmatrix} [\psi_u(w')]_{p,q} + [\psi_u(w')]_{p,q-1}\delta_{b_{q-1},e_i} & [\psi_u(w')]_{p,m} + [\psi_u(w')]_{p,m-1}\delta_{b_{m-1},e_i} \\ [\psi_u(w')]_{q,q} & [\psi_u(w')]_{q,m} + [\psi_u(w')]_{p,m-1}\delta_{b_{m-1},e_i} \end{pmatrix}$$

où m, p, q sont des entiers tels que $1 \leq p < q \leq |u| + 1$, $1 \leq p < m \leq |u| + 1$ et $1 \leq q < m \leq |u| + 1$. Dans cette situation, trois cas sont à envisager. Premièrement, si la lettre e_i diffère des lettres b_{q-1} et b_{m-1}, la matrice D est une sous-matrice de $\psi_u(w')$. Comme le mot w' est de longueur inférieure au mot w, nous pouvons appliquer l'hypothèse de récurrence à la matrice de Parikh de w' induite par u ce qui nous permet de conclure.

14. Vu que la matrice D ne contient pas d'éléments de la matrice $\psi_u(w)$ se trouvant au dessus de la diagonale principale de $\psi_u(w)$ et vu le Théorème 2.3.5, cette matrice est soit la matrice nulle soit la matrice constitué de 0 partout excepté dans son coin supérieur droit (dont la valeur s'avère être 1). Ainsi, son déterminant vaut toujours 0.

Deuxièmement, si la lettre e_i diffère de la lettre b_{q-1} (resp. b_{m-1}) et si la lettre e_i est égale à la lettre b_{m-1} (resp. b_{q-1}), nous pouvons en déduire que le déterminant de la matrice D est égal à la somme du déterminant de la matrice D_1 et du déterminant de la matrice D_2 où D_1 et D_2 sont respectivement les matrices suivantes :

$$\begin{pmatrix} [\psi_u(w')]_{p,q} & [\psi_u(w')]_{p,m} \\ [\psi_u(w')]_{q,q} & [\psi_u(w')]_{q,m} \end{pmatrix} \text{ (resp. } \begin{pmatrix} [\psi_u(w')]_{p,q} & [\psi_u(w')]_{p,m} \\ [\psi_u(w')]_{q,q} & [\psi_u(w')]_{q,m} \end{pmatrix} \text{)}$$

$$\begin{pmatrix} [\psi_u(w')]_{p,q} & [\psi_u(w')]_{p,m-1} \\ [\psi_u(w')]_{q,q} & [\psi_u(w')]_{q,m-1} \end{pmatrix} \text{ (resp. } \begin{pmatrix} [\psi_u(w')]_{p,q-1} & [\psi_u(w')]_{p,m} \\ [\psi_u(w')]_{q,q-1} & [\psi_u(w')]_{q,m} \end{pmatrix} \text{)}$$

Les matrices D_1 et D_2 sont des sous-matrices carrées de $\psi_u(w')$ étant de dimension 2. Étant donné que le mot w' est de longueur inférieure au mot w, nous pouvons appliquer l'hypothèse de récurrence à la matrice de Parikh de w' induite par u et en déduire que les déterminants de D_1 et D_2 sont des entiers positifs. Par conséquent, le déterminant de la matrice D est également un entier positif. Troisièmement, si les lettres e_i, b_{m-1} et b_{q-1} sont égales, nous pouvons en déduire que le déterminant de la matrice D est égal à la somme du déterminant de la matrice D_1, du déterminant de la matrice D_2, du déterminant de la matrice D_3 et du déterminant de la matrice D_4 où les matrices D_1, D_2, D_3 et D_4 sont respectivement représentées comme suit :

$$\begin{pmatrix} [\psi_u(w')]_{p,q} & [\psi_u(w')]_{p,m} \\ [\psi_u(w')]_{q,q} & [\psi_u(w')]_{q,m} \end{pmatrix}$$

$$\begin{pmatrix} [\psi_u(w')]_{p,q} & [\psi_u(w')]_{p,m-1} \\ [\psi_u(w')]_{q,q} & [\psi_u(w')]_{q,m-1} \end{pmatrix}$$

$$\begin{pmatrix} [\psi_u(w')]_{p,q-1} & [\psi_u(w')]_{p,m} \\ [\psi_u(w')]_{q,q-1} & [\psi_u(w')]_{q,m} \end{pmatrix}$$

$$\begin{pmatrix} [\psi_u(w')]_{p,q-1} & [\psi_u(w')]_{p,m-1} \\ [\psi_u(w')]_{q,q-1} & [\psi_u(w')]_{q,m-1} \end{pmatrix}$$

Les matrices D_1, D_2, D_3 et D_4 sont des sous-matrices carrées de $\psi_u(w')$ étant de dimension 2. Étant donné que le mot w' est de longueur inférieure au mot w, nous pouvons appliquer l'hypothèse de récurrence à la matrice de Parikh de w' induite par u et en déduire que les déterminants de D_1, D_2, D_3 et D_4 sont des entiers positifs. Par conséquent, le déterminant de la matrice D est également un entier positif.

Enfin, si la matrice D contient quatre éléments de la matrice $\psi_u(w)$ se trouvant au dessus de la diagonale principale de $\psi_u(w)$, la matrice D se représente comme suit :

$$\begin{pmatrix} [\psi_u(w')]_{p,q} + [\psi_u(w')]_{p,q-1}\delta_{b_{q-1},e_i} & [\psi_u(w')]_{p,m} + [\psi_u(w')]_{p,m-1}\delta_{b_{m-1},e_i} \\ [\psi_u(w')]_{l,q} + \psi_u(w')_{l,q-1}\delta_{b_{q-1},e_i} & [\psi_u(w')]_{l,m} + [\psi_u(w')]_{l,m-1}\delta_{b_{m-1},e_i} \end{pmatrix}$$

où p, q, m, l sont des entiers tels que $1 \leq p < q \leq |u| + 1$, $1 \leq p < m \leq |u| + 1$, $1 \leq l < q \leq |u| + 1$ et $1 \leq l < m \leq |u| + 1$. Dans cette situation trois cas sont à nouveau à envisager. Ces cas se traitent de manière analogue à la situation précédente ce qui conclut cette preuve.

$\square$

Le Théorème 2.4.1 nous permet de démontrer l'inégalité de Cauchy pour les mots à savoir la Proposition 2.4.2.

Proposition 2.4.2. *Soit $\Sigma = \{e_1, \ldots, e_k\}$ un aphabet ordonné. Pour tous mots $x, y, z, w \in \Sigma^*$, l'inégalité suivante est satisfaite et s'appelle "Inégalité de Cauchy pour les mots" :*

$$|w|_{xyz}|w|_y \leq |w|_{xy}|w|_{yz}.$$

Démonstration. Premièrement, nous allons considérer le cas où le mot z est égal au mot vide. Dans ce cas, nous avons l'égalité entre $|w|_{xyz}|w|_y$ et $|w|_{xy}|w|_y$ ainsi que l'égalité entre $|w|_{xy}|w|_{yz}$ et $|w|_{xy}|w|_y$. Par conséquent, nous relevons également l'égalité entre $|w|_{xyz}|w|_y$ et $|w|_{xy}|w|_{yz}$ ce qui signifie que l'inégalité de Cauchy est satisfaite si le mot z est égal au mot vide.

Deuxièmement, nous allons considérer le cas où le mot z diffère du mot vide. Dans ce cas, nous posons $u = xyz$ et $\psi_u(w) = (m_{i,j})_{1 \leq i,j \leq |u|+1}$. Vu le Théorème 2.3.6, nous pouvons en déduire l'égalité entre $m_{1,|u|+1}$ et $|w|_{xyz}$, l'égalité entre $m_{|x|+1,|xy|+1}$ et $|w|_y$, l'égalité entre $m_{1,|xy|+1}$ et $|w|_{xy}$ ainsi que l'égalité entre $m_{|x|+1,|xyz|+1}$ et $|w|_{yz}$. Par conséquent, la matrice suivante est une sous-matrice carrée de dimension 2 de la matrice de Parikh de w induite par le mot u sur l'alphabet Σ :

$$\begin{pmatrix} |w|_{xy} & |w|_{xyz} \\ |w|_y & |w|_{yz} \end{pmatrix}.$$

Vu le Théorème 2.4.1, le déterminant de cette matrice est un entier positif c'est-à-dire que l'entier $|w|_{xy}|w|_{yz} - |w|_{xyz}|w|_y$ est supérieur à 0 d'où l'inégalité de Cauchy sur les mots.

$\square$

Exemple 2.4.3. Si nous travaillons sur l'alphabet ordonné $\Sigma = \{a, b\}$ et si nous prenons les mots $x = a$, $y = b$, $z = ba$ et $w = abba$, nous obtenons :

- $|w|_{xyz} = |abba|_{abba} = 1$,
- $|w|_y = |abba|_b = 2$,
- $|w|_{xy} = |abba|_{ab} = 2$,
- $|w|_{yz} = |abba|_{bba} = 1$.

Ainsi, le membre de gauche de l'inégalité de Cauchy décrite dans la Proposition 2.4.2 vaut 2 et le membre de droite de cette inégalité vaut également 2. L'inégalité de Cauchy est donc satisfaite.

Une fois l'inégalité de Cauchy définie, nous pouvons généraliser cette inégalité à l'aide du théorème suivant.

Théorème 2.4.4. *Soient $\Sigma = \{e_1, \ldots, e_k\}$ un alphabet ordonné et m un entier tel que $m \geq 1$. Pour tous mots $w, x, y_1, \ldots, y_m, z \in \Sigma^*$, l'inégalité suivante est satisfaite :*

$$|w|_{y_1} \cdots |w|_{y_m} |w|_{xy_1 \cdots y_m z} \leq |w|_{xy_1} |w|_{y_1 y_2} \cdots |w|_{y_{m-1} y_m} |w|_{y_m z}.$$

Démonstration. Nous procédons par récurrence sur m. Si m vaut 1, nous obtenons l'inégalité de Cauchy vue à la Proposition 2.4.2. Par conséquent, le théorème est satisfait si m vaut 1 ce qui signifie que le cas de base est vérifié.

Nous supposons que le théorème est vrai pour $m \leq t$ ($t \in \mathbb{Z}_0$) et nous allons montrer qu'il reste vrai pour $t+1$.

Étant donné que y_t et y_{t+1} sont des mots sur l'alphabet Σ, leur concaténation est encore [15] un mot sur l'alphabet Σ. Ainsi, par hypothèse de récurrence, nous en déduisons ce qui suit :

$$|w|_{y_1} \cdots |w|_{y_{t-1}} |w|_{y_t y_{t+1}} |w|_{xy_1 \cdots y_t y_{t+1} z} \leq |w|_{xy_1} |w|_{y_1 y_2} \cdots |w|_{y_{t-1} y_t y_{t+1}} |w|_{y_t y_{t+1} z}.$$

En multipliant les deux membres de l'inégalité par $|w|_{y_t} |w|_{y_{t+1}}$ et en appliquant deux fois l'inégalité de Cauchy dans le membre de droite, nous obtenons :

$$|w|_{y_1} \cdots |w|_{y_{t+1}} |w|_{xy_1 \cdots y_t y_{t+1} z} |w|_{y_t y_{t+1}} \leq |w|_{xy_1} |w|_{y_1 y_2} \cdots |w|_{y_{t-1} y_t} |w|_{y_t y_{t+1}} |w|_{y_t y_{t+1}} |w|_{y_{t+1} z}$$

En simplifiant $|w|_{y_t y_{t+1}}$ dans les deux membres de l'inégalité ci-dessus, nous obtenons le résultat attendu.

□

Exemple 2.4.5. Si nous travaillons sur l'alphabet ordonné $\Sigma = \{a, b\}$ et, si nous considérons les mots $x = a$, $y_1 = b$, $y_2 = ab$, $z = ba$ et $w = abbaab$, nous obtenons :

- $|w|_{y_1} = |abbaab|_b = 3$,
- $|w|_{y_2} = |abbaab|_{ab} = 5$,
- $|w|_{xy_1 y_2 z} = |abbaab|_{ababba} = 0$,
- $|w|_{xy_1} = |abbaab|_{ab} = 5$,
- $|w|_{y_1 y_2} = |abbaab|_{bab} = 4$,
- $|w|_{y_2 z} = |abbaab|_{abba} = 2$.

Ainsi, le membre de gauche de l'inégalité décrite dans le Théorème 2.4.4 vaut 0 et le membre de droite de cette inégalité vaut 40. L'inégalité est donc bien satisfaite.

15. Voir la Définition 1.3.1.

Dans la proposition suivante, nous allons montrer une inégalité qui ressemble fort à l'inégalité de Cauchy décrite dans la Proposition 2.4.2. En effet, l'inégalité que nous allons démontrer s'obtient en inversant le mot w avec les mots xyz, xy, yz et y.

Proposition 2.4.6. *Soit $\Sigma = \{e_1, \ldots, e_k\}$ un alphabet ordonné. Pour tous mots $w, x, y, z \in \Sigma^*$, l'inégalité suivante est satisfaite :*

$$|xyz|_w |y|_w \leq |xy|_w |yz|_w.$$

Démonstration. La proposition s'avère être évidente dans le cas où xyz, y, xy, yz ou w est égal au mot vide ainsi que dans le cas où le nombre d'occurrences de w dans xyz, y, xy ou yz est nul. Ainsi, dans la suite de cette démonstration, nous allons nous intéresser au cas où xyz, y, xy, yz et w ne sont pas égaux au mot vide et où le nombre d'occurrences de w dans xyz, y, xy et yz est non nul. Dans ce cas, nous notons w_u une réprésentation vectorielle [16] de l'occurrence de w dans u et nous allons montrer qu'il existe une fonction f injective, définie dans l'ensemble des couples de la forme (w_y, w_{xyz}), pour laquelle l'image du couple (w_y, w_{xyz}) est le couple [17] (w_{xy}, w_{yz}).

Afin de démontrer ce théorème, nous notons $w_y = (j_1, \ldots, j_{|w|})$ avec $j_1, \ldots, j_{|w|} \in \{1, \ldots, |y|\}$ et nous traitons dans un premier temps les cinq cas suivants :

- Si $w_{xyz} = (i_1, \ldots, i_{|w|})$ où $i_1, \ldots, i_{|w|} \in \{1, \ldots, |x|\}$, alors l'image du couple (w_y, w_{xyz}) par la fonction f s'avère être le couple [18] (w_{xyz}, w_y).
- Si $w_{xyz} = (i_1, \ldots, i_{|w|})$ où $i_1, \ldots, i_{|w|} \in \{|xy|+1, \ldots, |xyz|\}$, alors l'image du couple (w_y, w_{xyz}) par la fonction f s'avère être le couple (w_y, w_{xyz}).
- Si $w_{xyz} = (i_1, \ldots, i_{|w|})$ où $i_1, \ldots, i_{|w|} \in \{|x|+1, \ldots, |xy|\}$, alors l'image du couple (w_y, w_{xyz}) par la fonction f s'avère être le couple (w_{xyz}, w_y) [19].
- Si $w_{xyz} = (i_1, \ldots, i_{|w|})$ où $i_1, \ldots, i_{|w|} \in \{1, \ldots, |xy|\}$, alors l'image du couple (w_y, w_{xyz}) par la fonction f s'avère être le couple (w_{xyz}, w_y).
- Si $w_{xyz} = (i_1, \ldots, i_{|w|})$ où $i_1, \ldots, i_{|w|} \in \{|x|+1, \ldots, |xyz|\}$, alors l'image du couple (w_y, w_{xyz}) par la fonction f s'avère être le couple (w_y, w_{xyz}).

16. Voir la Remarque 2.0.3.
17. Par définition de l'injectivité, une fonction allant de l'ensemble des couples de la forme (w_y, w_{xyz}) vers l'ensemble des couples de la forme (w_{xy}, w_{yz}) est injective si et seulement si tout élément de la forme (w_{xy}, w_{yz}) est l'image d'au plus un élément de la forme (w_y, w_{xyz}). Par conséquent, le nombre d'éléments de la forme (w_y, w_{xyz}) est plus petit que le nombre d'éléments de la forme (w_{xy}, w_{yz}), c'est-à-dire que $|xyz|_w|y|_w \leq |xy|_w|yz|_w$.
18. En effet, vu la forme de w_y, nous pouvons en déduire une occurrence du mot w dans le mot xyz à savoir $(|x|+j_1, \ldots, |x|+j_r)$. Etant donné que les entiers $i_1, \ldots, i_{|w|}$ (resp. $|x|+j_1, \ldots, |x|+j_{|w|}$) appartiennent à l'ensemble $\{1, \ldots, |x|\}$ (resp. $\{|x|+1, \ldots, |xy|\}$), ces entiers appartiennent également à l'ensemble $\{1, \ldots, |xy|\}$ (resp. $\{|x|+1, \ldots, |xyz|\}$).
19. Nous devons choisir entre le couple (w_{xyz}, w_y) et le couple (w_y, w_{xyz}) sinon, la fonction n'est pas injective. Effectivement, si nous ne choisissons pas entre ces deux couples, les couples (w_y, w_{xyz}) et (w_{xyz}, w_y) sont tous les deux envoyés sur le couple (w_y, w_{xyz}).

Dans les cinq cas traités ci-dessus, un couple de la forme (w_{xy}, w_{yz}) est l'image par la fonction f d'au plus un couple[20] de la forme (w_y, w_{xyz}). Ainsi, la fonction f, définie sur l'ensemble des couples (w_y, w_{xyz}) considérés dans les cinq cas ci-dessus, est injective.

Après avoir considéré les cinq cas ci-dessus, il nous reste à traiter deux cas à savoir le cas où $w_{xyz} = (i_1, \ldots, i_{|w|})$ avec $i_1, \ldots, i_{|w|} \in \{1, \ldots, |x|, |xy| + 1, \ldots, |xyz|\}$ et le cas où $w_{xyz} = (i_1, \ldots, i_{|w|})$ avec $i_1, \ldots, i_{|w|} \in \{1, \ldots, |xyz|\}$.

Premièrement, nous allons considérer le cas où $w_{xyz} = (i_1, \ldots, i_{|w|})$ avec $i_1, \ldots, i_{|w|} \in \{1, \ldots, |x|, |xy| + 1, \ldots, |xyz|\}$. Dans ce cas, la première partie du mot w apparaît dans x et la seconde partie de w apparaît dans z. Ainsi, nous constatons ce qui suit[21] :

$$w_{xyz} = (i_1, \ldots, i_p, i_{p+1}, \ldots, i_r)$$

où $p \geq 1$, $r \geq 2$, $i_p \leq |x|$, $i_{p+1} > |xy|$ et où $i_r \leq |xyz|$. Nous pouvons également écrire une représentation vectorielle de l'occurrence du mot w dans le mot y comme suit :

$$w_y = (j_1, \ldots, j_p, j_{p+1}, \ldots, j_r)$$

où $j_r \leq |y|$, $p \geq 1$ et $r \geq 2$. À l'aide de cette dernière représentation, nous pouvons en déduire une autre représentation vectorielle de w dans xyz :

$$(|x| + j_1, \ldots, |x| + j_p, |x| + j_{p+1}, \ldots, |x| + j_r).$$

En notant $j'_l = |x| + j_l$ pour chaque entier l compris entre 1 et r, nous pouvons construire les représentations vectorielles suivantes :

$$\begin{aligned} w_{xy} &= (i_1, \ldots, i_p, j'_{p+1}, \ldots, j'_r), \\ w_{yz} &= (j'_1, \ldots, j'_p, i_{p+1}, \ldots, i_r). \end{aligned}$$

Vu ce qui précède, un couple de la forme (w_{xy}, w_{yz}) est l'image d'au plus un couple de la forme (w_y, w_{xyz}) où $w_{xyz} = (i_1, \ldots, i_{|w|})$ avec $i_1, \ldots, i_{|w|} \in \{1, \ldots, |x|, |xy| + 1, \ldots, |xyz|\}$. De plus, l'image d'un couple de la forme (w_y, w_{xyz}) où $w_{xyz} = (i_1, \ldots, i_{|w|})$ avec $i_1, \ldots, i_{|w|} \in \{1, \ldots, |x|, |xy|+1, \ldots, |xyz|\}$ ne saurait pas[22] être égale à un couple de la forme (w_y, w_{xyz}) ou à un couple de la forme (w_{xyz}, w_y). Ainsi, la fonction f, définie sur l'ensemble des couples de la forme (w_y, w_{xyz}) considérés dans les cinq premiers cas ainsi que dans ce cas-ci, est injective.

20. Les occurrences du mot w dans le mot xyz ne peuvent pas être identiques si nous sommes dans deux cas distincts et un couple de la forme (w_y, w_{xyz}) est envoyé soit sur lui même soit sur le couple obtenu en échangeant ses composantes selon le cas dans lequel nous nous trouvons. Ainsi, un couple de la forme (w_{xy}, w_{yz}) ne peut pas être l'image de deux couples de la forme (w_y, w_{xyz}) appartenant à des cas distincts et selon le cas dans lequel nous nous trouvons, un couple de la forme (w_{xy}, w_{yz}) est l'image d'au plus un couple de la forme (w_y, w_{xyz}).

21. Le nombre p est supérieur à 1 et le nombre r est supérieur à 2 sinon, nous pouvons nous ramener à un des cinq premiers cas. Ainsi, dans ce cas, le mot w est de longueur supérieure à 2.

22. Vu la forme de w_{xyz}, c'est évident.

Deuxièmement, nous allons traiter le cas où $w_{xyz} = (i_1, \ldots, i_{|w|})$ avec i_1, …, $i_{|w|} \in \{1, \ldots, |xyz|\}$. Dans ce cas, le début du mot w apparaît dans x, le milieu de w apparaît dans y et la fin de w apparaît dans z. Ainsi, nous pouvons écrire ce qui suit[23] :

$$w_{xyz} = (i_1, \ldots, i_p, i_{p+1}, \ldots, i_{p+q}, i_{p+q+1}, \ldots, i_r)$$

où $p \geq 1$, $q \geq 1$, $r \geq 3$, $i_p \leq |x|$, $i_{p+1} > |x|$, $i_{p+q} \leq |xy|$, $i_{p+q+1} > |xy|$ et où $i_r \leq |xyz|$. Nous pouvons également écrire une représentation vectorielle de l'occurrence du mot w dans le mot y comme suit :

$$w_y = (j_1, \ldots, j_p, j_{p+1}, \ldots, j_{p+q}, j_{p+q+1}, \ldots, j_r)$$

où $p \geq 1$, $q \geq 1$, $r \geq 3$. À l'aide de cette dernière représentation, nous pouvons en déduire une autre représentation vectorielle de l'occurrence de w dans xyz :

$$(|x| + j_1, \cdots, |x| + j_p, |x| + j_{p+1}, \ldots, |x| + j_{p+q}, |x| + j_{p+q+1}, \ldots, |x| + j_r).$$

Dans la suite de cette démonstration, nous notons $j'_l = |x| + j_l$ pour chaque entier l compris entre 1 et r. Vu la forme des deux réprésentations vectorielles de w_{xyz} données ci-dessus, nous savons qu'il existe un entier t compris entre $p+1$ et r tel que j'_t est inférieur[24] à i_t. Ainsi, nous considérons le plus petit entier t satisfaisant ces conditions et nous construisons les représentations vectorielles suivantes[25] :

$$\begin{aligned} w_{xy} &= (i_1, \ldots, i_{t-1}, j'_t, \ldots, j'_r), \\ w_{yz} &= (j'_1, \ldots, j'_{t-1}, i_t, \ldots, i_r). \end{aligned}$$

Vu ce qui précède, un couple de la forme (w_{xy}, w_{yz}) est l'image d'au plus un couple de la forme (w_y, w_{xyz}) où $w_{xyz} = (i_1, \ldots, i_{|w|})$ avec i_1, …, $i_{|w|} \in \{1, \ldots, |xyz|\}$. De plus, l'image d'un couple de la forme (w_y, w_{xyz}) où $w_{xyz} = (i_1, \ldots, i_{|w|})$ avec i_1, …, $i_{|w|} \in \{1, \ldots, |xyz|\}$ ne saurait pas[26] être égale à un couple de la forme (w_y, w_{xyz}), à un couple de la forme (w_{xyz}, w_y) ou à un couple de la forme (w_{xy}, w_{yz}) défini comme dans le cas précédent. Par conséquent, la fonction f, définie sur l'ensemble des couples de la forme (w_y, w_{xyz}) pour laquelle l'image du couple (w_y, w_{xyz}) s'avère être le couple (w_{xy}, w_{yz}), est injective ce qui nous permet de conclure.

□

23. Les nombres p et q sont supérieurs à 1 et le nombre r est supérieur à 3 sinon, nous pouvons nous ramener à un des six cas traités auparavant. Ainsi, dans ce cas, la longueur du mot w est supérieure à 3.

24. Les entiers i_l, $1 \leq l \leq p$, sont inférieurs à $|x|$ et les entiers j'_l, $1 \leq l \leq r$ sont strictement supérieurs à $|x|$. Ainsi, si l'entier t existe, il doit être supérieur à $p+1$. De plus, les entiers i_l, $p+q+1 \leq l \leq r$, sont supérieurs à $|xy|+1$ et les entiers j'_l, $1 \leq l \leq r$ sont strictement inférieurs à $|xy|+1$. Ainsi, il existe un entier t compris entre $p+1$ et r tel que j'_t est inférieur à i_t.

25. L'occurrence du mot w dans le mot yz est bien définie étant donné que j'_{t-1} est inférieur à j'_t et que j'_t est inférieur à i_t. De plus, l'entier i_{t-1} est strictement inférieur à j'_{t-1} et l'entier j'_{t-1} est inférieur à j'_t. Ainsi, l'occurrence du mot w dans le mot xy est bien définie.

26. Vu la forme de w_{xyz}, c'est évident.

Une généralisation du Théorème 2.4.6 s'avère être le théorème suivant.

Théorème 2.4.7. *Soient* $\Sigma = \{e_1, \ldots, e_k\}$ *un alphabet ordonné,* m *un entier tel que* $m \geq 1$. *Pour tous mots* $w, x, y_1, \ldots, y_m, z \in \Sigma^*$, *l'inégalité suivante est satisfaite :*

$$|y_1|_w \cdots |y_m|_w |xy_1 \cdots y_m z|_w \leq |xy_1|_w |y_1 y_2|_w \cdots |y_{m-1} y_m|_w |y_m z|_w.$$

Démonstration. Ce théorème se démontre de manière analogue au Théorème 2.4.4.

□

Chapitre 3

Mots équivalents

Vu la Proposition 2.3.8, nous savons que plusieurs mots peuvent être associés à une même matrice de Parikh. Dans cette section, nous allons nous intéresser à ces mots en introduisant la notion d'équivalence de Parikh.

Définition 3.0.1. Soit $\Sigma = \{e_1, \ldots, e_k\}$ un alphabet ordonné. Deux mots w, $w' \in \Sigma^*$ sont é*quivalents* s'ils possèdent la même matrice de Parikh [1]. Si deux mots w et $w' \in \Sigma^*$ sont équivalents, nous notons $w \equiv_k w'$.

Exemple 3.0.2. Soit l'alphabet ordonné $\Sigma = \{a, b\}$. Les mots *abbaba* et *baabba* sont équivalents vu la forme de leur matrice de Parikh (cfr Proposition 2.3.8).

Vu la Définition 3.0.1, nous pouvons facilement vérifier si des mots sont équivalents étant donné qu'il nous suffit de calculer leur matrice de Parikh et de vérifier que ces matrices sont égales. Cependant, il s'avère plus difficile de trouver des mots équivalents d'où l'utilité de ce chapitre. En effet, dans ce chapitre, nous allons voir plusieurs méthodes afin de trouver rapidement des mots équivalents.

3.1 Concaténation et mots équivalents

La première méthode ayant pour but de trouver rapidement des mots équivalents est décrite à l'aide du Théorème 3.1.2 (trouvé dans l'article [1]). Avant d'énoncer et de démontrer ce théorème, il est nécessaire d'introduire la proposition suivante vérifiant à l'aide de la forme des matrices de Parikh que ψ_u est un morphisme de monoïdes.

Proposition 3.1.1. *Soit $\Sigma = \{e_1, \ldots, e_k\}$ un alphabet ordonné. Pour tous mots $u, v, w \in \Sigma^*$, la relation suivante est satisfaite :*

$$\psi_u(vw) = \psi_u(v)\psi_u(w).$$

1. Les mots w et w' sont équivalents si nous avons l'égalité entre les matrices $\psi_k(w)$ et $\psi_k(w')$.

Démonstration. Vu le Théorème 2.3.5 et vu la définition du produit matriciel, nous pouvons en déduire ce qui suit :

$$[\psi_u(v)]_{i,l}[\psi_u(w)]_{l,j} = \begin{cases} 1 \text{ si } i = j, \\ |w|_{e_{i,j-1}}\delta_{i,l} + |v|_{e_{i,j-1}}\delta_{j,l} + |v|_{e_{i,l-1}}|w|_{e_{l,j-1}} \text{ si } i+1 \leq l \leq j-1, \\ 0 \text{ sinon,} \end{cases}$$

$$[\psi_u(vw)]_{i,j} = \begin{cases} 1 \text{ si } i = j, \\ |vw|_{e_{i,j-1}} \text{ si } i \leq j-1, \\ 0 \text{ sinon.} \end{cases}$$

Le mot $e_{i,j-1}$ (où i et j sont des entiers tels que $i \leq j-1$) peut apparaître de trois manières différentes dans le mot vw. En effet, le mot $e_{i,j-1}$ peut être un sous-mot de v, le mot $e_{i,j-1}$ peut être un sous-mot de w ou le début du mot $e_{i,j-1}$ peut être un sous-mot de v et la fin du mot $e_{i,j-1}$ un sous-mot de w. Ainsi, nous obtenons la relation suivante :

$$|vw|_{e_{i,j-1}} = |w|_{e_{i,j-1}} + |v|_{e_{i,j-1}} + \textstyle\sum_{l=i+1}^{j-1} |v|_{e_{i,l-1}}|w|_{e_{l,j-1}}.$$

Vu cette relation, nous pouvons en déduire l'égalité entre les éléments $[\psi_u(vw)]_{i,j}$, $1 \leq i \leq |u|+1$ et $1 \leq j \leq |u|+1$, de la matrice $\psi_u(vw)$ et les éléments $\sum_{l=1}^{|u|+1}[\psi_u(v)]_{i,l}[\psi_u(w)]_{l,j}$ de la matrice $\psi_u(v)\psi_u(w)$ ce qui permet de conclure la preuve.

□

Le théorème suivant nous dit que la relation d'équivalence de Parikh, notée $\equiv_k$, est une congruence. Ainsi, si nous avons l'équivalence de Parikh entre les mots α_1 et β_1 ainsi qu'entre les mots α_2 et β_2, nous pouvons facilement trouver d'autres mots équivalents à savoir le mot obtenu en concaténant α_1 et α_2 ainsi que le mot obtenu en concaténant β_1 et β_2.

Théorème 3.1.2. *Soit $\Sigma = \{e_1, \ldots, e_k\}$ un alphabet ordonné et α_1, α_2, β_1, β_2 des mots sur l'alphabet à Σ. Si les mots α_1 et β_1 sont équivalents et si les mots α_2 et β_2 sont équivalents, alors les mots $\alpha_1\alpha_2$ et $\beta_1\beta_2$ sont aussi équivalents.*

Démonstration. Vu la Définition 3.0.1 et la Proposition 3.1.1, la première affirmation est évidente.

3.2 Règles d'équivalence de Parikh

À l'aide de l'article [1] et de l'article [11], nous allons énoncer et démontrer deux règles, nommées règle 1 et règle 2. Ces règles nous permetteront de trouver rapidement plusieurs mots équivalents.

Théorème 3.2.1. ***Règle 1*** *: Soient* $\Sigma = \{e_1, \ldots, e_k\}$ *un alphabet ordonné,* w *un mot sur l'alphabet* Σ*,* i *un entier tel que* $1 \leq i \leq k-2$ *et* j *un entier tel que* $2 \leq j \leq k-i$*. Si le mot* $e_i e_{i+j}$ *(resp.* $e_{i+j}e_i$*) est un facteur de* w *et si ce facteur est remplacé par le facteur* $e_{i+j}e_i$ *(resp.* $e_i e_{i+j}$*) dans* w*, un mot équivalent à* w *en résulte.*

Démonstration. Nous notons $w = xe_ie_{i+j}y$ et $w' = xe_{i+j}e_iy$ avec x et y deux mots sur l'alphabet Σ. Pour démontrer ce théorème, nous allons montrer que les mots w et w' sont équivalents. Etant donné que ψ_k est un morphisme, nous avons :

$$\begin{aligned} \psi_k(w) &= \psi_k(x)\psi_k(e_i)\psi_k(e_{i+j})\psi_k(y), \\ \psi_k(w') &= \psi_k(x)\psi_k(e_{i+j})\psi_k(e_i)\psi_k(y). \end{aligned}$$

Ainsi, vu la Définition 3.0.1 et vu le Théorème 2.3.6, les mots w et w' sont équivalents si nous obtenons ce qui suit [2] :

$$\psi_k(e_i)\psi_k(e_{i+j}) = \psi_k(e_{i+j})\psi_k(e_i).$$

Vu la Remarque 2.3.3 et vu la définition de produit matriciel, la matrice $\psi_k(e_i)\psi_k(e_{i+j})$ est une matrice carrée triangulaire supérieure de dimension $(k+1)$ dont les éléments sont donnés ci-dessous :

$$\begin{aligned} (\psi_k(e_i)\psi_k(e_{i+j}))_{n,l} &= \begin{cases} \sum_{p=1}^{k+1}(\psi_k(e_i))_{n,p}(\psi_k(e_{i+j}))_{p,l} \text{ si } l > n, \\ 1 \text{ si } n = l, \\ 0 \text{ sinon,} \end{cases} \\ &= \begin{cases} \delta_{n,i}\delta_{l,i+1} + \delta_{n,i+j}\delta_{l,i+j+1} \text{ si } l > n, \\ 1 \text{ si } n = l, \\ 0 \text{ sinon.} \end{cases} \end{aligned}$$

Par un raisonnement analogue, les éléments de la matrice $\psi_k(e_{i+j})\psi_k(e_i)$ sont identiques aux éléments de la matrice $\psi_k(e_i)\psi_k(e_{i+j})$. Par conséquent, les deux mots sont équivalents.

□

Remarque 3.2.2. Si nous travaillons sur un alphabet ordonné constitué de deux lettres, nous ne pouvons pas appliquer la règle 1.

Remarque 3.2.3. Soit $\Sigma = \{e_1, \ldots, e_k\}$ un alphabet ordonné. Nous notons $w = xe_ie_{i+1}y$ et $w' = xe_{i+1}e_iy$ avec x, y deux mots sur l'alphabet Σ et i un entier tel que $1 \leq i \leq k-1$. Si les mots w et w' sont équivalents, la relation suivante devrait être satisfaite [3] :

$$\psi_k(e_i)\psi_k(e_{i+1}) = \psi_k(e_{i+1})\psi_k(e_i).$$

2. Vu le Théorème 2.3.6, nous pouvons en déduire que les matrices de Parikh sont inversibles.
3. Vu la Définition 3.0.1 et le fait que ψ_k est un morphisme, cette relation est satisfaite si les mots w et w' sont équivalents.

Or, vu la Remarque 2.3.3 et vu la définition du produit matriciel, nous avons ce qui suit :

$$
\begin{aligned}
(\psi_k(e_i)\psi_k(e_{i+1}))_{n,l} &= \begin{cases} \delta_{n,i}\delta_{l,i+1} + \delta_{n,i}\delta_{l,i+2} + \delta_{n,i+1}\delta_{l,i+2} \text{ si } l > n, \\ 1 \text{ si } n = l, \\ 0 \text{ sinon,} \end{cases} \\
(\psi_k(e_{i+1})\psi_k(e_i))_{n,l} &= \begin{cases} \delta_{n,i}\delta_{l,i+1} + \delta_{n,i+1}\delta_{l,i+2} \text{ si } l > n, \\ 1 \text{ si } n = l, \\ 0 \text{ sinon.} \end{cases}
\end{aligned}
$$

Par conséquent, nous n'avons pas l'égalité entre les matrices $\psi_k(e_i)\psi_k(e_{i+j})$ et $\psi_k(e_{i+j})\psi_k(e_i)$ étant donné que les éléments se trouvant à la i^{e} ligne et à la $(i+2)^{\text{e}}$ colonne de ces matrices sont respectivement 1 et 0. Les mots w et w' ne sont donc pas équivalents et la règle 1 n'est donc pas satisfaite si j vaut 1.

Théorème 3.2.4. ***Règle 2*** *: Soient $\Sigma = \{e_1, \ldots, e_k\}$ un alphabet ordonné, i un entier tel que $1 \leq i \leq k-1$ et x, y, z, w, w' des mots sur l'alphabet Σ. Si $w = xe_ie_{i+1}ye_{i+1}e_iz$, si $w' = xe_{i+1}e_iye_ie_{i+1}z$, si $|y|_{e_{i-1}} = 0$ lorsque $i \neq 1$ et si $|y|_{e_{i+2}} = 0$ lorsque $i \neq k-1$, alors les mots w et w' sont équivalents.*

Démonstration. Étant donné que la fonction ψ_k est un morphisme (de monoïdes), nous avons ce qui suit :

$$
\begin{aligned}
\psi_k(w) &= \psi_k(x)\psi_k(e_i)\psi_k(e_{i+1})\psi_k(y)\psi_k(e_{i+1})\psi_k(e_i)\psi_k(z), \\
\psi_k(w') &= \psi_k(x)\psi_k(e_{i+1})\psi_k(e_i)\psi_k(y)\psi_k(e_i)\psi_k(e_{i+1})\psi_k(z).
\end{aligned}
$$

Vu la Définition 3.0.1 et vu le Théorème 2.3.6, il nous suffit de vérifier l'égalité entre la matrice $\psi_k(e_i)\psi_k(e_{i+1})\psi_k(y)\psi_k(e_{i+1})\psi_k(e_i)$ et la matrice $\psi_k(e_{i+1})\psi_k(e_i)\psi_k(y)\psi_k(e_i)\psi_k(e_{i+1})$ pour démontrer ce théorème. Cependant, cette égalité de matrices n'est pas évidente[4]. Pour démontrer cette égalité, nous allons tout d'abord utiliser la définition du produit matriciel et obtenir ce qui suit :

$$
\begin{aligned}
&(\psi_k(e_i)\psi_k(e_{i+1})\psi_k(y)\psi_k(e_{i+1})\psi_k(e_i))_{m,n} \\
&\qquad = \sum_{r=1}^{k+1}\sum_{s=1}^{k+1}(\psi_k(e_i)\psi_k(e_{i+1}))_{m,r}(\psi_k(y))_{r,s}(\psi_k(e_{i+1})\psi_k(e_i))_{s,n}
\end{aligned}
$$

4. Voir la Remarque 3.2.3.

Ensuite, vu le Corollaire 2.3.6 et vu la Remarque 3.2.3, nous pouvons en déduire ce qui suit :

$$[\psi_k(e_i)\psi_k(e_{i+1})]_{m,r} = \begin{cases} 1 \text{ si } m = r, \\ 1 \text{ si } m = i \text{ et } r = i+1, \\ 1 \text{ si } m = i \text{ et } r = i+2, \\ 1 \text{ si } m = i+1 \text{ et } r = i+2, \\ 0 \text{ sinon,} \end{cases}$$

$$[\psi_k(y)]_{r,s} = \begin{cases} 1 \text{ si } r = s, \\ |y|_{e_{r,s-1}} \text{ si } s > r, \\ 0 \text{ sinon,} \end{cases}$$

$$[\psi_k(e_{i+1})\psi_k(e_i)]_{s,n} = \begin{cases} 1 \text{ si } s = n, \\ 1 \text{ si } s = i \text{ et } n = i+1, \\ 1 \text{ si } s = i+1 \text{ et } n = i+2, \\ 0 \text{ sinon.} \end{cases}$$

Enfin, vu la définition du produit matriciel et vu les hypothèses de ce théorème [5], nous pouvons facilement vérifier que l'élément $[\psi_k(e_i)\psi_k(e_{i+1})\psi_k(y)\psi_k(e_{i+1})\psi_k(e_i)]_{m,n}$ de la matrice $\psi_k(e_i)\psi_k(e_{i+1})\psi_k(y)\psi_k(e_{i+1})\psi_k(e_i)$ et l'élément $[\psi_k(e_{i+1})\psi_k(e_i)\psi_k(y)\psi_k(e_i)\psi_k(e_{i+1})]_{m,n}$ de la matrice $\psi_k(e_{i+1})\psi_k(e_i)\psi_k(y)\psi_k(e_i)\psi_k(e_{i+1})$ sont égaux et valent :

$$\begin{cases} 1 \text{ si } m = n, \\ 2 + |y|_{e_i} \text{ si } m = i \text{ et } n = i+1, \\ 2 + |y|_{e_{i+1}} \text{ si } m = i+1 \text{ et } n = i+2, \\ 2 + |y|_{e_i} + |y|_{e_{i+1}} + |y|_{e_{i,i+1}} \text{ si } m = i \text{ et } n = i+2, \\ |y|_{e_{m,n-1}} \text{ si } i \neq 1, i+1 < m < n \text{ et } n \neq i+2, \\ |y|_{e_{m,n-1}} \text{ si } i \neq 1 \text{ et } m < n < i-1, \\ |y|_{e_{m,n-1}} \text{ si } i = 1, m > 2 \text{ et } n > 4, \\ |y|_{e_{m,n-1}} \text{ si } i \neq k-1 \text{ et } i+2 < m < n, \\ |y|_{e_{m,n-1}} \text{ si } i \neq k-1, m < n < i+1 \text{ et } m \neq i, \\ |y|_{e_{m,n-1}} \text{ si } i = k-1, m < k-1 \text{ et } n < k, \\ 0 \text{ sinon.} \end{cases}$$

5. Nous utilisons les hypothèses suivantes :

$$|y|_{e_{i-1}} = 0 \text{ si } i \neq 1,$$
$$|y|_{e_{i+2}} = 0 \text{ si } i \neq k-1.$$

Vu ce qui précède, les éléments de la matrice $\psi_k(e_{i+1})\psi_k(e_i)\psi_k(y)\psi_k(e_i)\psi_k(e_{i+1})$ sont identiques aux éléments de la matrice $\psi_k(e_i)\psi_k(e_{i+1})\psi_k(y)\psi_k(e_{i+1})\psi_k(e_i)$. Par conséquent, nous avons l'égalité entre ces matrices ainsi que l'égalité entre les matrices $\psi_k(w)$ et $\psi_k(w')$.

□

Remarque 3.2.5. Contrairement à la règle 1, nous pouvons appliquer la règle 2 lorsque nous travaillons sur un alphabet ordonné constitué de deux lettres. En effet, si nous travaillons sur l'alphabet ordonné $\Sigma = \{e_1, e_2\}$ et si les mots x, y et z appartiennent à Σ^*, les mots $xe_1e_2ye_2e_1z$ et $xe_2e_1ye_1e_2z$ sont équivalents. Cette équivalence de Parikh s'avère être un cas particulier du Théorème 3.2.4 et nous l'obtenons en prenant i qui vaut 1 dans ce théorème.

3.3 Equivalence de Parikh et permutation circulaire

Dans cette section, nous allons établir une autre méthode afin de trouver rapidement des mots équivalents. Plus précisément, nous allons nous intéresser aux articles [12] et [11] dans le but d'établir un lien entre l'équivalence de Parikh [6] et la notion de permutation circulaire, aussi appelée cycle [7].

Théorème 3.3.1. *Soient k un entier tel que $k \geq 2$ et $\Sigma = \{e_1, \ldots, e_k\}$ un alphabet ordonné. Si les mots v_l, $1 \leq l \leq k$, sont égaux aux mots $e_l \cdots e_k e_1 \cdots e_{l-1}$, alors les mots $w_1 = v_1 \cdots v_k$ et $w_2 = v_2 \cdots v_k v_1$ sont équivalents.*

Nous ne sommes pas en mesure de démontrer le théorème précédent sans introduire plusieurs lemmes.

Lemme 3.3.2. *Soient $\Sigma = \{e_1, \ldots, e_k\}$ un alphabet ordonné, i un entier tel que $1 \leq i \leq k$ et j un entier tel que $1 \leq j \leq k^2$. La lettre e_i apparaît à la position j dans le mot w_1 si et seulement si la lettre e_{k+1-i} apparaît à la position $k^2 + 1 - j$ dans le mot [8] w_2.*

Démonstration. Les mots v_n, $1 \leq n \leq k$, sont égaux aux mots $e_n \cdots e_k e_1 \cdots e_{n-1}$ par définition. Ainsi, nous pouvons en déduire que ces mots sont de longueur k, que le nombre d'occurrences d'une lettre e_i appartenant à Σ dans chacun de ces mots vaut 1 et que la position de la lettre e_i dans ces mots est déterminée à l'aide du tableau suivant :

---

6. Voir la Définition 3.0.1.

7. Soient l'ensemble $X = \{1, \ldots, n\}$, $t_1, \ldots, t_p$ des éléments de X avec $p \geq 1$ et $t_{p+1}, \ldots, t_n$ les $n-p$ autres éléments de X. Le *cycle* associé à $t_1, \ldots, t_p$ est la permutation qui remplace $t_1, \ldots, t_{p-1}$ par l'élément suivant, t_p par t_1 et qui laisse inchangés les autres éléments. On note ce cycle $\begin{pmatrix} t_1 & t_2 & \cdots & t_{p-1} & t_p \end{pmatrix}$ où $t_1, \ldots, t_p$ sont appelés les éléments du cycle et où p est la longueur du cycle.

8. Les mots w_1 et w_2 ont été définis dans le Théorème 3.3.1.

	v_1	v_2	$\cdots$	v_i	v_{i+1}	v_{i+2}	$\cdots$	v_{k-1}	v_k
e_i	i	$i-1$	$\cdots$	1	k	$k-1$	$\cdots$	$i+2$	$i+1$

Nous remarquons que, si la lettre e_i ne se trouve pas en première position dans le mot v_n (resp. v_k) où n est tel que $1 \leq n \leq k-1$, cette lettre se déplace d'une unité vers la gauche lorsque nous passons de v_n (resp. v_k) à v_{n+1} (resp. v_1). Par contre, si la lettre e_i se trouve en première position dans le mot v_n (resp. v_k), cette lettre se déplace de la première position à la dernière [9] position lorsque nous passons de v_n (resp. v_k) à v_{n+1} (resp. v_1).

Nous considérons n un entier positif et nous notons $P_k(n)$ le plus petit entier [10] strictement positif congru à n modulo k. Vu le tableau ci-dessus, nous pouvons en déduire que la position de la lettre e_i dans v_n ($1 \leq n \leq k$) vaut $P_k(i-(n-1))$. Ainsi, vu la définition des mots w_1 et w_2, nous pouvons en déduire que la lettre e_i apparaît dans les mots w_1 et w_2 respectivement en les positions $tk + P_k(i-t)$ et $tk + P_k(i-t-1)$, $0 \leq t \leq k-1$. En remplacant i par $k+1-i$, nous pouvons en déduire que la lettre e_{k+1-i} apparaît dans le mot w_2 en les positions $tk+P_k(k-i-t)$, $0 \leq t \leq k-1$. Ainsi, pour conclure la preuve, il nous suffit de démontrer que l'ensemble des valeurs $tk+P_k(k-i-t)$, $0 \leq t \leq k-1$, coïncide [11] avec l'ensemble des valeurs $k^2+1-tk-P_k(i-t)$, $0 \leq t \leq k-1$. Plus précisément, nous allons montrer l'égalité [12] entre $k^2+1-tk-P_k(i-t)$ et $(k-1-t)k+P_k(k-i-(k-1-t))$ pour chaque t fixé ($0 \leq t \leq k-1$). Autrement dit, nous allons montrer l'égalité entre $1+k-P_k(i-t)$ et $P_k(1+t-i)$. Pour prouver cette dernière égalité, nous allons envisager deux cas.

Premièrement, nous allons considérer le cas où $0 \leq t \leq i-1$. Dans ce cas, l'entier $i-t$ est tel que $1 \leq i-t \leq k$ et l'entier $1+t-i$ est tel que $1-k \leq 1+t-i \leq 0$. Ainsi, l'entier $P_k(i-t)$ vaut $i-t$ et l'entier $P_k(1+t-i)$ vaut $1+t-i+k$ d'où l'égalité entre $1+k-P_k(i-t)$ et $P_k(1+t-i)$.

Deuxièmement, nous allons considérer le cas où $i \leq t \leq k-1$. Dans ce cas, l'entier $i-t$ est tel que $1-k \leq i-t \leq 0$ et l'entier $1+t-i$ est tel que $1 \leq 1+t-i \leq k$. Ainsi, l'entier $P_k(i-t)$ vaut $i-t+k$ et l'entier $P_k(1+t-i)$ vaut $1+t-i$ d'où l'égalité entre $1+k-P_k(i-t)$ et $P_k(1+t-i)$.

□

9. La dernière position est la k^{e} position.

10. L'entier $P_k(n)$ est compris entre 1 et k.

11. La lettre e_i apparaît aux positions $tk+P_k(i-t)$, $0 \leq t \leq k-1$, dans le mot w_1 et l'entier $tk+P_k(i-t)$ est compris entre 1 et k^2. Ainsi, si nous montrons que la lettre e_{k+1-i} apparaît aux positions $k^2+1-tk-P_k(i-t)$ dans le mot w_2 et si nous remplaçons $tk+P_k(i-t)$ par j, il est évident que la lettre e_i apparaît à la position j dans le mot w_1 si et seulement si la lettre e_{k+1-i} apparaît à la position k^2+1-j dans le mot w_2.

12. Si nous remplaçons t par $k-1-t$ dans $tk+P_k(k-i-t)$, nous obtenons $(k-1-t)k+P_k(k-i-(k-1-t))$. De plus, comme nous prenons l'entier t compris entre 0 et $k-1$, l'entier $k-1-t$ est également compris entre 0 et $k-1$.

Définition 3.3.3. Soit $\Sigma = \{e_1, \ldots e_k\}$ un alphabet ordonné. La lettre e_{k+1-i} appartenant à Σ est la *lettre complémentaire* de la lettre e_i appartenant à Σ.

Remarque 3.3.4. Vu le Lemme 3.3.2, il existe une bijection entre l'ensemble des occurrences d'une lettre dans le mot w_1 et l'ensemble des occurrences de sa lettre complémentaire dans le mot w_2. Ainsi, le nombre d'occurrences d'une lettre dans le mot w_1 est égal au nombre d'occurrences de sa lettre complémentaire dans le mot w_2.

Vu la Remarque 3.3.4, nous pouvons étendre le Lemme 3.3.2 et une généralisation de ce lemme s'avère être le lemme suivant.

Lemme 3.3.5. *Soit $\Sigma = \{e_1, \ldots, e_k\}$ un alphabet ordonné. Pour tous les indices i et j tels que $1 \leq i \leq j \leq k$, le nombre d'occurrences du sous-mot $e_i \cdots e_j$ dans le mot w_1 est égal au nombre d'occurrences du sous-mot $e_{k+1-j} \cdots e_{k+1-i}$ dans le mot[13] w_2.*

Démonstration. Vu le Lemme 3.3.2, le mot $e_i \cdots e_j$ s'avère être un sous-mot de w_1 composé des lettres e_l, $i \leq l \leq j$, apparaissant respectivement aux positions p_l dans le mot w_1 si et seulement si le mot $e_{k+1-j} \cdots e_{k+1-i}$ s'avère être un sous-mot de w_2 composé des lettres e_{k+1-l}, $i \leq l \leq j$, apparaissant respectivement aux positions $k^2 + 1 - p_l$ dans le mot w_2. Ainsi, il existe une bijection entre l'ensemble des occurrences du sous-mot $e_i \cdots e_j$ de w_1 et l'ensemble des occurrences du sous-mot $e_{k+1-j} \cdots e_{k+1-i}$ de w_2 ce qui nous permet de conclure[14].

□

Remarque 3.3.6. En remplacant l'indice i par 1 et l'indice j par k dans le Lemme 3.3.5, nous pouvons en déduire que le nombre d'occurrences du mot $e_1 \cdots e_k$ dans le mot w_1 est égal au nombre d'occurrences du mot $e_1 \cdots e_k$ dans le mot w_2. Par conséquent, nous avons l'égalité entre $[\psi_k(w_1)]_{1,k+1}$ et $[\psi_k(w_2)]_{1,k+1}$.

À ce stade, nous n'avons toujours pas montré l'égalité entre la matrice de Parikh du mot w_1 et la matrice de Parikh du mot w_2. En effet, nous avons uniquement montré l'égalité[15] entre les éléments $[\psi_k(w_1)]_{1,k+1}$ et $[\psi_k(w_2)]_{1,k+1}$. Cependant, le lemme suivant va nous permettre d'égaler ces matrices.

Lemme 3.3.7. *Soit $\Sigma = \{e_1, \ldots, e_k\}$ un alphabet ordonné. Pour chaque entier i compris entre 1 et $k - 1$, les facteurs de longueur i du mot $e_1 \cdots e_k$ s'avèrent être des sous-mots de w_1 apparaissant tous le même nombre de fois dans[16] w_1.*

Démonstration. Pour démontrer ce lemme, nous allons procéder par récurrence sur l'indice k. Si $k \leq 2$, le lemme est évident[17].

13. Les mots w_1 et w_2 ont été définis dans le Théorème 3.3.1.

14. Vu que nous avons une bijection entre ces deux ensembles, le nombre d'occurrences du mot $e_i \cdots e_j$ dans le mot w_1 est égal au nombre d'occurrences du mot $e_{k+1-j} \cdots e_{k+1-i}$ dans le mot w_2.

15. Cette égalité à été montrée dans la remarque précédente

16. Le mot w_1 a été défini dans le Théorème 3.3.1.

17. Nous passons en revue toutes les possibilités.

Nous supposons que le lemme est vrai si nous travaillons sur un alphabet ordonné constitué de k ($k \in \mathbb{Z}_0$) lettres et nous allons vérifier qu'il reste vrai si nous travaillons sur un alphabet ordonné constitué de $k+1$ lettres.

Dans la suite de cette démonstration, nous prenons p et q des entiers tels que $1 \leq p \leq q \leq k+1$ et nous notons le mot w_1 (resp. w_2) construit sur l'alphabet ordonné $\{e_p, \ldots, e_q\}$ par $w_1(p,q)$ (resp. $w_2(p,q)$).

Tout d'abord, nous allons considérer le cas où l'entier i est compris entre 1 et $k-1$ et où les facteurs de longueur i du mot $e_1 \cdots e_{k+1}$ ne comportent pas la lettre e_{k+1}. Si les facteurs du mot $e_1 \cdots e_{k+1}$ ne comportent pas la lettre e_{k+1}, ces facteurs sont des facteurs du mot $e_1 \cdots e_k$. Ainsi, nous pouvons conclure grâce aux deux affirmations suivantes :

(1) Par hypothèse de récurrence, les facteurs de longueur i du mot $e_1 \cdots e_k$ sont des sous-mots du mot $w_1(1,k)$. Étant donné que le nombre d'occurrences des facteurs de longueur i du mot $e_1 \cdots e_k$ dans le mot $w_1(1,k+1)$ reste inchangé si nous remplaçons le mot $w_1(1,k+1)$ par le mot $w_1(1,k+1)$ auquel nous avons supprimé les lettres e_{k+1}, nous supprimons la lettre e_{k+1} du mot $w_1(1,k+1)$ afin d'obtenir le mot $w_1(1,k)e_1 \cdots e_k$. Les facteurs de longueur i du mot $e_1 \cdots e_{k+1}$ sont donc des sous-mots de $w_1(1,k+1)$.

(2) Si m est un entier compris entre 1 et $k-i+1$, le mot $e_m \cdots e_{m+i-1}$ est un facteur de longueur i du mot $e_1 \cdots e_k$. Étant donné que le nombre d'occurrences des facteurs de longueur i du mot $e_1 \cdots e_k$ dans le mot $w_1(1,k+1)$ reste inchangé si nous remplaçons le mot $w_1(1,k+1)$ par le mot $w_1(1,k+1)$ auquel nous avons supprimé les lettres e_{k+1}, nous supprimons la lettre e_{k+1} du mot $w_1(1,k+1)$ afin d'obtenir le mot $w_1(1,k)e_1 \cdots e_k$ et nous en déduisons ce qui suit :

$$|w_1(1,k+1)|_{e_m \cdots e_{m+i-1}} = 1 + |w_1(1,k)|_{e_m \cdots e_{m+i-1}} + \sum_{l=m}^{m+i-1} |w_1(1,k)|_{e_m \cdots e_l}.$$

Par hypothèse de récurrence, pour chaque indice j tel que $1 \leq j \leq k-1$, les facteurs de longueur j du mot $e_1 \cdots e_k$ apparaissent tous le même nombre de fois dans $w_1(1,k)$. Ainsi, vu la relation ci-dessus, nous pouvons en déduire que les facteurs de longueur i du mot $e_1 \cdots e_k$ apparaissent tous le même nombre de fois dans $w_1(1,k+1)$.

Ensuite, nous allons considérer le cas où l'entier i est compris entre 1 et $k-1$ et où les facteurs de longueur i du mot $e_1 \cdots e_{k+1}$ comportent la lettre e_{k+1}. Comme les facteurs de $e_1 \cdots e_{k+1}$ sont de longueur au plus $k-1$ et contiennent la lettre e_{k+1}, ces facteurs ne peuvent pas contenir la lettre e_1. Ainsi, nous procédons de manière analogue au cas précédent [18].

18. Etant donné que les facteurs de longueur i du mot $e_1 \cdots e_{k+1}$ ne contiennent pas la lettre e_1, ces facteurs sont des facteurs du mot $e_2 \cdots e_{k+1}$. Ainsi, nous pouvons supprimer la lettre e_1 du mot $w_1(1,k+1)$

Enfin, nous allons considérer le cas où l'entier i vaut k et nous allons montrer que les facteurs de longueur k du mot $e_1 \cdots e_{k+1}$ possèdent tous le même nombre d'occurrences dans $w_1(1, k+1)$. En d'autres termes, nous allons montrer ce qui suit :

$$|w_1(1, k+1)|_{e_1 \cdots e_k} = |w_1(1, k+1)|_{e_2 \cdots e_{k+1}}.$$

Vu le Lemme 3.3.5, nous avons l'égalité entre $|w_1(1, k+1)|_{e_2 \cdots e_{k+1}}$ et $|w_2(1, k+1)|_{e_1 \cdots e_k}$. De plus, le nombre d'occurrences du facteur $e_1 \cdots e_k$ dans le mot $w_1(1, k+1)$ reste inchangé si nous remplaçons le mot $w_1(1, k+1)$ par le mot $w_1(1, k+1)$ auquel nous avons supprimé les lettres e_{k+1}. Ainsi, il nous suffit de montrer ce qui suit :

$$|w_1(1, k)e_1 \cdots e_k|_{e_1 \cdots e_k} = |w_2(1, k)e_1 \cdots e_k|_{e_1 \cdots e_k}.$$

Vu le Théorème 2.3.6, nous pouvons en déduire l'égalité entre l'entier $|w_1(1, k)e_1 \cdots e_k|_{e_1 \cdots e_k}$ et l'entier $[\psi_k(w_1(1, k)e_1 \cdots e_k)]_{1,k+1}$ ainsi que l'égalité entre l'entier $|w_2(1, k)e_1 \cdots e_k|_{e_1 \cdots e_k}$ et l'entier $[\psi_k(w_2(1, k)e_1 \cdots e_k]_{1,k+1}$. De plus, vu le Théorème 2.3.6 et définition du produit matriciel, nous pouvons en déduire que les éléments de la matrice $\psi_k(w_1(1, k)e_1 \cdots e_k)$ et que les éléments de la matrice $\psi_k(w_2(1, k)e_1 \cdots e_k)$ sont donnés respectivement par [19] :

$$\begin{aligned}
[\psi_k(w_1(1, k)e_1 \cdots e_k)]_{p,q} &= \sum_{n=1}^{k+1} [\psi_k(w_1(1, k))]_{p,n} [\psi_k(e_1 \cdots e_k)]_{n,q}, \\
&= \begin{cases} 1 \text{ si } 1 \leq p = q \leq k+1, \\ \sum_{n=p+1}^{q} |w_1(1, k)|_{e_{p,n-1}} + 1 \text{ si } 1 \leq p \leq q-1 \leq k+1, \\ 0 \text{ sinon,} \end{cases}
\end{aligned}$$

$$\begin{aligned}
[\psi_k(w_2(1, k)e_1 \cdots e_k)]_{p,q} &= \sum_{n=1}^{k+1} [\psi_k(w_2(1, k)]_{p,n} [\psi_k(e_1 \cdots e_k]_{n,q}, \\
&= \begin{cases} 1 \text{ si } 1 \leq p = q \leq k+1, \\ \sum_{n=p+1}^{q} |w_2(1, k)|_{e_{p,n-1}} + 1 \text{ si } 1 \leq p \leq q-1 \leq k+1, \\ 0 \text{ sinon.} \end{cases}
\end{aligned}$$

Comme l'entier k est supérieur à 1, les entiers 1 et $k+1$ ne sauraient pas être identiques. Ainsi, vu la forme des éléments de la matrice $\psi_k(w_1(1, k)e_1 \cdots e_k)$ et vu la forme des éléments de la matrice $\psi_k(w_2(1, k)e_1 \cdots e_k)$, nous obtenons les égalités suivantes [20] :

$$\begin{aligned}
|w_1(1, k)e_1 \cdots e_k|_{e_1 \cdots e_k} &= 1 + |w_1(1, k)|_{e_{1,1}} + |w_1(1, k)|_{e_{1,2}} + \cdots |w_1(1, k)|_{e_{1,k}}, \\
|w_2(1, k)e_1 \cdots e_k|_{e_1 \cdots e_k} &= 1 + |w_2(1, k)|_{e_{1,1}} + |w_2(1, k)|_{e_{1,2}} + \cdots |w_2(1, k)|_{e_{1,k}}.
\end{aligned}$$

---

afin d'obtenir le mot $(e_2 \cdots e_{k+1})w_1(2, k+1)$. Nous pouvons donc conclure en appliquant l'hypothèse de récurrence sur k.

19. Pour tous les entiers i et j tels que $1 \leq i \leq j \leq k+1$, nous savons que le nombre d'occurrences du mot $e_{i,j}$ dans le mot $e_1 \cdots e_k$ vaut 1.

20. Pour rappel, les mots $e_{i,j}$ sont égaux aux mots $e_i \cdots e_j$.

Vu le Lemme 3.3.5, nous pouvons réécrire l'entier $|w_2(1,k)e_1 \cdots e_k|_{e_1 \cdots e_k}$ comme étant égal à la somme suivante :

$$|w_2(1,k)e_1 \cdots e_k|_{e_1 \cdots e_k} = 1 + |w_1(1,k)|_{e_{k,k}} + |w_1(1,k)|_{e_{k-1,k}} + \cdots + |w_1(1,k)|_{e_{1,k}}.$$

Par hypothèse de récurrence, les facteurs de longueur i du mot $e_1 \cdots e_k$ s'avèrent être des sous-mots de $w_1(1,k)$ apparaissant tous le même nombre de fois dans $w_1(1,k)$. Ainsi, nous pouvons conclure la preuve.

□

Grâce aux lemmes précédents, nous pouvons démontrer le Théorème 3.3.1 énoncé au début de cette sous-section. Pour rappel, ce théorème est le suivant.

Théorème 3.3.8. *Soient k un entier tel que $k \geq 2$ et $\Sigma = \{e_1, \ldots, e_k\}$ un alphabet ordonné. Si les mots v_l, $1 \leq l \leq k$, sont égaux aux mots $e_l \cdots e_k e_1 \cdots e_{l-1}$, alors les mots $w_1 = v_1 \cdots v_k$ et $w_2 = v_2 \cdots v_k v_1$ sont équivalents.*

Démonstration. Vu le Lemme 3.3.5 et le Lemme 3.3.7, nous pouvons facilement en déduire ce théorème.

□

Le Théorème 3.3.1 nous a permis de déterminer deux mots équivalents. Cependant, nous pouvons identifier d'autres mots équivalents à l'aide de ce théorème. Pour ce faire, il est nécessaire de définir la notion de projection.

Définition 3.3.9. Soient $\Sigma = \{e_1, \ldots, e_k\}$ un alphabet ordonné et w un mot sur l'alphabet Σ. La *projection* du mot w sur un alphabet ordonné $\Sigma_{<k}$ obtenu en enlevant certaines lettres de Σ est le mot u obtenu en supprimant les lettres n'appartenant pas à $\Sigma_{<k}$ du mot w.

Théorème 3.3.10. *Soit $\Sigma = \{e_1, \ldots, e_k\}$ un alphabet ordonné. Si les mots v_i, $1 \leq i \leq k$, sont égaux aux mots $e_i \cdots e_k e_1 \cdots e_{i-1}$ et si les mots x_j, $0 \leq j \leq k$, sont uniquement constitués des lettres e_j et e_{j+1} lorsque j est compris entre 1 et $k-1$, les mots $X_1 = x_0 v_1 x_1 v_2 x_2 \cdots x_{k-1} v_k x_k$ et $X_2 = x_0 v_2 x_1 v_3 x_2 \cdots x_{k-2} v_k x_{k-1} v_1 x_k$ sont équivalents.*

Démonstration. Pour démontrer ce théorème, nous allons considérer les suites de mots suivantes :

$$X_1^{(j)} = x_0 v_1 \cdots x_j v_{j+1} v_{j+2} \cdots v_k x_k, \quad 0 \leq j \leq k-1 \tag{3.1}$$

$$X_2^{(j)} = x_0 v_2 \cdots x_j v_{j+2} \cdots v_k v_1 x_k, \quad 0 \leq j \leq k-1 \tag{3.2}$$

et nous allons montrer par récurrence sur j que les mots $X_1^{(j)}$ et $X_2^{(j)}$ sont équivalents pour chaque indice j tel que $0 \leq j \leq k-1$.

Vu la forme de ces suites, nous obtenons l'égalité entre $X_1^{(0)}$ et $x_0 v_1 v_2 \cdots v_k x_k$ ainsi que l'égalité entre $X_2^{(0)}$ et $x_0 v_2 \cdots v_k v_1 x_k$. Par conséquent, vu le Théorème 3.3.1, les mots $X_1^{(0)}$ et $X_2^{(0)}$ sont équivalents, ce qui vérifie le cas de base.

Nous supposons que les mots $X_1^{(j-1)}$ et $X_2^{(j-1)}$ sont équivalents ($j \geq 1$) et nous allons montrer que les mots $X_1^{(j)}$ et $X_2^{(j)}$ sont équivalents. Afin de montrer que les mots $X_1^{(j)}$ et $X_2^{(j)}$ sont équivalents, nous notons ce qui suit :

$$X_1^{(j)} = L_1 x_j R_1, \tag{3.3}$$

$$X_2^{(j)} = L_2 x_j R_2 \tag{3.4}$$

où L_1 (resp. L_2) est la partie se trouvant à gauche de x_j dans $X_1^{(j)}$ (resp. $X_2^{(j)}$) et où R_1 (resp. R_2) est la partie se trouvant à droite de x_j dans $X_1^{(j)}$ (resp. $X_2^{(j)}$). Vu la forme du mot X_1^{j-1} donnée en (3.1), vu la forme du mot X_1^j donnée en (3.1) et (3.3), vu la forme du mot X_2^{j-1} donnée en (3.2) et vu la forme du mot X_2^j donnée en (3.2) et (3.4), nous obtenons les égalités suivantes :

$$X_1^{(j-1)} = L_1 R_1, \tag{3.5}$$

$$X_2^{(j-1)} = L_2 R_2. \tag{3.6}$$

Vu les égalités (3.3) à (3.6), si nous considérons deux entiers l et m satisfaisant les inégalités suivantes :

$$1 \leq l \leq k,\ 1 \leq m \leq k,\ l \leq m,$$

le nombre d'occurrences du mot [21] $e_{l,m}$ dans le mot $X_1^{(j)}$ (resp. $X_2^{(j)}$) est strictement supérieur au nombre d'occurrences du mot $e_{l,m}$ dans le mot $X_1^{(j-1)}$ (resp. $X_2^{(j-1)}$) lorsque :

- le mot L_1 (resp. L_2) contient le sous-mot $e_l \cdots e_{j-1}$ avec $l \leq j$, le mot x_j contient la lettre e_j et le mot R_1 (resp. R_2) contient le sous-mot $e_{j+1} \cdots e_m$ avec [22] $j \leq m$.
- le mot L_1 (resp. L_2) contient le sous-mot $e_l \cdots e_j$ avec $l \leq j+1$, le mot x_j contient la lettre e_{j+1} et le mot R_1 (resp. R_2) contient le sous-mot $e_{j+2} \cdots e_m$ avec $j+1 \leq m$.
- le mot L_1 (resp. L_2) contient le sous-mot $e_l \cdots e_{j-1}$ avec $l \leq j$, le mot x_j contient le sous-mot $e_j e_{j+1}$ et le mot R_1 (resp. R_2) contient le sous-mot $e_{j+2} \cdots e_m$ avec $j+1 \leq m$.

Par hypothèse de récurrence, les mots $X_1^{(j-1)}$ et $X_2^{(j-1)}$ sont équivalents. Ainsi, pour montrer que les mots $X_1^{(j)}$ et $X_2^{(j)}$ sont équivalents, il nous suffit de vérifier que le nombre supplémentaire d'occurrences de chacun des mots $e_{l,m}$, $1 \leq l \leq m \leq k$, obtenu en passant du mot $X_1^{(j-1)}$ au mot $X_1^{(j)}$ est identique au nombre d'occurrences de chacun des mots $e_{l,m}$ obtenu en passant du mot $X_2^{(j-1)}$ au mot $X_2^{(j)}$.

Premièrement, nous allons regarder le nombre d'occurrences des mots $e_{j+1} \cdots e_m$ et $e_{j+2} \cdots e_m$ dans R_1 et R_2. Vu la Définition 3.3.9 et vu les égalités (3.1) à (3.4), les projections de R_1 et de R_2 sur l'alphabet $\{e_{j+1}, \ldots, e_m\}$ sont respectivement les mots suivants :

21. Le mot $e_{l,m}$ est égal au mot $e_l \cdots e_m$.
22. Si l'indice l (resp. m) vaut j, le mot $e_l \cdots e_{j-1}$ (resp. $e_{j+1} \cdots e_m$) est égal au mot vide.

$$R'_1 = (e_{j+1}\cdots e_m)(e_{j+2}\cdots e_m e_{j+1})\cdots(e_m e_{j+1}\cdots e_{m-1})x'_k,$$
$$R'_2 = (e_{j+2}\cdots e_m e_{j+1})(e_{j+3}\cdots e_m e_{j+1}e_{j+2})\cdots(e_{j+1}\cdots e_m)x'_k,$$

où x'_k est la projection du mot x_k sur l'alphabet $\{e_{j+1},\dots,e_m\}$. Ainsi, en appliquant le Théorème 3.3.1 à l'alphabet $\{e_{j+1},\dots,e_m\}$, nous en déduisons que les mots R'_1 et R'_2 sont équivalents. Comme les mots $e_{j+1}\cdots e_m$ et $e_{j+2}\cdots e_m$ ne contiennent pas les lettres $e_1,\dots,e_j$, $e_{m+1},\dots,e_{k-1}$ et e_k, nous pouvons en déduire que le nombre d'occurrences de ces mots dans R_1 (resp. R_2) est égal au nombre d'occurrences de ces mots dans la projection de R_1 (resp. R_2) sur l'alphabet $\{e_{j+1},\dots,e_m\}$. Par conséquent, les mots $e_{j+1}\cdots e_m$ et $e_{j+2}\cdots e_m$ possèdent le même nombre d'occurrences dans R_1 et R_2.

Deuxièmement, nous allons examiner le nombre d'occurrences des mots $e_l\cdots e_j$ et $e_l\cdots e_{j-1}$ dans L_1 et L_2. Par hypothèse de récurrence, les mots $X_1^{(j-1)}$ et $X_2^{(j-1)}$ sont équivalents, c'est-à-dire que les mots L_1R_1 et L_2R_2 sont équivalents vu les égalités (3.5) et (3.6). Comme les mots $e_l\cdots e_j$ et $e_l\cdots e_{j-1}$ ne contiennent pas les lettres e_1, …, e_{l-1}, e_{j+1}, …, e_{k-1} et e_k, nous pouvons en déduire que le nombre d'occurrences de ces mots dans L_1R_1 (resp. L_2R_2) est égal au nombre d'occurrences de ces mots dans la projection de L_1R1 (resp. L_2R_2) sur l'alphabet $\{e_l,\dots,e_j\}$. De plus, vu les égalités (3.1) à (3.4), les projections de R_1 et R_2 sur l'alphabet $\{e_l,\dots,e_j\}$ sont égales et valent ce qui suit :

$$(e_l\cdots e_j)^{k-j}x''_k$$

où x''_k est la projection de x_k sur l'alphabet $\{e_l,\dots,e_j\}$. Ainsi, les mots $e_l\cdots e_j$ et $e_l\cdots e_{j-1}$ possèdent le même nombre d'occurrences dans la projection du mot L_1 sur l'alphabet $\{e_l,\dots,e_j\}$ que dans la projection du mot L_2 sur l'alphabet $\{e_l,\dots,e_j\}$. Vu que les mots $e_l\cdots e_j$ et $e_l\cdots e_{j-1}$ ne possèdent pas les lettres $e_1,\dots,e_{l-1},e_{j+1}\dots,e_{k-1}$ et e_k, ces mots possèdent le même nombre d'occurrences dans L_1 que dans L_2.

Comme les mots $e_{j+1}\cdots e_m$ et $e_{j+2}\cdots e_m$ possèdent le même nombre d'occurrences dans R_1 et R_2 et comme les mots $e_l\cdots e_j$ et $e_l\cdots e_{j-1}$ possèdent le même nombre d'occurrences dans les mots L_1 et L_2, le nombre supplémentaire d'occurrences de chaque mot $e_{l,m}$ $(1\le l\le m\le k)$ obtenu en passant du mot $X_1^{(j-1)}$ au mot $X_1^{(j)}$ est identique au nombre supplémentaire d'occurrences de chaque mot $e_{l,m}$ obtenu en passant du mot $X_2^{(j-1)}$ au mot $X_2^{(j)}$.

□

Une généralisation de ce théorème s'avère être le Théorème 3.3.11.

Théorème 3.3.11. *Soient $\Sigma=\{e_1,\dots,e_k\}$ un alphabet ordonné, l un entier tel que $0\le l\le k-2$, m un entier tel que $2\le m\le k-l$ et $\Sigma_{<k}$ un alphabet ordonné obtenu en enlevant certaines lettres de Σ notamment les lettres e_i, $l\le i\le l+m+1$. Si les mots x_0, x_m sont des mots arbitraires sur l'alphabet Σ et si les mots x_i, $1\le i\le m-1$, sont des*

mots uniquement constitués des lettres de l'alphabet $\Sigma_{<k}$ ainsi que des lettres e_{l+i} et e_{l+i+1}, les mots suivants sont équivalents :

$$W_1 = x_0 e_{l+1} e_{l+2} \cdots e_{l+m} x_1 e_{l+2} \cdots e_{l+m} e_{l+1} x_2 \cdots x_{m-1} e_{l+m} e_{l+1} \cdots e_{l+m-1} x_m,$$
$$W_2 = x_0 e_{l+2} \cdots e_{l+m} e_{l+1} x_1 e_{l+3} \cdots e_{l+1} e_{l+2} x_2 \cdots x_{m-1} e_{l+1} e_{l+2} \cdots e_{l+m} x_m.$$

Démonstration. En appliquant le Théorème 3.3.10 à l'alphabet ordonné $\{e_{l+1}, \ldots, e_{l+m}\}$, nous en déduisons que les mots [23] X_1 et X_2 sont équivalents si les mots x_j, $1 \leq j \leq m-1$, sont uniquement constitués des lettres e_{l+j} et e_{l+j+1}. De plus, les mots X_1 et X_2 restent équivalents si les mots x_j, $1 \leq j \leq m-1$, sont également constitués des lettres de l'alphabet $\Sigma_{<k}$. En effet, l'alphabet $\Sigma_{<k}$ ne comporte pas les lettres $e_l, e_{l+1}, \ldots, e_{l+m}$ et e_{l+m+1} ce qui signifie que les matrices de Parikh des mots X_1 et X_2 restent inchangées si nous remplaçons les mots x_j, $1 \leq j \leq m-1$, constitués uniquement des lettres e_{l+j} et e_{l+j+1} par des mots constitués uniquement des lettres de $\Sigma_{<k}$ ainsi que des lettres e_{l+j} et e_{l+j+1}.

$\square$

Remarque 3.3.12. Si nous prenons $l = 0$ et $m = k$ dans le théorème précédent, nous obtenons le Théorème 3.3.1.

3.4 Miroir et équivalence de Parikh

Cette section nous fournit quelques méthodes afin de trouver rapidement des mots équivalents et est basée sur l'article [2] ainsi que sur l'article [9]. Ces méthodes sont basées sur une nouvelle notion appelée image miroir. Ainsi, avant d'énoncer ces méthodes, nous allons définir l'image miroir et donner trois propriétés liant l'image miroir avec l'équivalence de Parikh [24]

Définition 3.4.1. Soient $\Sigma = \{e_1, \ldots, e_k\}$ un alphabet ordonné, w un mot sur l'alphabet Σ de longueur finie et n la longueur de ce mot. L'*image miroir* de w, aussi appelé le *retourné* de w ou le *miroir* de w, est le mot obtenu en remplacant la $(n-i+1)^{\text{e}}$ lettre de w par la i^{e} lettre de w. De plus, l'*image miroir* du mot w est notée w^R.

Exemple 3.4.2. Soit l'alphabet ordonné $\Sigma = \{a, b, c\}$. L'image miroir du mot $w = aababc$ s'avère être le mot $w' = cbabaa$.

Après avoir défini l'image miroir, nous allons énoncer deux propositions qui ont été démontrées lors des travaux pratiques du cours intitulé "Théorie des automates et langages formels" (cfr. article [6]).

23. Les mots X_1 et X_2 sont définis dans le Théorème 3.3.10.
24. Voir la Définition 3.0.1.

Proposition 3.4.3. *Soit $\Sigma = \{e_1, \ldots, e_k\}$ un alphabet ordonné. Pour tous mots u et v sur l'alphabet Σ, nous avons l'égalité suivante :*

$$(uv)^R = v^R u^R.$$

Proposition 3.4.4. *Soit $\Sigma = \{e_1, \ldots, e_k\}$ un alphabet ordonné. Pour tout mot w sur l'alphabet Σ, nous avons l'égalité suivante :*

$$(w^R)^R = w.$$

Une fois la Proposition 3.4.3 et la Proposition 3.4.4 vues, nous pouvons énoncer et démontrer plusieurs propriétés liant l'image miroir et l'équivalence de Parikh.

Lemme 3.4.5. *Soient $\Sigma = \{e_1 \cdots e_k\}$ un alphabet ordonné et u, v des mots sur l'alphabet Σ. Le mot v est obtenu en appliquant la règle 1 (resp. 2) au mot u si et seulement si le mot v^R est obtenu en appliquant la règle 1 (resp. 2) au mot[25] u^R.*

Démonstration. $\Rightarrow$ Premièrement, nous allons envisager le cas où le mot v est obtenu en appliquant la règle 1 au mot u. Vu le Théorème 3.2.1, nous pouvons écrire les mots u et v comme suit :

$$u = u_1 e_i e_{i+j} u_2,$$
$$v = u_1 e_{i+j} e_i u_2.$$

avec i un entier compris entre 1 et $k-2$, j un entier compris entre 2 et $k-i$ ainsi que u_1 et u_2 des mots sur l'alphabet Σ. Une fois la forme des mots u et v décrite, nous prenons l'image miroir de ces mots et nous en déduisons vu la Proposition 3.4.3 que le mot v^R s'obtient en appliquant la règle 1 au mot u^R.

Deuxièmement, nous allons envisager le cas où le mot v est obtenu en appliquant la règle 2 au mot u. Vu le Théorème 3.2.4, nous pouvons écrire les mots u et v comme suit :

$$u = x e_i e_{i+1} y e_{i+1} e_i z,$$
$$v = x e_{i+1} e_i y e_i e_{i+1} z.$$

avec i un entier compris entre 1 et $k-1$, x et z des mots quelconques sur l'alphabet Σ ainsi que y un mot sur l'alphabet Σ dans lequel le nombre d'occurrence de la lettre e_{i-1} est nul si i diffère de 1 et le nombre d'occurrence de la lettre e_{i+2} est nul si i diffère de $k-1$. Une fois la forme des mots u et v décrite, nous prenons l'image miroir de ces mots et nous en déduisons vu la Proposition 3.4.3 que nous devons appliquer la règle 2 au mot u^R pour obtenir le mot v^R.

$\Leftarrow$ Vu la Proposition 3.4.4, les mots u (resp. v) et $(u^R)^R$ (resp. $(v^R)^R$) sont égaux. Ainsi, nous pouvons procéder de manière analogue à ce qui précède et conclure la preuve.

□

25. Les règles 1 et 2 ont été respectivement définies dans le Théorème 3.2.1 et le Théorème 3.2.4.

Avant d'énoncer et de démontrer le Théorème 3.4.10 permettant de trouver des mots équivalents à l'aide d'autres mots équivalents, il est nécessaire de définir la notion d'équivalence élémentaire de Parikh ainsi que d'énoncer le Théorème 3.4.9 qui sera démontré dans le chapitre 6.

Définition 3.4.6. Soient $\Sigma = \{e_1, \ldots, e_k\}$ un alphabet ordonné et w, w' des mots sur l'alphabet Σ. Les mots w et w' sont *élémentairement équivalents*, ce que l'on note $w \equiv_k^E w'$, si le mot w' est obtenu en appliquant la règle 1 et/ou la règle 2 un nombre fini de fois au mot w.

Exemple 3.4.7. Si nous travaillons sur l'alphabet ordonné $\Sigma = \{a, b, c, d\}$, les mots $w = dabdba$ et $w' = bdadab$ sont élémentairement équivalents. En effet, en appliquant [26] une fois le Théorème 3.2.4 à w, nous obtenons le mot $v = dbadab$ et, en appliquant [27] une fois le Théorème 3.2.1 à v, nous obtenons w'.

Remarque 3.4.8. Si les mots w et w' sont élémentairement équivalents, leurs matrices de Parikh sont qualifiées de matrices élémentairement équivalentes.

Théorème 3.4.9. *Soient $\Sigma = \{e_1 \cdots e_k\}$ un alphabet ordonné et u, v des mots sur l'alphabet Σ. Les mots u et v sont équivalents si et seulement s'ils sont élémentairement équivalents.*

À l'aide du Théorème 3.4.9, nous pouvons démontrer le théorème suivant.

Théorème 3.4.10. *Soient $\Sigma = \{e_1 \cdots e_k\}$ un alphabet ordonné et u, v des mots sur l'alphabet Σ. Les mots u et v sont équivalents si et seulement si les mots u^R et v^R sont équivalents.*

Démonstration. $\Rightarrow$ Vu le Théorème 3.4.9, les mots u et v sont élémentairement équivalents. Par conséquent, vu la Définition 3.4.6, il existe un entier k positif et des mots $u_0, u_1, \cdots, u_k$ satisfaisant les conditions suivantes :

- le mot u_0 est égal au mot u,
- le mot u_k est égal au mot v,
- les mots u_{j+1}, $0 \leq j \leq k-1$, sont obtenus en appliquant soit la règle 1 soit la règle 2 aux mots u_j.

De plus, vu le Lemme 3.4.5, nous pouvons en déduire que les mots u_{j+1}^R, $0 \leq j \leq k-1$, sont obtenus en appliquant soit la règle 1 soit la règle 2 aux mots u_j^R. Ainsi, les mots u^R et v^R sont élémentairment équivalents et même équivalents vu le Théorème 3.4.9.

26. Nous prenons $i = 1$, $x = d$, $y = d$ et $z = \varepsilon$ dans l'énoncé de ce théorème.
27. Nous prenons $i = 2$ et $j = 2$ dans l'énoncé de ce théorème.

$\Leftarrow$ Vu la propostion 3.4.4, nous avons l'égalité entre les mots u et $(u^R)^R$ ainsi que l'égalité entre les mots v et $(v^R)^R$. Par conséquent, nous pouvons procéder de manière analogue à ce qui précède et conclure la preuve.

$\square$

Une dernière méthode afin de trouver rapidement des mots équivalents s'avère être le lemme suivant.

Lemme 3.4.11. *Soient $\Sigma = \{e_1 \cdots e_k\}$ un alphabet ordonné et i_j, $1 \leq j \leq k$, des entiers positifs. Si nous notons $w = e_1^{i_1} e_2^{i_2} \cdots e_k^{i_k}$, les mots ww^R et $w^R w$ sont équivalents.*

Démonstration. Pour démontrer ce lemme, nous allons procéder par récurrence sur k. Si k vaut 0 ou si k vaut 1, le lemme est satisfait c'est-à-dire que le cas de base est vérifié[28].

Nous supposons à présent que le lemme est satisfait pour k $(k \geq 1)$ et nous allons vérifier qu'il le reste pour $k+1$. Comme nous connaissons la forme du mot w, nous pouvons déterminer la forme du mot ww^R :

$$\underbrace{e_1^{i_1} \cdots e_{k+1}^{i_{k+1}}}_{w} \underbrace{e_{k+1}^{i_{k+1}} \cdots e_1^{i_1}}_{w^R}.$$

Une fois la forme du mot ww^R déterminée, nous appliquons $(i_k \times i_{k+1})$ fois le Théorème 3.2.4 à ce mot[29]. Par conséquent, nous obtenons le mot $w_1 w_1^R$ qui est équivalent au mot ww^R :

$$\underbrace{e_1^{i_1} \cdots e_{k-1}^{i_{k-1}} e_{k+1}^{i_{k+1}} e_k^{i_k}}_{w_1} \underbrace{e_k^{i_k} e_{k+1}^{i_{k+1}} e_{k-1}^{i_{k-1}} \cdots e_1^{i_1}}_{w_1^R}.$$

Après avoir appliqué un nombre fini de fois la règle 2 au mot ww^R, nous appliquons un nombre fini de fois la règle 1 au mot $w_1 w_1^R$. Plus précisément, nous procédons en $(k-1)$ étapes et lorsque nous sommes à la n^{e} étape[30], nous remplaçons le facteur $e_{k-n}e_{k+1}$ (resp. $e_{k+1}e_{k-n}$) par le facteur[31] $e_{k+1}e_{k-n}$ (resp. $e_{k-n}e_{k+1}$) dans w_1 (resp. w_1^R) jusqu'à ce que w_1

28. Si k vaut 0, le mot w est égal au mot vide et, si k vaut 1, le mot w est égal au mot $e_1^{i_1}$. Par conséquent, vu la définition 3.4.1, nous avons l'égalité entre les mots w et w^R ainsi que l'égalité entre les mots ww^R et $w^R w$. Les mots ww^R et $w^R w$ sont donc équivalents vu la Définition 3.0.1.

29. Dans le Théorème 3.2.4, nous remplaçons i par k, x par $e_1^{i_1} \cdots e_{k-1}^{i_{k-1}}$, y par ε et z par $e_{k-1}^{i_{k-1}} \cdots e_1^{i_1}$. Comme le mot y est égal au mot vide, le nombre d'occurrence de la lettre e_{i-1} dans y vaut 0 et le nombre d'occurrence de la lettre e_{i+2} dans y vaut 0.

30. Vu que nous procédons en $(k-1)$ étapes, l'entier n est compris entre 1 et $k-1$.

31. Nous remplaçons respectivement les entiers i et j du Théorème 3.2.1 par les entier $k-n$ et $n+1$. Comme n est compris entre 1 et $k-1$, l'entier $k-n$ est compris entre 1 et $k-1$ et l'entier $n+1$ est compris entre 2 et k.

(resp. w_1^R) ne contiennent plus le facteur $e_{k-n}e_{k+1}$ (resp. $e_{k+1}e_{k-n}$). Vu le Théorème 3.2.1, il en résulte un mot équivalent à $w_1w_1^R$ s'écrivant comme suit :

$$\underbrace{e_{k+1}^{i_{k+1}}e_1^{i_1}\cdots e_{k-1}^{i_{k-1}}e_k^{i_k}}_{w_2}\underbrace{e_k^{i_k}e_{k-1}^{i_{k-1}}\cdots e_1^{i_1}\cdots e_{k+1}^{i_{k+1}}}_{w_2^R}.$$

Si nous enlevons les i_{k+1} premières lettres et les i_{k+1} dernières lettres du mot $w_2w_2^R$, nous obtenons :

$$\underbrace{e_1^{i_1}\cdots e_{k-1}^{i_{k-1}}e_k^{i_k}}_{w_3}\underbrace{e_k^{i_k}e_{k-1}^{i_{k-1}}\cdots e_1^{i_1}\cdots e_k^{i_k}}_{w_3^R}.$$

Par hypothèse de récurrence, nous savons que le mot $w_3w_3^R$ est équivalent au mot $w_4w_4^R$ représenté comme suit :

$$\underbrace{e_k^{i_k}e_{k-1}^{i_{k-1}}\cdots e_1^{i_1}}_{w_4}\underbrace{e_1^{i_1}\cdots e_{k-1}^{i_{k-1}}e_k^{i_k}}_{w_4^R}.$$

En ajoutant $e_{k+1}^{i_{k+1}}$ au début et à la fin du mot $w_4w_4^R$, nous obtenons le mot suivant étant équivalent au mot [32] $w_2w_2^R$:

$$\underbrace{e_{k+1}^{i_{k+1}}e_k^{i_k}\cdots e_1^{i_1}}_{w^R}\underbrace{e_1^{i_1}\cdots e_k^{i_k}e_{k+1}^{i_{k+1}}}_{w}.$$

Vu ce qui précède, nous avons l'équivalence de Parikh entre les mots ww^R et $w_1w_1^R$, entre les mots $w_1w_1^R$ et $w_2w_2^R$ ainsi qu'entre les mots $w_2w_2^R$ et w^Rw. Par conséquent, les mots ww^R et w^Rw sont équivalents.

□

32. Comme les mots $w_3w_3^R$ et $w_4w_4^R$ sont équivalents, les mots $w_2w_2^R$ et w^Rw sont également équivalents.

Chapitre 4

Matrice à signes alternés et matrice retournée

Ce chapitre est essentiellement basé sur l'article [13]. Dans celui-ci, nous allons introduire des matrices particulières appelées matrices à signes alternés et matrices retournées. Nous pourrons par la suite établir un lien entre ces matrices qui nous sera utile dans le chapitre suivant.

4.1 Matrice à signes alternés

Nous allons commencer par introduire les matrices à signes alternés. Plus précisement, nous allons donner la définition d'une matrice à signes alternés ainsi que quelques propriétés relatives à ces matrices.

Définition 4.1.1. Soient k un entier supérieur à 1 et $A = (a_{i,j})_{1\leq i,j\leq k}$ une matrice carrée de dimension k. La *matrice à signes alternés* issue de A, notée $A^{(alt)} = (a'_{i,j})_{1\leq i,j\leq k}$, est la matrice dont les éléments $a'_{i,j}$, $1 \leq i \leq k$ et $1 \leq j \leq k$, sont donnés par $(-1)^{i+j}a_{i,j}$.

Exemple 4.1.2. Si nous considérons la matrice A représentée ci-dessous, nous pouvons facilement en déduire sa matrice à signes alternés (representée également ci-dessous).

$$A = \begin{pmatrix} 2 & 3 \\ 2 & 2 \end{pmatrix}, A^{(alt)} = \begin{pmatrix} 2 & -3 \\ -2 & 3 \end{pmatrix}.$$

Après avoir défini les matrices à signes alternés, nous pouvons donner quelques propriétés relatives à ces matrices.

Théorème 4.1.3. *Soient $\Sigma = \{e_1, \cdots, e_k\}$ un alphabet ordonné et u un mot sur l'alphabet Σ tel que $u = b_1 \cdots b_{|u|}$ ($b_1, \ldots, b_{|u|} \in \Sigma$). Pour tous mots $w, w' \in \Sigma^*$, nous pouvons écrire la relation suivante :*

$$[\psi_u(w)\psi_u(w')]^{(alt)} = [\psi_u(w)]^{(alt)}[\psi_u(w')]^{(alt)}.$$

Démonstration. Vu le Théorème 2.3.5 et vu la définition du produit matriciel, les éléments de la matrice $\psi_u(w)\psi_u(w') = (m'''_{i,j})_{1\leq i,j\leq |u|+1}$ sont donnés par ce qui suit :

$$m''_{i,j} = \begin{cases} 1 \text{ si } i = j, \\ |w|_{u_{i,j-1}} + |w'|_{u_{i,j-1}} + \sum_{k=1}^{|u|+1} |w|_{u_{i,k-1}}|w'|_{u_{k,j-1}} \text{ si } j > i, \\ 0 \text{ sinon.} \end{cases}$$

Vu le Théorème 2.3.5, vu la Définition 4.1.1 et vu la définition du produit matriciel, nous pouvons en déduire la forme des matrices $[\psi_u(w)]^{(alt)} = (m_{i,k})_{1\leq i,k\leq |u|+1}$, $[\psi_u(w')]^{(alt)} = (m'_{k,j})_{1\leq k,j\leq |u|+1}$ et $[\psi_u(w)\psi_u(w')]^{(alt)} = (m''_{i,j})_{1\leq i,j\leq |u|+1}$:

$$m_{i,k} = \begin{cases} 1 \text{ si } i = k, \\ (-1)^{i+k}|w|_{u_{i,k-1}} \text{ si } k > i, \\ 0 \text{ sinon,} \end{cases}$$

$$m'_{k,j} = \begin{cases} 1 \text{ si } k = j, \\ (-1)^{k+j}|w'|_{u_{k,j-1}} \text{ si } j > k, \\ 0 \text{ sinon,} \end{cases}$$

$$m''_{i,j} = \begin{cases} 1 \text{ si } i = k, \\ (-1)^{i+j}[|w|_{u_{i,j-1}} + |w'|_{u_{i,j-1}} + \sum_{k=1}^{|u|+1} |w|_{u_{i,k-1}}|w'|_{u_{k,j-1}}] \text{ si } j > i, \\ 0 \text{ sinon.} \end{cases}$$

Vu la définition du produit matriciel, vu la forme de la matrice $[\psi_u(w)]^{(alt)}$ et vu la forme de la matrice $[\psi_u(w')]^{(alt)}$, les éléments de la matrice $[\psi_u(w)]^{(alt)}[\psi_u(w')]^{(alt)} = (m'''_{i,j})_{1\leq i,j\leq |u|+1}$ sont représentés comme suit :

$$\begin{aligned} m'''_{i,j} &= \sum_{k=1}^{|u|+1} m_{i,k}m'_{k,j}, \\ &= \begin{cases} 1 \text{ si } i = j, \\ (-1)^{i+j}[|w|_{u_{i,j-1}} + |w'|_{u_{i,j-1}} + \sum_{k=i+1}^{j-1} |w|_{u_{i,k-1}}|w'|_{u_{k,j-1}}] \text{ si } j > i, \\ 0 \text{ sinon.} \end{cases} \end{aligned}$$

Ainsi, nous remarquons que les matrices $[\psi_u(w)\psi_u(w')]^{(alt)}$ et $[\psi_u(w)]^{(alt)}[\psi_u(w')]^{(alt)}$ sont égales, ce qui nous permet de conclure la preuve.

□

Théorème 4.1.4. *Soient $\Sigma = \{e_1, \ldots, e_k\}$ un alphabet ordonné et u un mot sur l'alphabet Σ tel que $u = b_1 \cdots b_{|u|}$ ($b_1, \ldots, b_{|u|} \in \Sigma$ et $b_i \neq b_{i+1}$ pour tout entier i compris entre 1 et $|u| - 1$). Pour tout mot $w \in \Sigma^*$, nous avons ce qui suit :*

$$[\psi_u(w)]^{-1} = [\psi_u(w^R)]^{(alt)}.$$

Démonstration. Nous allons procéder par récurrence sur la longueur du mot w. Si la longueur de w vaut 0, le mot w est égal au mot vide et le mot w^R est égal au mot w. Vu le Théorème 2.3.5 et vu la Définition 4.1.1, nous avons l'égalité entre les matrices $\psi_u(w)$ et $I_{|u|+1}$ ainsi que l'égalité entre les matrices $[\psi_u(w)]^{(alt)}$ et $I_{|u|+1}$. Le théorème est donc satisfait si la longueur de w vaut 0. Si la longueur de w vaut 1, le mot w s'avère être une lettre et vu la Définition 3.4.1, nous avons l'égalité entre w et w^R. Pour vérifier que le théorème est satisfait si la longueur de w vaut 1, nous posons les trois égalités suivantes :

- $\psi_u(w) = (m_{i,j})_{1 \leq i,j \leq (|u|+1)}$,
- $\psi_u(w)[\psi_u(w)]^{(alt)} = (n_{i,j})_{1 \leq i,j \leq (|u|+1)}$,
- $[\psi_u(w)]^{(alt)}\psi_u(w) = (n'_{i,j})_{1 \leq i,j \leq (|u|+1)}$.

Vu ces trois égalités et vu la Définition 4.1.1, nous obtenons ce qui suit pour tous les entiers i et j tels que $\leq i, j \leq |u| + 1$:

- $[\psi_u(w)]_{i,j} = m_{i,j}$
- $[\psi_u(w)]_{i,j}^{(alt)} = (-1)^{i+j} m_{i,j}$,
- $n_{i,j} = \sum_{l=1}^{|u|+1} m_{i,l}(-1)^{l+j} m_{l,j} = (-1)^j \sum_{l=1}^{|u|+1} (-1)^l m_{i,l} m_{l,j}$,
- $n'_{i,j} = \sum_{l=1}^{|u|+1} (-1)^{i+l} m_{i,l} m_{l,j} = (-1)^i \sum_{l=1}^{|u|+1} (-1)^l m_{i,l} m_{l,j}$.

Nous remarquons que, si les éléments $n_{i,j}$ et $n'_{i,j}$ sont distincts, ils diffèrent uniquement à un signe près. Cependant, nous allons montrer que les éléments $n_{i,j}$ et $n'_{i,j}$ sont toujours égaux. Vu ce qui précède et vu la Définition 2.3.1, nous obtenons ce qui suit :

$$\begin{aligned} n_{i,j} &= (-1)^j \sum_{i=1}^{|u|+1} (-1)^l m_{i,l} m_{l,j}, \\ &= \begin{cases} 1 \text{ si } i = j, \\ \delta_{b_i,w}\delta_{b_{i+1},w} \text{ si } j = i + 2, \\ 0 \text{ sinon.} \end{cases} \end{aligned}$$

De plus, nous savons que les lettres b_i, $1 \leq i \leq |u| + 1$, diffèrent des lettres b_{i+1}. Ainsi, nous pouvons en déduire que $n_{i,j}$ est égal à 1 si i est égal à j et que $n_{i,j}$ est égal à 0 si i diffère de j.

Par un raisonnement analogue, nous pouvons déterminer la forme des éléments $n'_{i,j}$, $1 \leq i, j \leq |u| + 1$, et remarquer que ces éléments sont égaux aux éléments $n_{i,j}$. Par conséquent, nous avons l'égalité entre les matrices $\psi_u(w)[\psi_u(w)]^{(alt)}$, $[\psi_u(w)]^{(alt)}\psi_u(w)$ et $I_{|u|+1}$. Le théorème est donc satisfait si la longueur de w vaut 1.

Nous supposons que le théorème est vrai si la longueur de w vaut n $(n \geq 1)$ et nous allons montrer que le théorème reste vrai si la longueur de w vaut $n + 1$.

Dans la suite de cette démonstration, nous notons $w = w'a$ avec w' un mot de longueur $(n - 1)$ sur l'alphabet Σ et a une lettre de Σ. Ainsi, vu que ψ_u est un morphisme, nous avons ce qui suit :

$$\begin{aligned} [\psi_u(w)]^{-1} &= [\psi_u(w'a)]^{-1}, \\ &= [\psi_u(w')\psi_u(a)]^{-1}, \\ &= [\psi_u(a)]^{-1}[\psi_u(w')]^{-1}. \end{aligned}$$

En appliquant l'hypothèse de récurrence ainsi que le Théorème 4.1.3, nous pouvons conclure cette preuve :

$$\begin{aligned} [\psi_u(a)]^{-1}[\psi_u(w')]^{-1} &= [\psi_u(a)]^{(alt)}\ [\psi_u(w'^R)]^{(alt)}, \\ &= [\psi_u(a)\psi_u(w'^R)]^{(alt)}, \\ &= [\psi_u(aw'^R)]^{(alt)}, \\ &= [\psi_u(w^R)]^{(alt)}. \end{aligned}$$

□

Corollaire 4.1.5. *Soient $\Sigma = \{e_1, \cdots, e_k\}$ un alphabet ordonné. Pour tout mot $w \in \Sigma^*$, nous avons la relation suivante :*

$$[\psi_k(w)]^{-1} = [\psi_k(w^R)]^{(alt)}.$$

Démonstration. Ce corollaire s'avère être un cas particulier du Théorème 4.1.4. En effet, vu la Remarque 2.3.3, la fonction ψ_k est égale à la fonction ψ_u où $u = e_1 \cdots e_k$.

□

4.2 Matrice retournée

À présent, nous allons donner une définition en ce qui concerne les matrices retournées et montrer plusieurs propriétés relatives à ce type de matrices.

Définition 4.2.1. Soient k un entier supérieur à 1 et $A = (a_{i,j})_{1\leq i,j\leq k}$ une matrice carrée de dimension k. La *matrice retournée* issue de A, notée $A^{(rev)} = (a'_{i,j})_{1\leq i,j\leq k}$, est la matrice dont les éléments $a'_{i,j}$, $1 \leq i \leq k$ et $1 \leq j \leq k$, sont donnés par $a_{k+1-j,k+1-i}$.

Exemple 4.2.2. Si nous considérons la matrice A représentée ci-dessous, nous pouvons facilement en déduire sa matrice retournée (représentée également ci-dessous) :

$$A = \begin{pmatrix} 2 & 3 \\ 1 & 4 \end{pmatrix}, \quad A^{(rev)} = \begin{pmatrix} 4 & 3 \\ 1 & 2 \end{pmatrix}.$$

Théorème 4.2.3. *Soient $\Sigma = \{e_1, \cdots, e_k\}$ un alphabet ordonné et u un mot sur l'alphabet Σ tel que $u = b_1 \cdots b_{|u|}$ ($b_1, \ldots, b_{|u|} \in \Sigma$). Pour tous mots $w, w' \in \Sigma^*$, nous avons la relation suivante :*

$$[\psi_u(w)\psi_u(w')]^{(rev)} = [\psi_u(w')]^{(rev)}[\psi_u(w)]^{(rev)}.$$

Démonstration. Vu le Théorème 2.3.5 et vu la Définition 4.2.1, nous pouvons en déduire la forme des matrices $[\psi_u(w')]^{(rev)}$ et $[\psi_u(w)]^{(rev)}$:

$$[\psi_u(w')]^{(rev)}_{i,m} = \begin{cases} 1 \text{ si } i = m, \\ |w'|_{u_{|u|+2-m,|u|+1-i}} \text{ si } m > i, \\ 0 \text{ sinon.} \end{cases}$$

$$[\psi_u(w)]^{(rev)}_{m,j} = \begin{cases} 1 \text{ si } m = j, \\ |w|_{u_{|u|+2-j,|u|+1-m}} \text{ si } j > m, \\ 0 \text{ sinon.} \end{cases}$$

Vu le Théorème 2.3.5, vu la Définition 4.2.1 et vu la définition du produit matriciel, nous savons que les éléments $[\psi_u(w)\psi_u(w')]^{(rev)}_{i,j}$, $1 \leq i \leq |u|+1$ et $1 \leq j \leq |u|+1$, sont égaux à 1 si i est égal à j, à 0 si j est strictement inférieur à i et à $|w|_{u_{|u|+2-j,|u|+1-i}} + |w'|_{u_{|u|+2-j,|u|+1-i}} + \sum_{l=1}^{|u|+1} |w|_{u_{|u|+2-j,l-1}} |w'|_{u_{l,|u|+1-i}}$ si j est strictement supérieur à i.

Il nous reste donc à calculer la forme de la matrice $[\psi_u(w')]^{(rev)}[\psi_u(w)]^{(rev)}$. Vu la définition du produit matriciel et vu la forme des matrices $[\psi_u(w')]^{(rev)}$ et $[\psi_u(w)]^{(rev)}$, nous pouvons facilement en déduire que la matrice $[\psi_u(w')]^{(rev)}[\psi_u(w)]^{(rev)}$ est égale à la matrice $[\psi_u(w)\psi_u(w')]^{(rev)}$.

□

Théorème 4.2.4. *Soient $\Sigma = \{e_1, \cdots, e_k\}$ un alphabet ordonné et u un mot sur l'alphabet Σ tel que $u = b_1 \cdots b_{|u|}$ ($b_1, \ldots, b_{|u|} \in \Sigma$). Pour tout mot $w \in \Sigma^*$, nous avons la relation suivante :*

$$\psi_u(w^R) = [\psi_{u^R}(w)]^{(rev)}.$$

Démonstration. Nous allons procéder par récurrence sur la longueur du mot w. Si la longueur de w vaut 0, le mot w est égal au mot vide et le mot w^R est égal au mot w. De plus, vu le Théorème 2.3.5 et vu la Définition 4.2.1, nous avons l'égalité entre les matrices $\psi_u(w)$ et $I_{|u|+1}$ ainsi qu'entre les matrices $[\psi_{u^R}(w)]^{(rev)}$ et $I_{|u|+1}$. Le théorème est donc satisfait si la longueur de w vaut 0. Si la longueur de w vaut 1, le mot w s'avère être une lettre et le mot w^R est égal au mot w. Pour montrer que le théorème est satisfait si la longueur de w vaut 1, nous posons les quatres égalités suivantes :

- $\psi_u(w) = (m_{i,j})_{1 \leq i,j \leq |u|+1}$,
- $\psi_{u^R}(w) = (n_{i,j})_{1 \leq i,j \leq |u|+1}$,
- $[\psi_{u^R}(w)]^{(rev)} = (n'_{i,j})_{1 \leq i,j \leq |u|+1}$,
- $u^R = c_1 \cdots c_{|u|}$ avec $c_i = b_{|u|+1-i}$ pour tout entier i compris entre 1 et $|u|$.

Vu ces quatre égalités, vu la Définition 4.2.1 et vu la Définition 2.3.1, nous pouvons en déduire que $n'_{i,j}$ est égal à $m_{i,j}$ pour chaque entier i compris entre 1 et $|u|$ et pour chaque entier j compris entre 1 et $|u|$. Ainsi, la matrice $\psi_u(w)$ est égale à la matrice $[\psi_{u^R}(w)]^{(rev)}$ et le théorème est satisfait si la longueur du mot w vaut 1.

A présent, nous supposons que le théorème est satisfait si la longueur de w vaut n ($n \geq 1$) et nous allons vérifier qu'il le reste si la longueur de w vaut $n+1$.

Etant donné que n est un entier supérieur à 1, nous pouvons poser $w = w'a$ avec w' un mot de longueur $(n-1)$ sur l'alphabet Σ et a une lettre de Σ. Ainsi, vu la Proposition 3.4.3 et vu le fait que ψ_u est un morphisme, nous obtenons les égalités suivantes :

$$\begin{aligned} \psi_u(w^R) &= \psi_u(aw'^R) \qquad , \\ &= \psi_u(a)\psi_u(w'^R). \end{aligned}$$

Comme la longueur de a et la longueur de w' sont strictement inférieures à la longueur de w, nous pouvons appliquer l'hypothèse de récurrence sur w' et conclure la preuve à l'aide du Théorème 4.2.3 :

$$\begin{aligned} \psi_u(w^R) &= [\psi_{u^R}(a)]^{(rev)}[\psi_{u^R}(w')]^{(rev)}, \\ &= [\psi_{u^R}(w')\psi_{u^R}(a)]^{(rev)}, \\ &= [\psi_{u^R}(w)]^{(rev)}. \end{aligned}$$

□

Théorème 4.2.5. *Soit $\Sigma = \{e_1, \cdots, e_k\}$ un alphabet ordonné. Pour tout mot $w \in \Sigma^*$, nous avons la relation suivante :*

$$\psi_k(w^R) = [\psi_{e_k \cdots e_1}(w)]^{(rev)}.$$

Démonstration. Ce théorème est un cas particulier du Théorème 4.2.4. En effet, vu la Remarque 2.3.3, le morphisme ψ_k est égal au morphisme ψ_u où $u = e_1 \cdots e_k$ et vu la Définition 3.4.1, le miroir de $e_1 \cdots e_k$ est égal à $e_k \cdots e_1$.

□

Remarque 4.2.6. Nous notons $\psi_{\Sigma',k}(w)$ la matrice de Parikh du mot w sur l'alphabet ordonné $\Sigma' = \{e_k, \ldots e_1\}$.

Corollaire 4.2.7. *Soit $\Sigma = \{e_1, \cdots, e_k\}$ un alphabet ordonné. Pour tout mot $w \in \Sigma^*$, nous avons la relation suivante :*

$$\psi_{\Sigma',k}(w)^{(rev)} = \psi_k(w^R).$$

Démonstration. Ce corollaire s'avère être un cas particulier du Théorème 4.2.5. En effet, vu la Remarque 4.2.6 et vu la Définition 2.3.6, nous avons l'égalité entre les matrices $\psi_{e_k \cdots e_1}(w)$ et $\psi_{\Sigma',k}(w)$ ce qui conclut la preuve.

□

4.3 Lien entre les matrices à signes alternés et les matrices retournées

Grâce aux deux sections précédentes, nous pouvons établir un lien entre les matrices à signes alternés et les matrices retournées.

Théorème 4.3.1. *Soient $\Sigma = \{e_1, \cdots, e_k\}$ un alphabet ordonné et u un mot sur l'alphabet Σ tel que $u = b_1 \cdots b_{|u|}$ ($b_1, \ldots, b_{|u|} \in \Sigma$ et $b_i \neq b_{i+1}$ pour tout entier i compris entre 1 et $|u|+1$). Pour tout mot $w \in \Sigma^*$, nous obtenons la relation suivante :*

$$[\psi_u(w)]^{-1} = [\psi_{u^R}(w)^{(rev)}]^{(alt)}.$$

Démonstration. Ce théorème découle du Théorème 4.1.4 et du Théorème 4.2.4.

□

Une fois le théorème 4.3.1 démontré, nous pouvons énoncé le théorème suivant qui s'avère être un cas particulier du Théorème 4.3.1.

Théorème 4.3.2. *Soit $\Sigma = \{e_1, \cdots, e_k\}$ un alphabet ordonné. Pour tout mot $w \in \Sigma^*$, nous avons la relation suivante :*

$$[\psi_k(w)]^{-1} = [\psi_{\Sigma',k}(w)^{(rev)}]^{(alt)}.$$

Démonstration. Ce théorème découle du Corollaire 4.1.5 et du Théorème 4.2.7.

□

Chapitre 5

Mot ambigu

Dans le chapitre intitulé "Mots équivalents", nous avons vu que plusieurs mots pouvaient posséder la même matrice de Parikh. Suite à cette constatation, nous allons nous intéresser aux mots ambigus définis à la Definition 5.0.1.

Définition 5.0.1. Soient $\Sigma = \{e_1, \ldots, e_k\}$ un alphabet ordonné et w un mot sur l'alphabet Σ. Le mot w est ambigu s'il existe un mot w' sur l'alphabet Σ satisfaisant les deux conditions suivantes :

- w' diffère de w,
- w' est équivalent à w.

Exemple 5.0.2. Si nous travaillons sur l'alphabet ordonné $\Sigma = \{a, b\}$, les mots *abbaba* et *baabba* sont ambigus étant donné qu'ils sont différents et qu'ils possèdent la même matrice de Parikh (cfr Proposition 2.3.8).

Remarque 5.0.3. Si le mot w est ambigu, sa matrice de Parikh $\psi_k(w)$ est qualifiée de matrice ambiguë.

Si nous travaillons sur un alphabet ordonné constitué de deux lettres, nous allons voir qu'il est facile de déterminer si un mot est ambigu ou non ambigu à l'aide de la section suivante. Dans le cas d'un alphabet ordonné constitué de plus de deux lettres, nous allons voir que ce n'est plus si simple [1]. Cependant, nous allons trouver des mots ambigus ou non ambigus à l'aide des sections "Mot ambigu et facteur", "Mot ambigu et empreinte", "Mot n-ambigu" et "Mot ambigu, miroir et fonction ϕ".

5.1 Mot ambigu sur un alphabet constitué de deux lettres

Dans cette section, nous allons travailler sur un alphabet ordonné constitué de deux lettres et déterminer si un mot est ambigu ou non ambigu à l'aide du théorème suivant.

1. Nous ne disposons pas, à ce jour, de critères afin de déterminer si un mot est ambigu ou non ambigu sur un alphabet ordonné constitué de plus de deux lettres.

Théorème 5.1.1. *Soient $\Sigma = \{e_1, e_2\}$ un alphabet ordonné et w un mot sur l'alphabet Σ. Le mot w est ambigu si et seulement s'il existe des mots $x, y, z \in \Sigma^*$ tel que $w = xe_1e_2ye_2e_1z$ ou $w = xe_2e_1ye_1e_2z$.*

Démonstration. $\Leftarrow$ Nous procédons par l'absurde en supposant que le mot w n'est pas ambigu. Vu la Définition 5.0.1, il n'existe pas de mot w' distinct de w et équivalent à w. Or, vu le Théorème 3.2.4, cette affirmation est absurde ce qui signifie que le mot w est ambigu.

$\Rightarrow$ Nous procédons par l'absurde en supposant qu'il n'existe pas de mots $x, y, z \in \Sigma^*$ tels que $w = xe_1e_2ye_2e_1z$ ou $w = xe_2e_1ye_1e_2z$. Ainsi, le mot w posséde une des formes suivantes :

$$e_1^{|w|},\ e_2^{|w|},\ e_1^n e_2^{|w|-n},\ e_2^n e_1^{|w|-n},\ e_1^n e_2 e_1^{|w|-n-1} \text{ ou } e_2^n e_1 e_2^{|w|-n-1}$$

où n s'avère être un entier compris entre 1 et $|w| - 1$. Nous pouvons facilement vérifier que, si w possède une des six formes citées ci-dessus, le mot w n'est pas ambigu ce qui est absurde. Il existe donc des mots $x, y, x \in \Sigma^*$ tels que $w = xe_1e_2ye_2e_1z$ ou $w = xe_2e_1ye_1e_2z$ si w est ambigu.

□

5.2 Mot ambigu et facteur

Dans cette section, nous allons travailler sur un alphabet ordonné quelconque et étudier l'ambiguïté d'un mot à l'aide de ses facteurs. Premièrement, nous allons montrer qu'un mot est ambigu s'il possède un facteur ambigu. Deuxièmement, nous allons nous intéresser à une caractéristique des mots non ambigus à savoir leurs facteurs de longueur 2. Troisièmement, nous allons montrer que certains facteurs du mot infini $(e_2e_3 \cdots e_ke_1)^w$ sont ambigus.

Lemme 5.2.1. *Soient $\Sigma = \{e_1, \ldots, e_k\}$ un alphabet ordonné et y un mot sur l'alphabet Σ. Si le mot y est ambigu, tous les mots possédant le facteur y sont ambigus.*

Démonstration. Par hypothèse, nous savons que le mot y est ambigu. Ainsi, vu la Définition 5.0.1, il existe un mot y' sur l'alphabet Σ tel que y' diffère de y et tel que $\psi_k(y')$ est égal à $\psi_k(y)$. Dans la suite de cette démonstration, nous prenons x et z des mots quelconques sur l'alphabet Σ. Le mot $xy'z$ diffère du mot xyz étant donné que le mot y' diffère du mot y et la matrice $\psi_k(xy'z)$ est égale à la matrice $\psi_k(xyz)$ vu la Proposition 3.1.1. Nous pouvons donc conclure la preuve grâce à la Définition 5.0.1 et grâce à la Définition 1.4.3.

□

Remarque 5.2.2. Dans le Lemme 5.2.1, le mot y doit être un facteur et non un sous-mot étant donné que nous pouvons facilement trouver un mot non ambigu possédant au moins un sous-mot ambigu. Par exemple, si nous travaillons sur l'alphabet ordonné $\Sigma = \{a, b, c\}$, le mot *abba* est un sous-mot ambigu du mot $w = abcba$ étant donné que les mots *abba* et *baab* diffèrent et possèdent la même matrice de Parikh. De plus, le mot w n'est pas ambigu étant donné qu'il n'existe pas de mots, différents de w, possédant la même matrice de Parikh que w. En effet, le nombre d'occurrences du mot *abc* dans w vaut 1, le nombre d'occurrences de la lettre c dans w vaut 1 et le nombre d'occurrences des lettres a et b dans w vaut 2. Ainsi, la lettre a n'apparaissant pas dans le sous-mot *abc* doit se positionner après la lettre b constituant le sous-mot *abc* et la lettre b n'apparaissant pas dans le sous-mot *abc* doit se positionner soit avant la lettre a constituant le sous-mot *abc* soit après la lettre c constituant le sous-mot *abc*. Cependant, le nombre d'occurrences du mot *bc* dans w vaut 1 ce qui signifie que la lettre b n'apparaissant pas dans le sous-mot *abc* doit se positionner après la lettre c constituant le sous-mot *abc* et le nombre d'occurrences du mot *ab* dans w vaut 1 ce qui signifie que la lettre a doit se trouver en dernière position. Nous avons donc trouvé un mot non ambigu possèdant un facteur ambigu.

Afin de prouver le Corollaire 5.2.5 déterminant la forme des facteurs de longueur 2 d'un mot non ambigu, nous avons besoin de la proposition suivante.

Proposition 5.2.3. *Soit $\Sigma = \{e_1, \ldots, e_k\}$ un alphabet ordonné. Les mots suivants sont ambigus :*

(1) *$e_i e_j$ où les entiers i et j sont tels que $1 \leq i, j \leq k$ et $|i-j| > 1$,*
(2) *$e_i e_j^{m+2} e_i$ et $e_j e_i e_j^m e_i e_j$ où les entiers i, j, m sont tels que $1 \leq i, j \leq k$, $|i-j| = 1$ et $m \geq 0$,*

Démonstration. Premièrement, nous savons par hypothèse que les entiers i et j sont tels que $|i-j| > 1$. Ainsi, nous pouvons appliquer le Théorème 3.2.1 aux mots $e_i e_j$ et en déduire que ces mots sont ambigus.

Deuxièmement, nous savons par hypothèse que l'entier j est soit égal à l'entier $i-1$ soit égal à l'entier $i+1$ et que l'entier m est positif. Ainsi, nous pouvons appliquer le Théorème 3.2.4 aux mots $e_i e_j^{m+2} e_i$ et $e_j e_i e_j^m e_i e_j$ afin d'en déduire que ces mots sont ambigus.

□

Remarque 5.2.4. Un mot est bien formé si aucun de ces facteurs n'est égal à un mot mentionné dans la Proposition 5.2.3. En d'autres termes, les mots n'étant pas bien formés possèdent au moins un mot ambigu de la Proposition 5.2.3 comme facteur. Par conséquent, vu le Lemme 5.2.1, les mots n'étant pas bien formés sont ambigus.

Grâce à la Proposition 5.2.3, nous pouvons démontrer le corollaire suivant.

Corollaire 5.2.5. *Soient $\Sigma = \{e_1, \cdots, e_k\}$ un alphabet ordonné et w un mot sur l'alphabet Σ. Si le mot w n'est pas ambigu et si le mot $e_i e_j$ avec $0 \leq i \leq k$ et $0 \leq j \leq k$ est un facteur de w, nous avons l'inégalité suivante :*

$$|i-j| \leq 1.$$

Par conséquent, les seuls facteurs de longueur 2 *pouvant apparaître dans* w *sont les mots* e_ie_i, e_ie_{i+1} *et* $e_{i+1}e_i$ *avec* $1 \leq i \leq k-1$.

Démonstration. Si le mot w n'est pas ambigu et si le mot e_ie_j est un facteur de w, nous pouvons affirmer en prenant la contraposée du Lemme 5.2.1 que le mot e_ie_j n'est pas ambigu. Or, vu la Proposition 5.2.3, nous savons que les mots e_ie_j sont ambigus si les entiers i et j sont tels que $1 \leq i \leq k$, $1 \leq j \leq k$ et $|i-j| > 1$. De plus, nous pouvons facilement vérifier que les mots e_ie_j sont non ambigus si les entiers i et j sont tels que $1 \leq i, j \leq k$ et $|i-j| \leq 1$ ce qui nous permet de conclure la preuve.

□

Si nous considérons le mot infini $(e_2e_3\cdots e_ke_1)^w$, nous pouvons en déduire que certains de ses facteurs sont ambigus.

Théorème 5.2.6. *Soient* k *un entier tel que* $k \geq 3$ *et* $\Sigma = \{e_1, \cdots, e_k\}$ *un alphabet ordonné. Il existe un entier positif, noté* $n(k)$*, tel que chaque facteur de longueur supérieure à* $n(k)$ *du mot infini* $(e_2e_3\cdots e_ke_1)^w$ *est ambigu.*

Démonstration. Vu le Théorème 3.2.1, nous pouvons en déduire que les mots contenant le facteur e_ke_1 sont ambigus. Etant donné que les facteurs du mot infini $(e_2e_3\cdots e_ke_1)^w$ contiennent le facteur e_ke_1 si ces derniers sont de longueur supérieure à k, nous pouvons conclure la preuve en prenant $n(k)$ qui vaut k.

□

5.3 Mot n-ambigu

A l'aide de l'article [14], nous allons introduire la notion de n-ambiguité. Cette notion nous sera utile dans la section suivante intitulée "Mot ambigu et empreinte" afin de trouver des mots ambigus.

Définition 5.3.1. Soient $\Sigma = \{e_1, \ldots, e_k\}$ un alphabet ordonné constitué d'au moins deux lettres, n un entier supérieur à 1 et w un mot sur l'alphabet Σ. Le mot w est 0-*ambigu* s'il est égal au mot $e_2e_1e_2$ et le mot w est n-*ambigu* s'il existe des mots w_{n-1} et z sur l'alphabet Σ tels que w_{n-1} est $(n-1)$-ambigu et tel que $w = w_{n-1}zw_{n-1}$.

Une fois la notion de n-ambiguité définie, nous pouvons énoncer le Théorème 5.3.2 relatif à cette notion.

Théorème 5.3.2. *Soient* k *un entier strictement supérieur à* 1, $\Sigma = \{e_1, \ldots, e_k\}$ *un alphabet ordonné et* w *un mot sur l'alphabet* Σ*. Si le mot* w *est* $(k-1)$*-ambigu, ce mot est ambigu.*

Nous ne démontrons pas ce théorème étant donné qu'il nous sera uniquement utile dans la section "Mot ambigu et empreinte" et que sa démonstration se révèle être longue et fastidieuse.

5.4 Mot ambigu et empreinte

Comme pour la section consacrée à la n-ambiguïté, cette section est basée sur l'article [14]. Afin de donner des informations supplémentaires sur l'ambiguïté d'un mot, il est nécessaire d'introduire une nouvelle notion à savoir l'empreinte. Comme nous allons le voir, l'empreinte d'un mot est le mot obtenu en remplaçant les lettres consécutives répétées en une seule occurrence.

Définition 5.4.1. Soient $\Sigma = \{e_1, \ldots, e_k\}$ un alphabet ordonné et w un mot sur l'alphabet Σ. L'*empreinte* du mot w, notée $print(w)$, est le mot obtenu en remplaçant tous les facteurs e_i^j, $1 \leq i \leq k$ et $j \geq 2$, du mot w par e_i.

Exemple 5.4.2. Si nous travaillons sur l'alphabet ordonné $\Sigma = \{a, b, c\}$ et si nous considérons le mot $aabbbaabccc$, l'empreinte de ce mot est égal au mot $ababc$.

À l'aide de la Définition 5.4.1 et à l'aide du Théorème 5.3.2, nous pouvons énoncer ainsi que démontrer deux propriétés nous permettant de trouver des mots ambigus à savoir le Corollaire 5.4.3 et le Théorème 5.4.4.

Corollaire 5.4.3. *Soient k un entier supérieur à 2 et $\Sigma = \{e_1, \ldots, e_k\}$ un alphabet ordonné. Il existe un entier positif $n(k)$ tel que pour chaque entier $n > n(k)$ et pour chaque mot $w \in \Sigma^*$ contenant au moins deux lettres distinctes de Σ, w^n est ambigu.*

Démonstration. Pour démontrer ce corollaire, nous allons procéder par récurrence sur k et considérer w un mot sur l'alphabet Σ contenant au moins deux lettres distinctes de Σ.

Si k vaut 2, le mot w contient les lettres e_1 et e_2 étant donné que le mot w contient au moins deux lettres distinctes sur l'alphabet Σ. Par conséquent, l'empreinte de w contient le facteur e_1e_2 et/ou le facteur e_2e_1 et l'empreinte du mot w^3 contient le facteur $(e_1e_2)^3$ et/ou le facteur $(e_2e_1)^3$ ce qui signifie que la longueur de l'empreinte du mot w^3 est supérieure à 6. Vu le Théorème 5.1.1, nous pouvons en déduire que l'empreinte d'un mot ne peut pas être ambiguë si cette empreinte est de longueur inférieure à 4. Effectivement, un mot ambigu doit avoir une longueur supérieure à 4 et une empreinte de longueur 4 n'est pas ambiguë [2]. Nous pouvons également déduire du Théorème 5.1.1 que les empreintes de longueur supérieure à 5 sont ambiguës. Effectivement, les empreintes de longueur 5 sont ambiguës [3] et les empreintes de longueur strictement supérieure à 5 sont également ambiguë étant donné que l'ensemble des empreintes de longueur $i+1$ ($i \geq 5$) est obtenu en concaténant les mots appartenant à l'ensemble des empreintes de longueur i se terminant par la lettre e_1 avec la lettre e_2 ou en concaténant les mots appartenant à l'ensemble des empreintes de longueur i se terminant par la lettre e_2 avec la lettre e_1. Vu ce qui précède, nous pouvons affirmer que la longueur maximale d'une empreinte non ambiguë vaut 4. Comme la longueur de

2. Si l'empreinte d'un mot est de longueur 4, cette empreinte est soit égal à $e_1e_2e_1e_2$ soit égal à $e_2e_1e_2e_1$. Vu le Théorème 5.1.1, il est évident que ces mots ne sont pas ambigus.

3. Si l'empreinte d'un mot est de longueur 5, cette empreinte est soit égal à $e_1e_2e_1e_2e_1$ soit égal à $e_2e_1e_2e_1e_2$. Ainsi, vu le Théorème 5.1.1, les empreintes de longueurs 5 sont ambigues.

l'empreinte du mot w^3 est supérieure à 6, l'empreinte du mot w^3 est ambiguë et le mot w^3 est également ambigu[4]. Par conséquent, vu le Théorème 5.2.1, nous pouvons facilement conclure que l'entier $n(2) = 3$ convient.

Nous supposons que le corollaire est satisfait pour $k < k'$ (k' un entier supérieur à 3) et nous allons montrer qu'il le reste pour $k = k'$ en envisageant trois cas.

Premièrement, si le mot w^2 n'est pas bien formé, nous pouvons en déduire vu la Remarque 5.2.4 que ce mot est ambigu. Ainsi, vu le Théorème 5.2.1, l'entier $n(k') = 2$ convient.

Deuxièmement, si le mot w^2 est bien formé et si le mot w ne contient pas la lettre e_1, le mot w s'avère être un mot sur l'alphabet ordonné $\{e_2, \ldots, e_{k'}\}$ contenant $(k'-1)$ lettres. Ainsi, nous pouvons appliquer l'hypothèse de récurrence et conclure qu'il existe un entier positif $n(k'-1)$ tel que pour chaque entier $n > n(k'-1)$ et pour chaque mot $w \in \Sigma^*$ contenant au moins deux lettres distinctes de Σ, w^n est ambigu.

Troisièmement, nous allons considérer le cas où le mot w^2 est bien formé et où le mot w contient la lettre e_1. Vu la Remarque 5.2.4, aucun facteur du mot w^2 ne peut être égal à un mot de la forme e_ie_j avec i et j des entiers tels que $1 \leq i, j \leq k'$ et $|i-j| > 1$. Par conséquent, la lettre e_1 peut être précédée ou suivie uniquement de la lettre e_1 ou de la lettre e_2 dans le mot w^2 et comme le mot w contient au moins deux lettres distinctes, le mot w contient également la lettre e_2 ce qui nous permet de dire que le mot w^2 contient un facteur de la forme $e_2e_1^ie_2$ avec $i \in \mathbb{Z}$. Cependant, vu la Remarque 5.2.4, aucun facteur du mot w^2 ne peut être égal à un mot de la forme $e_2e_1^le_2$ avec l un entier tel que $l \geq 2$ ce qui implique que i doit valoir 1. Vu ce qui précède, il existe x et y deux mots sur l'alphabet Σ tels que $w^2 = xe_2e_1e_2y$. Pour clore ce troisième cas, nous notons $w_0 = e_2e_1e_2$, $z_i = yx$ et $w_i = w_{i-1}z_iw_{i-1}$ pour chaque entier i compris entre 1 et $k'-1$. Vu ces notations, le mot w^2 est égal au mot xw_0y et le mot $w^{2^{k'}}$ est égal au mot $xw_{k'-1}y$ qui est ambigu vu la Définition 5.3.1, le Théorème 5.3.2 et la Proposition 5.2.1. Ainsi, vu le Théorème 5.2.1, nous pouvons conclure en considérant $n(k) = 2^{k'}$.

$\square$

Théorème 5.4.4. *Soient k un entier supérieur à 2 et $\Sigma = \{e_1, \ldots, e_k\}$ un alphabet ordonné. Il existe un entier positif $n(k)$ tel que pour tout mot $w \in \Sigma^*$ satisfaisant $|w| \geq n(k)$ et $w = print(w)$, w est ambigu.*

Démonstration. Nous allons montrer qu'il existe un entier positif $m(n, k)$ (avec n un entier compris entre 0 et $k-1$) tel que pour tout mot $w \in \Sigma^*$ satisfaisant $|w| \geq m(n, k)$ et $w = print(w)$, w est ambigu ou possède un facteur n-ambigu. Pour prouver cette affirmation, nous allons procéder par récurrence sur k et prendre un mot w tel que $|w| \geq m(n, k)$ et $w = print(w)$.

4. Voir la Définition 5.4.1 et le Théorème 5.1.1.

Si k vaut 2, le mot w est ambigu[5] lorsque nous remplaçons $m(n,k)$ par 5. Ainsi, le théorème est satisfait si k vaut 2 et si n est un entier quelconque compris entre 0 et $k-1$ ce qui signifie que le cas de base est vérifié.

Nous supposons à présent que le théorème est satisfait si $k < k'$ (k' un entier supérieur à 3) et nous allons montrer en procédant par récurrence sur n qu'il le reste si $k = k'$.

Si n vaut 0, les deux cas suivants nous permettent de dire que le mot w est ambigu ou que le mot w contient un facteur 0-ambigu lorsque nous remplaçons $m(0,k')$ par $m(k'-2,k'-1)+3$.

Premièrement, nous considérons le cas où le mot w possède un facteur u sur l'alphabet ordonné $\{e_2,\ldots,e_{k'}\}$ tel que $|u| \geq m(k'-2,k'-1)$. Dans ce cas, nous pouvons appliquer l'hypothèse de récurrence[6] sur $k'-1$ et en déduire que le mot u est ambigu ou que le mot u possède un facteur $(k'-2)$-ambigu. Premièrement, si le mot u est ambigu, le mot w est également ambigu vu le Théorème 5.2.1. Deuxièmement, si le mot u possède un facteur $(k'-2)$-ambigu, nous pouvons en déduire du Théorème 5.3.2 que le mot u possède un facteur ambigu. Par conséquent, les mots u et w sont ambigus vu le Théorème 5.2.1.

Deuxièmement, nous considérons le cas où le mot w est bien formé et où le mot w ne contient pas de facteur u sur l'alphabet ordonné $\{e_2,\ldots,e_{k'}\}$ tel que $|u| \geq m(k'-2,k'-1)$. Dans ce cas, il existe un entier i compris entre 2 et $|w|-1$ tel qu'une lettre e_1 se trouve à la position i dans le mot w. Vu la Remarque 5.2.4 et vu le fait que w est égal à $print(w)$, la lettre e_1 peut être suivie ou précédée uniquement de la lettre e_2. Ainsi, le mot w contient le facteur $e_2e_1e_2$ qui est 0-ambigu vu la Définition 5.3.1.

À présent, nous supposons que le théorème est satisfait pour $n < n'$ (n' un entier supérieur à 1) et nous allons montrer qu'il le reste pour $n = n'$. Plus précisément, nous allons montrer que le mot w est ambigu ou que le mot w possède un facteur n'-ambigu lorsque nous remplaçons $m(n',k')$ par $(w(n',k')+1)\cdot m(n'-1,k')$ avec $w(n',k')$ le nombre possible de mots de longueur[7] $m(n'-1,k')$ sur l'alphabet ordonné $\{e_1,\ldots,e_{k'}\}$.

Étant donné que le mot w est de longueur supérieure à $(w(n',k')+1)\cdot m(n'-1,k')$, nous pouvons diviser le mot w en facteurs de longueurs $m(n'-1,k')$ (où le dernier facteur

5. Vu le Théorème 5.1.1, l'empreinte d'un mot construit sur un alphabet ordonné constitué de deux lettres est ambiguë si la longueur de cette empreinte est supérieure à 5. Pour plus de détail, nous pouvons lire la preuve du Corollaire 5.4.3.

6. Comme le mot u est un facteur du mot w et comme le mot w est égal au mot $print(w)$, le mot u est égal au mot $print(u)$.

7. Le nombre $w(n',k')$ est égal à $k'^{m(n'-1,k')}$. En effet, comme nous travaillons sur l'alphabet $\{e_1,\ldots,e_{k'}\}$, la i^{e} lettre d'un mot de longueur $m(n'-1,k')$ peut prendre k' valeurs distinctes et en effectuant ce raisonnement pour toutes les lettres d'un mot de longueur $m(n'-1,k')$, nous pouvons en déduire qu'il existe $k'^{m(n'-1,k')}$ mots de longueur $m(n'-1,k')$.

de w peut être de longueur inférieure à $m(n'-1,k')$) et en déduire qu'il existe au moins $w(n',k')+1$ facteurs de longueur $m(n'-1,k')$ dans w. Vu la définition de $w(n',k')$, nous savons qu'au moins deux de ces facteurs sont identiques. Par conséquent, il existe un mot $v \in \Sigma^*$ de longueur $m(n'-1,k')$ ainsi que des mots $x, y, z \in \Sigma^*$ tels que $w = xvyvz$. En appliquant l'hypothèse de récurrence sur $n'-1$, nous pouvons en déduire que le mot v est ambigu ou que le mot v possède un facteur $(n'-1)$-ambigu. Vu le Théorème 5.2.1, le mot w est ambigu si le mot v est ambigu et vu la Définition 5.3.1, le mot vyv est n'-ambigu si le mot v est $(n'-1)$-ambigu, ce qui implique que le mot w possède un facteur n'-ambigu.

Vu ce qui précède, nous pouvons affirmer qu'il existe un entier positif $m(n,k)$ (avec n un entier compris entre 0 et $k-1$) tel que pour tout mot $w \in \Sigma^*$ satisfaisant $|w| \geq m(n,k)$ et $w = print(w)$, w est ambigu ou possède un facteur n-ambigu. Comme cette affirmation est vérifiée pour tout entier n compris entre 0 et $k-1$, elle est en particulier vraie si nous remplaçons n par $k-1$. Par conséquent, nous pouvons considérer $n(k)$ qui vaut $m(k-1,k)$ et conclure cette preuve grâce au Théorème[8] 5.3.2 et au Théorème 5.2.1.

□

5.5 Mot ambigu, miroir et fonction ϕ

Dans cette section basée sur l'article [13], nous allons regarder quatre mots bien particuliers et montrer que, si un de ces mots est ambigu, les trois autres mots sont également ambigus.

Définition 5.5.1. Soit $\Sigma = \{e_1, \ldots, e_k\}$ un alphabet ordonné. La *fonction* $\phi : \Sigma^* \to \Sigma^*$ est le morphisme (de monoïdes[9]) qui à chaque lettre e_i de Σ associe la lettre e_{k-i+1} de Σ.

Proposition 5.5.2. *Soit $\Sigma = \{e_1, \ldots, e_k\}$ un alphabet ordonné. Pour tout mot $w \in \Sigma^*$, nous pouvons écrire la relation suivante :*

$$\psi_k(\phi(w)) = \psi_{\Sigma',k}(w).$$

Démonstration. Cette proposition découle de la Définition 5.5.1, de la Remarque 4.2.6 et du Théorème 2.3.6.

□

8. Comme k est un entier supérieur à 2 et comme nous remplaçons n par $k-1$, les hypothèses du Théorème 5.3.2 sont satisfaites.

9. Vu la définition d'un monoïde, il est évident que $(\Sigma^*, \cdot, \varepsilon)$ est un monoïde. Effectivement, vu la Définition 1.3.1, la concaténation est une opération binaire, interne, partout définie, associative et dont le neutre est le mot vide.

Proposition 5.5.3. *Soit $\Sigma = \{e_1, \ldots, e_k\}$ un alphabet ordonné. Pour tout mot $w \in \Sigma^*$, les phrases suivantes sont équivalentes :*

(1) *le mot w est ambigu,*
(2) *le mot w^R est ambigu,*
(3) *le mot $\phi(w)$ est ambigu,*
(4) *le mot $(\phi(w))^R$ est ambigu.*

Démonstration. Premièrement, les phrases (1) et (2) sont équivalentes. En effet, vu le Théorème 2.3.6, vu le Théorème 4.1.4 et vu la Définition 4.1.1, nous pouvons en déduire ce qui suit :

$$\begin{aligned}\psi_k(w) = \psi_k(w') &\Leftrightarrow [\psi_k(w)]^{-1} = [\psi_k(w')]^{-1},\\ &\Leftrightarrow [\psi_k(w^R)]^{(alt)} = [\psi_k(w'^R)]^{(alt)},\\ &\Leftrightarrow \psi_k(w^R) = \psi_k(w'^R).\end{aligned}$$

Si nous supposons que le mot w (resp. w^R) est ambigu, il existe un mot w' (resp. w'^R) qui est différent de w (resp. w^R) et qui possède la même matrice de Parikh que w (resp. w^R). Ainsi, vu ce qui précède, les mots w^R (resp. w) et w'^R (resp. w') possèdent la même matrice de Parikh et comme les mots w (resp. w^R) et w' (resp. w'^R) sont différents, les mots w^R (resp. w) et w'^R (resp. w') sont aussi différents. Vu la Définition 5.0.1, le mot w^R (resp. w) est donc ambigu.

Deuxièmement, les phrases (1) et (3) sont équivalentes. En effet, vu le Théorème 4.3.1 et vu le Théorème 5.5.2, nous pouvons en déduire ce qui suit :

$$\begin{aligned}\psi_k(\phi(w)) = \psi_k(\phi(w')) &\Leftrightarrow \psi_{\Sigma',k}(w) = \psi_{\Sigma',k}(w'),\\ &\Leftrightarrow [\psi_{\Sigma',k}(w)]^{(alt)} = [\psi_{\Sigma',k}(w')]^{(alt)},\\ &\Leftrightarrow [(\psi_{\Sigma',k}(w))^{(alt)}]^{(rev)} = [(\psi_{\Sigma',k}(w'))^{(alt)}]^{(rev)},\\ &\Leftrightarrow [\psi_k(w)]^{-1} = [\psi_k(w')]^{-1},\\ &\Leftrightarrow \psi_k(w) = \psi_k(w').\end{aligned}$$

Si nous supposons que le mot w (resp. $\phi(w)$) est ambigu, il existe un mot w' (resp. $\phi(w')$) qui est différent de w (resp. $\phi(w)$) et qui possède la même matrice de Parikh que w (resp. $\phi(w)$). Ainsi, vu ce qui précède, les mots $\phi(w)$ (resp. w) et $\phi(w')$ (resp. w') possèdent la même matrice de Parikh et vu la Définition 5.5.1, les mots $\phi(w)$ (resp. w) et $\phi(w')$ (resp. w') sont différents étant donné que les mots w (resp. $\phi(w)$) et w' (resp. $\phi(w')$) sont différents. Vu la Définition 5.0.1, le mot $\phi(w)$ (resp. w) est donc ambigu.

Troisièmement, les phrases (1) et (4) sont équivalentes. En effet, étant donné que les phrases (1) et (2) sont équivalentes et que les phrases (1) et (3) sont équivalentes, nous pouvons en déduire ce qui suit :

$$\begin{aligned}\psi_k(w) = \psi_k(w') &\Leftrightarrow \psi_k(\phi(w)) = \psi_k(\phi(w')),\\ &\Leftrightarrow \psi_k((\phi(w))^R) = \psi_k((\phi(w'))^R).\end{aligned}$$

Si nous supposons que le mot w (resp. $(\psi(w))^R$) est ambigu, il existe un mot w' (resp. $(\psi(w'))^R$) qui est différent de w (resp. $(\psi(w))^R$) et qui possède la même matrice de Parikh que w (resp. $(\psi(w))^R$). Vu ce qui précède, les mots $(\psi(w))^R$ (resp. w) et $(\psi(w'))^R$ (resp. w') possèdent la même matrice de Parikh. De plus, vu la Définition 5.5.1 et vu la Définition 3.4.1, les mots $(\psi(w))^R$ (resp. w) et $(\psi(w'))^R$ (resp. w') sont différents étant donné que les mots w (resp. $(\psi(w))^R$) et w' (resp. $(\psi(w'))^R$) sont différents. Par conséquent, nous pouvons en déduire vu la Définition 5.0.1 que le mot $(\psi(w))^R$ (resp. w) est donc ambigu.

□

Chapitre 6

Mots élémentairement équivalents

Les règles 1 et 2 ont été respectivement définies à l'aide du Théorème 3.2.1 et à l'aide du Théorème 3.2.4. Ces règles nous permettent de définir la notion d'équivalence élémentaire de Parikh rappelée ci-dessous.

Définition 6.0.1. Soient $\Sigma = \{e_1, \ldots, e_k\}$ un alphabet ordonné et w, w' des mots sur l'alphabet Σ. Les mots w et w' sont é*lémentairement* é*quivalents*, ce que l'on note $w \equiv_k^E w'$, si le mot w' est obtenu en appliquant la règle 1 et/ou la règle 2 un nombre fini de fois au mot w.

Ce chapitre est essentiellement basé sur l'article [1] et sur l'article [2]. Dans celui-ci, nous allons d'abord montrer que deux mots élémentairement équivalents sont équivalents peu importe l'alphabet ordonné sur lequel nous travaillons. Puis, nous allons montrer que deux mots équivalents sont élémentairement équivalents si nous travaillons sur un alphabet ordonné constitué de deux lettres.

Théorème 6.0.2. *Soient $\Sigma = \{e_1, \ldots, e_k\}$ un alphabet ordonné et w, w' deux mots sur l'alphabet Σ. Si les mots w et w' sont élémentairement équivalents, ces mots sont équivalents.*

Démonstration. Par hypothèse, les mots w et w' sont élémentaitements équivalents. Ainsi, vu la Définition 6.0.1, le mot w' est obtenu en appliquant un nombre fini de fois la règle 1 et/ou la règle 2 au mot w. Vu le Théorème 3.2.1 et vu le Théorème 3.2.4, le mot w' est équivalent au mot w.

$\square$

À présent, nous allons travailler sur un alphabet ordonné constitué de deux lettres et montrer que, si deux mots sont équivalents, un de ces mots est obtenu en appliquant un nombre fini de fois la règle 2 à l'autre mot. Cependant, afin de prouver cette affirmation, il est nécessaire d'introduire la Définition 6.0.4, le Théorème 6.0.7 et la Proposition 6.0.8.

Remarque 6.0.3. Dans la Définition 6.0.4, il est nécessaire d'introduire la notion de graphe non orienté. Pour rappel, le graphe $G = (V, E)$ est la donnée du couple (V, E) où V est un ensemble (fini ou infini) et où E est une partie de $V \times V$ (i.e. une relation sur V). Les éléments de V sont appelés les sommets ou noeuds de G et les éléments de E sont appelés les arcs ou arêtes de G. Si E est une relation symétrique sur V, nous dirons que G est un graphe non orienté. Autrement dit, G est non orienté si : $\forall v_1, v_2 \in V : (v1, v2) \in E \Rightarrow (v2, v1) \in E$.

Définition 6.0.4. Soit M une matrice de Parikh. Le *graphe* Γ_M est le graphe non orienté, noté (V, E), où les ensembles V et E sont définis comme suit :

- V est l'ensemble des mots ayant M comme matrice de Parikh,
- le couple (α, β) appartient à l'ensemble E si et seulement si le mot β est obtenu en appliquant [1] un nombre fini de fois la règle 2 au mot α.

Remarque 6.0.5. Vu le Théorème 3.2.4, nous pouvons en déduire que les mots α et β introduits dans la Définition 6.0.4 sont équivalents.

Remarque 6.0.6. Soient $\Sigma = \{e_1, e_2\}$ un alphabet ordonné et α un mot sur l'alphabet Σ. Étant donné que Σ est un alphabet ordonné constitué de deux lettres, le mot α peut s'écrire $e_1^{x_1} e_2 e_1^{x_2} e_2 \cdots e_1^{x_n} e_2 e_1^{x_{n+1}}$ avec $x_1, \ldots, x_n$ et x_{n+1} des entiers positifs. De plus, la matrice représentée ci-dessous :

$$\begin{pmatrix} 1 & p & q \\ 0 & 1 & n \\ 0 & 0 & 1 \end{pmatrix}$$

avec n, p, q des entiers positifs est la matrice de Parikh associée au mot α si et seulement si $(x_1, \ldots, x_{n+1})$ est une solution du système suivant :

$$\begin{cases} x_1 + x_2 + \cdots + x_{n-1} + x_n + x_{n+1} = p, \\ nx_1 + (n-1)x_2 + \cdots + 2x_{n-1} + x_n = q. \end{cases}$$

Théorème 6.0.7. *Soient $\Sigma = \{e_1, e_2\}$ un alphabet ordonné et M une matrice de Parikh représentée comme suit :*

$$M = \begin{pmatrix} 1 & p & q \\ 0 & 1 & n \\ 0 & 0 & 1 \end{pmatrix}$$

où n, p et q sont des entiers positifs. Si n, p et q sont tels que $n \geq 2$, $p \geq 2$ et $q \in [2, n \cdot p - 2]$, il existe au moins deux mots associés à la matrice de Parikh M.

1. Voir le Théorème 3.2.4.

Démonstration. Si p est supérieur à q, il existe un entier positif s tel que p est égal à la somme de q et de s. Dans la suite de cette démonstration, nous considérons t un entier satisfaisant les inégalités suivantes :

$$0 \leq t \leq \frac{q}{2}.$$

Vu ces inégalités, nous pouvons en déduire que les nombres $t, q-2t$ et $t+s$ sont des entiers positifs et qu'une solution du système donné à la Remarque 6.0.6 s'avère être $(0, \ldots, 0, t, q-2t, t+s)$ où le chiffre 0 occupe les $(n-2)$ premières positions[2]. Ainsi, vu la Remarque 6.0.6, le mot suivant est associé à la matrice de Parikh M :

$$e_2^{n-2} e_1^t e_2 e_1^{q-2t} e_2 e_1^{t+s}.$$

Pour chaque t, nous obtenons une solution distincte et comme q est supérieur à 2 par hypothèse, il existe au moins deux mots distincts associés à la matrice[3] de Parikh M.

Il nous reste à traiter le cas où q est strictement supérieur à p. Dans ce cas, il existe des entiers positifs s et r tels[4] que $1 \leq s < n$, $0 \leq r < p$ et $q = s \cdot p + r$. Vu la définition des nombres r et s, les nombres r et $p - r$ sont des entiers positifs et une solution du système décrit à la Remarque 6.0.6 s'avère être $(0, \ldots, 0, r, p - r, 0, \ldots, 0)$ où le chiffre 0 occupe les $(n - s - 1)$ premières positions[5] ainsi que les s dernières positions. Ainsi, vu la Remarque 6.0.6, le mot suivant est associé à la matrice de Parikh M :

$$\alpha = e_2^{n-s-1} e_1^r e_2 e_1^{p-r} e_2^s.$$

Vu le Théorème 3.2.4, nous pouvons trouver un mot différent de α étant équivalent à α en envisageant deux possibilités à savoir le cas où r diffère de 0 et le cas où r est égal 0. Premièrement, si r est strictement supérieur à 0, nous pouvons appliquer[6] la règle 2 au mot α pour obtenir le mot suivant qui est différent de α et équivalent à α :

$$e_2^{n-s-1} e_1^{r-1} e_2 e_1^{p-r} e_2 e_1 e_2^{s-1}.$$

Deuxièmement, si r vaut 0, nous pouvons à nouveau appliquer[7] la règle 2 au mot α afin d'obtenir le mot suivant qui est différent de α et équivalent à α :

$$e_2^{n-s-1} e_1 e_2 e_1^{p-2} e_2 e_1 e_2^{s-1}.$$

2. Le nombre n est supérieur à 2.
3. Comme q est supérieur 2, l'entier t peut au moins prendre deux valeurs à savoir 0 et 1.
4. Il s'agit d'une division euclidienne. Cependant, l'entier positif s ne peut pas valoir $n-1$ lorsque r vaut $p-1$ et l'entier positif s ne peut pas valoir 1 lorsque r vaut $p-1$ et lorsque n vaut 2.
5. L'entier s est compris entre 1 et $n-1$. Ainsi, le nombre $n-s-1$ est un entier positif.
6. Les nombres r et $p-r$ sont des entiers positifs étant r est strictement inférieur à p et que r est strictement supérieur à 0.
7. L'entier p est supérieur à 2.

Ainsi, dans les deux configurations envisagés ci-dessus à savoir le cas où r diffère de 0 et le cas où r est égal à 0, nous avons trouvé un mot différent de α et équivalent à α. Par conséquent, il existe au moins deux mots distincts associés à la matrice M, ce qui nous permet de conclure.

□

La proposition suivante est une propriété du nombre d'occurrences qui nous sera utile afin de démontrer le Théorème 6.0.9.

Proposition 6.0.8. *Soit $\Sigma = \{e_1, \ldots, e_k\}$ un alphabet ordonné. Pour tout mot w sur l'alphabet Σ et pour toutes lettres distinctes a, b de Σ, la relation suivante est satisfaite :*

$$(|w|_a) \cdot (|w|_b) = |w|_{ab} + |w|_{ba}.$$

Démonstration. Pour démontrer cette proposition, nous allons procéder par récurrence sur la longueur du mot w. Si la longueur de w vaut 0, la proposition est satisfaite [8] ce qui signifie que le cas de base est vérifié.

Nous supposons à présent que la proposition est vraie si la longueur de w est strictement inférieure à n ($n \in \mathbb{Z}_0$) et nous allons vérifier qu'elle reste vraie si la longueur de w vaut n.

Nous notons $w = w'e_l$ avec w' un mot de longueur $(n-1)$ sur l'alphabet Σ et e_l une lettre appartenant à Σ. Comme la longueur de w' est strictement inférieure à celle de w, nous pouvons appliquer l'hypothèse de récurrence sur w' et comme le mot w est égal au mot $w'e_l$, nous pouvons en déduire l'égalité entre $|w|_a$ et $|w'|_a + \delta_{e_l,a}$, l'égalité entre $|w|_{ba}$ et $|w'|_{ba} + |w'|_b$, l'égalité entre $|w|_b$ et $|w'|_b + \delta_{e_l,b}$ ainsi que l'égalité entre $|w|_{ab}$ et $|w'|_{ab} + |w'|_a\delta_{e_l,b}$. Par conséquent, la proposition est satisfaite étant donné que les lettres a et b sont distinctes et que nous pouvons écrire ce qui suit :

$$\begin{aligned}
|w|_a \cdot |w|_b &= (|w'|_a + \delta_{e_l,a}) \cdot (|w'|_b + \delta_{e_l,b}),\\
&= (|w'|_a \cdot |w'|_b) + |w'|_a \cdot \delta_{e_l,b} + |w'|_b \cdot \delta_{e_l,a} + \delta_{e_l,a} \cdot \delta_{e_l,b},\\
&= |w'|_{ab} + |w'|_{ba} + |w'|_a \cdot \delta_{e_l,b} + |w'|_b \cdot \delta_{e_l,a} + \delta_{e_l,a} \cdot \delta_{e_l,b},\\
&= |w'|_{ab} + |w'|_{ba} + |w'|_a \cdot \delta_{e_l,b} + |w'|_b \cdot \delta_{e_l,a},\\
&= |w'|_{ba} + |w'|_b \cdot \delta_{e_l,a} + |w'|_{ab} + |w'|_a \cdot \delta_{e_l,b},\\
&= |w|_{ba} + |w|_{ab}.
\end{aligned}$$

□

8. Si la longueur du mot w vaut 0, le mot w est égal au mot vide. Vu la Définition 2.0.5, le nombre d'occurrences d'un mot quelconque dans le mot vide vaut 0.

Grâce à la Proposition 6.0.8, qui est une propriété relative aux nombres d'occurrences, nous pouvons démontrer le Théorème 6.0.9.

Théorème 6.0.9. *Soient $\Sigma = \{e_1, e_2\}$ un alphabet ordonné et α, β des mots sur l'alphabet Σ. Si les mots α et β sont équivalents, le mot β est obtenu en appliquant un nombre fini de fois la règle 2 au mot α.*

Démonstration. Vu la Définition 3.0.1, les mots α et β possèdent la même matrice de Parikh que nous notons M. Vu la Définition 6.0.4, le fait de montrer que des mots équivalents sont obtenus les uns par rapport aux autres en appliquant un nombre fini de fois la règle 2 revient à montrer que le graphe Γ_M est connexe. Si l'ensemble V contient un seul élément, il s'avère évident que le graphe Γ_M est connexe [9]. Ainsi, dans la suite de cette démonstration, nous allons considérer que l'ensemble V contient au moins deux éléments et nous allons montrer que le graphe Γ_M est connexe. Comme la règle 2 peut souvent s'appliquer de différentes manières [10], nous imposons la condition suivante :

Soient x_i, $1 \leq i \leq n$, des entiers positifs et $\alpha = e_1^{x_1} e_2 e_1^{x_2} e_2 \cdots e_1^{x_n} e_2 e_1^{x_{n+1}}$ un mot appartenant à l'ensemble V. Si i et j sont les deux premiers entiers tels que $1 \leq i+1 < j \leq n$, $x_i \neq 0$ et $x_j \neq 0$, alors le mot α peut se réécrire $\gamma_1 e_1^{x_i} e_2 \gamma_2 e_2 e_1^{x_j} \gamma_3$ avec γ_1, γ_2 et γ_3 des mots sur l'alphabet Σ et le mot α' obtenu en appliquant la règle 2 au mot α s'écrit $\gamma_1 e_1^{x_i - 1} e_2 e_1 \gamma_2 e_1 e_2 e_1^{x_j - 1} \gamma_3$.

Cependant, même si nous pouvons appliquer [11] la règle 2 au mot α, la condition ci-dessus n'est pas toujours applicable. En effet, nous ne pouvons pas appliquer cette condition au mot α, si nous sommes dans un des cas suivants :

(1) Si l'entier i est l'unique entier compris entre 1 et $n+1$ tel que $0 < x_i \leq |\alpha|_{e_1}$, nous ne pouvons pas appliquer la condition. Dans ce cas, l'entier positif x_i vaut $|\alpha|_{e_1}$ et le mot α s'écrit [12] $e_2^{i-1} e_1^{x_i} e_2^{n-i+1}$.

(2) Si l'entier i est tel que $1 \leq i \leq n$, $0 < x_i < |\alpha|_{e_1}$, $0 < x_{i+1} < |\alpha|_{e_1}$ et $x_j = 0$ pour tous les entiers j différents de i et $i+1$, nous ne pouvons pas appliquer la condition. Dans ce cas, l'entier positif x_{i+1} vaut $|\alpha|_{e_1} - x_i$ et le mot α s'écrit $e_2^{i-1} e_1^{x_i} e_2 e_1^{x_{i+1}} e_2^{n-i}$.

Dans la suite de cette démonstration, nous notons p au lieu de $|\alpha|_{e_1}$ et q au lieu de $|\alpha|_{e_1 e_2}$. En appliquant autant de fois que possible la règle 2 décrite dans la condition citée ci-dessus au mot α, nous obtenons le mot α_0 de la forme $e_2^{i-1} e_1^{x_i} e_2 e_1^{p-x_i} e_2^{n-i}$ avec $1 \leq i \leq n$ et $0 \leq x_i \leq p$. Vu le Théorème 3.2.4, les mots α et α_0 sont équivalents ce qui nous permet de

9. Un sommet est relié à lui même à l'aide d'un chemin de longueur 0.

10. Par exemple, si nous appliquons la règle 2 au mot $e_1e_2e_1e_2e_2e_1$, nous pouvons obtenir soit le mot $e_2e_1e_1e_2e_1e_2$ soit le mot $e_1e_2e_2e_1e_1e_2$.

11. Le mot α appartient à l'ensemble V et comme l'ensemble V contient au moins deux éléments, le mot α est ambigu vu la Définition 6.0.4. Ainsi, vu le Théorème 5.1.1, nous pouvons appliquer la règle 2 à α.

12. Comme nous pouvons appliquer la règle 2 au mot α, le nombre x_i doit être supérieur à 2.

dire que le nombre d'occurrences du mot e_1e_2 dans le mot α_0 vaut q. En d'autres termes, l'entier q est égal à $x_i + (n-i) \cdot p$ c'est-à-dire que x_i est le reste et $(n-i)$ le quotient [13] de la division euclidienne de q par p. Si nous procédons de manière analogue pour le mot β, qui est un mot de l'ensemble V distinct [14] de α, nous obtenons le mot suivant :

$$\beta_0 = e_2^{n-s-1} e_1^r e_2 e_1^{p-r} e_2^s$$

où r est le reste et s le quotient de la division euclidienne de $|\beta|_{e_1e_2}$ par $|\beta|_{e_1}$. Vu la Définition 6.0.4, les mots α et β sont équivalents ce qui implique que nous obtenons l'égalité entre $|\beta|_{e_1}$ et p, l'égalité entre $|\beta|_{e_2}$ et n ainsi que l'égalite entre $|\beta|_{e_1e_2}$ et q. Par conséquent, les mots α_0 et β_0 sont égaux et le graphe Γ_M est connexe [15].

□

Le Théorème 6.0.2 nous permet de démontrer le théorème suivant affirmant que deux mots équivalents sont élémentairement équivalents si nous travaillons sur un alphabet ordonné constitué de deux lettres.

Théorème 6.0.10. *Soit $\Sigma = \{e_1, e_2\}$ un alphabet ordonné. Si deux mots sur l'alphabet Σ sont équivalents, ils sont élémentairement équivalents.*

Démonstration. La preuve de ce théorème est triviale vu la Définition 6.0.1 et vu le Théorème 6.0.9.

□

Vu ce qui précède, nous pouvons conclure ce chapitre à l'aide du théorème 6.0.11 affirmant que, si nous travaillons sur un alphabet ordonné constitué de deux lettres, deux mots sont équivalents si et seulement si ces mots sont élémentairement équivalents.

Théorème 6.0.11. *Soient $\Sigma = \{e_1, e_2\}$ un alphabet ordonné et u, v des mots sur l'alphabet Σ. Les mots u et v sont équivalents si et seulement s'ils sont élémentairement équivalents.*

Démonstration. La preuve de ce théorème est immédiate vu le Théorème 6.0.2 et le Théorème 6.0.10.

□

13. Si q est strictement inférieur à p, le quotient de la division euclidienne de q par p vaut 0 c'est-à-dire que n vaut i.
14. Nous avons supposé que l'ensemble V contenait au moins deux mots.
15. Il existe un chemin reliant les mots α et α_0 ainsi qu'un chemin reliant les mots β et β_0. Comme les mots α_0 et β_0 sont égaux, il existe un chemin reliant les mots α et β.

Chapitre 7

Égalité de mots

Ce chapitre est essentiellement basé sur l'article [8] et sur l'article [13]. Dans celui-ci, nous allons énoncer des critères à satisfaire afin de déterminer un mot unique. Ces critères seront d'abord déterminés dans le cas d'un alphabet ordonné constitué de deux lettres puis, dans le cas d'un alphabet ordonné quelconque. Cependant, avant de citer les critères à satisfaire pour déterminer un mot unique sur un alphabet ordonné constitué de deux lettres, il est nécessaire d'introduire les formules de Newton-Girard.

7.1 Formules de Newton-Girard

Les formules de Newton [1]-Girard [2], aussi appelées identités de Newton, nous procurent une relation entre les polynômes symétriques élémentaires et les sommes de Newton. Avant d'énoncer cette relation, il est nécessaire de définir les polynômes symétriques élémentaires ainsi que les sommes de Newton.

Définition 7.1.1. Soit n un entier strictement positif. Pour tout entier k compris entre 1 et n, le k^{e} *polynôme symétrique élémentaire à n variables* est la somme des différents produits composés de k variables distinctes parmi les n variables. Mathématiquement, si nous notons $x_1, \ldots, x_n$ les n variables, le k^{e} *polynôme symétrique élémentaire à n variables* s'écrit comme suit :

$$s_k(x_1, \cdots, x_n) = \textstyle\sum_{1<i_1<\ldots<i_k<n} x_{i_1} \cdots x_{i_k}.$$

Exemple 7.1.2. Si n vaut 3 et si k vaut 2, le deuxième polynôme symétrique élémentaire à 3 variables, noté $s_2(x_1, x_2, x_3)$, s'avère être le polynôme $x_1x_2 + x_1x_3 + x_2x_3$.

Remarque 7.1.3. Par convention, nous notons $s_0(x_1, \cdots, x_n) = 1$ et $s_k(x_1, \ldots, x_n) = 0$ si k est strictement supérieur à n.

1. Isaac Newton est un physicien, philosophe, astronome et mathématicien anglais né en 1642 et mort en 1727.
2. Albert Girard est un mathématicien français né en 1595 et mort en 1632.

Définition 7.1.4. Soient n un entier strictement positif et $x_1, \cdots, x_n$ des variables. Pour tout entier strictement positif k, la k^e *somme de Newton* est égale à la somme des puissances k-ièmes des n variables. Mathématiquement, la k^e *somme de Newton* s'écrit comme suit :

$$p_k(x_1, \cdots, x_n) = \Sigma_{i=1}^n x_i^k.$$

Exemple 7.1.5. Si k vaut 2 et si n vaut 3, la deuxième somme de Newton s'écrit $x_1^2+x_2^2+x_3^2$.

Comme nous allons l'observer dans la Définition 7.1.6, nous pouvons établir un lien entre les polynômes symétriques élémentaires et les sommes de Newton définis respectivement à la Définition 7.1.1 et à la Définition 7.1.4.

Définition 7.1.6. Soient n un entier strictement positif et $x_1, \cdots, x_n$ des variables. Les formules suivantes sont les *formules de Newton-Girard* :

$$ks_k(x_1, \cdots, x_n) = \sum_{i=1}^k (-1)^{i-1} s_{k-i}(x_1, \cdots, x_n) p_i(x_1, \cdots, x_n), \ \forall k \in N_0.$$

Exemple 7.1.7. Afin de comprendre correctement la Définition 7.1.6, nous allons l'illustrer sur un exemple. Si nous considérons $n = 3$ et $k = 2$, la formule de Newton-Girard s'écrit comme suit :

$$2s_2(x_1, x_2, x_3) = s_1(x_1, x_2, x_3)p_1(x_1, x_2, x_3) - p_2(x_1, x_2, x_3).$$

Exemple 7.1.8. Si nous considérons $n = 3$ et $k = 1$, la formule de Newton-Girard s'écrit comme suit :

$$s_1(x_1, x_2, x_3) = p_1(x_1, x_2, x_3).$$

Remarque 7.1.9. Vu la Définition 7.1.6, nous pouvons exprimer les sommes de Newton en fonction des polynômes symétriques élémentaires et inversément. Afin de comprendre cette remarque, nous pouvons l'illustrer sur un exemple. Vu l'Exemple 7.1.7 et vu l'Exemple 7.1.8, nous pouvons en déduire les relations suivantes :

$$\begin{aligned} s_2(x_1, x_2, x_3) &= \frac{p_1^2(x_1, x_2, x_3) - p_2(x_1, x_2, x_3)}{2}, \\ p_2(x_1, x_2, x_3) &= s_1^2(x_1, x_2, x_3) - 2 \cdot s_2(x_1, x_2, x_3). \end{aligned}$$

Nous pouvons donc exprimer $s_2(x_1, x_2, x_3)$ en fonction de $p_1(x_1, x_2, x_3)$ et $p_2(x_1, x_2, x_3)$ ainsi qu'exprimer $p_2(x_1, x_2, x_3)$ en fonction de $s_1(x_1, x_2, x_3)$ et $s_2(x_1, x_2, x_3)$.

Remarque 7.1.10. Le nombre n de variables n'intervient pas directement dans les formules de Newton-Girard [3], mais intervient tout de même indirectement dans ces formules vu la Remarque 7.1.3.

3. En effet, vu la Définition 7.1.6, le fait de changer le nombre n ne modifie pas la forme de ces formules.

7.2 Égalité de mots sur un alphabet ordonné constitué de deux lettres

Vu la Définition 3.0.1, deux mots sont équivalents s'ils possèdent la même matrice de Parikh et vu la Remarque 2.3.8, la fonction qui à un mot associe sa matrice de Parikh n'est pas injective. Par conséquent, deux mots ayant la même matrice de Parikh ne sont pas forcément identiques ou en d'autres termes, deux mots équivalents ne sont pas forcément égaux. Ainsi, la notion d'égalité est plus forte que la notion d'équivalence.

Dans cette section, nous allons nous intéresser exclusivement à un alphabet ordonné constitué de deux lettres et donner des critères à satisfaire afin de déterminer un unique mot.

Lemme 7.2.1. *Soient $\Sigma = \{e_1, e_2\}$ un alphabet ordonné et w, w' des mots sur l'alphabet Σ. Si les mots w et w' possèdent le même vecteur de Parikh, noté*

$$\begin{pmatrix} r \\ t \end{pmatrix},$$

et si le nombre $|w|_{e_1e_2^i}$ est égal au nombre $|w'|_{e_1e_2^i}$ pour chaque entier i compris entre 1 et $\min(r,t)$, alors les mots w et w' sont égaux.

Démonstration. Pour chaque entier i compris entre 1 et r, nous notons x_i le nombre d'occurrences de la lettre e_2 se trouvant à droite de la i^{e} lettre e_1 dans w et x'_i le nombre d'occurrences de la lettre e_2 se trouvant à droite de la i^{e} lettre e_1 dans w'. Afin de prouver ce lemme, nous allons montrer que x_i est égal à x'_i pour chaque entier i compris entre 1 et r ce qui nous permettra de conclure la preuve. En effet, un mot sur l'alphabet Σ est entièrement déterminé par son vecteur de Parikh et par les nombres d'occurrences de la lettre e_2 se trouvant à droite de chacunes des lettres e_1 constituant ce mot. Dès lors, le mot w est égal à w' si x_i est égal à x'_i pour chaque entier i compris entre 1 et r.

Par hypothèse, les mots w et w' possèdent le même vecteur de Parikh. Ainsi, vu la Définition 2.2.1, nous obtenons :

$$|w|_{e_1} = |w'|_{e_1} = r \text{ et } |w|_{e_2} = |w'|_{e_2} = t.$$

Par hypothèse, nous savons également que le nombre $|w|_{e_1e_2^i}$ est égal au nombre $|w'|_{e_1e_2^i}$ pour chaque entier i compris entre 1 et $\min(r,t)$. Ainsi, nous pouvons en déduire ce qui suit :

$$|w|_{e_1e_2^i} = |w'|_{e_1e_2^i}, \ \forall i \in \{1, \ldots, r\}.$$

En effet, si r est inférieur à t, cette égalité est évidente étant donné que l'entier $\min(r,t)$ est égal à r. Si r est strictement supérieur à t, nous savons que l'entier $\min(r,t)$ est égal à t d'où l'égalité :

$$|w|_{e_1e_2^i} = |w'|_{e_1e_2^i}, \ \forall i \in \{1, \ldots, t\}$$

et nous avons également l'égalité suivante étant donné que le nombre d'occurences de la lettre e_2 (dans w comme dans w') vaut t :

$$|w|_{e_1 e_2^i} = |w'|_{e_1 e_2^i} = 0, \forall i \in \{t+1, \ldots, r\}.$$

Dans la suite de cette démonstration, nous notons α_j au lieu de $|w|_{e_1 e_2^j}$ pour chaque entier j compris entre 1 et r. Vu la définition du nombre α_j, ce nombre s'obtient en considérant chaque lettre e_1 dans w ainsi qu'en calculant le nombre de choix possibles de j lettres e_2 parmi les x_j lettres e_2 se trouvant à droite de la lettre e_1. Ainsi, pour chaque entier j compris entre 1 et r, nous obtenons la relation suivante :

$$\underbrace{\sum_{i=1}^{r} \binom{x_i}{j}}_{:=q_j(x_1,\cdots,x_r)} = \alpha_j$$

Comme le nombre $|w'|_{e_1 e_2^i}$ est égal au nombre $|w|_{e_1 e_2^i}$ pour chaque entier i compris entre 1 et r, nous pouvons en déduire l'égalité entre $q_j(x'_1, \ldots, x'_r)$ et α_j pour chaque entier j compris entre 1 et r. Par conséquent, nous pouvons écrire :

$$q_j(x_1, \ldots, x_r) = q_j(x'_1, \ldots, x'_r), \forall j \in \{1, \ldots, r\}.$$

Étant donné que nous disposons de l'égalité entre $q_j(x_1, \ldots, x_r)$ et α_j pour chaque entier j compris entre 1 et r, nous pouvons considérer le système suivant :

$$\begin{cases} q_1(x_1, \ldots, x_r) = \alpha_1 \\ \vdots \\ q_r(x_1, \ldots, x_r) = \alpha_r \end{cases}$$

Après diverses manipulations algébriques, le système ci-dessus peut se réécrire sous la forme suivante :

$$\begin{cases} p_1(x_1, \ldots, x_r) = P_1(\alpha_1, \ldots, \alpha_r) \\ \vdots \\ p_r(x_1, \ldots, x_r) = P_r(\alpha_1, \ldots, \alpha_r) \end{cases}$$

où $p_j(x_1, \cdots, x_r)$ $(1 \leq j \leq r)$ est la j^{e} somme de Newton décrite à la Définition 7.1.4 et où $P_j(x_1, \ldots, x_r)$ est un polynôme en $\alpha_1, \ldots, \alpha_r$. Vu que le passage entre les deux systèmes précédents peut sembler compliqué, nous allons l'illustrer sur un exemple afin de mieux le comprendre. Si nous considérons r qui vaut 3, nous obtenons ce qui suit [4] :

4. Vu la Remarque 1.1.1, nous pouvons facilement calculer q_1, q_2 et q_3. De plus, vu la Définition 7.1.4, nous pouvons trouver la valeurs des trois sommes de Newton.

$$
\begin{cases} q_1(x_1,\ldots,x_r)=\alpha_1\\ q_2(x_1,\ldots,x_r)=\alpha_2\\ q_3(x_1,\ldots,x_r)=\alpha_3 \end{cases}
\Leftrightarrow
\begin{cases} x_1+x_2+x_3=\alpha_1,\\ \dfrac{1}{2}(x_1^2+x_2^2+x_3^2-x_1-x_2-x_3)=\alpha_2,\\ \dfrac{1}{6}(x_1^3+x_2^3+x_3^3-3x_1^2-3x_2^2-3x_3^2+2x_1+2x_2+2x_3)=\alpha_3. \end{cases}
$$

$$
\Leftrightarrow
\begin{cases} x_1+x_2+x_3=\alpha_1,\\ x_1^2+x_2^2+x_3^2=2\alpha_2+\alpha_1,\\ x_1^3+x_2^3+x_3^3=6\alpha_3+6\alpha_2+\alpha_1. \end{cases}
$$

$$
\Leftrightarrow
\begin{cases} p_1(x_1,\ldots,x_r)=\alpha_1,\\ p_2(x_1,\ldots,x_r)=2\alpha_2+\alpha_1,\\ p_3(x_1,\ldots,x_r)=6\alpha_3+6\alpha_2+\alpha_1. \end{cases}
$$

Les polynômes $p_j(x_1,\ldots,x_r)$, $1\leq j\leq r$, sont obtenus à partir des polynômes $q_j(x_1,\ldots,x_r)$ étant donné que nous avons l'égalité entre $q_j(x_1,\ldots,x_r)$ et α_j ainsi que l'égalité entre $p_j(x_1,\cdots,x_r)$ et $P_j(\alpha_1,\ldots,\alpha_r)$ pour chaque entier j compris entre 1 et r. Les polynômes $p_j(x'_1,\ldots,x'_r)$ sont obenus quant à eux à partir des polynômes $q_j(x'_1,\ldots,x'_r)$ étant donné que nous avons l'égalité entre $q_j(x'_1,\ldots,x'_r)$ et α_j ainsi que l'égalité entre $p_j(x'_1,\cdots,x'_r)$ et $P_j(\alpha_1,\ldots,\alpha_r)$ pour chaque entier j compris entre 1 et r. De plus, vu l'égalité entre $q_j(x_1,\ldots,x_r)$ et $q_j(x'_1,\ldots,x'_r)$, nous pouvons en déduire :

$$p_j(x_1,\cdots,x_r)=p_j(x'_1,\cdots,x'_r),\ \forall j\in\{1,\ldots,r\}.$$

Afin de conclure cette démonstration, nous considérons le polynôme P de degré r ayant $x_1,\ldots,x_r$ comme racines et 1 comme coefficient dominant. Vu le théorème fondamental de l'algèbre [5], nous pouvons en déduire que le polynôme P se représente comme suit :

$$P(x)=x^r-s_1(x_1,\ldots,x_r)x^{r-1}-\cdots-s_{r-1}(x_1,\ldots,x_r)x-s_r(x_1,\ldots,x_r)$$

où $s_j(x_1,\ldots,x_r)$ $(1\leq j\leq r)$ est le j^{e} polynôme symétrique élémentaire. Nous considérons également le polynôme P' de degré r ayant $x'_1,\ldots,x'_r$ comme racines et 1 comme coefficient dominant. Vu le théorème fondamental de l'algèbre, ce polynôme peut s'écrire comme suit :

$$P'(x)=x^r-s_1(x'_1,\ldots,x'_r)x^{r-1}-\cdots-s_{r-1}(x'_1,\ldots,x'_r)x-s_r(x'_1,\ldots,x'_r)$$

où $s_j(x'_1,\ldots,x'_r)$ $(1\leq j\leq r)$ est le j^{e} polynôme symétrique élémentaire.

5. Théorème fondamental de l'algèbre : Tout polynôme P de degré $r\geq 1$ possède exactement r zéros si nous les comptons avec leur multiplicité. Ainsi, si $x_1,\ldots,x_k$ sont les k $(k\leq r)$ zéros de P de multiplicité respective $\alpha_1,\ldots,\alpha_k$, nous obtenons $\alpha_1+\ldots+\alpha_k=r$ et, si c_r est le coefficient dominant de P, nous constatons l'égalité entre $P(x)$ et $c_r\Pi_{i=1}^k(x-x_i)^{\alpha_i}$.

Les nombres α_j, $1 \leq j \leq r$, sont connus par hypothèse. D'une part, les polynômes $p_j(x_1, \ldots, x_r)$ $(= p_j(x'_1, \ldots, x'_r))$, $1 \leq j \leq r$, sont connus [6] vu le système précédent. D'autre part, les polynômes $s_j(x_1, \ldots, x_r)$ $(= s_j(x'_1, \ldots, x'_r))$ sont également connus vu la Remarque 7.1.9. Vu la forme des polynômes P et P', nous connaissons leurs coefficients, ce qui signifie que ces polynômes sont connus et identiques. Par conséquent, les polynômes P et P' possèdent les mêmes racines et nous pouvons même affirmer que :

$$x_i = x'_i, \forall i \in \{1, \ldots, r\}$$

étant donné que la définition des nombres x_i et x'_i nous permet d'en déduire les inégalités suivantes :

$$0 \leq x_r \leq \cdots \leq x_1 \leq t \text{ et } 0 \leq x'_r \leq \cdots \leq x'_1 \leq t$$

□

Lemme 7.2.2. *Le lemme 7.2.1 reste vrai si nous remplaçons le mot $e_1e_2^i$ par $e_1^ie_2, e_2e_1^i$ ou $e_2^ie_1$.*

Démonstration. Ce lemme affirme que nous pouvons remplacer le mot $e_1e_2^i$ dans le Lemme 7.2.1 par trois mots distincts. Ainsi, nous allons envisager trois cas afin de démontrer ce lemme.

Premièrement, nous remplaçons le mot $e_1e_2^i$ du Lemme 7.2.1 par le mot $e_2e_1^i$. Nous pouvons facilement conclure en procédant de manière analogue à la démonstration du Lemme 7.2.1.

Deuxièmement, nous remplaçons le mot $e_1e_2^i$ du Lemme 7.2.1 par le mot $e_1^ie_2$. En effectuant ce remplacement, nous avons par hypothèse l'égalité entre $|w|_{e_1^ie_2}$ et $|w'|_{e_1^ie_2}$ pour chaque entier i compris entre 1 et $\min(|w|_{e_1}, |w|_{e_2})$. De plus, vu la Définition 3.4.1, nous avons l'égalité entre $|w^R|_{e_2e_1^i}$ et $|w|_{e_1^ie_2}$ pour chaque entier i compris entre 1 et $\min(|w|_{e_1}, |w|_{e_2})$. Ainsi, nous pouvons en déduire l'égalité entre $|w^R|_{e_2e_1^i}$ et $|w'^R|_{e_2e_1^i}$ pour chaque entier i compris entre 1 et $\min(|w|_{e_1}, |w|_{e_2})$. Vu le cas précédent, les mots w^R et w'^R sont égaux et vu la Définition 3.4.1, les mots w et w' sont égaux.

Troisièmement, nous remplaçons le mot $e_1e_2^i$ du Lemme 7.2.1 par le mot $e_2^ie_1$. En effectuant ce remplacement, nous avons par hypothèse l'égalité entre $|w|_{e_2^ie_1}$ et $|w'|_{e_2^ie_1}$ pour chaque entier i compris entre 1 et $\min(|w|_{e_1}, |w|_{e_2})$. De plus, vu la Définition 3.4.1, nous avons l'égalité entre $|w^R|_{e_1e_2^i}$ et $|w|_{e_2^ie_1}$ pour chaque entier i compris entre 1 et $\min(|w|_{e_1}, |w|_{e_2})$. Ainsi, nous pouvons en déduire l'égalité entre $|w^R|_{e_1e_2^i}$ et $|w'^R|_{e_1e_2^i}$ pour chaque entier i compris entre 1 et $\min(|w|_{e_1}, |w|_{e_2})$. Vu le Lemme 7.2.1, les mots w^R et w'^R sont égaux et vu la Définition 3.4.1, les mots w et w' le sont aussi.

□

6. Pour chaque entier j compris entre 1 et r, le polynôme $p_j(x_1, \ldots, x_r)$ est égal au polynôme $P_j(\alpha_1, \ldots, \alpha_r)$.

7.3 Mot ambigu et projection

Dans cette section, nous allons énoncer des critères à satisfaire afin de déterminer un mot unique sur un alphabet ordonné quelconque. Avant d'énoncer ces critères, il est nécessaire d'introduire deux nouvelles notions (à savoir la notion de sous-alphabet et la notion de projection), le Théorème 7.3.7, le Lemme 7.3.9 ainsi que le Corollaire 7.3.10.

Définition 7.3.1. Soit $\Sigma = \{e_1, \ldots, e_k\}$ un alphabet ordonné. Un *sous-alphabet* de Σ est un alphabet ordonné obtenu en supprimant certaines lettres de Σ.

Exemple 7.3.2. Soit $\Sigma = \{a, b, c\}$ un alphabet ordonné. L'alphabet ordonné $\{a, c\}$ est un sous-alphabet de Σ étant donné que cet alphabet est obtenu en supprimant la lettre b de Σ.

Remarque 7.3.3. Soit $\Sigma = \{e_1, \ldots, e_k\}$ un alphabet ordonné. Si i et j sont des entiers tels que $1 \leq i \leq j \leq k$, l'alphabet ordonné $\{e_i, \ldots, e_j\}$ est un sous-alphabet de Σ.

Définition 7.3.4. Soient $\Sigma = \{e_1, \ldots, e_k\}$ un alphabet ordonné et Σ' un sous-alphabet de Σ. La *projection* de Σ^* sur Σ'^* est le morphisme (de monoïdes) qui envoit les lettres appartenant à Σ' sur elles-mêmes et qui envoit les lettres n'appartenant pas à Σ' sur le mot vide.

Exemple 7.3.5. Soient $\Sigma = \{a, b, c\}$ un alphabet ordonné, $\Sigma' = \{a, c\}$ un sous-alphabet [7] de Σ et $w = acbc$ un mot sur l'alphabet Σ. La projection du mot w sur Σ'^* est égale au mot [8] acc.

Remarque 7.3.6. Soient $\Sigma = \{e_1, \ldots, e_k\}$ un alphabet ordonné et $\Sigma' = \{e_i, \ldots, e_j\}$ $(1 \leq i \leq j \leq k)$ un sous-alphabet de Σ. La projection de Σ^* sur Σ'^* est notée $\pi_{i,j}$ et le morphisme qui à un mot associe sa matrice de Parikh induite par le mot $e_i \cdots e_j$ est noté $\psi_{i,j}$.

Une fois la Définition 7.3.1 et la Définition 7.3.4 données, nous pouvons énoncer et démontrer le théorème suivant.

Théorème 7.3.7. *Soient $\Sigma = \{e_1, \ldots, e_k\}$ un alphabet ordonné, u un mot sur l'alphabet Σ et p, q des entiers tels que $1 \leq p \leq q \leq k$. Pour chaque mot $w \in \Sigma^*$, nous pouvons écrire la relation suivante :*

$$\psi_{\pi_{p,q}(u)}(w) = \psi_{\pi_{p,q}(u)}(\pi_{p,q}(w)).$$

7. Voir l'Exemple 7.3.2.

8. Le mot vide est le neutre pour la concaténation.

Démonstration. Pour simplifier les notations, nous remplaçons $\pi_{p,q}(u)$ par u'. Vu la Remarque 7.3.6 et vu la Définition 2.0.5, nous pouvons en déduire l'égalité entre $|w|_{u'_{i,j}}$ et $|\pi_{p,q}(w)|_{u'_{i,j}}$ pour tous les entiers i et j tels que $1 \leq i \leq j \leq |u'|$. Ainsi, vu le Théorème 2.3.5 décrivant la forme des matrices de Parikh induite par un mot, nous pouvons conclure cette preuve.

□

Afin d'énoncer et de démontrer le Lemme 7.3.9 qui s'avère être un cas particulier du Théorème 7.3.7, nous introduisons la remarque suivante.

Remarque 7.3.8. Soient k un entier positif et $A = (a_{i,j})_{1\leq i,j\leq k+1}$ une matrice de Parikh étant de dimension $(k+1)$. Si p et q sont des entiers tels que $1 \leq p \leq q \leq k$, la matrice $A_{p,q}$ s'avère être la sous-matrice de A obtenue en préservant les éléments $a_{i,j}$, $p \leq i, j \leq q+1$, et en supprimant tous les autres éléments. Par exemple, nous avons :

$$A = \begin{pmatrix} 1 & 3 & 4 & 5 \\ 0 & 1 & 3 & 4 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix} \text{ et } A_{2,3} = \begin{pmatrix} 1 & 3 & 4 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}.$$

Lemme 7.3.9. *Soient $\Sigma = \{e_1, \ldots, e_k\}$ un alphabet ordonné et p, q des entiers tels que $1 \leq p \leq q \leq k$. Pour chaque mot $w \in \Sigma^*$, nous avons la relation suivante :*

$$[\psi_k(w)]_{p,q} = \psi_{p,q}(\pi_{p,q}(w)).$$

Démonstration. Vu la Remarque 7.3.6, nous avons l'égalité entre $\pi_{p,q}(e_1 \cdots e_k)$ et $e_p \cdots e_q$ ainsi que l'égalité entre $\psi_{e_p \cdots e_q}$ et $\psi_{p,q}$. De plus, vu le Corollaire 2.3.6 et vu la Remarque 7.3.8, nous avons l'égalité entre la matrice $\psi_{p,q}(w)$ et la matrice $[\psi_k(w)]_{p,q}$. Nous pouvons donc conclure grâce au Théorème 7.3.7.

□

À l'aide du Lemme 7.3.9, nous pouvons démontrer le corollaire suivant qui nous sera utile à la démonstration du Théorème 7.3.11.

Corollaire 7.3.10. *Soient $\Sigma = \{e_1, \ldots, e_k\}$ un alphabet ordonné et u, v des mots équivalents sur l'alphabet Σ. Si p et q sont des entiers tels que $1 \leq p \leq q \leq k$, nous avons les affirmations suivantes :*

- *Le mot $\pi_{p,q}(u)$ est équivalent au mot $\pi_{p,q}(v)$.*
- *Si le mot $\pi_{p,q}(v)$ n'est pas ambigu, nous avons l'égalité entre les mots $\pi_{p,q}(u)$ et $\pi_{p,q}(v)$.*

Démonstration. Par hypothèse, nous savons que les mots u et v sont équivalents ce qui signifie que ces mots possèdent la même matrice de Parikh. Vu le Lemme 7.3.9, nous pouvons en déduire ce qui suit :

$$\psi_{p,q}(\pi_{p,q}(v)) = \psi_{p,q}(\pi_{p,q}(u))$$

pour chaque entier p et pour chaque entier q tels que $1 \leq p \leq q \leq k$. Ainsi, vu le Théorème 2.3.6, la première affirmation est satisfaite.

Vu la première affirmation, les mots $\pi_{p,q}(u)$ et $\pi_{p,q}(v)$ sont équivalents pour chaque entier p et chaque entier q tels que $1 \leq p \leq q \leq k$. De plus, nous savons par hypothèse que le mot $\pi_{p,q}(v)$ n'est pas ambigu. Par conséquent, nous avons l'égalité entre les mots $\pi_{p,q}(u)$ et $\pi_{p,q}(v)$.

□

Le théorème suivant nous donne des critères à satisfaire afin de déterminer un unique mot.

Théorème 7.3.11. *Soient $\Sigma = \{e_1, \ldots, e_k\}$ un alphabet ordonné et u, w des mots équivalents sur l'alphabet Σ. Si tous les facteurs de longueur 2 des mots u et w possèdent une des formes suivantes :*

- *$e_i e_i$ où $1 \leq i \leq k$,*
- *$e_i e_{i+1}$ où $1 \leq i \leq k-1$,*
- *$e_{i+1} e_i$ où $1 \leq i \leq k-1$,*

s'il existe des entiers p et q tels que $1 < p \leq q < k$ et si les mots $\pi_{1,q}(w)$ et $\pi_{p,k}(w)$ ne sont pas ambigus, alors les mots u et w sont égaux

Démonstration. Nous procédons par l'absurde en supposant que les mots u et w diffèrent. Vu le Corollaire 7.3.10, nous avons l'égalité entre les mots $\pi_{1,q}(w)$ et $\pi_{1,q}(u)$ ainsi que l'égalité entre les mots $\pi_{p,k}(w)$ et $\pi_{p,k}(u)$.

Premièrement, nous allons supposer que le mot $\pi_{1,q}(w)$ (resp. $\pi_{p,k}(w)$) est égal au mot vide c'est-à-dire que le mot w ne contient pas les lettres $e_1, \ldots, e_q$ (resp. $e_p, \ldots, e_k$). Dans ce cas, nous avons l'égalité entre les mots w et $\pi_{p,k}(w)$ (resp. $\pi_{1,q}(w)$). Étant donné que $\pi_{p,k}(w)$ (resp. $\pi_{1,q}(w)$) est égal à $\pi_{p,k}(u)$ (resp. $\pi_{1,q}(u)$) et que u est équivalent à w, nous avons l'égalité entre u et w ce qui est absurde.

Deuxièmement, nous allons supposer que les mots $\pi_{1,q}(w)$ et $\pi_{p,k}(w)$ ne sont pas égaux au mot vide. Dans ce cas, nous faisons les hypothèses suivantes :

1) n la longueur des mots [9] u et w,
2) $w = b_1 \cdots b_n$ avec $b_1, \ldots, b_n \in \Sigma$,
3) $u = c_1 \cdots c_n$ avec $c_1, \ldots, c_n \in \Sigma$,
4) j le plus petit entier [10] tel que $1 \leq j \leq n$ et $b_j \neq c_j$,
5) m la longueur du mot $\pi_{1,q}(w)$ $(= \pi_{1,q}(u))$,
6) $\pi_{1,q}(w) = d_1 \cdots d_m$ $(= \pi_{1,q}(u))$ avec $d_1, \ldots, d_m \in \Sigma$,
7) r la longueur du mot $\pi_{p,k}(w)$ $(= \pi_{p,k}(u))$,
8) $\pi_{p,k}(w) = f_1 \cdots f_r$ $(= \pi_{p,k}(u))$ avec $f_1, \ldots, f_r \in \Sigma$,
9) l l'entier tel que $1 \leq l \leq j$ et [11] $\pi_{1,q}(b_1 \cdots b_{j-1}) = d_1 \cdots d_{l-1}$ $(= \pi_{1,q}(c_1 \cdots c_{j-1}))$,
10) s l'entier tel que $1 \leq s \leq j$ et $\pi_{p,k}(b_1 \cdots b_{j-1}) = f_1 \cdots f_{s-1}$ $(= \pi_{p,k}(c_1 \cdots c_{j-1}))$.

Afin de démontrer le corollaire lorsque les mots $\pi_{1,q}(w)$ et $\pi_{p,k}(w)$ ne sont pas égaux au mot vide, nous allons envisager quatre cas.

Tout d'abord, nous supposons que l'entier j est strictement supérieur à 1 et que la lettre b_j appartient à l'ensemble $\{e_1, \ldots, e_{p-1}\}$ (resp. $\{e_{q+1}, \ldots, e_k\}$). Vu l'hypothèse 6 (resp. 8), vu l'hypothèse 9 (resp. 10) et vu le fait que p est inférieur à q, nous avons l'égalité entre b_j et d_l (resp. f_s). Vu la forme des facteurs de longueur 2 du mot w, la lettre b_{j-1} appartient à l'ensemble $\{e_1, \ldots, e_p\}$ (resp. $\{e_q, \ldots, e_k\}$) et vu l'hypothèse 4, nous avons l'égalité entre les lettres b_{j-1} et c_{j-1}. Ainsi, la lettre c_j appartient à l'ensemble $\{e_1, \ldots, e_{p+1}\}$ (resp. $\{e_{q-1}, \ldots, e_k\}$) vu la forme des facteurs de longueur 2 du mot u. Les trois affirmations suivantes nous permettent de conclure ce cas :

- Si p est strictement inférieur à q, la lettre c_j appartient à l'ensemble $\{e_1, \ldots, e_q\}$ (resp. $\{e_p, \ldots, e_k\}$). Vu l'hypothèse 6 (resp. 8) et vu l'hypothèse 9 (resp. 10), les lettres c_j et d_l (resp. f_s) sont égales. Par conséquent, les lettres c_j et b_j sont égales, ce qui est absurde vu la définition de j.

- Si p est égal à q et si la lettre c_j diffère de la lettre e_{p+1} (resp. e_{p-1}), la lettre c_j appartient à l'ensemble $\{e_1, \ldots, e_p\}$ (resp. $\{e_p, \ldots, e_k\}$). Par conséquent, nous avons l'égalité entre la lettre c_j et la lettre d_l (resp. f_s) ainsi que l'égalité entre la lettre b_j et la lettre c_j, ce qui est absurde vu la définition de j.

9. Comme les mots u et w sont équivalents, ils ont la même longueur.
10. Cet entier existe vu que les mots u et w diffèrent par hypothèse.
11. Vu la définition de l'entier j, les mots $b_1 \cdots b_{j-1}$ et $c_1 \cdots c_{j-1}$ sont égaux.

- Si p est égal à q et si la lettre c_j est égale à la lettre e_{p+1} (resp. e_{p-1}), nous avons l'égalité[12] entre les lettres c_{j-1}, b_{j-1} et e_p. De même, nous notons l'égalité entre la première lettre c_i ($i \geq j$) du mot u appartenant à l'ensemble $\{e_1, \ldots e_p\}$ (resp. $\{e_p, \ldots, e_k\}$) et la lettre e_p, vu la forme des facteurs de longueur 2 du mot u. Ainsi, vu l'hypothèse 6 (resp. 8) et vu l'hypothèse 9 (resp. 10), la lettre d_l (resp. f_s) est égale à la lettre e_p, ce qui est absurde étant donné que b_j est égal à d_l et que b_j appartient à l'ensemble $\{e_1, \ldots, e_{p-1}\}$ (resp. $\{e_{p+1}, \ldots, e_k\}$).

Ensuite, nous supposons que l'entier j est strictement supérieur à 1 et que la lettre b_j appartient à l'ensemble $\{e_p, \ldots, e_q\}$. Si la lettre c_j appartient à l'ensemble $\{e_1, \ldots, e_q\}$, les lettres $\pi_{1,q}(b_j)$, $\pi_{1,q}(c_j)$ et d_l sont égales, ce qui est absurde vu la définition de j. Si la lettre c_j appartient à l'ensemble $\{e_{q+1}, \ldots, e_k\}$, les lettres $\pi_{p,k}(b_j)$, $\pi_{p,k}(c_j)$ et f_s sont égales, ce qui est absurde vu la définition de j.

De plus, nous supposons que l'entier j est égal à 1 et que la lettre b_1 appartient à l'ensemble $\{e_1, \ldots, e_{p-1}\}$ (resp. $\{e_{q+1}, \ldots, e_k\}$). Dans ce cas, la première lettre b_i ($i \geq 1$) du mot w n'appartenant pas à l'ensemble $\{e_1, \ldots, e_{p-1}\}$ (resp. $\{e_{q+1}, \ldots, e_k\}$) s'avère être la lettre e_p (resp. e_q) vu la forme des facteurs de longueur 2 du mot w. Par conséquent, le mot $\pi_{p,k}(w)$ (resp. $\pi_{1,q}(w)$) commence par la lettre e_p (resp. e_q). Vu l'égalité entre les mots $\pi_{p,k}(w)$ (resp. $\pi_{1,q}(w)$) et $\pi_{p,k}(u)$ (resp. $\pi_{1,q}(u)$), le mot u ne peut pas commencer par une lettre de l'ensemble[13] $\{e_{q+1}, \ldots, e_k\}$ (resp. $\{e_1, \ldots, e_{p-1}\}$). Ainsi, la lettre c_1 appartient à l'ensemble $\{e_1, \ldots, e_q\}$ (resp. $\{e_p, \ldots, e_k\}$) tout comme la lettre b_1. Vu l'hypothèse 6 (resp. 8) et vu l'hypothèse 9 (resp. 10), nous avons l'égalité entre les lettres b_1, c_1 et d_1 (resp. f_1) ce qui est absurde vu la définition de j.

Enfin, nous supposons que l'entier j est égal à 1 et que la lettre b_1 appartient à l'ensemble $\{e_p, \ldots, e_q\}$. Dans ce cas, les mots $\pi_{1,q}(w)$ et $\pi_{p,k}(w)$ commencent tous les deux par la lettre b_1. Étant donné que les mots $\pi_{1,q}(w)$ (resp. $\pi_{p,k}(w)$) et $\pi_{1,q}(u)$ (resp. $\pi_{p,k}(u)$) sont égaux, les mots $\pi_{1,q}(u)$ et $\pi_{p,k}(u)$ commencent tous les deux par b_1. Ainsi, la lettre b_1 est égale à la lettre c_1, ce qui est absurde vu la définition de j.

□

Le théorème suivant nous donne également des critères à satisfaire pour déterminer un mot unique. Contrairement au Théorème 7.3.11, les hypothèses de ce théorème ne font pas intervenir de conditions sur les facteurs de longueur 2 d'un mot. Cependant, si les hypothèses de ce théorème sont satisfaites, les hypothèses du Théorème 7.3.11 sont également satisfaites comme nous le verrons dans la démonstration de ce théorème.

12. Les lettres b_{j-1} et c_{j-1} appartiennent à l'ensemble $\{e_1, \ldots, e_p\}$ (resp. $\{e_p, \ldots, e_k\}$) étant donné que p est égal à q. Vu la forme des facteurs de longueur 2 des mots u et w, les lettres b_{j-1} et c_{j-1} sont égales à la lettre e_p.

13. Comme le mot $\pi_{p,k}(u)$ commence par la lettre e_p (resp. e_q), le mot u ne peut pas commencer par une lettre de l'ensemble $\{e_{p+1}, \ldots, e_k\}$ (resp. $\{e_1, \ldots, e_{q-1}\}$). De plus, puisque p est inférieur à q, le mot u ne peut pas commencer par une lettre de l'ensemble $\{e_{q+1}, \ldots, e_k\}$ (resp. $\{e_1, \ldots, e_{p-1}\}$).

Théorème 7.3.12. *Soient $\Sigma = \{e_1, \ldots, e_k\}$ un alphabet ordonné, w un mot sur l'alphabet Σ, u un mot sur l'alphabet Σ équivalent à w et p, q des entiers tels que $1 \leq p < q \leq k$. Si les mots $\pi_{1,q}(w)$ et $\pi_{p,k}(w)$ ne sont pas ambigus et si le mot $\pi_{p,q}(w)$ diffère du mot vide, nous avons l'égalité entre u et w.*

Démonstration. Pour commencer, nous allons montrer que tous les facteurs de longueur 2 du mot w sont de la forme e_ie_j avec $1 \leq i \leq k$, $1 \leq j \leq k$ et $|i - j| \leq 1$. Pour ce faire, nous procédons par l'absurde en supposant que le mot w possède au moins un facteur e_ie_j tel que $1 \leq i \leq k$, $1 \leq j \leq k$ et $|i - j| > 1$. Etant donné les mots $\pi_{1,q}(w)$ et $\pi_{p,k}(w)$ ne sont pas ambigus, les entiers i et j sont tels que $i < p$ et $j > q$ ou tels que $i > q$ et $j < p$. Effectivement, si les entiers i et j ne satisfont pas une de ces deux conditions, nous arrivons à une absurdité :

- Si les entiers i et j sont supérieurs à p, le mot e_ie_j est un facteur de $\pi_{p,k}(w)$, ce qui est absurde[14].
- Si les entiers i et j sont inférieurs à q, le mot e_ie_j est un facteur de $\pi_{1,q}(w)$, ce qui est absurde[14].
- Si l'entier i est tel que $p \leq i \leq q$ et si l'entier j est strictement supérieur à i, le mot e_ie_j est un facteur de $\pi_{p,k}(w)$ ce qui est absurde[14]. Si l'entier i est tel que $p \leq i \leq q$ et si l'entier j est strictement inférieur à i, le mot e_ie_j est un facteur de $\pi_{1,q}(w)$, ce qui est absurde[14].
- Si l'entier j est tel que $p \leq i \leq q$ et si l'entier i est strictement supérieur à j, le mot e_ie_j est un facteur de $\pi_{p,k}(w)$ ce qui est absurde[14]. Si l'entier j est tel que $p \leq i \leq q$ et si l'entier i est strictement inférieur à j, le mot e_ie_j est un facteur de $\pi_{1,q}(w)$, ce qui est absurde[14].

Sans perte de généralité, nous pouvons supposer que les entiers i et j sont tels que $i < p$ et $j > q$. Étant donné que le mot $\pi_{p,q}(w)$ diffère du mot vide, le mot w est composé d'au moins une lettre e_r avec r un entier compris entre p et q. Si cette lettre apparaît à droite (resp. gauche) de e_ie_j, le mot w peut s'écrire $xe_ie_jye_rz$ (resp. $xe_rye_ie_jz$) avec x, y, z des mots sur l'alphabet Σ. Ainsi, le mot $\pi_{1,q}(w)$ s'écrit $\pi_{1,q}(x)e_i\pi_{1,q}(y)e_r\pi_{1,q}(z)$ (resp. $\pi_{1,q}(x)e_r\pi_{1,q}(y)e_i\pi_{1,q}(z)$). Comme le mot $\pi_{1,q}(w)$ n'est pas ambigu, nous pouvons en déduire, vu le Corollaire 5.2.5, que la lettre e_p apparaît dans le mot $\pi_{1,q}(y)e_r$ ainsi que dans le mot ye_r [15] et que cette lettre est la première lettre d'indice strictement supérieur à $p - 1$ qui apparaît après (resp. avant) le facteur e_ie_j dans le mot w. Par conséquent, le mot e_je_p

14. Les mots $\pi_{1,q}(w)$ et $\pi_{p,k}(w)$ ne sont pas ambigus. Ainsi, vu le Corollaire 5.2.5, les facteurs de longueur 2 de ces mots devraient être de la forme e_me_n avec $1 \leq m \leq k$, $1 \leq n \leq k$ et $|m - n| \leq 1$. Or, le mot e_ie_j est un facteur de w ne satisfaisant pas cette condition.

15. La lettre e_p est égal à la lettre e_r ou apparaît dans le mot y

(resp. e_pe_j) est un facteur de $\pi_{p,k}(w)$ tel [16] que $|j-p| > 1$ ce qui est absurde étant donné que $\pi_{p,k}$ n'est pas ambigu [17]. Tous les facteurs de w sont donc de la forme e_ie_j avec $1 \leq i \leq k$, $1 \leq j \leq k$ et $|i-j| \leq 1$.

Ensuite, nous allons montrer que les facteurs de longueur 2 du mot u sont de la forme e_ie_j avec $1 \leq i \leq k$, $1 \leq j \leq k$ et $|i-j| \leq 1$. Vu le Corollaire 7.3.10, nous avons l'égalité entre les mots $\pi_{1,q}(w)$ et $\pi_{1,q}(u)$ ainsi que l'égalité entre les mots $\pi_{p,k}(w)$ et $\pi_{p,k}(u)$. Par conséquent, nous pouvons procéder de manière analogue à ce qui précède et en déduire que les facteurs de longueur 2 du mot u ont la forme attendue.

Enfin, les hypothèses du Théorème 7.3.11 sont satisfaites vu les hypothèses de ce théorème et la forme des facteurs de longueur 2 des mots u et w. Par conséquent, nous pouvons appliquer le Théorème 7.3.11 et conclure la preuve.

$\square$

16. Nous savons que q est strictement supérieur à p par hypothèse et nous avons considéré le cas où j est strictement supérieur à q. Ainsi, nous pouvons en déduire que $|j-p| = j-p > q-p > 1$.

17. Voir le corollaire 5.2.5.

Bibliographie

[1] A. Atanasiu, *Binary amiable words*, Int. J. Found. Comput. Sci. **18** (2007), 387-400.

[2] A. Atanasiu, R. Atanasiu, I. Petre, *Parikh matrices and amiable words*, Theoret. Comput. Sci. **390** (2008), 102-109.

[3] A. Mateescu, A. Salomaa, *Matrix indicators for subword occurrences and ambiguity*, Int. J. Found. Comput. Sci. **15** (2004), 277-292.

[4] M. Rigo, *Algèbre linéaire*, Notes de cours, Université de Liège, (2009-2010).

[5] M. Rigo, *Algorithmique et calculabilité*, Notes de cours, Université de Liège, (2009-2010).

[6] M.Rigo, *Théorie des automates et langages formels*, Notes de cours, Université de Liège, (2009-2010).

[7] A. Salomaa, *Counting (scattered) subwords*, EATCS Bull. **81** (2003), 165-179.

[8] A. Salomaa, *Connections between subwords and certain matrix mappings*, Theoret. Comput. Sci. **340** (2005), 188-203.

[9] A. Salomaa, *On the injectivity of Parikh matrix mappings*, Fund. Inform. **64** (2005), 391-404.

[10] A. Salomaa, *Subword histories and associated matrices*, Theoret. Comput. Sci. **407** (2008), 250-257.

[11] A. Salomaa, *Criteria for the matrix equivalence of words*, Theorical Computer Science, **411** (2010), 1818-1827.

[12] A. Salomaa, S. Yu, *Subword occurrences, Parikh matrices and Lyndon images*, Int. J. Found. Comp. Sci. **21** (2010), 91-111.

[13] V.G. Şerbănuţă, T.F. Şerbănuţă, *Injectivity of the Parikh matrix mappings revisited*, Fund. Inform. **73** (2006), 265-283.

[14] V.G. Şerbănuţă, *On Parikh matrices, ambiguity and prints*, Int. J. Found. Comp. Sci. **20** (2009), 151-165.