

Dynamique symbolique sur un monoïde finiment présenté

1.1	Monoïdes finiment présentés	10
1.1.1	Définition	10
1.1.2	Graphe de Cayley	11
1.1.3	Cas des groupes finiment présentés	12
1.2	Les décalages, deux points de vue complémentaires	13
1.2.1	Point de vue topologique	13
	L'espace des configurations	13
	Définition topologique des décalages	14
1.2.2	Définition par motifs interdits	15
	Motifs et langages	15
	Motifs interdits et langage d'un décalage sur M	16
1.2.3	Coïncidence des définitions	19
1.3	Classes de décalages	20
1.3.1	Décalages de type fini	20
1.3.2	Décalages sofiques	21
1.3.3	Décalages effectifs	23

Dans ce premier chapitre nous présentons les décalages dans le cas qui correspond au cadre le plus général de cette thèse, celui des des décalages définis sur un monoïde finiment présenté. La Partie 1.1 définit la notion de monoïde finiment présenté. Sur ces monoïdes il est possible de définir les décalages de deux façons équivalentes, qui correspondent à une approche topologique (Partie 1.2.1) ou combinatoire (Partie 1.2.2). La Partie 1.3 présente les trois classes de décalages dont il sera question dans la suite du manuscrit.

Ce choix de définir les décalages dans ce cadre très général permet de proposer un formalisme qui unifie les notations pour les décalages sur \mathbb{Z}^d (Chapitres 2 et 3) et des décalages d'arbre (Chapitre 4).

Dans ce chapitre nous utiliserons des notions élémentaires de théorie des groupes, et nous supposerons que le lecteur est familier de ces objets ; à défaut,

nous l’invitons à consulter [Rot94]. Nous supposerons de même pour la théorie des graphes pour laquelle nous conseillons [Die05].

1.1 Monoïdes finiment présentés

Cette partie est consacrée aux monoïdes finiment présentés, objets sur lesquels seront définis les décalages dans toute la suite. Notre étude se restreint à ces structures car elles constituent à nos yeux le juste milieu entre simplicité de la définition, puisque la description d’un tel monoïde est finie, et complexité de l’objet, puisque le problème du mot peut y être indécidable [Col86]. Un outil important pour l’étude de ces monoïdes est le graphe de Cayley (Partie 1.1.2), qui permet d’une part de visualiser leurs structures et d’autre part d’en caractériser certaines propriétés.

1.1.1 Définition

Un *monoïde* (\mathbb{M}, \circ) , ou plus simplement \mathbb{M} , est constitué d’un ensemble d’éléments \mathbb{M} stable par une loi de composition associative \circ , et qui possède un élément neutre $\varepsilon_{\mathbb{M}}$ pour cette loi. Nous prenons ici la convention de noter cette loi multiplicativement, en omettant la plupart du temps le symbole \circ .

Remarque. Dans toute la suite du manuscrit, nous ne considérons que des monoïdes dont le cardinal est infini.

Un monoïde \mathbb{M} est *libre* s’il existe un sous-ensemble \mathcal{G} de \mathbb{M} tel que tout élément de \mathbb{M} s’écrit d’une et une seule manière comme un produit fini d’éléments de \mathcal{G} . On dit alors que \mathbb{M} est *libre sur* \mathcal{G} . Tous les monoïdes libres sur un même ensemble \mathcal{G} sont isomorphes, c’est pourquoi on s’autorise à parler *du monoïde libre sur* \mathcal{G} . Dans le cas où \mathcal{G} est un ensemble fini à n éléments, on parle *du monoïde libre à n générateurs*, et on le note \mathbb{M}_n .

Soit \mathbb{M} le monoïde libre sur \mathcal{G} , et \mathcal{R} un sous-ensemble de $\mathbb{M} \times \mathbb{M}$, appelé ensemble de relations. On définit sur \mathbb{M} une relation symétrique $\mathcal{E} \subseteq \mathbb{M} \times \mathbb{M}$ de la façon suivant : on dit que $(g, h) \in \mathcal{E}$ si et seulement si $g = sut$ et $h = svt$ avec $v, s, t \in \mathbb{M}$ et $(u, v) \in \mathcal{R} \cup \{(m', m), (m, m') \in \mathcal{R}\}$. La clôture réflexive et transitive de la relation \mathcal{E} est alors une relation d’équivalence $\sim_{\mathcal{R}}$. En quotientant le monoïde libre \mathbb{M} par cette relation d’équivalence, on obtient le *monoïde de présentation* $\langle \mathcal{G} | \mathcal{R} \rangle$. On dit qu’un monoïde est *finiment présenté* s’il possède une présentation $\langle \mathcal{G} | \mathcal{R} \rangle$ pour laquelle \mathcal{G} et \mathcal{R} sont des ensembles finis. Le cardinal de l’ensemble \mathcal{G} est appelé le *rang du monoïde* \mathbb{M} .

Exemple 1.1.1. Le monoïde des mots finis sur l'alphabet $\{0, 1\}$ est un monoïde finiment présenté de rang 2 dont une présentation est $\langle 0, 1 | \emptyset \rangle$. L'opération dans ce monoïde est la concaténation, et l'élément neutre pour cette loi est le mot vide ε . Il est isomorphe au monoïde libre à deux générateurs \mathbb{M}_2 .

Exemple 1.1.2. Une présentation du groupe diédral d'ordre 4, qui est le groupe des isométries du carré, est :

$$\langle a, b | a^4 = b^2 = \varepsilon, ba = a^3b \rangle$$

où l'élément b peut être interprété comme une réflexion, et l'élément a une rotation d'angle $\frac{\pi}{4}$.

1.1.2 Graphe de Cayley

Le graphe de Cayley d'un groupe donne une représentation visuelle de la structure du groupe par un graphe. C'est aussi un moyen pour construire des graphes avec de fortes propriétés de régularité. On peut également étendre cette notion pour définir le graphe de Cayley d'un monoïde.

Définition 1. Le *graphe de Cayley du monoïde* $\mathbb{M} = \langle \mathcal{G} | \mathcal{R} \rangle$, noté $\Gamma_{\mathbb{M}} = (V_{\Gamma}, E_{\Gamma})$ est un graphe orienté construit de la manière suivante : l'ensemble des sommets est $V_{\Gamma} = \mathbb{M}$, et les arêtes sont

$$E_{\Gamma} = \{(g, g.g_i) \in \mathbb{M} \times \mathbb{M} \mid g_i \text{ est un générateur de } \mathbb{M}\}.$$

Notons que le graphe de Cayley du monoïde \mathbb{M} dépend donc de la représentation qu'on a choisie.

Deux exemples en sont donnés sur la Figure 1 ; pour plus de clarté on a attribué une couleur à chaque générateur du monoïde. Dans la suite nous identifierons un élément du monoïde \mathbb{M} avec le nœud qui le représente dans le graphe de Cayley $\Gamma_{\mathbb{M}}$. Lorsqu'aucune confusion n'est possible, le graphe de Cayley de \mathbb{M} sera simplement appelé Γ .

La *longueur* d'un élément g du monoïde \mathbb{M} est la longueur du plus court chemin entre les sommets correspondant à g et à l'élément neutre ε dans le graphe de Cayley Γ ; on la note $|g|$. Certaines propriétés du monoïde peuvent être lues directement sur le graphe de Cayley. De manière informelle, les cycles dans le graphe correspondent aux relations \mathcal{R} entre générateurs. En particulier, un graphe de Cayley est celui d'un monoïde libre si et seulement si le graphe est acyclique.

$$\mathbb{M}_2 = \langle 0, 1/\emptyset \rangle$$

$$\mathcal{D}_4 = \langle a, b/a^4 = b^2 = \varepsilon, b.a = a^3.b \rangle$$

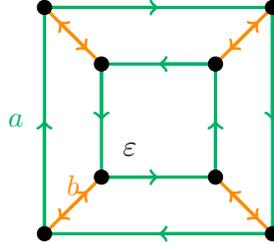
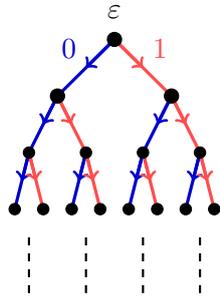


FIGURE 1: Les graphes de Cayley des monoïdes finiment présentés des Exemples 1.1.1 et 1.1.2.

1.1.3 Cas des groupes finiment présentés

Un groupe est un ensemble d'éléments muni d'une loi de composition interne associative, qui possède un élément neutre pour cette loi et tel que chaque élément possède un inverse. Cette structure gagne donc en rigidité par rapport à celle d'un monoïde car on impose la présence des inverses.

Remarque. Si \mathcal{G} est une partie d'un groupe, on notera par \mathcal{G}^{-1} l'ensemble des inverses des éléments de \mathcal{G} .

Nous avons choisi dans ce chapitre de d'abord traiter le cas des monoïdes finiment présentés. Toutes les notions définies précédemment sont toutefois compatibles avec les groupes finiment présentés, au sens où l'on peut voir un groupe finiment présenté $\mathbb{G} = \langle \mathcal{G} | \mathcal{R} \rangle$ comme un cas particulier de monoïde : si \mathcal{G} est un ensemble de générateurs pour le groupe, c'est-à-dire que tout élément du groupe s'écrit comme un produit fini d'éléments de \mathcal{G} et d'inverses d'éléments de \mathcal{G} , et que le groupe est défini par l'ensemble de relations \mathcal{R} sur des éléments de $\mathcal{G} \cup \mathcal{G}^{-1}$, alors le groupe \mathbb{G} peut être vu comme le monoïde finiment présenté suivant :

$$\mathbb{G} = \langle \mathcal{G} \cup \mathcal{G}^{-1} | \mathcal{R} \cup \{gg^{-1} = \varepsilon, g \in \mathcal{G}\} \rangle .$$

Le graphe de Cayley du groupe nous donne également un moyen très simple de définir une distance entre les éléments du groupe. Étant donnés deux éléments g et h du groupe \mathbb{G} , la *distance entre g et h* , notée $d_{\mathbb{G}}(g, h)$, est la longueur du plus court chemin entre les sommets correspondant à g et h dans le graphe de

Cayley $\Gamma(\mathbb{G})$ de \mathbb{G} . Cette distance est équivalente à la suivante :

$$d'(g, h) = \min \{ |j_1| + \dots + |j_k|, gh^{-1} = g_{i_1}^{j_1} \dots g_{i_k}^{j_k} \}.$$

De même que pour un monoïde, la *longueur* d'un élément g du groupe \mathbb{G} est la distance entre g et l'élément neutre ε ; on la note $|g| = d_{\mathbb{G}}(g, \varepsilon)$. Un résultat classique est que le groupe \mathbb{G} muni de la distance $d_{\mathbb{G}}$ est un groupe topologique. Notons qu'une fois encore la distance ainsi définie sur le groupe dépend de la présentation choisie, mais pour un groupe finiment présenté toutes les présentations donnent des distances équivalentes.

On notera \mathbb{F}_n le *groupe libre à n générateurs*. Par convention, lorsque l'on travaillera sur un groupe on omettra les inverses dans l'ensemble de générateurs.

1.2 Les décalages, deux points de vue complémentaires

Cette partie présente deux manières équivalentes de définir un décalage sur un monoïde finiment présenté. Ces deux définitions correspondent chacune à un certain point de vue, et sont donc plus ou moins adaptées selon le contexte.

1.2.1 Point de vue topologique

Dans cette partie, un décalage est vu comme un système dynamique, où l'on fait agir le monoïde \mathbb{M} sur un certain sous-ensemble de l'espace des configurations.

L'espace des configurations

L'espace des configurations $A^{\mathbb{M}}$ est défini comme un espace produit et à ce titre on peut le munir d'une topologie produit. Avec la topologie discrète (équivalente à celle donnée par la distance d définie par $d(x, y) = 0$ si $x = y$ et $d(x, y) = 1$ si $x \neq y$) sur l'alphabet A , on munit l'espace des configurations $A^{\mathbb{M}}$ d'une topologie produit (aussi appelée topologie pro-discrète) équivalente à celle donnée par la distance suivante :

$$\forall x, y \in A^{\mathbb{M}}, \begin{cases} d(x, y) = 2^{-\min\{|g|, x_g \neq y_g\}} & \text{si } x \neq y \\ d(x, y) = 0 & \text{si } x = y \end{cases}$$

Deux configurations sont d'autant plus proches qu'elles coïncident sur un motif central plus grand, et le plus court élément du monoïde pour lequel elles diffèrent donne la distance entre ces configurations.

Un résultat classique est que l'ensemble $A^{\mathbb{M}}$ des configurations, muni de cette topologie produit, est compact (il s'agit d'une application directe du théorème de Tychonoff).

Définition topologique des décalages

On définit la *transformation* $\sigma_{\mathbb{M}}$ comme l'action naturelle de \mathbb{M} sur l'espace des configurations, où un élément du monoïde $g \in \mathbb{M}$ agit sur $A^{\mathbb{M}}$ par translation (voir la Figure 2) via l'application $(\sigma_{\mathbb{M}})^g : A^{\mathbb{M}} \rightarrow A^{\mathbb{M}}$ (que l'on notera simplement σ^g lorsqu'aucune confusion n'est possible sur le monoïde \mathbb{M}) définie par :

$$(\sigma_{\mathbb{M}})^g(x)_h = x_{g,h} \text{ pour tout } i \in A^{\mathbb{M}}.$$

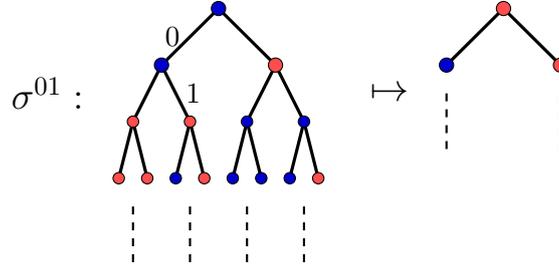


FIGURE 2: Action naturelle de $\mathbb{M}_2 = \{0, 1\}^*$ par translation sur l'espace des configurations $\{\bullet, \circ\}^{\mathbb{M}_2}$.

Si $x \in A^{\mathbb{M}}$ est une configuration et $\mathbb{S} \subseteq \mathbb{M}$ un ensemble de coordonnées, la restriction de x à \mathbb{S} notée $x_{\mathbb{S}}$ est un élément de $A^{\mathbb{S}}$ tel que :

$$\forall g \in \mathbb{S}, (x_{\mathbb{S}})_g = x_g.$$

Si \mathbb{S} est un sous-ensemble fini de $A^{\mathbb{M}}$ et $p \in A^{\mathbb{S}}$, on appelle *cylindre relatif* au motif p l'ensemble des configurations dans lesquelles le motif p apparaît en position ε :

$$[[p]] = \{x \in A^{\mathbb{M}} \mid x_{\mathbb{S}} = p\}.$$

Les cylindres sont des ouverts-fermés.

Définition 2. Un *décalage* est un sous-ensemble stable par la translation $\sigma_{\mathbb{M}}$ de $A^{\mathbb{M}}$, c'est-à-dire que $\sigma^g(\mathbf{T}) \subseteq \mathbf{T}$ pour tout $g \in \mathbb{M}$, et fermé pour la topologie produit.

On notera par $\mathcal{E}_A^{\mathbb{M}}$ l'ensemble de tous les décalages définis sur le monoïde \mathbb{M} et sur l'alphabet A . On notera aussi

$$\mathcal{E}^{\mathbb{M}} = \bigcup_{A \text{ alphabet fini}} \mathcal{E}_A^{\mathbb{M}}$$

l'ensemble des décalages sur le monoïde \mathbb{M} (on pourra omettre l'exposant \mathbb{M} lorsqu'aucune confusion n'est possible).

Remarque. Dans le cas d'un décalage \mathbf{T} sur un groupe \mathbb{G} , la stabilité par translation s'exprime de manière équivalente par $\sigma^g(\mathbf{T}) = \mathbf{T}$.

En particulier l'espace des configurations $A^{\mathbb{M}}$ est un décalage, que l'on appelle aussi *décalage plein*, ou *décalage trivial*.

1.2.2 Définition par motifs interdits

Dans cette partie nous optons pour le point de vue combinatoire pour définir les décalages. Toutes les définitions et propriétés exposées dans cette partie sont des généralisations de celles présentées dans l'ouvrage de référence de Lind et Marcus [LM95].

Dans toute cette partie nous considérons un monoïde finiment présenté $\mathbb{M} = \langle \mathcal{G} | \mathcal{R} \rangle$ de rang d , où $\mathcal{G} = \{g_1, \dots, g_d\}$ est un ensemble de d générateurs, ainsi qu'un alphabet fini A dont chaque élément est appelé *lettre* ou encore *couleur*.

Motifs et langages

On considère les coloriages de \mathbb{M} par A et on appelle *configuration* un élément de $A^{\mathbb{M}}$. Une configuration $x \in A^{\mathbb{M}}$ est donc une application $x : \mathbb{M} \rightarrow A$; on adopte dans la suite la notation x_g pour désigner l'image $x(g)$ de $g \in \mathbb{M}$ par l'application x . Un *support* est un ensemble fini $\mathbb{S} \subseteq \mathbb{M}$ qui contient le mot vide du monoïde ε . Étant donné un support \mathbb{S} , un *motif de support* \mathbb{S} est un élément de $A^{\mathbb{S}}$. De manière analogue, on désigne par p_g l'image de $g \in \mathbb{S}$ par l'application $p : \mathbb{S} \rightarrow A$.

On dit qu'un motif $p \in A^{\mathbb{S}}$ *apparaît* dans une configuration $x \in A^{\mathbb{M}}$ s'il existe un élément $g \in \mathbb{M}$ tel que pour tout $i \in \mathbb{S}$, $x_{g.i} = p_i$; dans ce cas on dit que le motif p apparaît dans x en position g , et on le note $p \sqsubset_g x$, ou simplement $p \sqsubset x$. On dit qu'un motif p de support \mathbb{S} apparaît dans un motif p' de support \mathbb{S}' s'il existe un élément $g \in \mathbb{S}'$ tel que pour tout $i \in \mathbb{S}$, $g.i \in \mathbb{S}'$ et $p'_{g.i} = p_i$; on note dans ce cas $p \sqsubseteq p'$. Si de plus pour tout $i \in \mathbb{S}$ et pour tout générateur $g_k \in \mathcal{G}$, $g.i.g_k \in \mathbb{S}'$, on dit que p *apparaît strictement* dans p' et on le note $p \sqsubset p'$. Le *support élémentaire de taille* n , noté \mathbb{S}_n , est l'ensemble des éléments

de \mathbb{M} situés à distance au plus n de l'élément neutre ε . Un *bloc de taille n* (ou *motif élémentaire de taille n*) est un motif dont le support est \mathbb{S}_n .

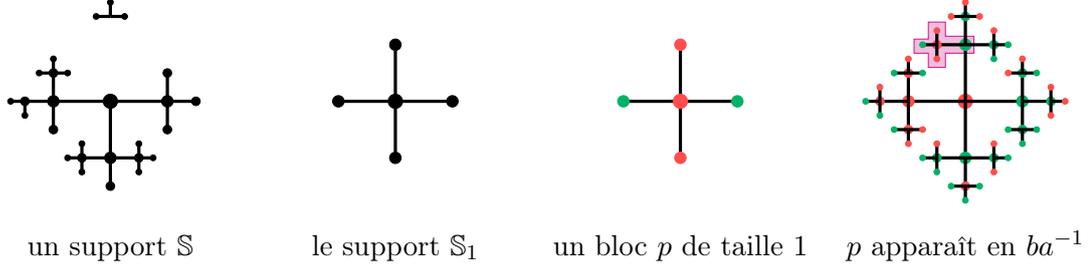


FIGURE 3: Sur le groupe libre à deux générateurs $\mathbb{F}_2 = \langle a, b | \emptyset \rangle$: un support, le support élémentaire de rayon 1, un bloc de taille 1 qui apparaît en position ba^{-1} .

Un *langage sur \mathbb{M}* est un ensemble de motifs dont le support est inclus dans \mathbb{M} , deux motifs du langage pouvant avoir des supports différents. Sans perte de généralité pour l'utilisation que nous allons en faire, comme nous le verrons dans la Partie 1.2.2 au moment de définir un décalage, on peut se ramener à ne considérer que des langages de motifs élémentaires.

Exemple 1.2.1. • Sur le monoïde \mathbb{N} et sur l'alphabet à deux éléments $\{ \blacksquare, \blacksquare \}$, on définit le langage \mathcal{L}_1 comme l'ensemble fini de motifs élémentaires suivants :

$$\mathcal{L}_1 = \left\{ \begin{array}{c} \blacksquare \blacksquare \blacksquare \\ \blacksquare \blacksquare \blacksquare \end{array}, \begin{array}{c} \blacksquare \blacksquare \blacksquare \\ \blacksquare \blacksquare \blacksquare \end{array}, \begin{array}{c} \blacksquare \blacksquare \blacksquare \\ \blacksquare \blacksquare \blacksquare \end{array}, \begin{array}{c} \blacksquare \blacksquare \blacksquare \\ \blacksquare \blacksquare \blacksquare \end{array} \right\}$$

- Sur le groupe \mathbb{Z}^2 et sur l'alphabet $\{ \blacksquare, \blacksquare \}$, on définit le langage \mathcal{L}_2 formé de tous les motifs élémentaires (dont le support est de la forme $[-n; n]^2$) présentant un \blacksquare au centre, entouré de \blacksquare , eux-mêmes entourés de \blacksquare , ...et ainsi de suite.

Motifs interdits et langage d'un décalage sur \mathbb{M}

Définition 3. Soit F un ensemble de motifs. On appelle *décalage défini par l'ensemble de motifs interdits F* l'ensemble des configurations

$$\mathbf{T}_F = \left\{ x \in A^{\mathbb{M}}, \text{ aucun motif de } F \text{ n'apparaît dans } x \right\}.$$

Tous les langages de motifs ne sont pas des langages de décalage. En effet, dès qu'un motif appartient au langage d'un décalage, alors ce motif apparaît dans une configuration. En conséquence tous ses sous-motifs, puisqu'ils apparaissent dans la même configuration, sont aussi dans le langage du décalage. De même, si un motif appartient au langage il apparaît dans une configuration infinie, donc *a fortiori* il apparaît strictement dans un autre motif du langage. Ces deux notions sont en fait suffisantes pour caractériser les langages de décalage, comme nous le montrons dans la proposition suivante.

Proposition 1.2.1. *Soit \mathcal{L} un langage de motifs sur un monoïde \mathbb{M} . Alors \mathcal{L} est le langage d'un décalage si et seulement si*

- \mathcal{L} est factoriel : $\forall m \in \mathcal{L}$, si $m' \sqsubseteq m$ alors $m' \in \mathcal{L}$;
- \mathcal{L} est prolongeable : $\forall m \in \mathcal{L}$, $\exists m' \in \mathcal{L}$, $m \sqsubset m'$.

Le langage \mathcal{L}_1 de l'Exemple 1.2.1 n'est ni factoriel ni prolongeable (comme tout autre langage fini pour ce dernier point), tandis que le langage \mathcal{L}_2 du même exemple est prolongeable mais pas factoriel car il ne contient pas le motif \blacksquare par exemple.

Démonstration. • Soit $\mathcal{L} = \mathcal{L}(\mathbf{T})$ le langage d'un décalage. Si $m \in \mathcal{L}$ et $m' \sqsubseteq m$, alors il existe $x \in \mathbf{T}$ tel que $m \sqsubseteq x$, et donc *a fortiori* $m' \sqsubseteq x$. Ainsi $m' \in \mathcal{L}$, c'est-à-dire que \mathcal{L} est stable par sous-mot élémentaire. D'autre part si $m \in \mathcal{L}$ tel que $m \in \mathcal{L}_n(\mathbf{T})$, il existe $x \in \mathbf{T}$ tel que $x|_{\mathbb{S}_n} = m$. On pose $m' = x|_{\mathbb{S}_{n+1}}$. Alors $m \sqsubset m'$ et $m' \in \mathcal{L}$ donc \mathcal{L} est prolongeable.

- Réciproquement soit $\mathcal{L} \subset \mathcal{E}$ un langage prolongeable et stable par sous-mot élémentaire. Considérons le décalage $\mathbf{T} = \mathbf{T}_{\mathcal{L}^c}$ et montrons que $\mathcal{L} = \mathcal{L}(\mathbf{T})$.
 - Si $m \in \mathcal{L}(\mathbf{T})$, alors $m \sqsubseteq x \in \mathbf{T}_{\mathcal{L}^c}$ donc $m \notin \mathcal{L}^c$ i.e. $m \in \mathcal{L}$. On a bien $\mathcal{L}(\mathbf{T}) \subseteq \mathcal{L}$.
 - Si $m \in \mathcal{L}$, alors il existe un entier n tel que $m \in \mathcal{A}^{\mathbb{S}_n}$. Comme le langage \mathcal{L} est prolongeable, il existe $m_1 \in \mathcal{A}^{\mathbb{S}_{n+1}} \cap \mathcal{L}$ tel que $m \sqsubset m_1$. En itérant le procédé, on construit une suite $(m_k)_{k \in \mathbb{N}}$ d'éléments de \mathcal{L} telle que $m \sqsubset m_1 \sqsubset \dots \sqsubset m_k \sqsubset m_{k+1} \sqsubset \dots$. Notons x_k un élément de $\mathcal{A}^{\mathbb{M}}$ tel que $(x_k)|_{\mathbb{S}_{n+k}} = m_k$. Par compacité de l'espace des configurations $\mathcal{A}^{\mathbb{M}}$, il existe ϕ une extraction telle que la suite extraite $(x_{\phi(k)})_{k \in \mathbb{N}}$ converge ; on appelle $x = \lim x_{\phi(k)}$, montrons que $x \in \mathbf{T}$. Soit m' un motif qui apparait dans x . Alors $\exists k \in \mathbb{N}$ tel que

$m' \sqsubseteq m_k$, donc $m' \notin \mathcal{L}^c$. On a bien $x \in \mathbf{T}$ donc m est autorisé dans \mathbf{T} c'est-à-dire $\mathcal{L} \subseteq \mathcal{L}(\mathbf{T})$. \square

Toute naturelle que semble être la notion de langage d'un décalage, elle n'en est pas moins extrêmement contraignante comme nous venons de le voir. C'est d'ailleurs un raison pour laquelle on définit habituellement un décalage en donnant un ensemble de motifs interdits et non son langage.

Le langage $\mathcal{L}(\mathbf{T})$ d'un décalage \mathbf{T} est donc exactement l'ensemble de tous les motifs autorisés (par opposition aux motifs interdits). Son complémentaire, que l'on notera $\mathcal{L}(\mathbf{T})^c$, est donc exactement l'ensemble de tous les motifs interdits c'est-à-dire le plus grand ensemble (au sens de l'inclusion) de motifs interdits qui définit \mathbf{T} .

Proposition 1.2.2. *Soit \mathbf{T} un décalage. Alors le complémentaire du langage $\mathcal{L}(\mathbf{T})^c$ de \mathbf{T} est un ensemble de motifs interdits qui définit $\mathbf{T} : \mathbf{T} = \mathbf{T}_{\mathcal{L}(\mathbf{T})^c}$.*

Démonstration. Soit x une configuration du décalage \mathbf{T} . Comme le langage $\mathcal{L}(\mathbf{T})$ contient par définition tous les motifs élémentaires apparaissant dans \mathbf{T} , tous les motifs élémentaires qui apparaissent dans x sont dans $\mathcal{L}(\mathbf{T})$, c'est-à-dire que x est bien une configuration du décalage défini par l'ensemble de motifs interdits $\mathcal{L}(\mathbf{T})^c$. Ainsi x est bien une configuration du décalage $\mathbf{T}_{\mathcal{L}(\mathbf{T})^c}$.

Réciproquement supposons que x soit une configuration du décalage $\mathbf{T}_{\mathcal{L}(\mathbf{T})^c}$. Alors tout motif élémentaire qui apparaît dans x est dans $\mathcal{L}(\mathbf{T})$. En particulier pour tout $n \in \mathbb{N}$, le motif élémentaire $x_{\mathbb{S}_n}$ de taille n est dans le langage de \mathbf{T} . Par conséquent il existe une configuration x_n dans le décalage \mathbf{T} qui coïncide avec x sur le support $\mathbb{S}_n : (x_n)_{|\mathbb{S}_n} = x_{|\mathbb{S}_n}$. On peut construire une configuration $x_n \in \mathbf{T}$ pour tout entier n , de sorte que la suite des x_n converge vers la configuration x . Le décalage \mathbf{T} étant fermé, on en déduit que x est bien une configuration de \mathbf{T} . \square

1.2.3 Coïncidence des définitions

Proposition 1.2.3. *Les définitions combinatoire et topologique coïncident.*

Démonstration. Supposons que \mathbf{T} est un décalage au sens combinatoire de la Définition 3. Alors il existe un ensemble de motifs interdits F tel que $\mathbf{T} = \mathbf{T}_F$. En réécrivant le décalage \mathbf{T} à l'aide de cylindres on a :

$$\mathbf{T}_F = \bigcap_{m \in F} \bigcap_{u \in \mathbb{M}} \sigma^u(\llbracket m \rrbracket^c) = \bigcap_{m \in F} \bigcap_{g \in \mathbb{M}} (\sigma_{\mathbb{M}}^g(\llbracket m \rrbracket))^c = A^{\mathbb{M}} \setminus \bigcup_{m \in F} \bigcup_{g \in \mathbb{M}} \sigma_{\mathbb{M}}^g(\llbracket m \rrbracket)$$

Cette écriture permet de retrouver facilement la définition topologique : la stabilité par translation $\sigma_{\mathbb{M}}$ apparaît clairement, et le décalage \mathbf{T} est un ensemble fermé comme intersection de fermés.

Réciproquement supposons que \mathbf{T} soit un décalage sur \mathbb{M} au sens topologique de la Définition 2. Alors par la Proposition 1.2.2 on sait que le complémentaire du langage de \mathbf{T} est un ensemble de motifs interdits qui le définit. \square

La première définition, topologique, (voir la Partie 1.2.1) provient directement de la théorie des systèmes dynamiques. Un décalage est vu comme un ensemble fermé et stable par l'action naturelle du monoïde (par translation). Cette définition donne aussi un critère simple pour vérifier qu'un ensemble de configurations est bien un décalage, sans toutefois avoir besoin d'exhiber un ensemble de motifs interdits.

La deuxième définition, combinatoire, (voir la Partie 1.2.2) est utilisable très concrètement. Elle permet notamment de construire un décalage en explicitant un ensemble de motifs interdits ; chaque motif interdit ou autorisé correspond en fait à une règle locale que les configurations du décalage doivent respecter. La plupart des exemples présentés dans ce manuscrit utilisent cette définition. Comme nous le verrons par la suite, c'est aussi la définition qui se rapproche le plus de celles des pavages du plan discret \mathbb{Z}^2 et des diagrammes espace-temps de machines de Turing, objets sur lesquels nous reviendrons plus en détails dans la Partie 2.3.1.

1.3 Classes de décalages

Cette partie introduit les trois classes de décalages dont il sera question dans la suite de cette thèse. Les deux premières, la classe des décalages de type fini et la classe des décalages sofiques, sont l'expression de contraintes locales sur les configurations. La troisième, celle des décalages effectifs, apparaîtra naturellement dans la Partie 2.2.3 en appliquant l'opération de sous-action projective aux deux classes précédentes.

1.3.1 Décalages de type fini

Les décalages triviaux sont les ensembles de configurations $A^{\mathbb{M}}$, on note par \mathcal{FS} (de l'anglais *full-shift*) la classe des décalages triviaux. Les décalages de type fini, dont les configurations vérifient un nombre fini de contraintes locales, sont les décalages les plus simples à définir et les plus aisément manipulables (décalages triviaux mis à part).

Définition 5. Un décalage est dit *de type fini* si on peut le définir par un ensemble fini de motifs interdits. On note par \mathcal{SFT} (de l'anglais *shift of finite type*) la classe des décalages de type fini.

Si \mathbf{T} est un décalage de type fini donné par l'ensemble F de motifs interdits, on peut supposer que tous les motifs de F ont le même support. Si ce n'est pas le cas, on transforme F en un ensemble de motifs interdits de même support de la manière suivante. Soit k la taille du plus grand support d'un motif de F . Pour tout motif $m \in F$, on remplace m par l'ensemble de motifs $\mathcal{L}_k([m])$ de support \mathbb{S}_k . Si F est un ensemble de motifs qui sont tous de support \mathbb{S}_k , on dit que l'ensemble F est d'ordre k . Étant donné un décalage de type fini \mathbf{T} , le plus petit entier pour lequel on peut trouver un ensemble d'ordre k qui définit \mathbf{T} est appelé *l'ordre* du décalage \mathbf{T} .

Exemple 1.3.1. Le décalage \mathbf{T}_1 de l'Exemple 1.2.2 est un décalage de type fini d'ordre 2, puisqu'il suffit d'interdire le motif  pour le définir.

1.3.2 Décalages sofiques

On considère un nouvel alphabet A' , pas nécessairement différent de A , ainsi que \mathbf{T} un décalage sur l'alphabet A et m un entier naturel. On définit tout d'abord les fonctions de bloc, qui permettent de recoder l'alphabet A en un nouvel alphabet A' , ce recodage se faisant en respectant des règles locales.

Définition 6. Une application $\Phi : \mathbf{T} \subseteq A^{\mathbb{M}} \rightarrow A'^{\mathbb{M}}$ est appelée *m-fonction de bloc* s'il existe une application locale $\phi : \mathcal{L}_m(\mathbf{T}) \rightarrow A'$ telle que

$$\text{pour tout } g \in \mathbb{M}, \Phi(x)_g = \phi(x_{|_{g.\mathbb{S}_m}}),$$

où $x_{|_{g.\mathbb{S}_m}}$ est le motif de support \mathbb{S}_m qui apparaît en position g dans la configuration x (voir la Figure 4). On dit dans ce cas que ϕ est la *fonction locale de* Φ . Le plus petit entier m vérifiant cette propriété est appelé *la mémoire* de la fonction de bloc Φ .

Dans le cas où $\mathbb{M} = \mathbb{Z}$, le théorème de Curtis-Lyndon-Hedlund (voir Théorème 3.4 de [Hed69]) nous apprend que les fonctions de blocs sont exactement les applications $\Phi : X \rightarrow Y$ qui sont continues et qui commutent avec l'action $\sigma_{\mathbb{Z}}$. Cette propriété reste vraie dans le cas d'un monoïde finiment présenté.

Proposition 1.3.1. *Les fonctions de bloc sur \mathbb{M} sont exactement les applications $\Phi : A^{\mathbb{M}} \rightarrow A'^{\mathbb{M}}$ qui sont continues et qui commutent avec la translation $\sigma_{\mathbb{M}}$.*

Cette caractérisation des fonctions de bloc nous permet de montrer facilement qu'elles transforment des décalages en décalages.

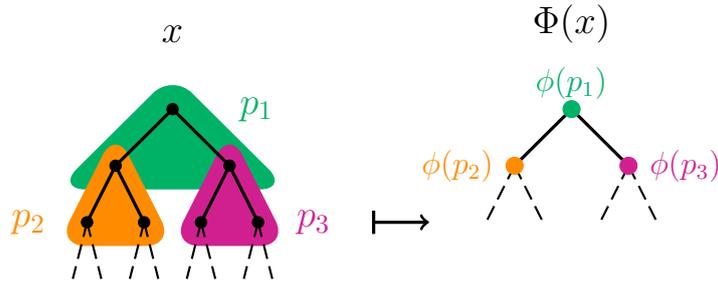


FIGURE 4: Une 1-fonction de bloc Φ appliquée à une configuration sur \mathbb{M}_2 . L'image de la configuration x par Φ est calculée grâce à sa fonction locale ϕ . Dans $\Phi(x)$, la lettre d'un nœud dépend de la lettre de ce nœud et de celles de ses deux fils dans la configuration x .

Proposition 1.3.2. *L'image d'un décalage par une fonction de bloc est également un décalage.*

Démonstration. Considérons un décalage \mathbf{T} et une fonction de bloc Φ de mémoire m . Il suffit de montrer que l'ensemble $\Phi(\mathbf{T})$ est fermé et invariant par σ^g pour tout élément $g \in \mathbb{M}$. Comme Φ commute avec σ^g , $\sigma^g(\Phi(\mathbf{T})) = \Phi(\sigma^g(\mathbf{T}))$. Le décalage \mathbf{T} commute avec σ^g , on a donc $\Phi(\sigma^g(\mathbf{T})) = \Phi(\mathbf{T})$. L'ensemble $\Phi(\mathbf{T})$ est donc invariant par σ^g . Enfin, $\Phi(\mathbf{T})$ étant l'image continue d'un ensemble compact, il s'agit d'un ensemble fermé et donc d'un décalage. \square

Proposition 1.3.3. *L'inverse d'une fonction de bloc bijective entre deux décalages est également une fonction de bloc.*

Démonstration. Soit Φ une fonction de bloc bijective. Alors Φ possède un inverse Φ^{-1} , et il est facile de voir que cette fonction inverse commute avec la translation $\sigma_{\mathbb{M}}$. Par compacité de l'espace des configurations $A^{\mathbb{M}}$, la fonction Φ^{-1} est continue. D'après la Proposition 1.3.1, la fonction Φ^{-1} est donc aussi une fonction de bloc. \square

Une fonction de bloc bijective entre \mathbf{T} et \mathbf{T}' est appelée une *conjugaison*. On dit dans ce cas que les décalages \mathbf{T} et \mathbf{T}' sont *conjugués*. Si la k -fonction de bloc Φ est une conjugaison, on dit que c'est une *k -conjugaison*.

Le résultat de la Proposition 1.3.2 nous permet de définir une nouvelle classe de décalages, en nous intéressant à l'image des SFT par les fonctions de bloc. Les SFT sont des objets que l'on peut décrire de manière très simple, en énumérant un ensemble fini de motifs interdits. De la même façon, une fonction de bloc

Définition 8. Une machine de Turing avec ruban \mathbb{M} est la donnée de $\mathcal{M} = (Q, \Gamma, \#, q_0, \delta, Q_F)$ où :

- Q est un l'ensemble fini des états possibles pour la tête de calcul ; $q_0 \in Q$ est l'état initial ;
- Γ est un alphabet fini ;
- $\# \notin \Gamma$ est le symbole blanc ;
- $\delta : Q \times \Gamma \rightarrow Q \times \Gamma \times (E \cup \{\varepsilon\})$ est la fonction de transition, où E est l'ensemble des arêtes du graphe non orienté $\mathcal{T}_{\mathbb{M}}$ associé au monoïde \mathbb{M} . Étant donné un état de la tête de calcul et la lettre inscrite sur le ruban à la position de la tête de calcul, cette lettre est remplacée par une nouvelle lettre, la tête de calcul se déplace sur une cellule adjacente (ou bien reste en place) et passe dans un nouvel état ;
- $F \subset Q_F$ est l'ensemble des états finaux. Lorsqu'un état final est atteint, la machine arrête son calcul.

Une *configuration* de la machine \mathcal{M} est la description de son ruban à une étape de calcul donnée. Il s'agit donc d'un élément de l'ensemble $(\Gamma \cup (Q \times \Gamma))^{\mathbb{M}}$, avec la contrainte de n'avoir qu'une seule tête de lecture, et seulement un nombre fini de cellules qui ne contiennent pas le symbole blanc $\#$. Un *calcul* de la machine \mathbb{M} sur le motif d'entrée p de support \mathbb{S} est la suite de configurations (c_n) , où c_0 est la configuration telle que pour tout $g \in \mathbb{M} \setminus \mathbb{S}$, $(c_0)_g = \#$ et pour tout $g \in \mathbb{S}$, $(c_0)_g = p_g$, et les configurations suivantes sont telles qu'on peut passer de la configuration c_n à la configuration c_{n+1} en appliquant la fonction de transition de la machine \mathcal{M} . Un calcul de \mathcal{M} sur un motif d'entrée p est fini si et seulement si un état final est atteint ; on dit dans ce cas que la machine \mathcal{M} *s'arrête* sur le motif d'entrée p . On appelle *domaine* de la machine \mathcal{M} l'ensemble des motifs d'entrée sur lesquels la machine \mathcal{M} s'arrête.

Lorsque le monoïde \mathbb{M} est \mathbb{N} (resp. \mathbb{Z}), on retrouve la définition classique des machines de Turing avec ruban semi-infini (resp. bi-infini) [Rog87]. La thèse de Church propose de formaliser la notion de calcul effectif par ces machine de Turing classiques. Cette hypothèse est largement confortée par le fait que les différents modèles de calcul proposés pour définir cette notion sont de puissance équivalente (programmes RAM, lambda-calcul, fonctions récursives). Cependant lorsque l'on définit des machines de Turing ayant comme ruban le graphe de Cayley d'un monoïde \mathbb{M} ou groupe \mathbb{G} finiment présenté quelconque, la puissance de calcul de ces machines peut être strictement supérieure à celle des machines de Turing classiques (une machine de Turing classique lit les lettres d'un motif sur \mathbb{M} ou \mathbb{G} dans l'ordre lexicographique de leur position).

Théorème 1.3.4 (Folklore). *Les machines de Turing sur \mathbb{G} ont même puissance de calcul que les machines de Turing classique si et seulement si le problème du mot est décidable sur \mathbb{G} .*

Dans les cas particuliers étudiés dans cette thèse (décalages sur \mathbb{Z}^d et décalages sur \mathbb{M}_2), les machines de Turing peuvent donc toujours être pensées comme des machines de Turing au sens classique.

Étant donné un langage de motifs \mathcal{L} , une machine de Turing \mathcal{M} peut être vue de deux façons comme un reconnaisseur de ce langage.

Définition 9. On dit que la machine \mathcal{M} *décide*, ou *reconnaît*, le langage \mathcal{L} si la machine \mathcal{M} s'arrête sur tout motif d'entrée, et répond de manière positive si le motif d'entrée appartient au langage \mathcal{L} , de manière négative sinon. On dit alors que le langage \mathcal{L} est *récuratif*.

Définition 10. On dit que la machine \mathcal{M} *énumère* le langage \mathcal{L} si la machine \mathcal{M} s'arrête exactement sur les motifs du langage \mathcal{L} . On dit dans ce cas que le langage \mathcal{L} est *récurativement énumérable*.

Remarque. Tout langage récuratif est aussi récurativement énumérable, mais la réciproque est fausse.

Définition 11. Un décalage \mathbf{T} est *effectif* s'il existe un ensemble de motifs récurativement énumérable F qui le définit, c'est-à-dire tel que $\mathbf{T} = \mathbf{T}_F$. On appelle \mathcal{RE} la classe des décalages effectifs.

On pourrait définir de manière analogue la classe des décalages récuratifs comme la classe des décalages qui peuvent être définis par un ensemble récuratif de motifs interdits. Cette classe est incluse dans la classe des décalages effectifs.

Proposition 1.3.5. *Pour tout décalage effectif \mathbf{T} , il existe un ensemble récuratif de motifs interdits F tel que $\mathbf{T} = \mathbf{T}_F$.*

Démonstration. Soit \mathbf{T} un décalage effectif défini sur un alphabet A . Alors il existe un ensemble récurativement énumérable F de motifs interdits tel que $\mathbf{T} = \mathbf{T}_F$. Sans perte de généralité on peut supposer que F est le complémentaire du langage de \mathbf{T} , c'est-à-dire $F = \mathcal{L}(\mathbf{T})^c$. Soit $(m_k)_{k \in \mathbb{N}}$ un énumération de l'ensemble F . On peut lui associer la suite $(n_k)_{k \in \mathbb{N}}$ des tailles des supports des motifs m_k . Si la suite $(n_k)_{k \in \mathbb{N}}$ est croissante, alors l'ensemble F est aussi récuratif (voir [Rog87, Théorème III p. 59]) et donc la proposition est démontrée. Sinon, on construit un nouvel ensemble \tilde{F} qui possède une énumération $(\tilde{m}_k)_{k \in \mathbb{N}}$ telle

que la suite des tailles de support $(\tilde{n}_k)_{k \in \mathbb{N}}$ est croissante, et qui définit le même décalage \mathbf{T} .

Pour tout motif m_k on s'intéresse à l'entier $\max_{i \leq k} n_i$. Si cet entier est égal à n_k , c'est-à-dire que tous les motifs qui précèdent m_k dans la suite sont définis sur des supports de taille inférieure ou égale à celle du support de m_k , alors le motif m_k apparaîtra tel quel dans la suite $(\tilde{m}_k)_{k \in \mathbb{N}}$, pas nécessairement au même rang. Si l'entier $\max_{i \leq k} n_i$ est différent de n_k , alors le motif m_k est précédé dans la suite par un motif dont le support est de taille strictement supérieure. Dans la nouvelle suite $(\tilde{m}_k)_{k \in \mathbb{N}}$, le motif m_k sera remplacé par une suite de motifs $\tilde{m}_{\phi(k)}, \dots, \tilde{m}_{\phi(k)+\psi(k)}$ de même support de taille $\max_{i \leq k} n_i$. Les fonctions ϕ et ψ dépendent de la suite $(n_k)_{k \in \mathbb{N}}$ et sont définies récursivement de la manière suivante :

- $\phi(0) = 0$ et $\psi(0) = 0$;
- pour tout entier n la fonction ϕ est donnée par $\phi(n+1) = \phi(n) + \psi(n) + 1$.
La fonction ψ est donnée par :

$$\psi(n+1) = \begin{cases} 0 & \text{si } n_k = \max_{i \leq k} n_i \\ (|\mathbb{S}_{\max_{i \leq k} n_i}| - |\mathbb{S}_{n_k}|)^A & \text{si } n_k \neq \max_{i \leq k} n_i. \end{cases}$$

Les motifs $\tilde{m}_{\phi(k)}, \dots, \tilde{m}_{\phi(k)+\psi(k)}$ sont l'ensemble des motifs de support $\mathbb{S}_{\max_{i \leq k} n_i}$ dans lesquels le motif m_k apparaît en position ε .

Ainsi le nouvel ensemble de motifs interdits \tilde{F} est récursif, et tel que $\tilde{F} \subseteq F$. Plus précisément, tout motif de F apparaît dans un motif de \tilde{F} . Il définit aussi le décalage \mathbf{T} . Soit x une configuration du décalage $\mathbf{T}_{\tilde{F}}$. Alors x ne peut pas contenir de motif de F , car sinon il contiendrait un motif de \tilde{F} ce qui est contradictoire. Donc $\mathbf{T} = \mathbf{T}_{\tilde{F}}$ ce qui montre la proposition. \square

La Proposition 1.3.5 montre qu'en fait la classe des décalages récursifs est exactement celle des décalages effectifs, c'est pourquoi dans la suite de cette thèse il ne sera plus question de décalages récursifs.

Exemple 1.3.3. On construit un exemple de décalage sur \mathbb{Z} effectif qui n'est pas sofique. On se place sur l'alphabet $A = \{0, a, b\}$, et on considère le décalage $\mathbf{T} = \mathbf{T}_{\{ba; 0a^m b^n 0; m \neq n\}}$. Alors son langage est

$$\mathcal{L}(\mathbf{T}) = \{m \in \{0, a, b\}^* \mid \text{les blocs contenant des } a \text{ et des } b \text{ sont de la forme } a^n b^n\},$$

qui est un exemple classique de langage non rationnel. Donc le décalage \mathbf{T} est un décalage effectif non sofique.

La classe des décalages effectifs correspond donc aux décalages calculables. Elle contient les deux classes définies précédemment, de manière stricte sur \mathbb{Z} comme le montrent les exemples 1.3.3 et 1.3.2 :

$$\mathcal{SFT} \subseteq \mathcal{Sofique} \subseteq \mathcal{RE}.$$

Dans la Partie 2.2.3 du Chapitre 2 consacré aux décalages sur \mathbb{Z}^d , l'étude de la sous-action projective fera apparaître les décalages effectifs comme ceux obtenus en appliquant cette opérations aux décalages sofiques.

