Projet de recherche

MÉTHODES PROFINIES, ERGODIQUES ET COMBINATOIRES EN DYNAMIQUE SYMBOLIQUE

HERMAN GOULET-OUELLET

Table des matières

Introduction

Ce projet de recherche s'inscrit dans le cadre de la dynamique symbolique et ses liens avec la théorie des semigroupes, la combinatoire des mots et la théorie ergodique. On y retrouve aussi des thèmes de théorie des automates et de théorie des groupes. La dynamique symbolique est, en résumé, l'étude des mots infinis vus comme systèmes dynamiques sous l'action de la fonction de décalage (dont la définition est rappelée plus loin dans l'introduction). Ces systèmes proviennent de constructions tant géométriques, initialement motivées par l'étude des orbites dans certains systèmes dynamiques (codages de rotations et d'échanges d'intervalles), que combinatoires (substitutions, systèmes S-adiques).

La dynamique symbolique trouve son origine dans les travaux éponymes de Hedlund et Morse sur les codages de rotations des cercles [48, 49]. Hedlund et Morse développent l'étude systématique des trajectoires symboliques obtenues en encodant les orbites d'un système donné grâce à une partition finie. Ils appliquent leurs méthodes à l'étude des rotations irationnelles d'un cercle unité (voir la figure 1). Leurs travaux mènent à la notion de mot sturmien, qui demeure un sujet de recherche actif à ce jour.

La dynamique symbolique est aujourd'hui un champs d'études à part entière ayant des connexions avec plusieurs autres domaines, comme la combinatoire, l'algèbre, la géométrie (fractales, pavages), et la logique. Les systèmes dynamiques apparaissent aussi naturellement, par exemple, dans l'étude des machines de Turing, dans les systèmes de numération, ou encore dans la vérification de modèles. Mes propres recherches se concentrent principalement sur des questions algébriques, combinatoires, et ergodiques.

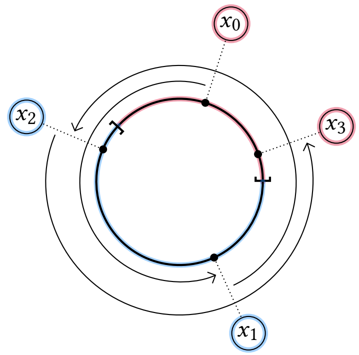

FIGURE 1 – Mot infini codant l'orbite d'un point sous une rotation d'un cercle. Le mot infini qui encode l'orbite $x_0, x_1, x_2, \dots$ commence par *baab*, où *a* représente la région bleue et *b* la région rouge.

Les questions étudiées dans ce projet portent plus précisément sur la notion de shift (parfois appelée sous-shift), notion centrale en dynamique symbolique. Soit A un alphabet, A^* le monoïde libre sur A et $A^{\mathbb{Z}}$ l'ensemble des mots bi-infinis sur A. On munit A de sa topologie discrète et $A^{\mathbb{Z}}$ de sa topologie produit. On considère l'action sur $A^{\mathbb{Z}}$ de la fonction $S \colon A^{\mathbb{Z}} \to A^{\mathbb{Z}}$, appelée fonction de décalage, définie par $(Sx)_n = x_{n+1}$. On appelle shift un sous-ensemble $X \subseteq A^{\mathbb{Z}}$ fermé topologiquement et invariant par S. Un shift minimal est un shift qui ne contient aucun sous-shift propre et non-vide.

L'exemple suivant est un objet d'étude classique en dynamique symbolique qui nous servira à illustrer différentes notions tout au long de ce document. Soit le mot infini à gauche $\mathbf{t} = (t_k)_{k \in \mathbb{N}}$ sur deux symboles $\{a, b\}$ défini de la manière suivante : pour chaque $k \in \mathbb{N}$, $t_k = a$ si le nombre de 1 dans la représentation binaire de k est pair, et $t_k = b$ sinon. Ainsi, on a par exemple, $t_5 = a$ puisque $(5)_2 = 101$, ou encore $t_{11} = b$ car $(11)_2 = 1011$. On obtient le mot binaire infini

$$\mathbf{t} = abbabaabbaababbabaababbaabbabaabbaababba \cdots$$

appelé mot de Thue–Morse. On peut également construire ce mot de manière combinatoire en itérant la substitution $\sigma \colon a \mapsto ab, b \mapsto ba$ sur le symbole a. Le mot de Thue–Morse possède de nombreuses propriétés intéressantes (voir le survol [1]). Par exemple, il évite le motif *uuu* (on dit qu'il est *sans cube*). Cette propriété est liée à plusieurs questions classiques en combinatoire des mots, notamment le théorème de Dejean sur les seuils de répétition. Le shift de Thue–Morse, c'est-à-dire la fermeture topologique de l'orbite du mot de Thue–Morse, est un exemple de shift minimal.

L'objectif du projet est d'étudier les aspects ergodiques et profinis des shifts en connexion avec leurs propriétés combinatoires. À plus long terme, le projet vise à mieux comprendre le lien entre les invariants profinis et d'autres types d'objets, comme les mesures ergodiques, les mots de retour, et les groupes de dimension. Le projet est séparé en trois volets.

1. **Invariants profinies.** Le premier volet vise à étudier les invariants profinis associés aux shifts minimaux dans un cadre S-adique, plus précisément les groupes de Schützenberger.

2. **Densité des langages rationels.** Le deuxième volet porte sur la densité des langages rationnels sous les mesures ergodiques des shifts, notamment grâce à une approche cohomologique (cocycles, cobords).
3. **Aspects combinatoires.** Le troisième et dernier volet du projet vise à étudier les propriétés combinatoires des shifts, en particulier à travers les notions de mots de retour et de facteurs bispéciaux, avec l'objectif de développer des outils utiles aux deux autres volets.

1. Invariants profinis

Le premier volet vise à étudier les shifts minimaux via l'adhérence de leur langage dans monoïde profini libre. L'étude de cette adhérence donne lieu à un invariant de conjugaison appelé groupe de Schützenberger, prenant la forme d'un groupe profini.

Les semigroupes profinis sont des semigroupes topologiques compacts et totalement déconnectés (comme l'espace de Cantor). Ils apparaissent naturellement en théorie des automates et dans l'étude des variétés de langages (voir par exemple [50]). On retrouve aussi des structures profinis, sous diverses formes, en théorie de Galois et en analyse p-adique [32, 52], dans l'étude des algèbres de Boole, ou encore en logique [22, 36].

Dans le contexte de la théorie des automates, les semigroupes profinis mènent à une version topologique de la correspondance d'Eilenberg, qui établit une bijection entre les variétés de langages et les variétés de semigroupes finis. Le monoïde profini libre, dénoté $\widehat{A^*}$, est défini comme la complétion de A^* sous la distance profinie. La distance profinie entre deux mots u et v est, très brièvement, l'inverse du plus petit monoïde de transition d'un automate qui distingue u et v. En d'autres mots, plus u et v sont difficiles à distinguer par un automate, plus ils sont proches pour la distance profinie. Il existe une dualité entre les langages rationnels et le monoïde profini libre, sous laquelle les ouverts-fermés du monoïde profini libre correspondent aux les langages rationnels. Cette dualité permet d'obtenir des versions topologiques de plusieurs énoncés sur les langages rationnels.

La connexion entre les semigroupes profinis et la dynamique symbolique remonte quant à elle aux travaux d'Almeida dans les années 2000 [2, 3, 4]. Ses travaux permettent notamment d'associer à tout shift minimal X un groupe profini, appelé groupe de Schützenberger, que l'on notera G(X) [7]. Notons au passage que la construction de ce groupe est liée aux classes de Green, dont la définition est rappelée à la section 2. La correspondance entre les semigroupes profinis et la dynamique symbolique a un double objectif : d'une part, elle permet d'étudier la structure des monoïdes profinis libres d'un point de vue dynamique, et d'autre part elle permet d'appliquer des méthodes profinies à l'étude de questions dynamiques et combinatoires. Ces idées ont été explorées par plusieurs chercheurs au cours des deux dernières décennies [5, 6, 7, 8, 19, 20, 21, 37, 38] et l'ouvrage [9] offre un exposé détaillé sur le sujet.

La structure du groupe G(X) a été étudiée principalement dans le cas des shifts substitutifs. Rappelons qu'une substitution est un endomorphisme $\sigma\colon A^* \to A^*$ sur le monoïde libre. L'application répétée d'une substitution σ définit un shift $X(\sigma)$, soit le plus grand shift dont tous les facteurs finis apparaissent dans les mots de la forme $\sigma^n(a)$, $a \in A$ et $n \geq 0$. Un shift de la forme $X(\sigma)$ est dit substitutif. Par exemple, le shift de Thue–Morse (décrit à l'introduction) est substitutif avec $\sigma\colon a \mapsto ab, b \mapsto ba$. Dans un article publié en 2013, Almeida et Costa [7] ont décrit des présentations du groupe G(X) par générateurs et relations dans le cas substitutif. Ces présentations se sont révélées particulièrement utiles pour étudier la structure de G(X) [37, 38].

Le premier objectif de ce volet est de généraliser les présentations d'Almeida et Costa au-delà du cas substitutif, plus précisément dans le cadre des systèmes S-adiques. Le formalisme S-adique est une méthode puissante et flexible de représentation des shifts, comparable par exemple aux tours de Kakutani–Rokhlin ou aux diagrammes de Bratteli. La puissance expressive du modèle S-adique éclipse largement celle du modèle substitutif. L'étude des représentations S-adiques constituent un sujet de recherche présentement très actif. On peut citer par exemple des caractérisations S-adiques obtenues récemment pour certaines familles de shifts, comme les shifts dendriques [33, 34] et les shifts à croissance linéaire [29], ce dernier résultat constituant une solution possible à la fameuse conjecture S-adique proposée dans l'ouvrage collectif [30].

Problème 1. *Développer une théorie S-adique profinie dans le but d'obtenir des présentations pour les groupes de Schützenberger en dehors du cas substitutif.*

Ces présentations pourront être appliqués à l'étude des groupes de Schützenberger de shifts non-substitutifs. Parmi les applications possibles, mentionnons certaines questions de décidabilité, par exemple décider si un groupe fini donné est un quotient de G(X). Cette question a été résolue dans le cas substitutif par Almeida et Costa [7] mais demeure ouverte dans le cas général. Cette partie du projet fait l'objet d'une collaboration en cours avec J. Almeida et A. Costa, dont une première partie sera disponible sous peu. Nos méthodes font appel à la théorie des semigroupoïdes profinis et utilisent plusieurs avancées récentes dans l'étude des suites S-adiques, notamment sur les propriétés de reconnaissabilité [11, 15].

En parallèle, je propose de développer l'étude des groupes de Schützenberger à travers leurs quotients maximaux. J'ai déjà mis en œuvre cette approche en étudiant les quotients maximaux pronilpotents dans le cas substitutif (voir la première section du rapport sur mes travaux antérieurs). Plus généralement, le quotient maximal déterminé par une propriété algébrique (abélien, nilpotent, p-groupe, etc.) est le plus grand quotient d'un groupe possédant cette propriété. Les quotients maximaux permettent d'obtenir un groupe plus simple à étudier, tout en préservant certaines informations clés sur le groupe original. Étudier d'autres types de quotients maximaux permettrait de mieux comprendre la relation entre le groupe de Schützenberger et différentes propriétés du shift.

Problème 2. *Étudier la structure des quotients maximaux prorésolubles des groupes de Schützenberger associés aux shifts minimaux.*

Un point de départ prometteur serait d'étudier d'abord les quotients prométabéliens dont la structure est plus simple. L'approche que j'ai en tête est inspirée du calcul de Fox [31] et d'un théorème de représentation de Magnus [45]. Cette ligne de recherche est également liée à certains problèmes ouverts en théorie des groupes sur les topologies profinies du groupe libre (voir [46]).

2. Densités des langages rationnels

Le deuxième volet vise à étudier la notion de densité des langages rationnels en lien avec les propriétés ergodiques des shifts. Cette partie du projet fait suite à certains de mes travaux récents,

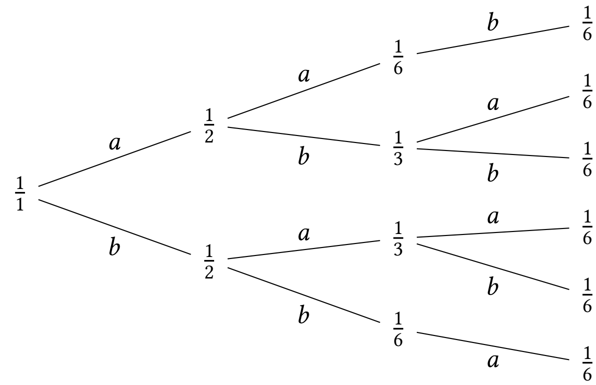

FIGURE 2 – L'unique mesure ergodique μ sur le shift de Thue–Morse. Les nœuds représentent la valeur de μ sur les ensembles $\{x \in A^{\mathbb{Z}} \mid x_0 \ldots x_{|w|-1} = w\}$, $w \in A^*$.

présentés à la section 2 du rapport sur mes travaux antérieurs. Rappelons que la densité naturelle d'un langage L sur un alphabet A, introduite par Berstel [12], est définie par

$$\delta(L) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \frac{|L \cap A^i|}{|A|^i}.$$

Il s'agit donc de la limite en moyenne de la probabilité qu'un mot de longueur n soit dans L. Notons qu'on peut voir cette définition comme une généralisation de la notion de densité naturelle étudiée en théorie des nombres. Il s'agit d'une notion classique et largement étudiée en théorie des langages et des codes [12, 13, 17, 39, 40, 42, 43, 44, 51, 53].

Dans la définition de densité naturelle, les mots finis sont tirés au hasard de manière uniforme, mais il est aussi naturelle de considérer le cas plus générale où μ est une mesure qui charge un certain shift X (par exemple minimal). On définit alors la densité de L relativement à μ par

$$\delta_\mu(L) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \mu(L \cap A^n).$$

Dans le cadre de ce projet, on s'intéresse plus précisément au cas où la mesure μ est une mesure ergodique. Une mesure est ergodique si, pour toute partie borélienne $B \subseteq X$ on a

$$\mu(B) = \mu(S^{-1}B) \qquad \text{et} \qquad B = S^{-1}B \implies \mu(B) = 0 \text{ ou } 1.$$

À titre de motivation, rappelons qu'un shift substitutif minimal admet une unique mesure ergodique [47] (voir la figure 2). Dans ce contexte, on peut interpréter la densité d'un langage rationnel comme la fréquence d'apparition d'une expression régulière dans X.

Dans une collaboration en cours avec V. Berthé et D. Perrin, nous avons entrepris l'étude des densités des langages rationnels sous les mesures ergodiques sur $A^{\mathbb{Z}}$. Un premier problème sur lequel nous nous penchons est celui de l'existence.

Problème 3. *Montrer l'existence des densités des langages rationnels sous toute mesure de* $A^{\mathbb{Z}}$.

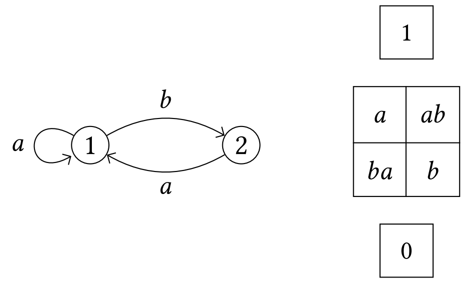

Figure 3 – Un automate (à gauche) et le diagramme en boîtes d'oeufs de son monoïde de transition (à droite). Le monoïde a six éléments répartis dans trois $\mathcal{J}$-classes, représentés par les grandes boîtes. Les petites boîtes représentent les $\mathcal{H}$-classes, qui dans cet exemple sont toutes des singletons.

Nos travaux nous ont déjà mené à une solution, non-encore publiée. Notre approche, similaire à celle adoptée dans notre article sur les langages à groupe [14], se base encore une fois cocycles. Dans le cas d'un langage rationnel général, la définition du cocycle devient beaucoup plus subtile : elle fait intervenir la structure du monoïde de transition, et plus particulièrement ses classes sous les relations de Green, un outil classique en théorie des semigroupes. On dit que deux éléments s, t d'un monoïde M sont $\mathcal{R}$-équivalents, $\mathcal{L}$-équivalents ou $\mathcal{J}$-équivalents si

$$s\mathrm{M} = t\mathrm{M}, \quad \mathrm{M}s = \mathrm{M}t \quad \text{ou} \quad \mathrm{M}s\mathrm{M} = \mathrm{M}t\mathrm{M},$$

respectivement. Les relations $\mathcal{R}$, $\mathcal{L}$ et $\mathcal{J}$, de même que l'intersection $\mathcal{H} = \mathcal{R} \cap \mathcal{L}$, sont appelés relations de Green. Ce sont des relations d'équivalences permettant de décomposer M de manière naturelle. Notons que dans le cas des groupes, toutes les relations de Green se réduisent à une seule classe, égale au groupe en entier. La décomposition d'un monoïde en classes de Green est traditionnellement représentée par un diagramme de « boîtes d'oeufs », comme illustré à la figure 3. Mentionnons au passage que le groupe de Schützenberger au projet volet est en fait isomorphe à une certaine $\mathcal{H}$-classe de $\widehat{\mathrm{A}^*}$.

Notre preuve de l'existence des densités repose sur un examen de la structure de la $\mathcal{J}$-classe minimale du monoïde de transition contenant les images des facteurs du shift. La preuve qu'on obtient est cependant non-constructive, c'est-à-dire qu'elle ne permet pas d'obtenir des formules closes pour les densités dans le cas général. Néanmoins, dans le cas d'une mesure ergodique supportée par un shift minimal, nos recherches suggèrent qu'il devrait être possible d'adapter la méthode des cobords, décrite à la section 2 du rapport sur mes travaux antérieurs, afin d'attaquer le problème suivant.

Problème 4. *Développer une méthode effective pour calculer les densités des langages rationnels sous les mesures ergodiques supportés par les shifts minimaux.*

Comme dans le cas des langages à groupes, il semble que les densités s'expriment en terme des fonctions cobordantes et que l'existence de ces fonctions soit liée aux propriétés des mots de retour, quoi que d'une manière plus subtile. On note par contre que l'approche cohomologique ne semble pas fonctionner dans certains cas (lorsque le produit croisé associé au cocycle admet plus

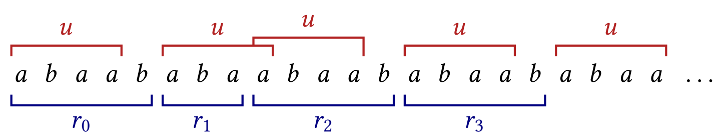

FIGURE 4 – Découpage du mot de Fibonacci en mots de retour à $u = abaa$.

d'une seule mesure ergodique qui se projette sur la mesure originale). Il s'agit d'un problème qui demeure en partie ouvert même dans le cas des langages à groupes.

Enfin, mentionnons une conséquence remarquable d'une solution positive au problème 3 : la fonction de densité δ_μ induit une mesure de probabilité sur chaque monoïde fini, et donc par passage à la limite, une mesure de probabilité sur le monoïde profini libre. Cela nous mène au problème suivant, qui fait le pont vers le premier volet du projet.

Problème 5. *Étudier les mesures de probabilités induites sur les monoïdes profinis libres par les mesures ergodiques sur les shifts.*

En particulier, il est naturel de demander s'il existe un lien entre ces mesures et les groupes de dimension. Dans le contexte de la dynamique symbolique, les groupes de dimension sont des groupes abéliens ordonnés qui capturent la structure combinatoire et dynamique des shifts. La structure du groupe de dimension est étroitement liée à l'ensemble des mesures ergodiques sur le shift ainsi qu'aux cocycles et cobords. Les mesures profinies induites par les densités pourraient permettre de clarifier les liens, encore mal compris, entre les groupes de dimension et les groupes de Schützenberger.

3. Aspects combinatoires

Le troisième et dernier volet porte sur les deux notions de mots de retour et facteurs bispéciaux. Les problèmes présentés ici font suite aux travaux présentés à la section 3 du rapport sur mes travaux antérieurs et sont motivés par leurs applications potentielles aux deux premiers volets du projet.

Les mots de retour constituent un thème central en dynamique symbolique. Ils permettent notamment de construire des représentations S-adiques [27], de caractériser plusieurs familles de shifts [10, 24, 29, 33, 54], de résoudre des problèmes de décidabilité [25, 26, 28], et d'étudier les codes bifixes [16]. Ils interviennent aussi dans des sujets plus combinatoires, comme dans l'étude des exposants critiques [23] et de la richesse palindromique [35]. Enfin, ils sont utiles pour obtenir les relations dans les présentations profinies évoquées au premier volet de même que les fonctions cobordantes rencontrées au deuxième volet.

Étant donné un facteur fini u dans un shift X, un mot de retour à u dans X est un mot $r \in A^*$ qui sépare deux occurrences consécutives de u dans X. Autrement dit, u est un mot tel que ru est un facteur de X et $ru \in uA^* \setminus A^+uA^+$ (où A^+ dénote l'ensemble des mots non-vides). La notion est illustrée à la figure 4. Dans cet exemple, les deux mots de retour aba, $abaab$ forment une base du groupe libre sur deux générateurs. Il s'agit d'un instance du Théorème du retour (voir la section 3 du rapport sur mes travaux antérieurs).

Le premier objectif de ce volet est de poursuivre l'étude générale de la stabilité algébrique des mots de retour, que j'ai déjà entâmée dans une collaboration récente (voir la section 3 du rapport sur mes travaux antérieurs). Dans le cadre de cette collaboration, nous avons établi la réciproque du Théorème du retour, ce qui donne une caractérisation algébrique de la dendricité. On peut reformuler cette caractérisation en terme de stabilité : un shift est dendrique exactement si tous ses ensembles de retour sont stables dans le groupe libre et ont la même cardinalité.

Problème 6. *Utiliser les conditions de stabilité algébrique pour caractériser d'autres familles de shifts, comme les shifts automatiques ou les shifts suffixes-connexes. Inversement, donner une interprétation combinatoire à différentes propriétés de stabilité algébrique.*

Le deuxième objectif du dernier volet est de développer des outils pour le calcul des facteurs bispéciaux. Un facteur u dans un shift X est spécial à gauche (ou à droite) s'il existe au moins deux lettres distinctes telles que au (ou ua) est facteur de X. Un facteur qui à la fois spécial à droite et à gauche est dit bispécial. L'étude des facteurs bispéciaux est lié aux mots de retour et aux graphes d'extensions. On peut observer par exemple que le graphe d'extension d'un facteur qui n'est pas bispécial est nécessairement un arbre (avec une ou deux branches) et donc les facteurs bispéciaux déterminent la dendricité.

Dans un article publié en 2012, K. Klouda [41] décrit un algorithme permettant de calculer les facteurs bispéciaux dans les systèmes D0L (que l'on peut voir comme une généralisation de la notion de shift substitutif). En collaboration avec K. Klouda et Š Starosta, nous avons commencer à implémenter cet algorithme dans le logiciel SageMath [1]. Nos tentatives d'implémentations ont révélées plusieurs problèmes de conception dans l'algorithme, de même que plusieurs pistes pour le généraliser.

Problème 7. *Corriger, généraliser, et implémenter l'algorithme de Klouda pour calculer les facteurs bispéciaux.*

Un tel algorithme pourra servir d'outil d'exploration et l'experimentation aux deux autres volets, en particulier en permettant d'obtenir de meilleures descriptions pour les ensembles de mots de retour dans le cadre substitutif.

Problème 8. *Utiliser l'algorithme de Klouda pour développer une approche systématiques au calculs des ensembles de mots de retour dans les shifts substitutifs.*

Parmi les autres retombées potentielles, mentionnons le calcul algorithmique de formules exactes pour les complexités en facteurs des shifts subsitutifs (via une formule de Cassaigne [18]), ainsi que l'étude des exposants critiques (grâce aux travaux de Dolce et al. [23]).

Affectations

IRIF. L'Institut de recherche en informatique fondamentale (IRIF) constitue un choix naturel pour poursuivre mes recherches. L'équipe Automates et applications s'intéresse à plusieurs thèmes couverts par ce projet : la théorie des automates et sa connexion avec les semigroupes finis, ainsi que la combinatoire des mots. Le projet est aussi connecté aux thèmes de l'équipe Combinatoire.

1. https://www.sagemath.org/

Les deux équipes comptent parmi leurs membres V. Berthé et W. Steiner, experts en dynamiques symboliques et combinatoire des mots, et avec qui je pourrai travailler sur de nombreuses questions liées au projet. En particulier, V. Berthé, avec qui j'ai d'ailleurs déjà collaboré, possède une solide expertise en théorie ergodique et sur les systèmes S-adiques. Nos collaborations en cours me confirment qu'elle est intéressée à développer une collaboration suivie sur différents thèmes, tant en lien avec le projet qu'avec ses propres intérêts de recherche. W. Steiner s'intéresse quant à lui aux systèmes de numération et aux pavages en lien avec la dynamique symbolique, ce qui m'offrira la possibilité d'enrichir les thématiques de mon projet à moyen ou long terme.

Le volet profini du projet est aussi susceptible d'intéresser O. Carton, spécialiste de la théorie des automates et de la théorie des semigroupes, de même que S. van Gool, qui travaille sur les applications des méthodes profinies en logique et en informatique théorique – ce qui ouvre encore d'autres possibilités. On peut mentionner de plus T. Colcombet, qui s'intéresse lui aussi à la théorie des automates et avec qui des collaborations seraient possibles.

LABRI. Mon projet de recherche touche à plusieurs thèmes étudiés au LaBRI (Laboratoire bordelais de recherche en informatique), autant au sein du département CombAlgo (Combinatoire et algorithmique), plus particulièrement dans l'équipe Combinatoire et interactions, qu'au département M2F (Méthodes et modèles formels), avec l'équipe Fondements logiques du calcul. Le département CombAlgo s'intéresse de manière large aux objets discrets, avec une affinité pour l'algèbre et la théorie des probabilités ; les thèmes du département M2F incluent les fondements des méthodes formelles, et en particulier les automates.

L'équipe Combinatoire et Interactions compte parmi ses membres plusieurs chercheurs dont les intérêts s'alignent avec les thèmes couverts par ce projet. Je pense en particulier à S. Labbé, V. Delecroix et P. Narbel, qui s'intéressent tous trois à la dynamique symbolique et à la combinatoire des mots. V. Delecroix possède pour sa part une expertise particulière sur les représentations S-adiques et les mots de retour, ce qui offrirait des occasions d'approfondir ces aspects du projet en les connectant avec d'autres questions ; collaborer avec S. Labbé permettrait des ouvertures vers la géométrie discrète, en particulier en lien avec les pavages, la géométrie fractale et la numération (via les fractions continues multi-dimensionnelles).

Mentionnons enfin quelques membres de l'équipe *Fondements logiques du calcul* susceptibles de collaborations à court ou moyen terme : S. Lombardy en dynamique symbolique et combinatoire des mots, T. Place en théorie des automates, et M. Zeitoun pour la théorie des automates et les langages formels avec des approches topologiques et algébriques.

LIGM. Mon projet s'inscrit en plein dans les thématiques étudiées par plusieurs membres de l'équipe *Bases de données, automates, analyses d'algorithmes et modèles* (BAAM) au Laboratoire d'informatique Gaspard-Monge (LIGM).

Parmi les membres de l'équipe, mentionnons tout d'abord D. Perrin, avec qui j'ai déjà collaboré, qui est spécialiste des semigroupes profinis, de la dynamique symbolique et de la combinatoire des mots. M.-P. Béal s'intéresse aussi à la dynamique symbolique, la combinatoire des mots et la théorie des automates en lien notamment avec la théorie des codes. Mes échanges récents avec elle me confirme l'intérêt qu'elle porte à mon candidature et nous entrevoyons déjà des collaborations autour de nos intérêts communs.

Parmi d'autres collaborations possibles, citons P. Rotondo, qui s'intéresse à la dynamique symbolique et la combinatoire des mots (notamment les mots sturmiens) en en lien avec des questions probabilistes, de même que C. Nicaud, avec une expertise en théorie des automates et en théorie combinatoire des groupes (avec un aspect probabiliste). Collaborer avec ce dernier serait susceptible d'enrichir mon projet, en particulier sur des thèmes liés à la théorie des groupes.

Mentionnons aussi deux autres membres ayant travaillé sur des sujets proches des thèmes de ce projet : V. Marsault (automates, numération, logique) et V. Jugé (théorie des semigroupes, combinatoire des mots). En particulier, d'éventuelles collaborations avec V. Jugé pourraient permettre des ouvertures thématiques sur d'autres sujets, comme les monoïdes de traces ou les groupes de tresses. Enfin des collaborations avec G. Kucherov sur les aspects algorithmiques de la combinatoire des mots offrent la possibilité d'une ouverture vers une seconde équipe (*Algorithmique Discrète et Applications*).

Références

[1] J.-P. Allouche et J. Shallit. « The ubiquitous Prouhet-Thue-Morse sequence ». Dans : *Sequences and their Applications*. Sous la dir. de C. Ding, T. Helleseth et H. Niederreiter. Springer London, 1999, p. 1-16. doi : 10.1007/978-1-4471-0551-0_1.

[2] J. Almeida. « Dynamics of implicit operations and tameness of pseudovarieties of groups ». *Trans. Amer. Math. Soc.* 354.1 (2001), p. 387-411. doi : 10.1090/s0002-9947-01-02857-4.

[3] J. Almeida. « Profinite groups associated with weakly primitive substitutions ». *J. Math. Sci.* 144.2 (2007), p. 3881-3903. doi : 10.1007/s10958-007-0242-y.

[4] J. Almeida. « Symbolic dynamics in free profinite semigroups ». Dans : *Algebraic Systems, Formal Languages and, Conventional and Unconventional Computation Theory*. Sous la dir. de M. Ito. T. 1366. RIMS Kôkyûroku Bessatsu. 2004, p. 1-12. url : http://hdl.handle.net/2433/25356.

[5] J. Almeida et A. Costa. « A geometric interpretation of the Schützenberger group of a minimal subshift ». *Ark. Math.* 54.2 (2016), p. 243-275. doi : 10.1007/s11512-016-0233-7. arXiv : 1507.06885.

[6] J. Almeida et A. Costa. « Infinite-vertex free profinite semigroupoids and symbolic dynamics ». *J. Pure Appl. Algebra* 213.5 (2009), p. 605-631. doi : 10.1016/j.jpaa.2008.08.009.

[7] J. Almeida et A. Costa. « Presentations of Schützenberger groups of minimal subshifts ». *Israel J. Math.* 196.1 (2013), p. 1-31. doi : 10.1007/s11856-012-0139-4. arXiv : 1001.1475.

[8] J. Almeida, A. Costa, R. Kyriakoglou et D. Perrin. « On the group of a rational maximal bifix code ». *Forum Math.* 32.3 (2020), p. 553-576. doi : 10.1515/forum-2018-0270. arXiv : 1811.03185.

[9] J. Almeida, A. Costa, R. Kyriakoglou et D. Perrin. *Profinite semigroups and symbolic dynamics*. Springer International Publishing, 2020. doi : 10.1007/978-3-030-55215-2.

[10] L. Balková, E. Pelantová et W. Steiner. « Sequences with Constant Number of Return Words ». *Monatsh. Math.* 155.3-4 (2008), p. 251-263. doi : 10.1007/s00605-008-0001-2.

[11] M.-P. Béal, D. Perrin, A. Restivo et W. Steiner. *Recognizability in S-adic shifts*. Preprint. 2023. doi : 10.48550/arxiv.2302.06258. arXiv : 2302.06258.

[12] J. Berstel. « Sur la densité asymptotique de langages formels ». Dans : *Automata Languages and Programming*. Sous la dir. de M. Nivat. North-Holland, 1972, p. 345-358.

[13] J. Berstel, D. Perrin et C. Reutenauer. *Codes and Automata*. Cambridge University Press, 2009. doi : 10.1017/cbo9781139195768.

[14] V. Berthé, H. Goulet-Ouellet, C.-F. Nyberg-Brodda, D. Perrin et K. Petersen. *Density of group languages in shift spaces*. Preprint. 2024. doi : 10.48550/arxiv.2403.17892. arXiv : 2403.17892.

[15] V. Berthé, W. Steiner, J. M. Thuswaldner et R. Yassawi. « Recognizability for sequences of morphisms ». *Ergodic Theory Dynam. Systems* 39.11 (2019), p. 2896-2931. doi : 10.1017/etds.2017.144. arXiv : 1705.00167.

[16] V. Berthé et al. « Maximal Bifix Decoding ». *Discrete Math.* 338.5 (2015), p. 725-742. doi : 10.1016/j.disc.2014.12.010.

[17] M. Bodirsky, T. Gärtner, T. von Oertzen et J. Schwinghammer. « Efficiently Computing the Density of Regular Languages ». Dans : *LATIN 2004 : Theoretical Informatics*. Springer Berlin Heidelberg, 2004, p. 262-270. doi : 10.1007/978-3-540-24698-5_30.

[18] J. Cassaigne. « Complexité et facteurs spéciaux ». *Bull. Bel. Math. Soc. Simon Stevin* 4.1 (1997), p. 67-88. doi : 10.36045/bbms/1105730624.

[19] A. Costa. « A profinite approach to complete bifix decodings of recurrent languages ». *Forum Math.* 35.4 (2023), p. 1021-1045. doi : 10.1515/forum-2022-0246. arXiv : 2208.11768.

[20] A. Costa et B. Steinberg. « Profinite groups associated to sofic shifts are free ». *Proc. London Math. Soc.* 102.3 (2011), p. 341-369. doi : 10.1112/plms/pdq024. arXiv : 0908.0439.

[21] A. Costa et B. Steinberg. « The Karoubi envelope of the mirage of a subshift ». *Comm. Algebra* 49 (2021), p. 4820-4856. doi : 10.1080/00927872.2021.1931263. arXiv : 2005.07490.

[22] M. De Berardinis et S. Ghilardi. « Profiniteness, monadicity and universal models in modal logic ». *Annals of Pure and Applied Logic* 175.7 (2024), p. 103454. doi : 10.1016/j.apal.2024.103454. arXiv : 2305.04592.

[23] F. Dolce, L. Dvořáková et E. Pelantová. « On balanced sequences and their critical exponent ». *Theor. Comput. Sci.* 939 (2023), p. 18-47. doi : 10.1016/j.tcs.2022.10.014. arXiv : 2108.07503.

[24] F. Durand. « A characterization of substitutive sequences using return words ». *Discrete Math.* 179.1-3 (1998), p. 89-101. doi : 10.1016/S0012-365X(97)00029-0. arXiv : 0807.3322.

[25] F. Durand. « Decidability of the HD0L Ultimate Periodicity Problem ». *RAIRO Theor. Inform. Appl.* 47.2 (2013), p. 201-214. doi : 10.1051/ita/2013035.

[26] F. Durand. « Decidability of Uniform Recurrence of Morphic Sequences ». *Int. J. Found. Comput.* 24.01 (2013), p. 123-146. doi : 10.1142/s0129054113500032.

[27] F. Durand, B. Host et C. Skau. « Substitutional Dynamical Systems, Bratteli Diagrams and Dimension Groups ». *Ergodic Theory Dynam. Systems* 19.4 (1999), p. 953-993. doi : 10.1017/S0143385799133947. arXiv : 0807.3621.

[28] F. Durand et J. Leroy. « Decidability of the isomorphism and the factorization between minimal substitution subshifts ». *Discrete Anal.* 2022.7 (2022), p. 1-66. doi : 10.19086/da.36901. arXiv : 1806.04891.

[29] B. Espinoza. *The structure of low complexity subshifts*. Preprint. 2023. doi : 10.48550/arXiv.2305.03096. arXiv : 2305.03096.

[30] N. P. Fogg. *Substitutions in Dynamics, Arithmetics and Combinatorics*. Sous la dir. de V. Berthé, S. Ferenczi, C. Mauduit et A. Siegel. Springer Berlin Heidelberg, 2002. doi : 10.1007/b13861.

[31] R. H. Fox. « Free Differential Calculus I : Derivation in the Free Group Ring ». *Ann. Math.* 57.3 (1953), p. 547-560. doi : 10.2307/1969736.

[32] M. D. Fried et M. Jarden. *Field Arithmetic*. Springer Berlin Heidelberg, 2008. doi : 10.1007/978-3-540-77270-5.

[33] F. Gheeraert, M. Lejeune et J. Leroy. « S-adic characterization of minimal ternary dendric shifts ». *Ergodic Theory Dynam. Systems* 42.11 (2022), p. 3393-3432. doi : https://doi.org/10.1017/etds.2021.84. arXiv : 2102.10092.

[34] F. Gheeraert et J. Leroy. *S-adic characterization of minimal dendric shifts.* Arxiv preprint. 2022. doi : 10.48550/arxiv.2206.00333. arXiv : 2206.00333.

[35] A. Glen, J. Justin, S. Widmer et L. Q. Zamboni. « Palindromic richness ». *European J. Combin.* 30.2 (2009), p. 510-531. doi : 10.1016/j.ejc.2008.04.006.

[36] S. J. van Gool et B. Steinberg. « Pro-aperiodic monoids via saturated models ». *Israel J. Math.* 234.1 (2019), p. 451-498. doi : 10.1007/s11856-019-1947-6. arXiv : 1609.07736.

[37] H. Goulet-Ouellet. « Freeness of Schützenberger groups of primitive substitutions ». *Int. J. Algebra Comput.* 32.06 (2022), p. 1101-1123. doi : 10.1142/S0218196722500473. arXiv : 2109.11957v1.

[38] H. Goulet-Ouellet. « Pronilpotent quotients associated with primitive substitutions ». *J. Algebra* 606 (2022), p. 341-370. doi : 10.1016/j.jalgebra.2022.05.021. arXiv : 2204.05706v1.

[39] G. Hansel et D. Perrin. « Codes and Bernoulli partitions ». *Math. Syst. Theory* 16.1 (1983), p. 133-157. doi : 10.1007/BF01744574.

[40] G. Hansel et D. Perrin. « Rational probability measures ». *Theor. Comput. Sci.* 65.2 (1989), p. 171-188. doi : 10.1016/0304-3975(89)90042-x.

[41] K. Klouda. « Bispecial factors in circular non-pushy D0L languages ». *Theor. Comput. Sci.* 445 (2012), p. 63-74. doi : 10.1016/j.tcs.2012.05.007.

[42] T. Koga. « On the Density of Regular Languages ». *Fundamenta Informaticae* 168.1 (2019), p. 45-49. doi : 10.3233/fi-2019-1823.

[43] J. Kozik. « Conditional densities of regular languages ». *Electron. Notes Theor. Comput. Sci.* 140 (2005), p. 67-79. doi : 10.1016/j.entcs.2005.06.023.

[44] J. F. Lynch. « Convergence laws for random words ». *Australas. J. Combin.* 7 (1993), p. 145-156.

[45] W. Magnus. « On a theorem of Marshall Hall ». *Ann. Math.* 40.4 (1939), p. 764-768. doi : 10.2307/1968892.

[46] C. Marion, P. V. Silva et G. Tracey. « The pro-k-solvable topology on a free group ». *J. Aust. Math. Soc.* 116.3 (2024), p. 363-383. doi : 10.1017/s1446788723000162. arXiv : 2304.10235.

[47] P. Michel. « Stricte ergodicité d'ensembles minimaux de substitution ». *C. R. Acad. Sci.* 278 (1974), p. 811-813.

[48] M. Morse et G. A. Hedlund. « Symbolic dynamics ». *Amer. J. Math.* 60.4 (1938), p. 815-866. doi : 10.2307/2371264.

[49] M. Morse et G. A. Hedlund. « Symbolic dynamics II : Sturmian trajectories ». *Amer. J. Math.* 62.1 (1940), p. 1-42. doi : 10.2307/2371431.

[50] J.-E. Pin. « Profinite Methods in Automata Theory ». Dans : *26th International Symposium on Theoretical Aspects of Computer Science*. Sous la dir. de S. Albers et J.-Y. Marion. Freiburg, Germany, 2009, p. 31-50. doi : 10.4230/LIPIcs.STACS.2009.1856. hal : inria-00359677.

[51] A. Salomaa et M. Soittola. *Automata-Theoretic Aspects of Formal Power Series.* Springer New York, 1978. doi : 10.1007/978-1-4612-6264-0.

[52] J.-P. Serre. *Cohomologie Galoisienne.* Springer Berlin Heidelberg, 1994. doi : 10.1007/bfb0108758.

[53] R. Sin'ya. « An Automata Theoretic Approach to the Zero-One Law for Regular Languages : Algorithmic and Logical Aspects ». *Electron. Proc. Theor. Comput. Sci.* 193 (2015), p. 172-185. doi : 10.4204/eptcs.193.13. arXiv : 1509.07209.

[54] L. Vuillon. « A Characterization of Sturmian Words by Return Words ». *European J. Combin.* 22.2 (2001), p. 263-275. doi : 10.1006/eujc.2000.0444.