Questions de localisabilité pour le calcul distribué

Ghazal Kachigar

THÈSE DE L'UNIVERSITÉ DE BORDEAUX

École Doctorale Mathématiques et Informatique

QUESTIONS DE LOCALISABILITÉ POUR LE CALCUL DISTRIBUÉ

Présentée par

Ghazal KACHIGAR

Le 10/12/2019

Sous la direction de

Cyril GAVOILLE & Gilles ZÉMOR

Pour obtenir le grade de

Docteur de l'Université de Bordeaux

Spécialité

Mathématiques

Devant le jury composé de

Cyril GAVOILLE	Professeur - LaBRI (Bordeaux)	Directeur de thèse
Frédéric MAGNIEZ	Directeur de recherche - IRIF (Paris Diderot)	Rapporteur
Jean-François MARCKERT	Directeur de recherche - LaBRI (Bordeaux)	Examinateur
Ion NECHITA	Chargé de recherche - LPT (Toulouse)	Examinateur
Simon PERDRIX	Chargé de recherche - LORIA (Nancy)	Rapporteur
Gilles ZÉMOR	Professeur - IMB (Bordeaux)	Directeur de thèse

Institut de Mathématiques de Bordeaux

LaBRI

Résumé

Cette thèse suit un plan à deux parties. Le point de départ en est la notion de *résistance à la localisation*, qui est importante en calcul distribué quantique.

Dans la première partie, qui est plutôt théorique, nous retraçons l'historique de certaines notions et résultats en information quantique et en calcul distribué, plus précisément le phénomène d'*intrication* et la *condition non-signalling* en information quantique et le *modèle LOCAL* et le *problème de coloration* en calcul distribué. Ensuite, nous évoquons le modèle ϕ-*LOCAL*, développé en 2009 comme adaptation de la condition non-signalling au modèle LOCAL dans le but d'étudier l'existence d'algorithmes distribués quantiques. Finalement, nous soulignons quelques limites du modèle ϕ-LOCAL à l'aide des notions de consistance globale et de consistance locale, et nous présentons une version plus adéquate de ce modèle.

La deuxième partie comporte les principaux résultats techniques obtenus au cours de cette thèse dans le domaine de la théorie des probabilités. Nous introduisons la notion de k-*localisabilité* qui est une traduction probabiliste du modèle ϕ-LOCAL. Nous montrons en quoi cette notion est proche, mais plus faible, que la notion de k-dépendance, largement étudiée dans la littérature probabiliste. Nous évoquons des résultats récents autour de la coloration 1-dépendante du chemin qui permettent de conclure au sujet de la coloration 1-localisable du chemin : elle est possible dès qu'il y a plus de quatre couleurs. Dans la suite, nous traitons la question de la possibilité de la coloration 1-localisable du chemin à l'aide de trois couleurs : nous verrons qu'elle n'est pas possible. Pour répondre à cette question, nous avons eu recours à la programmation linéaire et à la combinatoire : en particulier, nous démontrons un théorème qui donne la solution explicite d'un programme linéaire ayant une forme particulière, ainsi qu'une formule pour les nombres de Catalan.

Abstract

This thesis is divided in two parts. Its starting point is the concept of *resistance to localisation*, an important concept in distributed quantum computing.

In the first, theoretical part of this thesis, we go over the history of certain concepts and results in quantum information theory and distributed computing, such as the phenomenon of *entanglement* and the *non-signalling condition* in the first domain, and the *LOCAL model* and *the colouring problem* in the second domain. We then focus on the ϕ-*LOCAL* model, whose goal is to study the possibility of quantum distributed algorithms, and which was developed in 2009 by adapting the non-signalling condition to the LOCAL model. We introduce the concepts of global and local consistency in order to emphasise some shortcomings of this model. Finally, we present a more adequate version of the ϕ-LOCAL model.

The second part of this thesis contains our major technical results in probability theory. We define the concept of k-localisability which is a probabilistic translation of the ϕ-LOCAL model. We show that this concept is close to but weaker than the concept of k-dependence which is well-studied in the probabilistic literature. We mention recent results concerning 1-dependent colouring of the path graph and the conclusion they allow us to reach with regards to 1-localisable colouring of the path graph : that it is possible with four or more colours. The rest of this part is dedicated to answering the question of the possibility of 1-localisable colouring of the path graph using three colours which we will show to be impossible. In answering this question we have made use of methods in linear programming and combinatorics. In particular, we prove a theorem on the explicit solution of a linear programming problem having a certain form, and a formula for the Catalan numbers.

Remerciements

Je souhaiterais tout d'abord remercier mes directeurs de thèse pour leur disponibilité, leur patience et leurs conseils tant scientifiques que pratiques.

Je remercie également toutes les personnes qui ont pris le temps de discuter avec moi, me conseiller et répondre à mes questions, que ce soit lors d'entretiens en petits groupes ou lors d'une grande manifestation scientifique.

Je remercie enfin les membres du jury d'avoir accepté de lire et de critiquer le présent manuscrit dans le but de l'améliorer.

Table des matières

Introduction

Les travaux menés au cours de cette thèse ont mené à des réflexions et des résultats de nature différente, conceptuels et théoriques pour certains et techniques et calculatoires pour d'autres, ce qui nous a incité à présenter nos contributions en suivant un plan à deux parties.

La première partie est plutôt théorique. Nous y retraçons l'histoire de certaines idées et résultats en information quantique et en calcul distribué, ainsi que leur « mariage » dans l'article [GKM09], qui ont été l'inspiration pour les résultats plus techniques présentés dans la deuxième partie, et nous y présentons quelques contributions de nature théorique et numérique à ce domaine d'étude.

Le Chapitre 1 est consacré à l'information quantique. Nous commençons par retracer les questionnements autour des fondements de la mécanique quantique qui ont abouti au *théorème de Bell* [Bel64] sur les particularités des résultats statistiques issus d'une mesure quantique, et sa vérification expérimentale par le groupe mené par Aspect. Ce résultat, qui a permis de démarquer la mécanique quantique des théories classiques et relativistes via les particularités de l'*intrication quantique*, peut être présenté simplement sous forme de distributions de probabilité présentant diverses formes de corrélations, associées à un « jeu » à deux joueurs, le *jeu CHSH*. Après avoir présenté ce jeu, nous tournerons notre attention vers l'étude de ces diverses distributions de probabilités : celles issues de l'utilisation de ressources classiques, celles issues de l'utilisation de ressources quantiques, et des distributions supraquantiques vérifiant une certaine propriété, celle d'être *non-signalling*, ainsi que l'intérêt de ces dernières pour étudier de manière plus simple l'existence ou non d'avantages quantiques pour la résolution de certains problèmes.

Dans le Chapitre 2, nous commençons par présenter un bref historique du calcul distribué, avant de nous focaliser sur le problème qui a inspiré les résultats techniques présentés dans la deuxième partie de cette thèse, le problème de *coloration de graphe*. Nous présenterons l'algorithme de Cole et Vishkin [CV86] pour la coloration distribuée de graphes, et nous présenterons une preuve théorique de Linial [Lin87, Lin92] qui montre plus ou moins que

l'algorithme de Cole et Vishkin est optimal. Notre présentation de cette preuve fait usage d'un cadre conceptuel neuf qui, comme nous le verrons dans le Chapitre 4, subsume nombre d'autres concepts et résultats. Nous expliquerons ensuite en quoi la preuve de Linial ne passe plus dans un cadre quantique, en illustrant nos propos à l'aide d'un jeu quantique, le carré magique de Mermin. Cela nous amène à nous interroger sur la possibilité de battre l'algorithme de Cole et Vishkin en ayant recours à des ressources quantiques. Dès lors, plusieurs pistes de recherche s'ouvrent à nous : soit d'essayer de trouver un algorithme quantique qui le bat, soit d'essayer de trouver des arguments pour montrer s'il est possible ou non de le battre.

Le Chapitre 3 est consacré au modèle ϕ-*LOCAL*, développé dans [GKM09], et qui tente de répondre au problème ouvert mentionné dans le paragraphe précédent en suivant la deuxième voie, c'est-à-dire étudier la possibilité de faire mieux que l'algorithme de Cole et Vishkin en ayant accès à des ressources quantiques. Ce modèle fait cela en mariant des notions issues du calcul distribué et la propriété *non-signalling* étudiée dans le Chapitre 1, dont nous avons mentionné l'intérêt pour l'étude de l'existence ou non d'avantage quantique pour la résolution de certains problèmes. Nous étendons ensuite l'argument utilisé dans l'article en question afin d'étudier le problème de 2-coloration sur le cycle pair au problème de 2-coloration du chemin. Mais le modèle ϕ-LOCAL tel que présenté dans [GKM09] souffre d'un défaut : il s'avère qu'il existe une variante plus faible de ce modèle qui permet de répondre aux mêmes questions. Les notions de *consistance globale* et *consistance locale* sont introduites afin d'expliquer ce défaut et définir cette variante plus faible.

Dans le Chapitre 4, nous présentons quelques résultats associés que nous ne pouvions pas mettre dans les chapitres précédents sans casser le rythme du texte. Nous revenons ainsi sur le cadre conceptuel utilisé pour expliquer la preuve de Linial dans le Chapitre 2 et nous montrons qu'il permet également de subsumer la notion de *nombre chromatique quantique* [CMN$^+$06]. Nous présentons également une étude numérique menée sur un exemple simple dans ce même cadre conceptuel.

Nous passons ensuite à la deuxième partie de cette thèse, où des résultats plus techniques dans le domaine des probabilités et de la combinatoire sont présentés.

Dans le Chapitre 5, nous commençons par présenter les notions de k-*dépendance* et de *block factor* issues de la théorie des processus stochastiques, et nous expliquons en quoi elles transposent des notions que nous avions vues dans le domaine du calcul distribué au domaine des processus stochastiques : par exemple, la notion de block factor traduit parfaitement la calculabilité distribuée, et la notion de k-dépendance implique la calculabilité dans le modèle ϕ-LOCAL. Nous présentons ensuite une nouvelle notion, celle de la

k-*localisabilité*, qui traduit plus fidèlement la notion de calculabilité dans le modèle ϕ-LOCAL que la notion de k-dépendance puisqu'elle est plus faible. Nous présentons ensuite des résultats obtenus par Holroyd et Liggett autour de la coloration k-dépendante [HL15, HL16, Hol17] : construction de processus de q-coloration 1-dépendante pour $q \geq 4$ et de 3-coloration 2-dépendante, et l'impossibilité de la 3-coloration 1-dépendante. Nous verrons en particulier que l'étude d'un type de processus induit par la coloration, les processus *hard-core*, a été centrale pour démontrer l'impossibilité de la 3-coloration 1-dépendante. La k-dépendance impliquant la k-localisabilité, les résultats de Holroyd et Liggett permettent de clore la question pour la q-coloration 1-localisable et la 3-coloration 2-localisable. Mais la question de l'existence d'une 3-coloration 1-localisable reste ouverte : pour l'étudier, nous avons adopté la même stratégie que Holroyd et Liggett et nous avons étudié les processus hard-core 1-localisables.

Le Chapitre 6 présente les résultats de cette étude. Pour se familiariser avec le problème, nous commençons par l'étudier lorsqu'il y a un petit nombre fini de variables aléatoires : nous trouvons que l'existence d'un processus hard-core 1-localisable est équivalent à la résolubilité d'un système d'équations présentant certaines caractéristiques. Plus précisément, nous nous intéressons à la valeur maximale prise par les variables sous ces contraintes : il s'agit donc d'un problème de *programmation linéaire*. Après avoir développé une intuition pour le problème grâce à une étude sur de petits exemples, nous dérivons le problème de programmation linéaire à résoudre dans le cas général, et nous constatons que nous pouvons nous limiter à un sous-ensemble des contraintes présentant une régularité remarquable. Il s'avère que de tels problèmes admettent une solution dont nous pouvons donner une forme close : c'est un résultat que nous prouvons par la suite en nous servant du *théorème de dualité* en programmation linéaire. Finalement, lorsque nous appliquons ce résultat au problème des processus hard-core 1-localisables, la suite des *nombres de Catalan* fait son apparition sous forme d'une formule qui sera prouvée dans le Chapitre 7. Grâce à cette formule, nous démontrons que la 3-coloration 1-localisable n'est pas possible si le nombre de variables aléatoires est suffisamment élevé. Les résultats présentés jusqu'ici ont fait l'objet d'une publication dans [GKZ19]. La suite de ce chapitre est dédié à la généralisation de ces résultats : premièrement, à des processus « k-hard-core » k-localisables (liés au problème de *coloration à distance* comme les processus hard-core étaient liés au problème de coloration), où nous voyons la suite des *nombres de Fuss-Catalan* apparaître ; et deuxièmement, à des variables aléatoires indexées par les sommets d'un graphe quelconque.

Enfin, le Chapitre 7 est consacré à la preuve de la formule apparaissant dans le Chapitre 6 pour la suite des nombres de Catalan et les nombres de Fuss-Catalan. Nous démontrons que, pour les nombres de Catalan, il s'agit d'un cas

particulier d'une formule valable pour la *convolution k-ème des nombres de Catalan*. Quant au nombres de Fuss-Catalan, il s'agit d'une question ouverte de savoir si la formule est un cas particulier d'une formule plus générale. Les preuves dans les deux cas sont très similaires et utilisent l'interprétation de ces suites de nombres comme le nombre d'un certain type de chemin sur la grille, ainsi que le *principe d'inclusion-exclusion*.

Première partie

Calcul quantique, calcul distribué

CHAPITRE 1
De l'intrication à la non-localité

1.1 L'intrication quantique et le théorème de Bell

Le paradoxe EPR. Dans les années qui suivent la naissance de la mécanique quantique, les physiciens mettent en évidence plusieurs aspects contre-intuitifs de cette discipline, tel le problème de la mesure quantique illustré grâce au paradoxe du chat de Schrödinger [Sch35] en 1935. Un autre problème, soulevé la même année par Einstein, Podolsky et Rosen [EPR35], concerne le phénomène d'intrication quantique qui, semblerait-il, permettrait de transmettre de l'information à distance instantanément, contredisant ainsi une conséquence fondamentale de la théorie de la relativité, à savoir qu'aucun signal ne peut être transmis plus rapidement qu'à la vitesse de la lumière. Cependant, ce paradoxe n'a pas fait l'objet de beaucoup d'attention jusqu'à la publication en 1964 par John Bell d'un article en donnant un résultat expérimentalement vérifiable permettant de trancher la question [Bel64], et une mise en pratique expérimentale de ce résultat en 1981 par le groupe mené par Alain Aspect [AGR81].

Expliquons en quoi consiste le **paradoxe Einstein-Podolsky-Rosen (EPR)** : si l'on peut prédire avec certitude, sans perturber un système, la valeur d'une propriété mesurable de ce système à un point donné de l'espace-temps, il est naturel de considérer que cette propriété mesurable a une valeur intrinsèque (« il existe un élément de réalité physique correspondant à cette propriété », pour reprendre l'expression utilisée par Einstein, Podolsky et Rosen). Ce principe est appelé le **réalisme** dans la littérature. Un autre principe fondamental (qui est une conséquence de la théorie de la relativité), la **localité**, dit que deux événements éloignés dans l'espace (*space-like separated* en anglais) ne peuvent pas s'influencer l'un l'autre. Or, en mécanique quantique il est possible de lier des particules en les **intriquant**. Ce faisant, les états des particules seront reliés, et elles seront pour ainsi dire dans un état global sans que l'on puisse définir séparément des états locaux pour chacune d'entre elles. Il est ainsi possible de faire en sorte qu'en mesurant une propriété d'une des particules (par exemple, sa position) nous pourrons inférer avec certitude la valeur de cette propriété pour l'autre particule. Un aspect remarquable de l'intrication est qu'elle subsiste même si l'on sépare les deux particules de plusieurs années-lumière (garantissant ainsi qu'elles seront éloignées dans l'espace, ou *space-like separated*). Dès lors, en effectuant une mesure sur une particule, on obtiendra une valeur bien précise pour sa position. De plus, si l'on

mesurait l'autre particule à plusieurs années-lumière de la première particule, on obtiendrait une valeur pour la position qui dépend de la valeur prise par la première particule. Dès lors, si deux particules dans un état intriqué sont données chacune à deux individus, Alice et Bob, à plusieurs années-lumière l'un de l'autre, et qui ne connaissent pas l'état du système, alors [Wis06] :

- Si le réalisme et la localité tiennent, alors Alice peut déduire avec certitude la position de la particule de Bob en mesurant la position de la sienne. Or, cette mesure ne perturbe pas la particule de Bob, par conséquent, par le principe de réalisme, la position de la particule de Bob a une valeur intrinsèque. Le même argument vaut pour la quantité de mouvement de la particule de Bob : Alice peut la déduire avec certitude en mesurant la quantité de mouvement de sa particule et ceci ne perturbe pas la particule de Bob, donc la quantité de mouvement a également une valeur intrinsèque. Donc il est possible de déterminer avec certitude et la position, et la quantité de mouvement de la particule de Bob, ce qui contredit le **principe d'incertitude de Heisenberg** [1]. Donc la mécanique quantique dans sa formulation d'origine serait « incomplète » (ce qui ne signifie pas qu'elle serait incohérente). C'est la conclusion d'Einstein, Podolsky et Rosen, qui proposent d'ajouter des « variables cachées » correspondant aux positions et aux quantités de mouvement des particules afin de « compléter » la mécanique quantique.
- Si le réalisme tient, et que la mécanique quantique dans sa formulation d'origine est « complète », le fait qu'Alice puisse prédire avec certitude la position de la particule de Bob après avoir mesuré celle de sa particule implique que de l'information au sujet de la particule détenue par Bob a été transmise instantanément via cette mesure à Alice. La localité serait donc violée, ce qui n'est pas possible puisqu'elle est une conséquence de la théorie de la relativité, et donc, si elle est violée, la théorie de la relativité l'est aussi.
- Si la localité tient, et que la mécanique quantique dans sa formulation d'origine est « complète », alors, par le principe de localité, le fait qu'Alice mesure sa particule ne va pas perturber celle de Bob, puisque les deux particules sont à plusieurs années-lumière l'une de l'autre. Néanmoins, la position de la particule de Bob dépendrait alors de la mesure faite par Alice, et n'aurait donc pas une valeur intrinsèque. Einstein, Podolsky et

1. Le principe d'incertitude repose sur le fait que la mesure de la position et celle de la quantité de mouvement ne commutent pas, i.e. l'ordre des mesures va influer sur le résultat obtenu. On dit qu'elles ne sont pas simultanément mesurables. Or, notre argumentation repose sur une implication de type $A(p) \rightarrow B(p)$, où $A(p)$ dit que l'on peut déterminer avec certitude la valeur de la propriété p, et $B(p)$ dit que p a une valeur intrinsèque. Or, le fait que la position et la quantité de mouvement ne soient pas simultanément mesurables n'empêche pas que l'antécédent $A(p)$ soit vrai individuellement pour chacune d'entre elles.

Rosen n'admettent pas cette troisième alternative sans pour autant la réfuter avec un argument rigoureux.

Le théorème de Bell. Dans son article de 1964, John Bell propose un théorème sur les résultats statistiques d'une expérience qui seraient différentes selon que l'on suppose que les axiomes de la mécanique quantique sans « variables cachées » sont vraies, ou selon que l'on parte d'une théorie admettant des « variables cachées », montrant ainsi que la mécanique quantique est incompatible avec l'existence de variables cachées [2]. Autrement dit, il existe des distributions de probabilités que l'on peut obtenir en effectuant certaines mesures sur certains états quantiques si l'on en reste à la formulation originale de la mécanique quantique, mais qui ne peuvent plus s'obtenir dès lors que l'on fait l'hypothèse de l'existence de variables cachées.

Pour résumer, Einstein, Podolsky et Rosen avaient identifié et étudié trois alternatives dans leur article, dont une et seulement une peut être vraie :

1. **La localité et le réalisme tiennent, donc il y a des variables cachées.**

   Il s'agit de l'alternative préconisée par les auteurs de l'article.

2. **Le réalisme tient et il n'y a pas de variables cachées, donc la localité ne tient pas.**

   Cette alternative est rejetée par les auteurs puisqu'elle entre en contradiction avec la théorie de la relativité.

3. **La localité tient et il n'y a pas de variables cachées, donc le réalisme ne tient pas.**

   Cette alternative est rejetée par les auteurs, mais sans donner d'argument rigoureux pour ce rejet.

Dans son article, Bell propose un résultat expérimentalement vérifiable qui permettrait de trancher sur l'existence ou non de variables cachées. Or, la conclusion des expériences menées sur la base de ce résultat est qu'il n'y a pas de variables cachées, donc la première alternative ne peut pas être vraie. Donc, par élimination, la troisième alternative est vraie et c'est le réalisme qui ne tient pas, au grand dam d'Einstein.

2. C'est en effet en concevant une procédure expérimentale à partir de ce théorème que Alain Aspect et ses collègues montrent en 1981 que nous obtenons des résultats statistiques compatibles avec la mécanique quantique (et incompatibles avec l'existence de variables cachées), donnant ainsi une première réponse à la question posée par Einstein, Podolsky et Rosen. Cependant, on accuse à ces expériences de laisser ouvertes certaines « échappatoires » [Wika], et ce n'est qu'en 2015 qu'une expérience dont on admet qu'elle ferme toutes les échappatoires fut conduite [G+15, S+15].

L'explication ci-dessus, avec l'utilisation des expressions telles que « variables cachées », etc. peut de prime abord sembler obscure : illustrons donc ce que tout cela signifie en nous inspirant des « chaussettes de Bertlmann », un exemple inventé par John Bell lui-même pour expliquer le paradoxe EPR [Bel81]. Reinhold Bertlmann était un collègue de Bell au CERN dans les années 1970, et il avait la particularité de porter des chaussettes dépareillées tous les jours. Nous allons supposer que les règles suivantes sont valables :

1. Bertlmann n'a que des chaussettes de couleurs bleues et roses, et si l'une des chaussettes qu'il porte est bleue, alors l'autre est rose.
2. Bertlmann porte des chaussures hautes qui cachent la couleur de ses chaussettes, et on ne peut détecter la couleur d'une chaussette qu'en enlevant la chaussure correspondante.
3. Les chaussettes peuvent être de laine ou de soie, et si l'une est de laine, l'autre est de soie.
4. On ne peut détecter le type de tissu qu'au toucher.
5. La couleur d'une chaussette ne dépend pas de son type de tissu et inversement.
6. On n'a pas le droit de toucher les chaussettes si les chaussures sont enlevées, et on n'a pas le droit d'enlever une chaussure si l'on a déjà touché la chaussette portée en-dessous.

Dans cette analogie, mesurer une propriété d'une particule afin de trouver sa valeur, c'est la même chose qu'enlever une chaussure pour identifier la couleur de la chaussette portée en-dessous, ou bien toucher une chaussette afin d'identifier le type de tissu utilisé.

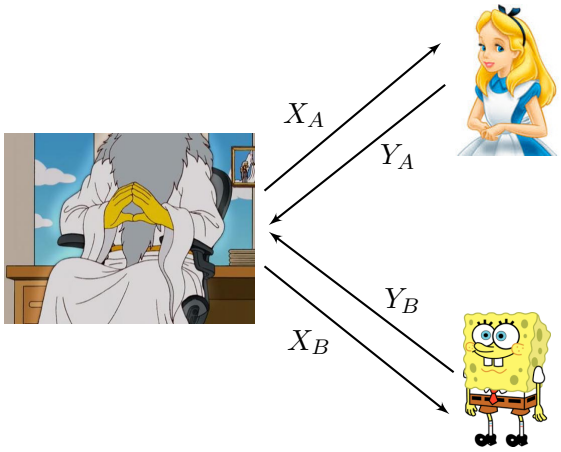
- Le principe d'incertitude de Heisenberg dit que l'on ne peut pas connaître à la fois la couleur et le type de tissu d'une chaussette : il s'agit en effet d'une conséquence de la dernière règle.
- La localité dit qu'enlever une chaussure ou toucher une chaussette n'aura aucune incidence sur la couleur ou le type de tissu de l'autre chaussette.
- Le réalisme dit que si l'on peut déterminer la couleur (resp. le type de tissu) d'une chaussette, alors il existe effectivement des propriétés ayant une valeur déterminée, la couleur et le type de tissu, dont cette chaussette est dotée.
- L'existence de variables cachées correspond au fait que le matin du jour où l'on s'interroge sur la couleur et le type de tissu des chaussettes de Bertlmann, Bertlmann a bien décidé de porter une chaussette rose d'un type de tissu donné sur un pied et une chaussette bleue de l'autre type

de tissu sur l'autre pied. On ne sait pas quelle chaussette est portée sur quel pied (c'est une information « cachée »), mais on sait que l'une des quatres configurations possibles a été choisi par Bertlmann au moment où il a enfilé ses chaussures.

Einstein, Podolsky et Rosen proposaient donc comme solution à leur problème l'hypothèse que Bertlmann a effectivement enfilé ses chaussettes d'une certaine manière le matin, ce qui semble tout à fait conforme à l'intuition. Or, dans son article, Bell montre que cette hypothèse est incompatible avec la mécanique quantique. Plus précisément, dans le monde quantique, Bertlmann n'a pas enfilé une configuration particulière de chaussettes le matin. À la place, il a enfilé une configuration globale de chaussettes qui fait qu'elles ont des valeurs opposées pour la couleur et le type de tissu sans que ces valeurs soient déterminées. Elles ne le seront en effet qu'au moment où une chaussure est enlevée (resp. une chaussette est touchée). Et alors, une fois que la valeur correspondante est connue, on sait que la chaussette sur l'autre pied aura la valeur opposée.

1.2 Le jeu CHSH et la condition non-signalling

Le jeu CHSH. En 1969, Clauser, Horne, Shimony et Holt reformulent le théorème de Bell sous forme d'un « jeu » [CHSH69].



Condition de réussite :
$Y_A \oplus Y_B = X_A \wedge X_B$

FIGURE 1.1 – Le jeu CHSH.

Dans le jeu CHSH, il y a deux joueurs, communément appelés Alice et Bob, et un arbitre. Avant le début du jeu, Alice et Bob peuvent discuter et convenir d'une stratégie, mais ils n'ont plus le droit de communiquer une fois que le jeu a débuté. Le jeu se déroule de la manière suivante : l'arbitre distribue un bit X_A à Alice et un bit X_B à Bob. Alice et Bob doivent émettre des bits Y_A et Y_B (rappelons-le, sans communication). Pour gagner, il faut que les entrées et les sorties vérifient la relation suivante : $X_A \wedge X_B = Y_A \oplus Y_B$.

Il est facile de voir (cf. l'encadré ci-dessous) que si la sortie de Alice (resp. Bob) est une fonction déterministe de son entrée, ils ne peuvent gagner à ce jeu qu'avec une probabilité de 0,75. Si la sortie de chaque joueur est une fonction probabiliste de son entrée, cela permet de se protéger du cas où l'arbitre serait malveillant et choisirait les entrées afin de faire échouer la fonction qui renvoie les entrées sur les sorties, mais pas de réussir avec une meilleure probabilité. Ces deux derniers cas (où les sorties sont fonction déterministe, resp. probabiliste des entrées) correspondent au cas où il y aurait des « variables cachées ». Or, dans cette reformulation, le résultat de Bell dit qu'en ayant accès à une certaine ressource quantique, on peut gagner avec probabilité $\cos^2(\pi/8) \approx 0,85$.

La stratégie déterministe pour le jeu CHSH

Détaillons un peu plus les conditions de réussite du jeu CHSH. Il faut avoir :

$$\begin{aligned}
&\text{(1) } Y_A \oplus Y_B = 0 \quad \text{si} \quad (X_A, X_B) = (0,0)\\
&\text{(2) } Y_A \oplus Y_B = 0 \quad \text{si} \quad (X_A, X_B) = (0,1)\\
&\text{(3) } Y_A \oplus Y_B = 0 \quad \text{si} \quad (X_A, X_B) = (1,0)\\
&\text{(4) } Y_A \oplus Y_B = 1 \quad \text{si} \quad (X_A, X_B) = (1,1)
\end{aligned}$$

Dans une stratégie déterministe, Y_A est fonction de X_A (et de même pour Y_B et X_B).

Or, si, par exemple, on décide de sortir $Y_A = 0$ lorsque $X_A = 0$, (1) et (2) nous obligent à sortir $Y_B = 0$ quel que soit X_B. Or, on doit sortir $Y_A = 1$ lorsque $X_A = 1$ pour satisfaire le cas (4). Mais alors on va sortir $(Y_A, Y_B) = (1,0)$ lorsque $(X_A, X_B) = (1,0)$, ce qui viole le cas (3).

Plus généralement, quelles que soient les affectations choisies, on va pouvoir satisfaire jusqu'à trois cas, mais alors on ne pourra plus satisfaire le cas qui reste, ce qui signifie que l'on ne réussira que trois fois sur quatre.

Avant de continuer, remarquons que dans le jeu CHSH, on parle d'entrées

et de sorties. Or, rappelons que ce jeu est une reformulation d'une expérience de physique. Concrètement, les entrées correspondent donc à des choix de mesure, et les sorties à des résultats de mesure.

Différents types de solution au jeu CHSH. Nous allons maintenant exprimer mathématiquement les différents cas de figure qui ont été considérés ci-dessus.

1. Si la sortie de chaque joueur est une **fonction déterministe** de son entrée, cela signifie que $Y_A = f(X_A)$ et $Y_B = g(X_B)$. Une autre formulation, qui est plus intéressante pour la suite, est que

   (a) $\Pr(Y_A = f(x_A) \mid X_A = x_A) = 1$

   (b) $\Pr(Y_B = g(x_B) \mid X_B = x_B) = 1$

   (c) $\Pr(Y_A \neq f(x_A) \mid X_A = x_A) = 0$

   (d) $\Pr(Y_B \neq g(x_B) \mid X_B = x_B) = 0$

   De plus, comme la sortie de chaque joueur n'est fonction que de son entrée, les probabilités conditionnelles sont indépendantes, i.e. pour tout y_A, y_B, x_A, x_B :

   $$\Pr(Y_A = y_A, Y_B = y_B \mid X_A = x_A, X_B = x_B)$$
   $$=$$
   $$\Pr(Y_A = y_A \mid X_A = x_A) \Pr(Y_B = y_B \mid X_B = x_B).$$

2. Maintenant, si la sortie de chaque joueur est une **fonction probabiliste** de son entrée, cela revient à dire qu'il existe un certain nombre de fonctions déterministes $f_1, \cdots, f_n$, $g_1 \cdots, g_m$ et des réels positifs $p_1, \cdots, p_n$, $q_1 \cdots, q_m$ avec

   $$\sum_{i=1}^{n} p_i = 1$$

   $$\sum_{j=1}^{m} q_j = 1$$

   telles que $Y_A = f_i(X_A)$ avec probabilité p_i et $Y_B = g_j(X_B)$ avec probabilité q_j.

   De plus, ici aussi la sortie de chaque joueur n'est fonction que de son entrée, par conséquent :

$$\Pr(Y_A = f_i(x_A), Y_B = g_j(y_B) \mid X_A = x_A, X_B = x_B)$$
$$=$$
$$\Pr(Y_A = f_i(x_A) \mid X_A = x_A) \Pr(Y_B = g_j(x_B) \mid X_B = x_B)$$

Quelques remarques :

- La situation précédente, où la sortie de chaque joueur était fonction déterministe de son entrée, est un cas particulier de cette situation : en effet, on retrouve le cas « fonction déterministe » en prenant $n = m = 1$, $f_1 = f$, $g_1 = g$ et $p_1 = q_1 = 1$.
- Une dernière chose à noter est que ce que nous avons présenté ici correspond à ce qui est appelé l'**aléa local** dans la littérature. On peut aussi avoir une source d'**aléa partagé**, c'est-à-dire des réels positifs $\lambda_1, \cdots, \lambda_{m\times n}$ avec

  $$\sum_{i=1}^{m\times n} \lambda_i = 1$$

  tel que pour tout $i \in \{1,\ldots,n\}$, $j \in \{1,\ldots,m\}$, il existe $k \in \{1,\ldots,m\times n\}$ tel que $Y_A = f_i(X_A)$ et $Y_B = g_j(X_B)$ avec probabilité λ_k.

  On aura alors

  $$\Pr(Y_A = f_i(x_A), Y_B = g_j(y_B) \mid X_A = x_A, X_B = x_B)$$
  $$=$$
  $$\Pr(Y_A = f_i(x_A) \mid X_A = x_A) \Pr(Y_B = g_j(x_B) \mid X_B = x_B)$$

  L'aléa partagé est plus fort que l'aléa local (dans la mesure où l'on peut simuler l'aléa local à l'aide de l'aléa partagé) mais ce n'est pas un grand avantage : en effet, il y a des résultats qui montrent que l'aléa partagé peut également être simulé à l'aide de l'aléa local pour un coût supplémentaire négligeable dans divers cas de figure [New91, GK19].

3. Nous ne donnons pas les détails de la stratégie quantique. Nous notons cependant que, si le réalisme ne tient pas mais que la localité tient, alors la sortie d'un joueur ne doit pas permettre d'obtenir des informations sur l'entrée de l'autre joueur (sinon de l'information au sujet de cette entrée pourrait être transmise instantanément, ou **signalée**, via la mesure) : on appelle cela la propriété de **non-signalling**.

Mathématiquement, cela revient à dire que la distribution de probabilité de la sortie d'un joueur en fonction de son entrée peut être définie indépendamment de l'entrée de l'autre :

Pour Alice :

$$\begin{aligned}\Pr(Y_A = y_A \mid X_A = x_A) &= \sum_{y_B=0,1} \Pr(Y_A = y_A, Y_B = y_B \mid X_A = x_A, X_B = 0) \\ &= \sum_{y_B=0,1} \Pr(Y_A = y_A, Y_B = y_B \mid X_A = x_A, X_B = 1)\end{aligned}$$

Pour Bob :

$$\begin{aligned}\Pr(Y_B = y_B \mid X_B = x_B) &= \sum_{y_A=0,1} \Pr(Y_A = y_A, Y_B = y_B \mid X_A = 0, X_B = x_B) \\ &= \sum_{y_A=0,1} \Pr(Y_A = y_A, Y_B = y_B \mid X_A = 1, X_B = x_B)\end{aligned}$$

Notons que toute distribution obtenue à l'aide de ressources classiques (donc où la sortie de chaque joueur est une fonction déterministe ou probabiliste de son entrée) ou de ressources quantique est non-signalling.

Une hiérarchie de distributions de probabilités. Nous pouvons tirer la conclusion suivante à partir de ce qui a été dit précédemment :

1. Tout ce qui peut être fait avec une fonction déterministe classique peut être fait avec une distribution de probabilités classique.
2. Toute distribution de probabilité classique peut être obtenue à l'aide de mesures appropriées sur une ressource quantique appropriée.
3. Toute distribution de probabilité obtenue en effectuant des mesures sur une ressource quantique vérifie la condition non-signalling.

De plus, même si cela est interdit par les règles du jeu, nous pouvons noter que si Alice et Bob avaient le droit de communiquer, et donc de connaître leurs entrées respectives et coordonner leurs sorties respectives, ils pourraient réaliser n'importe lesquelles de ces distributions (déterministe, probabiliste ou non-signalling).

Ou encore, en notant

- $\mathscr{D}_{\text{dét}}$ l'ensemble des distributions correspondant aux fonctions déterministes
- $\mathscr{D}_{\text{aléa}}$ l'ensemble des distributions correspondant aux fonctions probabilistes (avec l'aléa local ou partagé)

- $\mathscr{D}_{\text{quant}}$ l'ensemble des distributions issues d'une mesure sur une ressource quantique
- $\mathscr{D}_{\text{NS}}$ l'ensemble des distributions vérifiant la condition non-signalling
- $\mathscr{D}_{\text{S}}$ l'ensemble des distributions possibles si Alice et Bob avaient le droit de communiquer

on a :

$$\mathscr{D}_{\text{dét}} \subsetneq \mathscr{D}_{\text{aléa}} \subsetneq \mathscr{D}_{\text{quant}} \subset \mathscr{D}_{\text{NS}} \subset \mathscr{D}_{\text{S}} \tag{1.1}$$

Dans la littérature, le terme **non-local** est utilisée pour qualifier les distributions qui sont dans $\mathscr{D}_{\text{S}}$ sans être dans $\mathscr{D}_{\text{aléa}}$, i.e. qui ne sont pas réalisables à l'aide de ressources classiques si aucune communication n'est autorisée. Nous voyons alors, grâce à la hiérarchie ci-dessus, que certaines distributions quantiques et non-signalling (i.e. dans $\mathscr{D}_{\text{quant}}$ ou dans $\mathscr{D}_{\text{NS}}$) sont non-locales.

Non-localité et intrication. Revenons brièvement sur certaines notions introduites dans la Section 1.1 afin de faire le lien avec la notion de non-localité.

Les distributions dans $\mathscr{D}_{\text{quant}}$ sont issues d'une mesure sur une ressource quantique, et celle-ci peut être dans un état intriqué ou non. Il y a le lien suivant entre la nature de l'état de la ressource quantique en question et la non-localité de la distribution résultante [Bar07] :

- Toute distribution non-locale est nécessairement issue d'une mesure sur une ressource quantique dans un état intriqué, i.e. **la non-localité vient toujours de l'intrication**.
- Mesurer une ressource quantique dans un état intriqué ne donnera pas nécessairement lieu à une distribution non-locale, i.e. **l'intrication n'a pas toujours pour conséquence la non-localité**.

1.3 La boîte Popescu-Rohrlich

En 1994, Popescu et Rohrlich [PR94] remarquent qu'il existe une distribution de probabilité qui permet de gagner avec probabilité 1 au jeu CHSH et qui de surcroît a la propriété d'être non-signalling. Cette distribution de probabilité, communément appelée la **boîte Popescu-Rohrlich (PR)** dans la littérature, est la suivante :

$$\Pr(Y_A = 0, Y_B = 0 \mid X_A = x_A, X_B = x_B) = 1/2 \text{ si } x_A = 0 \text{ ou } x_B = 0$$

$$\Pr(Y_A = 1, Y_B = 1 \mid X_A = x_A, X_B = x_B) = 1/2 \text{ si } x_A = 0 \text{ ou } x_B = 0$$

$$\Pr(Y_A = 0, Y_B = 1 \mid X_A = x_A, X_B = x_B) = 1/2 \text{ si } x_A = 1 \text{ et } x_B = 1$$

$$\Pr(Y_A = 1, Y_B = 0 \mid X_A = x_A, X_B = x_B) = 1/2 \text{ si } x_A = 1 \text{ et } x_B = 1$$

$$\Pr(Y_A = y_A, Y_B = y_B \mid X_A = x_A, X_B = x_B) = 0 \text{ dans les autres cas}$$

Un argument algébrique dû à Boris Tsirelson permet de montrer qu'il est impossible de gagner au jeu CHSH avec une meilleure probabilité que $\cos^2(\pi/8)$ en employant des ressources quantiques [Tsi80]. De surcroît, il a été montré que si la boîte PR était physiquement implémentable (à l'aide de ressources physiques inconnues de la communauté scientifique au moment actuel, par exemple), cela aurait pour conséquence de rendre tous les problèmes de « complexité de la communication » [3] triviaux [vD13], ce qui va à l'encontre du consensus général.

Ainsi, il existe des distributions non-signalling qui ne peuvent être implémentées à l'aide d'aucune ressource quantique (ni d'ailleurs à l'aide d'autres ressources physiques inconnues de nous au moment actuel). Par conséquent, l'avant-dernière inclusion dans la hiérarchie (1.1) est en réalité une inclusion stricte, et une version plus exacte de cette hiérarchie serait :

$$\mathscr{D}_{\text{dét}} \subsetneq \mathscr{D}_{\text{aléa}} \subsetneq \mathscr{D}_{\text{quant}} \subsetneq \mathscr{D}_{\text{NS}} \subset \mathscr{D}_{\text{S}} \tag{1.2}$$

La boîte PR et d'autres « boîtes non-signalling » (avec plus d'entrées, plus de sorties, et/ou plus de joueurs) ont été depuis étudiées dans la littérature.

Leur étude est intéressante pour des raisons mathématiques : pour un nombre d'entrées, de sorties et de joueurs fixés, l'ensemble de telles boîtes définit en effet un polytope convexe dont il peut être intéressant de déterminer le nombre et la nature des points extrémaux [BLM+05].

Mais cela a aussi une application à la théorie quantique comme il a été noté par [Arf14] : puisque toute distribution quantique est non-signalling, alors, pour un problème donné, si aucune distribution non-signalling meilleure qu'une distribution de probabilité classique permettant

3. Il s'agit d'un domaine d'études dont le but est de calculer la quantité de communication nécessaire pour effectuer diverses tâches de manière distribuée entre deux partis.

de le résoudre n'existe, alors aucune distribution obtenue en ayant recours aux ressources quantiques ne sera meilleure que cette distribution de probabilité classique. Par conséquent, ce problème n'admettra pas d'avantage quantique. Cette technique permettrait d'étudier la résolvabilité quantique ou non de problèmes (ce qui demanderait de regarder un nombre infini d'états) par l'intermédiaire de l'étude des distributions non-signalling correspondantes (donc de regarder un polytope avec un nombre fini de points extrémaux), autrement dit par l'étude de ce qui est a priori un problème plus simple.

Par exemple, dans [Arf14] l'ensemble des boîtes à deux joueurs et à entrées et sorties binaires a été étudiée : il a été montré que les boîtes non-signalling confèrent un avantage seulement pour les problèmes de type $Y_A \oplus Y_B = f(X_A, X_B)$, où X_A et X_B sont les entrées des joueurs A et B, Y_A et Y_B leurs sorties, et f une fonction calculable quelconque; et que, par exemple, pour le problème $Y_A \wedge Y_B = \neg(X_A \wedge X_B)$, toute solution permettant une probabilité de réussite supérieure à $2/3$ est signalling, et, de surcroît, il existe une solution n'utilisant que des ressources classiques et permettant de gagner avec probabilité $2/3$.

Cela nous permet de terminer cette section par une remarque qui permet de raffiner davantage la hiérarchie 1.2 : n'importe quel problème calculable est résoluble avec probabilité 1 si nous avons droit à la communication (en effet, cela reviendrait à résoudre le problème de manière centralisée). Or, il existe des problèmes calculables tels qu'il n'y a aucune boîte non-signalling permettant de les résoudre avec probabilité 1, d'où :

$$\mathscr{D}_{\text{dét}} \subsetneq \mathscr{D}_{\text{aléa}} \subsetneq \mathscr{D}_{\text{quant}} \subsetneq \mathscr{D}_{\text{NS}} \subsetneq \mathscr{D}_{\text{S}} \tag{1.3}$$

1.4 Généralisation des boîtes non-signalling

Dans cette section, nous allons donner la généralisation de la condition non-signalling à plusieurs joueurs. Nous allons également énoncer un résultat de « redondance » qui permet de réduire considérablement le nombre d'équations intervenant dans cette condition.

Condition non-signalling à n-joueurs. Supposons que nous avons n joueurs $\{1, \ldots, n\}$, ayant chacun des entrées prises dans un ensemble fini I, et dont les sorties sont des éléments d'un ensemble fini O. Alors, si la distribution de leurs sorties en fonction de leurs entrées est non-signalling, cela signifie que un ou plusieurs joueurs, même s'ils forment une coalition, ne pourront pas utiliser leurs sorties pour déduire quelque chose sur les entrées des autres joueurs.

Autrement dit :

Pour tout $k \in \{1, \ldots, n-1\}$,

pour tout $S \subset \{1, \ldots, n\}$ sous-ensemble de taille k,

pour tout $x_S \in I^S$, $x_{\overline{S}},\ x'_{\overline{S}} \in I^{\overline{S}}$

pour tout $y_S \in O^S$

$$\begin{aligned}\Pr(Y_S = y_S \mid X_S = x_S) &= \sum_{y_{\overline{S}} \in O^{\overline{S}}} \Pr(Y_S = y_S, Y_{\overline{S}} = y_{\overline{S}} \mid X_S = x_S, X_{\overline{S}} = x_{\overline{S}}) \\ &= \sum_{y_{\overline{S}} \in O^{\overline{S}}} \Pr(Y_S = y_S, Y_{\overline{S}} = y_{\overline{S}} \mid X_S = x_S, X_{\overline{S}} = x'_{\overline{S}})\end{aligned} \tag{1.4}$$

En réalité, il a été démontré dans [BLM+05] qu'il suffit de regarder seulement les coalitions de $n-1$ joueurs.

Autrement dit, en notant $X_{-i} := (X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n)$ (et de même pour Y_{-i}, etc.) :

Pour tout $i \in \{1, \ldots, n\}$,

pour tout $x_{-i} \in I^{n-1}$, $x_i,\ x'_i \in I$

pour tout $y_{-i} \in O^{n-1}$

$$\begin{aligned}\Pr(Y_{-i} = y_{-i} \mid X_{-i} = x_{-i}) &= \sum_{y_i \in O} \Pr(Y_{-i} = y_{-i}, Y_i = y_i \mid X_{-i} = x_{-i}, X_i = x_i) \\ &= \sum_{y_i \in O} \Pr(Y_{-i} = y_{-i}, Y_i = y_i \mid X_{-i} = x_{-i}, X_i = x'_i)\end{aligned} \tag{1.5}$$

La preuve en est très simple : l'encadré suivant contient la preuve pour un cas particulier, ce qui en donne l'intuition de la preuve dans le cas général.

Preuve que (1.5) implique (1.4) : un cas particulier

Nous allons regarder le cas où $n = 3$, et démontrer que la condition pour $S = \{1\}$ suit des conditions (1.5) pour $i = 2$ et $i = 3$:

$$\begin{aligned}
&\sum_{y_2,y_3\in O} \Pr(Y_1 = y_1, Y_2 = y_2, Y_3 = y_3 \mid X_1 = x_1, X_2 = x_2, X_3 = x_3) \\
&= \sum_{y_2\in O}\sum_{y_3\in O} \Pr(Y_1 = y_1, Y_2 = y_2, Y_3 = y_3 \mid X_1 = x_1, X_2 = x_2, X_3 = x_3) \\
&= \sum_{y_2\in O}\sum_{y_3\in O} \Pr(Y_1 = y_1, Y_2 = y_2, Y_3 = y_3 \mid X_1 = x_1, X_2 = x_2, X_3 = x'_3) \\
&= \sum_{y_3\in O}\sum_{y_2\in O} \Pr(Y_1 = y_1, Y_2 = y_2, Y_3 = y_3 \mid X_1 = x_1, X_2 = x'_2, X_3 = x'_3) \\
&= \sum_{y_2,y_3\in O} \Pr(Y_1 = y_1, Y_2 = y_2, Y_3 = y_3 \mid X_1 = x_1, X_2 = x'_2, X_3 = x'_3)
\end{aligned}$$

Non-localité à n parties. Enfin, bien que ce soit la condition non-signalling qui nous intéressera principalement dans ce qui suit, nous souhaiterions terminer par une remarque brève à propos des phénomènes d'intrication et de non-localité qui y sont reliées.

En effet, elles se manifestent également lorsqu'il y a plus de deux parties, mais le tableau se complique : par exemple, on peut avoir de l'« intrication partielle », c'est-à-dire une bipartition (ou plus) des n parties telles que le « sous-état » de chaque sous-ensemble des joueurs est un état intriqué, mais les deux sous-états ne sont pas intriqués l'un avec l'autre. De même, il est possible d'avoir de la non-localité partielle, et c'est un phénomène plus subtil que ce qui a été pensé au début.

CHAPITRE 2
Calcul distribué et le modèle LOCAL

2.1 Le calcul distribué et le problème de coloration

Le calcul distribué : des problématiques diverses et variées. À la naissance de l'informatique, on s'intéressait aux questions de calculabilité sur une seule machine, ou encore des questions d'efficacité telles que comment réduire la taille d'une machine ou construire plus de machines à moindre coût. Au fur et à mesure que ces problèmes furent résolus et que les ordinateurs devinrent de plus en plus ubiquitaires, de nouvelles possibilités émergèrent : par exemple, la possibilité de stocker la même information sur plusieurs postes afin de se protéger contre les risques de perte de cette information et avoir ainsi plus de **robustesse**. Pour ce faire, il fallait s'assurer que les ordinateurs puissent communiquer efficacement, ce qui donna lieu à toute une série de nouvelles problématiques : est-il possible d'assurer une communication efficace, même si l'infrastructure physique sous-jacente ne garantit pas que l'information envoyée sera transmise en entier, au bout d'un intervalle de temps borné et sans erreur, comme c'est le cas pour le modèle de réseau à paquets sur lequel internet est basé ? Est-il possible d'assurer qu'un problème au niveau d'un des **nœuds** du réseau, qui fait qu'il ne participe plus au réseau (un *crash*) ou qu'il envoie des informations fausses (un *comportement byzantin*), ne remette pas en cause le bon fonctionnement de tout le réseau [LSP82, FLP85] ? De nombreux problèmes émergent ainsi car le réseau est supposé être **non-fiable** et **asynchrone**.

Mais on peut aussi supposer que le réseau est fiable (c'est-à-dire qu'il n'y a pas de crash, ni de comportement byzantin) et synchrone (c'est-à-dire qu'il y a une horloge « globale » pour tous les nœuds du réseau), et s'intéresser aux questions autour de la quantité de communication nécessaire pour résoudre une tâche donnée. Par exemple :

- Si chaque canal de communication entre deux nœuds du système a une bande passante limitée (par exemple, au plus B bits), combien de fois faut-il et suffit-il de communiquer pour résoudre un problème donné ?

  Le modèle CONGEST(B) [Pel00, page 27] a été développé afin de répondre à ce type de questions.

- Faut-il et suffit-il de communiquer avec ses voisins pour résoudre un problème donné, ou faut-il également communiquer avec les voisins

de ses voisins ? Est-il impossible de résoudre un problème donné sans communiquer avec tout le réseau, autrement dit, sans que chaque nœud reçoive des informations de la part de tout le réseau ?

Le modèle LOCAL [Lin87, Lin92] [Pel00, page 27] a été développé pour répondre à ce type de questions, où l'on s'intéresse principalement à la distance de communication nécessaire et/ou suffisante, donc sans imposer de limites sur la bande passante autorisée.

Le modèle LOCAL. Dans ce qui suit, nous nous intéresserons principalement au modèle LOCAL et à quelques extensions. Ainsi, nous commencerons par définir plus précisément ce modèle.

Nous avons un réseau de communication modélisé par un graphe (orienté ou non) $G = (V, E)$.

1. V est l'ensemble des **nœuds** du graphe, représentant les acteurs du réseau (ordinateurs ou processeurs).
2. E l'ensemble des **arêtes** du graphe, représentant les liens de communication entre les acteurs.
3. Le graphe est étiqueté (typiquement par des entiers), l'étiquette de chaque nœud étant son **identifiant**.
4. Les identifiants sont deux-à-deux distincts.
5. On définit une **ronde de communication** de la manière suivante : chaque nœud envoie et reçoit des messages à ses voisins et fait un calcul local.
6. Nous n'imposons aucune limite sur la bande passante des liens reliant les nœuds, ni sur la puissance de calcul local de chaque nœud.

Nous remarquerons aussi qu'il est équivalent (dans le monde régi par la physique classique du moins) de communiquer et faire des calculs pendant r rondes de communications, et de communiquer pendant r rondes de communication et faire les calculs une fois que les communications seront effectuées. Autrement dit, cela revient au même d'envoyer et de recevoir des messages à ses voisins immédiats et faire un calcul basé sur ces messages r fois, et d'envoyer et de recevoir des messages à ses voisins situés à une distance d'au plus r, et faire un seul calcul basé sur tous les messages reçus. En effet, pour chaque nœud, cela revient à faire plusieurs étapes de calcul progressivement en fonction des nouvelles informations obtenues à chaque ronde de communication, ou bien attendre d'avoir toutes les informations qu'il est possible d'obtenir à partir des nœuds à distance au plus r de ce

nœud et de faire le calcul en une seule fois. En tout cas, pour chaque nœud, les informations qui servent à faire le calcul sont exactement, sans lacune ni excédant, les informations obtenues des nœuds à distance au plus r d'un nœud donné. Nous pouvons résumer cela par

Principe de l'équivalence dans le modèle LOCAL

« r fois une ronde de communication = une r-ronde de communication »

La coloration de graphes dans le modèle LOCAL. Un des résultats les plus connus dans ce domaine est celui concernant la **coloration de graphes** (voir par exemple [BE13]) : il s'agit d'assigner une étiquette (une couleur) parmi un nombre q fixé d'étiquettes à chaque nœud telle que des nœuds voisins ont des étiquettes différentes. Ce problème a des utilisations pratiques, par exemple pour **l'ordonnancement de tâches**, où il s'agit de répartir diverses tâches entre plusieurs processeurs tels que des processeurs voisins n'effectuent pas la même tâche. Il est NP-difficile de déterminer le **nombre chromatique** $\chi(G)$ d'un graphe G quelconque, c'est-à-dire le nombre minimum de couleurs avec lequel il peut être colorié, et encore il est d'autant plus difficile de trouver une bonne coloration. Pour cette raison, pour la coloration distribuée on choisit souvent $q = \Delta + 1$, où Δ est le **degré du graphe**, c'est-à-dire la valeur maximum des degrés des sommets : en effet, un raisonnement par récurrence simple permet de démontrer que n'importe quel graphe est toujours $(\Delta + 1)$-coloriable, et cela peut être fait de façon efficace.

En 1986, Cole et Vishkin proposent un algorithme qui permet de colorier les chemins orientés, les cycles orientés et les arbres orientés enracinés avec $(\Delta + 1)$, en l'occurrence, 3 couleurs, en $\mathcal{O}(\log^*(n))$ rondes de communication [CV86], où $\log^*(n)$ est le nombre de fois qu'il faut appliquer la fonction $\log$ à n afin d'obtenir une valeur inférieure ou égale à 1 : c'est la fonction réciproque de la fonction de **tétration** ou la **tour d'exponentielle**.

La fonction tour d'exponentielle (binaire) [a]

$$\begin{aligned} Tr(0) &= 1 \\ Tr(n) &= 2^{Tr(n-1)},\ n \geq 1 \end{aligned}$$

a. Plus généralement, on peut définir la fonction tour d'exponentielle de base a, où l'on prend a comme base d'exponentiation à la place de 2.

C'est donc un nombre plus grand que n'importe quelle constante en théorie, même si en pratique (c'est-à-dire, pour $n \leq 2^{2^{16}}$, qui est une borne supérieure sur le nombre de particules dans l'univers), on a $\log^*(n) \leq 5$.

Il est possible d'étendre l'algorithme de Cole et Vishkin à n'importe quel graphe en utilisant une décomposition en forêts [PR01].

En 1987 puis 1992, Nathan Linial conçoit une méthode pour démontrer des bornes inférieures sur le nombre de rondes nécessaires pour colorier un graphe à l'aide d'un algorithme déterministe ou probabiliste de type Las Vegas (c'est-à-dire, qui utilise de l'aléa, potentiellement partagé, mais où l'obtention du bon résultat est garanti) [Lin87, Lin92].

En 1991, Moni Naor utilise une argumentation similaire pour démontrer une borne inférieure sur le nombre de rondes nécessaires avec un algorithme probabiliste de type Monte-Carlo (c'est-à-dire, qui est susceptible d'échouer avec une certaine probabilité) [Nao91].

Une conséquence de ces résultats est l'optimalité, à une constante additive près, de l'algorithme de Cole et Vishkin : Linial démontre en effet une borne inférieure de $\frac{1}{2}(\log^*(n) - 3)$ rondes de communications, et Naor une borne inférieure de $t := \frac{1}{2}\log^*(n) - b - 2$ pour obtenir une bonne coloration avec probabilité au moins $(1 - (1/\log^{(b)}(n)))^{n^{1/3}/2t} + 2t/n^{1/3}$ sur un chemin ou un cycle (non-orienté).

Plan de la suite. Dans ce qui suit, nous allons d'abord expliquer l'algorithme de Cole et Vishkin. Puis nous allons présenter l'argument de Linial après en avoir explicité les principaux ingrédients. Enfin, nous allons prendre l'exemple d'un jeu quantique, le **carré magique de Mermin-Peres**, pour expliquer en quoi l'argument de Linial ne fonctionne plus dans un contexte où l'on permet l'utilisation de ressources quantiques.

2.2 Algorithme de Cole et Vishkin

Algorithme de Cole et Vishkin. Comme il a été dit, nous supposons que nous disposons d'un cycle orienté de N sommets, dont chacun est étiqueté avec une valeur distincte prise dans $\{1, \ldots, N\}$ (pour simplifier). Notons $n := \lceil \log_2(N) \rceil$, et remarquons que les identifiants fournissent une coloration initiale : en effet, tous les identifiants sont distincts donc a fortiori les identifiants de deux nœuds voisins sont distincts.

Nous définissons la procédure suivante, qui nécessite une ronde de communication, pour déduire une nouvelle coloration : chaque sommet u obtient à partir de son voisin de l'arête entrante l'identifiant ou la couleur de ce dernier. Il considère ensuite l'écriture binaire $\mathbf{b} = (b_n, \ldots, b_1)$ de sa couleur et de la couleur de son voisin $\mathbf{c} = (c_n, \ldots, c_1)$. Ces écritures binaires sont distinctes puisque les deux couleurs le sont. Alors, il existe $1 \leq i \leq n$ tel

que $(b_i, \ldots, b_1) = (c_i, \ldots, c_1)$ et $c_{i+1} \neq b_{i+1}$. Alors la nouvelle couleur de u est $F_u(\mathbf{b}, \mathbf{c}) = (c_{i+1}, i_k, \ldots, i_1)$, où $(i_k, \ldots, i_1)$ est l'écriture binaire de i. Nous constatons que :

1. $|F_u(\mathbf{b}, \mathbf{c})| \leq \lceil \log_2(n) \rceil + 1$, puisque $1 \leq i \leq n$ donc il y a au plus $\lceil \log_2(n) \rceil$ bits dans l'écriture binaire de i. Par conséquent, nous passons de N couleurs à au plus $2^{\lceil \log_2(n) \rceil + 1} \approx 2n \approx 2\lceil \log_2(N) \rceil$ couleurs dans la nouvelle coloration.

2. Si u a comme couleur $\mathbf{b}$ et son voisin v a comme couleur $\mathbf{c}$, et que le voisin de v a comme couleur $\mathbf{d}$, et que ces trois couleurs sont deux-à-deux distinctes, alors les nouvelles couleurs de u et de v sont distinctes aussi.

   En effet, ces nouvelles couleurs sont $F_u(\mathbf{b}, \mathbf{c}) = (c_{i+1}, i_k, \ldots, i_1)$ et $F_v(\mathbf{c}, \mathbf{d}) = (d_{j+1}, j_\ell, \ldots, j_1)$, où i et j sont les premiers indices où $\mathbf{b}$ et $\mathbf{c}$ (resp. $\mathbf{c}$ et $\mathbf{d}$) diffèrent, et $(i_k, \ldots, i_1)$ est l'écriture binaire de i et $(j_\ell, \ldots, j_1)$ celle de j. Alors, si $\mathbf{b} \neq \mathbf{c}$ et $\mathbf{c} \neq \mathbf{d}$, cela signifie que :

   - Soit $i \neq j$, auquel cas $(i_k, \ldots, i_1) \neq (j_\ell, \ldots, j_1)$ et donc $F_u(\mathbf{b}, \mathbf{c}) \neq F_v(\mathbf{c}, \mathbf{d})$.
   - Soit $i = j$, mais alors, puisque $\mathbf{c} \neq \mathbf{d}$ et par définition de j, $c_{i+1} = c_{j+1} \neq d_{j+1}$, d'où $F_u(\mathbf{b}, \mathbf{c}) \neq F_v(\mathbf{c}, \mathbf{d})$.

Nous voyons que dans l'argumentation ci-dessus, seul le fait que des sommets voisins aient des couleurs distinctes sert à garantir que les nouvelles couleurs des sommets voisins seront distinctes aussi. Ainsi, en appliquant une première fois cette technique, nous passons de N couleurs à $2\lceil \log_2(N) \rceil$ couleurs, et nous pourrons réappliquer cette technique afin de passer de $2\lceil \log_2(N) \rceil$ couleurs à $2\lceil \log_2(2\lceil \log_2(N) \rceil) \rceil = 2\lceil \log_2(\lceil \log_2(N) \rceil) + 1 \rceil$ couleurs, etc. Ainsi, cette technique est appelée la **réduction de couleur itérative** dans la littérature (voir par exemple [KW06] pour une étude de quelques généralisations de cette technique).

Notons que la réduction de couleurs s'arrête lorsqu'il y a 6 couleurs, ce qui arrive en $\mathcal{O}(\log_2^*(N))$ rondes de communications par définition de la fonction $\log^*$. En effet, $2\lceil \log_2(6) \rceil = 6$. Donc, à partir du moment où l'on atteint 6 couleurs (il y a des techniques pour détecter cela, par exemple, calculer le nombre d'itérations nécessaires si l'on connaît N), pour réduire le nombre de couleurs à 3, il faut adopter une autre procédure qui est la suivante :

Chaque nœud reçoit de son voisin de l'arête entrante la couleur de ce dernier. Ensuite, si sa couleur est la couleur la plus grande (donc, si sa couleur est 6, pour commencer), il change de couleur pour avoir la plus petite couleur qui n'est pas la couleur de son voisin. Sinon, il conserve sa couleur. Cette

procédure va ainsi transformer une bonne coloration initiale en une bonne coloration avec une couleur en moins. Donc, en appliquant cette procédure trois fois, on passera de 6 couleurs à 3 couleurs.

D'autres algorithmes, d'autres techniques : bref aperçu. Nous avons dit que l'algorithme de Cole et Vishkin et ses extensions reposent sur la technique de réduction de couleur itérative. Bien que nous ne l'étudions pas davantage dans la suite de ce manuscrit, nous souhaiterions mentionner qu'une autre technique, la **fragmentation de graphe** (*graph shattering* en anglais), a été récemment développée pour résoudre la coloration distribuée et d'autres problèmes de type « rupture de symétrie » [BEPS16]. Il s'agit d'algorithmes se déroulant en deux phases : une première phase probabiliste, inspirée de la version distribuée du lemme local de Lovász [CPS14], où il s'agit de répéter $\mathcal{O}(\text{polylog}(\Delta))$ fois une procédure afin de résoudre le problème sur une partie du graphe avec une probabilité d'échec négligeable. On résout ensuite le problème sur les *fragments* de graphe qu'il reste à l'aide du meilleur algorithme déterministe existant, sachant que l'on a la garantie que chaque fragment de graphe ainsi obtenu est de taille au plus $\mathcal{O}(\text{polylog}(n))$.

Dans un autre article récent [CKP16], il a été montré qu'une borne inférieure sur la complexité d'un algorithme distribué probabiliste pour résoudre un problème donné sur une instance de taille n est la complexité d'un algorithme distribué déterministe pour le même problème sur une instance de taille $\log(n)$, ce qui est très proche de la taille des fragments obtenus via la technique du *graph shattering*, et laisserait penser que cette technique est fondamentale pour le passage entre algorithmes déterministes et algorithmes probabilistes.

2.3 Preuve de Linial

Un jeu multijoueur sur les graphes. Avant de détailler la preuve de Linial sur le nombre minimal de rondes de communications nécessaires pour la coloration, nous allons définir un type de jeu à plusieurs joueurs sur un graphe $G = (V_G, E_G)$.

Notons pour cela $N := |V|$ le nombre de sommets du graphe et m le nombre de joueurs, et J_i le joueur numéro i pour $1 \leq i \leq m$.

Dans ce jeu, y a un arbitre dont le rôle est de placer les m joueurs sur $s \leq m$ sommets du graphe G (donc, potentiellement, plusieurs joueurs peuvent être placés sur le même sommet), et les joueurs doivent résoudre un problème P (qui peut différer d'une version du jeu à l'autre). Pour cela, ils peuvent

communiquer et convenir d'une stratégie avant le début du jeu, mais ils n'ont plus le droit de communiquer une fois qu'ils sont placés sur les sommets du graphe.

Plus précisément, on donne

- un graphe $G = (V_G, E_G)$
- un ensemble $\mathscr{G}$ de sous-graphes de G de taille $\leq m$
- pour chaque sous-graphe $H \in \mathscr{G}$, un ensemble $\Pi(H) = \{\pi : \{J_1, \ldots, J_m\} \to V_H\}$ de **fonctions de placement** des joueurs (ces fonctions étant surjectives mais non nécessairement injectives).
- un ensemble $\mathscr{Y} = \{Y_{H,\pi} \mid H \in \mathscr{G}, \pi \in \Pi(H)\}$ de sorties autorisées pour chaque sous-graphe H et chaque placement de joueurs sur ce sous-graphe.

1. Au moment où le jeu commence, l'arbitre choisit un sous-graphe $H \in \mathscr{G}$ et une fonction de placement $\pi \in \Pi$, et place chaque joueur J_i sur le sommet $\pi(J_i)$ du sous-graphe H.
2. Chaque joueur i doit sortir une valeur y_i telle que $(y_1, \ldots, y_m) \in Y_{H,\pi}$.

Un cas particulier : le graphe des voisinages. Nous nous intéresserons plus particulièrement au cas particulier suivant du jeu général défini ci-dessus :

Nous considérons l'ensemble $\mathscr{H}$ de tous les graphes étiquetés qu'il est possible d'obtenir à partir d'un même graphe et d'un même ensemble d'étiquettes (par exemple, l'ensemble de tous les cycles de longueur n étiquetés à l'aide d'une permutation σ_n de $\{1, \ldots, n\}$). Nous fixons un paramètre $t \leq n$ et nous définissons le graphe G de la manière suivante : les sommets de G sont les voisinages de taille t (ou encore t-voisinages) des graphes dans $\mathscr{H}$, et deux sommets de G sont reliés si et seulement s'ils sont des t-voisinages de deux sommets adjacents d'un graphe $H \in \mathscr{H}$.

Autrement dit (en utilisant la notation $N_t(v)$ pour le t-voisinage d'un sommet v :

$V_G = \{N_t(u) \mid u \in V_H, H \in \mathscr{H}\}$

$E_G = \{\{N_t(u), N_t(v)\} \mid u, v \in V_H, \{u, v\} \in E_H, H \in \mathscr{H}\}$

Le graphe G ainsi défini s'appelle **le graphe des voisinages à distance** t des graphes de $\mathscr{H}$.

$$\mathscr{H} = \{1 \to 2 \to 3, 1 \to 3 \to 2, 2 \to 1 \to 3, 2 \to 3 \to 1, 3 \to 1 \to 2, 3 \to 2 \to 1\}$$

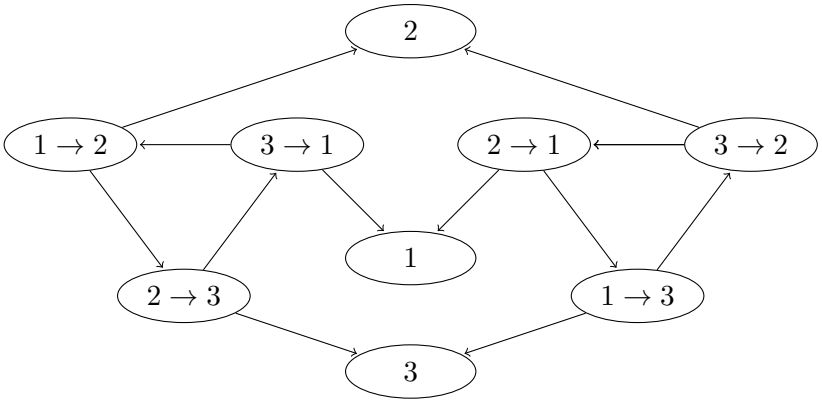


FIGURE 2.1 – Graphe des voisinages à distance 1 du 3-chemin orienté.

Nous considérons ensuite qu'il y a autant de joueurs que de sommets dans les graphes de $\mathscr{H}$, et nous prenons l'ensemble des sous-graphes $\mathscr{G}$ sur lesquels les joueurs seront placés comme étant tout simplement identique à $\mathscr{H}$.

L'ensemble des sorties autorisées $\mathscr{Y}$ sera quant à elle l'ensemble de toutes les bonnes colorations de chaque sous-graphe.

Résultats de Linial. L'article de Linial comporte les résultats suivants (entre autres) :

1. Un graphe est q-coloriable en t rondes si et seulement si son graphe de voisinages à distance t est q-coloriable.
2. Pour les chemins et les cycles de longueur n, il faut prendre $t = \Omega(\log^*(n))$ afin que le nombre chromatique du graphe des voisinages à distance t soit 3.

Nous allons maintenant donner l'idée de la preuve de ces deux points :

1. Un graphe est q-coloriable en t rondes si et seulement si son graphe de voisinages à distance t est q-coloriable.

Démonstration. Notons qu'il s'agit d'une équivalence.

Commençons donc par démontrer la direction : si le graphe des voisinages à distance t d'un graphe d'une famille donnée est q-coloriable, alors ce graphe est q-coloriable en t rondes. Pour cela, rappelons le principe selon lequel t fois une ronde de communication est identique à une t-ronde de communication.

En effet, au bout d'une t-ronde de communication, chaque nœud connaîtra son t-voisinage, et il doit décider de sa couleur en fonction de ce t-voisinage. Mais une q-coloration du graphe de voisinage fournit justement une fonction allant des voisinages possibles de chaque nœud à une couleur parmi q couleurs possibles. Donc, si nous avons une q-coloration du graphe des voisinages à distance t d'un graphe d'une famille donnée, il est aisé d'en déduire une q-coloration pour n'importe quel graphe de cette famille.

Démontrons maintenant l'autre direction : si chaque graphe d'une famille donnée est q-coloriable en t rondes, alors le graphe des voisinages à distance t est q-coloriable aussi. Pour cela, notons qu'un sommet doit décider de sa couleur une fois qu'il connaît son t-voisinage. Autrement dit, il y a une fonction qui renvoie chaque t-voisinage sur une couleur parmi q, et on peut ainsi colorier le graphe des voisinages à distance t en appliquant cette fonction à chaque nœud de ce graphe (i.e. à chaque t-voisinage).

□

2. Pour les chemins et les cycles de longueur n, il faut prendre $t = \Omega(\log^*(n))$ afin que le nombre chromatique du graphe des voisinages à distance t soit 3.

Démonstration. Pour prouver ce point, Linial considère, pour un $n \in \mathbb{N}$ fixé, une famille de graphes paramétrés $(B_{s,n})_{s\in\mathbb{N}}$, proches du graphe de de Bruijn, dont les sommets et les arêtes sont comme suit :

$$V(B_s) = \{(v_1, \ldots, v_s) \mid v_i \in V_G \text{ et } v_i \neq v_j \text{ pour } i \neq j \in \{1, \ldots, s\}\}$$
$$E(B_s) = \{\{(v_1, \ldots, v_s), (v_2, \ldots, v_{s+1})\} \mid (v_1, \ldots, v_s), (v_2, \ldots, v_{s+1}) \in V(B_s), v_1 \neq v_{s+1}\}$$

Et où le paramètre n donne le nombre de v_i distincts possibles.

Son argument utilise également la notion de **line graph** $L(G)$ d'un graphe orienté G, dont nous rappelons la définition : il s'agit du graphe dont les sommets sont les arêtes de G, i.e. $V(L(G)) = E(G)$, et dont deux sommets sont adjacents si les arêtes correspondantes dans G partagent une extrémité.

Sa preuve repose alors sur les arguments suivants :

1. $B_{2t+1,n}$ est un sous-graphe de $\mathcal{N}_t(P_n)$, donc $\chi(B_{2t+1}) \leq \chi(\mathcal{N}_t(P_n))$ (cela reste vrai si l'on remplace P_n par C_n).
2. $B_{1,n}$ s'obtient à partir du graphe complet K_n en remplaçant chaque arête par deux arêtes orientées.
3. $B_{s+1,n} = L(B_{s,n})$.

4. $\chi(L(G)) \geq \log \chi(G)$

Le point (2) a pour conséquence que $\chi(B_{1,n}) = n$.

En itérant les points (3) et (4), on a que

$B_{s,n} = L^{s-1}(B_{1,n})$

et

$\chi(L^s(G)) \geq \log^s \chi(G)$

d'où finalement

$$\begin{aligned} \chi(B_{2t+1,n}) &= \chi(L^{2t}(B_{1,n})) \\ &\geq \log^{2t} \chi(B_{1,n}) \\ &= \log^{2t}(n) \end{aligned}$$

On utilise ceci et le point (1) pour conclure que $\log^{2t}(n) \leq \chi(\mathcal{N}_t(P_n))$.

Donc, pour avoir $\chi(\mathcal{N}_t(P_n)) = 3$, il faut prendre $2t \geq \log^*(n) - 1$.

□

2.4 La preuve de Linial : passage au quantique

La raison pour laquelle la preuve de Linial n'est plus valable lorsque l'on passe à un cadre quantique réside dans l'étape

Si un graphe est q-coloriable en t rondes, alors son graphe des voisinages à distance t est q-coloriable aussi.

Ou plutôt sa contraposée

Si, pour un graphe donné, son graphe des voisinages à distance t n'est pas q-coloriable, alors il existe une manière d'étiqueter ce graphe qui fera échouer un algorithme de q-coloration classique.

La raison fondamentale derrière cela est que dans le monde quantique, il est possible d'assigner des valeurs à des « morceaux » d'une totalité sans que

l'on puisse les « recoller » ensemble de façon cohérente. Des liens entre ce phénomène et les notions de pré-faisceau, de faisceau et de section globale en théorie des catégories ont été relevés dans [Man13]. Nous allons nous contenter d'illustrer ce que cela signifie à l'aide d'un autre « jeu », le **carré magique de Mermin-Peres**.

Carré magique de Mermin-Peres. Il s'agit d'un jeu à deux joueurs, Alice et Bob, ainsi qu'un arbitre. Alice et Bob peuvent discuter avant le début du jeu afin de convenir d'une stratégie, mais ils n'ont plus le droit de communiquer une fois que le jeu a commencé.

Le jeu est le suivant : il y a un grille 3×3. L'arbitre distribue un numéro de ligne à Alice et un numéro de colonne à Bob. Alice et Bob doivent chacun émettre trois valeurs binaires pour chaque case de leur ligne (resp. colonne) telles que le XOR de ces valeurs vaut 0 ou 1 selon le numéro de ligne ou de colonne. De plus, il y aura exactement une case qui sera un point d'intersection de la ligne donnée à Alice et la colonne donnée à Bob : la valeur émise par Alice pour cette case doit être la même que celle émise par Bob.

y_1	y_2	y_3
y_4	y_5	y_6
y_7	y_8	y_9

Contraintes sur les lignes :
$y_1 \oplus y_2 \oplus y_3 = 0$
$y_4 \oplus y_5 \oplus y_6 = 0$
$y_7 \oplus y_8 \oplus y_9 = 0$
Contraintes sur les colonnes :
$y_1 \oplus y_4 \oplus y_7 = 0$
$y_2 \oplus y_5 \oplus y_8 = 0$
$y_3 \oplus y_6 \oplus y_9 = 1$

FIGURE 2.2 – Carré magique de Mermin-Peres.

Récapitulons : il y a une grille 3x3. On donne à Alice et à Bob chacun un morceau de cette grille (une ligne resp. une colonne), et ils doivent remplir le morceau de la grille qui leur est donné, et ce morceau seulement, tel que certaines contraintes sont vérifiées.

Carré magique de Mermin-Peres : un exemple

Si l'arbitre distribue la ligne numéro 1 à Alice et la colonne numéro 3 à Bob

- Alice doit émettre des valeurs pour les bits suivants : y_1, y_2, y_3, et ces valeurs doivent vérifier $y_1 \oplus y_2 \oplus y_3 = 0$.
- Bob doit émettre des valeurs pour les bits suivants : y_3, y_6, y_9, et ces valeurs doivent vérifier $y_3 \oplus y_6 \oplus y_9 = 1$.
- Alice et Bob doivent émettre la même valeur pour y_3.

Par exemple, nous voyons que $(y_1, y_2, y_3) = (0, 0, 0)$ et $(y_3, y_6, y_9) = (0, 0, 1)$ constitue une solution correcte.

Regardons maintenant ce qui se passe lorsque les sorties d'Alice et Bob sont fonction déterministe de leurs entrées. Pour une ligne donnée à Alice, il y a trois entrées possibles pour Bob, et pour chacune de ces trois colonnes en entrée, il faut qu'Alice et Bob émettent la même valeur pour la case qui est le point d'intersection de leur ligne et leur colonne. Autrement dit, pour chaque ligne donnée à Alice, sa solution va fixer les valeurs des trois cases de cette ligne (il en va de même pour les colonnes données à Bob). Ainsi, Alice et Bob peuvent gagner à ce jeu à l'aide d'une fonction déterministe si et seulement s'il est possible de remplir toutes les cases de la grille avec une valeur binaire précise de sorte que les contraintes soient vérifiées. Or, il est simple de voir que ceci est impossible : en effet, si l'on XOR les trois contraintes sur les lignes ensemble, nous obtenons :

$$y_1 \oplus y_2 \oplus y_3 \oplus y_4 \oplus y_5 \oplus y_6 \oplus y_7 \oplus y_8 \oplus y_9 = 0$$

Tandis que les contraintes sur les colonnes donnent :

$$y_1 \oplus y_2 \oplus y_3 \oplus y_4 \oplus y_5 \oplus y_6 \oplus y_7 \oplus y_8 \oplus y_9 = 1$$

Ce qui est contradictoire.

Plus généralement, on peut remplir toutes les cases sauf une de sorte à remplir les contraintes. Ainsi, il est possible de gagner pour 8 entrées sur les 9 possibles. Le fait de permettre à Alice et Bob d'utiliser de l'aléa classique leur permettra de randomiser la case qui pose problème (leur permettant de résister ainsi à un arbitre malveillant qui leur donnerait exprès les entrées qui

feront échouer leur stratégie en les faisant tomber sur cette case), mais cela ne leur permettra pas de gagner à coup sûr.

Nous venons de voir qu'une stratégie classique pour Alice et Bob consisterait à se mettre d'accord sur une grille remplie d'avance d'entiers binaires, ou éventuellement sur plusieurs grilles s'ils disposent d'une source d'aléa partagée pour en choisir une au moment où il faut jouer. Cependant, il est impossible de remplir la grille de sorte à vérifier toutes les contraintes, et aucune stratégie classique ne permet de gagner avec une probabilité meilleure que 8/9.

Or, il existe une stratégie quantique permettant de gagner à ce jeu avec probabilité 1 : dans cette stratégie, on a préalablement distribué un état quantique particulier à Alice et Bob, et la grille est pré-remplie non pas avec des entiers binaires, mais avec des **opérateurs de Pauli** : ce sont des matrices correspondant à des mesures différentes qu'il est possible de faire sur l'état quantique partagé.

La valeur à mettre dans chaque case est le résultat obtenu en mesurant cet état à l'aide de l'opérateur de Pauli indiqué dans la case en question. Puisque le résultat d'une mesure quantique est aléatoire, les valeurs obtenues seront différentes d'un déroulement du jeu à un autre. Cependant, les propriétés algébriques vérifiées par opérateurs de Pauli tels qu'ils sont mis dans les cases garantissent que les contraintes seront toujours vérifiées. De plus, la grille est remplie d'opérateurs de Pauli, mais à chaque déroulement de l'algorithme nous n'obtenons que les valeurs binaires pour la ligne et la colonne qui ont été données en entrée.

Lien avec la preuve de Linial. Expliquons maintenant le lien avec le jeu de coloration du graphe des voisinages dans la preuve de Linial : le premier résultat de Linial repose essentiellement sur le même type de raisonnement que nous avons fait pour le cas d'une solution déterministe au carré magique de Mermin-Peres, à savoir le fait que les solutions locales (i.e. le remplissage d'une ligne/colonne dans le cas du carré magique, et la coloration d'un graphe étiqueté d'une façon particulière dans le cas du graphe des voisinages) doivent pouvoir être « recollées » de façon cohérente afin de fournir la solution globale. En ce qui concerne une solution probabiliste (de type Las Vegas), cela revient essentiellement, comme pour le carré magique, à utiliser une source d'aléa partagé afin de choisir l'une des solutions déterministes.

Enfin, dans le cas d'une solution quantique, on aura distribué au préalable un état quantique aux joueurs, qui doivent se servir d'une mesure sur cet état quantique afin de déterminer la couleur. Or, la mesure en question va dépendre (et ne va dépendre que) de ce que le joueur voit au bout de t rondes de

communication, autrement dit, de son t-voisinage. Autrement dit, une solution quantique consisterait à remplir le graphe des voisinages à distance t à l'aide de mesures quantiques appropriées qui :

- donneront un résultat correct à chaque déroulement de l'algorithme.
- donneront un résultat aléatoire, et donc différent, à chaque déroulement de l'algorithme.
- ne donneront un résultat que pour les voisinages qui sont effectivement vus par les joueurs, i.e. même si l'on a assigné une mesure à chaque nœud du graphe des voisinages, seule la partie de ce graphe qui est vue par les joueurs sera colorée.

CHAPITRE 3
LE MODÈLE ϕ-LOCAL

Dans cette section, nous allons décrire le modèle ϕ-LOCAL proposé par [GKM09] comme adaptation de la condition non-signalling au modèle LOCAL.

3.1 Motivation et définition

Il existe des versions quantiques de nombreux algorithmes fondamentaux en calcul distribué : par exemple, pour le problème des généraux byzantins [BOH05], où il s'agit de se mettre d'accord sur la valeur d'un bit malgré la présence de comportements byzantins chez certains nœuds, ou encore pour le problème de l'élection du leader [TKM12] qui a été utilisé récemment afin de développer un algorithme de calcul de diamètre quantique plus efficace que le meilleur algorithme classique possible dans le modèle CONGEST($\log(n)$) [LGM18].

Néanmoins, comme nous l'avions constaté dans le chapitre précédent en étudiant le rapport entre le fait de trouver une solution quantique à un problème et le fait de trouver une boîte non-signalling qui permet de le résoudre, il faudrait regarder une infinité d'états quantiques pour voir si l'un d'entre eux permet de résoudre le problème en question, tandis que pour un nombre fixé de joueurs, d'entrées et de sorties, l'ensemble des boîtes non-signalling forme un polytope convexe avec un nombre fini de points extrémaux. Dès lors, si aucun état quantique ne se présente comme étant particulièrement utile pour la résolution d'un problème, il est bien plus simple de regarder s'il existe une solution non-signalling à ce problème meilleure que sa solution classique : si ce n'est pas le cas, a fortiori aucune solution quantique ne peut battre la solution classique. De même, si l'on cherche à étudier si l'accès à des ressources quantiques permettrait de résoudre tel problème dans le modèle LOCAL avec moins de rondes de communications que si l'on ne disposait que de ressources classiques, il faudrait a priori regarder un nombre infini d'états quantiques. Il serait donc intéressant de trouver un équivalent aux boîtes non-signalling qui permettraient de décider plus facilement de la possibilité d'un éventuel avantage quantique. Le modèle ϕ-LOCAL a donc été proposé pour trouver des bornes inférieures sur le nombre de rondes nécessaires pour résoudre un problème à l'aide de ressources non-signalling et a fortiori à l'aide de ressources quantiques.

Revenons maintenant dans le cadre du jeu sur le graphe des voisinages défini dans la Section 2.3. On considère donc une famille de graphes étiquetés (orientés ou non) $\mathscr{H}$, t un entier, G le graphe des voisinages à distance t de cette famille de graphes. Pour un graphe $H \in \mathscr{H}$ et un sommet $v \in V$, on dénote par $N_t(H,v)$ ce que voit le sommet d'identifiant v dans un rayon de taille t autour de lui dans le graphe H, i.e. son t-voisinage dans ce graphe. On considère une version du jeu où les sorties sont probabilistes, i.e. où chaque nœud v émet une sortie y_v avec une certaine probabilité en fonction de son entrée qui est $N_t(H,v)$. On dit alors que la distribution des sorties en fonction des entrées est ϕ-LOCAL(t) si pour chaque sous-ensemble de sommets de V, la distribution de leur sortie est fonction seulement de l'union de leurs entrées et est indépendante des entrées des autres sommets. Mathématiquement :

Définition 1. *On dit que qu'une distribution sur les sorties $(y_v)_{v\in V}$ conditionnée par les entrées $(N_t(H,v))_{v\in V}$ est ϕ-**LOCAL(t)** si*

pour tout

$S \subset V$

$H' \neq H \in \mathscr{H}$ *vérifiant*

$\mathscr{N} := \{N_t(H,u)\}_{u\in S} = \{N_t(H',u)\}_{u\in S}$

en notant

$\mathscr{N}_1 = \{N_t(H,v)\}_{v\in V\setminus S}$

$\mathscr{N}_2 = \{N_t(H',v)\}_{v\in V\setminus S}$

on a :

$$\sum_{y_{\overline{S}}} \Pr(Y_S = y_S, Y_{\overline{S}} = y_{\overline{S}} \mid X_S = \mathscr{N}, X_{\overline{S}} = \mathscr{N}_1)$$

$$=$$

$$\sum_{y_{\overline{S}}} \Pr(Y_S = y_S, Y_{\overline{S}} = y_{\overline{S}} \mid X_S = \mathscr{N}, X_{\overline{S}} = \mathscr{N}_2)$$

Notons que sous les conditions de cette définition, on a

$$\bigcup_{N\in\mathscr{N}\cup\mathscr{N}_1} N = H$$

et

$$\bigcup_{N\in\mathscr{N}\cup\mathscr{N}_2} N = H'.$$

La 2-coloration dans le modèle ϕ-LOCAL. Dans [GKM09], un premier résultat dans ce modèle a été démontré au sujet de la 2-coloration d'un cycle non-orienté. Il est aisé d'étendre le raisonnement derrière ce résultat afin d'obtenir un résultat analogue pour les n-chemins non-orientés. Nous allons donc énoncer le résultat prouvé dans [GKM09] puis son analogue pour les chemins, et ensuite prouver ce dernier résultat.

Tout d'abord, on rappelle que (dans le cas où le nombre de sommets est connu des nœuds), il faut au moins $n/2-1$ rondes de communications pour colorier un n-cycle non-orienté (où n est pair) avec 2 couleurs, et au moins $n-1$ rondes de communications pour colorier un n-chemin non-orienté avec 2 couleurs : en effet, pour un nombre de rondes de communications inférieur, le graphe de voisinages correspondant comporte un cycle de longueur impair et n'est donc pas 2-coloriable [Lin92].

Or

- Un n-cycle non-orienté (avec n pair) n'est pas 2-coloriable dans le modèle ϕ-LOCAL(t) pour $t \leq n/4-1$.
- Un n-cycle non-orienté (avec n pair) est 2-coloriable dans le modèle ϕ-LOCAL($n/4$).

Comme on a dit, ces résultats ont été prouvés dans [GKM09]. Voici leurs analogues pour le n-chemin :

- Un n-chemin non-orienté n'est pas 2-coloriable dans le modèle ϕ-LOCAL(t) pour $t \leq \lfloor n/3 \rfloor - 1$.
- Un n-chemin non-orienté est 2-coloriable dans le modèle ϕ-LOCAL($\lfloor n/3 \rfloor$).

Nous allons commencer par prouver un résultat simple utile pour la preuve de ces points avant de les prouver un par un.

> Toute distribution autre que la distribution uniforme sur les deux 2-colorations possibles du n-chemin ne peut pas être ϕ-LOCAL(t), pour $t \leq \lfloor n/2-1 \rfloor$.

Démonstration. Considérons la distribution des sorties en fonction des entrées d'un seul sommet. Pour $t \leq \lfloor n/2-1 \rfloor$, il existe au moins un sommet qu'il ne voit pas dans son voisinage. Or, si la distribution des entrées n'est pas uniforme sur les deux 2-colorations possibles du n-chemin, l'une ou l'autre couleur est plus probable pour un sommet en fonction de sa distance à une extrémite ou l'autre du chemin, et va donc révéler des informations à ce sujet.

□

Un n-chemin non-orienté n'est pas 2-coloriable dans le modèle ϕ-LOCAL(t) pour $t \leq \lfloor n/3 \rfloor - 1$.

Démonstration. Nous allons montrer que ce n'est pas le cas pour $t = \lfloor n/3 \rfloor - 1$ et ce sera a fortiori vrai pour des valeurs plus faibles de t.

Pour cela, nous considérons deux t-voisinages N_1 et N_2 avec les spécificités suivantes :

1. N_1 est un chemin de longueur $t+1$, i.e. il est clair qu'il s'agit du t-voisinage d'un des sommets à son extrémité qui est aussi à l'une des extrémités du n-chemin de base (sans perte de généralité, nous supposerons qu'il s'agit du sommet « à gauche » dans le dessin du graphe).
2. N_2 est un chemin de longueur $2t+1$, c'est donc le t-voisinage du sommet au milieu qui est aussi un sommet à distance au moins t des extrémités du n-chemin de base.
3. N_1 et N_2 n'ont aucun sommet en commun.

On voit donc un total de $(t + 1) + (2t + 1)$ sommets dans $N_1 \cup N_2$, soit $3\lfloor n/3 \rfloor - 1$ sommets, donc il y a au moins un sommet u que l'on ne voit pas dans ces deux voisinages. Alors, on peut « compléter » ces vues partielles de deux façons afin d'obtenir un n-chemin.

1. Soit on insère le sommet manquant u entre N_1 et N_2, auquel cas on tombe sur le chemin $N_1 - u - N_2$. Dans ce cas, il y aura $t+1+t$ sommets entre le premier sommet de N_1 et le sommet au milieu de N_2, ces deux sommets seront donc coloriés avec la même couleur.
2. Soit on insère le sommet manquant u entre N_1 et N_2, auquel cas on tombe sur le chemin $N_1 - N_2 - u$. Dans ce cas, il y aura $t + t$ sommets entre le premier sommet de N_1 et le sommet au milieu de N_2, ces deux sommets seront donc coloriés avec deux couleurs différentes.

Le raisonnement est valable dans le sens inverse. Par conséquent, si les deux sommets en question voient qu'ils ont la même couleur, ils savent qu'ils sont dans le premier cas, et s'ils voient qu'ils ont des couleurs différentes, ils savent qu'ils sont dans le deuxième cas. Par conséquent, la distribution des sorties va toujours dépendre de, et donc révéler des informations sur quelque chose qui n'est pas dans l'union des t-voisinages N_1 et N_2.

Nous allons maintenant justifier les points suivants :

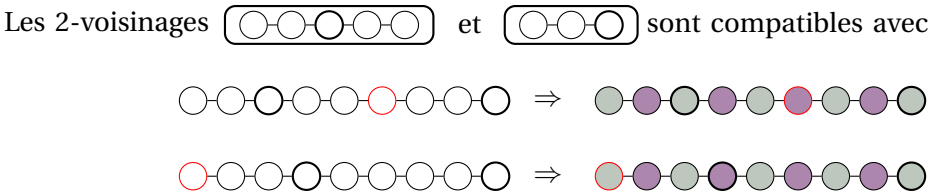


FIGURE 3.1 – Illustration dans le cas $n = 9$.

- Pourquoi avons-nous choisi N_1 et N_2 de cette manière ?

  Remarquons que N_1 est le voisinage le plus minimal possible et N_2 le voisinage le plus maximal possible. Si N_1 était plus grand, nous verrions plus de sommets, par conséquent nous aurions plus d'informations sur le n-chemin complet et le problème serait plus facile à résoudre. Par contre, une fois N_1 fixé de la sorte, nous avons intérêt à prendre N_2 le plus grand possible : en effet, nous savons que le premier sommet de N_1 est à une extrémité du n-chemin. Or, si nous avions pris N_2 plus petit, cela signifierait que nous connaîtrions la distance à l'autre extrémité du n-chemin du sommet dont c'est le voisinage. Par conséquent, même si nous ne connaissons pas les identifiants des sommets qui ne sont ni dans N_1, ni dans N_2, nous savons à quelle distance se trouvent les deux sommets dont c'est le voisinage et donc s'ils auront la même couleur ou non.

  En réalité, dans ce dernier cas, nous pourrions penser que même si l'identité ou la différence des couleurs n'est susceptible de révéler aucune information, le fait que les sommets aient une couleur plutôt que l'autre pourrait révéler des informations sur les sommets que nous ne voyons pas, car les deux couleurs auraient des probabilités différentes d'apparaître. Mais comme nous le verrons dans la preuve du prochain point, ce problème ne se pose pas car, comme nous l'avons vu, seule la distribution uniforme sur les deux 2-colorations possibles du n-chemin peut être ϕ-LOCAL, donc ce cas de figure n'arrive pas.

- Pourquoi avons-nous choisi de regarder les voisinages de deux sommets seulement et pas plus ?

  Pour une raison analogue à celle qui dit que pour une boîte non-signalling à n joueurs, il suffit de regarder les plus grandes coalitions (cf. equation 1.5) : nous pouvons en effet montrer que les conditions sur les petites coalitions découlent des conditions sur les plus grandes. Donc, en utilisant la contraposée de ce résultat, si les conditions ne sont pas vérifiée pour la plus petite coalition possible, i.e. celle à deux joueurs, elles ne sont a fortiori pas vérifiées pour les coalitions plus grandes.

□

Un n-chemin non-orienté est 2-coloriable dans le modèle ϕ-LOCAL($\lfloor n/3 \rfloor$).

Démonstration. Il y a deux 2-colorations possibles du n-chemin complet. Nous pouvons montrer que la distribution qui choisit uniformément l'une de ces deux 2-colorations est ϕ-LOCAL($\lfloor n/3 \rfloor$).

En effet, il suffit de considérer le cas de deux voisinages : s'il y a plus de voisinages, on aura plus d'informations et le problème sera a fortiori plus simple à résoudre.

Donc, pour deux voisinages N_1, N_2, nous avons deux cas de figure :

- Soit les deux voisinages comportent moins de $2t + 1$ sommets, auquel cas nous connaissons la distance à l'extrémité des sommets dont ce sont les voisinages, par conséquent leur couleur ne va révéler aucune information en dehors de ce qui peut être connu en regardant les voisinages.
- Soit au moins un voisinage comporte $2t + 1$ sommets et l'autre au moins $t + 1$ sommets, mais dans ce cas nous voyons au plus $3t + 2$ sommets dans ces voisinages, soit $3\lfloor n/3 \rfloor + 2$ sommets. Or, $3\lfloor n/3 \rfloor + 2 > n$, ce qui signifie que les deux voisinages se chevauchent, et les deux sommets voient à quelle distance ils sont l'un par rapport à l'autre (même si nous ne voyons pas le n-chemin complet). Il n'y a donc aucune information extérieure qui puisse être révélée par la distribution des couleurs.

□

3.2 Consistance locale et consistance globale

Le but de cette section est de préparer une clarification conceptuelle au sujet du modèle ϕ-LOCAL en soulignant quelques subtilités dans l'interaction entre la spécification d'un problème distribué, ce problème « en situation réelle », et la solution de ce problème.

Consistance locale et consistance globale. Nous commencerons par introduire une distinction entre entre ce que nous appelons la **consistance locale** et la **consistance globale** : cette distinction a été relevée lors de discussions avec Pierre Fraigniaud et Frédéric Magniez. En particulier, l'exemple que nous utilisons afin de l'illustrer est dû à Pierre Fraigniaud.

Il s'agit d'une situation qui émerge lorsque nous considérons d'un côté la spécification théorique d'un problème distribué et de l'autre ce qui peut se passer dans les faits lorsque nous avons affaire à ce problème : il s'agit de la raison d'être des exceptions dans les langages de programmation. Plus précisément, nous donnons toujours dans la spécification d'un problème l'ensemble des entrées autorisées, mais il est possible, en pratique, d'avoir une entrée autre que celles qui sont autorisées, et il faut gérer cette situation de manière adéquate. Si dans le cadre du calcul centralisé, l'entrée est soit dans l'ensemble des entrées autorisées, soit elle n'y est pas, c'est un peu plus subtil dans le cadre du calcul distribué : il y a trois situations possibles, que nous appellerons la consistance globale, la consistance locale et « l'absence de contraintes ». La consistance globale correspond au cas où les entrées possibles sont exactement celles autorisées par la spécification du problème, « l'absence de contraintes » correspond au cas de figure où les entrées sont n'importe quoi, et la consistance locale est intermédiaire entre les deux.

Nous allons expliquer cela plus adéquatement à l'aide de l'exemple suivant :

Exemple (consistance globale et consistance locale)

Nous considérons un jeu à deux joueurs, Alice et Bob, où le but est de résoudre le problème dont la spécification est la suivante :

- Nous donnons en entrée x_A à Alice et x_B à Bob avec la condition $(x_A, x_B) \in \{(0,0),(1,1)\}$.
- Alice et Bob doivent sortir $(y_A, y_B) = (x_A + x_B, x_A + x_B)$.

- **Consistance globale** : les entrées possibles sont exactement celles autorisées par la spécification du problème. Autrement dit, si Alice a 0 (resp. 1), elle sait que Bob a 0 (resp. 1) et inversement. Dans ce cas, chacun sort 0 s'il voit 0, et 2 s'il voit 1.
- **Consistance locale** : les entrées possibles sont celles telles que Alice et Bob, localement, ne peuvent pas détecter de problème. Autrement dit, les entrées possibles sont $\{(0,0),(0,1),(1,0),(1,1)\}$. Dans ce cas, si Alice voit 0, elle ne sait pas si l'entrée globale est légale, c'est-à-dire si c'est $(0,0)$, ou bien illégale, c'est-à-dire $(0,1)$. Il en va de même pour Bob. Mais puisque les entrées illégales ne font pas partie de la spécification du problème, Alice et Bob ne sont pas dans l'obligation de sortir la « bonne » réponse qui leur correspond. Autrement dit, si l'entrée globale est $(0,1)$, Alice et Bob ne sont pas dans l'obligation de sortir $(1,1)$, même si c'est bien la somme de leurs entrées, puisque leurs entrées ne font pas partie des entrées légales du problème.
- **« L'absence de contraintes »** : n'importe quelle entrée est autorisée. Par exemple, nous pouvons donner $(-1,-1)$ à Alice et Bob, voire

(« grenouille », « grenouille »). Mais alors Alice et Bob seront en mesure de détecter chacun localement qu'il y a un problème. Et même si l'entrée d'un des deux partis est consistante avec la spécification (par exemple, si nous leur donnons $(1, -1)$), au moins l'un d'entre eux détectera un problème.

Nous voyons que la différence la plus intéressante est celle entre la consistance globale et la consistance locale, et si nous avons également regardé le cas de « l'absence de contraintes », c'était principalement dans le but de souligner la différence avec la condition de consistance locale.

Rapport avec le quantique et le non-signalling. Le problème que nous avons utilisé à titre d'exemple admet une solution classique (chacun sort 0 s'il voit 0, et 2 s'il voit 1). Néanmoins, regardons ce qui se passerait si les joueurs décidaient de recourir à des ressources quantiques (ou non-signalling) pour résoudre le problème (remarquons que, puisqu'il existe une solution classique permettant de gagner à coup sûr, il existe aussi des solutions quantiques et non-signalling permettant de le faire). Chaque entrée localement consistante avec la spécification (donc, 0 ou 1) correspond à une mesure spécifique sur la ressource en question.

- Dans une situation de consistance globale, chaque joueur fait la mesure qu'il doit faire en fonction de son entrée, et le résultat de la mesure donnera la bonne sortie.
- Dans une situation de consistance locale, encore une fois, chaque joueur fait la mesure qu'il doit faire en fonction de son entrée. Si l'entrée globale est légale, le résultat de la mesure donnera la bonne sortie. Mais si l'entrée globale n'est pas légale (par exemple, si elle est $(0, 1)$), il n'y a pas de « bonne sortie » que les mesures effectuées par les deux joueurs doivent donner. Néanmoins, il y aura une sortie (y_A, y_B), et elle sera le résultat d'une mesure sur une ressource quantique ou une ressource abstraite non-signalling. Pour cette raison, cette sortie ne peut pas être n'importe quoi : en particulier, la distribution des sorties conditionnées par les entrées doit vérifier la condition non-signalling.
- Dans une situation où il n'y a pas de contrainte, chaque joueur peut émettre un message d'erreur si son entrée locale ne correspond pas à la spécification : en effet, seules les entrées localement légales correspondent à une mesure sur la ressource utilisée pour résoudre le problème.

Spécification et solution d'un problème. Une autre subtilité reliée qui mérite d'être soulignée est la suivante : un problème sur n joueurs peut très bien être spécifié à l'aide d'une fonction partielle (c'est-à-dire, qui ne doit sortir des valeurs que pour un sous ensemble $S \subset E_1 \times \ldots \times E_n$, où E_i est l'ensemble des entrées localement autorisées pour le joueur i) : cela n'empêche qu'il faut, dans la solution de type boîte non-signalling, spécifier les sorties pour toutes les entrées. Autrement dit, nous pouvons avoir affaire à une fonction partielle au niveau de la **spécification du problème**, mais il faut que la **solution** proposée assigne des valeurs de sortie pour n'importe quelle entrée possible. Illustrons cela à l'aide du jeu suivant [1] :

Guess Your Neighbour's Input (GYNI) [ABB+10]

Il y a n joueurs $J_1, \ldots, J_n$. Le joueur i reçoit en entrée un bit $x_i \in \{0,1\}$, et doit sortir y_i tel que $y_i = x_{(i+1) \mod n}$.

Nous pouvons visualiser cela en supposant que les joueurs sont disposés dans le sens des aiguilles d'une montre sur un cycle, et chaque joueur doit deviner l'entrée du joueur à sa droite (d'où le nom du jeu, « Devinez l'entrée de votre voisin »).

Dans [ABB+10], les résultats suivants sont montrées :

- Si les entrées $(x_1, \ldots, x_n)$ sont données avec la probabilité suivante :

$$\begin{cases} \frac{1}{2^{n-1}} & \text{si} \quad x_1 \oplus \ldots \oplus x_n = 0 \text{ et } n \text{ est impaire} \\ \frac{1}{2^{n-1}} & \text{si} \quad x_1 \oplus \ldots \oplus x_{n-1} = 0 \text{ et } n \text{ est paire} \\ 0 \text{ sinon} \end{cases}$$

  Alors, le fait de disposer de ressources quantiques n'octroie aucun avantage par rapport à la stratégie classique. Néanmoins, il est possible de gagner avec une probabilité $p_{NS} = \frac{4}{3} p_C$, où p_C est la meilleure probabilité de réussite classique, à l'aide de deux boîtes non-signalling incommensurables (c'est-à-dire, telles que l'une ne peut pas être ramenée à l'autre en permutant ou inversant des entrées et des sorties).

  En particulier, pour chaque entrée des boîtes non-signalling, des probabilités sont spécifiées pour chaque sortie, et ce même si l'entrée n'est pas légale selon la spécification du problème.

- Si chaque entrée globale $(x_1, \ldots, x_n) \in \{0,1\}^n$ apparaît avec une probabilité uniforme $\frac{1}{2^n}$, alors ni les ressources quantiques, ni les ressources non-signalling ne permettent d'avoir un avantage par rapport à la stratégie classique.

1. Notons que ce jeu a été étudié dans un autre but que celui pour lequel nous l'évoquons ici.

GYNI, consistance locale et consistance globale. Le rapport entre ce que nous venons de dire sur le jeu GYNI, et la consistance locale et la consistance globale est assez subtil.

Nous pouvons avoir l'impression que le cas de figure où il y a un avantage non-signalling correspond à une situation de consistance globale : en effet, il y a une distribution sur les entrées qui en interdit certaines. Néanmoins, comme nous l'avons dit, au niveau de la solution qui est proposée, une distribution de probabilités sur les sorties est donnée pour chaque entrée, même celles qui sont illégales. Par conséquent, contrairement à ce que nous pourrions penser de prime abord, il s'agit bel et bien d'une situation de consistance locale. Cela aurait été une situation de consistance globale si seules des distributions de sorties conditionnées par les entrées légales avaient été données au niveau de la **solution** proposée.

3.3 Limites du modèle ϕ-LOCAL

Nous allons maintenant utiliser ce qui a été dit dans la section précédente afin de souligner en quoi le modèle ϕ-LOCAL n'est pas tout à fait adéquat.

Nous rappelons d'abord la définition d'une distribution de probabilités qui est ϕ-LOCAL(t) pour un $t \in \mathbb{N}$:

Définition 1. *On dit que qu'une distribution sur les sorties $(y_v)_{v\in V}$ conditionnée par les entrées $(N_t(H,v))_{v\in V}$ est ϕ-**LOCAL(t)** si*

pour tout

$S \subset V$

$H' \neq H \ \in \mathscr{H}$ *vérifiant*

$\mathscr{N} := \{N_t(H,u)\}_{u\in S} = \{N_t(H',u)\}_{u\in S}$

en notant

$\mathscr{N}_1 = \{N_t(H,v)\}_{v\in V\setminus S}$

$\mathscr{N}_2 = \{N_t(H',v)\}_{v\in V\setminus S}$

on a :

$$\sum_{y_{\overline{S}}} \Pr(Y_S = y_S, Y_{\overline{S}} = y_{\overline{S}} \mid X_S = \mathscr{N}, X_{\overline{S}} = \mathscr{N}_1)$$

$$=$$

$$\sum_{y_{\overline{S}}} \Pr(Y_S = y_S, Y_{\overline{S}} = y_{\overline{S}} \mid X_S = \mathscr{N}, X_{\overline{S}} = \mathscr{N}_2)$$

Nous avions remarqué que cette définition signifie en particulier que

$$\bigcup_{N \in \mathscr{N} \cup \mathscr{N}_1} N = H$$

et

$$\bigcup_{N \in \mathscr{N} \cup \mathscr{N}_2} N = H'$$

Il s'agit d'une forme de consistance globale imposée sur les entrées. En effet, regardons ce que donnent les trois situations définies dans la section précédente (i.e. la consistance globale, la consistance locale et « l'absence de contraintes ») dans le cas d'un problème distribué sur les graphes où l'entrée de chaque nœud est son t-voisinage :

- **Consistance globale** : les entrées possibles sont exactement celles autorisées par la spécification du problème. Souvent, la spécification du problème indique que les nœuds forment un graphe d'une famille de graphes donnée, même si individuellement ils peuvent ne pas savoir de quel graphe de cette famille il s'agit précisément.

  Donc, par exemple, si la famille des graphes dans la spécification du problème est la famille des chemins d'une certaine longueur $n > 5$, et que le nœud d'identifiant 1 voit le 2-voisinage « 5-2-1-3-4 », il est sûr que l'entrée du nœud d'identifiant 2 est un 2-voisinage de la forme «?-5-2-1-3 ». Un autre 2-voisinage du nœud d'identifiant 2, même s'il est un « bon » 2-voisinage en soi, ne serait pas compatible avec le 2-voisinage qu'a en entrée le nœud d'identifiant 1 et la contrainte dans la spécification du problème que les nœuds forment un n-chemin.

- **Consistance locale** : les entrées possibles sont celles telles que chaque nœud, localement, ne peut pas détecter de problème. Ceci signifie que chaque nœud v voit un sous-graphe d'un graphe de la famille de graphes dans la spécification du problème tel que ce sous-graphe est un bon t-voisinage pour le nœud en question, i.e. dans le sous-graphe en question :
  - il y a tous les nœuds à distance au plus t de v, et il n'y a aucun nœud à distance plus de t de v.
  - il n'y a aucune répétition des identifiants : chaque nœud a un identifiant distinct de tous les autres.

  Autrement dit, chaque nœud voit ce qui est localement un voisinage qui est compatible avec l'une des entrées globales autorisées par la

spécification, même si l'union des voisinages peut ne correspondre à aucune des entrées globales autorisées par la spécification.

Chaque nœud doit alors répondre comme si l'ensemble des voisinages formait un graphe global de la famille donnée dans la spécification, même si ce n'est pas le cas, car il n'est dans l'obligation de fournir une bonne réponse que si l'entrée globale est légale.

- « L'absence de contraintes » : n'importe quelle entrée est autorisée. Un nœud peut avoir en entrée un « voisinage » comportant des identifiants double, un s-voisinage pour un $s \neq t$, un sous-graphe d'un graphe de la famille de graphes donnée dans la spécification qui n'est pas un t-voisinage, etc. Mais dans ce cas, chaque nœud est en mesure de détecter localement qu'il y a un problème.

Conclusion : une version plus faible du modèle ϕ-LOCAL. Par conséquent, le modèle ϕ-LOCAL proposé dans [GKM09] est trop fort puisqu'elle impose une consistance globale sur les entrées. Une version plus faible de ϕ-LOCAL, compatible seulement avec la consistance locale, serait la suivante :

Définition 2. *Nous disons que qu'une distribution sur les sorties $(y_v)_{v\in V}$ conditionnée par les entrées $(N_t(H,v))_{v\in V}$ est ϕ-**LOCAL(t) faible** si*

pour tout

$S \subset V$

$H_{v_i}, H'_{v_i} \in \mathscr{H}$ *non tous identiques pour* $i = 1, \ldots, |V \setminus S|$

$\mathscr{N} := \{N_t(H,u)\}_{u\in S}$

en notant

$\mathscr{N}_1 = \{N_t(H_{v_i}, v_i)\}_{v_i \in V\setminus S}$

$\mathscr{N}_2 = \left\{N_t(H'_{v_i}, v_i)\right\}_{v_i \in V\setminus S}$

on a :

$$\sum_{y_{\overline{S}}} \Pr(Y_S = y_S, Y_{\overline{S}} = y_{\overline{S}} \mid X_S = \mathscr{N}, X_{\overline{S}} = \mathscr{N}_1)$$

$$=$$

$$\sum_{y_{\overline{S}}} \Pr(Y_S = y_S, Y_{\overline{S}} = y_{\overline{S}} \mid X_S = \mathscr{N}, X_{\overline{S}} = \mathscr{N}_2)$$

Toute distribution qui est ϕ-LOCAL(t) faible est ϕ-LOCAL(t), puisque si les conditions d'égalité des probabilités sont valables pour tous les H_{v_i} et H'_{v_i},

elles le sont a fortiori dans le cas particulier où il existe un $H' \in \mathscr{H}$ vérifiant $\{N_t(H,u)\}_{u\in S} = \{N_t(H',u)\}_{u\in S}$ et tel que $H_{v_i} = H, H'_{v_i} = H' \; \forall i$.

Par contre, toute distribution ϕ-LOCAL(t) n'est pas nécessairement ϕ-LOCAL(t) faible : il faut en effet vérifier qu'elle peut s'étendre de façon cohérente si nous autorisons également des t-voisinages $\mathscr{N}_1$ et $\mathscr{N}_2$ en entrée qui ne forment pas nécessairement un graphe de $\mathscr{G}$ pris ensemble avec le t-voisinage $\mathscr{N}$.

CHAPITRE 4
AUTRES RÉSULTATS

Le but de ce chapitre est d'approfondir davantage certaines notions présentées dans les chapitres précédents et de présenter des résultats obtenus en lien avec ces notions, et qui ne pouvaient pas être présentés dans les chapitres précédents sans casser le rythme du texte.

4.1 Jeu multijoueur sur les graphes

Dans cette section, nous allons revenir sur le jeu présenté dans la Section 2.3 et regarder des cas particuliers de ce jeu autre que le graphe des voisinages. Nous allons d'abord montrer que ce jeu permet de subsumer des notions existantes dans la littérature. Ensuite, nous présenterons des résultats numériques obtenus pour un cas particulier de ce jeu.

Commençons donc par rappeler la définition du jeu.

Définition du jeu : rappel. Il s'agir d'un jeu défini sur un graphe $G = (V_G,\ E_G)$. Nous noterons m le nombre de joueurs, $N := |V|$ le nombre de sommets du graphe, et J_i le joueur numéro i pour $1 \leq i \leq m$.

Il y a un arbitre dont le rôle est de placer les m joueurs sur $s \leq m$ sommets du graphe G (donc plusieurs joueurs peuvent être placés sur le même sommet). Les joueurs doivent résoudre un problème P (qui peut différer d'une version du jeu à l'autre). Pour cela, ils peuvent communiquer et convenir d'une stratégie avant le début du jeu, mais ils n'ont plus le droit de communiquer une fois qu'ils sont placés sur les sommets du graphe.

Plus précisément, nous donnons

- un graphe $G = (V_G, E_G)$
- un ensemble $\mathscr{G}$ de sous-graphes de G de taille $\leq m$
- pour chaque sous-graphe $H \in \mathscr{G}$, un ensemble $\Pi(H) = \{\pi : \{J_1, \ldots, J_m\} \to V_H\}$ de **fonctions de placement** des joueurs (ces fonctions étant surjectives mais non nécessairement injectives).

- un ensemble $\mathscr{Y} = \{Y_{H,\pi} \mid H \in \mathscr{G}, \pi \in \Pi(H)\}$ de sorties autorisées pour chaque sous-graphe H et chaque placement de joueurs sur ce sous-graphe.

1. Au moment où le jeu commence, l'arbitre choisit un sous-graphe $H \in \mathscr{G}$ et une fonction de placement $\pi \in \Pi$, et place chaque joueur J_i sur le sommet $\pi(J_i)$ du sous-graphe H.
2. Chaque joueur i doit sortir une valeur y_i telle que $(y_1, \ldots, y_m) \in Y_{H,\pi}$.

Cas particulier : le nombre chromatique quantique [CMN+06]. Le nombre chromatique d'un graphe G admet plusieurs caractérisations : par exemple, le plus petit q tel qu'il existe un homomorphisme de graphes $G \to K_q$, où K_q est le graphe complet non-orienté sur q sommets.

Nous allons nous intéresser à la caractérisation suivante à l'aide d'un jeu distribué à deux joueurs et un arbitre (il est possible de montrer qu'elle est équivalente à la caractérisation à l'aide d'un homomorphisme) : il s'agit d'un cas particulier du jeu ci-dessus sur le graphe G, avec $m = 2$ et $s \in \{1, 2\}$: en effet, l'arbitre doit placer les deux joueurs sur deux sommets distincts de G ou un même sommet de G. Les joueurs doivent alors sortir une couleur parmi q couleurs possibles, avec les contraintes suivantes :

- Si $s = 1$, i.e. les deux joueurs sont placés sur le même sommet, alors ils doivent sortir la même couleur.
- Si $s = 2$, et les deux joueurs sont placés sur des sommets adjacents, alors ils doivent sortir des couleurs distinctes.
- Si $s = 2$, et les deux joueurs ne sont pas placés sur des sommets adjacents, il n'y a pas de contrainte sur les couleurs qu'ils peuvent sortir.

Alors, si les deux joueurs disposent de ressources classiques, il est possible de gagner à ce jeu (avec probabilité 1) si et seulement si $\chi(G) \leq q$.

En effet, si $\chi(G) \leq q$, alors il existe une q-coloration de G. Les deux joueurs ont alors à se mettre d'accord sur une q-coloration de G et choisir la couleur dans cette coloration du sommet qui leur est donné : cette solution vérifie les contraintes du problème. De même, si $\chi(G) > q$, cela signifie que G n'est pas q-coloriable. Or, s'il était possible de gagner au jeu dans ce cas, nous n'aurions qu'à « recoller » les solutions pour chaque sommet pour obtenir une q-coloration du graphe, ce qui contredit l'hypothèse de départ.

Or, il a été montré [CMN+06] que si les joueurs disposent d'une certaine ressource quantique, il existe des graphes pour lesquels nous pouvons gagner

au jeu (avec probabilité 1) pour un nombre $q < \chi(G)$ de couleurs. Le **nombre chromatique quantique** $\chi_q(G)$ est alors défini comme étant le plus petit nombre q de couleurs tel que nous pouvons gagner au jeu avec q couleurs sur le graphe G si les joueurs disposent de ressources quantiques [1].

Notre jeu subsume donc le jeu utilisé pour définir le nombre chromatique quantique. De plus, le résultat cité ci-dessus montre qu'il existe déjà un exemple d'un tel jeu où la solution quantique bat la meilleure solution classique, ce qui laisse espérer que ce n'est pas la seule situation de ce type.

Cas particulier : résolution numérique. Rappelons que dans la version du jeu utilisée pour l'argument de Linial, le but est de placer les joueurs sur le graphe des voisinages d'un graphe de sorte qu'ils y soient répartis sur une version étiquetée de ce graphe. Or, pour que le problème ait un intérêt, il faut considérer un graphe d'une certaine taille. Or, même dans les cas les plus simples, le graphe des voisinages comporte un nombre non-négligeable de sommets, et un nombre non-négligeable de configurations selon lesquelles les sommets peuvent être placées (par exemple, dans l'exemple 2.3 (A CORRIGER), le graphe des voisinages comporte 9 sommets, et il y a 6 configurations selon lesquelles les sommets peuvent être placés).

Nous avons donc décidé d'ajouter une couche d'abstraction, en considérant un problème où les sommets sont nécessairement placés sur un parmi plusieurs sous-graphes homomorphes, mais pas nécessairement tous les sous-graphes homomorphes d'un graphe donné (et où les sommets ne communiquent pas entre eux). L'exemple le plus simple que nous avons trouvé et que nous avons réussi à étudier est le suivant :

Le graphe G est la grille 3×3 sur la bouteille de Klein, les sommets de la grille étant numérotés comme sur le dessin ci-dessous. Le nombre chromatique de ce graphe est 4.

Nous nous sommes intéressés à deux cas de figure :

1. Lorsque les joueurs sont placés sur l'un des quatre cycles $0-1-4-3, 4-5-8-7, 2-6-3-5, 7-6-2-1$ ou $8-0-3-5$.
2. Lorsque les joueurs sont placés sur l'un des neuf cycles $0-1-4-3, 4-5-8-7, 2-6-3-5, 7-6-2-1, 8-0-3-5, 0-6-4-7, 1-8-2-4, 3-6-8-7, 0-1-5-2$.

Nous avons étudié ces deux cas de figure numériquement sous l'hypothèse

1. La définition est en réalité plus subtile car nous pouvons définir un nombre chromatique par *rang* de l'état quantique qui est utilisé pour résoudre le problème, et il faut prendre le minimum des nombres chromatiques ainsi définis.

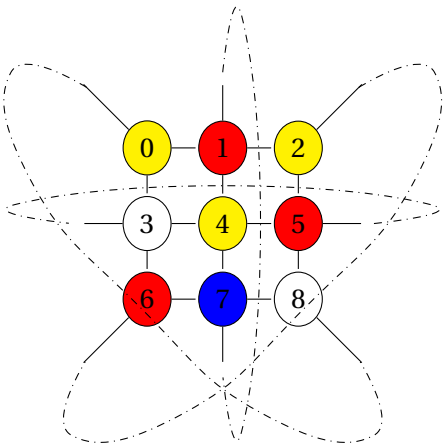


FIGURE 4.1 – Grille 3x3 sur la bouteille de Klein avec une 4-coloration.

de la consistance globale, c'est-à-dire que chaque joueur a la garantie que les autres joueurs sont placés sur l'un des cycles compatibles avec le sommet qui lui est donné. Par exemple, dans le premier cas de figure, si un joueur a le sommet numéro 0, il sait que tout le monde est placé sur le cycle $0-1-4-3$ ou sur le cycle $8-0-3-5$. En particulier, il sait qu'aucun joueur n'est placé sur le sommet 7 (par exemple).

Nous avons utilisé le logiciel SageMath pour écrire les contraintes non-signalling correspondant au problème de la 3-coloration dans ces deux cas de figure, en faisant l'hypothèse que toutes les bonnes colorations sont possibles et qu'aucune mauvaise coloration n'est possible, puis nous avons étudié le polytope résultant de ces contraintes : il s'est avéré que ce polytope est non-vide. Autrement dit, il existe une distribution non-signalling pour la 3-coloration.

Lorsque nous avons tenté d'étudier ce même problème sous l'hypothèse de la consistance locale (c'est-à-dire que les joueurs ne sont pas nécessairement placés sur l'un des cycles donnés dans la spécification du problème, et ils doivent sortir une bonne coloration s'ils sont tous placés sur un de ces cycles, mais pas s'ils sont placés autrement), le nombre de variables et de contraintes a explosé, et le problème ne pouvait plus être résolu numériquement dans un temps raisonnable.

Plus précisément, pour la consistance globale, nous avons trouvé $\mathcal{O}(2.10^3)$ **variables et** $\mathcal{O}(3.10^3)$ **contraintes non-signalling** pour le premier cas de figure (où les joueurs sont placés sur l'un parmi quatre cycles) ; et $\mathcal{O}(4.10^3)$ **variables et** $\mathcal{O}(1.10^4)$ **contraintes non-signalling** pour le deuxième (où les joueurs sont placés sur l'un parmi neuf cycles). Or, sous l'hypothèse de la

consistance locale, pour le premier cas de figure nous passons à $\mathcal{O}(2.10^5)$ **variables et** $\mathcal{O}(2.10^6)$ **contraintes non-signalling**!

Deuxième partie

k-localisabilité et k-dépendance

CHAPITRE 5
PRÉLIMINAIRES

Le but de ce chapitre et de définir et expliquer des notions utilisées dans la suite de cette partie, et de mettre en évidence des liens avec ce qui a été étudié dans la partie précédente.

5.1 La k-dépendance

Nous supposons que les définitions des notions de base en théorie des probabilités telles que univers, distribution de probabilité, variable aléatoire sont connues.

Définitions.

Définition 3. *Un **processus stochastique** est un ensemble de variables aléatoires indexées par* $\mathbb{Z}$ *:* $(X_n)_{n\in\mathbb{Z}}$.

Nous écrirons qu'un processus stochastique $(X_n)_{n\in\mathbb{Z}}$ est **i.i.d** lorsque les X_n sont indépendantes et identiquement distribuées.

Nous pouvons également voir un processus stochastique comme un ensemble de variables aléatoires indexées par les nœuds du chemin infini. En effet, cela serait un cas particulier d'un ensemble de variables aléatoires $(X_v)_{v\in V}$, où V est l'ensemble des sommets d'un graphe G avec un nombre dénombrable de sommets.

Définition 4. *Un processus stochastique* $(X_n)_{n\in\mathbb{Z}}$ *est **stationnaire** si* $(X_n)_{n\in\mathbb{Z}}$ *a la même distribution que* $(X_{n+1})_{n\in\mathbb{Z}}$.

Nous pourrions donc dire qu'un processus stationnaire est un processus dont la distribution est invariante par décalage d'indices.

Définition 5. *Un processus stochastique* $(X_n)_{n\in\mathbb{Z}}$ *est un* r***-block factor** s'il existe un processus stochastique i.i.d* $(Y_n)_{n\in\mathbb{Z}}$ *et une fonction*[1] *tels que* $X_n = f(Y_{n+1},\dots,Y_{n+r})\ \forall n \in \mathbb{Z}$.

1. Plus précisément, une fonction mesurable, mais ce détail n'a pas d'importance pour ce qui suit.

Autrement dit, la valeur de chaque X_n est fonction d'un « bloc » de r variables consécutives de $(Y_n)_{n\in\mathbb{Z}}$.

Définition 6. *Un processus stochastique est k-**dépendant** si pour tout $n \in \mathbb{Z}$, $(X_{\leq n}) := (\dots, X_{n-1}, X_n)$ et $(X_{>n}) := (X_{n+k+1}, X_{n+k+2}, \dots)$ sont indépendantes.*

La k-dépendance est une généralisation de l'indépendance : en effet, il est facile de voir qu'un processus 0-dépendant est tout simplement un processus indépendant. C'est une notion qui a été développée dans les années 1940-1950, où des variantes du théorème central limite pour les variables aléatoires k-dépendantes ont été démontrées [HR48, Dia55].

Liens avec des notions étudiées en calcul distribué. Nous allons maintenant faire quelques remarques sur le rapport entre ces notions et des choses que nous avons vues dans le domaine du calcul distribué.

- Il est facile de voir que la notion de r-block factor est très proche de celle de calculabilité distribuée classique en un nombre $f(r)$ de rondes (par exemple, $f(r) = r$ s'il s'agit d'un chemin orienté, ou $f(r) = r/2$ si r est pair et qu'il s'agit d'un chemin non-orienté). En effet, si nous reprenons l'exemple de la coloration, en considérant que la répartition des identifiants est donnée par un processus i.i.d $(Y_n)_{n\in\mathbb{Z}}$ et la répartition des couleurs par un processus $(X_n)_{n\in\mathbb{Z}}$, dire que $(X_n)_{n\in\mathbb{Z}}$ est un r-block factor de $(Y_n)_{n\in\mathbb{Z}}$ signifie que la coloration (aléatoire) d'un nœud est calculée à partir des identifiants (répartis aléatoirement) de r nœuds consécutifs (dont le nœud en question) sur le chemin.
- De même, la notion de k-dépendance (plus la stationnarité) est proche de celle de ϕ-LOCAL en un nombre $f(k)$ de rondes. En effet, dans une distribution ϕ-LOCAL, la sortie d'un nœud ne dépend que des entrées des variables à une certaine distance de lui. Donc, un ou plusieurs nœuds ne peuvent déduire aucune information autre que celle qui est disponible dans un rayon limité autour d'eux. Or, dans une distribution k-dépendant stationnaire, un ou plusieurs nœuds à distance plus de k l'un de l'autre ne peuvent pas connaître leurs indices à partir de leurs valeurs : autrement dit, ils ne peuvent connaître ni leur position absolue, ni leur distance, ni leur ordre.
- De plus, il est facile de voir que tout processus $(X_n)_{n\in\mathbb{Z}}$ qui est un r-block factor est nécessairement $(r-1)$-dépendant. En effet, pour un tel processus on a :

  $X_n = f(Y_{n+1}, \dots, Y_{n+r})$

  $X_{n+r-1} = f(Y_{n+r}, \dots, Y_{n+2r-1})$

$X_{n+r} = f(Y_{n+r+1}, \ldots, Y_{n+2r})$

On voit que la variable Y_{n+r} apparaît à la fois dans l'écriture de X_n et celle de X_{n+r-1}, donc il est possible que ces deux variables ne soient pas indépendantes[2], auquel cas le processus ne peut pas être $(r-2)$-dépendant. Par contre, l'ensemble des variables intervenant dans l'écriture de X_n et l'ensemble de celles intervenant dans l'écriture de X_{n+r} ont une intersection vide, et puisque le processus $(Y_n)_{n\in\mathbb{Z}}$ est i.i.d, cela signifie que X_n et X_{n+r} sont indépendantes. Cette indépendance vaut a fortiori pour des variables à une distance plus grande (c'est-à-dire, X_i avec $i < n$ et X_j avec $j > n+r$) d'où nous concluons que le processus est $(r-1)$-dépendant.

Pour faire le lien avec le calcul distribué, cette remarque pourrait s'interpréter de la manière suivante pour dire quelque chose de trivial : tout ce qui est calculable classiquement de façon distribuée en un certain nombre de ronde l'est aussi dans le modèle ϕ-LOCAL en le même nombre de rondes.

Lien entre k-dépendance et r-block factor. Nous avons montré plus haut que tout processus $(X_n)_{n\in\mathbb{Z}}$ qui est un r-block factor est nécessairement $(r-1)$-dépendant. Une question plus intéressante est celle de savoir, pour un processus k-dépendant, s'il existe un $r \in \mathbb{N}$ tel que ce processus est nécessairement un r-block factor[3].

La réponse à cette question est donnée pour la première fois par [AGKdV89]. Dans cet article, les auteurs construisent des processus 1-dépendant qui ne sont pas des 2-block factors, démontrant ainsi qu'au moins la contraposée de la remarque précédente (i.e. tout r-block factor est $(r-1)$-dépendant) est fausse. On a trouvé d'autres exemples répondant à la question plus générale depuis, cependant tous ces exemples partagent la caractéristique d'être plutôt « artificiels ».

En 2015 et 2016, Alexander Holroyd et Thomas Liggett construisent des processus de 3-coloration 2-dépendant et de q-coloration 1-dépendant pour $q \geq 4$ à titre d'exemples « naturels » de processus k-dépendant qui ne sont des r-block factor pour aucun $r \in \mathbb{N}$ [HL15, HL16]. Pour voir ce dernier point (i.e. un processus de coloration ne peut pas être un r-block factor quelque

2. En effet, elles peuvent être indépendantes si f est telle que Y_{n+r} n'intervient pas dans le calcul du résultat ou si son influence est en quelque sort annulée. Mais le fait que Y_{n+r} apparaît comme paramètre de la fonction signifie que cela n'est pas toujours le cas.

3. La réponse à cette question n'a toutefois pas beaucoup d'importance par rapport au calcul distribué. En effet, cela *aurait* l'implication que tout ce qui est ϕ-LOCAL en un nombre constant de rondes est ou n'est pas calculable classiquement de façon distribuée en un autre nombre constant de rondes *si* la ϕ-LOCALité impliquait la k-dépendance, ce qui n'est pas le cas (même si sa contraposée est vraie).

soit $r \in \mathbb{N}$), notons qu'il est aisé de traduire la preuve de Naor [Nao91] dans un langage probabiliste pour démontrer que le « rayon » d'un tel bloc est plus grand que n'importe quel entier r avec une probabilité positive s'exprimant comme une tour d'exponentielle dépendant de r [HSW16].

Nous décrirons le processus construit par Holroyd et Liggett plus longuement dans la dernière section de ce chapitre.

Enfin, notons que même si nous avons défini toutes les notions présentées dans cette section pour des variables indicées par $\mathbb{Z}$, ces notions ou des notions analogues existent pour des ensembles de variables aléatoires indicées par un intervalle fini de $\mathbb{Z}$ (par exemple $\{1,\dots,n\}$) ou les sommets d'un graphe aussi.

5.2 La k-localisabilité

Nous avons vu précédemment que tout processus k-dépendant et stationnaire est ϕ-LOCAL($f(k)$) pour une certaine fonction f de k, mais que la contraposée ne tient pas nécessairement. Nous avons donc essayé de définir une notion probabiliste qui traduit avec plus de justesse la notion de ϕ-LOCALité, et que nous avons appelée la k-localisabilité.

Pour cela, nous supposons que nous disposons d'un ensemble d'indices I, $I = \{1,\dots,n\}$ ou $I = \mathbb{Z}$, de variables aléatoires $(X_i)_{i\in I}$ indexées par I, et d'une notion de distance sur I.

Définition 7. *Une distribution de probabilités sur les $(X_i)_{i\in I}$ est dite k-**localisable** si pour tout $J,K \subseteq I$ intervalles à distance au moins k, la distribution de (X_J,X_K) ne dépend que de $\{|J|,|K|\}$.*

Autrement dit, cette distribution ne dépend ni des indices dans J et K (donc de la position absolue des intervalles), ni de la distance exacte entre J et K, ni de l'ordre entre J et K, c'est-à-dire du fait que les indices de l'un soient inférieurs aux indices de l'autre.

De plus, nous sommes autorisés à prendre J (ou K) égal à l'ensemble vide. Dans ce cas, la définition dit que la distribution de X_K ne dépend que de $|K|$. Par conséquent, toute distribution k-localisable est par définition stationnaire.

De plus, cette définition a aussi pour conséquence que la distribution de (X_J,X_K,X_L), pour des intervalles J, K, L mutuellement à distance au moins k les uns des autres ne dépend que de $\{|J|,|K|,|L|\}$. En effet, nous pouvons par exemple prendre J' l'union de J, K et des sommets entre les deux, et $K' := L$. Alors, la distribution de $(X_{J'},X_{K'})$ ne dépend que de $\{|J'|,|K'|\}$, donc également celle de (X_J,X_K,X_L). Le même raisonnement s'applique si

nous prenons $J'' := J$ et K'' l'union de K, L et des sommets entre les deux, ce qui permet de conclure. Nous pouvons généraliser ce raisonnement pour un nombre quelconque d'intervalles.

k-localisabilité et k-dépendance. Il est facile de voir qu'un processus k-dépendant stationnaire est k-localisable. En effet, la k-dépendance garantit que X_J et X_K sont indépendantes pour n'importe quels intervalles J et K à distance au moins k l'un de l'autre, i.e. la distribution de (X_J, X_K) ne dépend que de $\{J, K\}$; et la stationnarité garantit que $\Pr(X_J = x_J) = \Pr(X'_J = x_J)$ dès que $|J| = |J'|$. Ainsi, la distribution de (X_J, X_K) ne dépend que de $\{|J|, |K|\}$.

Par contre, un processus k-localisable n'est pas nécessairement k-dépendant, ou même ℓ-dépendant pour un $\ell \neq k$. Par exemple, si nous considérons n variables $X_1, \ldots, X_n$ à valeurs dans $\{1, \ldots, n\}$ et telles que $(X_1, \ldots, X_n) = (\sigma(1), \ldots, \sigma(n))$ où σ est une permutation sur $\{1, \ldots, n\}$ prise uniformément au hasard, alors cette distribution est 0-localisable puisque quels que soient I, J à distance au moins k l'un de l'autre, et $x_I \subset \{1, \cdots, n\}^{|I|}$, $x_J \subset \{1, \cdots, n\}^{|J|}$:

- $\Pr(X_I = x_I, X_J = x_J) = \frac{(n-|I|-|J|)!}{n!}$ si tous les éléments apparaissant dans x_I et x_J sont distincts.
- $\Pr(X_I = x_I, X_J = x_J) = 0$ si un élément apparaît au moins deux fois dans x_I et/ou x_J (puisque σ est une permutation).

Par contre, elle n'est k-dépendante pour aucun $k \leq n$: en effet, si nous sommes dans le premier cas de figure (c'est-à-dire tous les éléments de x_I et de x_J sont distincts)

$$\begin{aligned}\Pr(X_J = x_J)\Pr(X_I = x_I) &= \frac{(n-|I|)!}{n!}\frac{(n-|J|)!}{n!} \\ &\neq \frac{(n-|I|-|J|)!}{n!} \\ &= \Pr(X_J = x_J, X_I = x_I)\end{aligned}$$

k-localisabilité et ϕ-LOCAL. La notion de k-localisabilité traduit plus ou moins la notion de ϕ-LOCALité telle que présentée dans [GKM09], c'est-à-dire la ϕ-LOCALité avec la condition de consistance globale. Pour voir cela, rappelons que dans le cas de la consistance globale, en considérant les voisinages de tous les sommets ensemble, nous voyons qu'ils forment le type de graphe sur lequel on cherche à résoudre le problème (en l'occurrence, un

chemin dans le cas qui nous intéresse), alors que dans le cas de la consistance locale, ce n'est pas nécessairement le cas.

Si nous disposons d'une distribution k-localisable sur un graphe donné, cela nous donne une distribution vérifiant la condition ϕ-LOCAL(k) si le graphe est orienté et la condition ϕ-LOCAL($k/2$) si k est pair et le graphe est non-orienté [4]. En effet, si la distribution de sorties de deux ensembles I et J de sommets à distance au moins k l'un de l'autre ne dépend que de leurs tailles, en particulier, il ne dépend pas de ce qui se passe en dehors des k-voisinages de ces ensembles de sommets. Par conséquent, elle sera la même pour des k-voisinages identiques.

Par contre, cette distribution n'a aucune raison de dépendre des particularités des k-voisinages de I et J, et pour cette raison la k-localisabilité est un peu plus forte que la ϕ-LOCALité. Nous allons illustrer cela à l'aide de deux exemples.

Premièrement, considérons un chemin fini orienté de longueur n, un ensemble I tel que $|I| < n - k - 1$ et un ensemble J tel que $|J| = 1$ (i.e. J comporte un seul sommet) à distance au moins k l'un de l'autre. Alors, sous l'hypothèse de la k-localisabilité, la distribution conjointe de I et de J est la même que le sommet dans J se trouve à la fin du chemin ou non. Alors que dans le modèle ϕ-LOCAL(k), le sommet à la fin du chemin a un voisinage différent des sommets qui ne sont pas à la fin du chemin, donc la distribution conjointe de I et de J peut être différente dans le cas où J comporte le sommet à la fin du chemin et dans le cas où il comporte un sommet au milieu du chemin.

Deuxièmement, prenons un chemin fini orienté de longueur n et des intervalles I et J de taille 2 à distance au moins 1 l'un de l'autre. Supposons que les éléments de I ont pour identifiants 1 resp. 2 et ceux de J 5 resp. 4. Alors, sous l'hypothèse de la 1-localisabilité, la distribution conjointe de I et de J est la même indépendamment de l'identifiant du sommet qui suit les sommets 1 et 2 dans le chemin : autrement dit, indépendamment de si nous avons $1-2-3$, $1-2-4$, ou encore autre chose. Or, l'ensemble des sommets $I = \{1,2\}$ n'a pas le même voisinage dans ces deux cas, et puisque la distribution conjointe de I et J dépend du 1-voisinage dans le modèle ϕ-LOCAL(1), elle est susceptible d'être différente selon le voisinage.

k-localisabilité et échangeabilité. Enfin, notons que la k-localisabilité généralise la notion de l'**échangeabilité** tout comme la k-dépendance généralise l'indépendance :

4. Si k est impair, cela reviendrait à supposer que chaque sommet voit ses voisins à distance $\frac{k-1}{2}$ d'un côté et ses voisins à distance $\frac{k+1}{2}$ de l'autre ...

Définition 8. *Soit $(X_n)_{n\in\mathbb{Z}}$ un processus stochastique.*

1. *Nous disons que ce processus est k-**échangeable**, pour un $k \in \mathbb{N}$, si pour tout ensemble de k variables $X_{i_1}, \dots, X_{i_k}$, toute permutation σ, et toutes valeurs $x_1, \dots, x_k$ pouvant être prises par les variables,*

   $$\Pr(X_{i_1} = x_1, \dots, X_{i_k} = x_k) = \Pr(X_{\sigma(i_1)} = x_1, \dots, X_{\sigma(i_k)} = x_k).$$

2. *Nous disons que ce processus est **échangeable** s'il est k-échangeable pour tout $k \in \mathbb{N}$.*

Autrement dit, un processus est k-échangeable si la distribution d'un ensemble de k variables dépend seulement de la taille de cet ensemble et non de l'indice précis des variables. Il est ainsi facile de voir que la 0-localisabilité est la même chose que l'échangeabilité. De même, la k-localisabilité signifie que la distribution d'un certain nombre d'ensembles de variables dépend seulement de la taille de ces ensemble et non de leur position ou ordre.

Il est facile de voir que tout processus i.i.d est échangeable. Pour la contraposée, il existe un lien entre l'échangeabilité et l'indépendance sous certaines hypothèses sur les variables aléatoires sous-jacentes. Le premier résultat de ce type a été établi par de Finetti [Dia77, Kir19] :

Théorème 1. *Soit $(X_n)_{n\in\mathbb{Z}}$ un processus de Bernoulli de paramètre p échangeable. Alors il existe une distribution de probabilités μ sur $[0,1]$ tel que pour tout $n \in \mathbb{N}$ et $(X_1, \dots, X_n) \in \{0,1\}^n$*

$$\Pr(X_1 = x_1, \dots, X_n = x_n) = \int p^k (1-p)^{n-k} d\mu(p)$$

où k est le nombre de x_i égaux à 1.

Autrement dit, tout processus de Bernoulli de paramètre p échangeable est obtenu via un « mélange » de processus de Bernoulli de paramètre p indépendants dans le sens où le paramètre p est échantillonné selon la distribution μ.

Le théorème de de Finetti a été généralisé de plusieurs manières : par exemple, pour des variables prenant des valeurs dans un ensemble fini de taille supérieure à deux. Ou encore, il existe aussi une généralisation de ce théorème [Pet90] qui établit un lien similaire entre la k-dépendance et la k-localisabilité (appelée la k-symétrie dans l'article en question) : il s'agit de la seule instance de cette notion que nous avons trouvée dans la littérature existante.

5.3 Définitions auxiliaires

Le but de cette section est de servir de glossaire pour certains termes apparaissant dans les résultats énoncés plus tard, mais qui ne sont pas en eux-mêmes centraux pour notre sujet d'étude.

Processus de renouvellement. Nous commencerons par définir une notion importante pour comprendre l'énoncé du Théorème 2.

Définition 9. *Un processus $(S_n)_{n\in\mathbb{N}}$ à valeurs dans $\mathbb{N}$ est un **processus d'arrivées** si $0 < S_1$ et $S_i < S_{i+1}\ \forall i \in \mathbb{N}$. Les S_i sont appelés des **époques d'arrivée**.*

Un processus d'arrivées représente par exemple le temps (discret) d'arrivées d'une nouvelle personne dans une file d'attente. Mais il est plus générale que son nom l'indique, il peut également représenter le temps de départ d'une file d'attente.

Mais il existe deux autres façons de représenter le même type de phénomène.

Définition 10. *Étant donné un processus d'arrivées $(S_n)_{n\in\mathbb{N}}$, le **processus des temps d'arrivée**[5] associé est le processus $(T_n)_{n\in\mathbb{N}}$ tel que $T_t = 1$ s'il existe $i \in \mathbb{N}$ tel que $S_i = t$, et $T_t = 0$ sinon*[6].

Autrement dit, $T_t = 1$ indique qu'il y a eu une arrivée au temps t.

Définition 11. *Étant donné un processus d'arrivées $(S_n)_{n\in\mathbb{N}}$, le **processus des inter-arrivées** associé est le processus $(I_n)_{n\in\mathbb{N}}$ où $I_i = S_{i+1} - S_i$.*

En particulier, puisque $S_i < S_{i+1}$ par définition d'un processus d'arrivées, $I_i > 0\ \forall i \in \mathbb{N}$.

Puisque la donnée d'un processus d'arrivées est équivalente à la donnée d'un processus des temps d'arrivée, il est facile de voir que le processus des inter-arrivées peut également être défini à partir de ce dernier processus, même si la définition est plus fastidieuse à écrire.

Définition 12. *Un processus d'arrivées $(S_n)_{n\in\mathbb{N}}$ est un **processus de renouvellement** si les variables d'inter-arrivées associées sont i.i.d.*

5. Il s'agit d'une appellation non-standard.
6. Ou encore $T_t = \mathbb{1}_{\{S_i, i\geq 0\}}(t)$

Ensemble stable, polynôme de stabilité. Quelques résultats présentés plus loin vont au-delà des processus indexés par $\mathbb{Z}$ et concernent les ensembles de variables aléatoires indexés par les sommets d'un graphe, c'est pourquoi quelques notions de théorie des graphes sont nécessaires pour les comprendre.

Définition 13. *Soit $G = (V, E)$ un graphe. Un **ensemble stable** ou **ensemble indépendant** de G est un sous-ensemble $U \subset V$ de sommets mutuellement non-adjacents, i.e. tels que $\forall u, v \in U$, $\{u, v\} \notin E$.*

Définition 14. *Soit $G = (V, E)$ un graphe. Un **ensemble indépendant maximum** I de G est un ensemble indépendant tel que pour tout ensemble indépendant J de G, $|J| \leq |I|$.*

Définition 15. *Soit $G = (V, E)$ un graphe. Un **ensemble indépendant maximal** I de G est un ensemble indépendant tel qu'il n'existe aucun ensemble indépendant J de G tel que $I \subsetneq J$.*

Nous utiliserons également l'abréviation MIS (de l'anglais *maximal independent set*) pour désigner un ensemble indépendant maximal.

Nous allons maintenant définir le **polynôme d'indépendance** d'un graphe qui est bien connu dans la littérature, ainsi qu'une variante à plusieurs variables utilisée dans l'un des résultats de cette thèse.

Pour cela, nous allons considérer un graphe $G = (V, E)$. Nous notons

- $s(G, i)$ le nombre d'ensembles indépendants de taille i de G, pour un $i \leq |V|$,
- M la taille d'un ensemble indépendant maximum de G.

Définition 16. *Le polynôme d'indépendance I_G de G est l'élément suivant de l'anneau $\mathbb{Z}[z]$:*

$$I_G(z) := \sum_{i=0}^{M} s(G, i) z^i$$

Définition 17. *Le polynôme d'indépendance à plusieurs variables I_G^M de G est l'élément suivant de l'anneau $\mathbb{Z}[z_1, \ldots, z_M]$:*

$$I_G^M(z_1, \ldots, z_M) := \sum_{i=0}^{M} s(G, i) z_i$$

Coloration à distance d, graphe puissance [KK08]. Nous présentons ici quelques notions utiles pour comprendre la Section 6.4.1.

Définition 18. *Soient $G = (V, E)$ un graphe et $q, d \in \mathbb{N}$ et $(X_v)_{v \in V}$ un ensemble de variables aléatoires indexées par les sommets de G et à valeurs dans $\{1, \dots, q\}$. On parle de q-**coloration à distance** d si $X_u \neq X_v$ pour tout $u, v \in V$ à distance inférieure ou égale à d l'un de l'autre.*

Ainsi, le fait qu'une variable prenne une certaine couleur interdit à toutes les variables dans un rayon d de cette variable de prendre cette même couleur. Ainsi, une q-coloration « classique » est une q-coloration à distance 1.

Définition 19. *Soit $G = (V, E)$ et $d \in \mathbb{N}$. Le plus petit $q \in \mathbb{N}$ tel qu'il existe une q-coloration à distance d de G est appelé **nombre chromatique à distance** d de G. Nous le noterons $\chi(G, d)$.*

Il existe un lien entre la coloration à distance d d'un graphe et la coloration « classique » d'un graphe dérivé de ce même graphe.

Définition 20. *Soient $G = (V, E)$ un graphe et $d \in \mathbb{N}$. Nous appelons **graphe puissance** d **de** G, et on note G^d, le graphe ayant le même ensemble de sommets que G et dont l'ensemble des arêtes est :*

Si $d = 1$: E.

Si $d \geq 2$

$$\begin{aligned} \{\{u, v\}, \exists w_0, \dots, w_t,\ & t \leq d, \\ & w_0 = u, w_d = v, \\ & \{w_i, w_{i+1}\} \in E \text{ pour } i = 0, \dots, t\} \end{aligned}$$

Autrement dit, dans le graphe puissance d de G, tous les sommets qui sont à distance au moins d les uns des autres dans G sont reliés.

Proposition 1. *Soit $G = (V, E)$ un graphe et $d \in \mathbb{N}$. Alors $\chi(G, d) = \chi(G^d)$.*

5.4 Autour de la coloration k-dépendante

Dans ce qui suit, nous allons commencer par décrire brièvement les processus de coloration construits par Holroyd et Liggett. Mais nous nous intéresserons plus longuement à leurs arguments afin de démontrer l'impossibilité d'une 3-coloration stationnaire et 1-dépendante et plus tard, d'autres arguments d'impossibilité et de « borne inférieure » pour des processus plus généraux définis sur des graphes.

Parmi toutes les constructions possibles de leurs processus de coloration proposés par Holroyd et Liggett, nous en exposerons principalement deux : une première qui permet de « visualiser » de façon assez intuitive le processus en en donnant une construction récursive, et une deuxième qui consiste à donner une formule récursive permettant de calculer la probabilité de chaque coloration d'un chemin de longueur n à partir des probabilités de chaque coloration d'un chemin de longueur $n-1$.

Enfin, dans ce qui suit, tous les processus que nous étudierons seront stationnaires, mais nous omettrons de le mentionner à chaque fois par souci de concision. Par conséquent, il faut entendre « processus k-dépendant stationnaire » lorsque nous écrirons « processus k-dépendant ».

Processus de coloration : une description visuelle. Nous commençons par le chemin de longueur 1, et on choisit la couleur de l'unique sommet de ce chemin en prenant uniformément un élément de $\{1,\dots,q\}$, où $q = 3$ ou 4 est le nombre de couleurs. Ensuite, nous insérons un par un de nouveaux sommets en choisissant la position et la couleur du nouveau sommet de la manière suivante (nous notons n la longueur du chemin après avoir inséré ce nouveau sommet) :

Le nouveau sommet est inséré entre deux sommets du chemin de longueur $(n-1)$, donc dans l'une de $(n-2)$ positions possibles, avec probabilité $\frac{q-2}{n(q-2)+2}$ ou à l'une des deux extrémités de ce chemin avec probabilité $\frac{q-1}{n(q-2)+2}$. Ensuite, sa couleur est choisie uniformément au hasard parmi celles qui sont disponibles (c'est-à-dire, celles qui ne sont pas prises par ses voisins s'il est au milieu du chemin ou son voisin s'il est à l'extrémité).

Processus de coloration : une formule récursive. Nous allons donner une formule récursive qui permet de construire une q-coloration k-dépendante (avec $(q,k) \in \{(4,1),(3,2)\}$) d'un chemin de longueur n à partir d'une telle coloration d'un chemin de longueur $(n-1)$. La coloration sur le chemin infini s'obtient à partir de cette coloration puisqu'elle satisfait les conditions du théorème d'extension de Kolmogorov [7].

Avant de donner la formule, nous allons expliquer les notations qui y apparaissent. Nous écrirons

- $x := (x_1,\dots,x_n)$
- $\Pr(x)$ pour $\Pr(X_1 = x_1,\dots,X_n = x_n)$
- $\hat{x_i} := (x_1,\dots,x_{i-1},x_{i+1},\dots,x_n)$

7. Il s'agit d'un théorème qui donne les conditions pour qu'une distribution de probabilités définie sur des segments finis puisse être étendue sur $\mathbb{Z}$.

- Et donc

  $\Pr(\hat{x_i})$ pour $\Pr(X_1 = x_1, \ldots, X_{i_1} = x_{i-1}, X_{i+1} = x_{i+1}, \ldots, X_n = x_n)$,

  c'est-à-dire la coloration du chemin de longueur $(n-1)$ obtenu en recollant les deux chemins qui résultent après avoir enlevé le i-ème nœud d'un chemin de longueur n.

Nous donnons maintenant les formules pour la q-coloration, $q = 3$ ou 4 [HL16].

Lorsque le chemin est de longueur 1 (i.e. il y a un seul nœud), $\Pr(x) = \Pr(x_1) = 1/q$.

Pour un chemin de longueur n, la probabilité d'apparition d'une coloration se calcule à l'aide des formules récursives suivantes :

- **Pour la 4-coloration** :

$$\Pr(x) = \frac{1}{2(n+1)} \sum_{i=1}^{n} \Pr(\hat{x_i})$$

- **Pour la 3-coloration** :

$$\Pr(x) = \frac{1}{(n+2)} \sum_{i=1}^{n} \Pr(\hat{x_i})$$

Dans l'encadré de la page suivante, nous avons détaillé l'application de la formule pour la 4-coloration aux chemins de longueur 1 à 3.

Pour finir, nous remarquons qu'il n'est pas connu si ces colorations constituent les uniques 3- et 4-colorations 2-dépendantes resp. 1-dépendantes de $\mathbb{Z}$, même si pour la 4-coloration en particulier certains faits laissent croire qu'il ne peut pas en exister d'autres. La question de l'unicité de ces colorations k-dépendantes, $k = 1, 2$, est une donc une question ouverte.

Exemple : application de la formule de 4-coloration

- **$\mathbf{n = 1}$**

  Pour un chemin composé d'un seul nœud, chaque couleur apparaît avec probabilité 1/4.

- **$\mathbf{n = 2}$**

  Il y a une seule « classe » de coloration, $x = (a, b)$ avec $a \neq b$.

  Nous appliquons la formule

$$\begin{aligned}\Pr(ab) &= \frac{1}{2(1+2)}(\Pr(b) + \Pr(a)) \\ &= \frac{1}{6}(\frac{1}{4} + \frac{1}{4}) \\ &= \frac{1}{12}\end{aligned}$$

  Donc chaque coloration apparaît avec probabilité $\frac{1}{12}$.

- **$\mathbf{n = 3}$** Il y a deux « classes » de coloration :
  - $x = (a, b, a)$ avec $a \neq b$.

$$\begin{aligned}\Pr(aba) &= \frac{1}{2(2+2)}(\Pr(ba) + \Pr(aa) + \Pr(ab)) \\ &= \frac{1}{8}(\frac{1}{12} + 0 + \frac{1}{12}) \\ &= \frac{1}{48}\end{aligned}$$

  - $x = (a, b, c)$ avec a, b, c tous distincts.

$$\begin{aligned}\Pr(abc) &= \frac{1}{2(2+2)}(\Pr(bc) + \Pr(ac) + \Pr(ab)) \\ &= \frac{1}{8}(\frac{1}{12} + \frac{1}{12} + \frac{1}{12}) \\ &= \frac{1}{32}\end{aligned}$$

Et ainsi de suite ...

Impossibilité d'une 3-coloration 1-dépendante stationnaire. Un autre résultat de Holroyd et Liggett est une borne inférieure sur le nombre de couleurs q nécessaire pour colorer un chemin de façon 1-dépendante. Il avait déjà été montré par Schramm qu'il était impossible de le faire avec 3 couleurs [HSW16], le résultat de Holroyd et Liggett est plus général. Dans cette partie, nous allons donc expliquer leur stratégie et les résultats obtenus.

Leur démarche consiste à étudier un processus binaire plus simple $(Y_n)_{n\in\mathbb{Z}}$ induit par le processus de coloration $(X_n)_{n\in\mathbb{Z}}$ de la manière suivante : nous considérons la couleur c qui apparaît avec la plus grande fréquence.

Alors $Y_n = 1$ si $X_n = c$ et $Y_n = 0$ sinon.

Le processus $(Y_n)_{n\in\mathbb{Z}}$ a les propriétés suivantes :

1. Il s'agit d'un processus de la forme $Y_n = f(X_n)$, avec la même fonction f quel que soit n. Cette transformation conserve donc la 1-dépendance (plus généralement, elle conserve la k-dépendance) et la stationnarité.
2. Puisque $(X_n)_{n\in\mathbb{Z}}$ est un processus de coloration, deux variables voisines ne valent jamais toutes les deux c. Par conséquent, deux variables voisines de $(Y_n)_{n\in\mathbb{Z}}$ ne valent jamais toutes les deux 1 non plus. Ce type de processus binaires, où nous ne pouvons jamais avoir plusieurs « 1 » voisins, est appelé **processus à particules dures** (*hard-core process* en anglais) dans la littérature [8].
3. Quelle que soit la distribution des couleurs, la couleur la plus fréquente apparaît avec probabilité au moins $1/q$ (rappelons que q est le nombre de couleurs). Par conséquent, $\Pr(Y_n = 1) \geq 1/q$ quel que soit $n \in \mathbb{Z}$, ou encore :

$$q \geq 1/Pr(Y_n = 1) \tag{5.1}$$

   Nous en déduisons qu'étudier la valeur de $\Pr(Y_n = 1)$ permet d'obtenir une borne inférieure sur le nombre de couleurs nécessaires pour une coloration 1-dépendante.

Holroyd et Liggett démontrent alors le résultat suivant

Théorème 2. *Le processus $(Y_n)_{n\in\mathbb{Z}}$ est un processus de renouvellement dont la suite des inter-arrivées a comme fonction génératrice la fonction*

$$G(z) = \mathbb{E}(z^T) = \frac{pz^2}{1 - z + pz^2}$$

où $p := \Pr(Y_0 = 1)$

8. Et, d'un point de vue théorie des graphes, si nous regardons l'ensemble des variables qui valent 1, il s'agit d'un ensemble indépendant.

Ce théorème a pour corollaire que $p \leq 1/4$. En effet, si nous considérons le dénominateur de $G(z)$ comme un polynôme en z, son déterminant est $\Delta = 1 - 4p$. Et si $p > 1/4$, alors $\Delta < 0$ et cela contredit un théorème qui dit qu'une série de Taylor à coefficients réels non-négatifs admet une singularité au niveau du rayon de convergence réel, puisque les singularités de $G(z)$ sont imaginaires dans ce cas.

Nous pouvons alors utiliser ce corollaire et la propriété (3) du processus $(Y_n)_{n\in\mathbb{Z}}$ notée ci-dessus (à savoir que le nombre de couleurs q vérifie $q \geq 1/p$) afin de déduire que $q \geq 4$.

Processus à particules dures sur les graphes. Holroyd et Liggett ont également démontré des résultats sur les processus à particules dures 1-dépendants surles graphes plus généraux, plus précisément, des graphes non-orientés simples avec un nombre dénombrable de sommets et dont les degrés sont finis.

Nous allons considérer un graphe $G = (V, E)$ ayant ces caractéristiques, le processus à particules dures qui nous intéresse est alors celui indexé par les sommets de ce graphe : $(Y_v)_{v\in V}$.

Notons

$$p_h(G) := \max_{p\in\mathbb{R}^+}\{\exists (Y_v)_{v\in V} \text{ à particules dures 1-dépendant tel que } \Pr(Y_v = 1) = p\, \forall v \in V\}.$$

Nous avons alors :

Lemme 1. *Pour tout $p \leq p_h(G)$ il existe un unique processus à particules dures 1-dépendant $(Y_v)_{v\in V}$ avec $\Pr(Y_v = 1) = p\ \forall v \in V$. De plus, la loi de ce processus est invariant par automorphismes de G.*

Les résultats qui suivent utilisent la notion de **polynôme d'indépendance** I_G d'un graphe G, qui a été définie dans la Section 5.3.

Proposition 2. *Soit $G = (V, E)$ un graphe et $p \in [0, 1]$. Alors, $p \leq p_h(G)$ ssi pour tout sous-ensemble fini de sommets $S \subset G$, et $H := G_{|S}$ le graphe induit par ce sous-ensemble, nous avons*

$$I_H(-p) \geq 0$$

La preuve de ce résultat repose sur le principe d'inclusion-exclusion.

Nous pouvons utiliser ce résultat et l'expression des zéros du polynôme d'indépendance du chemin fini pour affiner le résultat obtenu grâce au

Théorème 2 : l'utilisation de 4 couleurs pour avoir une coloration 1-dépendante devient nécessaire même pour des chemins de petite longueur.

En effet, nous remarquons d'abord que tout sous-graphe H d'un chemin vérifiant les conditions de la Proposition 2 est lui-même un chemin ou l'union de plusieurs chemins disjoints. Or, on peut montrer que le polynôme d'indépendance de l'union disjointe de plusieurs graphes est le produit de leurs polynômes d'indépendance. Il suffit donc d'étudier le polynôme d'indépendance d'un chemin de longueur finie.

Nous avons $I_G(0) = 1$ pour n'importe quel graphe G puisqu'il a toujours un seul ensemble indépendant de taille 0, l'ensemble vide. Le polynôme d'indépendance prend donc une valeur positive entre 0 et ses racines les plus proches de l'origine. Nous cherchons de plus à avoir une borne inférieure p sur $p_h(G)$, c'est pourquoi nous nous intéressons à la valeur des racines les plus proches de l'origine du polynôme d'indépendance du chemin.

Nous pouvons montrer que les zéros du polynôme d'indépendance d'un chemin de longueur n ont pour expression [DSDS05]

$$r_{k,n} = -\frac{1}{4\cos^2\left(\frac{\pi k}{n+2}\right)}, k = 1, \ldots, \lfloor \frac{n+1}{2} \rfloor$$

Notons r_n^* la racine la plus proche de l'origine pour le chemin de longueur n. Nous pouvons montrer que $r_5^* < -1/3$ (et donc qu'il faut plus de 3 couleurs pour avoir une coloration 1-dépendante du 5-chemin). Et, plus généralement, nous pouvons montrer que r_n^* converge par le bas vers $-1/4$ lorsque n tend vers l'infini.

Le graphe étoile. Récemment, Liggett et Tang [MLT18] ont étudié les processus à particules dures et la coloration 1-dépendante sur les graphes étoiles à d branches infinies, i.e. les graphes ayant un sommet distingué v_0 de degré d dont émanent des chemins infinis dans un sens $(v_{i,1}, v_{i,2}, \ldots)$, $1 \leq i \leq d$.

Pour $d = 3$, le graphe étoile n'admet pas de 4-coloration 1-dépendante.

Résumé et conclusion. Nous pouvons résumer les résultats de Holroyd et Liggett sur la coloration du chemin de longueur n et sur $\mathbb{Z}$ de la manière suivante :

1. Il existe une q-coloration 1-dépendante stationnaire pour $q \geq 4$.
2. Il existe une 3-coloration 2-dépendante stationnaire.

3. La 3-coloration 1-dépendante stationnaire n'est pas possible pour n suffisamment grand.

Puisque, comme nous l'avons vu, une distribution k-localisable stationnaire est k-localisable, mais que la contraposée n'est pas vrai, les points (1) et (2) ci-dessus impliquent :

1. Il existe une q-coloration 1-localisable pour $q \geq 4$.
2. Il existe une 3-coloration 2-localisable.

Néanmoins, les résultats de Holroyd et Liggett ne permettent pas de répondre à la question de savoir s'il existe une 3-coloration 1-localisable. C'est pourquoi nous avons choisi d'étudier cette question, en adoptant une stratégie similaire à la leur passant par l'étude des processus à particules dures 1-localisables : il s'agit de l'objet du prochain chapitre.

CHAPITRE 6

Processus à particules dures 1-localisables

Le but de ce chapitre est de présenter le résultat technique principal obtenu pendant de cette thèse : l'étude des processus à particules dures 1-localisables (sur le chemin fini puis étendu à l'infini), suivi d'une généralisation aux processus à particules dures « de rayon k » k-localisables (nous définirons plus précisément cette notion dans la partie qui y est consacrée).

Rappelons que, tout comme l'étude des processus à particules dures 1-dépendantes a permis d'obtenir des résultats sur la coloration 1-dépendante (via l'inégalité (5.1) qui donne une borne inférieure), nous pouvons démontrer un résultat analogue via un argument similaire : à savoir que la probabilité maximale d'apparition d'un « 1 » quelque part donne une borne inférieure sur le nombre de couleurs nécessaires pour une coloration 1-localisable.

Commençons par rappeler la définition de la k-localisabilité, étant donné $I = \{1, \dots, n\}$ ou $I = \mathbb{Z}$ un ensemble d'indices, des variables aléatoires $(X_I)_{i\in I}$ et une notion de distance sur I.

Définition 7. *Une distribution de probabilités sur les $(X_i)_{i\in I}$ est dite k-**localisable** si pour tout $J, K \subseteq I$ intervalles à distance au moins k, la distribution de (X_J, X_K) ne dépend que de $\{|J|, |K|\}$.*

En particulier, rappelons que cela a pour conséquence que des variables aléatoires k-dépendantes sont stationnaires.

Définitions et notations.

Définition 21. *Nous disons qu'une famille de variables $(X_i)_{i\in I}$ est **à particules dures** s'il s'agit de variables binaires et s'il est impossible que deux variables voisines valent toutes les deux 1.*

Définition 22. *Soit*

$$\mathcal{I}_n := \{x = x_1x_2\dots x_n \mid x \in \{0,1\}^n, \\ \forall 1 < i < n\, x_i = 1 \Rightarrow (x_{i-1} = 0 \textit{ et } x_{i+1} = 0)\}.$$

*Un élément de $\mathcal{I}_n$ est appelé un **mot binaire à particules dures de longueur** n* [1].

1. Si $x_1 = 1$, alors $x_2 = 0$ car sinon nous aurions $x_1 = x_2 = 1$ ce qui violerait la condition pour $i = 2$. De même, si $x_n = 1$, alors $x_{n-1} = 0$ car sinon cela violerait la condition pour $i = n-1$. Par conséquent, tous les cas possibles sont traités dans cette définition.

Dans ce qui suit, les familles de variables ou les mots considérés seront toujours à particules dures. De même, lorsque nous parlerons de distribution de probabilité, il sera toujours entendu qu'il s'agit d'une distribution de variables à particules dures.

Nous notons $p_1 := \Pr(X_1 = 1)$, $p_2 := \Pr(X_1 = 1, X_3 = 1)$, et plus généralement $p_i := \Pr(X_1 = 1, \ldots, X_{2i-1} = 1)$ pour $1 \leq i \leq \lceil \frac{n}{2} \rceil$.

Notons $\mathsf{c}_n := \frac{1}{n+1}\binom{2n}{n}$ le n-ème nombre de Catalan (pour plus de détails, voir le Chapitre 7).

Résultats principaux. Nous allons démontrer le résultat suivant :

Théorème 3. *Soit $n \in \mathbb{N}$ et $\ell := \lceil \frac{n}{2} \rceil$.*

i. *Toute distribution de probabilité 1-localisable sur $\mathcal{I}_n$ vérifie $p_i \leq \mathsf{c}_{\ell-i+1}/\mathsf{c}_{\ell+1}$ $(1 \leq i \leq \ell)$.*

ii. *Il existe une distribution de probabilités 1-localisable sur $\mathcal{I}_n$ telle que $p_i = \mathsf{c}_{\ell-i+1}/\mathsf{c}_{\ell+1}$ $(1 \leq i \leq \ell)$.*

En prenant les probabilités marginales, nous montrons facilement que $p_i \leq \mathsf{c}_{\lceil \frac{n}{2}-i+1 \rceil}/\mathsf{c}_{\lceil \frac{n}{2}+1 \rceil}$ quel que soit n.

En particulier, en constatant que $\frac{\mathsf{c}_m}{\mathsf{c}_{m+1}} = \frac{m+2}{4m+2}$, on montre facilement que $p_i \underset{n\to\infty}{\longrightarrow} 1/4^i$, ce qui nous donne notre résultat principal.

Théorème 4. *Pour n assez grand, il n'existe pas de distribution de probabilité 1-localisable de q-coloration des $(X_i)_{i\in I}$ pour $q < 4$.*

6.1 Étude sur les petits cas

Le but de cette section est d'étudier « à la main » le problème étudié dans le Théorème 3 pour quelques valeurs de n petites dans deux buts : une meilleure clarification de ce résultat, et rendre familières les techniques qui seront utilisées pour résoudre le problème dans le cas général.

Rappelons pour cela que nous avons défini des variables

$$p_i := \Pr(X_1 = 1, \ldots, X_{2i-1} = 1) \text{ pour } 1 \leq i \leq \ell := \lceil \frac{n}{2} \rceil$$

1. **$\mathbf{n = 1}$**

   Dans ce cas, il n'y a qu'une seule variable, et nous avons le système suivant :

x_1	$\Pr(X_1 = x_1)$
1	p_1
0	$1 - p_1$

   Pour qu'une distribution 1-localisable existe, il faut et il suffit que toutes les probabilités soient positives. Il n'est pas compliqué de voir que dans ce cas de figure, cela équivaut à $0 \leq p_1 \leq 1$. Ce cas est donc trivial.

2. **$\mathbf{n = 2}$**

   Une fois de plus il n'y a qu'une seule variable, et nous avons le système suivant

(x_1, x_2)	$\Pr(X_1 = x_1, X_2 = x_2)$
(1,0)	p_1
(0,1)	p_1
(0,0)	$1 - 2p_1$

   Pour la première ligne, nous remarquons que si $X_1 = 1$, puisqu'il s'agit de variables à particules dures, nécessairement $X_2 = 0$ et donc $\Pr(X_1 = 1, X_2 = 0) = \Pr(X_1 = 1) = p_1$.

   Pour la deuxième ligne, nous utilisons un argument similaire : si $X_2 = 1$, nécessairement $X_1 = 0$, donc $\Pr(X_1 = 0, X_2 = 1) = \Pr(X_2 = 1)$. Ensuite, nous utilisons le fait que la distribution soit 1-localisable : cette propriété entraîne en effet que $\Pr(X_2 = 1) = \Pr(X_1 = 1)$, puisqu'il s'agit de variables indexées par des ensembles ayant la même taille (à savoir 1).

   Finalement, pour la dernière ligne, nous utilisons le fait que la somme des probabilités doit faire 1.

   Pour qu'une telle distribution existe, il faut et il suffit que toutes les probabilités soient positives, ce qui donne $0 \leq p_1 \leq 1/2$.

3. **$\mathbf{n = 3}$**

Maintenant, il y a deux variables p_1 et p_2. Nous avons le système suivant :

(x_1, x_2, x_3)	$\Pr(X_1 = x_1, X_2 = x_2, X_3 = x_3)$
(0,1,0)	p_1
(1,0,1)	p_2
(1,0,0)	$p_1 - p_2$
(0,0,1)	$p_1 - p_2$
(0,0,0)	$1 - 3p_1 + p_2$

Pour la première ligne, il s'agit de $\Pr(X_2 = 1)$, qui est la même chose que $\Pr(X_1 = 1) = p_1$, puisque la distribution est 1-localisable.

Pour la deuxième ligne, il s'agir de $\Pr(X_1 = 1, X_3 = 1) = p_2$ par définition.

Pour la troisième ligne, nous utilisons le fait que $\Pr(X_1 = 1, X_2 = 0, X_3 = 1) = p_2$ et

$$\Pr(X_1 = 1, X_2 = 0, X_3 = 0) + \Pr(X_1 = 1, X_2 = 0, X_3 = 1)$$

$$=$$

$$\Pr(X_1 = 1, X_2 = 0) = \Pr(X_1 = 1) = p_1$$

C'est plus ou moins le même raisonnement pour la quatrième ligne.

Enfin, pour la dernière ligne, nous utilisons le fait que la somme des probabilités doit faire 1.

Pour qu'une telle distribution existe, il faut et il suffit que toutes les probabilités soient positives, ce qui donne $0 \leq p_2 \leq p_1$ et $0 < p_1 < 1/2$.

Nous pourrions remarquer plusieurs choses au sujet de ces contraintes de positivité :

Tout d'abord, la troisième et la quatrième ligne imposent exactement les mêmes contraintes : il y a répétition et donc **redondance**.

Enfin, si nous ignorons cette redondance et les deux premières lignes (qui disent essentiellement que les variables doivent être positives), les contraintes restantes peuvent s'écrire sous forme matricielle comme suit :

$$\begin{bmatrix} 3 & -1 \\ -1 & 1 \end{bmatrix} \cdot \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} \leq \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

4. **$\mathbf{n = 4}$**

Nous en restons toujours à deux variables p_1 et p_2, et nous avons le système suivant :

(x_1, x_2, x_3, x_4)	$\Pr(X_1 = x_1, X_2 = x_2, X_3 = x_3, X_4 = x_4)$
(0,1,0,1)	p_2
(1,0,1,0)	p_2
(1,0,0,1)	p_2
(1,0,0,0)	$p_1 - 2p_2$
(0,0,0,1)	$p_1 - 2p_2$
(0,1,0,0)	$p_1 - p_2$
(0,0,1,0)	$p_1 - p_2$
(0,0,0,0)	$1 - 4p_1 + 3p_2$

Pour les trois premières lignes, il s'agit de $\Pr(X_1 = 1, X_3 = 1)$ (c'est direct pour la deuxième ligne et nous pouvons nous y ramener grâce à la 1-localisabilité pour les deux autres lignes).

Pour la quatrième (et la cinquième) ligne, nous utilisons le fait que

$$\begin{aligned} p_1 &= \Pr(X_1 = 1) \\ &= \Pr(X_1 = 1, X_2 = 0, X_3 = 0, X_4 = 0) \\ &+ \Pr(X_1 = 1, X_2 = 0, X_3 = 1, X_4 = 0) \\ &+ \Pr(X_1 = 1, X_2 = 0, X_3 = 0, X_4 = 1) \\ &= \Pr(X_1 = 1, X_2 = 0, X_3 = 0, X_4 = 0) + p_2 + p_2 \end{aligned}$$

Pour la sixième (et la septième) ligne, nous utilisons le fait que

$$\begin{aligned} p_1 &= \Pr(X_1 = 1) \\ &= \Pr(X_2 = 1) \\ &= \Pr(X_1 = 0, X_2 = 1, X_3 = 0, X_4 = 0) \\ &+ \Pr(X_1 = 0, X_2 = 1, X_3 = 0, X_4 = 1) \\ &= \Pr(X_1 = 0, X_2 = 1, X_3 = 0, X_4 = 0) + p_2 \end{aligned}$$

Enfin, pour la dernière ligne, nous utilisons le fait que la somme des probabilités doit faire 1.

Pour qu'une telle distribution existe, il faut et il suffit que toutes les probabilités soient positives.

Nous pouvons remarquer que, comme pour le cas $n = 3$ il y a **redondance** puisque certaines contraintes sont répétées mais aussi

parce que certaines contraintes sont plus fortes que d'autres : par exemple, $p_1 - 2p_2 \geq 0$ est plus fort que $p_1 - p_2 \geq 0$.

Enfin, en ignorant ces redondances, la contrainte $p_2 \geq 0$ et en prenant les contraintes les plus fortes, le système à résoudre peut s'écrire sous sous forme matricielle :

$$\begin{bmatrix} 4 & -3 \\ -1 & 2 \end{bmatrix} \cdot \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} \leq \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

6.2 Dérivation d'un système linéaire d'équations

Après avoir étudié le système pour des petites valeurs de n, nous allons dériver le système pour n (pair) quelconque dans cette section.

Plus précisément, nous prendrons $n = 2\ell$ pour un $\ell \in \mathbb{N}$. Nous commençons par introduire un formalisme et des règles algébriques qui seront très utiles pour la dérivation du système.

Formalisme algébrique et dérivation du système. Rappelons d'abord la définition de l'ensemble des mots binaires à particules dures de longueur n :

$$\begin{aligned} \mathcal{I}_n := \{x = x_1 x_2 \dots x_n \mid x \in \{0,1\}^n, \\ \forall 1 < i < n \; x_i = 1 \Rightarrow (x_{i-1} = 0 \text{ et } x_{i+1} = 0)\}. \end{aligned}$$

Dans ce qui suit, nous allons considérer des sous-ensemble de $\mathcal{I}_n$, notamment des sous-ensembles de type

$$\bigcup_{u \in \{0,1\}^i} \{sut \mid sut \in \mathcal{I}_n,\ |s| + |u| + |t| = n\}.$$

Nous utiliserons l'abbréviation (la notation) $s \star^i t$ pour un tel sous-ensemble. Nous utiliserons également cette notation pour les probabilités sur les éléments de $\mathcal{I}_n$. En particulier :

$$\sum_{u \in \{0,1\}^i} \Pr(sut) =: \Pr(s \star^i t)$$

Nous considérons ℓ variables $p_1, \dots, p_\ell$, et une fonction $\Lambda_n : \{0,1\}^n \to \mathbb{Z}[p_1, \dots, p_\ell]$.

Soit $p_0 := \sum_{s \in \{0,1\}^n} \Lambda_n(s)$.

Nous étendons le domaine de Λ_n à $\{0,1,\star\}^n$ à l'aide de la règle suivante :

(R0) $\Lambda_n(s\star t) = \Lambda_n(s0t) + \Lambda_n(s1t)$ pour tout s,t tels que $|s|+|t| = n-1$.

En répétant la règle (R0) jusqu'à ce qu'il ne reste plus que le symbole $\star$, nous obtenons

$$\Lambda_n(\star^n) = \sum_{s\in\{0,1\}^n} \Lambda_n(s) = p_0$$

Nous définissons également les propriétés suivantes :

(R1) $\Lambda_n(s) = 0$ si $s \in \{0,1\}^n \setminus \mathcal{I}_n$.

(R2) $\Lambda_n(s\star t\star) = \Lambda_n(s\star\star t) = \Lambda_n(\star s\star t)$ pour s,t tels que $|s|+|t| = n-2$.

(R3) $\Lambda_n(s\star\star t) = \Lambda_n(t\star\star s)$ pour s,t tels que $|s|+|t| = n-2$.

(R4) $\Lambda_n((1\star)^i\star^{n-2i}) = p_i$ pour $i \in \{1,\dots,\ell\}$.

À partir de maintenant, nous ne nous intéresserons qu'aux fonctions Λ_n étendues sur $\{0,1,\star\}^n$ grâce à (R0) et vérifiant les règles (R1), (R2) et (R4) [2], et nous introduisons le système suivant :

Système 1. *$p_i \geq 0$ et $\Lambda_{2\ell}(s) \geq 0$, pour tout $i \in \{1,\dots,\ell\}$ et $s \in \mathcal{I}_{2\ell}$.*

Nous avons alors le résultat suivant :

Théorème 5. *Soient $p_1,\dots,p_\ell \in [0,1]$. Il existe une distribution de probabilité 1-localisable $\Pr$ sur $\mathcal{I}_{2\ell}$ telle que $\Pr((1\star)^i\star^{2\ell-2i}) = p_i$ pour tout $i \in \{1,\dots,\ell\}$ si et seulement si le Système 1 est satisfait avec $\Lambda_{2\ell}(s) = \Pr(s)$ et $p_0 = 1$.*

Pour prouver ce théorème, nous aurons besoin des résultats intermédiaires suivants :

Lemme 2. *La valeur $\Lambda_n(s)$ s'exprime comme une fonction linéaire de $p_1,\dots,p_\ell$. De plus, elle est uniquement déterminée par $p_0,\dots,p_\ell$ pour tout $s \in \mathcal{I}_n$.*

Démonstration. Nous prouvons ce résultat par récurrence sur le nombre de 0 dans le mot s. S'il n'y a pas de 0 dans s, alors $\Lambda_n(s) = p_i$ pour un $i \in \{1,\dots,\ell\}$. Nous pouvons le voir grâce aux règles (R2) et (R4). S'il y a un 0 dans s, nous pouvons écrire $s = u0v$ avec u et v des mots (potentiellement vides) qui peuvent contenir ou non des 0. Alors, en utilisant la règle (R0), nous avons :

$$\Lambda_n(s) = \Lambda_n(u0v) = \Lambda_n(u\star v) - \Lambda_n(u1v)\,.$$

2. Comme nous le verrons, (R2) implique (R3) dans le cas qui nous intéresse.

Les mots $u\star v$ et $u1b$ ont au moins un 0 de moins que le mot $u0v$. Par conséquent, l'hypothèse de récurrence garantit que la valeur de $\Lambda_n(u\star v)$ et de $\Lambda_n(u1v)$ est uniquement déterminée.

□

Lemme 3. *L'unique fonction Λ_n définie dans le Lemme 2 satisfait la Propriété (R3).*

Démonstration. Nous devons vérifier que la fonction Λ_n définie dans le Lemme 2 satisfait $\Lambda_n(s\star\star t) = \Lambda_n(t\star\star s)$. Comme dans le Lemme 2, nous allons prouver ceci par récurrence double sur le nombre de 0 dans s et dans t. S'il n'y a aucun 0 dans s ni dans t, pour les mêmes raisons que ci-dessus, nous aurons $\Lambda_n(s) = p_i$ pour un certain p_i. S'il n'y a aucun 0 dans s mais que $t = r0v$, où r et v sont des mots potentiellement vides qui peuvent contenir des 0 ou non, nous avons, en utilisant la règle (R0) et l'hypothèse de récurrence sur r et v :

$$\begin{aligned}\Lambda_n(s\star\star r0v) &= \Lambda_n(s\star\star r\star v) - \Lambda_n(s\star\star r1v)\\ &= \Lambda_n(r\star v\star\star s) - \Lambda_n(r1v\star\star s)\\ &= \Lambda_n(r0v\star\star s)\\ &= \Lambda_n(t\star\star s)\end{aligned}$$

Ainsi, la propriété est vérifiée pour un s ne contenant pas de 0 et pour t quelconque. Enfin, il faut prouver par récurrence sur le nombre de 0 dans s que la propriété est vérifiée pour s et t quelconques. Nous pouvons prouver ceci en utilisant essentiellement la même procédure.

□

Remarquons que le Lemme 3 signifie que pour les éléments de $\mathcal{I}_n$, la propriété (R2)(qui correspond à une résistance translationnelle à la localisation) entraîne la propriété (R3) (qui correspond à une résistance permutationnelle à la localisation). Mais cela n'est pas nécessairement vrai en général et pour d'autres problèmes il serait peut-être nécessaire d'imposer les deux règles (R2) et (R3) dès le départ.

Un sous-système plus simple. Nous définissions maintenant

$$\mathcal{S}_n := \left\{(10)^k 0^{n-2k} : k \in \{0, \dots, \ell\}\right\}$$

Remarquons que $\mathcal{S}_n \subset \mathcal{I}_n$. Nous considérons le sous-système suivant du Système 1 :

Système 2. $p_i \geq 0$ *et* $\Lambda_{2\ell}(s) \geq 0$, *pour tout* $i \in \{1, \dots, \ell\}$ *et* $s \in \mathcal{S}_{2\ell}$.

Nous avons alors le résultat suivant :

Lemme 4. *Pour tout* $s \in \mathcal{I}_n$, *il existe* $(a_t)_{t\in\mathcal{S}_n}$, $a_t \in \mathbb{N}$, *telle que* $\Lambda_n(s) = \sum_{t\in\mathcal{S}_n} a_t \Lambda_n(t)$.

Démonstration. Nous commençons par définir les deux opérations suivantes :

(O1) Remplacer "010" par "⋆1⋆".

(O2) Déplacer "1⋆" au début du mot.

Nous pouvons mettre n'importe quel mot $s \in \mathcal{I}_{2\ell}$ sous une forme normale, que nous dénoterons par $\mathrm{NF}(s)$, de la manière suivante : répéter l'opération (O1) jusqu'à ce qu'il ne reste plus aucune instance de "010" dans le mot, ensuite répéter l'opération (O2) jusqu'à ce que tous les sous-mots de la forme "1⋆" se trouvent au début du mot.

Alors nous avons $\Lambda_{2\ell}(s) = \Lambda_{2\ell}(\mathrm{NF}(s))$. En effet, en appliquant (O1), nous ne changeons pas la valeur de $\Lambda(s)$ à cause des règles (R0) et (R1). De même pour (O2) à cause de la règle ((R3)). Il est facile à voir que pour n'importe quel s, il y a $a, b_1, \dots, b_k, c_1, \dots, c_k$ avec $2a + \sum_{i=1}^{k}(b_i + c_i) = 2\ell$, tels que $\mathrm{NF}(s) = (1\star)^a 0^{b_1} \star^{c_1} \dots 0^{b_k} \star^{c_k}$.

Remarquons que si nous tronquons $\mathrm{NF}(s)$ à la longueur $2a + b_1$, nous obtenons un élément de $\mathcal{S}_{2a+b_1}$ [3]. Étant donné un mot $\mathrm{NF}(s)$, nous appelons la *longueur de correction* la longueur maximale K à laquelle nous pouvons tronquer $\mathrm{NF}(s)$ tel que le mot ainsi obtenu est un élément de $\mathcal{S}_K$. La longueur de correction admet comme borne supérieure 2ℓ. Nous avons, en utilisant la règle (R0) :

3. En effet, cette troncation nous donne $(1\star)^a 0^{b_1}$ qui est la même chose que $(10)^a 0^{b_1} \in \mathcal{S}_{2a+b_1}$ en utilisant (R0) et (R1).

Si $c_1 \geq 2$:

$$\begin{aligned}
\Lambda_{2\ell}((1\star)^a 0^{b_1}\star^{c_1}\cdots 0^{b_k}\star^{c_k}) &= \Lambda_{2\ell}((1\star)^a 0^{b_1} 0\star^{c_1-1}\cdots 0^{b_k}\star^{c_k}) \\
&\quad + \Lambda_{2\ell}((1\star)^a 0^{b_1} 1\star^{c_1-1}\cdots 0^{b_k}\star^{c_k}) \\
&= \Lambda_{2\ell}((1\star)^a 0^{b_1+1}\star^{c_1-1}\cdots 0^{b_k}\star^{c_k}) \\
&\quad + \Lambda_{2\ell}((1\star)^a 0^{b_1-1} 01\star\star^{c_1-2}\cdots 0^{b_k}\star^{c_k}) \\
&= \Lambda_{2\ell}((1\star)^a 0^{b_1+1}\star^{c_1-1}\cdots 0^{b_k}\star^{c_k}) \\
&\quad + \Lambda_{2\ell}((1\star)^a 0^{b_1-1}\star 1\star\star^{c_1-2}\cdots 0^{b_k}\star^{c_k}) \quad \text{par (O1)} \\
&= \Lambda_{2\ell}((1\star)^a 0^{b_1+1}\star^{c_1-1}\cdots 0^{b_k}\star^{c_k}) \\
&\quad + \Lambda_{2\ell}((1\star)^{a+1} 0^{b_1-1}\star\star^{c_1-2}\cdots 0^{b_k}\star^{c_k}) \quad \text{par (O2)} \\
&= \Lambda_{2\ell}((1\star)^a 0^{b_1+1}\star^{c_1-1}\cdots 0^{b_k}\star^{c_k}) \\
&\quad + \Lambda_{2\ell}((1\star)^{a+1} 0^{b_1-1}\star^{c_2-1}\cdots 0^{b_k}\star^{c_k})
\end{aligned}$$

Si $c_1 = 1$:

$$\begin{aligned}
\Lambda_{2\ell}((1\star)^a 0^{b_1}\star 0^{b_2}\cdots 0^{b_k}\star^{c_k}) &= \Lambda_{2\ell}((1\star)^a 0^{b_1} 00^{b_2}\cdots 0^{b_k}\star^{c_k}) \\
&\quad + \Lambda_{2\ell}((1\star)^a 0^{b_1} 10^{b_2}\cdots 0^{b_k}\star^{c_k}) \\
&= \Lambda_{2\ell}((1\star)^a 0^{b_1+b_2+1}\star^{c_2}\cdots 0^{b_k}\star^{c_k}) \\
&\quad + \Lambda_{2\ell}((1\star)^a 0^{b_1-1}\star 1\star 0^{b_2-1}\cdots 0^{b_k}\star^{c_k}) \quad \text{par (O1)} \\
&= \Lambda_{2\ell}((1\star)^a 0^{b_1+b_2+1}\star^{c_2}\cdots 0^{b_k}\star^{c_k}) \\
&\quad + \Lambda_{2\ell}((1\star)^{a+1} 0^{b_1-1}\star 0^{b_2-1}\cdots 0^{b_k}\star^{c_k}) \quad \text{par (O2)}
\end{aligned}$$

Les deux termes dans la somme sont toujours de longueur 2ℓ. La longueur de correction du premier terme est au moins $2a + b_1 + 1 = K + 1$, et celle du deuxième terme est au moins $2a + 2 + b_1 - 1 = K + 1$. Ainsi, la longueur de correction augmente à chaque itération de la procédure ci-dessus et finira par atteindre la borne 2ℓ. Nous remarquons de plus qu'il y a des additions mais aucune soustraction à chaque itération, ce qui nous permet de conclure.

□

Le Lemme 4 nous permet de prouver la proposition suivante :

Proposition 3. *Le Système 1 est équivalent au Système 2, i.e., toute solution de l'un est aussi solution de l'autre.*

Démonstration. Il est évident que toute solution du Système 1 est aussi solution du Système 2. Le fait que cette implication vaut aussi dans le sens

inverse découle de cette conséquence immédiate du Lemme 4 : pour tout $s \in \mathcal{I}_n$ il y a un $t \in \mathcal{S}_n$ tel que $\Lambda(s) \geq \Lambda(t) \geq 0$.

□

Cela nous permet de focaliser notre attention sur le Système 2. Nous avons le résultat suivant sur la fonction Λ_n appliquée aux éléments de $\mathcal{S}_n$:

Lemme 5. $\Lambda_n((10)^k 0^{n-2k}) = \sum_{i=0}^{\ell-k} (-1)^i \binom{2\ell-2k+1-i}{i} p_{k+i}$, *pour* $k \in \{0, \ldots, \ell\}$.

Démonstration. Soit $Q_{v,u} = \Lambda_n((1\star)^v \star^{n-2v-u} 0^u)$. Nous allons prouver que :

$$Q_{v,u} = \sum_{i=0}^{\lceil u/2 \rceil} (-1)^i \binom{u+1-i}{i} p_{v+i} \,. \tag{6.1}$$

Ensuite, nous remarquons que $\Lambda_n((10)^k 0^{n-2k}) = Q_{k,2\ell-2k}$ et $n = 2\ell$, et nous utilisons l'équation (6.1) pour conclure.

Prouvons donc l'équation (6.1).

Fixons v. Nous commençons par montrer que l'équation (6.1) vaut pour $u = 0$ et pour $u = 1$. Nous avons $Q_{v,0} = \Lambda_n((1\star)^v \star^{n-2v}) = p_v$, ceci est vrai même pour $v = 0$, car $\Lambda_n(\star^n) = p_0$ en itérant la règle (R0).

$$\begin{aligned} Q_{v,1} &= \Lambda_n((1\star\cdots\star1)^v \star^{n-2v-1} 0) \\ &= \Lambda_n((1\star\cdots\star1)^v \star^{n-2v-1} \star) - \Lambda_n((1\star\cdots\star1)^v \star^{n-2v-1} 1) \\ &= \Lambda_n((1\star\cdots\star1)^v \star^{n-2v}) - \Lambda_n((1\star\cdots\star1)^{v+1} \star^{n-2v-1}) \text{ par la règle (R2)} \\ &= p_v - p_{v+1} \end{aligned}$$

Nous pouvons montrer la suite par récurrence, en supposant que l'égalité est vraie pour tout $w < u$. Ensuite, en utilisant le même raisonnement que pour $Q_{v,1}$, nous avons

$$\begin{aligned} \Lambda_n((1\star)^v \star^{n-2v-u} 0^u) &= \Lambda_n((1\star)^v \star^{n-2v-u} \star 0^{u-1}) - \Lambda_n((1\star)^v \star^{n-2v-u} 10^{u-1}) \\ &= \Lambda_n((1\star)^v \star^{n-2v-u+1} 0^{u-1}) - \Lambda_n((1\star)^v \star^{n-2v-u} 1\star 0^{u-2}) \\ &= \Lambda_n((1\star)^v \star^{n-2v-u+1} 0^{u-1}) - \Lambda_n((1\star)^{v+1} \star^{n-2v-u} 0^{u-2}) \\ &= Q_{v,u-1} + Q_{v+1,u-2} \end{aligned}$$

Nous utilisons ceci pour prouver la formule pour $Q_{v,2u+1}$ (la preuve pour $Q_{v,2u}$

étant analogue).

$$
\begin{aligned}
Q_{v,2u+1} &= Q_{v,2u} - Q_{v+1,2u-1} \\
&= \sum_{i=0}^{u} (-1)^i \binom{2u+1-i}{i} p_{v+i} - \left(\sum_{i=0}^{u} (-1)^i \binom{2u-i}{i} p_{v+i+1} \right) \\
&= \sum_{i=0}^{u} (-1)^i \binom{2u+1-i}{i} p_{v+i} + \sum_{i=1}^{u} (-1)^i \binom{2u+1-i}{i-1} p_{v+i} - (-1)^u p_{v+u+1} \\
&= p_v + \sum_{i=1}^{u} (-1)^i \left(\binom{2u+1-i}{i} + \binom{2u+1-i}{i-1} \right) p_{v+i} + (-1)^{u+1} p_{v+u+1} \\
&= p_v + \sum_{i=1}^{u} (-1)^i \binom{2u+2-i}{i} p_{v+i} + (-1)^{u+1} p_{v+u} \\
&= \sum_{i=0}^{u+1} (-1)^i \binom{2u+2-i}{i} p_{v+i} \,.
\end{aligned}
$$

Ainsi, nous avons prouvé l'équation (6.1).

□

Pour conclure sur cette section, nous avons démontré les points suivants que nous avions déjà relevés en étudiant le problème pour des petites valeurs de n dans la Section 6.1 :

- Tout d'abord, dans le Théorème 5, que l'existence d'une distribution de probabilités 1-localisable sur $\mathcal{I}_{2\ell}$ est équivalente à la résolubilité d'un système de $O(|\mathcal{I}_n|)$, soit $\exp(O(n))$ inégalités

  $$\Lambda_{2\ell}(s) \geq 0 \text{ pour } s \in \mathcal{I}_{2\ell}$$

  où chaque $\Lambda_{2\ell}(s)$ est une fonction **linéaire** des $p_i = \Lambda_{2\ell}((1\star)^i \star^{n-2i})$, $1 \leq i \leq \ell$.

- Ensuite, dans la Proposition 3, nous avons démontré qu'il y a beaucoup de **redondance** dans ce système, et qu'il existe un sous-ensemble $\mathcal{S}_{2\ell} \subset \mathcal{I}_{2\ell}$ de taille ℓ qui « résume » les contraintes, dans la mesure où les inégalités imposées par ce sous-ensemble sont au moins aussi fortes que les inégalités imposées par $\mathcal{I}_{2\ell}$.

- Finalement, nous nous retrouvons donc avec un système à ℓ variables et ℓ contraintes.

6.3 Résolution du système linéaire

6.3.1 Formulation du problème et stratégie

Précisons davantage le problème que nous souhaiterions résoudre : nous nous intéressons aux valeurs possibles pour les variables $p_1, \dots, p_\ell$ (qui, rappelons le, représentent la probabilité d'apparition de i « 1 » à distance au moins un les uns des autres, pour $i = 1, \dots, \ell$. Plus précisément, nous nous intéressons à la valeur maximale qui peut être prise par ces variables, en particulier p_1 (puisque, comme nous l'avons vu, cela nous permet d'obtenir le nombre minimum de couleurs nécessaires pour le problème de coloration 1-localisable du chemin).

Nous souhaiterions donc maximiser p_1 sous les contraintes imposées par le Système 2. Plus précisément, en utilisant le Lemme 5 et le fait que $p_0 = 1$ par définition, le problème qui nous intéresse est le suivant :

$$\text{Maximiser} \quad p_1 \quad \text{sous les contraintes}$$

$$\begin{cases} p_i & \geq 0, \quad i \in \{1, \dots, \ell\} \\ \sum_{i=0}^{\ell-k} (-1)^i \binom{2\ell-2k+1-i}{i} p_{k+i} & \geq 0, \quad k \in \{0, \dots, \ell-1\}\,. \end{cases} \tag{6.2}$$

Il s'agit donc d'un problème qui fait intervenir ℓ variables $p_1, \dots, p_l$, où le but est de maximiser une fonction f de ces variables (ici, $f(p_1, \dots, p_\ell) = p_1$) sous certaines contraintes linéaires de ces variables. Un tel problème est appelé un **problème de programmation linéaire** ou **problème LP** (de l'anglais *linear programming*), dont voici une définition formelle [Chv83] :

Définition 23. *Soient $m, n \in \mathbb{N}$, $c_i, b_j, a_{i,j} \in \mathbb{R}$ pour $1 \leq i \leq m$ et $1 \leq j \leq n$. Soient $\mathbf{c} = (c_1, \dots, c_n)^\mathsf{T}$, $\mathbf{b} = (b_1, \dots, b_m)^\mathsf{T}$, $\mathbf{x} = (x_1, \dots, x_n)^\mathsf{T}$ et soit $\mathbf{A} = (a_{i,j})$ une matrice $m \times n$. Un problème ayant la forme*

$$\textit{Maximiser } \mathbf{c}^\mathsf{T}\mathbf{x}, \textit{ sous les contraintes } \mathbf{A}\mathbf{x} \leq \mathbf{b} \textit{ et } \mathbf{x} \geq \mathbf{0},$$

est appelé un problème de programmation linéaire sous forme standard *(ou problème LP sous forme standard).*

*L'expression linéaire $\mathbf{c}^\mathsf{T}\mathbf{x}$ est appelée la **fonction objectif**.*

*$\mathbf{A}\mathbf{x} \leq \mathbf{b}$ et $\mathbf{x} \geq \mathbf{0}$ sont appelés les **contraintes**, les dernières étant plus précisément les contraintes de non-négativité.*

Nous pouvons montrer que le Problème (6.2) se formule comme suit en utilisant les notations introduites dans la Définition 23

Proposition 4. *Le problème LP $\mathbf{P}_\ell$ de la maximisation de la valeur de p_1 pour la 1-localisabilité sur $\mathcal{I}_{2\ell}$ est associé à la matrice et aux vecteurs $(\mathbf{A}_\ell, \mathbf{c}_\ell, \mathbf{b}_\ell, \mathbf{x}_\ell)$, où*

— $\mathbf{x}_\ell = (p_1, \ldots, p_\ell)$

— $\mathbf{c}_\ell = (1, 0, \ldots, 0)$

— $\mathbf{b}_\ell = (p_0, 0, \ldots, 0)$

— $a_{i,j} = (-1)^{i+j}\binom{2\ell+2-(i+j)}{j-i+1}$ *($a_{i,j} = 0$ if $j - i + 1 < 0$).*

Démonstration. Il y a ℓ variables $p_1, \ldots, p_\ell$ et ℓ contraintes associées au problème. Par conséquent, $\mathbf{x}_\ell = (p_1, \ldots, p_\ell)$. Comme le but est de maximiser p_1, la fonction objectif est p_1, et donc $\mathbf{c}_\ell = (1, 0, \ldots, 0)$.

Il reste à donner les coefficients de la matrice $\mathbf{A}_\ell$ et ceux du vecteur $\mathbf{b}_\ell$. Pour ce faire, nous allons regarder les contraintes autres que celles de la non-négativité :

$$\Lambda_{2\ell}((10)^k 0^{n-2k}) \geq 0 \quad \text{pour } k = 0, \ldots, \ell - 1 \,.$$

Nous avons, via le lemme 5, $\Lambda_{2\ell}((10)^k 0^{n-2k}) = \sum_{j=0}^{\ell-k}(-1)^j\binom{2\ell-2k+1-j}{j}p_{k+j}$, ou encore $\sum_{j=1}^{\ell}(-1)^{j+1}\binom{2\ell+1-j}{j}p_j \leq p_0$ et $\sum_{j=0}^{\ell-k}(-1)^{j+1}\binom{2\ell-2k+1-j}{j}p_{k+j} \leq 0$ pour $1 \leq k \leq \ell - 1$. Par conséquent, $\mathbf{b}_\ell = (p_0, 0, \ldots, 0)$.

Enfin, $a_{i,j}$ est le coefficient devant p_j dans la i-ème ligne. Pour la première ligne, $i = 1$, $a_{1,j}$ vaut $(-1)^{i+1}\binom{2\ell+1-i}{j}$. Pour la ligne $i = k + 1$, le coefficient devant p_{k+j} est $(-1)^{j+1}\binom{2\ell-2k+1-j}{j}$, i.e. le coefficient devant p_j est $(-1)^{i+j}\binom{2\ell+2-(i+j)}{j-i+1}$.

Par conséquent, $a_{i,j} = (-1)^{i+j}\binom{2\ell+2-(i+j)}{j-i+1}$ ($a_{i,j} = 0$ si $j - i + 1 < 0$).

□

Après avoir résolu ce problème linéaire et donc obtenu la valeur maximale de p_1, nous imposons que la valeur de p_1 soit une valeur inférieure ou égale à cette valeur maximale et cherchons ensuite à maximiser p_2. Il se trouve que cela donne lieu à un autre problème LP très similaire « dans la forme » (nous verrons dans quelques paragraphes ce que cela signifie) au Problème (6.2). Ensuite, nous répétons cette procédure, i.e. pour $i \geq 1$, nous imposons une valeur inférieure ou égale à la valeur maximale possible pour p_j, $j < i$, puis cherchons à maximiser p_i sous les contraintes résultantes.

La question est de savoir comment nous faisons pour résoudre le problème linéaire et trouver la valeur maximale de p_1. La stratégie la plus courante consiste à avoir recours à l'ordinateur. Cependant, il se trouve que dans ce cas

particulier, le problème a une forme spéciale qui nous permet de dériver une formule pour la valeur maximale de p_1. Cette dérivation repose sur **le théorème de dualité faible**. Avant d'énoncer ce théorème, il est nécessaire de donner la définition du **problème dual** d'un problème de programmation linéaire, que nous appelons aussi le **problème primal** par opposition au problème dual.

Définition 24. *Considérons un problème LP sous forme standard comme dans la Définition 23. Son* problème LP dual *est le problème suivant sur m variables $(y_1, \ldots, y_m)^\mathsf{T} = \mathbf{y}$:*

$$\textit{Minimiser } \mathbf{b}^\mathsf{T}\mathbf{y}, \textit{ sous les contraintes } \mathbf{A}^\mathsf{T}\mathbf{y} \geq \mathbf{c} \textit{ et } \mathbf{y} \geq \mathbf{0}.$$

Nous sommes maintenant en mesure d'énoncer le théorème de dualité faible (voir [Chv83, chapitre 5]).

Théorème 6. *Si $\mathbf{x}$ et $\mathbf{y}$ sont solutions du problème primal resp. dual, alors $\mathbf{c}^\mathsf{T}\mathbf{x} \leq \mathbf{b}^\mathsf{T}\mathbf{y}$.*

Nous allons utiliser une conséquence particulière de ce théorème :

Corollaire 1. *Si $\mathbf{x}$ et $\mathbf{y}$ sont solutions du problème primal resp. dual, telles que $\mathbf{c}^\mathsf{T}\mathbf{x} = \mathbf{b}^\mathsf{T}\mathbf{y}$, alors il existe une solution optimale et cette valeur optimise les fonctions objectif des deux problèmes.*

Nous utilisons ce résultat afin de démontrer le théorème général suivant :

Théorème 7. *Soit $\mathbf{A_n}$ une matrice $n \times n$ de la forme*

$$\begin{bmatrix} a_{1,1} & a_{1,2} & \ldots & \ldots & a_{1,n-1} & a_{1,n} \\ -1 & a_{2,2} & \ldots & \ldots & a_{2,n-1} & a_{2,n} \\ 0 & -1 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & -1 & a_{n-1,n-1} & a_{n-1,n} \\ 0 & 0 & \ldots & 0 & -1 & a_{n,n} \end{bmatrix}$$

Considérons le problème de maximisation LP $\mathbf{P_n}$ associé à $(\mathbf{A_n}, \mathbf{c_n}, \mathbf{b_n}, \mathbf{x_n})$, avec $\mathbf{b_n} = (b, 0, \ldots, 0)^\mathsf{T}$, $\mathbf{c_n} = (c, 0, \ldots, 0)^\mathsf{T}$ et $\mathbf{x_n} = (x_1, \ldots, x_n)^\mathsf{T}$ des vecteurs de longueur n. Alors la valeur optimale de la fonction objectif de P_n est obtenue en résolvant le cas particulier $\mathbf{A_n x_n} = \mathbf{b_n}$. De plus, cette valeur optimale est $\frac{u_n}{u_{n+1}} bc$, où la suite $(u_k)_{k\geq 1}$ est donnée par $u_1 = 1$ et $u_{k+1} = \sum_{i=1}^{k} a_{n-k+1,n-k+i}\, u_{k+1-i}$.

Pour prouver ce théorème, nous allons d'abord considérer le problème primal, et résoudre le cas particulier $\mathbf{A_n x_n} = \mathbf{b_n}$. Ensuite, nous considérerons

le problème dual associé que nous résoudrons dans le cas particulier où $\mathbf{A_n y_n} = \mathbf{c_n}$. Comme nous le verrons, nous obtiendrons la même valeur pour la fonction objectif dans ces deux cas, i.e. $\mathbf{c_n}^\mathsf{T}\mathbf{x_n} = \mathbf{b_n}^\mathsf{T}\mathbf{y_n}$, ce qui nous permettra de conclure qu'il s'agit de la valeur optimale grâce au théorème de dualité. Les deux sections suivantes sont dédiées à la résolution du problème primal puis du problème dual.

Mais avant d'embarquer dans cette aventure, revenons au problème (6.2) qui était notre point de départ :

- Il est facile de voir, notamment grâce à la Proposition 4, que ce problème est exactement de la forme donnée dans le Théorème 7, avec $b = c = 1$.
- Une fois que nous avons utilisé le Théorème 7 pour trouver la valeur maximale de p_1, nous imposons que la valeur de p_1 soit inférieure ou égale à cette valeur. Or, cela rendra redondante la première contrainte. Nous nous retrouverons donc avec le problème de maximiser p_2 sous les contraintes restantes. En faisant passer le « -1 » au début de l'autre côté, cela nous donne un problème qui a exactement la même forme, avec $c = 1$ et $b = p_1$.

  Plus généralement, en répétant cette procédure et en imposant des valeurs pour p_j où $j < i$ pour un $i \geq 2$, nous nous retrouverons avec un problème LP de maximisation de LP ayant la forme donnée dans le Théorème 7 et $c = 1$ et $b = p_{i-1}$.

Ainsi, en appliquant le Théorème 7 à chacun de ces problèmes et en utilisant un résultat prouvé dans le Chapitre 7, nous montrerons que, pour le Problème (6.2) nous avons :

Théorème 8. *Toute solution* $(p_1, \ldots, p_\ell) \in \mathbb{R}^\ell$ *du système d'inégalités* (6.2) *satisfait* $p_i \leq (c_\ell / c_{\ell+1}) \cdot p_{i-1} \leq (c_{\ell-i+1}/c_{\ell+1}) \cdot p_0$. *Il est possible d'avoir égalité.*

Ce qui a pour conséquence le Théorème 3.

6.3.2 Étude du problème primal

Comme nous l'avons dit, la première étape de la preuve du Théorème 7 consiste à regarder un problème LP primal

Problème 1. *Maximiser* $\mathbf{c_n}^\mathsf{T}\mathbf{x_n}$, *sous les contraintes* $\mathbf{A_n x_n} \leq \mathbf{b_n}$ *et* $\mathbf{x_n} \geq \mathbf{0}$.

où les $(\mathbf{A_n}, \mathbf{c_n}, \mathbf{b_n}, \mathbf{x_n})$ sont tels que donnés dans l'énoncé de ce théorème, et de résoudre ce problème dans le cas particulier où $\mathbf{A_n x_n} = \mathbf{b_n}$.

Nous avons alors le résultat suivant :

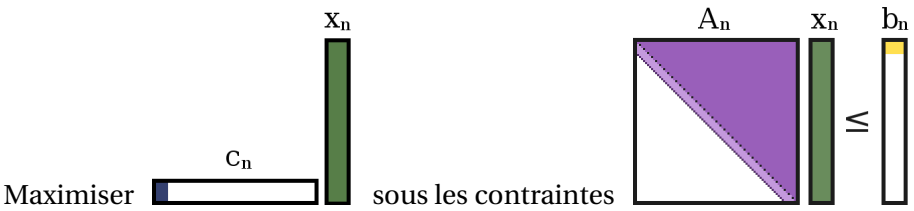


FIGURE 6.1 – Le problème LP 1.

Proposition 5. *La valeur de la fonction objectif du Problème 1 dans le cas* $\mathbf{A_n x_n} = \mathbf{b_n}$ *est* $\frac{u_n}{u_{n+1}} bc$.

La preuve de cette proposition repose sur plusieurs résultats intermédiaires.

Lemme 6. *Dans le cas* $\mathbf{A_n x_n} = \mathbf{b_n}$, *il existe* $(\mu_j)_{1 \leq j \leq n}$ *tel que* $x_{n-j} = \mu_j x_{n-j+1}$ *pour* $j \neq n$, *et* $b = \mu_n x_1$.

Démonstration. Nous allons faire une démonstration par récurrence. Notons d'abord que cela est vrai lorsque $j = 1$ puisque :

$$0 \;=\; \sum_{j=0}^{n} a_{n,j} x_j \;=\; -x_{n-1} + a_{n,n} x_n \;.$$

Ainsi $x_{n-1} = \mu_1 x_n$ avec $\mu_1 = a_{n,n}$.

Maintenant, prenons $k \geq 0$ et supposons qu'il est vrai que $x_{n-j} = \mu_j x_{n-j+1}$ pour $0 \leq j \leq k$. En particulier, nous avons $x_{n-k-1+i} = \mu_{k+1-i} x_{n-k+i}$ pour tout $1 \leq i \leq k+1$.

Nous allons maintenant regarder ce que la contrainte donnée par la $(n-k)$-ème ligne de la matrice entraîne lorsque $k \neq n-1$:

$$
\begin{aligned}
0 &= \sum_{j=1}^{n} a_{n-k,j} x_j \\
&= \sum_{j=n-k-1}^{n} a_{n-k,j} x_j \\
&= \sum_{i=0}^{k+1} a_{n-k,n-k-1+i} x_{n-k-1+i} \\
&= -x_{n-k-1} + \sum_{i=1}^{k} a_{n-k,n-k-1+i} x_{n-k-1+i} + a_{n-k,n} x_n \\
&= -x_{n-k-1} + \sum_{i=1}^{k} a_{n-k,n-k-1+i} \mu_{k+1-i} x_{n-k+i} + a_{n-k,n} x_n \\
&= \sum_{i=1}^{k+1} a_{n-k,n-k-1+i} \prod_{j=1}^{k+1-i} \mu_j x_n - x_{n-k-1} \qquad \text{where } \prod_{j=1}^{0} \mu_j = 1 \text{ par convention}
\end{aligned}
$$

Par conséquent,

$$
\begin{aligned}
x_{n-k-1} &= \sum_{i=1}^{k+1} a_{n-k,n-k-1+i} \prod_{j=1}^{k+1-i} \mu_j x_n \\
x_{n-k-1} &= \frac{\sum_{i=1}^{k+1} a_{n-k,n-k-1+i} \prod_{j=1}^{k+1-i} \mu_j}{\mu_1} x_{n-1} \quad \text{(en utilisant l'hypothèse de récurrence)} \\
x_{n-k-1} &= \frac{\sum_{i=1}^{k+1} a_{n-k,n-k-1+i} \prod_{j=1}^{k+1-i} \mu_j}{\prod_{j=1}^{k} \mu_j} x_{n-k} \quad \text{(en utilisant l'hypothèse de récurrence)}
\end{aligned}
$$

Enfin, si $k = n-1$, la première ligne de la matrice donne

$$\begin{aligned} b &= \sum_{j=1}^{n} a_{n-k,j} x_j \\ &= \sum_{i=1}^{n} a_{1,n-i} \prod_{j=1}^{n-i} \mu_j x_n \quad \text{(comme ci-dessus)} \\ &= \left(\sum_{i=1}^{n} a_{1,i} \prod_{j=1}^{n-i} \mu_j \right) x_n \\ b &= \frac{\sum_{i=1}^{n} a_{1,i} \prod_{j=1}^{n-i} \mu_j}{\prod_{j=1}^{n-1} \mu_j} x_1 \quad \text{(comme ci-dessus)} \end{aligned}$$

Par conséquent, $x_{n-k-1} = \mu_{k+1} x_{n-k}$, $b_1 = \mu_n x_1$ avec

$$\mu_{k+1} = \frac{\sum_{i=1}^{k+1} a_{n-k,n-k-1+i} \prod_{j=1}^{k+1-i} \mu_j}{\prod_{j=1}^{k} \mu_j}$$

□

Corollaire 2. *Soit $u_1 = 1$ et $u_i = \prod_{j=1}^{i-1} \mu_i$ pour $i \in \{2, \ldots, n+1\}$. Alors nous avons la relation de récurrence $u_{k+1} = \sum_{i=1}^{k} a_{n-k+1,n-k+i}\, u_{k+1-i}$.*

Démonstration. Il est facile à vérifier que $\mu_1 = u_2 = \frac{u_2}{u_1}$ et, plus généralement, $\mu_i = \frac{u_{i+1}}{u_i}$. Nous allons maintenant réécrire la relation de récurrence trouver pour les μ_k dans la preuve du Lemme 6 en utilisant les u_i :

$$\frac{u_{k+2}}{u_{k+1}} = \mu_{k+1} = -\frac{\sum_{i=1}^{k+1} a_{n-k,n-k-1+i}\, u_{k+2-i}}{u_{k+1}}.$$

On conclut que $u_{k+2} = \sum_{i=1}^{k+1} a_{n-k,n-k-1+i}\, u_{k+2-i}$.

□

Corollaire 3. *La suite $(u_j)_{1 \le j \le n+1}$ définie dans le Corollaire 2 satisfait $x_{n-j} = \frac{u_{j+1}}{u_j} x_{n-j+1}$ pour $j \neq n$ et $b = \frac{u_{n+1}}{u_n} x_1$.*

Démonstration. C'est une conséquence du Lemme 6 en remplaçant μ_j par $\frac{u_{j+1}}{u_j}$ (grâce au Corollaire 2).

□

Preuve de la Proposition 5. Nous utilisons le fait que $b = \frac{u_{n+1}}{u_n} x_1$ (Corollaire 3), et le fait que la fonction objectif est cx_1.

□

6.3.3 Étude du problème dual

Nous arrivons maintenant à la deuxième étape de la preuve du Théorème 7, qui consiste à regarder le problème dual du Problème 1 :

Problème 2. *Minimiser* $\mathbf{b_n}^\mathsf{T} \mathbf{y_n}$, *sous les contraintes* $\mathbf{A_n}^\mathsf{T} \mathbf{y_n} \geq \mathbf{c_n}$ *et* $\mathbf{y_n} \geq \mathbf{0}$.

Plus précisément, nous allons nous intéresser au cas particulier où $\mathbf{A_n}^\mathsf{T} \mathbf{y_n} = \mathbf{c_n}$, et démontrer le résultat suivant :

Proposition 6. *La valeur de la fonction objectif du Problème 2 dans le cas* $\mathbf{A_n}^\mathsf{T} \mathbf{y_n} = \mathbf{c_n}$ *est* $\frac{u_n}{u_{n+1}} bc$.

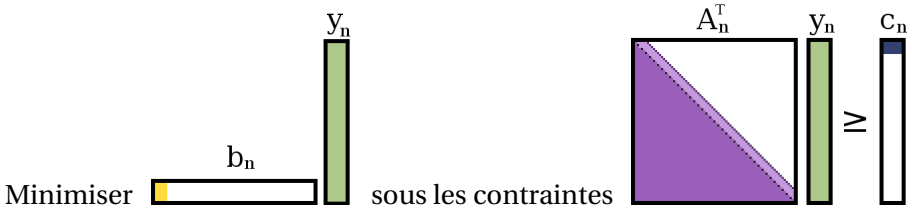


FIGURE 6.2 – Le problème LP dual 2.

Avant de continuer, il est nécessaire d'introduire quelques notations : nous allons considérer une suite de problèmes LP $(\mathbf{P_k})_{k \leq n}$ associés au problème P_n donné par $(\mathbf{A_n}, \mathbf{c_n}, \mathbf{b_n}, \mathbf{x_n})$, où $\mathbf{A_{k-1}}$ est la sous-matrice $(k-1) \times (k-1)$ en bas à droite de $\mathbf{A_k}$. En d'autres termes, $\mathbf{A_k} = (a^k_{i,j})$ avec

$$a^n_{i,j} = a_{i,j} \quad \text{et} \quad a^{k-1}_{i,j} = a^k_{i+1,j+1}.$$

De plus,

$$\mathbf{b_{k-1}} := (x_{n-k}, 0, \ldots, 0)^\mathsf{T}$$

et

$$\mathbf{x_{k-1}} = (x_{n-k}, \ldots, x_n)^\mathsf{T}$$

À partir de maintenant, nous écrirons $a^n_{i,j}$ au lieu de $a_{i,j}$ car nous allons établir des résultats faisant intervenir différentes valeurs de cet exposant.

Nous sommes maintenant en mesure d'énoncer et de prouver nos résultats.

Lemme 7. *Il existe des fonctions $f_2, \ldots, f_{n+1}$ telles que $y_i = f_i(y_1, \ldots, y_{i-1})$ pour $2 \leq i \leq n$ et $f_n(y_1, \ldots, y_{n+1}) = 0$.*

Démonstration. Il s'agit d'une réécriture de l'équation donnée par la $(i-1)$-ème ligne de la matrice, où nous avons isolé la variable y_i. Ainsi, par exemple, la ligne 1 donne pour la variable y_2 : $a^n_{1,1}y_1 - y_2 = c$, i.e., $y_2 = f_1(y_1)$ où $f_1(y_1) = a^n_{1,1}y_1 - c$.

La $(i-1)$-ème ligne donne pour y_i : $a^n_{1,i-1}y_1 + a^n_{2,i-1}y_2 + \cdots + a^n_{i-1,i-1}y_{i-1} - y_i = 0$, i.e., $y_i = f_i(y_1, \ldots, y_{i-1})$ avec $f_i(y_1, \ldots, y_{i-1}) = \sum_{k=1}^{i-1} a^n_{k,i-1}y_k$.

Finalement, la dernière ligne de la matrice donne $f_{n+1}(y_1, \ldots y_n) = 0$ avec $f_{n+1}(y_1, \ldots, y_n) = \sum_{k=1}^{n} a^n_{k,n}y_k$.

□

Proposition 7. *Pour tout $1 \leq k \leq n+1$, il y a $U_{k,n}$ et $V_{k,n}$ tels que $f_k(y_1, \ldots, y_{k-1}) = U_{k,n}\, y_1 + V_{k,n}$. De plus,*

1. $V_{k,n} = -cU_{k-1,n-1}$.
2. $$U_{k,n} = \sum_{\substack{0 \leq j \leq k-1 \\ 0 = i_0 < \cdots < i_j = k-1}} \prod_{0 \leq u \leq j-1} a^n_{i_u+1, i_{u+1}}, \text{ pour } k \geq 2.$$

Démonstration. Nous avons $y_2 = f_1(y_1) = a^n_{1,1}y_1 - c$, d'où $U_{1,n} = 1$, $V_{1,n} = 0$, $U_{2,n} = a^n_{1,1}$ $V_{2,n} = -c$. La propriété est donc vraie pour $k = 1$ et $k = 2$, et nous avons bien $V_{2,n} = -cU_{1,n-1}$.

Maintenant, prenons un $k \geq 2$, et supposons que nous avons $f_j(y_1, \ldots, y_{j-1}) = U_{j,n}y_1 + V_{j,n}$ pour tout $1 \leq j \leq k$. Alors

$$
\begin{aligned}
f_{k+1}(y_1,\ldots,y_k) &= \sum_{j=1}^{k} a_{j,k}^n y_j \quad \text{(par définition de } f_{k+1} \text{ dans le Lemme 7)} \\
&= \sum_{j=1}^{k} a_{j,k}^n \left(U_{j,n} y_1 + V_{j,n}\right) \quad \text{(en utilisant l'hypothèse de récurrence)} \\
&= U_{k+1,n} y_1 + V_{k+1,n}
\end{aligned}
$$

où $U_{k+1,n} = \sum_{j=1}^{k} a_{j,k}^n U_{j,n}$ et $V_{k+1,n} = \sum_{j=1}^{k} a_{j,k}^n V_{j,n}$.

Nous allons démontrer point (1) par récurrence et en constatant que :

$$
\begin{aligned}
\text{(*)} \quad U_{k,n-1} &= \sum_{j=1}^{k-1} a_{j,k-1}^{n-1} U_{j,n-1} \\
V_{k+1,n} &= \sum_{j=1}^{k} a_{j,k}^n V_{j,n} \\
&= \sum_{j=2}^{k} a_{j,k}^n V_{j,n} \quad \text{(parce que } V_{1,n} = 0\text{)} \\
&= \sum_{j=2}^{k} a_{j,k}^n c U_{j-1,n-1} \quad \text{(hypothèse de récurrence)} \\
&= c \sum_{j=1}^{k-1} a_{j+1,k}^n U_{j,n-1} \\
&= c \sum_{j=1}^{k-1} a_{j,k-1}^{n-1} U_{j,n-1} \quad \text{(en utilisant } a_{i,j}^{n-1} = a_{i+1,j+1}^n\text{)} \\
&= c U_{k,n-1} \quad \text{(en utilisant (*))}
\end{aligned}
$$

Nous allons maintenant démontrer le point (2), encore par récurrence.

Il est facile de voir que c'est vrai pour $k = 2$. En effet,

$$
\sum_{\substack{1 \le j \le 1 \\ 0 = i_0 < \cdots < i_j = 1}} \prod_{0 \le u \le j-1} a_{i_u+1,i_{u+1}}^n = a_{1,1}^n = U_{2,n} \,.
$$

Maintenant, prenons un $\ell \geq 3$ et supposons que l'égalité est vraie pour tous les $U_{k,n}$, $k < \ell$. Nous utilisons ensuite la relation de récurrence obtenue précédemment dans cette preuve.

$$
\begin{aligned}
U_{\ell,n} &= \sum_{k=1}^{\ell-1} a^n_{k,\ell-1} U_{k,n} \\
&= \sum_{k=1}^{\ell-1} a^n_{k,\ell-1} \left(\sum_{\substack{0\le j\le k-1 \\ 0=i_0<\cdots<i_j=k-1}} \prod_{0\le u\le j-1} a^n_{i_u+1,i_{u+1}} \right) \\
&= \sum_{k=1}^{\ell-1} \left(\sum_{\substack{0\le j\le k-1 \\ 0=i_0<\cdots<i_j=k-1}} \left(\prod_{0\le u\le j-1} a^n_{i_u+1,i_{u+1}} \right) a^n_{i_j+1,\ell-1} \right) \\
&= \sum_{\substack{0\le j\le \ell-2 \\ 0=i_0<\cdots<i_j<i_{j+1}=\ell-1}} \left(\prod_{0\le u\le j-1} a^n_{i_u+1,i_{u+1}} \right) a^n_{i_j+1,i_{j+1}} \\
&= \sum_{\substack{0\le j\le \ell-2 \\ 0=i_0<\cdots<i_j<i_{j+1}=\ell-1}} \prod_{0\le u\le j} a^n_{i_u+1,i_{u+1}} \\
&= \sum_{\substack{0\le j'\le \ell-1 \\ 0=i_0<\cdots<i_{j'}=\ell-1}} \prod_{0\le u\le j'-1} a^n_{i_u+1,i_{u+1}} \,.
\end{aligned}
$$

□

Corollaire 4. *Pour le Problème 2,* $y_1 = \frac{U_{n,n-1}}{U_{n+1,n}} c$ *dans le cas où* $\mathbf{A_n}^\mathsf{T}\mathbf{y_n} = \mathbf{c_n}$.

Démonstration. Nous utilisons la dernière contrainte $f_{n+1}(y_1,\ldots,y_n) = 0$, que l'on peut réécrire de la manière suivante : $U_{n+1,n}y_1 + V_{n+1,n} = 0$ en utilisant la Proposition 7. Encore une fois en utilisant la Proposition 7, nous utilisons $V_{n+1,n} = -cU_{n,n-1}$ et donc $U_{n+1,n}y_1 - cU_{n,n-1} = 0$.

□

Lemme 8. $U_{n+1,n} = u_{n+1}$, *où la suite* $(u_n)_{n\ge1}$ *est celle donnée dans le Corollaire 2.*

Démonstration. Le Corollaire 2 donne :

$$
u_{k+1} = \sum_{i=1}^{k} a^n_{n-k+1,n-k+i} u_{k+1-i} = \sum_{i=0}^{k-1} a^n_{n-k+1,n-i} u_{i+1} \,.
$$

La preuve consiste à réappliquer cette égalité à u_{i+1} et ainsi de suite. En effet, nous allons montrer par récurrence que :

$$u_{k+1} = \sum_{\substack{0\le j\le k\\ 0=i_j<\cdots<i_0=k}} \prod_{0\le u\le j-1} a^n_{n-i_u+1,n-i_{u+1}}\,.$$

Ceci est vrai pour $k=1$ puisque $\sum_{\substack{0\le j\le 1\\ 0=i_j<\cdots<i_0=1}} \prod_{0\le u\le j-1} a^n_{n-i_u+1,n-i_{u+1}} = a^n_{n,n} = u_2$. Prenons maintenant un $\ell\ge 1$ et supposons que l'égalité est vraie pour tous les $k\le\ell$. Alors, nous avons :

$$\begin{aligned}
u_{\ell+1} &= \sum_{k=0}^{\ell-1} a^n_{n-\ell+1,n-k} u_{k+1}\\
&= \sum_{k=0}^{\ell-1} a^n_{n-\ell+1,n-k} \sum_{\substack{0\le j\le k\\ 0=i_j<\cdots<i_0=k}} \prod_{0\le u\le j-1} a^n_{n-i_u+1,n-i_{u-1}} \text{ par hypothèse de récurrence}\\
&= \sum_{k=0}^{\ell-1} \sum_{\substack{0\le j\le k\\ 0=i_j<\cdots<i_0=k}} \left(\prod_{0\le u\le j-1} a^n_{n-i_u+1,n-i_{u+1}}\right) a^n_{n-\ell+1,n-i_0}\\
&= \sum_{\substack{0\le j\le \ell-1\\ 0=i_j<\cdots<i_0<i_{-1}=\ell}} \left(\prod_{0\le u\le j-1} a^n_{n-i_u+1,n-i_{u+1}}\right) a^n_{n-i_{-1}+1,n-i_0}\\
&= \sum_{\substack{0\le j\le \ell-1\\ 0=i_j<\cdots<i_0<i_{-1}=\ell}} \prod_{-1\le u\le j-1} a^n_{n-i_u+1,n-i_{u+1}}\\
&= \sum_{\substack{1\le j'\le \ell\\ 0=i_{j'}<\cdots<i_0=\ell}} \prod_{0\le u\le j'-1} a^n_{n-i_u+1,n-i_{u+1}}\\
&= \sum_{\substack{0\le j'\le \ell\\ 0=i_{j'}<\cdots<i_0=\ell}} \prod_{0\le u\le j'-1} a^n_{n-i_u+1,n-i_{u+1}} \text{ (pour } j'=0 \text{ le produit est vide)}
\end{aligned}$$

Ceci conclut la récurrence.

Ainsi :

$$u_{n+1} = \sum_{\substack{0\le j\le n\\ 0=i_j<\cdots<i_0=n}} \prod_{0\le u\le j-1} a^n_{n-i_u+1,n-i_{u+1}}\,. \tag{6.3}$$

Il reste à prouver que ceci est identique à $U_{n+1,n}$. Nous avons, par la Proposition 7,

$$U_{k,n} = \sum_{\substack{0\le j\le k-1\\ 0=i_0<\cdots<i_j=k-1}} \prod_{0\le u\le j-1} a^n_{i_u+1,i_{u+1}}\,.$$

Ainsi

$$U_{n+1,n} = \sum_{\substack{0\le j\le n\\ 0=i_0<\cdots<i_j=n}} \prod_{0\le u\le j-1} a^n_{i_u+1,i_{u+1}} .$$

Nous considérons à nouveau l'équation (6.3).

$$\begin{aligned} u_{n+1} &= \sum_{\substack{0\le j\le n\\ 0=i_j<\cdots<i_0=n}} \prod_{0\le u\le j-1} a^n_{n-i_u+1,n-i_{u+1}} \\ &= \sum_{\substack{0\le j\le n\\ 0=n-i_0<\cdots<n-i_j=n}} \prod_{0\le u\le j-1} a^n_{n-i_u+1,n-i_{u+1}} \\ &= \sum_{\substack{0\le j\le n\\ 0=k_0<\cdots<k_j=n}} \prod_{0\le u\le j-1} a^n_{k_u+1,k_{u+1}} \\ &= U_{n+1,n} . \end{aligned}$$

□

Preuve de la Proposition 6. Nous utilisons

$$y_1 = \frac{U_{n,n-1}}{U_{n+1,n}} c \quad \text{(Corollaire 4)}$$

$$U_{n+1,n} = u_{n+1} \quad \text{(Lemme 8)}$$

et le fait que la fonction objectif est by_1. □

6.3.4 Conclusion

Nous sommes maintenant en mesure de prouver le Théorème 7 dont nous rappelons l'énoncé :

Théorème 7. *Soit* $\mathbf{A_n}$ *une matrice* $n \times n$ *de la forme*

$$\begin{bmatrix} a_{1,1} & a_{1,2} & \dots & \dots & a_{1,n-1} & a_{1,n} \\ -1 & a_{2,2} & \dots & \dots & a_{2,n-1} & a_{2,n} \\ 0 & -1 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & -1 & a_{n-1,n-1} & a_{n-1,n} \\ 0 & 0 & \dots & 0 & -1 & a_{n,n} \end{bmatrix}$$

Considérons le problème de maximisation LP $\mathbf{P_n}$ *associé à* $(\mathbf{A_n}, \mathbf{c_n}, \mathbf{b_n}, \mathbf{x_n})$*, avec* $\mathbf{b_n} = (b, 0, \dots, 0)^\mathsf{T}$*,* $\mathbf{c_n} = (c, 0, \dots, 0)^\mathsf{T}$ *et* $\mathbf{x_n} = (x_1, \dots, x_n)^\mathsf{T}$ *des*

vecteurs de longueur n. Alors la valeur optimale de la fonction objectif de P_n est obtenue en résolvant le cas particulier $\mathbf{A_n x_n} = \mathbf{b_n}$. De plus, cette valeur optimale est $\frac{u_n}{u_{n+1}}bc$, où la suite $(u_k)_{k\geq 1}$ est donnée par $u_1 = 1$ et $u_{k+1} = \sum_{i=1}^{k} a_{n-k+1,n-k+i}\, u_{k+1-i}$.

Démonstration. Le Corollaire 2 donne une valeur particulière de la fonction objectif dans le Problème 1 : $cx_1 = \frac{u_n}{u_{n+1}}bc$.

Le Corollaire 4 et le Lemme 8 donnent une valeur particulière de la fonction objectif dans le Problème 2 : $by_1 = \frac{u_n}{u_{n+1}}bc$.

Nous utilisons le théorème de dualité (ou plutôt son corollaire 1) pour conclure qu'il s'agit de la valeur optimale de la fonction objectif. □

Enfin, il reste à prouver

Théorème 8. *Toute solution $(p_1, \ldots, p_\ell) \in \mathbb{R}^\ell$ du système d'inégalités (6.2) satisfait $p_i \leq (\mathsf{c}_\ell/\mathsf{c}_{\ell+1}) \cdot p_{i-1} \leq (\mathsf{c}_{\ell-i+1}/\mathsf{c}_{\ell+1}) \cdot p_0$. Il est possible d'avoir égalité.*

Or, nous avons,

Proposition 8. *La suite (u_n) associée au Problème (6.2) est exactement la suite des nombres de .*

Démonstration. Nous avons $u_1 = 1 = \mathsf{c}_1$ et $u_2 = a_{\ell,\ell} = 2 = \mathsf{c}_2$. Il reste à prouver que $u_{k+1} = c_{k+1}$ si pour tout $j \leq k$ nous avons$u_j = \mathsf{c}_j$.

Encore une fois, en utilisant le Corollaire 2, nous avons

$$\begin{aligned} u_{k+1} &= \sum_{i=1}^{k} a_{\ell-k+1,\ell-k+i} u_{k+1-i} \\ &= \sum_{i=1}^{k} (-1)^{i+1} \binom{2k+1-i}{i} \mathsf{c}_{k+1-i} \\ &= \sum_{i=0}^{k-1} (-1)^{i} \binom{2k-i}{i+1} \mathsf{c}_{k-i} \\ &= \mathsf{c}_{k+1} \quad \text{(voir la formule (7.2), Chapitre 7)} \end{aligned}$$

□

Preuve du Théorème 8. Pour p_1, il s'agit d'une application du Théorème 7 avec $b = c = 1$ et $(u_n) = (\mathsf{c}_n)$.

Pour p_i, $i \geq 2$, nous pouvons démontrer facilement par récurrence qu'il s'agit d'une application du Théorème 7 avec $b = p_{i-1}$, $c = 1$, $(u_n) = (\mathsf{c}_n)$ et en utilisant une hypothèse de récurrence sur p_{i-1}. □

6.4 Généralisations

6.4.1 Processus k-hard-core k-localisables

Les résultats présentés dans ce chapitre se généralisent sans grande difficulté aux processus k-localisables $(X_n)_{n\in\mathbb{N}}$ ayant une propriété que nous avons appelée « k-hard-core » (une généralisation de la propriété d'être à particules dures ou *hard-core* en anglais). Nous allons commencer par définir ce que nous entendons par « k-hard-core », puis discuter des liens entre cette propriété et la coloration, avant de présenter la version k-hard-core de quelques résultats principaux.

Nous commençons par rappeler la définition de la k-localisabilité pour un ensemble de variables aléatoires $(X_I)_{i\in I}$ indexées par un ensemble $I = \{1,\dots,n\}$ ou $I = \mathbb{Z}$ et une notion de distance sur I :

Définition 7. *Une distribution de probabilités sur les $(X_i)_{i\in I}$ est dite k-**localisable** si pour tout $J, K \subseteq I$ intervalles à distance au moins k, la distribution de (X_J, X_K) ne dépend que de $\{|J|, |K|\}$.*

Variables ou mots h-hard-core.

Définition 25. *Nous disons que $(X_i)_{i\in I}$ sont des **variables h-hard-core** s'il s'agit de variables binaires et s'il est impossible qu'une variable vale 1 s'il est dans un rayon de taille h autour d'une variable valant 1.*

Définition 26. *Soit*

$$\mathcal{I}_n^h := \{x = x_1x_2\dots x_n \mid x \in \{0,1\}^n, \\ \forall 1 < i < n\, x_i = 1 \Rightarrow x_j = 0, j = \max(1, i-h), \dots, \min(h+i, n)\}.$$

*Un élément de $\mathcal{I}_n^h$ est appelé un **mot binaire h-hard-core de longueur** n* [4].

Il est facile de voir que des variables (ou un mot) 1-hard-core sont tout simplement des variables (resp. un mot) à particules dures.

4. Cette expression paraît complexe car elle prend en compte le fait que si $i \leq h$ alors nous devons nous arrêter à l'indice $j = 1$ à droite et si $i \geq n - h$ alors nous devons nous arrêter à l'indice $j = n$ à gauche.

Si la propriété d'être à particules dures interdit la présence de « 1 » dans le voisinage immédiat d'un « 1 », la propriété d'être h-hard-core interdit la présence de « 1 » dans le voisinage à distance h d'un « 1 ». Une autre façon de voir serait de considérer les motifs interdits : dans un mot hard-core, « 11 » est le seul motif interdit. Dans un mot 2-hard-core, « 11 » et « 101 » sont les seuls motifs interdits. Plus généralement, dans un mot h-hard-core, l'ensemble des motifs interdits est $\{10^i1, i = 0, \dots, h-1\}$.

Terminons ce paragraphe par une remarque sur la k-localisabilité et le fait d'être h-hard-core.

- Si $k < h$, il n'existe pas de distribution k-localisable sur les mots h-hard-core de longueur n [5]. En effet, $\Pr(x_i = x_{i+h} = 1) = 0$ mais il existe un $0 < j \leq n - i - h$ tel que $\Pr(x_i = x_{i+h+j} = 1) > 0$ par définition de la propriété d'être h-hard-core. Or, les ensembles $I = \{i\}$, $J = \{i + h\}$ et $J' = \{i + h + j\}$ sont à distance plus de k les uns des autres.
- Toute distribution k-localisable est nécessairement $(k + 1)$-localisable. En particulier, s'il existe une distribution k-localisable sur les mots h-hard-core, cette distribution est aussi ℓ-localisable pour tout $\ell > k$.
- Par conséquent, le seul cas de figure intéressant à étudier est celui des mots k-hard-core k-localisables.

Coloration à distance d. Nous avions vu qu'un intérêt d'étudier les variables aléatoires à particules dures ayant la propriété d'être 1-localisables ou 1-dépendantes étaient qu'elles fournissaient une borne inférieure sur le nombre de couleurs nécessaires pour une coloration 1-localisable resp. 1-dépendante. Nous allons maintenant voir qu'il existe un lien similaire entre un processus h-hard-core et la coloration à distance d, une notion que nous avons définie pour les graphes dans le Chapitre 5, Section 5.3. Puisque nous avons principalement affaire à des variables définies sur un chemin fini ou le chemin infini ici, nous allons redéfinir la coloration à distance d pour ce cas particulier afin de permettre une meilleure compréhension.

Définition 18. *Soient $G = (V, E)$ un graphe et $q, d \in \mathbb{N}$ et $(X_v)_{v \in V}$ un ensemble de variables aléatoires indexées par les sommets de G et à valeurs dans $\{1, \dots, q\}$. On parle de q-**coloration à distance** d si $X_u \neq X_v$ pour tout $u, v \in V$ à distance inférieure ou égale à d l'un de l'autre.*

Définition 27. *Soient $q, d \in \mathbb{N}$ et $(X_i)_{i \in I}$ un ensemble de variables aléatoires à valeurs dans $\{1, \dots, q\}$. Nous parlons de q-**coloration à distance** d si $X_i \neq X_j$ pour tout $i, j \in I$ à distance inférieure ou égale à d l'un de l'autre.*

5. Autre que la distribution triviale où le mot « 0^n » apparait avec probabilité 1.

Nous voyons facilement qu'un processus d-hard-core est induit par un processus de coloration à distance d de façon identique à celle qui fournit un processus hard-core à partir d'un processus de coloration « classique ». Plus précisément, étant donné une q-coloration à distance d $(X_i)_{i\in I}$, nous fixons une couleur $q^* \in \{1,\dots,q\}$ et définissons $(Y_i)_{i\in I}$ via $Y_i := 1$ si $X_i := q^*$ et $Y_i = 0$ sinon. Puisque $(X_i)_{i\in I}$ est une q-coloration à distance d, aucune variable dans un rayon d d'une variable ayant pris la valeur q^* ne peut prendre cette valeur. Par conséquent, si $Y_i = 1$, alors $Y_{i+j} = 0$ pour tout $j = 1,\dots,d$, et $(Y_i)_{i\in I}$ sont des variables aléatoires d-hard-core.

Nous remarquons aussi que cette transformation conserve la propriété d'être d-localisable (ou d-dépendant).

Un raisonnement similaire à celui qui a été fait pour les processus hard-core 1-localisables (resp. d-dépendants) permet de montrer qu'étudier les processus d-hard-core d-localisables (resp. d-dépendants) fournit une borne inférieure sur le nombre q de couleurs nécessaires pour une coloration à distance d d-localisable (resp. d-dépendante), à savoir :

$$q \geq 1/Pr(Y_n = 1) \tag{6.4}$$

Dans le Chapitre 5, Section 5.3, nous avons également définie la notion de graphe puissance d d'un graphe. En particulier, nous avons vu que le nombre chromatique à distance d d'un graphe est identique au nombre chromatique du graphe puissance d de ce graphe. Nous pouvons donc nous intéresser à la question des liens entre des variables aléatoires de coloration à distance d sur un chemin, des variables aléatoires de coloration (à distance 1) sur la puissance d d'un chemin, des variables aléatoires d-hard-core sur un chemin e des variables aléatoires hard-core sur la puissance d de ce chemin, ainsi que la propriété d'être k-localisable.

1. Une q-coloration à distance d sur un chemin est transformable en une q-coloration sur la puissance d d'un chemin (et inversement).
2. Une « assignation de valeurs » d-hard-core sur un chemin est transformable en une assignation de valeurs hard-core sur la puissance d de ce chemin (et inversement).
3. En étudiant les variables aléatoires hard-core, nous obtenons une borne inférieure sur le nombre de couleurs nécessaires pour une coloration 1-localisable (grâce à la borne (5.1)).
4. En étudiant les variables aléatoires d-hard-core, nous obtenons une borne inférieure sur le nombre de couleurs nécessaires pour une coloration à distance d d-localisable (grâce à la borne (6.4)).

5. (1) et (4) ensemble entraînent qu'en étudiant les variables aléatoires d-hard-core, nous obtenons une borne inférieure sur le nombre de couleurs nécessaires pour une coloration 1-dépendante de la puissance d d'un chemin. Néanmoins, rien de ce qui a été dit permet de garantir que cette borne inférieure est optimale.

Résultats. Nous présentons maintenant des analogues de résultats obtenus dans ce chapitre pour des variables aléatoires (ou mots) k-hard-core k-localisables.

Nous reprenons des notations, notamment la notation $\star$ pour « toutes les valeurs possibles sur une position », avec la différence que dans cette partie nous ne considérerons pas des suites de moins de k $\star$ consécutifs.

Nous commençons par donner les équivalents des règles (R0) à (R4) pour une fonction $\Lambda_n : \{0,1\}^n \to \mathbb{Z}[p_1,\dots,p_\ell]$, où les p_i, $i = 1,\dots,\ell$ sont des variables.

(RR0) $\Lambda_n(s\star^k t) = \sum_{u\in\{0,1\}^k} \Lambda_n(sut)$ pour tout s,t tels que $|s|+|t| = n-k$.

(RR1) $\Lambda_n(s) = 0$ si $s \in \{0,1\}^n \setminus \mathcal{I}_n^k$.

(RR2) $\Lambda_n(s\star^k t\star) = \Lambda_n(s\star^k\star t) = \Lambda_n(\star s\star^k t)$ pour s,t tels que $|s|+|t| = n-k-1$.

(RR3) $\Lambda_n(s\star^k\star t) = \Lambda_n(t\star^k\star s)$ pour s,t tels que $|s|+|t| = n-k-1$.

(RR4) $\Lambda_n((1\star^k)^i\star^{n-(k+1)i}) = p_i$ pour $i \in \{1,\dots,\ell\}$.

Nous limitant à $n = (k+1)\ell$ et des fonctions Λ_n vérifiant les règles (RR1), (RR2), (RR4), nous avons alors l'équivalent du Système 1 :

Système 3. *$p_i \geq 0$ et $\Lambda_{(k+1)\ell}(s) \geq 0$, pour tout $i \in \{1,\dots,\ell\}$ et $s \in \mathcal{I}^k_{(k+1)\ell}$.*

Nous avons les équivalents des Lemmes 2 et 3 :

Lemme 9. *La valeur $\Lambda_n(s)$ s'exprime comme une fonction linéaire de $p_1,\dots,p_\ell$. De plus, elle est uniquement déterminée par $p_0,\dots,p_\ell$ pour tout $s \in \mathcal{I}_{k,n}$.*

La preuve de ce lemme repose également sur une induction sur le nombre de « 0 » consécutifs dans s, en prenant en compte la subtilité qu'il est impossible d'avoir strictement moins de k « 0 » consécutifs puisque s est k-hard-core par hypothèse.

Lemme 10. *L'unique fonction Λ_n définie dans le Lemme 9 satisfait la Propriété (RR3).*

La preuve de ce lemme est identique modulo la subtilité soulignée ci-dessus (nous considérons des suites de k « 0 » consécutifs au lieu d'un seul « 0 »).

Grâce à ces deux lemmes, nous montrons l'équivalent du Théorème 5 :

Théorème 9. *Soient $p_1, \dots, p_\ell \in [0,1]$. Il existe une distribution de probabilité k-localisable $\Pr$ sur $\mathcal{I}_{k,(k+1)\ell}$ telle que $\Pr((1\star^k)^i\star^{(k+1)\ell-(k+1)i}) = p_i$ pour tout $i \in \{1,\dots,\ell\}$ si et seulement si le Système 3 est satisfait avec $\Lambda_{(k+1)\ell}(s) = \Pr(s)$ et $p_0 = 1$.*

Ensuite, nous démontrons qu'il suffit de regarder un sous-système du Système 3 en nous limitant au sous-ensemble suivant de $\mathcal{I}_n^k$:

$$\mathcal{S}_n^k = \left\{(10^k)^j 0^{n-(k+1)j} : j \in \{0,\dots,\ell\}\right\}$$

Système 4. *$p_i \geq 0$ et $\Lambda_{(k+1)\ell}(s) \geq 0$, pour tout $i \in \{1,\dots,\ell\}$ et $s \in \mathcal{S}_{(k+1)\ell}$.*

Nous avons alors un résultat similaire au Lemme 4.

Lemme 11. *Pour tout $s \in \mathcal{I}_{k,n}$, il existe $(a_t)_{t\in\mathcal{S}_n^k}$, $a_t \in \mathbb{N}$, telle que $\Lambda_n(s) = \sum_{t\in\mathcal{S}_n^k} a_t\Lambda_n(t)$.*

La preuve de ce lemme s'obtient facilement à partir de la preuve du Lemme 4 en substituant $\star^k$ à $\star$.

Cela nous permet de montrer que

Proposition 9. *Le Système 3 est équivalent au Système 4, i.e., toute solution de l'un est aussi solution de l'autre.*

Et nous démontrons un analogue du Lemme 5 pour les éléments de $\mathcal{S}_n^k$:

Lemme 12. *$\Lambda_n((10^k)^j 0^{n-2j}) = \sum_{i=0}^{\ell-j}(-1)^i\binom{(k+1)\ell-k(i-1)}{i}p_{j+i}$, pour $j \in \{0,\dots,\ell\}$.*

La preuve de ce Lemme est similaire à celui du Lemme 5 mais assez fastidieuse à rédiger puisqu'il faut considérer toutes les valeurs de $u \mod (k+1)$, là où dans le Lemme de base (i.e. pour $k = 1$) il n'y avait que deux cas à considérer (u pair ou impair).

Nous formulons ensuite le problème de trouver la valeur maximale de p_1 sous les contraintes données par le Système 4 comme un problème LP $(\mathbf{A_n}, \mathbf{c_n}, \mathbf{b_n}, \mathbf{x_n})$ avec :

$$\begin{cases} \mathbf{x}_\ell &= (p_1, \ldots, p_\ell) \\ \mathbf{c}_\ell &= (1, 0, \ldots, 0) \\ \mathbf{b}_\ell &= (1, 0, \ldots, 0) \\ a_{i,j} &= 0 \quad \text{si} \quad j - i + 1 < 0 \\ a_{i,j} &= (-1)^{i+j} \binom{(k+1)(\ell+1)-(kj+i)}{j-i+1} \quad \text{sinon} \end{cases} \tag{6.5}$$

Ce problème vérifie les conditions de forme nécessaires pour pouvoir appliquer le Théorème 7, et la seule chose qu'il reste à faire est de déterminer la nature de la suite $(u_n)_{n\in\mathbb{N}}$.

Lemme 13. *La suite $(u_n)_{n\in\mathbb{N}}$ associé au problème LP (6.5) est la suite des nombres de Fuss-Catalan $(\mathsf{FC}(k+1,n))_{n\in\mathbb{N}}$.*

La preuve de ce lemme est similaire à celui du Lemme 13. Pour la définition des nombres de Fuss-Catalan et la formule de récurrence utilisée dans la preuve, voir le Chapitre 7, Section 7.3.

Démonstration. Nous avons $u_1 = 1 = \mathsf{FC}(k+1,1)$ et $u_2 = a_{\ell,\ell} = k+1 = \mathsf{FC}(k+1,2)$. Il reste à prouver que $u_{k+1} = \mathsf{FC}(k+1,m+1)$ si pour tout $j \leq m$ nous avons $u_j = \mathsf{FC}(k+1,j)$.

Nous utilisons la relation de récurrence vérifiée par la suite $(u_n)_{n\in\mathbb{N}}$:

$$\begin{aligned} u_{m+1} &= \sum_{i=1}^{m} a_{\ell-m+1,\ell-m+i} u_{m+1-i} \\ &= \sum_{i=1}^{m} (-1)^{i+1} \binom{(k+1)(\ell+1) - (k\ell - km + ki + \ell - m + 1)}{i} \\ &\qquad \mathsf{FC}(k+1, m+1-i) \\ &= \sum_{i=1}^{m} (-1)^{i+1} \binom{(k+1)m - ki + k}{i} \mathsf{FC}(k+1, m+1-i) \\ &= \sum_{i=0}^{m-1} (-1)^{i} \binom{(k+1)m - ki}{i+1} \mathsf{FC}(k+1, m-i) \\ &= \mathsf{FC}(k+1, m+1) \quad \text{(Théorème 13)} \end{aligned}$$

□

Nous obtenons enfin que :

Théorème 10. *Toute solution* $(p_1, \dots, p_\ell) \in \mathbb{R}^\ell$ *du Système 3 satisfait* $p_i \leq \frac{\mathsf{FC}(k+1,\ell)}{\mathsf{FC}(k+1,\ell+1)} \cdot p_{i-1} \leq \frac{\mathsf{FC}(k+1,\ell-i+1)}{\mathsf{FC}(k+1,\ell+1)} \cdot p_0$. *Il est possible d'avoir égalité.*

Pour conclure, rappelons que, en ce qui concerne les variables k-hard-core, seule l'étude de celles qui sont k-localisables est intéressante, puisque toute famille de variables k-localisables est h-localisable lorsque $k < h$ et qu'il n'existe pas de variables k-hard-core h-localisables autre qu'une solution triviale lorsque $h < k$. Nous rappelons aussi le lien entre les processus k-hard-core et la coloration à distance k, et le fait que l'étude de la valeur de p_1 est utile pour trouver le nombre minimum de couleurs q nécessaires pour avoir une coloration à distance k k-localisable. En utilisant donc la relation $q \geq \frac{1}{p_1}$ et le Théorème 10, nous pouvons déduire qu'il faut au moins $\lceil \frac{\mathsf{FC}(k+1,\ell+1)}{\mathsf{FC}(k+1,\ell)} \rceil$ couleurs.

6.4.2 Processus k-localisables sur les graphes

Nous allons ici étendre la notion de k-localisabilité aux graphes quelconques à un nombre dénombrables de sommets et des degrés finis. Pour cela, nous allons considérer un tel graphe $G = (V, E)$ et des variables aléatoires $(X_v)_{v \in V}$. Nous allons noter $n := |V|$.

Définition 28. *Une distribution de probabilités sur les* $(X_v)_{v \in V}$ *est dite* k*-localisable si pour tout* $I, J \subseteq V$ *à distance au moins* k *et tels que* $G_{|I|}$ *et* $G_{|J|}$ *sont connexes, la distribution de* (X_I, X_J) *ne dépend que de la classe d'isomorphisme de* $\left\{G_{|I|}, G_{|J|}\right\}$.

Nous allons nous intéresser aux variables aléatoires hard-core 1-localisables, ou, de manière équivalente, aux distributions 1-localisables sur les motifs hard-core sur les graphes, et voir dans quelle mesure des résultats prouvés au début de ce chapitre se généralisent. Nous commençons donc par donner quelques définitions :

Nous reprenons la notation $N_t(v)$ introduite dans le Chapitre 2, Section 2.1 pour la voisinage de taille t d'un sommet $v \in V$.

Définition 29. *Nous disons que* $(X_v)_{v \in V}$ *sont des* ***variables à particules dures*** *s'il s'agit de variables binaires et si* $X_u = 1$ *pour un* $u \in V$ *implique* $X_w = 0$ *pour tout* $w \in N_1(u)$.

Définition 30. *Soit* $N = 2^n$ *et* $L_1, \dots, L_N : V \to \{0,1\}$ *deux-à-deux distincts*[6].

$$\mathcal{I}_G := \{L_i(V) \mid i \in \{1, \dots, N\}, L_i(u) = 1 \Rightarrow L_i(w) = 0 \text{ pour tout } w \in N_1(u)\}.$$

6. Ce sont donc tous les motifs binaires possibles sur G.

Un élément de $\mathcal{I}_G$ est appelé un ***motif ou label binaire à particules dures sur*** G.

Nous allons nous intéresser aux questions suivantes :

1. **Généralisation du Théorème 5**. Existe-t-il des variables $p_1, \dots, p_\ell$, pour un $\ell \in \mathbb{N}$, et un système d'équations en ces variables, telles que l'existence d'une distribution 1-localisable sur les motifs binaires à particules dures est équivalente à la résolubilité de ce système ?
2. **Généralisation du Lemme 5** Comment s'expriment certaines (ou l'ensemble des) équations de ce système ?
3. **Généralisation de la Proposition 3**. S'il existe un tel système, existe-t-il un sous-système plus simple de ce système qui donne les mêmes contraintes ?

Généralisation du Théorème 5. Nous allons commencer par formuler le problème de la 1-localisabilité comme un ensemble de règles algébriques qui doivent être vérifiées par les motifs, comme nous l'avions fait pour les mots sur le chemin.

Nous considérons pour cela des familles de motifs $L : V \to \{0,1\}$ resp. $L^\star : V \to \{0,1,\star\}$ tels que deux sommets voisins ne peuvent pas avoir tous les deux comme valeur 1. La première famille de motifs est exactement $\mathcal{I}_G$, nous utiliserons la notation $\mathcal{I}_G^\star$ pour la deuxième.

Nous considérons des variables $p_1, \dots, p_M$, où M est le nombre de stabilité de G (ce choix sera justifié plus tard), et une fonction $\Lambda_V : \mathcal{I}_G \to \mathbb{Z}[p_1, \dots, p_M]$. Nous définissons $p_0 := \sum_{L \in \mathcal{I}_G} \Lambda_V(L)$.

Les règles suivantes permettent d'étendre le domaine de Λ_V à des motifs sur $\{0,1,\star\}^n$, i.e. à $\mathcal{I}_G^\star$:

(RG0) $\Lambda_V(L_1 \cup \{v \mapsto \star\} \cup L_2) = \Lambda_V(L_1 \cup \{v \mapsto 0\} \cup L_2) + \Lambda_V(L_1 \cup \{v \mapsto 1\} \cup L_2)$ où il existe $V_1, V_2 \subset V$ disjoints tel que L_i est la restriction d'un motif L à V_i, pour $i = 1, 2$, et $V_1 \cup V_2 \cup \{v\} = V$.

En itérant la règle (RG0) jusqu'à ce qu'il ne reste plus que le symbole $\star$ comme label à gauche, on obtient :

$$\Lambda_V(\star^V) = \sum_{L \in \mathcal{I}_G} \Lambda_V(L) = p_0$$

Nous définissons de plus les propriétés suivantes :

(RG1) $\Lambda_V(L) = 0$ si $L \in \{0,1\}^V \setminus \mathcal{I}_G$.

(RG2) Si nous prenons deux décompositions suivantes de l'ensemble des sommets de G,

$V_1 \cup V_2 \cup \{v_1\} \cup \{v_2\}$

et

$V_1' \cup V_2' \cup \{v_1'\} \cup \{v_2'\}$,

telles que $G_{|V_i}$ et $G_{|V_i'}$ sont isomorphes pour $i = 1, 2$,

et si nous notons $L_i := L(V_i)$ et $L_i' := L(V_i')$, $i = 1, 2$,

et nous supposons $L_i = L_i'$ (à isomorphisme près) pour $i = 1, 2$,

alors

$\Lambda_V(L_1 \cup L_2 \cup \{v_1 \mapsto \star\} \cup \{v_2 \mapsto \star\}) = \Lambda_V(L_1' \cup L_2' \cup \{v_1' \mapsto \star\} \cup \{v_2' \mapsto \star\})$

(RG3) $\Lambda_V(L_i) = p_i$ pour $1 \leq i \leq M$,

où $L_i \in \mathcal{I}_G^\star$ est comme suit : il existe un stable S de G de taille i tel que $L_i(v) = 1$ pour tout $v \in S$ et $L_i(v) = \star$ pour tout $v \in V \setminus S$.

Une conséquence importante de la règle (RG3) est la suivante : des motifs $L_i \neq L_i'$ qui valent « 1 » sur deux stables différents de taille i vérifient tous les deux $\Lambda_V(L_i) = \Lambda_V(L_i') = p_i$, i.e.

> Quel que soit le stable de taille i et L_i valant « 1 » sur ce stable, $\Lambda_V(L_i) = p_i$.

Par conséquent, il y a autant de variables que de tailles possibles pour les stables, ce qui justifie le choix plus haut d'en prendre autant que M, le nombre de stabilité du graphe.

Dorénavant, nous ne considérerons que des fonctions Λ_V vérifiant les Propriétés (RG0), (RG1) et (RG3).

Lemme 14. *Toute fonction Λ_V vérifiant les Propriétés (RG0) et (RG3) satisfait également la Propriété (RG2).*

Démonstration. Le but est de prouver que :

$$\Lambda_V(L_1 \cup L_2 \cup \{v_1 \mapsto \star\} \cup \{v_2 \mapsto \star\}) = \Lambda_V(L_1' \cup L_2' \cup \{v_1' \mapsto \star\} \cup \{v_2' \mapsto \star\})$$

Pour L_i, L_i', v_i, v_i', $i = 1, 2$ vérifiant les conditions données dans la règle (RG2).

Pour cela, nous procédons par récurrence sur le nombre de « 0 » dans $L_1 \cup L_2$ (ou, de manière équivalente, dans $L'_1 \cup L'_2$, puisque $L_i = L'_i$ pour $i = 1, 2$ par hypothèse).

S'il n'y a pas de « 0 » dans $L_1 \cup L_2$, i.e. il n'y a que des « 1 » et des « $\star$ » dedans, alors nous utilisons la règle (RG3) :

$$\Lambda_V(L_1 \cup L_2 \cup \{v_1 \mapsto \star\} \cup \{v_2 \mapsto \star\}) = p_i \text{ pour un } i \in \{0, M\}$$

S'il y a un « 0 » dans $L_1 \cup L_2$ (sans perte de généralité, supposons qu'il est sur le sommet u dans L_1, et notons L_1^- le motif résultant en enlevant u) :

$$\begin{aligned}
&\Lambda_V(L_1 \cup L_2 \cup \{v_1 \mapsto \star\} \cup \{v_2 \mapsto \star\}) \\
=&\Lambda_V(L_1^- \cup L_2 \cup \{u \mapsto 0\} \cup \{v_1 \mapsto \star\} \cup \{v_2 \mapsto \star\}) \\
=&\Lambda_V(L_1^- \cup L_2 \cup \{u \mapsto \star\} \cup \{v_1 \mapsto \star\} \cup \{v_2 \mapsto \star\}) \\
-&\Lambda_V(L_1^- \cup L_2 \cup \{u \mapsto 1\} \cup \{v_1 \mapsto \star\} \cup \{v_2 \mapsto \star\}) \text{ règle (RG1)} \\
=&\Lambda_V(L_1'^- \setminus \{u\} \cup L'_2 \cup \{u \mapsto \star\} \cup \{v'_1 \mapsto \star\} \cup \{v'_2 \mapsto \star\}) \\
-&\Lambda_V(L_1'^- \setminus \{u\} \cup L'_2 \cup \{u \mapsto 1\} \cup \{v'_1 \mapsto \star\} \cup \{v'_2 \mapsto \star\}) \text{ (HR)} \\
=&\Lambda_V(L'_1 \setminus \{u\} \cup L'_2 \cup \{v'_1 \mapsto \star\} \cup \{v'_2 \mapsto \star\}) \text{ règle (RG1)}
\end{aligned}$$

Nous pouvons nous servir de l'hypothèse de récurrence dans la dernière étape puisqu'il y a un « 0 » de moins dans $L_1 \setminus \{u\} \cup L_2 \cup \{u \mapsto \star\}$ et dans $L_1 \setminus \{u\} \cup L_2 \cup \{u \mapsto 1\}$ que dans $L_1 \cup L_2$.

□

Lemme 15. *Pour tout $L \in \mathcal{I}_G$, la valeur de $\Lambda_V(L)$ est fonction linéaire de, et dépend uniquement de, $p_1, \ldots, p_M$.*

Démonstration. Nous prouvons ce résultat par récurrence sur le nombre de 0 dans le motif L. S'il n'y a pas de 0 dans L, alors $\Lambda_V(L) = p_i$ pour un $i \in \{1, \ldots, M\}$ (règle (RG3)). S'il y a un 0 dans L, nous pouvons écrire $L = L_1 \cup \{v \mapsto 0\} \cup L_2$ pour un sommet v tel que $L(v) = 0$, et où L_1 et L_2 sont des labels potentiellement vides qui peuvent contenir ou non des 0. Alors, en utilisant la règle (RG0), nous avons :

$$\begin{aligned}
\Lambda_V(L) &= \Lambda_V(L_1 \cup \{v \mapsto 0\} \cup L_2) \\
&= \Lambda_V(L_1 \cup \{v \mapsto \star\} \cup L_2) - \Lambda_V(L_1 \cup \{v \mapsto 1\} \cup L_2)
\end{aligned}$$

Les labels $L_1 \cup \{v \mapsto \star\} \cup L_2$ et $L_1 \cup \{v \mapsto 1\} \cup L_2$ comportent au moins un 0 de moins que le mot $L = L_1 \cup \{v \mapsto 0\} \cup L_2$. Par conséquent, l'hypothèse de récurrence garantit que la valeur de $\Lambda_V(L_1 \cup \{v \mapsto \star\} \cup L_2)$ et de $\Lambda_V(L_1 \cup \{v \mapsto 1\} \cup L_2)$ est uniquement déterminée. □

Nous définissons maintenant un système d'équations :

Système 5. $p_i \geq 0$ *et* $\Lambda_V(L) \geq 0$, *pour tout* $i \in \{1, \ldots, M\}$ *et* $L \in \mathcal{I}_G$.

Il est alors facile de voir que nous avons le résultat suivant :

Théorème 11. *Soient* $p_1, \ldots, p_M \in [0,1]$. *Il existe une distribution de probabilité* 1*-localisable* $\Pr$ *sur* $\mathcal{I}_G$ *telle que* $\Pr(L_i) = p_i$, *où*

$$L_i := \{v \mapsto 1, v \in S \text{ stable de taille } i\}$$

pour tout $i \in \{1, \ldots, M\}$, *si et seulement si le Système 5 est satisfait avec* $\Lambda_V(L) = \Pr(L)$ *et* $p_0 = 1$.

Généralisation du Lemme 5. Rappelons l'énoncé du Lemme que nous cherchons à généraliser (rappel : ici, $n = 2\ell$) :

Lemme 5. $\Lambda_n((10)^k 0^{n-2k}) = \sum_{i=0}^{\ell-k} (-1)^i \binom{2\ell-2k+1-i}{i} p_{k+i}$, *pour* $k \in \{0, \ldots, \ell\}$.

Nous faisons maintenant une remarque au sujet du coefficient apparaissant dans la somme : nous pouvons montrer, en utilisant par exemple la méthode des étoiles et des barres [Wikb], que le nombre $s(P_{n-2k}, i)$ d'ensembles indépendants de taille i (voir la Section 5.3) sur un chemin de longueur $n - 2k = 2\ell - 2k$ est exactement $\binom{2\ell-2k+1-i}{i}$. Autrement dit, nous pouvons reformuler le Lemme 5 de la manière suivante :

Le Lemme 5 en termes d'ensembles indépendants

$\Lambda_n((10)^k 0^{n-2k}) = \sum_{i=0}^{\ell-k} (-1)^i s(P_{n-2k}, i) p_{k+i}$, pour $k \in \{0, \ldots, \ell\}$.

Autrement dit, il semblerait que le coefficient devant p_{k+i} soit le nombre d'ensembles indépendant de taille i sur le sous-chemin de taille $n-2k$ où il n'y a que des « 0 ».

En particulier, pour $k = 0$ nous avons

$$\Lambda_n(0^n) = \sum_{i=0}^{\ell} (-1)^i s(P_n, i) p_i = I_{P_n}^M(-p_1, \ldots, (-1)^\ell p_\ell)$$

(Voir la Définition 17 du polynôme d'indépendance à plusieurs variables)

Nous pouvons d'abord remarquer que cela rapproche notre résultat du résultat suivant de Holroyd et Liggett sur l'existence d'une distribution 1-dépendant sur les mots hard-core :

Proposition 2. *Soit $G = (V, E)$ un graphe et $p \in [0, 1]$. Alors, $p \leq p_h(G)$ ssi pour tout sous-ensemble fini de sommets $S \subset G$, et $H := G_{|S}$ le graphe induit par ce sous-ensemble, nous avons*

$$I_H(-p) \geq 0$$

En effet, lorsque nous regardons la preuve de ce résultat de près, nous voyons que les probabilités des motifs, qui doivent être positives, s'expriment comme fonction des divers polynômes de stabilité, et ce fait qui nous permet d'affirmer une proximité entre ce résultat et le notre.

Nous allons maintenant prouver une généralisation de cette formule aux graphes, non pas pour un sous-ensemble restreint de motifs mais pour l'ensemble des motifs possibles. Il s'agit donc d'un résultat bien plus général.

Pour cela, nous considérons un graphe $G = (V, E)$ et un motif $L : V \to \{0, 1, \star\}$ ayant les caractéristiques suivantes :

- $L(V)$ comporte u « 1 ».
- Il n'y a aucun « 0 » adjacent à un « 1 » dans $L(V)$: en effet, s'il y en a, nous pouvons toujours les remplacer par des « $\star$ » sans modifier $\Lambda_n(L(V))$.

Nous allons distinguer deux parties du motif :

- une partie $U(L) \subset V$ dans laquelle il y a l'ensemble des « 1 » et l'ensemble des étoiles.
- une partie $Z(L) \subset V$ dans laquelle il n'y a que des suites de « 0 ».

Nous notons $G_U(L)$ et $G_Z(L)$ les sous-graphes correspondants.

Nous allons prouver le résultat suivant :

Proposition 10. *Soit M le nombre de stabilité de $G_Z(L)$ et $n := |V(G)|$. Nous avons :*

$$\Lambda_V(L(V)) \;=\; \sum_{i=0}^{M}(-1)^i s(G_Z(L), i)p_{u+i} \;=\; \sum_{i=0}^{n}(-1)^i s(G_Z(L), i)p_{u+i} \qquad (6.6)$$

Nous avons besoin du résultat bien connu suivant pour prouver cette proposition (voir par exemple [SSW10]) :

Lemme 16. *Soit $G = (V, E)$ un graphe et $v \in V$. Notons $N(v)$ le voisinage immédiat de v et $N[v] := N(v) \cup \{v\}$. Alors nous avons*

$$s(G, i) = S(G \setminus \{v\}, i) + S(G \setminus N[v], i - 1)$$

Preuve de la Proposition 10. Tout d'abord, nous notons que pour $i > M$, $s(G_Z(L), i) = 0$, ce qui justifie le fait que nous pouvons remplacer M par n dans la somme.

Nous commençons par montrer que c'est vrai lorsqu'il n'y a aucun « 0 » dans le motif : $\Lambda_V(L(V)) = \Lambda_V(U(L)) = p_u$ par définition, et une conséquence de la règle (RG0) qui dit que $\Lambda_V(\star^V) = p_0$ garantit que c'est vrai même lorsque $u = 0$.

Maintenant, nous montrons que cela est vrai aussi lorsqu'un seul sommet $v \in V$ a pour label 0, i.e. $G_Z(L) = \{v\}$.

$$\begin{aligned}\Lambda_V(L(V)) &= \Lambda_V(L(U(L)) \cup \{v \mapsto 0\}) \\ &= \Lambda_V(L(U(L)) \cup \{v \mapsto \star\}) - \Lambda_V(L(U(L)) \cup \{v \mapsto 1\}) \\ &= p_u - p_{u+1}\end{aligned}$$

Nous avons le droit de faire cela puisque par hypothèse sur le motif L, aucun « 0 » n'est adjacent à un « 1 ».

Nous démontrons la suite par récurrence sur le nombre de « 0 » dans le label, c'est-à-dire en faisant l'hypothèse que l'égalité (6.6) est vraie pour tout motif L' tel qu'il y a moins de « 0 » dans $L'(V)$ que dans $L(V)$.

Alors, nous considérons un sommet $v \in V$ appartenant à $Z(G)$, donc ayant pour label 0, et nous utilisons le même raisonnement :

$$\begin{aligned}\Lambda_V(L(V)) &= \Lambda_V(L(V \setminus \{v\}) \cup \{v \mapsto 0\}) \\ &= \Lambda_V(L(V \setminus \{v\}) \cup \{v \mapsto \star\}) - \Lambda_V(L(V \setminus \{v\}) \cup \{v \mapsto 1\}) \\ &= \Lambda_V(L(V \setminus \{v\}) \cup \{v \mapsto \star\}) - \Lambda_V(L(V \setminus N[v]) \cup \{v \mapsto 1\} \cup \{N(v) \mapsto \star\}) \\ &= \Lambda_V(L_\star) - \Lambda_V(L_1)\end{aligned}$$

Où nous notons

$N(v)$ le voisinage immédiat de v,

$N[v] := N(v) \cup \{v\}$,

$L_\star := L(V \setminus \{v\}) \cup \{v \mapsto \star\})$,

et $L_1 := L(V \setminus \{v\}) \cup \{v \mapsto 1\})$.

Ces deux labels contiennent moins de « 0 » que L, donc nous pouvons appliquer l'hypothèse de récurrence.

$$
\begin{aligned}
\Lambda_V(L(G)) &= \Lambda_V(L_\star) - \Lambda_V(L_1) \\
&= \sum_{i=0}^{n}(-1)^i s(G_Z(L_\star), i)p_{u+i} - \sum_{i=0}^{n}(-1)^i s(G_Z(L_1), i)p_{u+1+i} \quad (1) \\
&= \sum_{i=0}^{n}(-1)^i s(G_Z(L_\star), i)p_{u+i} - \sum_{i=0}^{n-1}(-1)^i s(G_Z(L_1), i)p_{u+1+i} \quad (2) \\
&= p_u + \sum_{i=1}^{n}(-1)^i s(G_Z(L_\star), i)p_{u+i} - \sum_{i=1}^{n}(-1)^{i+1} s(G_Z(L_1), i-1)p_{u+i} \\
&= p_u + \sum_{i=1}^{n}(-1)^i [s(G_Z(L_\star), i) + s(G_Z(L_1), i-1)]p_{u+i} \\
&= p_u + \sum_{i=1}^{n}(-1)^i s(G_Z(L), i)p_{u+i} \quad (3) \\
&= \sum_{i=0}^{n}(-1)^i s(G_Z(L), i)p_{u+i}
\end{aligned}
$$

Nous justifions maintenant certaines étapes du calcul précédent :

(1) Les variables dans la deuxième somme sont p_{u+1+i} puisqu'il y a un « 1 » de plus dans L_1 que dans $L(G)$.

(2) Dans la deuxième somme, nous remplaçons n par $(n-1)$ puisqu'il y a au moins une arête de G qui n'est pas dans $G_Z(L_1)$, donc son nombre de stabilité est au plus $(n-1)$.

(3) Remarquons d'abord que :

— $L_\star$ est le motif qui diffère de L dans la mesure où il attribue au sommet v la valeur « $\star$ » au lieu de « 0 ». $G_Z(L_\star)$ est le graphe induit par les sommets dont le label est « 0 » dans le motif $L_\star$, donc v n'est pas dans $G_Z(L_\star)$ mais tous les autres sommets et arêtes de G à qui L attribue la valeur « 0 » y sont, par conséquent $G_Z(L_\star) = G_Z(L) \setminus \{v\}$.

— De même, L_1 est le motif qui diffère de L dans la mesure où il attribue au sommet v la valeur « 1 » au lieu de « 0 », et à tous les voisins de v la valeur « $\star$ ». $G_Z(L_1)$ est le graphe induit par les sommets dont le label est « 0 » dans le motif L_1, donc v et ses voisins ne sont pas dans $G_Z(L_1)$, mais les autres sommets et arêtes de G à qui L attribue la valeur « 0 » y sont, par conséquent $G_Z(L_1) = G_Z(L) \setminus N[v]$.

Ces deux remarques ainsi que le Lemme 16 appliqué au graphe $G_Z(L)$ permettent de conclure :

$$\begin{aligned} s(G_Z(L_\star), i) + s(G_Z(L_1), i-1) &= s(G_Z(L) \setminus \{v\}, i) + s(G_Z(L) \setminus N[v], i-1) \\ &= s(G_Z(L), i) \end{aligned}$$

□

Généralisation de la Proposition 3. Nous allons maintenant nous intéresser à la question de savoir s'il existe un sous-ensemble de motifs dont les valeurs Λ_V permettent d'obtenir un sous-système du Système 5 équivalent à ce système tout en étant plus simple, voire ayant exactement les mêmes caractéristiques que le sous-système obtenu pour le chemin (par exemple, autant de variables que de contraintes, voire la même forme plus ou moins « triangulaire »), ce qui nous permettrait d'appliquer le reste de nos résultats. Nous allons voir, sans donner de preuve technique, qu'il existe bien un sous-système plus simple, mais qu'il n'est pas sûr que ce sous-système ait les caractéristiques souhaitées pour pouvoir appliquer le reste de nos résultats, i.e. être échelonné de la même manière et avoir autant de contraintes que de variables.

Pour cela, prenons note de quelques caractéristiques du sous-système obtenu pour le chemin :

- Il y a exactement une contrainte pour un nombre de « 1 » donné dans le motif.
- Dans tous les motifs il y a exactement une suite de « 0 » contigus.
- Il y a le minimum d'étoiles « $\star$ » possibles et elles sont toutes contiguës aux « 1 ».

En effet, dans la preuve du Lemme 4 qui est utilisé pour prouver ce théorème, nous considérons un motif composé d'un certain nombre de « $(1\star)$ » et de plusieurs suites de « 0 » séparées par des suites de « $\star$ », et nous montrons qu'il est possible de diminuer petit à petit le nombre d'étoiles jusqu'à n'avoir que des suites de « $(1\star)$ » et une seule suite de « 0 ». Cela nous permet d'affirmer que la contrainte donnée par un motif du deuxième type est toujours plus forte qu'une contrainte d'un motif du premier type.

Si nous essayons d'appliquer le même type de raisonnement aux motifs sur les graphes, cela fonctionne à la différence près qu'il est possible d'avoir plusieurs composantes formées uniquement de « 0 » sans pouvoir fusionner

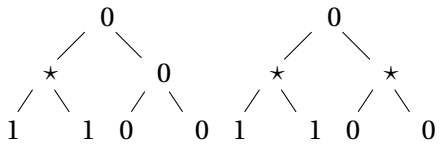


FIGURE 6.3 – Le motif de gauche donne une contrainte plus forte que le motif de droite.

ces composantes. Il est ainsi possible d'avoir des « sous-graphes des zéros » qui ne sont pas connexes.

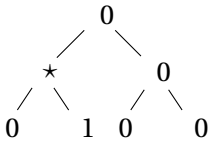


FIGURE 6.4 – Un « sous-graphe des zéros » non-connexe.

Introduisons le sous-ensemble suivant de l'ensemble $\mathcal{I}_G$ des motifs à particules dures sur G :

$$S_G := \{(V) \mid L(u) = \star \implies \exists v \in N(u), L(v) = 1 \text{ et } \forall v \in N(v), L(v) \neq \star\}.$$

Autrement dit, il s'agit des motifs où chaque « $\star$ » est contigu à au moins un « 1 » et il y a le minimum de « $\star$ » possibles.

Étant donné $L(V) \in S_G$, nous notons $V_0(L)$ l'ensemble des sommets prenant la valeur 0 par $L(V)$.

Nous considérons que $L(V), L'(V) \in \mathcal{S}_G$ sont équivalents si les graphes induits par $V_0(L)$ resp. $V_0(L')$ sont isomorphes.

Nous définissons $\mathcal{S}_G$ comme étant l'ensemble quotient de S_G par cette relation d'équivalence.

Et nous définissons le sous-système suivant du Système 5 :

Système 6. $p_i \geq 0$ *et* $\Lambda_V(L) \geq 0$, *pour tout* $i \in \{1, \ldots, M\}$ *et* $L \in \mathcal{S}_G$.

Nous avons alors (grâce au raisonnement informel ci-dessus) :

Proposition 11. *Le Système 5 est équivalent au Système 6, i.e., toute solution de l'un est aussi solution de l'autre.*

L'intérêt d'avoir dérivé ce théorème dans le cas du chemin résidait dans l'application possible du Théorème 7 à ce sous-système.

Pour pouvoir appliquer ce Théorème dans le cas général, il faudrait que le Système 6 comporte autant de contraintes que de variables. Rappelons que le nombre de variables est le nombre de stabilité M du graphe G. Donc, pour appliquer le Théorème dans le cas général, il faudrait avoir

$$|\mathcal{S}_G| = M$$

Il est également possible qu'un argument supplémentaire nous permette de considérer un sous-ensemble de $\mathcal{S}_G$ et le système correspondant : dans ce cas, il faut que ce sous-ensemble soit de taille M.

6.5 Résultats et considérations complémentaires

6.5.1 Processus 1-localisables vs. 1-dépendants : unicité

Dans cette partie, nous souhaiterions discuter ce qu'il est possible de déduire sur le rapport entre les processus hard-core resp. de coloration 1-localisables et 1-dépendants (et stationnaires). Rappelons que nous avions défini des variables p_i comme la probabilité d'apparition de i « 1 » dans un mot, i.e. $p_i := \Pr(X_1 = 1, X_3 = 1, \ldots, X_{2i-1} = 1)$. Nous rappelons également l'énoncé de notre résultat principal sur la valeur des p_i :

Théorème 3. *Soit* $n \in \mathbb{N}$ *et* $\ell := \lceil \frac{n}{2} \rceil$.

i. *Toute distribution de probabilité* 1*-localisable sur* $\mathcal{I}_n$ *vérifie* $p_i \leq \mathsf{c}_{\ell-i+1}/\mathsf{c}_{\ell+1}$ $(1 \leq i \leq \ell)$.

ii. *Il existe une distribution de probabilités* 1*-localisable sur* $\mathcal{I}_n$ *telle que* $p_i = \mathsf{c}_{\ell-i+1}/\mathsf{c}_{\ell+1}$ $(1 \leq i \leq \ell)$.

En particulier, puisque $\frac{\mathsf{c}_m}{\mathsf{c}_{m+1}} = \frac{m+2}{4m+2}$, ce résultat a pour conséquence que $\sup p_i \underset{n\to\infty}{\longrightarrow} 1/4^i$.

En particulier, lorsque n tend vers l'infini $\sup p_i$ tends vers $(\sup p_1)^i$. Autrement dit, si l'on prend $p_1 = 1/4$, à l'infini nous aurons :

$$\Pr(X_1 = 1, \ldots, X_{2i-1} = 1) = p_i \approx p_1^i = \Pr(X_1 = 1)^i = \prod_{i=1,3,\ldots,2i-1} \Pr(X_i = 1)$$

Cela signifie qu'il existe une **unique** distribution sur des variables hard-core $(X_n)_{n\in\mathbb{Z}}$ vérifiant $\Pr(X_i = 1) = 1/4$ qui soit 1-localisable, et ce processus se confond avec le processus 1-dépendant.

Mais si l'on prend $p_1 = \Pr(X_i = 1) < 1/4$, il est aisé de voir que p_i peut peut prendre une infinité de valeurs réelles tout en vérifiant $p_i \leq 1/4^i$. En particulier, il est possible d'avoir $p_i \neq p_1^i$. Ainsi, pour $p_1 < 1/4$, il existe une infinité non-dénombrable de distributions 1-localisables qui ne sont pas toutes 1-dépendantes.

Cependant, nous ne pouvons rien déduire de tout cela au sujet des processus de colorations 1-localisables. Rappelons qu'il est inconnu si le processus de 4-coloration 1-dépendant découvert par Holroyd et Liggett [HL16] est unique. Ainsi, même si nos résultats entraînent qu'il faut 4 couleurs pour que la coloration 1-localisable soit possible, il n'est pas clair s'il existe un unique processus de 4-coloration 1-localisable.

6.5.2 MIS 2-localisable

Outre les résultats exhaustifs obtenus ci-dessus, nous sommes également parvenus à démontrer l'impossibilité d'une distribution 2-localisable pour le problème de l'ensemble indépendant maximal que nous allons présenter dans cette section.

Nous commençons par rappeler les définitions d'un ensemble indépendant et celui d'un ensemble indépendant maximal (MIS) :

Définition 13. *Soit $G = (V, E)$ un graphe. Un **ensemble stable** ou **ensemble indépendant** de G est un sous-ensemble $U \subset V$ de sommets mutuellement non-adjacents, i.e. tels que $\forall u, v \in U$, $\{u, v\} \notin E$.*

Définition 31. *Soit $G = (V, E)$ un graphe. Un **ensemble indépendant maximal** I de G est un ensemble indépendant tel qu'il n'existe aucun ensemble indépendant J de G tel que $I \subsetneq J$.*

Nous considérons n variables aléatoires binaires $X_1, \ldots, X_n$ telles que les valeurs prises par $(X_1, \ldots, X_n)$ représentent un MIS sur le chemin (ou le cycle) de longueur n, i.e. si X_i vaut 1 X_{i-1} et X_{i+1} valent 0 (c'est donc un ensemble indépendant) et il n'y a jamais plus de deux « 0 » consécutifs (modulo n s'il s'agit du cycle) : sinon, nous pourrions changer l'un des « 0 » en « 1 » afin d'obtenir un ensemble indépendant qui inclut le premier qui ne serait ainsi pas maximal.

Rappelons que pour qu'une distribution de probabilités $\Pr$ sur les MIS d'un chemin ou d'un cycle de longueur n soit 2-localisable, $\Pr(X_I = x_I, Y_J =$

x_J) ne doit dépendre que de $\{|I|, |J|\}$ pour tous sous-ensembles I et J de $\{1, \ldots, n\}$ à distance au moins 2 l'un de l'autre.

Soit $1 \leq i \leq n-4$. Prenons $I = \{i, i+1\}$ et $J = \{i+3, i+4\}$. Regardons $p := \Pr(X_I = (0,0), X_J = (0,0)) = \Pr(X_i = 0, X_{i+1} = 0, X_{i+3} = 0, X_{i+4} = 0)$. Puisque I et J sont à distance exactement 2 l'un de l'autre, cela signifie qu'il y a deux sommets entre eux. Nous avons trois cas de figure :

- Soit aucun de ces sommets n'est pas dans le MIS, mais alors nous avons six sommets consécutifs qui ne sont pas dans le MIS : contradiction.
- Soit un seul de ces sommets n'est pas dans le MIS, mais alors nous avons toujours trois sommets consécutifs qui ne sont pas dans le MIS : contradiction.
- Il est impossible que les deux sommets soient à la fois dans le MIS puisqu'il sont voisins.

Puisque aucune de ces situations n'est possible, nous en concluons que $p = 0$. Or, puisque la distribution est supposée être 2-localisable, cela signifie que $\Pr(X_I = (0,0), X_J = (0,0))$ également pour I et J à distance plus de 2, i.e. pour I {i,i+1} et $J = \{i+j, i+j+1\}$ avec $1 \leq i \leq n-j-1$ et $4 \leq j \leq n-i-1$.

Nous en déduisons qu'une distribution 2-localisable sur les MIS d'un chemin ou d'un cycle de longueur n attribue une probabilité non-nulle seulement aux MIS tels que au plus une paire de sommets consécutives n'y appartient pas, i.e. s'il y a plus d'un $i \in \{1, \ldots, n-1\}$ tel que $x_i = x_{i+1} = 0$, alors $\Pr(X_1 = x_1, \ldots, X_n = x_n) = 0$.

Nous considérons maintenant $i \leq j \leq k \in \{1, \ldots, n\}$ tels que $i \leq j-3$ et $j \leq k-3$. Notons $d_1 = j-i$, $d_2 = k-j$, et $d_3 = k-i$ les distances entre les trois variables. Nous allons montrer que la valeur de $q := \Pr(X_i = 1, X_j = 1, X_k = 1)$ dépend de (d_1, d_2, d_3) en utilisant le constat ci-dessus, à savoir qu'il existe au plus une paire de « 0 » consécutifs.

La parité des (d_1, d_2, d_3) peut en théorie prendre quatre valeurs différentes :

- Tous les d_i sont pairs.
- Deux d_i sont pairs et l'autre est impair.
- Deux d_i sont impairs et l'autre est pair.
- Tous les d_i sont impairs.

Ensuite, nous pouvons remarquer que d_i, $i = 1, 2, 3$ est impair si et seulement si l'unique paire de « 0 » consécutifs apparaît entre les deux variables qui y interviennent. Par conséquent :

- Si la paire de « 0 » consécutifs apparaît quelque part entre X_i et X_k, l'un de d_1 ou d_2 est impair, l'autre étant pair, et d_3 est impaire.
- Si la paire de « 0 » consécutifs n'apparaît pas quelque part entre X_i et X_k, tous les d_i sont pairs.

Pour une distribution 2-localisable, $q = \Pr(X_i = 1, X_j = 1, X_k = 1)$ ne doit pas dépendre de i, j, k, par conséquent elle ne doit pas dépendre des d_i. Or, les remarques précédentes ont pour conséquence que $q = 0$ si toutes les d_i sont impairs ou si deux d'entre eux sont pairs.

Par conséquent :

- Soit nous excluons qu'une paire de « 0 » consécutifs puisse apparaître, auquel cas $(X_1, \ldots, X_n)$ peut prendre deux valeurs possibles qui correspondent à des 2-colorations (i.e. un « 0 » est toujours suivi d'un « 1 » et inversement). Mais alors $\Pr(X_i \neq X_j)$ dépend de la parité de $|i - j|$ qui peut valoir bien plus grand que 2, par conséquent elle ne peut pas être 2-localisable.
- Soit nous autorisons qu'une paire de « 0 » consécutifs puisse apparaître, mais alors nous devons exclure la possibilité que plus de deux « 1 » apparaissent. Or, ceci est impossible dès que $n > 6$, et il n'existe pas de distribution 2-localisable sur les MIS d'un chemin ou d'un cycle de longueur > 6.

6.5.3 La matrice du Théorème 7

Dans cette section, nous souhaiterions attirer l'attention sur un résultat récent [APT19] dans le domaine de la combinatoire des mots sur les matrices ayant plus ou moins la même forme que la matrice apparaissant dans le Théorème 7. Plus précisément, il s'agit de matrices dont les coefficients sont la valeur absolue des coefficients de cette matrice :

$$\begin{bmatrix} a_{1,1} & a_{1,2} & \ldots & \ldots & a_{1,n-1} & a_{1,n} \\ 1 & a_{2,2} & \ldots & \ldots & a_{2,n-1} & a_{2,n} \\ 0 & 1 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & a_{n-1,n-1} & a_{n-1,n} \\ 0 & 0 & \ldots & 0 & 1 & a_{n,n} \end{bmatrix}$$

Dans l'article cité, la matrice en question est construite à partir d'un mot x sur un alphabet ordonné $A = \{a_1 < \cdots < a_k\}$ à k lettres de la manière

suivante : le coefficient $a_{i,j}$ est égal au nombre d'occurrences non-contiguës du sous-mot $a_i a_{i+1} \dots a_j$ dans le mot x [7].

De nombreux résultats sont obtenus sur les déterminants de telles matrices. Notamment, il est montré que le déterminant se calcule à l'aide d'une formule très proche de la formule pour $U_{k,n}$ dans la Proposition 7, et que ce déterminant donne le nombre d'occurrences non-contiguës du mot $y = a_k a_{k-1} \dots a_1$ dans le mot x. En utilisant cette formule, nous pouvons facilement montrer que le déterminant d'une matrice $n \times n$ correspondant au problème de l'existence d'une distribution 1-localisable sur les mots hard-core de longueur $2n$ est le n-ème nombre de Catalan c_n. Puisque les signes des coefficients de cette matrice sont alternés, le déterminant ne change pas en prenant la valeur absolue de chaque coefficient. Nous pouvons alors nous interroger sur les mots sur un alphabet de k lettres qui donneraient lieu à une telle matrice $k \times k$, sachant qu'une même matrice peut correspondre à plusieurs mots différents.

Par exemple, la matrice 2×2 à coefficients positifs correspondant à notre problème est le suivant :

$$\begin{bmatrix} 4 & 3 \\ 1 & 2 \end{bmatrix}$$

Un mot sur l'alphabet $A = \{a < b\}$ ayant 4 occurrences de a, 3 occurrences de ab et 2 occurrences de b est le suivant : $baaaba$, et il est aisé de vérifier que le mot ba apparaît $\mathsf{c}_3 = 5$ fois dans ce mot.

La matrice 3×3 à coefficients positifs correspondant à notre problème est le suivant :

$$\begin{bmatrix} 6 & 10 & 4 \\ 1 & 4 & 3 \\ 0 & 1 & 2 \end{bmatrix}$$

Un mot sur l'alphabet $A = \{a < b < c\}$ ayant

— 6 occurrences de a

— 10 occurrences de ab

— 4 occurrences de abc

7. C'est-à-dire, pour $k \leq \ell$, a_k apparaît avant a_ℓ dans x, mais a_k et a_{k+1} ne sont pas nécessairement contiguës. Par exemple, le nombre d'occurrences non-contiguës de ab dans $babbab$ est 4.

— 4 occurrences de b

— 3 occurrences de bc

— 2 occurrences de c

est le suivant : $cabbabcaaaab$. Encore une fois, nous pouvons facilement vérifier que le mot cba apparaît $cc_4 = 14$ fois dans ce mot.

Bien que cela nous éloigne des principales questions étudiées dans cette thèse, il serait intéressant d'étudier davantage cette question afin de déterminer l'**existence** et l'**unicité** d'un mot sur un alphabet à k lettres correspondant à notre matrice $k \times k$, et le cas échéant, étudier les caractéristiques de ces mots et notamment déterminer s'il y a un lien entre un mot sur un alphabet à k lettres et un mot sur un alphabet à $(k+1)$ lettres.

CHAPITRE 7
Nombres de Catalan

Le but de ce chapitre est de démontrer une formule utilisée dans le Chapitre 6 sur les nombres de Catalan et leurs convolutions, ainsi qu'une généralisation de cette formule aux nombres de Fuss-Catalan.

7.1 Définitions et théorème principal

La suite $(\mathsf{c}_n)_{n\in\mathbb{N}}$ des nombres de Catalan est donnée par $\mathsf{c}_n := \frac{1}{n+1}\binom{2n}{n}$. Elle admet de multiples interprétations. Dans ce qui suit, nous allons utiliser une interprétation particulière en termes de chemins sur une grille.

Nous supposons que nous avons une grille à deux dimensions composée de $n \times n$ carrés. Le but est de tracer un chemin du point $(0,0)$ en bas à gauche jusqu'au point (n,n) en haut à droite en respectant les règles suivantes : nous avons le droit de faire uniquement des pas vers la droite ou vers le haut sur les arêtes de la grille, et de sorte à ce que le chemin ne descend jamais en-dessous de la diagonale, c'est-à-dire la droite passant par les points (x,x). Les chemins vérifiant ces propriétés sont appelés les **chemins de Dyck de taille** n. Il est facile de voir qu'un tel chemin a toujours $2n$ arêtes. De plus, le nombre de chemins de Dyck de taille n est donné par le n-ème nombre de Catalan c_n. Il a été montré dans [Ted11] que $\mathsf{c}_{k,n}$ donne le nombre de chemins de Dyck de longueur $n+k-1$ qui commencent avec $k-1$ arêtes verticales, c'est-à-dire, les chemins qui commencent à $(0,0)$, remontent jusqu'à $(0,k-1)$ et continuent ensuite jusqu'à $(n+k-1,n+k-1)$ sans descendre en-dessous de la diagonale.

Nous aurons également besoin de la définition de la k-ème convolution d'une suite d'entiers :

Définition 32. *La k-ème convolution d'une suite d'entiers $(v_n)_{n\in\mathbb{N}}$ que nous dénotons par $(v_{k,n})_{n\in\mathbb{N}}$, est définie de la manière suivante :*

$$v_{k,n} = \sum_{\substack{i_1+\cdots+i_k=n \\ i_1,\ldots,i_k\geq 0}} \prod_{j=1}^{k} v_{i_j}\,.$$

Il est bien connu que les nombres de Catalan (c_n) vérifient

$$\mathsf{c}_{n+1} = \sum_{i+j=n} \mathsf{c}_i \cdot \mathsf{c}_j.$$

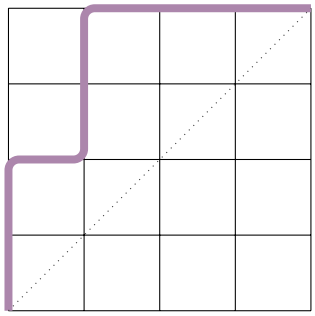


FIGURE 7.1 – Exemple d'un chemin de Dyck de taille 4.

Autrement dit :

$$\mathsf{c}_{n+1} \;=\; \mathsf{c}_{2,n} \tag{7.1}$$

Nous allons prouver le théorème suivant dans ce chapitre :

Théorème 12. *La suite* $(\mathsf{c}_{k,n})_{n\in\mathbb{N}}$ *de la* k*-ème convolution des nombres de Catalan* $(\mathsf{c}_n)_{n\in\mathbb{N}}$ *satisfait la relation de récurrence suivante par rapport à* n *:*

$$\mathsf{c}_{k,n+1} \;=\; \sum_{i=0}^{n}(-1)^i\binom{2n+k-i}{i+1}\mathsf{c}_{k,n-i}\,.$$

En particulier, grâce à l'égalité (7.1), nous pouvons déduire du Théorème 12 le résultat utile suivant :

$$\mathsf{c}_{n+1} \;=\; \sum_{i=0}^{n-1}(-1)^i\binom{2n-i}{i+1}\mathsf{c}_{n-i}\,. \tag{7.2}$$

Concernant la nouveauté de ces résultats, nous avons trouvé une mention de la formule **??** dans [Kos09], cependant aucune preuve ni aucune référence vers une preuve de cette formule n'y est donnée, et il n'y a non plus aucune mention de généralisations possibles tel notre Théorème 12 ou le Théorème 13 démontré plus tard dans ce chapitre.

7.2 Preuve du théorème

Nous allons utiliser les notations suivantes dans ce qui suit. Nous dénoterons par $\mathscr{D}_n$ l'ensemble de tous les chemins de Dyck de longueur n. Nous dénoterons par $\mathscr{C}_{k,n}$ le sous-ensemble de $\mathscr{D}_{n+k-1}$ correspondant aux chemins

qui commencent par $k-1$ arêtes verticales. En particulier, $\mathscr{C}_{1,n} = \mathscr{D}_n$. De plus, $\mathsf{c}_n = |\mathscr{D}_n|$ et $\mathsf{c}_{k,n} = |\mathscr{C}_{k,n}|$. Nous utiliserons la notation P_n pour un élément de $\mathscr{D}_n$ et la notation $P_{k,n}$ pour un élément de $\mathscr{C}_{k,n}$. Remarquons que, d'après ce que nous avons dit plus haut, tout $P_{k,n}$ est identique à un P_{n+k-1}. Les chemins « vides » P_0 et $P_{k,0}$ représentent un seul sommet sans aucune arête.

Définition 33 (Crochet)**.** *Nous appelons* crochet *l'unique chemin de Dyck de taille 1 P_1, formé d'une arête montante puis d'une arête allant vers la droite.*

Nous disons que nous *augmentons* un chemin de Dyck P_n au niveau du sommet i si nous insérons un crochet partant de ce sommet, ce qui donnera un chemin de Dyck de longueur P_{n+1}. Nous noterons cette opération $\mathrm{Aug}(P_n, i)$.

Lemme 17. *Les ensembles $\mathscr{D}_{n+1}$ et $\mathscr{C}_{k,n+1}$ vérifient les propriétés suivantes :*

1. $\mathscr{D}_{n+1} = \{\mathrm{Aug}(P_n, i) : P_n \in \mathscr{D}_n, 1 \leq i \leq 2n+1\}$
2. $\mathscr{C}_{k,n+1} = \{\mathrm{Aug}(P_{n+k-1}, i) : P_{n+k-1} \in \mathscr{C}_{k,n}, k \leq i \leq 2n+2k-1\}$

Démonstration. Il est facile de voir que le chemin résultant d'une augmentation d'un chemin de Dyck P_n de taille n au niveau de n'importe quel sommet i reste au-dessus de la diagonale et est donc un chemin de Dyck de taille $n+1$.

De plus, n'importe quel chemin de Dyck de taille $n+1$ contient au moins un crochet, donc en « contractant » ce crochet, nous obtenons un chemin de Dyck de taille n. Ceci prouve le point (1).

Pour montrer le point (2), nous considérons un chemin de Dyck P_{n+k-1} qui commence à $(0,0)$ (sommet de label 1), qui monte jusqu'à $(0, k-1)$ (sommet de label k) et puis continue jusqu'à $(n+k-1, n+k-1)$ (sommet de label $2n+2k-1$) sans descendre en-dessous de la diagonale. Le point (1) garantit qu'en augmentant P_{n+k-1} au niveau de n'importe quel sommet, nous obtiendrons un chemin de Dyck de taille $n+k$. De plus, si le sommet où nous augmentons est l'un des sommets de label $k, \ldots, 2n+2k-1$, nous obtiendrons un chemin qui monte jusqu'à $(0, k-1)$ et va ensuite jusqu'à $((n+1)+k-1, (n+1)+k-1)$ sans descendre en dessous de la diagonale, c'est-à-dire un chemin de $\mathscr{C}_{k,n+1}$.

Nous prenons maintenant un chemin de $\mathscr{C}_{k,n+1} \subset \mathscr{D}_{n+k}$. En contractant un crochet, nous obtenons un élément de $\mathscr{D}_{n+k-1}$ (en utilisant le point(1). De plus, comme le chemin est un élément de $\mathscr{C}_{k,n+1}$, il n'y a pas de crochet dont le sommet au milieu est un sommet de label inférieur à k. Or, contracter un crochet au niveau d'un sommet dont le label est au moins k donnera un chemin de taille $n+k-1$ qui va jusqu'à $(0, k-1)$, et ensuite jusqu'à $(n+k-1, n+k-1)$, i.e., un élément de $\mathscr{C}_{k,n+1}$, ce qui prouve le point (2).

□

Nous pouvons faire le constat suivant au sujet du nombre de crochets d'un chemin de Dyck de taille $n \geq 2$: s'il n'y a qu'un crochet, il se trouve nécessairement tout en haut à gauche, i.e. le chemin est un « agrandissement de facteur n » du crochet. Or, la seule façon d'obtenir ce chemin est via une augmentation d'un chemin de Dyck de taille $(n-1)$ qui est lui-même un agrandissement de facteur $(n-1)$.

Dans les autres cas, nous avons le résultat suivant :

Lemme 18. *Soit $n \geq 2$. Si $P_n \in \mathscr{D}_n$ contient plus de deux crochets, alors il existe $P_{n-1}, P'_{n-1} \in \mathscr{D}_{n-1}$, non nécessairement distincts, et $i \neq j$ tels que $P_n = \mathrm{Aug}(P_{n-1}, i) = \mathrm{Aug}(P'_{n-1}, j)$.*

Maintenant, supposons que nous augmentons un chemin $P_{k,n} \in \mathscr{C}_{k,n}$ non pas une fois, mais j fois au niveau de sommets de labels entre k et $2n+2k-1$. Ceci donnera lieu à un chemin $P_{k,n+j} \in \mathscr{C}_{k,n+j}$. Nous allons montrer qu'il existe une bijection entre l'ensemble des chemins $P_{k,n+j}$ qui peuvent être obtenus en augmentant j fois un chemin $P_{k,n}$ et l'ensemble des chemins $P_{k,n+j}$ qui peuvent être obtenus en augmentant une fois un chemin $P_{k,n+j-1}$ au niveau de i sommets différents. Avant de faire ceci, nous introduisons quelques notations :

Définition 34. *Soient $n, j \in \mathbb{N}$, $P_{k,n} \in \mathscr{C}_{k,n}$, et soit $I_j \subset \{k, \ldots, 2n+2k-1\}$ un multi-ensemble de taille j. Nous définissons $\mathrm{Aug}(P_{k,n}, I_j)$ comme étant l'insertion simultanée de j crochets dans le chemin $P_{k,n}$ au niveau des sommets dont les labels sont donnés par I_j.*

Les résultats suivants montrent que cette opération est cohérente.

Lemme 19. *Soient $i \geq 0$ et $P_{k,i} \in \mathscr{C}_{k,i}$. Alors, n'importe quel $P_{k,i+2} \in \mathscr{C}_{k,i+2}$ s'obtient en effectuant $P_{k,i+1} = \mathrm{Aug}(P_{k,i}, u)$ pour un $k \leq u \leq 2i+2k-1$, puis $P_{k,i+2} = \mathrm{Aug}(P_{k,i+1}, v)$ pour un $k \leq v \leq 2i+2k+1$.*

Démonstration. Nous pouvons constater qu'en augmentant un $P_{k,i} \in \mathscr{C}_{k,i}$ au niveau d'un sommet d'indice u, $k \leq u \leq 2i+1$, nous obtenons un $P_{k,i+1} \in \mathscr{C}_{k,i+1}$ dont les labels des sommets se déduisent de la manière suivante à partir des labels des sommets de $P_{k,i}$:

— Les labels $\leq u$ ne changent pas ;

— Les deux nouveaux sommets ont comme labels $u+1$ et $u+2$;

— Le label des autres sommets est incrémenté de 2..

□

Par conséquent, modulo ces nouveaux labels, nous avons le résultat suivant :

Corollaire 5. *Soient $i \geq 0$ et $P_{k,i} \in \mathscr{C}_{k,i}$. N'importe quel chemin $P_{k,i+2} \in \mathscr{C}_{k,i+2}$ s'obtient en effectuant $P_{k,i+2} = \mathrm{Aug}(P_{k,i}, I_2)$ pour un multi-ensemble $I_2 \subset \{k, \ldots, 2i+2k-1\}$ de taille 2.*

Corollaire 6. *Soient $n \geq 1$ et $1 \leq j \leq n$. Tout chemin $P_{k,n} \in \mathscr{C}_{k,n}$ s'obtient à partir d'un chemin $P_{k,n-j} \in \mathscr{C}_{k,n-j}$ en effectuant $\mathrm{Aug}(P_{k,n-j}, I_j)$ pour un multi-ensemble $I_j \subset \{k, \ldots, 2n+2k-2j-1\}$ de taille j.*

Démonstration. Nous pouvons le montrer par récurrence en utilisant des arguments similaires à ceux utilisés dans la preuve du Lemme 19.

□

Ainsi, nous avons démontré que la Définition 34 est bel et bien cohérente. Nous introduisons maintenant des notations pour les ensembles de sommets obtenus via des augmentations simples et multiples.

Définition 35.

1. *Soit $k \leq i \leq 2n+2k-3$. Nous définissons $W_{i,k,n} := \{\mathrm{Aug}(P_{k,n-1}, i) : P_{k,n-1} \in \mathscr{C}_{k,n-1}\}$, i.e., $W_{i,k,n}$ est le sous-ensemble de $\mathscr{C}_{k,n}$ obtenu en augmentant un élément de $\mathscr{C}_{k,n-1}$ au niveau du sommet de label i.*

2. *Soit $I_j \subset \{k, \ldots, 2n+2k-2j-1\}$ un multi-ensemble de taille j, $1 \leq j \leq n$. Nous définissons $W_{I_j,k,n} := \{\mathrm{Aug}(P_{k,n-j}, I_j) : P_{k,n-j} \in \mathscr{C}_{k,n-j}\}$, i.e., $W_{I_j,k,n}$ est le sous-ensemble de $\mathscr{C}_{k,n}$ obtenu à travers une augmentation multiple d'un élément de $\mathscr{C}_{k,n-j}$ à des positions indiquées par I_j.*

Proposition 12. *Soient $k \leq i_1 < \cdots < i_j \leq 2n+2k-3$. Alors il existe $k \leq i'_1 \leq \cdots \leq i'_j \leq 2n+2k-2j-1$ tels que $\bigcap_{u=1}^{j} W_{i_u,k,n} = W_{I_j,k,n}$, où I_j est le multi-ensemble $\left\{i'_1, \ldots, i'_j\right\}$.*

Démonstration. C'est trivialement vrai pour $j = 1$.

Pour $j = 2$, prenons i_1 et i_2 tels que $k \leq i_1 < i_2 \leq 2n+2k-3$. L'ensemble $W_{i_1,k,n} \cap W_{i_2,k,n}$ contient des chemins $P_{k,n} \in \mathscr{C}_{k,n}$ tels que $P_{k,n} = \mathrm{Aug}(P_{k,n-1}, i_1)$ et $P_n = \mathrm{Aug}(P'_{k,n-1}, i_2)$ avec $P_{n-1}, P'_{n-1} \in \mathscr{C}_{k,n-1}$. Par conséquent, ce chemin comporte deux crochets aux positions où les augmentations ont été effectuée. Maintenant, si nous contractons les deux crochets, cela nous donne un chemin $P_{k,n-2} \in \mathscr{C}_{k,n-2}$.

Rappeler que le fait que contracter un crochet change les labels des sommets de la manière suivante : si les sommets sur le crochets sont les sommets u, $u+1$, et $u+2$, les sommets dont le label est $\leq u$ gardent leur

labels, les sommets de labels $u+1$ et $u+2$ sont supprimés et le nouveau label de chaque sommet dont le label est $\geq u+2$ est l'ancien label diminué de 2.

Par conséquent, puisque $i_1 < i_2$, il est facile de voir que peut importe l'ordre dans lequel nous contractons les deux crochets, en prenant $P_{k,n-2}$ et en augmentant simultanément aux positions $i'_1 = i'_1$ et $i'_2 = i_2 - 2 \leq 2n+2k-5$, nous obtenons le chemin original $P_{k,n}$. De plus, si $i_2 = i_1+2$ (ce qui correspond à la situation où les deux crochets se suivent dans P_n), alors $i'_1 = i'_2$. C'est-à-dire que les indices peuvent être répétés.

Le cas général ressemble au cas $j = 2$. Pour $i_1, \ldots, i_j$ tels que $2 \leq i_1 < \cdots < i_j \leq 2n+2k-3$, un chemin P_n dans $\bigcap_{u=1}^{j} W_{i_u,j,n}$ contient j crochets. Si nous contractons tous ces crochets, nous obtenons un chemin $P_{n-j} \in \mathscr{D}_{n-j}$. En prenant ce chemin et en augmentant simultanément aux positions $i'_1 = i_1$, $i'_2 = i_2 - 2, \ldots,$ et $i'_j = i_2 - 2(j-1) \leq 2n+2k-2j-1$, nous retombons sur P_n.

□

Nous avons maintenant tous les éléments nécessaires pour prouver le Théorème 12.

Preuve du Théorème 12. Le Lemme 17 peut se reformuler de la manière suivante :

$$\mathscr{C}_{k,n+1} = \bigcup_{i=k}^{2n+2k-1} W_{i,k,n+1}\,.$$

Par conséquent

$$\mathsf{c}_{k,n+1} = |\mathscr{C}_{k,n+1}| = \left| \bigcup_{i=k}^{2n+2k-1} W_{i,k,n+1} \right|\,.$$

Nous utilisons le principe d'inclusion-exclusion, en nous arrêtant à $n+1$ au niveau des indices de la somme externe puisqu'un chemin de longueur $(n+1)+k-1$ (i.e., $n+k$) qui ne contient pas de crochets aux $k-1$ premières positions a au plus $n+1$ crochets.

$$\mathsf{c}_{k,n+1} = \sum_{i=1}^{n+1} (-1)^{i+1} \sum_{k \leq j_1 < j_2 < \cdots < j_i \leq 2n+2k-1} |W_{j_1,k,n+1} \cap W_{j_2,k,n+1} \cap \cdots \cap W_{j_i,k,n+1}|\,. \tag{7.3}$$

Ensuite, nous utilisons la Proposition 12 :

$$W_{j_1,k,n+1} \cap W_{j_2,k,n+1} \cap \cdots \cap W_{j_i,k,n+1} = W_{I_i,k,n+1}$$

pour un multi-ensemble $I_i \subset \{k, \ldots, 2n+2k-2i+1\}$ de taille i. Par conséquent, nous pouvons simplifier l'équation (7.3) :

$$\mathsf{c}_{k,n+1} = \sum_{i=1}^{n+1} (-1)^{i+1} \sum_{I_i \subset \{k,\ldots,2n+2k-2i+1\} \text{ multi-ensemble}} |W_{I_i,k,n+1}| \ . \tag{7.4}$$

Mais par définition, $|W_{I_i,k,n+1}|$ est le nombre d'éléments de $\mathscr{C}_{k,n+1}$ qui s'obtiennent en augmentant i fois chaque élément de $\mathscr{C}_{k,n+1-i}$ aux positions indiquées par I_i. Il y a donc autant de tels chemins qu'il y a d'éléments de $\mathscr{C}_{k,n+1-i}$, par conséquent $|W_{I_i,k,n+1}| = \mathsf{c}_{k,n+1-i}$ pour n'importe quel I_i. Et il y a $\binom{(2n-2i+k+2)+i-1}{i}$ multi-ensembles I_i possibles. Donc, finalement :

$$\sum_{I_i \subset \{k,\ldots,2n+2k-2i+1\} \text{ multi-ensemble}} |W_{I_i,k,n+1}| = \binom{2n+k+1-i}{i} \mathsf{c}_{k,n-i+1} \ .$$

Nous injections cela dans l'équation (7.4) afin d'obtenir

$$\mathsf{c}_{k,n+1} = \sum_{i=1}^{n+1} (-1)^{i+1} \binom{2n+k+1-i}{i} \mathsf{c}_{k,n-i+1}$$

ce qui donne le résultat souhaité à la substitution $i \mapsto i+1$ près.

□

7.3 Généralisation : nombres de Fuss-Catalan

Soit $s \in \mathbb{N}$. Nous considérons la suite de nombres $(\mathsf{FC}(s,n))_{n\in\mathbb{N}}$ définie de la manière suivante :

$$\mathsf{FC}(s,n) := \frac{1}{(s-1)n+1} \binom{sn}{n}$$

Cette suite s'appelle la *suite des nombres de Fuss-Catalan*, et nous voyons aisément que la suite $(\mathsf{c}_n)_{n\in\mathbb{N}}$ des nombres de Catalan est un cas particulier de cette suite en prenant $s = 2$. De plus, cette suite admet une interprétation similaire à celle des nombres de Catalan en termes de chemin sur la grille, à savoir la suivante [HP91] :

Nous supposons que nous disposons d'une grille à deux dimensions composée de $n \times (s-1)n$ carrés, c'est-à-dire un rectangle dont le côté horizontal est $(s-1)$ fois plus grand que son côté vertical. La diagonale de cette grille est la droite passant par les points $(x, (s-1)x)$. Nous considérons les chemins que nous pouvons tracer en partant du point $(0,0)$ en bas à gauche

jusqu'au point $(n,(s-1)n)$ en haut à droite en faisant uniquement des pas vers la droite ou vers le haut sur les arêtes de la grille, et tels que ce chemin ne descend jamais en dessous de la diagonale. Le nombre $\mathsf{FC}(s,n)$ donne alors le nombre de tels chemins sur une grille $n \times (s-1)n$: nous pouvons appeler ces chemins les *chemins de Dyck rectangulaires de taille* n *et de paramètre* s. Nous utiliserons la notation $\mathscr{D}_{s,n}$ pour l'ensemble de ces chemins, et nous pouvons noter qu'un élément de $\mathscr{D}_{s,n}$ comporte toujours $n+(s-1)n = sn$ arêtes et $n+(s-1)n+1 = sn+1$ sommets. Nous numérotons les sommets de 1 à $sn+1$ en commençant par en bas à gauche.

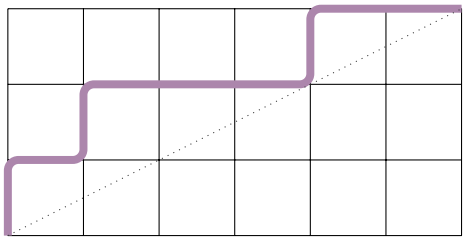


FIGURE 7.2 – Exemple d'un chemin de Dyck rectangulaire de taille 3 et de paramètre 3.

Dans ce qui suit, nous allons supposer que $s \geq 2$. En effet, $FC(1,n) = 1$ quel que soit $n \in \mathbb{N}$, et même si cela colle avec l'interprétation ci-dessus en termes de chemins [1], bien des éléments que nous introduisons par la suite ne sont valables que lorsque $s \geq 2$.

Nous avons alors le résultat suivant, qui est l'équivalent de l'équation (7.2) :

Théorème 13. *La suite* $(\mathsf{FC}(s,n))_{n\in\mathbb{N}}$ *des nombres de Fuss-Catalan satisfait la relation de récurrence suivante par rapport à* n *:*

$$\mathsf{FC}(s,n+1) = \sum_{i=0}^{n}(-1)^i \binom{s(n+1)-(s-1)(i+1)}{i+1} \mathsf{FC}(s,n-i).$$

Nous expliquons maintenant comment adapter la preuve du Théorème 12 afin de démontrer ce résultat.

Définition 36 (Crochet)**.** *Nous appelons* crochet *l'unique chemin de Dyck rectangulaire de paramètre* s *et de taille 1* P_1*, formé d'une arête montante puis de* $(s-1)$ *arêtes allant vers la droite.*

1. En effet, quel que soit $n \in \mathbb{N}$, il existe un seul chemin allant de $(0,0)$ à $(n,0)$ en n'utilisant que des pas vers la droite et vers le haut sur les arêtes de la grille.

Nous disons que nous *augmentons* un chemin $P_n \in \mathscr{D}_{s,n}$ au niveau du sommet i si nous insérons un crochet partant de ce sommet, ce qui donnera un chemin $P_{n+1} \in \mathscr{D}_{s,n+1}$. Nous noterons cette opération $\mathrm{Aug}(P_n, i)$.

Lemme 20. *L'ensemble $\mathscr{D}_{s,n+1}$ vérifie*

$$\mathscr{D}_{s,n+1} = \{\mathrm{Aug}(P_n, i) : P_n \in \mathscr{D}_{s,n}, 1 \leq i \leq sn+1\}$$

Démonstration. Le chemin résultant d'une augmentation d'un chemin $P_n \in \mathscr{D}_{s,n}$ de taille n au niveau de n'importe quel sommet i reste au-dessus de la diagonale et est donc un chemin de Dyck de taille $n+1$.

De plus, n'importe quel chemin $P_{n+1} \in \mathscr{D}_{s,n+1}$ terminent par au moins deux arêtes horizontales : s'ils terminaient par une arête verticale puis une arête horizontale, le point $(n, (s-1)n-1)$, qui est en-dessous de la diagonale passant par les points $(x, (s-1)x)$, serait sur le chemin. Par conséquent, il y a au moins deux arêtes horizontales après la dernière arête verticale du chemin, i.e. il y a un crochet. Et en enlevant ce crochet, nous obtenons un chemin $P_{n+1} \in \mathscr{D}_{s,n+1}$.

□

Contrairement aux chemins de Dyck « classiques », avec les chemins de Dyck de paramètre s et de taille $n \geq 2$, il est possible d'avoir un seul crochet sans que le chemin soit un « agrandissement de facteur n ». Néanmoins, lorsqu'il y a plus de deux crochets, nous avons le résultat suivant :

Lemme 21. *Soit $n \geq 2$. Si $P_n \in \mathscr{D}_n$ contient plus de deux crochets, alors il existe $P_{n-1}, P'_{n-1} \in \mathscr{D}_{n-1}$, non nécessairement distincts, et $i \neq j$ tels que $P_n = \mathrm{Aug}(P_{n-1}, i) = \mathrm{Aug}(P'_{n-1}, j)$.*

En particulier, nous pouvons constater que dans un chemin de longueur $n \geq 2$, s'il y a un crochet au niveau du sommet i, alors le chemin comporte un deuxième crochet. Dans ce cas, le lemme précédent nous assure qu'il existe un $i > 1$ tel que ce chemin peut également être obtenu en augmentant un chemin de longueur $(n-1)$ au niveau du sommet i, i.e. nous pouvons « oublier » les augmentations au niveau du sommet 1. Cela nous permet de préciser le Lemme 20 :

Corollaire 7. *L'ensemble $\mathscr{D}_{s,n+1}$ vérifie*

$$\mathscr{D}_{s,n+1} = \{\mathrm{Aug}(P_n, i) : P_n \in \mathscr{D}_{s,n}, 2 \leq i \leq sn+1\}$$

Ensuite, nous définissons l'augmentation multiple, et nous démontrons que l'ensemble des chemins $P_n \in \mathscr{D}_{s,n}$ obtenus à travers une augmentation

multiple d'un chemin $P_{n-j} \in \mathscr{D}_{s,n-j}$ est identique à l'ensemble des chemins $P_n \in \mathscr{D}_{s,n}$ qui peuvent être obtenus en augmentant un chemin $P_{n-1} \in \mathscr{D}_{s,n-1}$ au niveau de j sommets différents.

Définition 37. *Soient $n, j \in \mathbb{N}$, $P_n \in \mathscr{D}_{s,n}$, et soit $I_j \subset \{1, \ldots, sn+1\}$ un multi-ensemble de taille j. On définit $\mathrm{Aug}(P_n, I_j)$ comme étant l'insertion simultanée de j crochets dans le chemin P_n au niveau des sommets dont les labels sont donnés par I_j.*

Les résultats suivants montrent que cette définition est cohérente.

Lemme 22. *Soient $i \geq 0$ et $P_i \in \mathscr{D}_{s,i}$. Alors, n'importe quel $P_{i+2} \in \mathscr{D}_{s,i+2}$ s'obtient en effectuant $P_{i+1} = \mathrm{Aug}(P_i, u)$ pour un $1 \leq u \leq si+1$, puis $P_{i+2} = \mathrm{Aug}(P_{i+1}, v)$ pour un $1 \leq v \leq s(i+1)+1$.*

La preuve de ce lemme repose, comme pour le lemme 19, sur une renumérotation des sommets. Et modulo cette renumérotation, nous avons le résultat suivant :

Corollaire 8. *Soient $i \geq 0$ et $P_i \in \mathscr{D}_{s,i}$. N'importe quel chemin $P_{i+2} \in \mathscr{D}_{s,i+2}$ s'obtient en effectuant $P_{s,i+2} = \mathrm{Aug}(P_i, I_2)$ pour un multi-ensemble $I_2 \subset \{1, \ldots, si+1\}$ de taille 2.*

Corollaire 9. *Soient $n \geq 1$ et $1 \leq j \leq n$. Tout chemin $P_n \in \mathscr{D}_{s,n}$ s'obtient à partir d'un chemin $P_{n-j} \in \mathscr{D}_{s,n-j}$ en effectuant $\mathrm{Aug}(P_{n-j}, I_j)$ pour un multi-ensemble $I_j \subset \{1, \ldots, s(n-j)+1\}$ de taille j.*

Nous introduisons maintenant des notations pour les ensembles de sommets obtenus via des augmentations simples et multiples.

Définition 38.

1. *Soit $1 \leq i \leq s(n-1)-1$. On définit $W_{i,n} := \{\mathrm{Aug}(P_{n-1}, i) : P_{n-1} \in \mathscr{D}_{s,n-1}\}$, i.e., $W_{i,n}$ est le sous-ensemble de $\mathscr{D}_{s,n}$ obtenu en augmentant un élément de $\mathscr{D}_{s,n-1}$ au niveau du sommet numéro i.*
2. *Soit $I_j \subset \{1, \ldots, s(n-j)+1\}$ un multi-ensemble de taille j, $1 \leq j \leq n$. On définit $W_{I_j,n} := \{\mathrm{Aug}(P_{n-j}, I_j) : P_{n-j} \in \mathscr{D}_{s,n-j}\}$, i.e., $W_{I_j,n}$ est le sous-ensemble de $\mathscr{D}_{s,n}$ obtenu à travers une augmentation multiple d'un élément de $\mathscr{D}_{s,n-j}$ à des positions indiquées par I_j.*

Proposition 13. *Soient $1 \leq i_1 < \cdots < i_j \leq s(n-1)+1$. Alors il existe $1 \leq i'_1 \leq \cdots \leq i'_j \leq s(n-j)+1$ tels que $\bigcap_{u=1}^{j} W_{i_u,n} = W_{I_j,n}$, où I_j est le multi-ensemble $\left\{i'_1, \ldots, i'_j\right\}$.*

Cette proposition se démontre de façon similaire à la Proposition 12.

Nous avons maintenant tous les éléments nécessaires pour prouver le Théorème 13.

Preuve du Théorème 13. Le Corollaire 7 peut se reformuler de la manière suivante :

$$\mathscr{D}_{s,n+1} \;=\; \bigcup_{i=2}^{sn+1} W_{i,n+1}\,.$$

Par conséquent

$$\mathsf{FC}(s,n+1) = |\mathscr{D}_{s,n+1}| = \left|\bigcup_{i=2}^{sn+1} W_{i,n+1}\right|\,.$$

Nous utilisons le principe d'inclusion-exclusion, en nous arrêtant à $n+1$ au niveau des indices de la somme externe puisqu'un chemin de longueur $(n+1)$ a au plus $(n+1)$ crochets.

$$\mathsf{FC}(k,n+1) \;=\; \sum_{i=1}^{n+1}(-1)^{i+1} \sum_{2\le j_1<j_2<\cdots<j_i\le sn+1} |W_{j_1,n+1}\cap W_{j_2,n+1}\cap\cdots\cap W_{j_i,n+1}|\,. \tag{7.5}$$

Ensuite, nous utilisons la Proposition 13 :

$$W_{j_1,n+1}\cap W_{j_2,n+1}\cap\cdots\cap W_{j_i,n+1} \;=\; W_{I_i,n+1}$$

pour un multi-ensemble $I_i \subset \{2,\dots,s(n+1-i)+1\}$ de taille i. Par conséquent, nous pouvons simplifier l'équation (7.5) :

$$\mathsf{FC}(s,n+1) \;=\; \sum_{i=1}^{n+1}(-1)^{i+1} \sum_{I_i\subset\{2,\dots,s(n+1-i)+1\}\text{ multi-ensemble}} |W_{I_i,n+1}|\,. \tag{7.6}$$

Mais par définition, $|W_{I_i,n+1}|$ est le nombre d'éléments de $\mathscr{D}_{s,n+1}$ qui s'obtiennent en augmentant i fois chaque élément de $\mathscr{D}_{s,n+1-i}$ aux positions indiquées par I_i. Il y a donc autant de tels chemins qu'il y a d'éléments de $\mathscr{D}_{s,n+1-i}$, par conséquent $|W_{I_i,n+1}| = \mathsf{FC}(k,n+1-i)$ pour n'importe quel I_i. Et il y a $\binom{s(n+1-i)+i-1}{i}$ multi-ensembles I_i possibles. Donc, finalement :

$$\sum_{I_i\subset\{2,\dots,s(n+1-i)+1\}\text{ multi-ensemble}} |W_{I_i,n+1}| = \binom{s(n+1-i)+i-1}{i}\mathsf{FC}(k,n-i+1)\,.$$

Nous injectons cela dans l'équation (7.6) afin d'obtenir

$$\mathsf{FC}(s,n+1) \;=\; \sum_{i=1}^{n+1}(-1)^{i+1}\binom{sn-(s-1)(i-1)}{i}\mathsf{FC}(k,n-i+1)$$

Nous pouvons remarquer que pour $i = n+1$, le terme binomial correspondant vaut $\binom{sn-(s-1)n}{n+1} = \binom{n}{n+1} = 0$, donc il suffit d'aller jusqu'à $i = n$ dans la somme. Enfin, la substitution $i \mapsto i+1$ donne le résultat souhaité :

$$\mathsf{FC}(s, n+1) \;=\; \sum_{i=0}^{n-1} (-1)^i \binom{sn-(s-1)i}{i+1} \mathsf{FC}(k, n-i)$$

□

Concluons par une remarque sur cette preuve : pour obtenir la formule souhaitée, nous avons utilisé le Corollaire 7 et non le Lemme 20. Si nous avions utilisé ce lemme à la place, la preuve se serait déroulé de la même manière avec les différences suivantes : nous aurions pris des multi-ensembles $I_i \subset \{1, \ldots, s(n+1-i)+1\}$ et nous n'aurions pas pu supprimer le terme de la somme pour $i = n+1$. Nous aurions alors obtenu la formule

$$\mathsf{FC}(s, n+1) \;=\; \sum_{i=0}^{n} (-1)^i \binom{sn-(s-1)i+1}{i+1} \mathsf{FC}(k, n-i).$$

En effet, dans ce cas .

7.4 Une question ouverte

Nous concluons ce chapitre sur la remarque suivante au sujet d'une question ouverte : le Théorème 12 est plus général que le Théorème 13 dans la mesure où il fait intervenir les convolutions des nombres de Catalan. Nous pouvons donc nous demander s'il existe un résultat similaire pour les convolutions des nombres de Fuss-Catalan. Cela nécessiterait d'étudier si le résultat prouvé dans [Ted11] dont nous nous sommes servi se généralise aux nombres de Fuss-Catalan. Nous n'avons pas étudié cette question puisqu'elle était périphérique à nos sujets d'étude principaux. Néanmoins, si cela est vrai, il est vraisemblable que le résultat suivant l'est également :

Conjecture 1. *Pour tout $s, k \in \mathbb{N}$, $s \geq 2$ et $k \geq 1$, la suite $(\mathsf{FC}_k(s,n))_{n\in\mathbb{N}}$ de la convolution k-ème des nombres de Fuss-Catalan satisfait la relation de récurrence suivante par rapport à n :*

$$\mathsf{FC}_k(s, n+1) \;=\; \sum_{i=0}^{n} (-1)^i \binom{sn-(s-1)(i+1-k)+1}{i+1} \mathsf{FC}_k(s, n-i).$$

Le Théorème 12 prouve cette conjecture pour $s = 2$ et k quelconque, alors qu'il est facile d'adapter la preuve du Théorème 13 pour prouver cette conjecture pour s quelconque et $k = 1$ (cf. la remarque après la preuve de ce théorème).

Conclusion

L'apport principal de la première partie de cette thèse est théorique, en clarifiant certaines subtilités autour du modèle ϕ-LOCAL et en proposant un modèle plus adéquat qui ne souffre pas des limites du modèle ϕ-LOCAL. Mais ce nouveau modèle est plus complexe que le modèle ϕ-LOCAL classique, et les problèmes y sont plus difficiles à étudier, ce qui va à l'encontre de la motivation derrière le modèle ϕ-LOCAL, à savoir pouvoir résoudre plus facilement la question de la possibilité de certains algorithmes quantiques distribués.

Dans la deuxième partie de cette thèse, nous avons donné une description complète des distributions de probabilités 1-localisables sur les mots hard-core. Ce faisant, nous avons également démontré que la 3-coloration 1-localisable est impossible. Nous avons vu que ce résultat s'étend facilement aux distributions de probabilités k-localisables sur ce que nous avons appelés des mots k-hard-core, et qui sont liés au problème de coloration à distance k comme les mots hard-core le sont au problème de coloration. Cependant, il n'est pas clair dans quelle mesure nos résultats se généralisent aux « motifs sur les graphes », qu'il s'agisse de motifs de coloration ou motifs hard-core : nous n'avons réussi à généraliser qu'une partie de nos résultats. Enfin, pour prouver nos résultats, nous avons démontré des résultats intéressants en soi : un théorème sur la solution explicite d'un programme linéaire ayant une forme particulière, et une formule sur les nombres de Catalan et une généralisation de ces nombres, les nombres de Fuss-Catalan.

Bibliographie

[ABB+10] M. L. Almeida, J.-D. Bancal, N. Brunner, A. Acín, N. Gisin, and S. Pironio, *Guess your neighbor's input : A multipartite nonlocal game with no quantum advantage*, Physical Review Letters, 104 (2010), p. 230404. doi : 10.1103/PhysRevLett.104.230404.

[AGKdV89] J. Aaronson, D. Gilat, M. Keane, and V. de Valk, *An Algebraic Construction of a Class of One-Dependent Processes*, Annals of Probability, 17 (1989), pp. 128–143. doi : 10.1214/aop/1176991499.

[AGR81] A. Aspect, P. Grangier, and G. Roger, *Experimental Tests of Realistic Local Theories via Bell's Theorem*, Physical Review Letters, 47 (1981), pp. 460–463. doi : 10.1103/PhysRevLett.47.460.

[APT19] A. Atanasiu, G. Poovanandran, and W. C. Teh, *Parikh Determinants*, 2019, pp. 68–79. doi : 10.1007/978-3-030-28796-2_5.

[Arf14] H. Arfaoui, *Local Distributed Decision and Verification*, PhD thesis, Université Paris Diderot - Paris 7, July 2014.

[Bar07] J. Barrett, *Information processing in generalized probabilistic theories*, Physical Review A, 75 (2007). doi : 10.1103/PhysRevA.75.032304.

[BE13] L. Barenboim and M. Elkin, *Distributed Graph Coloring : Fundamentals and Recent Developments*, no. 1 in Synthesis Lectures on Distributed Computing Theory, 4, Morgan & Claypool, July 2013. doi : 10.2200/S00520ED1V01Y201307DCT011.

[Bel64] J. S. Bell, *On the Einstein-Podolsky-Rosen paradox*, Physics, 1 (1964), pp. 195–200. doi : 10.1103/PhysicsPhysiqueFizika.1.195.

[Bel81] J. S. Bell, *Bertlmann's socks and the nature of reality*, J. Phys. Colloques, 2 (1981), pp. 41–62. doi : 10.1051/jphyscol:1981202.

[BEPS16] L. Barenboim, M. Elkin, S. Pettie, and J. Schneider, *The Locality of Distributed Symmetry Breaking*, J. ACM, 63 (2016), pp. 20 :1–20 :45. doi : 10.1145/2903137.

[BLM+05] J. Barrett, N. Linden, S. Massar, S. Pironio, S. Popescu, and D. Roberts, *Nonlocal correlations as an information-*

theoretic resource, Physical Review A, 71 (2005). DOI : 10.1103/PhysRevA.71.022101.

[BOH05] M. Ben-Or and A. Hassidim, *Fast quantum byzantine agreement*, STOC '05, New York, NY, USA, 2005, ACM, pp. 481–485. DOI : 10.1145/1060590.1060662, http://doi.acm.org/10.1145/1060590.1060662.

[CHSH69] J. Clauser, M. Horne, A. Shimony, and R. Holt, *Proposed experiment to test local hidden-variable theories*, Physical Review Letters, 23 (1969), pp. 880–884. DOI : 10.1103/PhysRevLett.23.880.

[Chv83] V. Chvátal, *Linear Programming*, W. H. Freeman, 1983. DOI : 10.2307/2008305.

[CKP16] Y.-J. Chang, T. Kopelowitz, and S. Pettie, *An exponential separation between randomized and deterministic complexity in the local model*, SIAM Journal on Computing, 48 (2016). DOI : 10.1137/17M1117537.

[CMN+06] P. J. Cameron, A. Montanaro, M. W. Newman, S. A. Severini, and A. Winter, *On the quantum chromatic number of a graph*, Electr. J. Comb., 48 (2006).

[CPS14] K.-M. Chung, S. Pettie, and H.-H. Su, *Distributed algorithms for the lovász local lemma and graph coloring*, Proceedings of the Annual ACM Symposium on Principles of Distributed Computing, (2014). DOI : 10.1145/2611462.2611465.

[CV86] R. Cole and U. Vishkin, *Deterministic Coin Tossing and Accelerating Cascades : Micro and Macro Techniques for Designing Parallel Algorithms*, STOC '86, New York, NY, USA, 1986, ACM, pp. 206–219. DOI : 10.1145/12130.12151.

[Dia55] P. H. Diananda, *The central limit theorem for m-dependent variables*, Proceedings of the Cambridge Philosophical Society, 51 (1955), pp. 92–95. DOI : 10.1017/S0305004100029959.

[Dia77] P. Diaconis, *Finite forms of de Finetti's theorem on exchangeability*, Synthese, 36 (1977), pp. 271–281. DOI : 10.1007/BF00486116.

[DSDS05] A. D. Scott and A. D. Sokal, *The Repulsive Lattice Gas, the Independent-Set Polynomial, and the Lovász Local Lemma*, Journal of Statistical Physics, 118 (2005), pp. 1151–1261. DOI : 10.1007/s10955-004-2055-4.

[EPR35] A. Einstein, B. Podolsky, and N. Rosen, *Can quantum-mechanical description of physical reality be considered complete?*, Physical Review, 47 (1935), pp. 777–780. DOI : 10.1103/PhysRev.47.777.

[FLP85] M. J. Fischer, N. A. Lynch, and M. S. Paterson, *Impossibility of Distributed Consensus with One Faulty Process*, J. ACM, 32 (1985), pp. 374–382. DOI : 10.1145/3149.214121.

[G+15] M. Giustina et al., *Significant-Loophole-Free Test of Bell's Theorem with Entangled Photons*, Physical Review Letters, 115 (2015), p. 250401. doi : 10.1103/PhysRevLett.115.250401.

[GK19] M. Ghaffari and F. Kuhn, *On the use of randomness in local distributed graph algorithms*, 07 2019, pp. 290–299. doi : 10.1145/3293611.3331610.

[GKM09] C. Gavoille, A. Kosowski, and M. Marcin, *What Can Be Observed Locally? Round-based Models for Quantum Distributed Computing*, International Symposium on Distributed Computing, (2009), pp. 243–257. doi : 10.1007/978-3-642-04355-0_26.

[GKZ19] C. Gavoille, G. Kachigar, and G. Zémor, *Localisation-Resistant Random Words with Small Alphabet*, Words 2019, (2019), pp. 193–206. doi : 10.1007/978-3-030-28796-2_15.

[HL15] A. E. Holroyd and T. Liggett, *Symmetric 1-dependent colorings of the integers*, Electronic Communications in Probability, 20 (2015), pp. 1–8. doi : 10.1214/ECP.v20-4070.

[HL16] A. Holroyd and T. Liggett, *Finitely dependent coloring*, Forum of Mathematics, Pi, 4 (2016). doi : 10.1017/fmp.2016.7.

[Hol17] A. E. Holroyd, *One-dependent coloring by finitary factors*, Annales de Instut Henri Poincaré Probab. Stat., 53 (2017), pp. 753–765. doi : 10.1214/15-AIHP735.

[HP91] P. Hilton and J. Pedersen, *Catalan numbers, their generalization, and their uses*, The Mathematical Intelligencer, 13 (1991), pp. 64–75. doi : 10.1007/BF03024089.

[HR48] W. Hoeffding and H. Robbins, *The central limit theorem for dependent random variables*, Duke Mathematical Journal, 15 (1948). doi : 10.1215/S0012-7094-48-01568-3.

[HSW16] A. Holroyd, O. Schramm, and D. Wilson, *Finitary coloring*, Annals of Probability, (2016). doi : 10.1214/16-AOP1127.

[Kir19] W. Kirsch, *An elementary proof of de Finetti's theorem*, Statistics & Probability Letters, 151 (2019), pp. 84–88. doi : 10.1016/j.spl.2019.03.014.

[KK08] F. Kramer and H. Kramer, *A survey on the distance-colouring of graphs*, Discrete Mathematics, 308 (2008), pp. 422–426. doi : 10.1016/j.disc.2006.11.059. Combinatorics04.

[Kos09] T. Koshy, *Catalan numbers with applications*, Catalan Numbers with Applications, (2009). doi : 10.1093/acprof:oso/9780195334548.001.0001.

[KW06] F. Kuhn and R. Wattenhofer, *On the Complexity of Distributed Graph Coloring*, PODC '06, New York, NY, USA, 2006, ACM, pp. 7–15. doi : 10.1145/1146381.1146387.

[LGM18] F. Le Gall and F. Magniez, *Sublinear-time quantum computation of the diameter in congest networks*, PODC '18, New York, NY, USA, 2018, ACM, pp. 337–346. DOI : 10.1145/3212734.3212744, http://doi.acm.org/10.1145/3212734.3212744.

[Lin87] N. Linial, *Distributive Graph Algorithms Global Solutions from Local Data*, SFCS '87, Washington, DC, USA, 1987, IEEE Computer Society, pp. 331–335. DOI : 10.1109/SFCS.1987.20.

[Lin92] N. Linial, *Locality in Distributed Graph Algorithms*, SIAM J. Comput., 21 (1992), pp. 193–201. DOI : 10.1137/0221015.

[LSP82] L. Lamport, R. Shostak, and M. Pease, *The Byzantine Generals Problem*, ACM Trans. Program. Lang. Syst., 4 (1982), pp. 382–401. DOI : 10.1145/357172.357176.

[Man13] S. Mansfield, *The Mathematical Structure of Non-locality & Contextuality*, PhD thesis, Oxford University, 2013.

[MLT18] T. M. Liggett and W. Tang, *One-dependent hard-core processes and colorings of the star graph*, (2018).

[Nao91] M. Naor, *A Lower Bound on Probabilistic Algorithms for Distributive Ring Coloring*, SIAM J. Discrete Math., 4 (1991), pp. 409–412. DOI : 10.1137/0404036.

[New91] I. Newman, *Private vs. common random bits in communication complexity*, Inf. Process. Lett., 39 (1991), pp. 67–71. DOI : 10.1016/0020-0190(91)90157-D.

[Pel00] D. Peleg, *Distributed Computing : A Locality-sensitive Approach*, SIAM Monographs on Discrete Mathematics and Applications, Philadelphia, PA, USA, 2000. ISBN : 0-89871-464-8. DOI : 10.1137/1.9780898719772.

[Pet90] D. Petz, *A de Finetti-type theorem with m-dependent states*, Probability Theory and Related Fields, 85 (1990), pp. 65–72. DOI : 10.1007/BF01377629.

[PR94] S. Popescu and D. Rohrlich, *Nonlocality as an axiom*, Foundations of Physics, 24 (1994), pp. 379–385. DOI : 10.1007/BF02058098.

[PR01] A. Panconesi and R. Rizzi, *Some Simple Distributed Algorithms for Sparse Networks*, Distrib. Comput., 14 (2001), pp. 97–100. DOI : 10.1007/PL00008932.

[S+15] L. K. Shalm et al., *Strong Loophole-Free Test of Local Realism*, Physical Review Letters, 115 (2015), p. 250402. DOI : 10.1103/PhysRevLett.115.250402.

[Sch35] E. Schrödinger, *Die Gegenwärtige Situation in der Quantenmechanik*, Naturwissenschaften, 23 (1935), p. 807–812. DOI : 10.1007/978-3-663-14179-2_9.

[SSW10] L. Song, W. Staton, and B. Wei, *Independence polynomials of k-tree related graphs*, Discrete Applied Mathematics, 158 (2010), pp. 943–950. doi : 10.1016/j.dam.2010.01.002.

[Ted11] S. J. Tedford, *Combinatorial interpretations of convolutions of the Catalan numbers*, Integers, 11 (2011), pp. 35–45. doi : 10.1515/integ.2011.003.

[TKM12] S. Tani, H. Kobayashi, and K. Matsumoto, *Exact quantum algorithms for the leader election problem*, ACM Trans. Comput. Theory, 4 (2012), pp. 1 :1–1 :24. doi : 10.1145/2141938.2141939, http://doi.acm.org/10.1145/2141938.2141939.

[Tsi80] B. S. Tsirelson, *Quantum generalizations of Bell's inequality*, Letters in Mathematical Physics, 4 (1980), pp. 93–100. doi : 10.1007/BF00417500.

[vD13] W. van Dam, *Implausible Consequences of Superstrong Nonlocality*, 12 (2013), pp. 9–12. doi : 10.1007/s11047-012-9353-6.

[Wika] Wikipedia, *Loopholes in Bell test experiments*. https://en.wikipedia.org/wiki/Loopholes_in_Bell_test_experiments. Version du 07/05/2019.

[Wikb] Wikipedia, *Stars and bars (combinatorics)*. https://en.wikipedia.org/wiki/Stars_and_bars_(combinatorics). Version du 14/08/2019.

[Wis06] H. Wiseman, *From Einstein's Theorem to Bell's Theorem : A History of Quantum Nonlocality*, Contemporary Physics, 47 (2006). doi : 10.1080/00107510600581011.