HAL open science

Problèmes d'identification dans les graphes

Aline Parreau

To cite this version:

Aline Parreau. Problèmes d'identification dans les graphes. Mathématiques générales [math.GM]. Université de Grenoble, 2012. Français. NNT : 2012GRENM028 . tel-00745054

HAL Id: tel-00745054

https://theses.hal.science/tel-00745054

Submitted on 24 Oct 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ DE GRENOBLE

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Mathématiques - Informatique**

Arrêté ministériel : 7 août 2006

Présentée par

Aline Parreau

Thèse dirigée par **Sylvain Gravier**

préparée au sein de l' **Institut Fourier**
et de l'école doctorale **EDMSTII**

Problèmes d'identification dans les graphes

Thèse soutenue publiquement le **5 juillet 2012**,
devant le jury composé de :

M. Sylvain Gravier
CNRS - Université Joseph Fourier, Directeur de thèse
M. Olivier Hudry
Télécom ParisTech, Rapporteur
M. Ralf Klasing
CNRS-Université Bordeaux 1, Rapporteur
M. Sandi Klavžar
Université de Ljubljana, Examinateur
M. Julien Moncel
Université Toulouse 1 Capitole, Examinateur
Mme. Myriam Preissmann
CNRS - Grenoble INP, Examinatrice

Remerciements

Mes premiers remerciements vont à mon directeur de thèse, Sylvain Gravier, pour m'avoir permis d'effectuer cette thèse dans des conditions idéales. Merci Sylvain pour la confiance que tu m'as accordée, de m'avoir guidée et supportée pendant ces quelques années. Merci aussi pour la gestion de Maths à Modeler qui est une source de motivation pour moi. J'espère que tu ne regrettes pas trop de ne pas avoir suivi les conseils de Laurent...

Je voudrais remercier ensuite les membres de mon jury. Merci tout d'abord à Olivier Hudry et Ralf Klasing d'avoir accepté avec enthousiasme d'être rapporteurs de ma thèse et de l'avoir relue avec attention. Merci à Sandi Klavžar, Julien Moncel et Myriam Preissmann d'avoir accepté de faire partie de mon jury. Enfin merci à Michel Mollard qui avait aussi accepté d'être membre de ce jury mais qui n'a pas pu être disponible à cette date.

Durant ma thèse, j'ai eu la possibilité de partir en Finlande quelques mois (grâce à la région Rhône-Alpes que je remercie aussi). Je voudrais remercier tous les finlandais qui m'ont accueillie là-bas et qui ont fait de ce séjour une chouette expérience. "Kittos paljon" en particulier à Tero Laihonen et Iiro Honkala, l'accueil que vous m'avez reservé était aussi chaleureux que les températures étaient basses !

J'accorde une importance toute particulière au travail en groupe et je tiens ainsi à remercier les nombreuses personnes avec qui j'ai eu l'occasion de travailler. Un merci particulier à Louis Esperet qui a subi mes nombreuses questions existentielles à G-Scop. Merci Louis pour tous tes conseils, ta disponibilité et les jolis problèmes que tu m'as fait découvrir. Une grande partie de ma thèse a été réalisée dans le cadre du projet ANR IDEA. Si je ne suis pas très favorable aux principes des projets ANR, je dois bien avouer que celui-ci m'a permis de travailler dans de bonnes conditions, grâce notamment aux rencontres biannuelles du projet et au "workshop" final. Merci à Ralf d'avoir si bien géré le projet. Merci aux doctorants et post-doctorants qui ont fait vivre le projet et en particulier Florent, Petru et Reza qui de par leur enthousiasme et leur bonne humeur ont créé une ambiance sympathique et propice à la recherche. J'ai beaucoup apprécié les moments où nous avons travaillé ensemble (mais il me semble que vous me devez encore quelques bières...). Merci à tous les autres membres de l'ANR et en particulier Olivier Delmas, Paul Dorbec, Éleonora Guerrini, Julien Moncel, Mickaël Montassier, André Raspaud, Eric Sopena. Merci aux montpelliérains, Daniel Gonçalves et Alexandre Pinlou, et aux barcelonais, Oriol Serra et Guillem Perarnau, de m'avoir invitée dans vos laboratoires respectifs.

Merci à toute l'équipe du Master ROCO, et en particulier à Nadia Brauner, qui outre son plaisir communicatif à enseigner, se soucie en permanence de ma petite personne.

Un grand merci à mes "frère et soeur de thèse". Élise, ta présence pendant ces six derniers mois a été d'un grand soutien. Merci pour ta joie de vivre et ton énergie inépuisable. Merci d'avoir "su" relire des bouts de ma thèse, d'avoir assisté à quelques soutenances, d'avoir supporté mes ralages quotidiens, de m'avoir fait goûter la soupe à la banane et de m'avoir nourrie en chocolats à l'orange. J'espère bien pouvoir te rendre la pareille un jour ! Un peu plus haut au troisième étage, merci à Simon Modeste, ta bonne humeur, tes blagues, tes nouvelles idées quotidiennes pour travailler ensemble (il faudrait qu'on s'y mette un jour) et tes bavardages sur tout et rien ont égayé mes journées passées à l'Institut Fourier. Merci à ma cobureau, Arrrriadna, avec qui j'ai pu partager les joies de la dernière année de thèse. Je crois qu'on devient assez douées en concours de ralages. Merci pour tous les thés partagés et ces quelques jours passés dans le grand nord. Merci à Maxime d'avoir veillé à ce que je sois bien nourrie (mais économiquement !) en produits frais de l'Estacade, d'avoir supporté Caravan Palace et de ne pas m'avoir (trop) forcée à manger du fromage. Merci aux membres de l'Institut Fourier, de G-Scop et de Maths à Modeler, merci aux doctorants et anciens doctorants d'ici ou d'ailleurs avec qui j'ai passé de bons moments. Je pense en particulier à Bashar, Thomas, Jean, Camille, Marianne, Gunnar, Mathieu, Laurent, Alix, Nicolas, Fred, Grégoire, Marion,... Merci au personnel administratif qui nous simplifie grandement la vie.

Un immense merci à la Compagnie Rire et Soupleté et à ses associés. Les croûtes à la violette m'ont bien changé les idées pendant ces quatre années grenobloises. Un merci particulier à Liz (pour les relectures et les pique-niques), Blebête (pour les frites à la mayo), Rachtouf (juste trois mots) et Olivier (pour m'avoir fait perdre mon temps en cousant des costumes éléphants à des balles de jonglage). Merci à tous mes amis éparpillés un peu partout. Bzou et Bzette, les nombreux moments que l'on a passés ensemble étaient "juste trop bien". Merci aux amis montreuillois (ou assimilés), toujours fidèles aux postes : Pauline, Mathieu, Charlotte, Nicolas. Merci à ces deux derniers d'avoir fait le déplacement, cela me touche beaucoup.

Les derniers remerciements (et non des moindres) vont à ma famille. De la plus petite aux plus grands, merci à vous tous d'être près de moi. Merci à mes parents, sans qui je ne serais pas là aujourd'hui, d'avoir corrigé (quelques) (grosses) fautes et d'avoir fait des feuilles de vignes. Merci à Anne et Sylvain pour le Cerdon et le houmous (entre autres) et à Cécile d'avoir relu l'introduction et de faire autre chose que des maths, ça change ! (Je ne désespère cependant pas de te faire participer à Maths à Modeler un jour...) Olivier, merci d'avoir voulu assister à la soutenance et désolée du changement de date ! Merci à mes quatre neveux préférés de me ramener aux choses essentielles. Vincent, merci pour ta patience infinie (il en faut pour me supporter) et ton soutien sans faille, même à distance. Te savoir à mes côtés m'aide à avancer chaque jour.

Table des matières

Introduction **9**

1 Préliminaires **13**
1.1 Graphes . . . 14
1.1.1 Graphe, voisinage, degré . . . 14
1.1.2 Opérations sur les graphes . . . 15
1.1.3 Exemples et classes de graphes . . . 16
1.1.4 Coloration de graphes et graphes parfaits . . . 19
1.1.5 Hypergraphes . . . 19
1.1.6 Graphe orienté . . . 19
1.1.7 Représentation à l'aide de graphes bipartis . . . 20
1.2 Identification . . . 22
1.2.1 Codes identifiants dans les graphes non orientés . . . 22
1.2.2 Codes discriminants dans les graphes bipartis d'incidence . . . 24
1.2.3 Variantes . . . 26
1.3 Premiers résultats . . . 28
1.3.1 Existence . . . 28
1.3.2 Hérédité . . . 28
1.3.3 Bornes . . . 28
1.3.4 Complexité . . . 30

2 Familles extrémales **31**
2.1 Outils préliminaires . . . 32
2.1.1 Théorème de Bondy . . . 32
2.1.2 Autres outils . . . 34
2.2 Cas orienté . . . 35
2.2.1 Graphes finis . . . 35
2.2.2 Graphes infinis . . . 41
2.3 Cas non orienté . . . 44
2.3.1 Graphes finis . . . 44
2.3.2 Graphes infinis . . . 51
2.3.3 Conséquences . . . 54
2.3.4 Codes r-identifiants . . . 55
2.4 Conclusion . . . 56

3 Classes de graphes **59**
3.1 Graphes de Sierpiński . . . 60
3.1.1 Définition des graphes de Sierpiński . . . 61
3.1.2 Codes identifiants . . . 62
3.1.3 Codes localisateur-dominateurs . . . 63
3.1.4 Ensembles résolvants . . . 64
3.1.5 Conclusion sur les graphes de Sierpiński . . . 67
3.2 Graphes d'intervalles . . . 68
3.2.1 Borne inférieure générale . . . 68
3.2.2 Graphes d'intervalles unitaires . . . 69
3.2.3 Discussion sur la complexité . . . 70
3.3 Graphes adjoints . . . 71
3.3.1 Codes arête-identifiants . . . 71
3.3.2 Premiers résultats . . . 73
3.3.3 Bornes inférieures . . . 75
3.3.4 Bornes supérieures . . . 82
3.3.5 Complexité . . . 86
3.4 Grille du roi . . . 95
3.4.1 Méthode . . . 96
3.4.2 Détails de la preuve . . . 101
3.4.3 Preuves des lemmes techniques . . . 108
3.4.4 Conclusion sur la grille du roi . . . 115
3.5 Conclusion du chapitre . . . 116

4 Modifier les voisinages des sommets **117**
4.1 Codes identifiants évolutifs . . . 118
4.1.1 Définitions et premières remarques . . . 118
4.1.2 Sans mémoire dans les cycles . . . 123
4.1.3 Avec mémoire dans les cycles . . . 128
4.1.4 Limitons le nombre de rayons par sommet . . . 131
4.2 Codes identifiants tolérants . . . 134
4.2.1 Cadre : le plan euclidien . . . 134
4.2.2 Codes identifiants tolérants . . . 135
4.2.3 Premier exemple . . . 137
4.2.4 Existence d'un code identifiant tolérant . . . 137
4.2.5 Étude du motif horizontal . . . 139
4.2.6 Bornes générales . . . 141
4.2.7 Amélioration des bornes pour des valeurs fixées . . . 144
4.2.8 Étude des petites valeurs de r . . . 154
4.2.9 Conclusion sur les codes identifiants tolérants . . . 159
4.3 Conclusions et perspectives du chapitre . . . 160

5 Coloration identifiante **163**
5.1 Coloration globalement identifiante . . . 164
5.1.1 Comment identifier avec des couleurs ? . . . 164
5.1.2 Application pratique . . . 165
5.1.3 Lien avec d'autres colorations . . . 165

5.1.4 Lien avec les codes identifiants 165
5.1.5 Bornes générales 167
5.1.6 Cycles et lien avec les cycles universels 171
5.1.7 Conclusion sur les colorations identifiantes globales 174
5.2 Du global au local 174
5.3 Bornes générales 176
5.3.1 En fonction du nombre de sommets 176
5.3.2 En fonction du nombre chromatique 177
5.3.3 En fonction du degré maximum 180
5.4 Étude dans les graphes bipartis et complexité 186
5.4.1 Motivations 186
5.4.2 Coloration en quatre couleurs des graphes bipartis 187
5.4.3 Lien avec la coloration d'hypergraphe et complexité 188
5.4.4 Cas des arbres 190
5.4.5 Produit cartésien de graphes bipartis 191
5.5 Étude dans les graphes parfaits 191
5.5.1 k-arbres 192
5.5.2 Graphes d'intervalles 193
5.5.3 Graphes scindés 195
5.5.4 Cographes 197
5.5.5 Bilan sur les graphes parfaits et conjecture 199
5.6 Étude dans les graphes planaires 201
5.6.1 Graphes planaires extérieurs 201
5.6.2 Graphes planaires de grande maille 204
5.7 Conclusion sur les colorations identifiantes 205

Conclusion **207**

Bibliographie **208**

Introduction

J'ai pour la première fois entendu parler de mon sujet de thèse en 2006 lorsque j'interrogeais Julien Moncel sur ses recherches et qu'il m'a répondu, avec un sourire mystérieux, « je fais des codes ». Je n'osais alors pas en demander beaucoup plus et m'imaginais Julien écrivant des messages codés à longueur de journée. Trois ans plus tard, durant mon stage de master, Sylvain Gravier me propose d'assister aux journées du projet ANR IDEA autour des codes identifiants. Je me plonge alors dans ce problème et m'y retrouve rapidement embarquée. Ce sujet récent mais simple à appréhender me séduit et je décide d'y consacrer ma thèse. J'avais en fait rencontré les codes identifiants bien plus tôt en jouant au jeu « Qui est-ce ? »...

Le jeu « Qui est-ce ? » est un jeu qui se joue à deux. Chaque joueur a devant lui la même série d'une quarantaine de personnages. L'un des joueurs choisit un des personnages sans le dévoiler. L'autre joueur doit deviner quel est ce personnage en posant des questions auxquelles on peut répondre par oui ou non, comme « Ton personnage a-t-il les yeux bleus ? ». Le but du jeu est de trouver le personnage choisi le plus vite possible. Considérons maintenant un tout autre problème. Imaginons que nous soyons gérant d'un musée (ou de tout autre bâtiment comportant des salles) et que nous souhaitions protéger notre bâtiment des incendies. Nous avons à notre disposition des capteurs de fumée. Ces capteurs sont capables de détecter un feu qui se déclenche dans la pièce où ils sont situés ou bien dans une pièce voisine, c'est-à-dire une pièce qui communique avec la pièce où est situé le capteur. Les capteurs doivent être placés de manière stratégique pour que tout incendie soit automatiquement détecté. Il serait de plus souhaitable de pouvoir localiser précisément la pièce où l'incendie a lieu afin d'intervenir efficacement. Quelle que soit la pièce où l'incendie se déclenche, l'ensemble des capteurs s'allumant doit donc permettre de déterminer de manière unique le lieu de l'incendie. Les capteurs étant très chers, nous voulons en poser le moins possible.

Quel est le lien entre ces deux problèmes ? Dans les deux cas, des éléments doivent être identifiés : les personnages ou les pièces du musée. Nous appellerons ces éléments les *sujets*. De plus, chaque sujet a des *caractéristiques* qui permettent de l'identifier parmi les autres sujets. Ces caractéristiques sont évidentes pour les personnages (yeux bleus, cheveux roux, moustache,...). Pour une pièce du musée, ce sont ses pièces voisines qui permettent de l'identifier parmi les autres pièces et qui forment les caractéristiques de la pièce. Dans les deux cas, un des sujets joue un rôle particulier : le personnage choisi ou la pièce où a lieu l'incendie. Enfin, le but du problème est de sélectionner des caractéristiques (en posant des questions ou en plaçant des

capteurs) qui vont permettre d'identifier le sujet particulier. Bien entendu, dans les deux situations, il faut choisir ces caractéristiques pour que l'identification soit la plus efficace possible.

Ces deux problèmes ont aussi des différences de taille. Par exemple, dans le jeu « Qui est-ce ? » les questions sont posées de manière *dynamique*, les unes après les autres. La question posée dépend des réponses obtenues aux questions précédentes. Cela dit, nous pourrions imaginer une version statique du jeu « Qui est-ce ? » où toutes les questions sont posées en même temps et le but est de retrouver le personnage avec toutes les réponses obtenues. Le jeu perdrait beaucoup de son intérêt et ressemblerait plus au travail du médecin qui doit retrouver la maladie d'un patient à partir de tous ses symptômes. Ce problème rentre lui aussi dans notre cadre.

La différence qui nous intéresse plus ici concerne la relation entre les caractéristiques et les sujets. Dans le jeu « Qui est-ce ? », les caractéristiques peuvent être attribuées aux sujets de toutes les manières possibles (ou presque). Nous pourrions créer un jeu en choisissant des sujets, des caractéristiques et en attribuant les caractéristiques de manière aléatoire. Dans le problème du musée, les caractéristiques sont fixées par la disposition du musée. Pour étudier le problème, il est nécessaire de connaître cette disposition.

Dans cette thèse, nous nous intéressons particulièrement au problème du musée. Pour étudier ce problème, nous avons seulement besoin de connaître les pièces et quelles pièces sont voisines. Pour représenter un musée, nous utilisons un *graphe* : les pièces sont *les sommets* du graphe et deux pièces voisines sont représentées par une *arête* qui relie les deux sommets correspondants. Nous pouvons maintenant oublier le reste du musée et ne garder que le graphe. Le problème du musée revient à sélectionner un sous-ensemble de sommets du graphe (les pièces où sont posés les capteurs) de telle sorte que pour chaque sommet, l'ensemble des sommets proches parmi les sommets sélectionnés, soit non vide et unique. Un tel sous-ensemble porte le nom de *code identifiant*. Les codes identifiants permettent d'identifier les sommets du graphe avec un sous-ensemble de sommets. Ils ont été introduits en 1998 pour modéliser les détections de pannes dans des réseaux d'ordinateurs. Les codes identifiants ne sont pas la seule manière d'identifier les sommets d'un graphe et nous nous intéresserons aussi à d'autres problèmes du même type, parfois antérieurs aux codes identifiants. Les codes identifiants font aussi partie de la grande famille des problèmes de *domination* : un ensemble dominant correspondant au cas où l'on souhaite simplement détecter l'incendie mais sans le localiser précisément. Après avoir rappelé et fixé les notations utilisées, nous présentons dans le **chapitre 1** les codes identifiants en détail ainsi que d'autres problèmes étudiés dans cette thèse. Nous donnons quelques résultats essentiels sur les codes identifiants que nous utiliserons par la suite.

La question centrale que l'on se pose autour des codes identifiants est la suivante : étant donné un graphe, quelle est la taille minimum d'un code identifiant de ce graphe ? Plusieurs pistes se dessinent autour de cette question. On peut par exemple essayer de trouver une procédure automatique qui, lorsqu'on lui donne un graphe

quelconque, retourne un code identifiant optimal. Malheureusement, il n'existe pas de telle procédure efficace : trouver un code identifiant de taille minimale est un problème $\mathcal{NP}$-difficile, autrement dit, difficile pour un ordinateur. Qu'à cela ne tienne, au lieu de chercher la valeur *exacte* de la taille minimale d'un code identifiant, nous pouvons chercher des bornes sur cette taille. Si certains graphes sont très bons pour les codes identifiants, c'est-à-dire que leur code identifiant minimal est petit (logarithmique) par rapport au nombre de sommets, d'autres sont très mauvais et nécessitent presque tous les sommets pour être identifiés. Dans le **chapitre 2**, nous classifions tous les graphes dont le plus petit code identifiant est l'ensemble des sommets sauf un, ce qui correspond au cas le plus défavorable. Nous caractérisons aussi les graphes infinis et les graphes orientés dont le seul code identifiant est l'ensemble de tous les sommets.

Étudier le problème des codes identifiants dans des graphes généraux n'est donc pas un problème facile et les résultats obtenus, comme les bornes sur la taille optimale des codes identitifants, ne peuvent être très précis. Aussi, au **chapitre 3**, nous restreignons la classe des graphes étudiés. Beaucoup de classes de graphes ont déjà été considérées pour le problème des codes identifiants. Nous nous intéressons ici à quatre classes de graphes. La première est celle des graphes de Sierpiński. Ces graphes ont une structure particulièrement efficace pour étudier les problèmes d'identification : nous sommes capables de déterminer la taille exacte d'un code identifiant optimal ainsi que celle de deux autres variantes des codes identifiants. La deuxième classe étudiée est celle des graphes d'intervalles. Nous ne savons pas calculer exactement les tailles optimales des codes identifiants dans cette classe mais nous pouvons donner des bornes plus restreintes que dans le cas général. Pour la troisième classe, celle des graphes adjoints, le problème des codes identifiants reste $\mathcal{NP}$-difficile, mais nous pouvons là encore donner des bornes plus restreintes. Étudier ce problème revient à identifier les arêtes du graphe avec les arêtes qui lui sont adjacentes. Enfin, la quatrième classe est en fait un graphe unique : la grille du roi. Nous étudions des codes identifiants permettant d'identifier aussi des paires de sommets et améliorons la borne inférieure connue sur la densité d'un tel code.

Jusqu'à présent, les sommets du code identifiant, c'est-à-dire les capteurs, ne sont capables de détecter qu'un ensemble de sommets figés : ceux qui sont dans leur voisinage proche, c'est-à-dire leur voisinage à distance 1. Nous introduisons, dans le **chapitre 4**, de la flexibilité pour les capteurs, et ce, de deux manières différentes. La première possibilité étudiée est de considérer une identification des sommets en plusieurs étapes : les capteurs commencent par tester leur voisinage à distance 1 puis à chaque étape, ils testent un voisinage un peu plus grand, jusqu'à atteindre un rayon limite. Nous donnons les tailles optimales de tels codes pour la famille des cycles en considérant deux cas selon que le procédé ait une mémoire ou non.

La seconde possibilité considérée est de tolérer une marge d'erreur sur les capteurs. Pour cela, nous nous plaçons dans le plan et considérons des capteurs qui détectent les erreurs (incendies) jusqu'à un rayon dont nous connaissons la valeur avec une petite incertitude. Nous exhibons des constructions de codes permettant cette identification pour différentes valeurs du rayon et de son incertitude et donnons

des bornes inférieures sur la taille de ces codes et leur valeur asymptotique.

Dans le dernier chapitre, le **chapitre 5**, nous cherchons à identifier les sommets du graphe en utilisant des couleurs. Pour cela, nous introduisons le concept de *colorations identifiantes*. Les sommets du graphe sont coloriés et chaque sommet est identifié par les couleurs présentes dans son voisinage. Si l'on reprend l'exemple du médecin qui cherche à déterminer la maladie de son patient, cela revient à poser plusieurs questions en même temps (correspondant aux sommets (questions) ayant la même couleur). Nous montrons des liens avec le problème des codes identifiants ainsi qu'avec un problème de théorie des nombres, les cycles universels. Nous étudions particulièrement les *colorations localement identifiantes*. Ces colorations se comparent aux colorations *propres* d'un graphe (deux sommets adjacents ont des couleurs différentes). Pour plusieurs classes de graphes parfaits, graphes considérés comme « bons » pour la coloration propre, le nombre de couleurs nécessaire dans une coloration localement identifiante est au plus deux fois le nombre de couleurs nécessaire dans une coloration propre.

La plupart des résultats présentés dans cette thèse sont le fruit de collaborations, en grande partie dans le cadre du projet de l'ANR IDEA. Les co-auteurs et références des articles écrits sont précisés en début de chapitre ou partie.

Chapitre 1

Préliminaires

Sommaire

1.1 Graphes **14**
1.1.1 Graphe, voisinage, degré 14
1.1.2 Opérations sur les graphes 15
1.1.3 Exemples et classes de graphes 16
1.1.4 Coloration de graphes et graphes parfaits 19
1.1.5 Hypergraphes 19
1.1.6 Graphe orienté 19
1.1.7 Représentation à l'aide de graphes bipartis 20
1.2 Identification **22**
1.2.1 Codes identifiants dans les graphes non orientés 22
1.2.2 Codes discriminants dans les graphes bipartis d'incidence 24
1.2.3 Variantes 26
1.3 Premiers résultats **28**
1.3.1 Existence 28
1.3.2 Hérédité 28
1.3.3 Bornes 28
1.3.4 Complexité 30

1.1 Graphes

Nous donnons dans cette partie les définitions de base sur les graphes qui nous seront utiles tout au long de cette thèse. Les définitions plus spécifiques seront données au fur et à mesure. Pour en savoir plus, nous renvoyons aux livres de Berge [7] et Diestel [34] [1].

1.1.1 Graphe, voisinage, degré

Un *graphe simple non orienté* G est la donnée d'un couple (V, E) où E est un sous-ensemble de paires (non ordonnées) de V. Les éléments de V sont appelés les *sommets* et les éléments de E les *arêtes*. Lorsque les ensembles V et E ne sont pas précisés, $V(G)$ (resp. $E(G)$) désigne l'ensemble des sommets (resp. les arêtes) de G. Pour désigner l'arête $\{x, y\}$, nous écrirons indifféremment xy ou yx. Une arête xx est appelée *boucle*. Sauf mention contraire, les graphes considérés ici seront sans boucles. L'*ordre* d'un graphe G est son nombre de sommets et sera souvent désigné par la lettre n.

Deux sommets u et v sont dits *adjacents* ou *voisins* si uv est une arête. Soit $e = uv$ une arête. Les sommets u et v sont appelées *extrémités* de e et l'arête e est *incidente* à u et v. Deux arêtes sont *adjacentes* si elles ont une extrémité commune. Le *voisinage ouvert* d'un sommet u est l'ensemble de ses voisins : $N(u) = \{v \in V \mid uv \in E\}$. Le *voisinage fermé* inclus de plus le sommet u : $N[u] = N(u) \cup \{u\}$. Ces notions s'étendent à des sous-ensembles de sommets U : $N(U) = \bigcup_{u \in U} N(u)$ et $N[U] = N(U) \cup U$. Un sommet sans voisins est dit *isolé*. Un sommet avec un seul voisin est un sommet *pendant*. Un sommet voisin de tous les autres sommets est un sommet *universel*. Le *degré* du sommet u dans le graphe G, noté $d_G(u)$ ou $d(u)$ si le contexte est clair, est le nombre de voisins de u. Un graphe est *k-régulier* lorsque tous ses sommets sont de degré k. Le degré maximum d'un graphe G est noté $\Delta(G)$ ou Δ lorsque le contexte est clair.

Un *chemin* entre deux sommets u et v est une suite de sommets $u_0 = u, u_1, ..., u_k = v$ telle que $u_i u_{i+1}$ soit une arête pour tout $i \in \{0, ..., k-1\}$. Un graphe non orienté est dit *connexe* s'il existe un chemin entre toute paire de sommets. La *distance* entre deux sommets u et v de G, notée $d_G(u, v)$ ou $d(u, v)$ si le contexte est clair, est le nombre minimum d'arêtes sur un chemin entre u et v. Étant donné un entier r, la *boule de rayon r centrée en u*, notée $\mathcal{B}_r(u)$, désigne les sommets à distance au plus r de u :

$$\mathcal{B}_r(u) = \{v \in V \mid d(u, v) \leq r\}.$$

1. Notons que la terminologie utilisée dans cette thèse est la terminologie anglo-saxone, à titre d'exemple, les *chaînes* seront appelées *chemins*.

1.1.2 Opérations sur les graphes

Relations entre les graphes

Soient $G_1 = (V_1, E_1)$ et $G_2 = (V_2, E_2)$ deux graphes. Une fonction $\Phi : V(G_1) \to V(G_2)$ est un *morphisme* entre G_1 et G_2 si pour chaque arête xy de G_1, $\Phi(x)\Phi(y)$ est une arête de G_2. Deux graphes G_1 et G_2 sont *isomorphes* s'il existe un morphisme bijectif entre G_1 et G_2 dont la réciproque est un morphisme. On note alors $G_1 \simeq G_2$. Le morphisme Φ est appelé *isomorphisme* ou bien *automorphisme* dans le cas spécial où $G_1 = G_2$. Nous ne distinguerons en général pas deux graphes isomorphes et écrirons souvent $G_1 = G_2$ à la place de $G_1 \simeq G_2$.

Si $V_1 \subseteq V_2$ et $E_1 \subseteq E_2$, G_1 est un *sous-graphe* de G_2. Si de plus E_1 est exactement l'ensemble des arêtes de E_2 dont les extrémités sont dans V_1, alors G_1 est le sous-graphe de G_2 *induit par les sommets* V_1 et l'on note $G_1 = G_2[V_1]$. Un graphe G est dit *sans* G' (resp. *sans* G' *induit*) s'il n'y aucun sous-graphe (resp. sous-graphe induit) de G qui est isomorphe à G'.

Complémentaire

Le *complémentaire* d'un graphe G, noté $\overline{G}$, est le graphe dont les sommets sont ceux de G et où xy est une arête si et seulement si xy n'est pas une arête de G : $\overline{G} = (V(G), \overline{E(G)})$

Union, joint, produit cartésien

Soient $G_1 = (V_1, E_1)$ et $G_2 = (V_2, E_2)$ deux graphes sur des ensembles disjoints de sommets.

L'*union disjointe* de G_1 et G_2, notée $G_1 \sqcup G_2$, est le graphe ayant pour sommets l'ensemble $V_1 \cup V_2$ et pour arêtes l'ensemble $E_1 \cup E_2$.

Le *joint* de G_1 et G_2, noté $G_1 \bowtie G_2$ est le graphe ayant pour sommets l'ensemble $V_1 \cup V_2$ et pour arêtes l'ensemble $E_1 \cup E_2 \cup \{uv \mid u \in V_1, v \in V_2\}$ (voir la figure 1.1a).

Le *produit cartésien* de G_1 et G_2, noté $G_1 \square G_2$, est le graphe ayant pour sommets l'ensemble $V_1 \times V_2$ et où deux sommets (u_1, u_2) et (v_1, v_2) sont adjacents si $u_1 = v_1$ et $u_2v_2 \in E_2$ ou bien si $u_2 = v_2$ et $u_1v_1 \in E_1$ (voir la figure 1.1b).

(a) *Joint* **(b)** *Produit cartésien*

Figure 1.1 – *Joint et produit cartésien de deux graphes.*

Supprimer un ou plusieurs sommets

Soit S un sous-ensemble de sommets d'un graphe $G = (V, E)$. Le graphe $G - S$ désigne le sous-graphe de G induit par les sommets de l'ensemble $V \setminus S$: ses sommets sont $V \setminus S$ et ses arêtes sont celles de G ne contenant pas de sommets de S. Lorsque S n'est composé que d'un seul élément x, nous noterons plutôt $G - x$ le graphe $G - \{x\}$.

Puissance de graphe

Étant donné un entier t, la *puissance t-ième* d'un graphe $G = (V, E)$, notée G^t, est le graphe dont les sommets sont l'ensemble V et où deux sommets sont adjacents dans G^t s'ils sont à distance au plus t dans G :

$$E(G^t) = \{uv \in E(G) \mid d_G(u, v) \leq t\}.$$

Graphe adjoint

Le *graphe adjoint* (*line graph* en anglais) d'un graphe G, noté $\mathcal{L}(G)$, est le graphe dont les sommets sont les arêtes de G et où deux sommets de $\mathcal{L}(G)$ sont adjacents si les arêtes correspondantes sont adjacentes (voir la figure 1.2).

K_4 $\mathcal{L}(K_4)$

Figure 1.2 – *Le graphe K_4 et son graphe adjoint $\mathcal{L}(K_4)$.*

1.1.3 Exemples et classes de graphes

Nous donnons ici quelques graphes classiques et classes de graphes que nous utiliserons au cours de cette thèse.

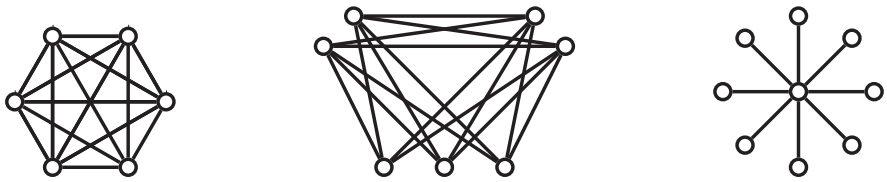
Graphes complets et stables

Le graphe *complet* à n sommets, noté K_n, a pour ensemble de sommets l'ensemble $\{v_1, \ldots, v_n\}$ et comme ensemble d'arêtes toutes les arêtes possibles. Le complémentaire du graphe complet $\overline{K_n}$ est appelé *stable* ou *indépendant*. Un ensemble de sommets d'un graphe G qui induit un sous-graphe complet (resp. stable) est appelé *clique* (resp. *stable*). Une clique de taille k est appelée *k-clique*. La taille de la plus grande clique de G est notée $\omega(G)$.

Un graphe *multiparti complet* défini à partir d'une suite d'entiers $n_1, ..., n_k$, noté $K_{n_1,...,n_k}$, est le graphe dont les sommets sont partitionnés en k parties : $V = V_1 \cup V_2 \cup$

$...\cup V_k$ avec V_i de taille n_i et où chaque sommet est adjacent à tous les sommets hors de sa partie. On peut aussi le voir comme le joint de k stables de taille respectives $n_1,...,n_k$.

Nous utiliserons en particulier l'*étoile* à n branches qui est le graphe $K_{1,n}$.



(a) K_6 **(b)** $K_{3,2,2}$ **(c)** $K_{1,8}$

Figure 1.3 – *Quelques graphes complets ou multiparti complets*

Couplage

Un *couplage* d'un graphe G est un sous-ensemble d'arêtes deux à deux non adjacentes. Un couplage est dit *parfait* si tout sommet du graphe est une extrémité d'une arête du couplage.

Chemins et cycles

Le *chemin* (élémentaire) à n sommets, noté P_n, a pour ensemble de sommets l'ensemble $\{v_1, \ldots, v_n\}$ et ses arêtes sont les paires v_iv_{i+1} avec $1 \leq i < n$. Les sommets v_1 et v_n sont appelées *extrémités* du chemin. Le *chemin infini* P_∞ a pour ensemble de sommets les entiers relatifs $\mathbb{Z}$ et deux sommets x et y sont adjacents si $|x - y| = 1$.

Le *cycle* sur n sommets, noté $\mathcal{C}_n$, a pour ensemble de sommets $\{v_1, \ldots, v_n\}$ et ses arêtes sont les paires v_iv_{i+1} avec $1 \leq i < n$ ainsi que l'arête v_nv_1.

La *longueur* d'un chemin fini ou d'un cycle est son nombre d'arêtes (ainsi le chemin P_n est de longueur $n - 1$).

La *maille* d'un graphe G est la taille du plus petit cycle contenu dans G.

Arbres et forêts

Un graphe sans cycle est une *forêt*. Si le graphe est de plus connexe, c'est un *arbre*. Un arbre à n sommets possède exactement $n - 1$ arêtes. Les sommets de degré 1 sont appelés les *feuilles*. Un arbre a toujours au moins deux feuilles et s'il en a exactement deux, c'est un chemin. Pour toute paire de sommets x et y d'un arbre, il existe un unique chemin reliant x à y.

Parfois un des sommets joue le rôle particulier de *racine* et l'on parle alors d'*arbre enraciné*. Chaque sommet x autre que la racine à un unique *père* (ou *prédécesseur*) qui est le premier sommet sur le chemin reliant x à la racine. Les autres voisins de x sont ses *fils*. Les sommets d'un arbre sont souvent appelés des *nœuds*.

Hypercube

L'*hypercube de dimension* n, noté $\mathcal{H}_n$, a pour ensemble de sommets $\{0,1\}^n$ et deux sommets sont adjacents s'ils diffèrent sur exactement une coordonnée. L'hypercube de dimension n peut être construit en effectuant des produits cartésiens par K_2 successifs : $\mathcal{H}_1 = K_2$ et $\mathcal{H}_n = \mathcal{H}_{n-1} \square K_2$.

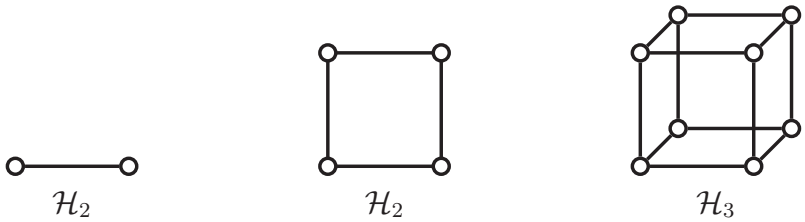
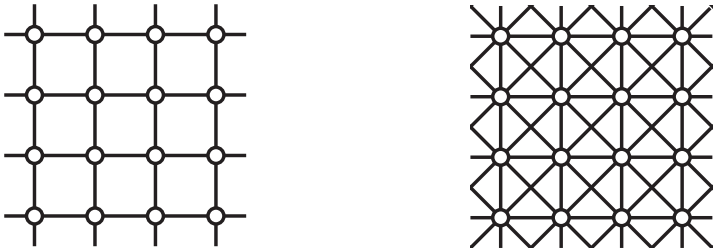


Figure 1.4 – *Les hypercubes de dimension 1, 2 et 3.*

Grilles

Une *grille finie* est le produit cartésien de deux chemins finis. La *grille infinie carrée*, ou simplement *grille infinie*, est le produit cartésien de deux chemins infinis. La *grille infinie du roi* a pour sommets $\mathbb{Z}^2$ et deux sommets (x, y) et (x', y') sont adjacents si $|x - x'| \leq 1$ et $|y - y'| \leq 1$. Autrement dit, le voisinage d'un sommet correspond aux déplacements possibles d'un roi sur un échiquier. La figure 1.5 montre des portions de grille infinie carrée et de grille infinie du roi.



(a) *Grille carrée* **(b)** *Grille du roi*

Figure 1.5 – *Portions de grille infinie carrée et de grille infinie du roi.*

Graphes d'intervalles

À partir d'une famille finie d'intervalles $\mathcal{I} = \{I_1, ..., I_n\}$ sur la droite réelle, le graphe d'*intersection* de la famille $\mathcal{I}$ est le graphe dont les sommets sont les intervalles et deux sommets sont adjacents lorsque les intervalles correspondants s'intersectent (voir la figure 1.6). Un graphe est un *graphe d'intervalles* s'il existe une famille d'intervalles dont c'est le graphe d'intersection. Les intervalles peuvent être ouverts ou fermés. Il est possible de passer d'une représentation avec des intervalles ouverts à une représentation avec des intervalles fermés, en gardant le même graphe d'intersection, et réciproquement.

Figure 1.6 – *Une famille d'intervalles et son graphe d'intersection.*

Graphes bipartis

Un graphe G est *biparti* si l'ensemble de ses sommets peut être partionné en deux ensembles stables S, T. Par abus de notations, nous écrirons souvent $G = (S \cup T, E)$ pour indiquer les deux ensembles stables du graphe biparti. Nous utiliserons surtout les graphes bipartis pour représenter des structures d'incidence (voir la partie 1.1.7).

1.1.4 Coloration de graphes et graphes parfaits

Une *coloration* des sommets d'un graphe G est une fonction $c : V(G) \to \mathbb{N}$. Pour un sommet u, $c(u)$ est la *couleur* de u. Une *k-coloration* est une coloration n'utilisant que k couleurs. Une coloration est *propre* si pour toute arête uv, les couleurs attribuées à u et v sont différentes. Le *nombre chromatique* d'un graphe G, noté $\chi(G)$, est le plus petit nombre de couleurs nécessaire dans une k-coloration propre de G.

Dans une coloration propre de G, tous les sommets d'une même clique sont de couleurs différentes. Par conséquent :

$$\chi(G) \geq \omega(G).$$

La réciproque n'est pas vraie : il existe des graphes G sans triangle ($\omega(G) = 2$) dont le nombre chromatique est arbitrairement grand. Un graphe G est *parfait* si pour tout sous-graphe induit H de G, $\omega(H) = \chi(H)$. Parmi les graphes définis précédemment, les graphes bipartis, les graphes d'intervalles et les arbres sont des graphes parfaits.

1.1.5 Hypergraphes

Un *hypergraphe* est la donnée d'un couple $(V, \mathcal{E})$ où V est un ensemble et $\mathcal{E}$ est un sous-ensemble de l'ensemble des parties de V (voir la figure 1.7 page 21). Les éléments de V sont les *sommets* et les éléments de $\mathcal{E}$ sont les *hyperarêtes* de l'hypergraphe : ce sont des sous-ensembles de sommets.

Un hypergraphe est *r-régulier* si tout sommet est dans exactement r hyperarêtes et *k-uniforme* si toutes ses arêtes sont de taille k. La notion d'hypergraphe généralise la notion de graphe : les graphes sont exactement les hypergraphes 2-uniformes.

1.1.6 Graphe orienté

Un *graphe orienté* D est la donnée d'un couple (V, A) où V est un ensemble et A est un sous-ensemble de paires *ordonnées* de V. Les éléments de V sont toujours

appelés les *sommets* et ceux de A sont les *arcs*. L'arc (x, y) sera noté $\overrightarrow{xy}$. Le *graphe sous-jacent* d'un graphe orienté est le graphe obtenu en supprimant les orientations des arêtes (il y a une arête xy dans le graphe sous-jacent si $\overrightarrow{xy}$ ou $\overrightarrow{yx}$ sont des arcs dans le graphe de départ). Deux arcs $\overrightarrow{xy}$ et $\overrightarrow{yx}$ sont des *arcs symétriques*.

Si $\overrightarrow{xy}$ est un arc de D, x est alors un *voisin entrant* de y et y est un *voisin sortant de* x. L'ensemble des voisins entrants (resp. sortants) de x est noté $N^+(x)$ (resp. $N^-(x)$), et son cardinal est le *degré entrant* (resp. *degré sortant*) de x, noté $d^+(x)$ (resp. $d^-(x)$). On note $\mathcal{B}_1^+(x)$ (resp. $\mathcal{B}_1^-(x)$) l'ensemble $N^+(x) \cup \{x\}$ (resp. $N^-(x) \cup \{x\}$).

Un sommet sans voisin sortant est appelé *puits* et un sommet sans voisin entrant est une *source*.

Un *chemin orienté* dans un graphe orienté est une suite de sommets $x_1, ..., x_n$ telle que $\overrightarrow{x_i x_{i+1}}$ soit un arc pour tout $1 \leq i < n$. Si $\overrightarrow{x_n x_1}$ est aussi un arc alors nous avons un *cycle orienté*. Un *arbre enraciné orienté* est un graphe orienté dont le graphe sous-jacent est un arbre, avec un sommet source unique, *la racine*, et tel que tous les arcs soient orientés depuis la racine, de sorte que pour chaque sommet x il y ait un chemin orienté de la racine vers le sommet x. Les terminologies sont alors les mêmes que pour les arbres enracinés dans le cas non-orienté.

La *fermeture transitive* d'un graphe orienté D est obtenue en ajoutant au graphe D un arc $\overrightarrow{xy}$ s'il existe un chemin orienté partant de x et allant à y. Étant donnés deux graphes orientés D_1 et D_2 sur des ensembles de sommets disjoints, l'*union disjointe* de D_1 et D_2, notée $D_1 \sqcup D_2$, dénote le graphe orienté dont l'ensemble des sommets est $V(D_1) \cup V(D_2)$ et dont l'ensemble des arcs est $A(D_1) \cup A(D_2)$. Le *joint orienté* de D_1 et D_2, noté $D_1 \overrightarrow{\triangleleft} D_2$, est le graphe orienté dont les sommets sont $V(D_1) \cup V(D_2)$ et dont les arcs sont $A(D_1) \cup A(D_2) \cup \{\overrightarrow{uv}, u \in D_1, v \in D_2\}$. Si D_1 est un seul sommet x, nous utiliserons la notation $x \overrightarrow{\triangleleft} D_2$ à la place de $D_1 \overrightarrow{\triangleleft} D_2$.

Nous utiliserons les graphes orientés principalement dans le chapitre 2. Aussi, lorsque cela n'est pas précisé, les graphes considérés seront non orientés.

1.1.7 Représentation à l'aide de graphes bipartis

Les graphes bipartis permettent de représenter des structures d'incidence comme les hypergraphes ou les graphes (orientés ou non orientés) et donne ainsi un cadre commun. Nous aurons souvent recours à cette représentation au cours de cette thèse.

Graphe biparti d'incidence d'un hypergraphe

Étant donné un hypergraphe $\mathcal{H} = (V, \mathcal{E})$, le *graphe biparti d'incidence de* $\mathcal{H}$, noté $B(\mathcal{H})$ est défini de la manière suivante (voir la figure 1.7 pour un exemple) :

- Il y a un sommet pour chaque sommet de l'hypergraphe, l'ensemble de ces sommets est noté S.
- Il y a un sommet pour chaque hyperarête de l'hypergraphe, l'ensemble de ces sommets est noté T.
- Il y a une arête st, $s \in S$, $t \in T$ si s est un sommet de l'hyperarête t.

Notons que tout graphe biparti $B = (S \cup T, E)$ est le graphe d'incidence d'un hypergraphe et en fait de deux hypergraphes qui sont duaux l'un de l'autre, dépendant

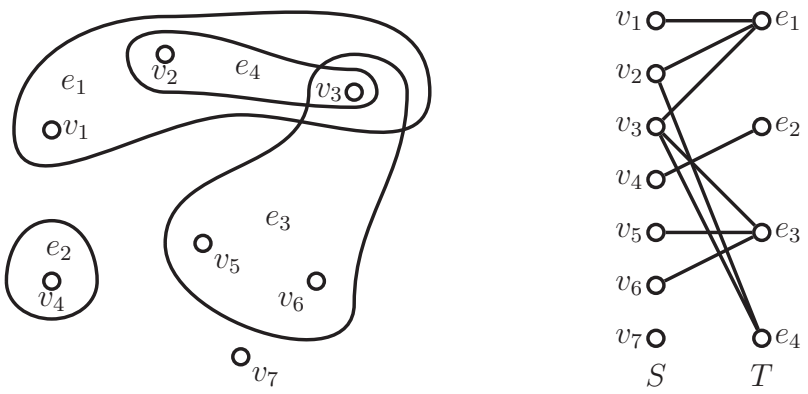


Figure 1.7 – *Exemple d'hypergraphe et de son graphe biparti d'incidence.*

du choix des rôles pour S et T. Nous supposerons en général que la partie S correspond aux sommets et la partie T aux hyperarêtes.

Graphe biparti d'incidence d'un graphe orienté

Étant donné un graphe orienté $D = (V, A)$, le *graphe biparti d'incidence de* D, noté $B(D)$ est défini de la manière suivante (voir la figure 1.8 pour un exemple) :

- À chaque sommet du graphe D correspond deux sommets dans $B(G)$, l'un dans S et l'un dans T.
- Il y a une arête st dans $B(D)$, $s \in S$, $t \in T$, si $\overrightarrow{ts}$ est un arc dans D ou bien si s et t correspondent au même sommet.

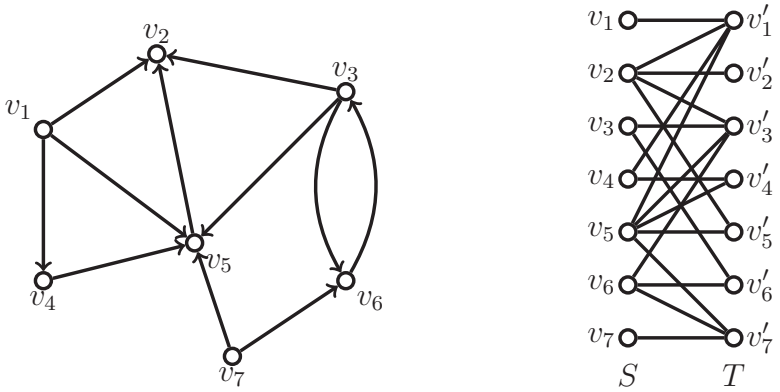


Figure 1.8 – *Exemple de graphe orienté et de son graphe biparti d'incidence.*

Cette définition n'est pas tout à fait conventionnelle, il n'y a en général pas d'arête ajoutée entre les deux copies d'un même sommet. Nous utiliserons néanmoins cette définition, de la même manière que par *voisinage* nous entendrons généralement *voisinage fermé*.

Pour qu'un graphe biparti $B = (S \cup T, E)$ soit le graphe biparti d'incidence d'un graphe orienté, il est nécessaire et suffisant que le graphe B possède un couplage

parfait. Cette dernière condition implique entre autre que les deux parties S et T de B doivent avoir le même cardinal. Pour chaque couplage parfait de B, il existe deux graphes orientés dont B est le graphe d'incidence, dépendant du rôle joué par S et T. Ces deux graphes sont obtenus l'un à partir de l'autre en changeant le sens des arcs. Nous supposerons implicitement que les rôles des parties S et T sont ceux joués dans la définition que l'on a donné ici.

Graphe biparti d'incidence des voisinages fermés d'un graphe non orienté

Étant donné un graphe $G = (V, E)$, le *graphe biparti d'incidence des voisinages fermés de* G, noté $B(G)$ est défini ainsi (voir la figure 1.9) :

- À chaque sommet du graphe G correspond deux sommets dans $B(G)$, l'un dans S et l'un dans T.
- Il y a une arête st dans $B(G)$, $s \in S$, $t \in T$ si le sommet correspondant à s est dans le voisinage fermé de t.

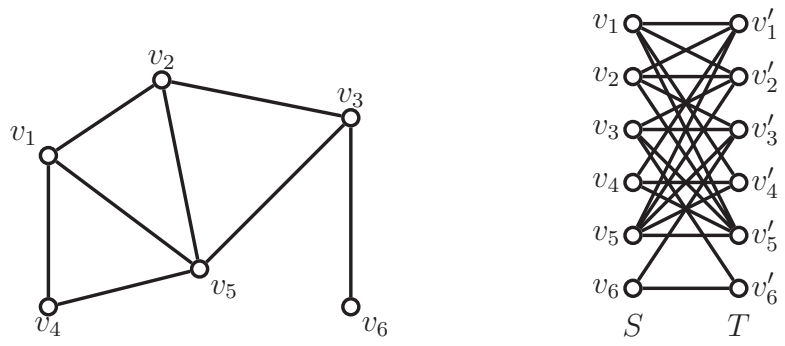


Figure 1.9 – *Exemple de graphe non orienté et de son graphe biparti d'incidence des voisinages fermés.*

Tout graphe biparti B n'est pas le graphe biparti d'incidence des voisinages fermés d'un graphe non orienté. Il est nécessaire que B contienne un couplage parfait mais il faut de plus que le graphe soit symétrique par rapport à ce couplage.

1.2 Identification

Nous définissons dans cette partie quelques problèmes d'identification : ceux que nous étudierons. Pour une liste plus exhaustive, nous invitons le lecteur à consulter la bibliographie en ligne maintenue par Antoine Lobstein [72]. Le problème central de cette thèse est celui des codes identifiants.

1.2.1 Codes identifiants dans les graphes non orientés

Un *code* désigne simplement un sous-ensemble de sommets du graphe.

Ensemble dominant

Un code C d'un graphe G *domine* un sous-ensemble de sommets V' de G si tout sommet de V' est à distance au plus 1 d'un sommet de C. Autrement dit, pour tout sommet v dans V', $N[v] \cap C \neq \emptyset$.

Souvent, l'ensemble V' est l'ensemble de tous les sommets du graphe. Le code C est alors appelé *ensemble dominant* (pour les théoriciens des graphes) ou *code couvrant* (pour les théoriciens des codes). Lorsque le graphe est fini, $\gamma(G)$ désigne la taille du plus petit ensemble dominant de G. La domination est un thème largement étudié en théorie des graphes. Pour en savoir plus, nous renvoyons le lecteur au livre de Haynes, Hedetniemi et Slater [52].

Ensemble séparant

Un sommet x *sépare* deux sommets u et v de G s'il est dans la différence symétrique de leur voisinage fermé : $x \in N[x] \Delta N[y]$.

Un code C *sépare* un sous-ensemble de sommets V' si toutes les paires de sommets dans V' sont séparées par un élément de C. Autrement dit, tous les ensembles $N[v] \cap C$ pour $v \in V'$ sont différents.

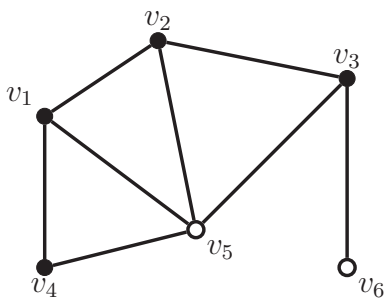
Si $V' = V$, C est simplement un *ensemble séparant* de G. La taille du plus petit ensemble séparant de G, lorsqu'il existe, est noté $\gamma^{\text{S}}(G)$.

Codes identifiants

Un code C *identifie* un sous-ensemble de sommets V' s'il domine et sépare V'. Autrement dit, tous les ensembles $N[v] \cap C$ pour $v \in V'$ sont uniques et non vides.

Lorsque $V' = V$, C est appelé *code identifiant*. Chaque sommet v du graphe est alors identifié par l'ensemble unique et non vide $I(v, C) = N[v] \cap C$, appelé *identifiant* ou *étiquette* du sommet v.

La figure 1.10 donne un exemple de code identifiant dans un graphe. Le tableau sur la droite de la figure indique les identifiants de chaque sommet : toutes les lignes sont différentes et non vides.



$V \backslash C$	v_1	v_2	v_3	v_4
v_1	●	●	-	●
v_2	●	●	●	-
v_3	-	●	●	-
v_4	●	-	-	●
v_5	●	●	●	●
v_6	-	-	●	-

Figure 1.10 – *Les sommets noirs forment un code identifiant. Le tableau indique les identifiants de chaque sommet.*

Si G possède un code identifiant, la taille du plus petit code identifiant est notée $\gamma^{\text{ID}}(G)$. Un code identifiant de taille $\gamma^{\text{ID}}(G)$ est dit optimal.

Motivations

Les codes identifiants ont été introduit en 1998 par Karpovsky, Chakrabarty et Levitin dans [64] pour modéliser un problème d'identification de processeurs défectueux dans des réseaux multiprocesseurs.

Les processeurs sont organisés sous forme de graphe et chaque processeur est capable de tester si un processeur dans son voisinage fermé est défectueux, et retourne une information binaire en fonction : 0 s'il n'y a pas de défaillance et 1 s'il y en a une. Le problème est de déterminer un sous-ensemble C de processeurs tel que :

1. Si tous les processeurs de C renvoient l'information 0, cela signifie qu'il n'y a pas de défaillance.
2. Si un des processeurs renvoie l'information 1, il y a un processeur défectueux et on peut le localiser de manière unique.

Si l'on suppose qu'il y a à tout moment au plus un processeur défectueux, alors l'ensemble C recherché correspond exactement à un code identifiant. La première condition, liée à la domination, nous permet de savoir s'il y a un processeur défectueux ou non, tandis que la deuxième nous permet de localiser l'élément lorsqu'il y en a un. Cette application justifie la propriété de domination et permet de considérer le problème des codes identifiants comme faisant partie des problèmes de domination.

D'autres applications sont apparues par la suite comme dans [84] pour localiser des incendies dans des bâtiments (problème présenté dans l'introduction) ou dans [83] pour modéliser un problème de localisation par des réseaux de capteurs. Pour motiver certaines variantes des codes identifiants nous aurons recours à la terminologie de ces applications : par exemple, un *capteur* correspond à un élément du code, qui *détecte* des *erreurs* : les processeurs défectueux.

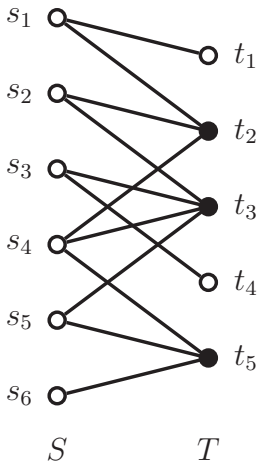
1.2.2 Codes discriminants dans les graphes bipartis d'incidence

Les codes identifiants permettent d'identifier les sommets d'un graphe avec les autres sommets. Des problèmes d'identification plus généraux peuvent être modélisés aussi à l'aide de graphes, notamment de graphes bipartis.

Dans un graphe biparti $B = (S \cup T, E)$, les sommets de S peuvent être vus comme les éléments à identifier (S comme "sujets") et les sommets de T comme les caractéristiques (T comme "traits") qui permettent d'identifier les éléments de S. La structure du graphe donne alors les caractéristiques de chaque sujet et l'on cherche des sous-ensembles de caractéristiques qui permettent d'identifier chaque sujet. Le graphe biparti est ici simplement un outil pour représenter une collection de sous-ensembles.

Nous reprenons la terminologie de Charbit *et al.* [15] en définissant un *code discriminant* dans un graphe biparti comme un sous-ensemble C de sommets de T qui identifie l'ensemble S. Les sous-ensembles $I(s, C) = N[s] \cap C$, $s \in S$, appelés ici aussi identifiants, sont non vides et uniques. Ils permettent donc d'identifier les sommets de S. La figure 1.11 donne un exemple de code discriminant.

La définition d'un code discriminant sous-entend que les deux parties du graphe biparti sont différenciées. Pour un graphe biparti fini B, la taille minimale d'un code



$S \backslash C$	t_2	t_3	t_5
s_1	•	-	-
s_2	•	•	-
s_3	-	•	-
s_4	•	•	•
s_5	-	•	•
s_6	-	-	•

Figure 1.11 – *Les sommets noirs forment un code discriminant. Le tableau indique les identifiants des sommets de S.*

discriminant dans B, si cela existe, est notée $\gamma^{\mathrm{D}}(B)$.

Cette définition généralise celle des codes identifiants : un code identifiant dans un graphe G est exactement un code discriminant dans le graphe biparti d'incidence des voisinages fermés de G (défini en partie 1.1.7). Dans cette thèse, nous étudierons principalement les codes identifiants. Le contexte plus général donné par les codes discriminants nous permet d'avoir des résultats plus globaux, comme nous le verrons dans le chapitre 2.

La notion de *code discriminant* correspond à la notion de *codes identifiants dans les hypergraphes* donnée par Moncel dans sa thèse [76] en considérant l'hypergraphe correspondant au graphe biparti B [2].

Lorsque l'hypothèse de domination est supprimée, le problème se ramène à un *problème de couverture par test*. Dans un problème de couverture par test, nous disposons d'une matrice dont toutes les lignes sont différentes (voir la figure 1.12). Le problème est de déterminer un sous-ensemble de colonnes de la matrice, de taille minimale, pour lequel toutes les lignes résultantes restent différentes. Lorsque la matrice de départ est binaire, cela revient exactement au problème des codes séparants dans un graphe biparti. En effet, chaque ligne de la matrice peut être vue comme un sommet de S, chaque colonne comme un sommet de T et les coefficients de la matrice donne les relations d'adjacence du graphe. L'opération étant bijective, les deux problèmes sont équivalents. Les problèmes de couverture par tests trouvent des applications dans de nombreux domaines comme en médecine pour identifier les symptômes de chaque maladie [68, 80].

2. Dans la thèse de Moncel, cela correspond plutôt aux codes identifiants dans l'hypergraphe dual : les éléments à identifier sont les hyperarêtes tandis que ce sont les sommets qui permettent d'identifier. Un passage au dual permet de se ramener d'une définition à l'autre.

$$\begin{array}{c c} & \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{array}{c} A \\ B \\ C \\ D \\ E \end{array} & \begin{pmatrix} 1 & 0 & 1 & 0 \\ 2 & 1 & 2 & 2 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 2 & 2 \\ 2 & 1 & 2 & 1 \end{pmatrix} \end{array}$$

Figure 1.12 – *Un exemple de problème de couverture par test. Les colonnes* 2 *et* 3 *ne permettent pas de séparer chaque ligne, contrairement aux colonnes* 1 *et* 4.

1.2.3 Variantes

De nombreuses variantes existent autour des codes identifiants. Nous définissons ici celles que nous utiliserons par la suite.

Graphes orientés

Il est possible de définir les codes identifiants dans un graphe orienté. Une manière naturelle est de les définir comme les codes discriminants dans le graphe biparti d'incidence du graphe orienté. Cela donne la définition suivante : un code identifiant dans un graphe orienté D est un sous-ensemble C de sommets tels que les ensembles $I^+(u, C) = B_1^+(u) \cap C$ soient distincts et non vide (la convention choisie étant qu'un sommet c domine un sommet u si $c = u$ ou si c est un sommet entrant de u).

La taille du plus petit code identifiant de D, si D est fini et possède un code identifiant, est notée $\overrightarrow{\gamma^{\text{ID}}}(D)$. La terminologie (ensemble séparant, dominant,...) pour les graphes orientés se déduit naturellement de celle des graphes non orientés. Les codes identifiants dans les graphes orientés ont été considérés dans [18, 20, 86]. Nous étudierons les codes identifiants dans les graphes orientés uniquement au chapitre 2.

Codes r-identifiants

Une manière naturelle d'étendre la notion de code identifiant est de permettre aux capteurs de détecter des erreurs sur des voisinages plus étendus : au lieu de détecter à distance 1, ils peuvent détecter à distance r. Un tel code est appelé *code r-identifiant* et correspond simplement à un code identifiant dans G^r, la puissance r-ième de G. Autrement dit, un *code r-identifiant* d'un graphe G est un sous-ensemble de sommets C tel que les ensembles $I_r(u, C) = \mathcal{B}_r(u) \cap C$ soient tous distincts et non vide. Si G a un code r-identifiant, la taille du plus petit code r-identifiant de G est notée $\gamma_r^{\text{ID}}(G)$.

Cette variante a été introduite dans l'article fondateur [64] et a été largement étudiée. Nous la considérons dans les chapitres 2 et 4.

Codes $(r, \leq \ell)$-identifiants

Les *codes $(r, \leq \ell)$-identifiants* permettent de modéliser des problèmes où il n'y a non pas une seule erreur, mais jusqu'à ℓ erreurs possibles. Il faut alors un en-

semble identifiant pour chaque ensemble X de taille inférieure où égale à ℓ. Soient X un ensemble et C un code, la trace de C sur X est l'ensemble $I_{r,\ell}(X,C) = (\bigcup_{x\in X} \mathcal{B}_r(x)) \cap C = \bigcup_{x\in X} I_r(x,C)$. Un code C est $(r, \leq \ell)$-*identifiant* si tous les ensembles $I_{r,\ell}(X,C)$ sont distincts, avec $|X| \leq \ell$.

Un code $(r, \leq 1)$-identifiant est simplement un code r-identifiant. Les codes $(r, \leq \ell)$-identifiant ont aussi été introduits dans [64]. Nous les étudierons uniquement dans la partie 3.4. Des variantes similaires ont été étudiées : en supposant que les erreurs sont situées parmi des sous-ensembles possibles d'erreurs (Honkala *et al.* [54]) ou encore en autorisant plus de ℓ erreurs avec une réponse particulière pour ce cas (Honkala et Laihonen [53]).

Codes localisateur-dominateurs

Pour certaines applications, il parait raisonnable de supposer que les sommets qui sont dans le code ne sont pas défectueux, il n'est donc pas nécessaire de les identifier. Un sous-ensemble de sommets C est un *code localisateur-dominateur* s'il domine tous les sommets et sépare toutes les paires de sommets qui ne sont pas dans C. La taille du plus petit code localisateur-dominateur d'un graphe G est notée $\gamma^{\text{LD}}(G)$.

Cette variante a été définie avant les codes identifiants par Slater et Rall [88, 89] et a été beaucoup étudiée (voir la bibliographie de Lobstein [72]). Nous la rencontrerons aux chapitres 3 et 4.

Ensembles résolvants

Les *ensembles résolvants* ou *ensembles localisateurs* permettent de modéliser des situations où les capteurs sont capables de déterminer la distance qui les sépare de chaque sommet. Un sommet est alors identifié par ses distances à chaque capteur. Plus formellement, un ensemble résolvant $C = \{c_1, ..., c_k\}$ est un sous-ensemble de sommets du graphe tels que tous les k-uplets $(d(u,c_1), ..., d(u,c_k))$ avec u un sommet du graphe soient distincts. La figure 1.13 donne un exemple d'ensemble résolvant dans le cycle de taille 6. La taille du plus petit ensemble résolvant d'un graphe G est appelé *dimension métrique de* G et notée $\dim(G)$. Un ensemble résolvant de taille minimale est appelé *base métrique*.

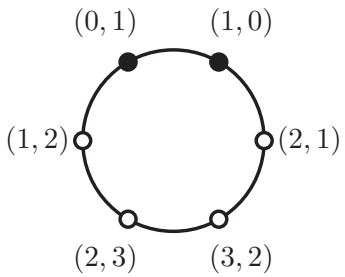


Figure 1.13 – *Les sommets noirs forment un ensemble résolvant. Pour chaque sommet, les distances aux deux sommets noirs sont indiquées.*

Les ensembles résolvants ont été introduits indépendamment par Slater en 1975 [87] et par Harary et Melter en [50]. C'est une des premières modélisations de

problèmes de localisation. Chartrand *et al.* [24] proposent une revue détaillée sur le sujet. Dans le chapitre 4, nous introduirons une notion qui correspond aux ensembles résolvants avec des capteurs ayant un rayon d'action borné.

1.3 Premiers résultats

Nous donnons maintenant quelques résultats essentiels sur les codes identifiants que nous utiliserons dans la suite.

1.3.1 Existence

Il n'existe pas toujours de code identifiant dans un graphe. En effet, si deux sommets ont exactement le même voisinage fermé, alors aucun sommet ne pourra les séparer. Deux sommets u et v tels que $N[u] = N[v]$ sont appelés des sommets *jumeaux* (voir la figure 1.14). C'est la seule restriction puisque si un graphe n'a pas de jumeaux alors l'ensemble de tous ses sommets est clairement un code identifiant.

Proposition 1.1 *Un graphe admet un code identifiant si et seulement s'il est sans jumeaux.*

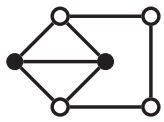


Figure 1.14 – *Les deux sommets noirs forment une paire de jumeaux.*

Lorsque seulement une partie T des sommets du graphe est utilisée pour l'identification, comme dans les codes discriminants, nous utiliserons le terme *T-jumeaux* pour désigner deux sommets u et v ayant le même voisinage fermé selon T : $N[u] \cap T = N[v] \cap T$. Dans les graphes orientés, deux sommets u et v sont jumeaux s'ils ont le même voisinage fermé entrant : $\mathcal{B}_1^+(u) = \mathcal{B}_1^+(v)$.

1.3.2 Hérédité

La propriété d'être un code identifiant est héréditaire. Ce résultat sera utilisé instinctivement tout au long de cette thèse.

Proposition 1.2 *Soient G un graphe sans jumeaux et C un code identifiant de G. Alors tout ensemble C' contenant C est un code identifiant de G.*

1.3.3 Bornes

Étant donné un graphe sans jumeaux G, il est facile de trouver un code identifiant de G : l'ensemble de tous les sommets en est un. Ce qui nous intéresse est de trouver des codes identifiants efficaces, c'est-à-dire comportant le moins d'éléments du code

possible. Une partie du problème des codes identifiants consiste donc à chercher la cardinalité optimale $\gamma^{\text{ID}}(G)$ d'un code identifiant. Nous ne sommes en général pas en mesure de déterminer la valeur exacte de $\gamma^{\text{ID}}(G)$ (voir la partie 1.3.4), mais nous pouvons borner cette valeur.

Une première borne inférieure est obtenue en remarquant que chaque sommet du graphe à un identifiant différent qui correspond à un sous-ensemble non vide de sommets du code. Si le code C considéré contient k éléments, il y a donc au plus $2^k - 1$ sous-ensembles non vide de C, ce qui mène à la borne suivante :

Théorème 1.3 (Karpovsky, Chakrabarty, Levitin [64]) *Soit G un graphe fini sans jumeaux. Alors :*

$$\gamma^{ID}(G) \geq \lceil \log(|V(G)| + 1) \rceil.$$

Cette borne est atteinte par une infinité de graphes, la figure 1.15 en montre un exemple. Moncel [77] donne une classification des graphes l'atteignant. La preuve et le résultat du théorème 1.3 se généralisent bien dans le cadre des codes discriminants pour lesquels nous obtenons la borne

$$\gamma^{\text{D}}(B) \geq \lceil \log(|S| + 1) \rceil$$

où S est la partie des sommets à identifier du graphe biparti B.

$\gamma^{\text{ID}}(G) = 3 = \lceil \log(|V(G)| + 1) \rceil$ $\qquad$ $\gamma^{\text{ID}}(G) = 3 = |V(G)| - 1$

Figure 1.15 – *Deux graphes atteignant les bornes inférieure et supérieure en fonction du nombre de sommets.*

La borne suivante est une borne supérieure en fonction du nombre de sommets.

Théorème 1.4 (Bertrand [9] et Gravier, Moncel [48])
Soit G un graphe sans jumeaux possédant au moins une arête. Alors :

$$\gamma^{ID}(G) \leq |V(G)| - 1.$$

Nous donnons une preuve de ce résultat dans le chapitre 2 et classifions tous les graphes atteignant cette borne. La figure 1.15 en donne un exemple. Il se trouve que tous les graphes atteignant cette borne ont un degré maximum élevé (supérieur à $|V(G)| - 2$). Foucaud *et al.* [40] proposent une conjecture permettant de borner la taille d'un code identifiant en fonction du degré maximum du graphe.

Conjecture 1.5 (Foucaud, Klasing, Kosowski, Raspaud [40])
Soit G un graphe non trivial sans jumeaux de degré maximum Δ. Alors :

$$\gamma^{ID}(G) \leq |V(G)| - \frac{|V(G)|}{\Delta} + O(1).$$

1.3.4 Complexité

Un autre sujet d'étude sur le problème des codes identifiants est sa complexité : peut-on trouver de manière efficace un code identifiant optimal ? Le problème de décision IDCODE se pose ainsi :

IDCODE
INSTANCE : Un graphe G sans jumeaux, un entier k
QUESTION : G possède-t-il un code identifiant de taille inférieure ou égale à k ?

On ne connaît pas de manière efficace pour résoudre ce problème. En effet, le problème IDCODE est $\mathcal{NP}$-complet [3] (Cohen *et al.* [27]), même lorsque le graphe G appartient à des classes restreintes de graphes comme la classe des graphes bipartis de degré maximum 3 (Charon *et al.* [21]) ou la classe des graphes planaires de degré maximum 4 et de maille arbitrairement grande (Auger *et al.* [2]). Dans le chapitre 3 nous montrons qu'il est $\mathcal{NP}$-complet pour la classe des graphes adjoints. Le problème IDCODE est par contre résolvable en temps linéaire sur la classe des arbres (Auger [1]).

3. Nous n'avons pas défini ici cette notion, nous renvoyons le lecteur au livre de Garey et Johnson [43], référence sur le sujet.

Chapitre 2

Familles extrémales

Sommaire

2.1 Outils préliminaires **32**
2.1.1 Théorème de Bondy 32
2.1.2 Autres outils 34
2.2 Cas orienté **35**
2.2.1 Graphes finis 35
2.2.2 Graphes infinis 41
2.3 Cas non orienté **44**
2.3.1 Graphes finis 44
2.3.2 Graphes infinis 51
2.3.3 Conséquences 54
2.3.4 Codes r-identifiants 55
2.4 Conclusion **56**

Dans ce chapitre, nous répondons à la question suivante : quels sont les graphes, orientés ou non orientés, dont les codes identifiants optimaux sont les plus grands, en terme de nombre de sommets ? Nous cherchons en quelque sorte à classer les graphes "mauvais" pour l'identification. Le théorème 1.4 nous dit, dans le cas non orienté, que leurs codes identifiants sont de taille $n-1$ où n est le nombre de sommets du graphe. Nous affinons ce résultat en donnant l'ensemble des graphes non orientés vérifiant cette propriété et donnons une classification similaire pour les graphes orientés.

Ce type de questions a déjà été résolu pour d'autres problèmes d'identification comme les ensembles résolvants (Chartrand *et al.* [24]) ou bien les codes localisateur-dominateurs (Slater [88]). Pour les codes identifiants, seulement une partie de ces graphes était connue avant notre étude.

Les outils que nous allons utiliser sont communs à l'étude des graphes orientés et non orientés et trouvent un sens dans le cadre des codes discriminants (ou de manière équivalente dans les hypergraphes). Dans une première partie, nous donnons des outils pour la suite de l'étude, exprimés en termes de codes discriminants. Dans la deuxième partie, nous considérons le cas des graphes orientés. Nous montrons que les graphes orientés dont le seul code identifiant est l'ensemble des sommets sont essentiellement les fermetures transitives d'arbres orientés. Dans la troisième partie, nous considérons les graphes non orientés. Pour ceux-ci, les exemples extrémaux finis sont composés de puissances de chemins particulières, jointes entre elles. Ces résultats ont des conséquences sur des bornes pour les codes identifiants que nous exposons au fur et à mesure de l'étude.

Ces travaux sont issus des articles [38] et [42] qui ont été réalisés avec Florent Foucaud, Eleonora Guerrini, Matjaž Kovše, Reza Naserasr et Petru Valicov.

2.1 Outils préliminaires

2.1.1 Théorème de Bondy et utilisation pour les codes discriminants

Le théorème suivant, attribué à Bondy [12], nous permet de trouver des bornes supérieures pour des ensembles séparants dans les codes discriminants, ce qui aura des applications pour certains problèmes d'identification. Ce théorème a par la suite reçu beaucoup d'attentions et de nombreuses preuves ont été données (voir par exemple [11, 74, 94]). Nous rappelons son énoncé ainsi que la preuve originelle de ce théorème.

Théorème 2.1 (Bondy [12])
Soit $\mathcal{A} = \{A_1, \dots, A_n\}$ une famille de n sous-ensembles distincts d'un ensemble X à n éléments. Il existe un sous-ensemble X' de X contenant $n-1$ éléments tel que les ensembles $A_i \cap X'$, $1 \leq i \leq n$, soient tous distincts.

Preuve : Soit $G_{\mathcal{A}}$ le graphe étiqueté suivant : les sommets de $G_{\mathcal{A}}$ sont les ensembles $A_1, \dots, A_n$ et pour chaque $x \in X$, s'il existe deux ensembles A_i et A_j tels que $A_i \Delta A_j = \{x\}$, alors, *pour une seule telle paire* $\{A_i, A_j\}$, il y a une arête entre A_i et A_j étiquetée x.

Le graphe $G_{\mathcal{A}}$ a au plus n arêtes et toutes les arêtes ont des étiquettes différentes. Supposons que $G_{\mathcal{A}}$ contienne exactement n arêtes. Alors $G_{\mathcal{A}}$ contient un cycle. Sans perte de généralité, les sommets du cycle sont A_1,...,A_k et les arêtes correspondantes sont étiquetées x_1,...,x_k, avec $A_i \Delta A_{i+1} = \{x_i\}$, pour $i < k$, et $A_k \Delta A_1 = \{x_k\}$. Nous pouvons supposer que $x_1 \in A_2$ et $x_1 \notin A_1$. Alors $x_1 \in A_3$ car A_3 et A_2 ne diffèrent que sur l'élément x_2 et $x_2 \neq x_1$. De la même manière, $x_1 \in A_4, A_5, ...$ Finalement, $x_1 \in A_k$ et comme $x_k \neq x_1$, nous avons $x_1 \in A_1$, ce qui est contradictoire.

Le graphe $G_{\mathcal{A}}$ contient donc au plus $n-1$ arêtes. Cela signifie qu'un des éléments de X, appelons-le y, n'apparaît sur aucune arête. Par définition du graphe, nous avons donc pour tous $i \neq j$, $A_i \Delta A_j \neq \{y\}$ et $X' = X \setminus \{y\}$ convient. □

La preuve du théorème s'adapte naturellement à un ensemble X de taille plus grande :

Corollaire 2.2 *Soit $\mathcal{A} = \{A_1, \ldots, A_n\}$ une famille de n sous-ensembles distincts d'un ensemble X à m éléments, avec $m \geq n$. Alors il existe un sous-ensemble X' de X contenant $n-1$ éléments tel que les ensembles $A_i \cap X'$, $1 \leq i \leq n$, soient tous distincts.*

Ce théorème se traduit en terme d'ensemble séparant dans les graphes bipartis.

Corollaire 2.3 (Charbit, Charon, Cohen, Hudry, Lobstein [16])
Soit $B = (S \cup T, E)$ un graphe biparti sans T-jumeaux au sein de S. Il existe un ensemble séparant S, inclus dans T, de taille inférieure ou égale à $|S| - 1$.

Preuve : C'est une application directe du corollaire 2.2 en prenant $X = T$ et $\mathcal{A} = \{N(i_k), i_k \in S\}$. □

Charbit *et al.* [16] caractérisent les graphes bipartis atteignant cette borne en utilisant un graphe étiqueté construit sur le modèle du graphe $\mathcal{A}$ de la preuve du théorème 2.1. Ce graphe, noté $H(B, C)$, est construit à partir du graphe B et d'un sous-ensemble C de T séparant S. Les sommets de $H(B, C)$ sont les sujets S et pour chaque caractéristique $c \in C$, s'il existe une paire de sujets i,j dans S tels que

$$(N[i] \cap C)\Delta(N[j] \cap C) = \{c\},$$

alors pour *une seule* telle paire, l'arête ij, étiquetée par c, est ajoutée au graphe $H(B, C)$. La figure 2.1 donne un exemple de construction.

Lorsque C est pris minimal pour l'inclusion, le lemme suivant garantit l'existence d'une arête pour chaque élément du code.

Lemme 2.4 (Charbit et al. [16]) *Dans un graphe biparti $B = (S \cup T, E)$, un ensemble $C \subseteq T$ séparant S est minimal pour l'inclusion si et seulement si*

$$\forall c \in C, \exists i, j \in S : (N[i] \cap C)\Delta(N[j] \cap C) = \{c\}.$$

En utilisant les mêmes arguments que dans la preuve du théorème de Bondy, il ne peut pas y avoir de cycles dans $H(B, C)$. Notons aussi que la construction du graphe $H(B, C)$ n'est pas unique : dans l'exemple de la figure 2, l'arête $s_2 s_4$ dans

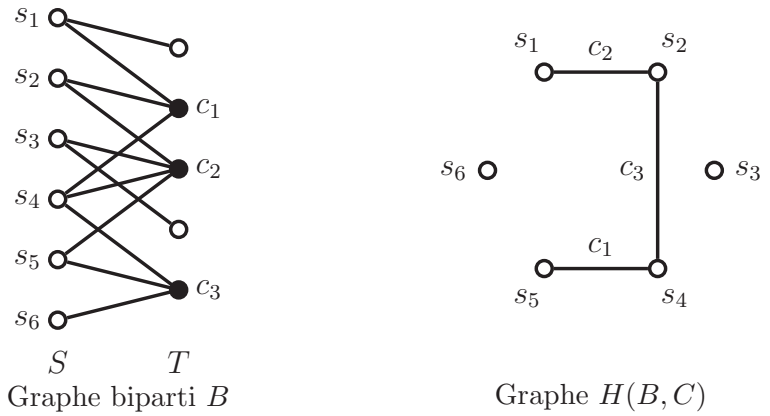


Figure 2.1 – *Construction du graphe $H(B,C)$ à partir d'un graphe biparti B et d'un code discriminant C.*

le graphe $H(B,C)$ pour être remplacée par l'arête s_3s_5. Dans le cas où C est un ensemble séparant S de taille minimale égale à $|S|-1$, le lemme suivant implique que la construction du graphe $H(B,C)$ est unique et possède $n-1$ arêtes, c'est donc un arbre.

Lemme 2.5 (Charbit et al. [16]) *Soit $B = (S \cup T, E)$ un graphe biparti. Soit $C \subseteq T$ un ensemble séparant S. Si $|C| = |S|-1$, alors pour tout élément de C, il existe une unique paire i,j d'individus tels que $(N[i] \cap C)\Delta(N[j] \cap C) = \{c\}$.*

Par construction l'arête étiquetée c de l'arbre $H(B,C)$ sépare alors les individus en deux sous-ensembles, l'un contenant l'élément c, l'autre ne le contenant pas. À partir de là, il est possible de compter et caractériser les graphes extrémaux pour le problème des codes discriminants dans les graphes bipartis. Nous utiliserons dans la partie 2.2.1 cette méthode pour décrire précisément les graphes orientés atteignant la borne supérieure.

2.1.2 Autres outils

La proposition suivante nous permettra de raisonner par induction.

Proposition 2.6 *Soit $B = (S \cup T, E)$ un graphe biparti possédant un code discriminant et soit S' un sous-ensemble de S. Alors :*

$$\gamma^D(G) \leq \gamma^D(G - S') + |S'|.$$

Preuve : Par hypothèse, le graphe B admet un code discriminant. En enlevant des sommets de S, cette propriété reste vraie et le graphe $B - S'$ admet donc aussi un code discriminant. Soit $C_0 \subseteq T$ un code discriminant de $B - S'$ de taille minimale. Soit $(x_1, \ldots, x_{|s'|})$ un ordre arbitraire des sommets de S'. De proche en proche, nous allons étendre C_0 à des sous-ensembles de sommets C_i de T de telle sorte que C_i identifie tous les sommets de $S_i = S - \{x_{i+1}, \ldots, x_{|S'|}\}$.

Soit $i \geq 1$ et supposons que C_{i-1} soit construit. Si C_{i-1} identifie les sommets de S_i, alors $C_i = C_{i-1}$. Sinon, comme C_{i-1} identifie les sommets de S_{i-1}, soit x_i n'est pas dominé par C_{i-1}, soit il existe un unique sommet y dans S_{i-1} tel que $N[x_i] \cap C_{i-1} = N[y] \cap C_{i-1}$. Dans le premier cas, il existe un élément c dans T qui domine x_i car B admet un code discriminant. Alors $C_i = C_{i-1} \cup \{c\}$ identifie les sommets de S_i. Dans le deuxième cas, x_i est dominé par un sommet de C_{i-1} et x_i et y ne sont pas T-jumeaux dans le graphe B, il existe donc un sommet $c \in T$ qui les sépare. Alors $C_i = C_{i-1} \cup \{c\}$ est un code discriminant.

À l'étape $|S'|$, $C_{|S'|}$ est un code discriminant de B de taille au plus $|C_0| + |S'|$ car au plus un sommet est ajouté à chaque étape. □

2.2 Cas orienté

Nous nous intéressons maintenant aux codes identifiants dans les graphes orientés. Contrairement au cas non orienté où le théorème 1.4 implique qu'il est toujours possible de retirer un des sommets du graphe pour former un code identifiant, il existe des graphes orientés non triviaux pour lesquels le seul code identifiant est l'ensemble de tous les sommets. Nous allons classifier tous ces graphes, pour les cas fini et infini.

2.2.1 Graphes finis

Dans cette partie nous classifions tous les graphes orientés finis D tels que $\overrightarrow{\gamma^{\text{ID}}}(D) = |V(D)|$. Nous allons montrer que ces graphes sont des fermetures transitives de forêts orientées. Une conséquence importante de cette classification est que ces graphes sont acycliques. En particulier, ils n'ont pas d'arcs symétriques. Dans un premier temps, nous donnons des outils nécessaires à la preuve, essentiellement conséquences de résultats de la partie 2.1. Nous construisons ensuite une famille de graphes extrémaux et montrons que ce sont les seuls. Nous proposons une preuve alternative de ce résultat en utilisant les graphes $H(B, C)$ de la partie 2.1.1. Finalement, nous donnons quelques conséquences dues à cette classification pour le théorème de Bondy.

Outils

La proposition suivante est une conséquence de la proposition 2.6 :

Proposition 2.7 *Soit D un graphe fini orienté sans jumeaux et soit S un ensemble de sommets de D tel que $D - S$ soit sans jumeaux. Alors*

$$\overrightarrow{\gamma^{\text{ID}}}(D) \leq \overrightarrow{\gamma^{\text{ID}}}(D - S) + |S|.$$

Preuve : Soient $B = B(D)$ le graphe d'incidence de D (défini dans la partie 1.1.7 du chapitre 1) et $B' = B(D) - S$ (les sommets ne sont enlevés que du côté des sujets S). D'après la proposition 2.6, $\gamma^{\text{D}}(B) \leq \gamma^{\text{D}}(B') + |S|$. Nous ne pouvons faire

que de meilleurs codes discriminants en ajoutant des caractéristiques dans T, donc $\gamma^{\mathrm{D}}(B') \leq \overrightarrow{\gamma^{\mathrm{ID}}}(D - S)$. De plus, $\gamma^{\mathrm{D}}(B) = \overrightarrow{\gamma^{\mathrm{ID}}}(D)$, ce qui permet de conclure. □

Le corollaire 2.3 se traduit directement dans les termes suivants pour le cas des graphes orientés :

Proposition 2.8 *Soit D un graphe fini orienté sans jumeaux et soit S un ensemble séparant D. Alors :*

$$\gamma^{s}(D) \leq |V(D)| - 1.$$

Description de la famille

Nous construisons des graphes extrémaux en utilisant les opérateurs $\overrightarrow{\lhd}$ et $\sqcup$ définis dans le chapitre 1 et illustrés sur la figure 2.2.

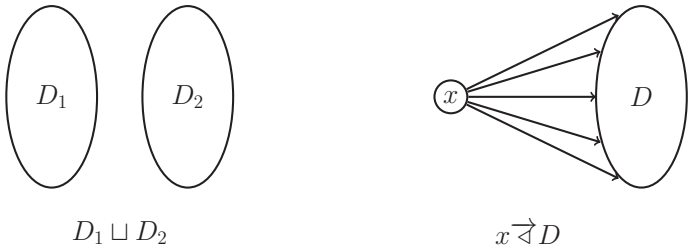


Figure 2.2 – *Illustration des deux opérateurs utilisés*

Soit $(K_1, \sqcup, \overrightarrow{\lhd})$ la famille de graphes orientés contenant K_1 (vu comme un graphe orienté contenant un seul sommet) et close par les opérateurs $\sqcup$ et $\overrightarrow{\lhd}$. Rappelons que dans un arbre orienté enraciné, toutes les arêtes sont orientées depuis la racine et qu'une forêt orientée enracinée est l'union d'arbres orientés enracinés. Nous omettons la preuve du lemme suivant qui peut se faire par induction.

Lemme 2.9 *Tout élément de $(K_1, \sqcup, \overrightarrow{\lhd})$ est la fermeture transitive d'une forêt orientée enracinée.*

Pour un graphe orienté D dans $(K_1, \sqcup, \overrightarrow{\lhd})$, nous notons $\overrightarrow{F}(D)$ la forêt orientée enracinée dont la fermeture transitive est D.

Proposition 2.10 *Pour tout graphe orienté D dans $(K_1, \sqcup, \overrightarrow{\lhd})$, $\overrightarrow{\gamma^{\mathrm{ID}}}(D) = |V(D)|$.*

De plus, si un sommet x est un sommet source de D, alors $V(D) \setminus \{x\}$ est un ensemble séparant de D. Sinon, le sommet x et son père dans $\overrightarrow{F}(D)$ sont les seuls sommets à ne pas être séparés l'un de l'autre par l'ensemble $V(D) \setminus \{x\}$.

Preuve : Soit C un code identifiant de D. Exceptés les sommets sources (qui sont les racines des arbres de $\overrightarrow{F}(D)$), chaque sommet doit être dans C pour être séparé de son père. Les sommets sources doivent aussi être dans C pour être dominés. □

Il n'y a pas d'autres graphes extrémaux

Nous allons maintenant prouver que ce sont les seuls graphes orientés à atteindre cette borne. Nous donnons tout d'abord une preuve directe. Nous verrons dans la suite comment obtenir ce résultat à partir des codes discriminants.

Théorème 2.11 *Soit D un graphe fini orienté sans jumeaux. $\overrightarrow{\gamma^{\text{ID}}}(D) = |V(D)|$ si et seulement si $D \in (K_1, \sqcup, \overrightarrow{\lhd})$.*

Preuve : Il suffit de montrer que si $\overrightarrow{\gamma^{\text{ID}}}(D) = |V(D)|$ alors $D \in (K_1, \sqcup, \overrightarrow{\lhd})$, l'autre sens étant la proposition 2.10. Supposons par l'absurde qu'il existe un graphe orienté fini sans jumeaux D tel que $\overrightarrow{\gamma^{\text{ID}}}(D) = |V(D)|$ mais que $D \notin (K_1, \sqcup, \overrightarrow{\lhd})$. Choisissons D avec le moins de sommets possible. Nous considérons deux cas.

Cas a. Il existe un sommet x de D tel que x n'ait pas de voisins sortants (x est un puits). Alors $D - x$ est sans jumeaux. D'après la proposition 2.7, nous avons $\overrightarrow{\gamma^{\text{ID}}}(D - x) = |V(D - x)|$. Par minimalité de D, $D - x \in (K_1, \sqcup, \overrightarrow{\lhd})$, et $\overrightarrow{F}(D - x)$ est bien définie. Comme $V(D) \setminus \{x\}$ n'est pas un code identifiant de D, soit x n'est pas dominé par $V(D) \setminus \{x\}$ soit il y a un sommet y qui n'est pas séparé de x par $V(D) \setminus \{x\}$.

Dans le premier cas, x est un sommet isolé et donc D est l'union disjointe de deux membres de $(K_1, \sqcup, \overrightarrow{\lhd})$, D est dans la famille, ce qui mène à une contradiction.

Nous pouvons donc supposer qu'il y a un sommet y qui n'est pas séparé de x par $V(D) \setminus \{x\}$. Alors, $N^+(x) = N^+(y) \cup \{y\}$ et D est la fermeture transitive de la forêt orientée construite à partir de $\overrightarrow{F}(D - x)$ en ajoutant x comme fils de y, ce qui mène à une contradiction.

Cas b. Tous les sommets du graphe ont au moins un voisin sortant. D'après la proposition 2.8, il existe un sommet x tel que $V(D) \setminus \{x\}$ soit un ensemble séparant, ce qui signifie que $D - x$ est sans jumeaux. En utilisant la proposition 2.7, $\overrightarrow{\gamma^{\text{ID}}}(D - x) = |V(D - x)|$ et $\overrightarrow{F}(D - x)$ est bien définie.

Comme chaque sommet de D, et en particulier chaque feuille t de $\overrightarrow{F}(D - x)$, a un sommet sortant de D, nous avons $\overrightarrow{tx} \in A(D)$. Le degré entrant de x, $d^+(x)$, est donc d'au moins un et $V(D) \setminus \{x\}$ est un ensemble dominant. Ce n'est pas un code identifiant de D, donc un sommet $y \neq x$ n'est pas séparé de x par $V(D) \setminus \{x\}$: $N^+(x) = N^+(y) \cup \{y\}$.

Nous allons maintenant montrer que y est une feuille, et que c'est l'unique feuille de $\overrightarrow{F}(D - x)$. Si $t \neq y$ est une feuille, alors $t \in N^+(x)$ et donc $t \in N^+(y)$ ce qui est une contradiction (le seul arc sortant de t est l'arc $\overrightarrow{tx}$). Comme il y a au moins une feuille dans $\overrightarrow{F}(D - x)$, y doit être cette feuille et il n'y en a qu'une. Cela implique que $\overrightarrow{F}(D - x)$ est un chemin dirigé.

Tous les sommets du graphe ont au moins un voisin sortant. En particulier, x a un voisin sortant t. Nécessairement $t \neq y$ sinon x et y seraient jumeaux. Supposons que t puisse être choisi de sorte que le père de t dans $\overrightarrow{F}(D - x)$ ne soit pas un voisin sortant de x. Alors $C = V(D) \setminus \{t\}$ est un code identifiant. En effet, x est le seul sommet dominé par tous les sommets de C, Le sommet t et son père sont séparés

par x. Toutes les autres paires de sommets sont séparées par le sommet qui est le descendant de l'autre dans $\overrightarrow{F}(D-x)$, s'il est différent de t, et par le père de t sinon.

Supposons maintenant que pour tout sommet $t \in N^-(x)$, le père de t soit aussi dans $N^-(x)$. Cela implique que la racine r de $\overrightarrow{F}(D-x)$ est dans $N^-(x)$. Alors $C = V(D) \setminus \{r\}$ est un code identifiant. En effet, x est le seul sommet dominé par tous les sommets de C. De plus chaque paire de sommet de $V(D) \setminus \{x\}$ est séparée par le sommet qui est le descendant de l'autre dans $\overrightarrow{F}(D-x)$ et enfin r est dominé par x .

Dans tous les cas, nous avons exhibé un code identifiant de cardinalité $|V(D)|-1$, ce qui mène à une contradiction. □

Les graphes de $(K_1, \sqcup, \overrightarrow{\lhd})$ sont acycliques ce qui mène au résultat suivant.

Corollaire 2.12 *Soit D un graphe fini orienté sans jumeaux contenant un cycle orienté. Alors $\overrightarrow{\gamma^{ID}}(G) \leq |V(D)| - 1$.*

Preuve alternative via les codes discriminants

À chaque graphe orienté D est associé son graphe biparti d'incidence $B(D)$ et $\overrightarrow{\gamma^{\text{ID}}}(D) = \gamma^{\text{D}}(B(D))$. Ce graphe biparti $B(D)$ a la particularité de posséder un couplage parfait. C'est aussi le cas des graphes bipartis B tels que $\gamma^{\text{D}}(B) = |T| = |S|$:

Lemme 2.13 *Soit $B = (S \cup T, E)$ un graphe biparti sans jumeaux tel que $\gamma^{D}(B) = |T| = |S|$. Alors B possède un couplage parfait.*

Preuve : Supposons que B ne possède pas de couplage parfait. D'après le théorème des mariages de Hall (voir par exemple [11]), comme $|S| = |T|$, il existe un sous-ensemble S' de S tel que $|N(S')| < |S'|$.

Soient $T' = N(S')$ et $S'' = S \setminus S'$. T' est un code discriminant de $B - S''$ donc $\gamma^{\text{D}}(B - S'') \leq |T'|$. D'après la proposition 2.6, $\gamma^{\text{D}}(B) \leq |T'| + |S''| \leq |S| - 1$, une contradiction. □

Pour tout graphe biparti B possédant un couplage parfait, il existe un graphe orienté D tel que $B = B(D)$. Chercher les graphes orientés D tels que $\overrightarrow{\gamma^{\text{ID}}}(D) = |V(D)|$ revient donc à chercher parmi les graphes bipartis d'attributs $B = (S \cup T, E)$ avec $|S| = |T|$ ceux qui vérifient $\gamma^{\text{D}}(B) = |T|$.

Cette étude a été menée par Charbit *et al.* dans [16] mais le lien avec les graphes orientés n'a pas été fait. Pour traiter ce cas, les auteurs de [16] ajoutent un sommet isolé s_0 à l'ensemble S. Soit $B + s_0$ le graphe obtenu. Alors un code C séparant S de $B + s_0$ correspond à un code discriminant dans le graphe B car tous les éléments de S doivent être dominés pour être séparés de s_0. D'après la proposition 2.5, le graphe $H(B, C)$ est alors un arbre. Chaque sommet de l'arbre correspond à un élément de S ou bien au sommet isolé s_0. Une fois le choix de s_0 fait, nous allons voir que la reconstruction du graphe biparti est unique, ce graphe contient un unique couplage parfait et le graphe orienté qui lui correspond est la fermeture transitive d'une forêt orientée enracinée.

Lemme 2.14 *Soit H un arbre sur l'ensemble de sommets $S + s_0$ dont les n arêtes sont étiquetées par des éléments de $T = \{t_1, ..., t_n\}$, de sorte que chaque étiquette soit différente des autres. Il existe un unique graphe biparti $B = (S \cup T, E)$ tel que $H = H(B + s_0, T)$, $|S| = |T|$ et dont le seul code discriminant S est l'ensemble T. Le graphe B possède un unique couplage parfait et le graphe orienté G associé à ce couplage parfait est la fermeture transitive d'une forêt orientée.*

Preuve : Soit $B = (S \cup T, E)$ un graphe biparti tel que $H = H(B + s_0, T)$ et dont le seul code discriminant S est l'ensemble T. Alors le graphe $H(B + s_0, T)$ est construit de manière unique (c'est le graphe H), et chaque arête de H sépare les sommets de $B + s_0$ en deux parties : ceux ayant l'étiquette de l'arête comme voisin et les autres. De plus, le graphe B possède un couplage parfait, noté M.

Le sommet s_0 est fixé dans l'arbre H et c'est un sommet isolé de $B + s_0$. Il n'est donc relié à aucun élément de T dans B. Soit e une arête incidente à s_0. Sans perte de généralité, supposons que e est étiquetée par t_1. L'arête e sépare les sommets de H en deux parties S_1 et $S_1' \cup \{s_0\}$. Tous les éléments de S_1 sont adjacents à t_1 dans B alors que tous les éléments de S_2 ne le sont pas, car s_0 n'est pas adjacent à t_1. Ainsi $N_B(t_1) = S_1$. Soit s_1 l'autre extrémité de e. Les sommets s_1 et s_0 ne sont séparés dans B que par le sommet t_1, donc $N_B(s_1) = \{t_1\}$. L'arête s_1t_1 est donc nécessairement dans le couplage M. Nous pouvons raisonner de la même manière avec tous les autres voisins dans H de s_0. Prenons maintenant un voisin s_2 de s_1 dans H. Soit t_2 l'étiquette de l'arête s_1s_2. Comme précédemment, l'arête s_1s_2 sépare le graphe H en deux composantes S_2 et S_2', avec $s_2 \in S_2$. Nous avons alors $N(t_2) = S_2$ et comme $N_B(s_1) = \{t_1\}$, $N_B(s_2) = \{t_1, t_2\}$. Alors l'arête s_2t_2 est dans le couplage M.

En continuant ce procédé, le couplage M est complètement fixé et donc unique. Le graphe D orienté tel que $B = B(D)$ est donc lui aussi unique. Or le graphe orienté D obtenu en prenant la fermeture transitive de la forêt orientée enracinée basée sur la forêt $H - s_0$ où tous les arbres ont comme racine le sommet adjacent à s_0, vérifie $B = B(D')$. Nous avons donc $D = D'$ ce qui termine la preuve. □

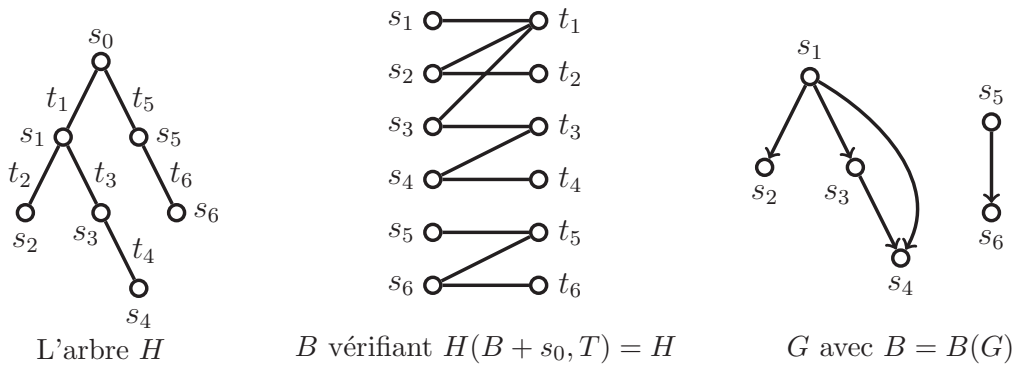
La figure 2.3 donne un exemple de passage de l'arbre H au graphe biparti puis au graphe orienté.

Corollaire 2.15 *Soit $B = (S \cup T, E)$ un graphe biparti vérifiant $|S| = |T|$ et $\gamma^D(B) = |T|$. Alors B est le graphe d'incidence de la fermeture transitive d'une forêt orientée.*

Conséquence pour le théorème de Bondy

Dans le théorème de Bondy, est-il possible de choisir l'ensemble X' de telle sorte que tous les ensembles $A_i \cap X'$ soient non vides, en supposant que tous les ensembles A_i étaient non vides au départ ? Notre caractérisation nous permet de répondre à cette question en fonction des ensembles de départ. Nous appelons un couple $(\mathcal{A}, X)$ avec $\mathcal{A}$ une collection de sous-ensembles de X un *système d'ensembles*. Un système d'ensembles $(\mathcal{A}, X)$ est dit *extrémal* si :

– Tous les ensembles de $\mathcal{A}$ sont distincts et non vides ;



L'arbre H — B vérifiant $H(B + s_0, T) = H$ — G avec $B = B(G)$

Figure 2.3 – *À partir d'un arbre H, il existe un unique graphe biparti B extrémal tel que $H = H(B + s_0, T)$ correspondant au graphe d'incidence de la fermeture transitive d'une forêt orientée.*

- $|\mathcal{A}| = |X|$;
- Pour tout $x \in X$, $\{x\} \in \mathcal{A}$ ou bien il existe deux éléments A_i et A_j dans $\mathcal{A}$ tels que $A_i \Delta A_j = \{x\}$.

À un système d'ensembles $(\mathcal{A}, X)$ est naturellement associé un graphe biparti $B(\mathcal{A}, X)$ représentant les incidences des ensembles de $\mathcal{A}$ avec X : l'ensemble S de B correspond à l'ensemble $\mathcal{A}$ et T correspond à l'ensemble X. Il y a une arête entre A_i et x si $x \in A_i$. Le lemme suivant reformule la notion de système extrémal en terme de codes discriminants.

Lemme 2.16 *$(\mathcal{A}, X)$ est extrémal si et seulement si $|\mathcal{A}| = |X|$ et le seul code discriminant de $B(\mathcal{A}, X)$ est l'ensemble T tout entier.*

En utilisant le théorème 2.11 :

Théorème 2.17 *Un ensemble de système $(\mathcal{A}, X)$ est extrémal si et seulement si $B(\mathcal{A}, X)$ est le graphe biparti d'incidence d'un graphe de $(K_1, \sqcup, \overrightarrow{\lhd})$.*

Cela peut se traduire directement dans le langage des systèmes d'ensembles, sans faire de référence aux graphes bipartis :

Corollaire 2.18 *Un ensemble de système $(\mathcal{A}, X)$ avec $|\mathcal{A}| = |X|$ est extrémal si et seulement si*
- $\bigcup_{A \in \mathcal{A}} A = X$
- *Pour tout ensemble $A \in \mathcal{A}$ avec au moins deux éléments, il existe un élément x de A tel que $A \setminus \{x\} \in \mathcal{A}$.*

Preuve : Si $(\mathcal{A}, X)$ est extrémal, alors le graphe biparti $B = B(\mathcal{A}, X)$ est le graphe d'incidence d'un élément D de $(K_1, \sqcup, \overrightarrow{\lhd})$ (théorème 2.17). En particulier, B possède un couplage parfait ce qui assure que chaque élément de X est présent dans au moins un ensemble de $\mathcal{A}$. Soit maintenant A un ensemble avec au moins deux éléments. L'ensemble A correspond à un sommet x de D qui n'est pas une source. Soit y le père

de x dans l'arbre sous-jacent à la forêt orientée D. Alors $A = \mathcal{B}_1^+(x)$ et l'ensemble $\mathcal{B}_1^+(y) = A \setminus \{x\}$ est dans la famille $\mathcal{A}$.

Soit maintenant un système d'ensemble $(\mathcal{A}, X)$ ayant les propriétés énoncées dans le corollaire. Supposons qu'il existe un élément x tels que tous les ensembles $A \setminus \{x\}$ pour $A \in \mathcal{A}$ soient non vides et distincts. Soit A un sous-ensemble de taille minimale contenant x (il existe car $\bigcup_{A \in \mathcal{A}} A = X$). Alors $A \neq \{x\}$ donc il existe un élément y et un ensemble A' tels que $A \setminus \{y\} = A'$. Nécessairement $x \neq y$ et A' contient x. Cela contredit la minimalité de A. Le système $(\mathcal{A}, X)$ est donc extrémal. □

2.2.2 Graphes infinis

Dans cette partie nous considérons les graphes infinis orientés, sans jumeaux, dont le seul code identifiant est l'ensemble de tous les sommets. Nous nous restreignons aux graphes orientés sans arcs symétriques. Classifier tous les graphes infinis orientés extrémaux nous semble difficile étant donné que tous les graphes infinis non orientés fournissent déjà des exemples de tels graphes, en remplaçant chaque arête xy par deux arcs $\overrightarrow{xy}$ et $\overrightarrow{yx}$, et que cette classe est déjà riche comme nous le verrons dans la partie 2.3.2.

Nous allons montrer que les graphes orientés infinis sans arcs symétriques dont le seul code identifiant est l'ensemble de tous les sommets sont composés de fermetures transitives d'arbres infinis orientés. Nous définissons tout d'abord ce que nous entendons par arbre infini orienté.

Un arbre infini orienté est un graphe orienté dont le graphe sous-jacent est un arbre (connexe et sans cycle) et dont chaque sommet a au plus une arête entrante. Nous considérons ici que tout chemin entre deux sommets fixés est fini.

S'il existe un sommet source dans un arbre infini orienté, il y a en exactement un, qui est appelé *racine*. Un arbre infini possédant une racine sera appelé *arbre de source finie*. Par opposition, un arbre sans source est appelé *arbre de source infinie*. La figure 2.4 donne des exemples d'arbres orientés infinis ou finis.

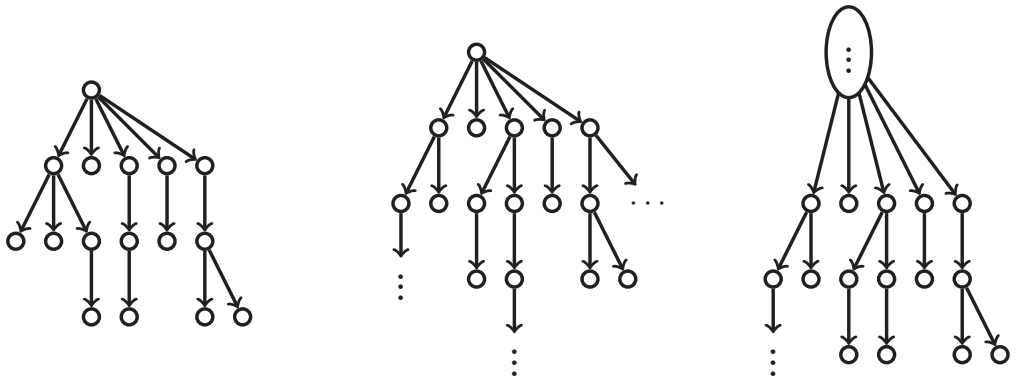


Figure 2.4 – *Exemples d'arbres orientés finis ou infinis.*

Une manière équivalente de définir les fermetures transitives d'arbres orientés infinis est la suivante. Un graphe orienté connexe D est la fermeture transitive d'un arbre orienté infini si et seulement si :

(1) Pour chaque sommet x de D, $\mathcal{B}_1^+(x)$ induit la fermeture transitive d'un chemin orienté, P_x, avec x comme extrémité.

(2) Pour chaque paire de sommets x, y, il existe un sommet z tel que $P_x \cap P_y = P_z$.

Les arbres de source finie sont ceux pour lesquels un des chemins P_x est fini, et de fait, tous les chemins P_x sont alors finis, et ont la même extrémité : la racine de l'arbre. A l'inverse, pour les arbres de sources infinies, tous les chemins sont semi-infinies et leur branche infinie est commune : il y a en quelque sorte une unique source infinie.

Notons que les arbres définis ainsi qui ont un nombre fini de sommets sont exactement les éléments connexes de $(K_1, \sqcup, \overrightarrow{\lhd})$.

Proposition 2.19 *Le seul code identifiant de la fermeture transitive d'un arbre orienté infini est l'ensemble de tous ses sommets.*

Preuve : Soit $D = (V, A)$ la fermeture transitive d'un arbre orienté infini et soit x un sommet de D. Si x est un sommet source, alors x doit être dans n'importe quel code identifiant pour être dominé. Si x n'est pas un sommet source, alors pour séparer x de son père dans l'arbre sous-jacent de D, x lui-même doit être dans tout code identifiant. □

Nous pouvons maintenant construire la famille complète des graphes orientés sans arcs symétriques qui ont besoin de tous leurs sommets pour être identifié. Étant donné un graphe orienté sans arcs symétriques H, nous construisons une famille de graphe $\Psi(H)$ comme suit.

1. Pour chaque sommet x de H, si x est un sommet source de H, nous attribuons à x la fermeture d'un arbre T_x de source finie. Si x n'est pas un sommet source nous attribuons à x la fermeture d'un arbre T_x de source infinie à x. Le choix de T_x est libre mais chaque T_x est construit sur son propre ensemble de sommets.
2. Pour chaque arc $\overrightarrow{xy}$ de H, nous attribuons un sous ensemble $V_{\overrightarrow{xy}}$ de $V(T_x)$, le choix de $V_{\overrightarrow{xy}}$ étant libre lui aussi.
3. Un élément de $\Psi(H)$ est construit en prenant comme ensemble de sommets $\bigcup_{x \in V(H)} V(T_x)$ et comme ensemble d'arcs, tous les arcs appartenant à l'un des arbres T_x ainsi que tous les arcs $\overrightarrow{zt}$ avec $z \in V_{\overrightarrow{xy}}$ et $t \in V(T_y)$.

La famille $\Psi(H)$ est l'ensemble des graphes pouvant être construits de cette manière. Notons qu'il n'y a pas d'arcs symétriques dans un graphe de $\Psi(H)$.

Proposition 2.20 *Étant donné un graphe orienté sans arcs symétriques H, tout graphe dans la famille $\Psi(H)$ n'est identifié que par l'ensemble de tous ses sommets.*

Preuve : Soit $D \in \Psi(H)$. Les sommets sources de D sont exactement les sources des arbres T_x de source finie attribués aux sommets sources de H, et doivent être dans tout code identifiant pour des raisons de domination.

Si un sommet u de D n'est pas une source, alors u appartient à la fermeture transitive d'un arbre T_x et, comme dans la preuve de la proposition 2.19, il y a un sommet v de T_x tel que $\mathcal{B}_1^+(u) \cap V(T_x) = (\mathcal{B}_1^+(v) \cap V(T_x)) \cup \{u\}$ (v est le père de u dans l'arbre T_x). Par construction, tout voisin entrant de u qui n'est pas dans T_x est aussi un voisin entrant de v. Donc $\mathcal{B}_1^+(u) = \mathcal{B}_1^+(v) \cup \{u\}$ et u doit être dans tout code identifiant de D. □

Théorème 2.21 *Soit D un graphe infini orienté sans arcs symétriques. L'ensemble de tous les sommets de D est le seul code identifiant de D si et seulement si $D \in \Psi(H)$ pour un graphe fini ou infini orienté sans arcs symétriques H.*

Preuve : La proposition 2.20 montre la condition suffisante du théorème. Nous montrons maintenant la condition nécessaire.

Soit D un graphe orienté infini sans arcs symétriques dont le seul code identifiant est l'ensemble de tous ses sommets. Soit x un sommet de D. L'ensemble $V(D) \setminus \{x\}$ n'est pas un code identifiant. Cela signifie x n'est pas dominé, et alors x est un sommet source, ou bien il existe deux sommets tels que $\mathcal{B}_1^+(v) = \mathcal{B}_1^+(u) \cup \{x\}$. Dans ce dernier cas, si x était différent de v, alors nous aurions deux arcs symétriques entre u et v. Cela n'est pas possible, donc $x = v$ et $\mathcal{B}_1^+(x) = \mathcal{B}_1^+(u) \cup \{x\}$. De plus, il existe un unique sommet u vérifiant $\mathcal{B}_1^+(x) = \mathcal{B}_1^+(u) \cup \{x\}$.

Ainsi, pour tout sommet x de D qui n'est pas une source, il existe un unique sommet, que nous notons x_{-1} tel que $\mathcal{B}_1^+(x) = \mathcal{B}_1^+(x_{-1}) \cup \{x\}$. Nous pouvons répéter cet argument sur x_{-i} pour obtenir x_{-i-1} pour $i = 1, 2, \cdots$ tant que x_{-i} n'est pas une source. Cela construit de manière unique un ensemble $\{\ldots, x_{-i}, \ldots, x_0 = x\}$ qui induit dans D la fermeture transitive d'un chemin fini ou infini P_x. Si x est une source, nous posons $P_x = \{x\}$.

Supposons maintenant que pour deux sommets x et y, $P_x \cap P_y \neq \emptyset$. Soit x_{-i} le premier sommet (dans l'ordre défini par P_x en partant de x) dans $P_x \cap P_y$. Par unicité de la construction, $P_x \cap P_y = P_{x_{-i}}$.

Nous définissons maintenant une relation d'équivalence sur les sommets de D : $x \equiv y$ si et seulement si $P_x \cap P_y \neq \emptyset$. La classe d'équivalence de x induit par définition la fermeture d'un arbre, noté T_x.

Si $u \notin V(T_x)$, il y a deux possibilités :
- soit $\overrightarrow{uv}$ est un arc de D pour tout sommet $v \in V(T_x)$,
- soit il n'y a aucun arc $\overrightarrow{uv}$ avec $v \in V(T_x)$.

En effet, supposons qu'il existe un arc $\overrightarrow{uv}$ pour un certain sommet $v \in V(T_x)$. Soit z un autre sommet de $V(T_x)$. Alors il existe un sommet t tel que $P_v \cap P_z = P_t$, et $\mathcal{B}_1^+(v) = \mathcal{B}_1^+(t) \cup \{v, v_{-1}, v_{-2}, ..., v_{-i+1}\}$. Comme $u \in \mathcal{B}_1^+(v)$ et u n'est pas dans l'ensembme $\{v, v_{-1}, v_{-2}, ..., v_{-i+1}\}$, $u \in \mathcal{B}_1^+(t)$. Mais $\mathcal{B}_1^+(t)$ est inclus dans $\mathcal{B}_1^+(z)$, donc $\overrightarrow{uz}$ est un arc de D.

Soit H le graphe dont les sommets sont les classes d'équivalence T_x et avec un arc entre T_x et T_y s'il y au moins un arc entre un sommet de T_x et un sommet de T_y dans D. H est un graphe orienté sans arcs symétriques et $D \in \Psi(H)$. □

Nous concluons cette partie par les deux remarques suivantes :
- La preuve précédente redonne la classification des graphes orientés finis, sans arcs symétriques, tels que $\overrightarrow{\gamma}^{\text{ID}}(D) = |V(D)|$. Cependant, le théorème 2.11 est

plus fort car concerne tous les graphes orientés et nous permet notamment d'obtenir le corollaire 2.12.
– Les graphes orientés dont le seul ensemble séparant est l'ensemble des sommets du graphe sont les graphes de $\Psi(H)$ pour lesquels H n'a pas de sommet source.

2.3 Cas non orienté

Nous nous intéressons maintenant aux graphes non orientés. Dans un premier temps, nous étudions les graphes finis. Classifier l'ensemble des graphes G pour lesquels $\gamma^{\text{ID}}(G) = |V(G)| - 1$ avait déjà été étudié et a fait l'enjeu d'une conjecture et de questions que nous rappelons en début de partie. Nous donnons ensuite la description complète de cette famille de graphe. Enfin, nous étudions le cas des graphes infinis et le cas des codes r-identifiants.

2.3.1 Graphes finis

Préliminaires

Dans le cas des graphes finis non orientés, il existe dans tout graphe sans jumeaux possédant au moins une arête un code identifiant de taille $n-1$ où n est le nombre de sommets (théorème 1.4). Autrement dit, il est toujours possible d'enlever à un graphe sans jumeaux un sommet et que le graphe reste sans jumeaux.

Plusieurs preuves de ce résultat existent, la première étant attribuée à Bertrand [9] ne semble pas accessible, la plus courte revient à Gravier et Moncel [48]. Nous exprimons ici ce résultat comme une conséquence directe du corollaire 2.3 du théorème de Bondy.

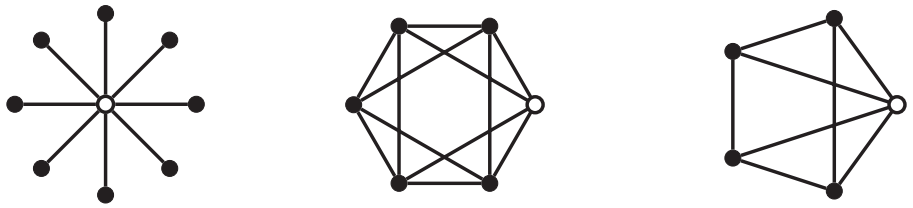
Théorème 2.22 (Bertrand [9] et Gravier, Moncel [48])
Soit G un graphe sans jumeaux possédant au moins une arête. Alors :

$$\gamma^{ID}(G) \leq |V(G)| - 1.$$

Preuve : Si le graphe G n'est pas connexe, un code identifiant de G peut être formé en réunissant des codes identifiants de chaque composante connexe. Nous supposons donc que G est connexe. En appliquant le corollaire 2.3 sur le graphe biparti $B(G)$ d'incidence des voisinages fermés de G, il existe un code séparant de $B(G)$ de taille $|V(G)| - 1$ qui ramené à G donne un code séparant de G de taille $|V(G)| - 1$. Ce code est un ensemble dominant car le graphe G est connexe. C'est donc un code identifiant. □

Cette borne est serrée, par exemple pour les graphes suivants (voir la figure 2.5) :
– les étoiles $K_{1,n}$, pour $n \geq 2$;
– les graphes complets privés d'un couplage maximal, notés $K_n \setminus M$.

Charbit *et al.* [16] ont conjecturé que ce sont les seuls graphes connexes atteignant la borne. Dans sa thèse [86], Skaags pose les questions suivantes :



$\gamma^{\text{ID}}(K_{1,8}) = 8$ $\gamma^{\text{ID}}(K_6 \setminus M) = 5$ $\gamma^{\text{ID}}(K_5 \setminus M) = 4$

Figure 2.5 – *Exemples de graphes pour lesquels $\gamma^{ID}(G) = |V(G)| - 1$.*

Question 2.23 (Skaggs [86]) 1. *Existe-t-il des graphes k-réguliers G d'ordre n avec $\gamma^{ID}(G) = n - 1$ pour $k < n - 2$?*

2. *Existe-t-il des graphes G d'ordre n impair et de degré maximum $\Delta < n - 1$ avec $\gamma^{ID}(G) = n - 1$?*

Nous allons voir dans la suite qu'il existe tout une autre classe de graphes atteignant la borne : les graphes de la forme P_{2k}^{k-1}, et qu'il est possible de faire des opérations entre ces graphes tout en continuant d'atteindre la borne. Cela nous donnera une classification complète permettant de répondre par la négative aux questions précédentes.

Outils pour la classification

En utilisant le graphe biparti d'incidence des voisinages fermés d'un graphe G (voir la partie 1.1.7), les résultats sur les codes discriminants nous donnent directement des résultats sur les codes identifiants. La proposition 2.6 se traduit pour les graphes non orientés de la manière suivante :

Proposition 2.24 *Soit G un graphe fini non orienté sans jumeaux et soit S un ensemble de sommets de G tel que $G - S$ soit sans jumeaux. Alors*

$$\gamma^{ID}(G) \leq \gamma^{ID}(G - S) + |S|.$$

La preuve est similaire à celle de la proposition 2.7 dans le cas orienté et est omise ici. Notons que ce résultat implique qu'en enlevant des sommets à un graphe, la taille du code identifiant optimal diminue d'au plus du nombre de sommets ajoutés. Charon, Hudry et Lobstein [23] ont récemment étudié les variations de $\gamma^{\text{ID}}(G)$ lorsque un sommet de G est supprimé et ont montré que l'augmentation de $\gamma^{\text{ID}}(G)$ après une suppression d'un sommet pouvait être arbitrairement grande.

Nous utiliserons la proposition 2.24 surtout sous la forme du corollaire suivant :

Corollaire 2.25 *Soit G un graphe connexe avec $\gamma^{ID}(G) = |V(G)| - 1$ et $|V(G)| \geq 4$. Si x est un sommet de G tel que $G - x$ soit sans jumeaux et contiennent au moins une arête, alors $\gamma^{ID}(G - x) = |V(G - x)| - 1$. De plus, il existe un tel sommet x pour lequel $G - x$ est connexe.*

Preuve : Supposons qu'il existe un sommet x de G tel que $G-x$ soit sans jumeaux et contiennent au moins une arête. Alors comme $G-x$ n'est pas le graphe $\overline{K_n}$, $\gamma^{\text{ID}}(G-x) \leq |V(G-x)|-1$ (théorème 2.22). D'après la proposition 2.24, $\gamma^{\text{ID}}(G) \leq \gamma^{\text{ID}}(G-x)+1 \leq |V(G)|-1$. Comme d'autre part $\gamma^{\text{ID}}(G) = |V(G)|-1$, nous avons $\gamma^{\text{ID}}(G-x) = |V(G-x)|-1$.

Si G est une étoile $K_{1,t}$ avec $t > 2$, alors n'importe quelle feuille x vérifie $G-x$ connexe et sans jumeaux.

Supposons donc que G ne soit pas une étoile. D'après le théorème 2.22, il existe un sommet x de G tel que $V(G-x)$ soit un code identifiant. Cela signifie donc que $G-x$ est sans jumeaux. Comme $G-x$ n'est pas le graphe $\overline{K_n}$, nous avons $\gamma^{\text{ID}}(G-x) = |V(G-x)|-1$.

Nous montrons maintenant que x peut être choisi de manière à ce que $G-x$ soit connexe. Supposons que ce ne soit pas le cas, alors comme $\gamma^{\text{ID}}(G-x) = |V(G-x)|-1$, toutes les composantes de $G-x$ sauf une sont des sommets isolés. Il y a donc au moins un sommet isolé y. Si $G-y$ est sans jumeaux, alors y peut être choisi à la place de x. Sinon il existe un sommet x' tel que $N[x] = N[x'] \cup \{y\}$. Alors $G-x'$ est connexe et sans jumeaux. □

Lemme 2.26 *Soit G un graphe sans jumeaux et soit v un sommet de G. Supposons que x et y soient deux jumeaux de $G-v$. Alors si $G-x$ ou $G-y$ a une paire de jumeaux, v est nécessairement un des sommets de la paire.*

Preuve : Comme v sépare x et y, v est adjacent à exactement l'un deux sommets, disons x. Supposons qu'il y ait des jumeaux z et t dans $G-x$ et que z soit adjacent à x et t ne le soit pas. Si $z \neq v$ alors y est aussi adjacent à z, et donc t est adjacent à y, ce qui implique que t est adjacent à x. Cela contredit le fait que x sépare z et t.

L'autre cas est prouvé de la même manière. □

La classe des graphes extrémaux est dans une certaine mesure stable par l'opération joint $\bowtie$.

Proposition 2.27 *Soient G_1 et G_2 des graphes sans jumeaux tels que pour tout ensemble séparant S de taille minimale il y ait un sommet qui soit adjacent à tous les sommets de S. Si $G_1 \bowtie G_2$ est sans jumeaux, alors*

$$\gamma^S(G_1 \bowtie G_2) = \gamma^S(G_1) + \gamma^S(G_2) + 1.$$

De plus, pour tout ensemble séparant S de $G_1 \bowtie G_2$ de taille $\gamma^S(G_1) + \gamma^S(G_2) + 1$, il y a un sommet de $G_1 \bowtie G_2$ adjacent à tout l'ensemble S.

Preuve : Soient S'_1 un ensemble séparant G_1 et S'_2 un ensemble séparant G_2, tous deux de taille minimale. Soit x_1 (resp. x_2) un sommet de G_1 (resp. G_2) adjacent à tout S'_1 (resp. S'_2). Le graphe $G_1 \bowtie G_2$ est sans jumeaux donc il existe un sommet y qui sépare x_1 et x_2 et alors $S'_1 \cup S'_2 \cup \{y\}$ est un ensemble séparant de $G_1 \bowtie G_2$, et donc $\gamma^{\text{S}}(G_1 \bowtie G_2) \leq \gamma^{\text{S}}(G_1) + \gamma^{\text{S}}(G_2) + 1$.

Soit maintenant S un ensemble séparant de taille minimale de $G_1 \bowtie G_2$. Les sommets de G_2 ne séparent aucun sommet de G_1, donc $S_1 = S \cap V(G_1)$ est un ensemble séparant de G_1. De la même manière, $S_2 = S \cap V(G_2)$ est un ensemble

séparant de G_2. Nous avons donc $|S| \geq \gamma^{\mathrm{S}}(G_1) + \gamma^{\mathrm{S}}(G_2)$. Comme $|S| \leq \gamma^{\mathrm{S}}(G_1) + \gamma^{\mathrm{S}}(G_2) + 1$, au moins un des deux ensembles S_1 ou S_2 est un ensemble séparant minimal. Supposons que S_1 soit un séparant de taille minimale. Il existe donc un sommet x_1 dans G_1 qui est adjacent à tout S_1, et par conséquent à tout l'ensemble S dans $G_1 \bowtie G_2$. Par conséquent S_2 ne peut pas être un ensemble minimal sinon il y aurait un autre sommet x_2 adjacent à tout S. Cela montre que $|S| = \gamma^{\mathrm{S}}(G_1) + \gamma^{\mathrm{S}}(G_2) + 1$ et qu'il existe un sommet du graphe qui est adjacent à tout l'ensemble S. □

Dans la proposition 2.27, si les graphes G_1 et G_2 ont au moins deux sommets, alors $\gamma^{\mathrm{ID}}(G_1 \bowtie G_2) = \gamma^{\mathrm{S}}(G_1 \bowtie G_2) = \gamma^{\mathrm{S}}(G_1) + \gamma^{\mathrm{S}}(G_2) + 1$ car la domination est assurée.

Description de la famille extrémale

Nous noterons dans la suite A_k le graphe P_{2k}^{k-1} lorsque $k \geq 2$, et $A_1 = \overline{K_2}$. La figure 2.6 montre les graphes A_k pour $k \leq 3$. Les sommets de A_k seront notés $x_1, ...x_{2k}$ dans l'ordre naturel donné par le chemin. Le seul automorphisme non trivial de A_k est celui qui envoie x_i sur x_{2k+1-i}. Le graphe A_k est sans jumeaux et lorsque $k \geq 2$, A_k est de degré maximum $2k+2$, connexe et $\overline{A_k}$ est aussi connexe.

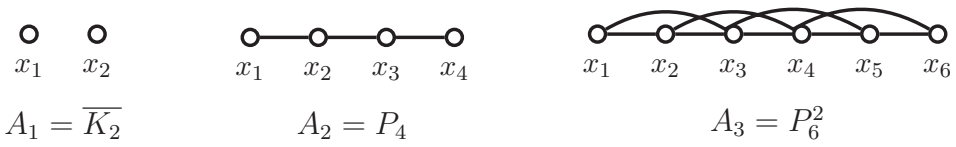


Figure 2.6 – *Les graphes A_1, A_2 et A_3.*

Les graphes A_k sont des graphes extrémaux pour les codes identifiants :

Proposition 2.28 *Pour tout $k \geq 1$, $\gamma^{S}(A_k) = 2k-1$. De plus, $N[x_k]$ et $N[x_{k+1}]$ sont les seuls ensembles séparants de taille $2k-1$ de A_k. Pour $k \geq 2$, $\gamma^{ID}(A_k) = 2k-1 = |V(A_k)| - 1$.*

Preuve : Soit S un ensemble séparant de A_k. Pour $i < k$, x_{i+k} est le seul élément dans la différence symétrique de x_i et x_{i+1}, donc $x_{i+k} \in S$. Par symétrie, on a donc $\{x_2, ..., x_{2k-1}\} \subseteq S$.

Pour séparer x_k et x_{k+1}, il faut ajouter x_1 ou x_{2k}. Les ensembles $V(A_k) \setminus \{x_1\} = N[x_{2k}]$ et $V(A_k) \setminus \{x_{2k}\} = N[x_1]$ sont des ensembles séparants A_k, de taille $2k-1$ et ceux sont donc les seuls.

Lorsque $k > 1$, ce sont des ensembles dominants et donc identifiants. □

Corollaire 2.29 *Soit $k \geq 1$. Pour tout ensemble séparant S de taille minimale de A_k, il existe un sommet qui est adjacent à tout l'ensemble S.*

Soit $\mathcal{A}$ la fermeture de l'ensemble $\{A_i \mid i = 1, 2, \ldots\}$ par l'opération $\bowtie$.

Proposition 2.30 *Soit $G \in \mathcal{A}$. Alors $\gamma^S(G) = |V(G)| - 1$. De plus, pour tout ensemble minimum S séparant G, il existe un sommet qui est adjacent à tous les sommets de S.*

Preuve : La proposition est vraie pour les éléments de base de $\mathcal{A}$ d'après la proposition 2.28 et d'après le corollaire 2.29. Elle reste vraie au passage de l'opération $\bowtie$ d'après la proposition 2.27. □

Corollaire 2.31 *Pour tout graphe $G \in \mathcal{A}$, $G \neq A_1$,*

$$\gamma^{ID}(G) = |V(G)| - 1.$$

Nous pouvons obtenir encore d'autres exemples de graphes extrêmes en ajoutant un sommet universel à un graphe de $\mathcal{A}$:

Proposition 2.32 *Pour tout graphe G de $\mathcal{A} \bowtie K_1$, nous avons :*

$$\gamma^{ID}(G) = \gamma^S(G) = |V(G)| - 1.$$

Preuve : Soit $G = G_1 \bowtie K_1$ avec $G_1 \in \mathcal{A}$. Notons u le sommet universel.

Soit S un ensemble séparant G de taille minimum. Nécessairement, $S \subseteq V(G_1)$ car u ne peut pas séparer de sommets. De plus S est un ensemble séparant G_1, donc $|S| \geq |V(G_1)| - 1$.

Si $|S| = |V(G_1)| - 1$, alors d'après la proposition 2.30, il existe un sommet x dans G_1 qui est adjacent à tout S. Alors x et u ne sont pas séparés. Donc $S = V(G_1)$ et $\gamma^S(G) = |V(G)| - 1$. S est un ensemble dominant de G, c'est donc aussi un code identifiant, ce qui termine la preuve. □

Nous avons mentionné dans le paragraphe précédent qu'un graphe complet privé d'un couplage maximum était aussi extrémal. Ce graphe se retrouve dans notre construction : lorsque n est pair, cela correspond au joint de $\frac{n}{2}$ graphes A_1, et lorsque n est impair, le graphe s'obtient en ajoutant à la construction précédente un sommet universel.

Jusqu'à présent, les graphes connexes G tels que $\gamma^{ID}(G) = |V(G)| - 1$ rencontrés sont donc les suivants :

- les étoiles $K_{1,t}$, pour $t \geq 2$;
- la famille $\mathcal{A}$ privée de A_1 ;
- la famille $\mathcal{A} \bowtie K_1$.

Il n'y a pas d'autres graphes extrémaux

Théorème 2.33 *Soit G un graphe connexe sans jumeaux. Nous avons $\gamma^{ID}(G) = |V(G)| - 1$ si et seulement si $G \in \{K_{1,t} \mid t \geq 2\} \cup \mathcal{A} \cup (\mathcal{A} \bowtie K_1)$ et $G \neq A_1$.*

Preuve : Remarquons tout d'abord que les seuls graphes sans sommet universel parmi les graphes de $\{K_{1,t} \mid t \geq 2\} \cup \mathcal{A} \cup (\mathcal{A} \bowtie K_1)$ sont ceux appartenant à la famille $\mathcal{A}$.

La partie “si” du théorème a déjà été montrée. D'après ce qu'il précède, il reste seulement à montrer que si un graphe G connexe vérifie $\gamma^{\text{ID}}(G) = |V(G)| - 1$, alors $G \in \{K_{1,t} \mid t \geq 2\} \cup \mathcal{A} \cup (\mathcal{A} \bowtie K_1)$ et $G \neq A_1$.

Procédons par l'absurde. Soit G le plus petit graphe connexe (en nombre de sommets), sans jumeaux, tel que $\gamma^{\text{ID}}(G) = |V(G)|-1$ et $G \notin \{K_{1,t} \mid t \geq 2\} \cup \mathcal{A} \cup (\mathcal{A} \bowtie K_1)$. Nécessairement, G a plus de cinq sommets.

D'après le corollaire 2.25, il existe un sommet x de G tel que $G - x$ soit connexe, sans jumeaux et $\gamma^{\text{ID}}(G-x) = |V(G-x)|-1$. Par minimalité de G, nous avons donc $G - x \in \{K_{1,t} \mid t \geq 2\} \cup \mathcal{A} \cup (\mathcal{A} \bowtie K_1)$, $G - x \neq A_1$. Nous considérons trois cas selon l'ensemble auquel $G - x$ appartient.

Cas 1, $G - x \in \{K_{1,t} \mid t \geq 2\}$.

Comme G a plus de cinq sommets, nous avons en fait $t \geq 3$. Le graphe G n'est pas une étoile, donc x est adjacent à au moins une feuille y. Soit C un code identifiant de $G - x$ contenant y et le sommet central u de $G - x$ (c'est toujours possible de trouver un tel code identifiant). L'ensemble C est un ensemble dominant qui n'est pas un code identifiant de G. Ce n'est donc pas un ensemble séparant et x et un des sommets z ne sont pas séparés par C.

Le sommet z contient nécessairement y dans son voisinage fermé. C'est donc obligatoirement le sommet u ou le sommet y. Comme les deux sommets sont dans C, x est adjacent à u. Si x n'est pas séparé de y, cela signifie que x est adjacent à l'unique feuille v de $G - x$ qui n'est pas dans le code C et à aucune autre feuille de $G - x$. Si x et u ne sont pas séparés, alors cela signifie que x est adjacent à tous les sommets de $G - x$ sauf le sommet v qui n'est pas dans le code. Dans les deux cas, en remplaçant v par u dans le code C, C devient un code identifiant de G de taille $|V(G)| - 2$, ce qui mène à une contradiction.

Cas 2, $G - x \in \mathcal{A}$.

Nous considérons deux sous-cas : soit $G - x = A_k$ pour un certain $k \geq 2$, soit $G - x = G_1 \bowtie G_2$, avec $G_1, G_2 \in \mathcal{A}$.

(1) $G - x = A_k$, pour $k \geq 2$. Le graphe G n'est pas un graphe de la famille $\mathcal{A} \bowtie K_1$, donc x n'est pas adjacent à tous les sommets de A_k. Il existe donc deux sommets consécutifs de A_k (par rapport au chemin P_{2k}), x_i et x_{i+1}, tels que l'un soit adjacent à x et l'autre non. Par symétrie de A_k, nous pouvons supposer que $i \leq k$. Nous allons montrer que $C = V(G) \setminus \{x_{2k}, x\}$, $C' = V(G) \setminus \{x_1, x\}$ ou bien $C'' = V(G) \setminus \{x_k, x_{k+1}\}$ est un code identifiant de G, ce qui sera contradictoire.

Notons tout d'abord que C est un code identifiant de A_k et que x_i et x_{i+1} sont dans C. Si x n'est pas dominé par C, cela signifie que x est seulement adjacent à x_{2k}, et alors C' est code identifiant de G. Nous pouvons donc supposer que x est dominé par C.

Si x est séparé de tous les sommets de $G - x$ par C, alors C est un code identifiant. Sinon il y a deux possibilité selon si x est adjacent à x_i ou à x_{i+1}. Dans la première possibilité, x est adjacent à x_{i+1} mais pas à x_i. Dans ce cas, chaque sommet x_j avec $j > i+k$ est séparé de x par x_{i+1} et chaque sommet x_j avec $j < i + k$ est séparé de x par x_i. Le seul sommet qui n'est pas séparé de

x est le sommet x_{i+k}. Alors x est exactement adjacent aux sommets allant de x_{i+1} à x_{2k-1}. Si C' n'est pas un code identifiant, cela signifie donc, en utilisant un raisonnement similaire, que x_k n'est pas séparé de x et donc que $i = 1$. Alors C'' est un code identifiant de G.
Dans la deuxième possibilité, x est adjacent à x_i mais pas à x_{i+1} et alors x est séparé de tous les sommets excepté peut-être x_1 lorsque $i = k$. Alors C' est un code identifiant.

(2) $G-x = G_1 \bowtie G_2$ avec $G_1, G_2 \in \mathcal{A}$. Le graphe G n'est pas dans $\mathcal{A} \bowtie K_1$, donc x n'est pas adjacent à tous les sommets de $G - x$. Il existe donc un sommet y qui n'est pas adjacent à x. Sans perte de généralité, $y \in V(G_1)$.
Soit C_1 un code identifiant de G_1 de taille $\gamma^{\text{ID}}(G_1) = |V(G_1)| - 1$ qui contient y. L'existence d'un tel code provient de la symétrie des graphes A_k et de la proposition 2.28.
L'ensemble $C = C_1 \cup V(G_2)$ est un code identifiant $V(G_1 \bowtie G_2)$ de taille $|V(G)| - 2$. Ce n'est donc pas un code identifiant de G. C est nécessairement un ensemble dominant x, et donc un sommet z de $G_1 \bowtie G_2$ n'est pas séparé de x. Comme $y \in C$, z n'est pas adjacent à y, et donc $z \in V(G_1)$.
Cela implique que z est adjacent à tous les sommets de G_2, donc x est aussi adjacent à tous les sommets de G_2. Nous pouvons donc écrire $G = (G_1') \bowtie G_2$ avec G_1' le graphe obtenu à partir de G_1 en ajoutant x et les arêtes reliant x à G_1. Un code identifiant minimum de G_1' accompagné de tous les sommets de G_2 forme un code identifiant de G. Cela signifie donc que $\gamma^{\text{ID}}(G_1') = |V(G_1')| - 1$. Par minimalité de G, le graphe G_1' ayant strictement moins de sommets que G, est un des graphes de l'ensemble $\{K_{1,t} \mid t \geq 2\} \cup \mathcal{A} \cup (\mathcal{A} \bowtie K_1)$, et $G_1' \neq A_1$. Or $G_1 \in \mathcal{A}$, et x n'est pas adjacent à l'un des sommets de G_1, donc G_1' n'a pas de sommet universel. Le graphe $G_1' + x$ est donc lui aussi dans la famille $\mathcal{A}$. Cela n'est pas possible car les graphes de $\mathcal{A}$ ont tous un nombre pair de sommets.

Cas 3, $G - x \in \mathcal{A} \bowtie K_1$.

Supposons que $G - x = A_{i_1} \bowtie A_{i_2} \bowtie \ldots \bowtie A_{i_s} \bowtie K_1$ et notons u le sommet correspondant à K_1.

Si x est adjacent à u, alors u est un sommet universel dans G, et $G - u$ est sans jumeaux. Le graphe G a au moins cinq sommets, donc $G - \{x, u\} = A_{i_1} \bowtie A_{i_2} \bowtie \ldots \bowtie A_{i_s}$ est connexe. Si $G - u$ n'est pas connexe, cela signifie que x est seulement adjacent à u. Soit alors C un code identifiant de $G-\{x, u\}$ de taille $|V(G-\{x, u\})|-1$. Le code $C \cup \{x\}$ est un code identifiant G de taille $|V(G)| - 2$, ce qui contredit le fait que $\gamma^{\text{ID}}(G) = |V(G)| - 1$. Nous pouvons donc supposer que $G - u$ est connexe. D'après le corollaire 2.25, $\gamma^{\text{ID}}(G - u) = |V(G - u)| - 1$ et par minimalité de G, $G - u \in \{K_{1,t} \mid t \geq 2\} \cup \mathcal{A} \cup (\mathcal{A} \bowtie K_1)$. Si $G - u \in \{K_{1,t} \mid t \geq 2\} \cup (\mathcal{A} \bowtie K_1)$, il y aurait deux sommets universels, qui seraient donc jumeaux. Ainsi $G - u \in \mathcal{A}$ et $G \in \mathcal{A} \bowtie K_1$, ce qui est contradictoire.

Nous pouvons donc supposer que x n'est pas adjacent à u. Si le graphe $G - u$ est sans jumeaux, nous reprenons le même raisonnement. Dans ce cas, si $G - u \in \{K_{1,t} \mid t \geq 2\} \cup \mathcal{A}$, nous appliquons le cas 1 ou 2. Si $G - u \in \mathcal{A} \bowtie K_1$ et que u' désigne le sommet universel de $G - u$, alors u' n'est pas universel dans G (sinon on

applique le même raisonnement). Les sommets u et u' induisent alors une copie de A_1 et $G \in \mathcal{A} \bowtie A_1 = \mathcal{A}$.

Si $G - u$ contient des jumeaux, alors x doit être un des sommets jumeaux car tous les autres sommets sont adjacents à u. Soit x' son jumeau et supposons que $x' \in V(A_{i_1})$. Notons $\{z_1, ..., z_{2k}\}$ les sommets de A_{i_1}. Sans perte de généralité, supposons que $x' = z_\ell$ avec $\ell \leq k$. Si $\ell \geq 2$, nous allons montrer que $C = V(G) \setminus \{z_\ell, z_{2k}\}$ est un code identifiant de G, ce qui est contradictoire. Pour cela, nous remarquons tout d'abord que les sommets de $A_{i_2} \bowtie \cdots \bowtie A_{i_s}$ sont déjà identifiés entre eux et avec les autres sommets du graphe par C. Chaque paire de sommets de A_{i_1} est séparée par un sommet de $V(A_{i_1}) \cap C$, sauf $z_{\ell+k-1}$ et $z_{\ell+k}$ qui sont séparés par x. Le sommet x est aussi séparé des sommets de A_{i_1} : il est séparé des sommets de $\{z_1, \ldots, z_{l+k-1}\}$ par x et des sommets de $\{z_{k+1}, \ldots, z_{2k}\}$ par z_1 ($\ell \geq 2$).

Cela signifie donc que $x' = x_1$. Alors le sous-graphe induit par $V(A_{i_1})$, u et x est isomorphe à A_{i_1+1}, et donc $G = A_{i_1+1} \bowtie A_{i_2} \bowtie \ldots \bowtie A_{i_s}$, ce qui mène à une contradiction. □

Le théorème 2.33 donne une réponse négative aux questions 2.23. En effet, pour la première question, les seuls graphes réguliers de la classification sont les graphes complets d'ordre pair privés d'un couplage parfait (qui correspondent aux copies disjointes de A_1), et sont $(n-2)$-régulier avec $n = |V(G)|$.

Quant à la deuxième question, chaque graphe de $\mathcal{A}$ est d'ordre pair, les graphes d'ordre impair de la caractérisation ne sont pas dans la famille $\mathcal{A}$ et ont donc tous un sommet universel : $\Delta(G) = n - 1$.

Nous avons en fait le résultat suivant qui se déduit du théorème 2.33 :

Corollaire 2.34 *Soit G un graphe fini connexe sans jumeaux d'ordre n de degré maximum Δ. Si $\gamma^{\text{ID}}(G) = n - 1$, alors $\Delta(G) \geq n - 2$ si n est pair et $\Delta(G) = n - 1$ si n est impair.*

Ce corollaire donne du sens à la conjecture 1.5. En effet, les graphes exhibés dans cette partie ayant un degré maximum élevé, il est naturel de chercher une borne supérieure sur $\gamma^{\text{ID}}(G)$ utilisant le degré maximum.

2.3.2 Graphes infinis

Dans Charon *et al.* [22], il est montré que le théorème 2.22 n'a pas d'extension directe pour les graphes infinis non orientés. En effet, il existe des graphes infinis sans jumeaux non triviaux pour lesquels le seul code identifiant est l'ensemble de tous les sommets du graphe. Dans cette partie, nous classifions tous les graphes infinis non orientés ayant cette propriété. Cela renforce un théorème de Gravier et Moncel [48] comme quoi aucun de ces graphes n'a tous ses sommets de degré fini.

Nous commençons par définir un tel graphe infini, qui sera notre exemple de base. Ce graphe a originellement été défini dans [22].

Soient $X = \{\ldots, x_{-1}, x_0, x_1, \ldots\}$ et $Y = \{\ldots, y_{-1}, y_0, y_1, \ldots\}$ des copies de $\mathbb{Z}$. Nous notons A_∞ le graphe dont l'ensemble des sommets est $X \cup Y$ et dont l'ensemble des arêtes E est :

$$E = \{x_i x_j \mid i \neq j\} \cup \{y_i y_j \mid i \neq j\} \cup \{x_i y_j \mid i < j\}.$$

Les sommets de X induisent une clique infinie, tout comme ceux de Y. De plus, le graphe induit par $\{y_1, y_2, \ldots, y_k, x_1, x_2 \ldots, x_k\}$ est isomorphe au graphe $A_k = P_{2k}^{k-1}$. La figure 2.7 donne une illustration de ce graphe.

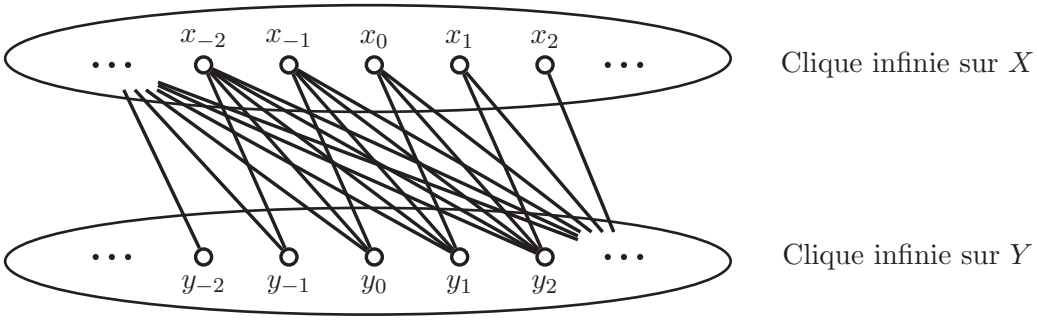


Figure 2.7 – *Le graphe A_∞. Son seul code identifiant est l'ensemble de tous ses sommets.*

Charon *et al.* [22] ont montré que le seul ensemble séparant de A_∞ est $V(A_\infty)$, c'est donc aussi le cas pour les codes identifiants. Nous rappelons ici le raisonnement menant à cette affirmation. Pour tout $i \in \mathbb{Z}$, x_i et x_{i+1} sont séparés uniquement par le sommet y_{i+1}, tandis que x_i est le seul sommet qui sépare y_i de y_{i+1}. Ainsi tous les sommets du graphe doivent être dans l'ensemble séparant.

Ce raisonnement reste vrai si nous ajoutons un nouveau sommet au graphe qui est adjacent à tous les sommets de X ou à aucun, et de manière similaire pour Y. Cela mène à la famille suivante :

Soit H un graphe simple fini ou infini muni d'un couplage parfait ρ considéré comme une involution $x \to \rho(x)$ de $V(H)$, où $x\rho(x)$ est une arête de H. À partir de H et de ρ, nous construisons le graphe $\Psi(H, \rho)$ de la manière suivante.

- Pour chaque sommet x de H nous assignons une copie $\Phi(x)$ de $\mathbb{Z}$.
- L'ensemble des sommets de $\Psi(H, \rho)$ est $\bigcup_{x \in V(H)} \Phi(x)$.
- Pour chaque arête du couplage $x\rho(x)$ de H, les arêtes entre $\Phi(x)$ et $\Phi(\rho(x))$ sont celles de A_∞.
- Pour toutes les autres arêtes xy de H, tous les sommets de $\Phi(x)$ et de $\Phi(y)$ sont adjacents.

Un exemple d'une telle construction est donné dans la figure 2.8.

Tous les graphes ainsi construits ne sont identifiables qu'avec tous les sommets du graphe :

Proposition 2.35 *Pour tout graphe simple H, fini ou infini, possédant un couplage parfait ρ, le graphe $\Psi(H, \rho)$ n'est identifié que par l'ensemble $V(\Psi(H, \rho))$.*

Preuve : Soit A_x la copie de A_∞ qui correspond à l'arête $x\rho(x)$. Alors pour tout sommet y de $V(\Psi(H, \rho)) \setminus V(A_x)$, soit y est adjacent à tous les sommets de $\Phi(x)$, soit à aucun d'entre eux. Aussi, deux sommets de $\Phi(x)$ ne peuvent être identifiés que par des sommets de $V(A_x)$ (en pratique par des sommets de $\Phi(\rho(x))$). Deux sommets x_i et x_{i+1} de $\Phi(x)$ sont alors uniquement séparés par le sommet y_{i+1} correspondant dans $\Phi(\rho(x))$, et donc tous les sommets de $\Phi(\rho(x))$ sont dans tout ensemble séparant. De la même manière, tous les sommets de $\Phi(x)$ sont dans tous les ensembles séparants.

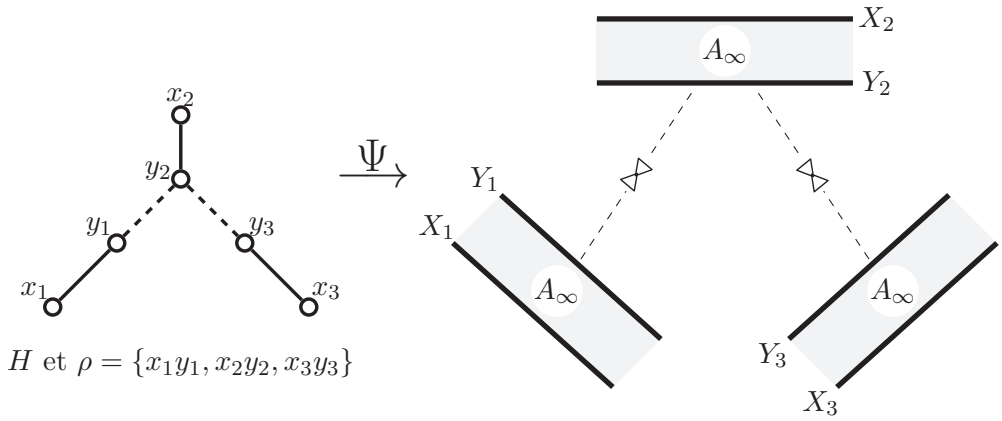


Figure 2.8 – *Construction de $\Psi(H, \rho)$ à partir de (H, ρ).*

Comme x a été choisi de manière arbitraire, cela signifie que dans un ensemble séparant, tous les sommets de $V(\Psi(H, \rho))$ sont présents. □

Tout graphe infini connexe dont le seul code identifiant est l'ensemble de ses sommets est construit de cette manière.

Théorème 2.36 *Soit G un graphe infini connexe sans jumeaux. Le seul code identifiant de G est l'ensemble de tous ses sommets si et seulement si $G = \Psi(H, \rho)$ où H est un graphe fini ou infini et ρ est couplage parfait de H.*

Preuve : La partie « si » du théorème est donnée par la proposition 2.35.

Soit G un graphe infini connexe, sans jumeaux, et tel que le seul code identifiant de G soit $V(G)$. Pour tout sommet v de G, $G - v$ possède donc nécessairement une paire de jumeaux. Pour montrer la partie « seulement si » du théorème, il suffit de montrer que tout sommet v de G appartient à un unique sous-graphe induit A_v isomorphe à A_∞ et que si un sommet hors de A_v est adjacent à une des parties X ou Y de A_v alors il est adjacent à toute la partie correspondante.

Soit x_1 un sommet de G, le graphe $G - x_1$ a une paire de jumeaux, y_1 et y_2. Sans perte de généralité, x_1 est adjacent à y_1 mais pas à y_2. D'après le lemme 2.26, x_1 doit être un des sommets d'une paire de jumeaux de $G - y_2$. Soit x_2 le jumeau de x_1 dans $G - y_2$. Notons que le choix de x_2 est unique. Dans le graphe $G - y_1$, il y a aussi une paire de jumeaux, et x_1 doit être l'un deux, notons x_0 l'autre (là aussi, le choix de x_0 est unique).

En continuant ce processus dans les deux directions (abscisse positives et négatives), nous construisons le graphe A_{x_1} isomorphe à A_∞ comme sous graphe de G. Comme chaque paire de sommets consécutifs de $X \subseteq V(A_{x_1})$ est séparée par un unique sommet dans $Y \subseteq V(A_{x_1})$, tous les sommets dans X sont jumeaux dans $G - Y$. Chaque sommet qui n'est pas dans $V(A_{x_1})$ est donc soit adjacent à tous les sommets de X soit à aucun d'entre eux. De manière similaire, chaque sommet qui n'est pas dans $V(A_{x_1})$ est adjacent à tous les sommets de Y ou à aucun.

La construction de A_{x_1} est unique, ce qui termine la preuve. □

2.3.3 Conséquences

Le théorème 2.36 nous dit en particulier que si un graphe G ne contient pas A_∞ comme sous-graphe induit, alors il existe un sommet y tel que $V(G) \setminus \{y\}$ soit un code identifiant. Nous pouvons en dire plus lorsque nous interdisons le graphe A_∞^+, c'est-à-dire, le graphe A_∞ restreint aux sommets d'indices positifs dans X et Y. C'est le résultat du lemme suivant, qui dit qu'alors pour chaque sommet x de G il existe un sommet y *dans le voisinage fermé de* x tel que $V(G) \setminus \{y\}$ soit un code identifiant.

Ce résultat est plus fort que le théorème 2.22 puisque nous avons une information sur le sommet que l'on peut retirer. Il est attribué à Bertrand [9]. Nous donnons une preuve indépendante ici car le manuscrit [9] n'est pas accessible.

Lemme 2.37 (Bertrand [9]) *Soit G un graphe sans jumeaux, infini ou non, ne contenant pas A_∞^+ comme sous-graphe induit. Alors, pour tout sommet x de G, il existe un sommet $y \in N[x]$ tel que $G - y$ soit sans jumeaux.*

Preuve : Supposons par l'absurde qu'il existe un sommet x_1 de G tel que pour tout $y \in N[x_1]$, $G - y$ contiennent des jumeaux.

En particulier, $G - x_1$ a une paire de jumeaux : y_1,y_2 et sans perte de généralité, x_1 est adjacent à y_2 mais pas à y_1. Alors $y_2 \in N[x_1]$, donc $G - y_2$ a une autre paire de jumeaux, u, u'. D'après le lemme 2.26, x_1 est l'un des sommets u, u'. Nous appelons l'autre élément x_2. Alors $x_2 \in N[x_1]$. Le sous-graphe induit par x_1, x_2, y_1, y_2 est isomorphe à A_2.

Nous montrons par induction que A_∞^+ est un sous graphe induit de G, obtenant une contradiction. Pour cela, supposons que A_k soit induit par les sommets $\{y_1, \ldots, y_k, x_1, \ldots, x_k\}$ de telle sorte que x_{k-1}, x_k soient jumeaux dans $G - y_k$ et que y_{k+1}, y_k soient jumeaux dans $G - x_{k-1}$.

Alors $x_k \in N[x_1]$. Donc $G - x_k$ contient une paire de jumeaux, et d'après le lemme 2.26, y_k doit être l'un d'eux. Notons y_{k+1} l'autre. Comme y_k et y_{k+1} sont jumeaux dans $G - x_k$, y_{k+1} est adjacent à $x_1, \ldots, x_k$ et à $y_1, \ldots, y_k$. En particulier, $y_{k+1} \in N[x_1]$ donc $G - y_{k+1}$ contient une paire de jumeaux et ici encore, x_k doit être l'un d'eux. Notons x_{k+1} l'autre. Alors x_{k+1} est adjacent à $x_1, \ldots, x_k$ et n'est pas adjacent à $y_1, \ldots, y_k$. Le sous-graphe induit par $\{y_1, \ldots, y_{k+1}, x_1, \ldots, x_{k+1}\}$ est donc isomorphe à A_{k+1} avec la propriété que x_k et x_{k+1} soient jumeaux dans $G - y_{k+1}$ et y_k, y_{k+1} soient jumeaux dans $G - x_k$.

Comme ce processus ne se termine pas, A_∞^+ est un sous graphe induit de G. □

Ainsi, si nous considérons un graphe fini, au voisinage de chaque sommet, il y a un sommet qui ne crée pas de jumeaux lorsqu'il est retiré. En choisissant des sommets assez éloignés, nous allons donc pouvoir créer des ensembles 4-indépendants I (tout paire de sommets de I est à distance au moins 4) tels que pour tout sommet x de I, le graphe $G - x$ soit sans jumeaux. Cela nous donne un code identifiant :

Lemme 2.38 *Soient G un graphe fini connexe sans jumeaux et I un ensemble de sommets 4-indépendant. Si pour tout x dans I, l'ensemble $V(G) \setminus \{x\}$ est un code identifiant de G, alors $C = V(G) \setminus I$ est un code identifiant de G.*

Preuve : L'ensemble C est clairement un ensemble dominant de G. Soient x et y deux sommets de G. S'ils sont tous les deux dans l'ensemble 4-indépendant I, alors $N[x] \cap C$ et $N[y] \cap C$ sont disjoints et non vides, donc x et y sont identifiés.

Sinon, supposons par exemple que $x \in C$. Alors si x et y ne sont pas séparés, cela signifie que x et y sont adjacents. Alors $N[x] \cup N[y]$ contient au plus un sommet u de I et $N[x] \Delta N[y] = \{u\}$. Une contradiction car $G - u$ est sans jumeaux. □

Nous obtenons alors une borne qui conforte la conjecture 1.5 :

Proposition 2.39 *Soit G un graphe connexe sans jumeaux sur n sommets de degré maximum Δ. Alors :*

$$\gamma^{ID}(G) \leq n\left(1 - \frac{\Delta - 2}{\Delta(\Delta-1)^5 - 2}\right) = n - \frac{n}{\Theta(\Delta^5)}.$$

Preuve : Soit I un ensemble 6-indépendant maximal de G. Alors $|I| \geq \frac{n(\Delta-2)}{\Delta(\Delta-1)^5-2}$ car pour tout sommet x de G, $|B_5(x)| \leq \frac{\Delta(\Delta-1)^5-2}{\Delta-2}$.

Pour chaque sommet $x \in I$, il existe d'après le lemme 2.37 un sommet $f(x)$ adjacent à x tel que $G - f(x)$ soit sans jumeaux. Soit $f(I) = \{f(x) \mid x \in I\}$.

Comme I est un ensemble 6-indépendant, $f(I)$ est un ensemble 4-indépendant de G, et $|f(I) = |I|$. D'après le lemme 2.38, $C = V(G) \setminus f(I)$ est un code identifiant de G. La borne s'obtient alors en prenant n'importe quel ensemble 6-indépendant maximal de G. □

Il est à noter que lorsque le graphe est régulier, nous pouvons obtenir un meilleur résultat. En effet, si G est régulier, pour tout sommet x, $G - x$ est sans jumeaux. Le complémentaire d'un ensemble 4-indépendant quelconque est donc toujours un code identifiant. La borne obtenue est alors de la forme $\gamma^{\text{ID}}(G) \leq n - \frac{n}{\Theta(\Delta^3)}$.

Dans Gravier et Moncel [48] il est montré que dans tout graphe infini dont les sommets sont tous de degré fini, il existe un sommet x tel que $V(G) \setminus \{x\}$ soit un code identifiant de G. Nous pouvons étendre ce résultat en utilisant le lemme 2.37 et une preuve similaire à la proposition 2.39 :

Théorème 2.40 *Soit G un graphe connexe infini sans jumeaux dont tous les sommets sont de degré fini. Alors il existe un ensemble infini de sommets I tel que $V(G) \setminus I$ soit un code identifiant de G.*

De plus, d'après le théorème 2.36, nous avons :

Corollaire 2.41 *Pour tout graphe infini possédant un sommet de degré fini, il existe un sommet x tel que $V(G) \setminus \{x\}$ soit un code identifiant de G.*

2.3.4 Codes r-identifiants

Dans cette partie, nous discutons le cas des codes r-identifiants. Un code r-identifiant d'un graphe G est équivalent à un code identifiant dans le graphes G^r, et $\gamma_r^{\text{ID}}(G) = \gamma^{\text{ID}}(G^r)$. Les graphes G tels que $\gamma_r^{\text{ID}}(G) = |V(G)| - 1$ sont donc les racines r-ièmes des graphes G tels que $\gamma^{\text{ID}}(G) = |V(G)| - 1$ que l'on vient de caractériser.

Trouver la racine r-ième d'un graphe H est en général un problème difficile (Motwani et Sudan [79]). Il ne semble pas que cela soit plus facile dans notre cas. Nous connaissons au moins quelques unes des racines r-èmes des graphes extrémaux.

Lorsque $k-1$ est divisible par r, alors $G = P_{2k}^s$, avec $s = \frac{k-1}{r}$ est une racine r-ième de A_k. Cependant, dans la plupart des cas, nous pouvons retirer des arêtes de G et toujours avoir $G^r = A_k$. De plus, une racine r-ième de A_k n'est pas toujours un sous-graphe de P_{2k}^s, comme le montre l'exemple de la figure 2.9 qui nous a été donné par Adrian Kosowski.

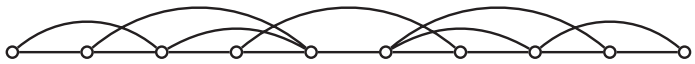


Figure 2.9 – *Une racine carrée de A_5 qui n'est pas un sous graphe de P_{10}^2*

Charon *et al.* [22] mentionnent aussi l'existence de racine r-ième du graphe complet privé d'un couplage maximal lorsque le nombre de sommets n du graphe vérifie $n \geq 3r^2$, donnant des exemples d'autres graphes atteignant la borne.

Dans le cas des graphes infinis, il existe une racine carrée de A_∞. Ce graphe est défini de la manière suivante : ses sommets sont l'ensemble $X \cup Y$, les arêtes entre X et Y sont les mêmes que pour A_∞ mais il n'y a aucune arête au sein de X ni au sein de Y. Nous ne savons pas s'il existe d'autres racines des graphes décrits dans le théorème 2.36.

Par ailleurs, un $(3r+1)$-ensemble indépendant de G devient un 4-ensemble indépendant dans G^r. Aussi nous pouvons généraliser le lemme 2.38, la proposition 2.39 et le théorème 2.40 :

Lemme 2.42 *Soit G un graphe connexe sur n sommets tel que G^r soit sans jumeaux. Soit I un ensemble $(3r+1)$-indépendant de G tel que pour chaque sommet v de I l'ensemble $V(G) \setminus \{v\}$ est un code r-identifiant de G. Alors $C = V(G) \setminus I$ est un code r-identifiant de G.*

Proposition 2.43 *Soit G un graphe connexe sur n sommets de degré maximum Δ, tel que G^r soit sans jumeaux. Alors :*

$$\gamma_r^{ID}(G) \leq n\left(1 - \frac{\Delta - 2}{\Delta(\Delta-1)^{5r} - 2}\right) = n - \frac{n}{\Theta(\Delta^{5r})}.$$

Proposition 2.44 *Soit G un graphe connexe infini dont les sommets sont tous de degré fini et tel que G^r soit sans jumeaux. Alors il existe un ensemble infini de sommets I tel que $V(G) \setminus I$ soit un code r-identifiant de G.*

2.4 Conclusion

Nous avons classifié dans ce chapitre les graphes « mauvais » pour les codes identifiants au sens où tous les sommets (sauf un dans le cas non orienté fini) sont nécessaires pour les identifier. Nous retrouverons au chapitre 5 certains d'entre eux, les graphes P_{2k}^{k-1}, qui sont aussi des pires cas pour une variante des codes identifiants.

Les familles de graphes exhibées ont une forte structure. Par exemple, dans le cas fini non orienté, les graphes d'ordre n dont le plus petit code identifiant est de taille $n-1$, sont essentiellement des joints de graphes P_{2k}^{k-1}. Aussi, cette famille est reconnaissable en temps polynomial. En effet, décomposer un graphe en facteurs joints est très facile – cela revient à chercher les composantes connexes du complémentaires – et tester si un graphe est isomorphe à P_{2k}^{k-1} est polynomial. En fait, pour tout k fixé, décider si un graphe G d'ordre n possède un code identifiant de taille $n-k$ est un problème polynomial puisqu'il suffit de tester pour chaque sous-ensemble X de taille k si $V(G) \setminus X$ est un code identifiant. Peut-on cependant trouver une structure pour les graphes G d'ordre n vérifiant $\gamma^{\text{ID}}(G) = n-k$? Par ailleurs, pour quelles fonctions f de n, le problème de décider si $\gamma^{\text{ID}}(G) = n - f(n)$ reste-t-il polynomial ?

Chapitre 3

Études dans des classes de graphes particulières

Sommaire

3.1 Graphes de Sierpiński 60
3.1.1 Définition des graphes de Sierpiński 61
3.1.2 Codes identifiants 62
3.1.3 Codes localisateur-dominateurs 63
3.1.4 Ensembles résolvants 64
3.1.5 Conclusion sur les graphes de Sierpiński 67
3.2 Graphes d'intervalles 68
3.2.1 Borne inférieure générale 68
3.2.2 Graphes d'intervalles unitaires 69
3.2.3 Discussion sur la complexité 70
3.3 Graphes adjoints 71
3.3.1 Codes arête-identifiants 71
3.3.2 Premiers résultats 73
3.3.3 Bornes inférieures 75
3.3.4 Bornes supérieures 82
3.3.5 Complexité 86
3.4 Grille du roi 95
3.4.1 Méthode 96
3.4.2 Détails de la preuve 101
3.4.3 Preuves des lemmes techniques 108
3.4.4 Conclusion sur la grille du roi 115
3.5 Conclusion du chapitre 116

Étudier les codes identifiants pour un graphe quelconque est un problème difficile lorsque nous n'avons aucune information sur le graphe. Il n'existe [1] pas d'algorithme polynomial qui prend en entrée un graphe quelconque et renvoie la taille du meilleur code identifiant de ce graphe (voir la partie 1.3.4 du chapitre 1). Nous cherchons alors des bornes sur la taille minimale d'un code identifiant mais les résultats obtenus sont très généraux. Pour ces raisons, il est naturel de restreindre la classe des graphes étudiés afin de pouvoir donner des résultats du même type mais plus précis. Ainsi, les codes identifiants ont par exemple été étudiés pour de très nombreuses restrictions comme les cycles, les hypercubes, les arbres, les graphes planaires, les grilles,... Nous considérons dans cette partie quatre classes de graphes. Les techniques utilisées et les résultats obtenus sont différents pour chacune d'entre elles.

La première classe étudiée est celle des graphes de Sierpiński. Pour ces graphes nous sommes capables de donner les valeurs optimales exactes des codes identifiants, des ensembles localisateur-dominateurs et des ensembles résolvants.

Nous considérons ensuite les graphes d'intervalles. Nous améliorons la borne inférieure $\log(|V(G)|+1)$ du théorème 1.3 en montrant qu'un code identifiant dans un graphe d'intervalles est de taille au moins $\Theta(\sqrt{|V(G)|})$.

Dans une troisième partie, nous étudions les codes identifiants dans les graphes adjoints. Cela revient à identifier les arêtes d'un graphe avec ses arêtes. Pour la classe des graphes adjoints aussi, la borne inférieure $\log(|V(G)|+1)$ n'est plus valable et devient une borne de l'ordre de $\sqrt{|V(G)|}$. Nous montrons de plus que le problème des codes identifiants est $\mathcal{NP}$-complet dans une sous-classe des graphes adjoints.

Enfin, la dernière classe étudiée est en fait un seul graphe : la grille infinie du roi. Nous améliorons la borne inférieure connue sur la densité minimale d'un code $(2, \leq 2)$-identifiant dans ce graphe.

3.1 Graphes de Sierpiński

Klavžar et Milutinović [66] ont introduit en 1997 les graphes $S(n,k)$ pour généraliser des graphes issus du problème des tours de Hanoï. Ces graphes ont ensuite pris le nom de *graphes de Sierpiński* car ils ont été introduits pour des raisons topologiques [67, 71, 75].

Ces graphes, définis sur l'ensemble $\{0,...,k-1\}^n$, proposent une métrique alternative aux métriques usuelles. Les distances dans ce graphe sont relativement faciles à calculer et ils ont de bonnes propriétés pour les problèmes de code. Il existe par exemple des codes parfaits (ensembles dominants tels que chaque sommet est dominé exactement une fois) pour tous les graphes $S(n,k)$ ([30, 67, 70]). D'autres types de codes ont aussi été étudiés sur ces graphes [4, 46].

Nous étudions dans cette partie quelques problèmes d'identification sur les graphes de Sierpiński : nous donnons les valeurs exactes de la taille minimale d'un code identifiant, localisateur-dominateur et d'un ensemble résolvant pour ces graphes. Ces travaux sont issus de travaux réalisés avec Sylvain Gravier, Matjaž Kovše, Michel Mollard et Julien Moncel [47] pour les parties 3.1.2 et 3.1.3 et d'un travail personnel pour la partie 3.1.4.

1. Sauf si $\mathcal{P} = \mathcal{NP}$.

3.1.1 Définition des graphes de Sierpiński

Définition directe

Les graphes de Sierpiński sont définis par deux paramètres entiers k et n. Le *graphe de Sierpiński* $S(n,k)$ a pour ensemble de sommets $\{0,1,\ldots,k-1\}^n$. Nous adopterons les notations utilisées en théorie des mots pour représenter les sommets. Deux sommets $u = i_1...i_n$, et $v = j_1...j_n$ de $S(n,k)$ sont adjacents s'il existe un indice $h \in \{1,...,n\}$ tel que :

1. $i_t = j_t$ pour $t \leq h-1$,
2. $i_h \neq j_h$,
3. $i_t = j_h$ et $j_t = i_h$ pour $t = h+1,...,n$.

La figure 3.1 montre les graphes $S(2,5)$ et $S(3,3)$.

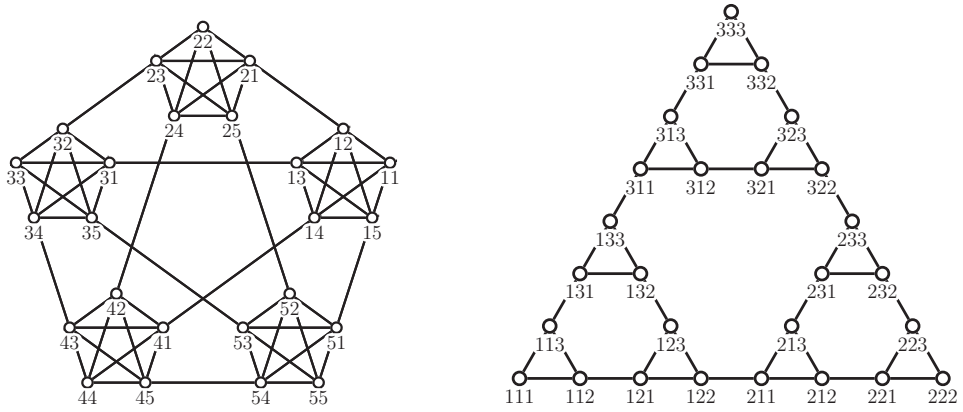


Figure 3.1 – *Le graphe $S(2,5)$ (à gauche) et le graphe $S(3,3)$ (à droite).*

Définition récursive

Le graphe $S(n,k)$ peut être construit de manière inductive. Le graphe $S(1,k)$ est une clique de taille k dont les sommets sont étiquetés de 1 à k. Le graphe $S(n,k)$ est obtenu à partir de k copies de $S(n,k-1)$. Devant les étiquettes de la copie k, la lettre k est ajoutée. Enfin, pour chaque paire d'entier i,j, l'arête $\{ij...j, ji...i\}$ est ajoutée.

Tours de Hanoï généralisées

Lorsque $k = 3$, les graphes $S(n,k)$ correspondent aux graphes du problème des Tours de Hanoï. Dans ce problème, il y a trois pics et n disques de tailles toutes différentes. Au départ, tous les disques sont positionnés par ordre de taille décroissante sur un des pics. Le but du jeu est de déplacer toute la pile de disques sur un autre pic, sachant qu'on ne peut déplacer qu'un seul disque à la fois, qu'il doit être en haut d'une pile et qu'il ne peut être déplacé que sur un disque plus grand. Ce problème peut être représenté par un graphe : un sommet correspond à une configuration des disques. Deux configurations sont reliées si l'on peut passer

de l'une à l'autre. Le graphe ainsi construit est isomorphe à $S(n, 3)$. Exprimer en terme de graphes, le problème des tours de Hanoï se résume à trouver un chemin entre deux sommets particuliers du graphe.

Lorsque $k > 3$, les graphes $S(n, k)$ peuvent modéliser une variante des jeux de Hanoï avec k pics, n disques et une règle pour passer d'une configuration à une autre modifiée [66].

Terminologie

Un sommet de la forme $u = i...i$ de $S(n, k)$ est appelé *sommet extrême* i et tous les autres sommets de $S(n, k)$ sont appelés *sommets internes*. Les sommets extrêmes de $S(n, k)$ sont de degré $k - 1$ alors que les sommets internes sont tous de degré k. Il y a exactement k sommets extrêmes dans $S(n, k)$ et $S(n, k)$ a k^n sommets.

Soit u un mot quelconque de longueur $d \leq n - 1$. Les sommets commençant par u induisent le graphe $S(n - d, k)$. Nous appelons ce graphe la *copie* u de $S(n - d, k)$.

Soit $u = i_1...i_n$ un sommet arbitraire de $S(n, k)$. Nous notons $K(u)$ la clique de taille k induite par les sommets de la forme $i_1...i_{n-1}j$ avec $1 \leq j \leq k$. Le voisinage d'un sommet extrême u est $K(u) \setminus \{u\}$ et un sommet interne u a exactement un voisin qui n'est pas dans $K(u)$.

3.1.2 Codes identifiants

Théorème 3.1 *Soient $k \geq 3$ et $n \geq 2$. Le plus petit code identifiant de $S(n, k)$ est de cardinalité $k^{n-1}(k - 1)$.*

Preuve : Nous montrons tout d'abord que pour tout code identifiant C de $S(n, k)$, $|C| \geq k^{n-1}(k - 1)$. Soit C un code identifiant $S(n, k)$. Pour chaque sommet interne u de $S(n, k)$, soit $m(u)$ l'unique voisin de u qui n'est pas dans $K(u)$. Pour chaque k-clique K de $S(n, k)$, soit

$$M(K) = \{m(u) \mid u \text{ est un sommet interne de } K\}.$$

Il y a k^{n-1} ensembles $M(K)$ et ils sont tous disjoints. Soient u et v deux sommets internes d'une même clique K. Pour séparer u et v, $m(u)$ ou $m(v)$ doit être dans le code identifiant C. Si K contient un sommet extrême, alors tous les sommets de $M(K)$ doivent être dans le code C. Ainsi, pour chaque k-clique, il y a au moins $k - 1$ sommets de $M(K)$ qui sont dans C. Comme il y a k^{n-1} ensembles $M(K)$ tous disjoints, C est de taille au moins $k^{n-1}(k - 1)$.

Nous construisons maintenant un code identifiant de cette taille. Pour $n = 2$, l'ensemble de tous les sommets internes est un code identifiant de $S(2, k)$ de taille $k(k - 1)$ (voir la figure 3.2). Pour $n \geq 3$ quelconque, soit C le code identifiant construit en prenant sur chaque copie de $S(2, k)$ les sommets internes de $S(2, k)$ (voir la figure 3.2). Alors C est un code identifiant de cardinalité $k^{n-1}(k - 1)$. □

Ces graphes sont des exemples arbitrairement grands de graphes atteignant la borne de la conjecture 1.5 pour toutes valeurs du degré maximum.

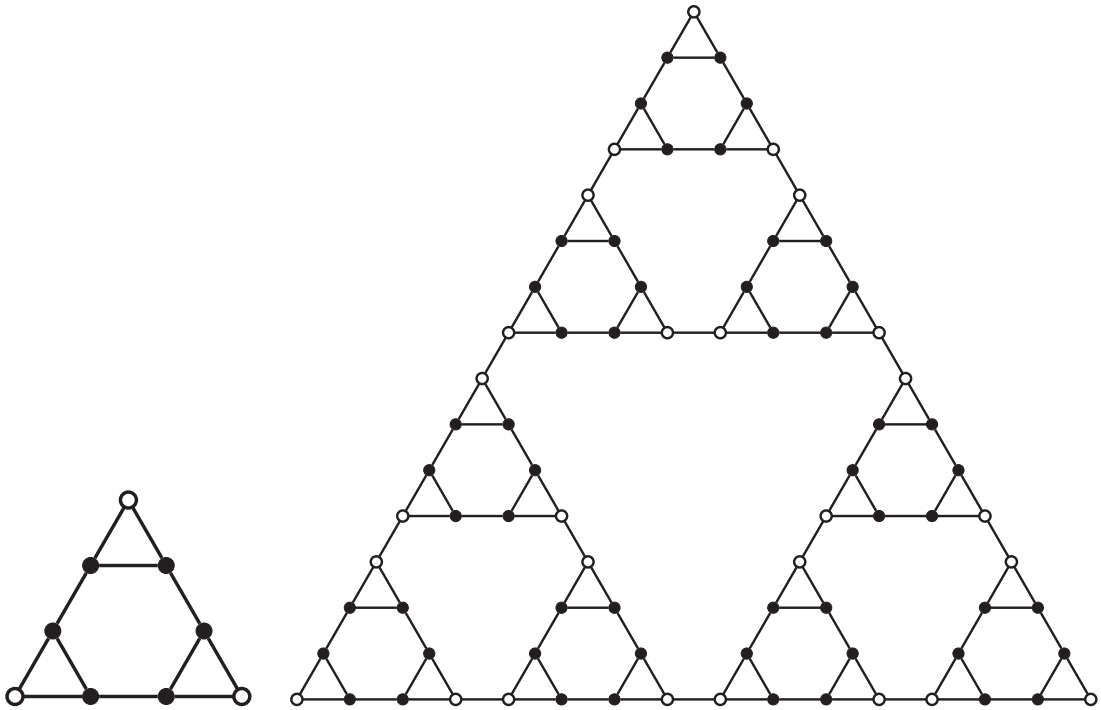


Figure 3.2 – *Des codes identifiants optimaux pour $S(2,3)$ et $S(4,3)$. Le code identifiant de $S(4,3)$ est construit en répétant sur chaque copie de $S(2,3)$ le code identifiant de $S(2,3)$.*

3.1.3 Codes localisateur-dominateurs

Nous déterminons maintenant la cardinalité minimale d'un code localisateur-dominateur dans les graphes de Sierpiński. Rappelons qu'un code localisateur-dominateur C d'un graphe G est un ensemble dominant de G qui sépare les sommets de $V(G)\backslash C$.

Théorème 3.2 *Soient $n \geq 2$ et $k \geq 3$. La cardinalité minimum d'un code localisateur-dominateur dans le graphe de Sierpiński $S(n,k)$ est $\frac{k^{n-1}(k-1)}{2}$.*

Preuve : Soit C un code localisateur-dominateur de $S(n,k)$. Une arête uv est *traversante* si $K(u) \neq K(v)$, autrement dit, l'arête uv n'appartient pas à une k-clique. L'ensemble de toutes les arêtes traversantes est un couplage et chaque sommet interne appartient à exactement une arête traversante. Considérons la partition suivante $\mathcal{P}$ des sommets de $S(n,k)$: chaque sommet extrême est un élément isolé dans la partition et chaque paire de sommets $\{u,v\}$ formant une arête traversante est un élément de la partition. Pour un sommet u de $S(n,k)$, $P(u)$ désigne l'élément de $\mathcal{P}$ qui contient u.

Si deux sommets u et v d'une même k-clique K sont tels que l'ensemble $P(u) \cup P(v)$ ne contienne aucun élément du code C, alors u et v ne sont pas dans le code et ne sont pas séparés par C, ce qui est contradictoire. Ainsi, pour chaque k-clique K, au moins $k-1$ ensembles parmi les k ensembles de $\mathcal{P}$ contenant un sommet de K doivent contenir un élément du code C. Un ensemble de $\mathcal{P}$ contribue pour au plus deux k-cliques différentes et il y a k^{n-1} cliques de taille k. Nécessairement, $|C| \geq \frac{k^{n-1}(k-1)}{2}$.

Nous allons maintenant construire un code localisateur-dominateur atteignant la borne. Il suffit de construire un code C tel que pour chaque k-clique K, un des éléments de K soit dans le code (pour la domination) et $k-1$ ensembles parmi les ensembles de $\mathcal{P}$ intersectant K, contiennent au moins un élément de C. En effet, soit u et v deux sommets qui ne sont pas dans C. Si u et v sont dans la même clique, alors un des ensembles $P(u)$ ou $P(v)$ contient un élément de C. Supposons que ce soit $P(u)$. Alors $P(u)$ contient alors deux éléments, u et u', et $u' \in C$ n'est pas adjacent à v donc sépare u et v. Si u et v ne sont pas dans la même clique, alors chacun d'entre eux est dominé par un élément de leur clique qui le sépare de l'autre.

Pour $n = 2$, nous construisons un tel code en prenant comme code C tout d'abord les sommets $i(i+1)$ avec $i \in 1, \ldots, k-1$ et le sommet $k1$ et en ajoutant ensuite un sommet par arête traversante qui n'a pas encore d'extrémité dans C (voir la figure 3.3). Le code C obtenu est dominant et dans la partition $\mathcal{P}$, toutes les arêtes traversantes contiennent exactement un sommet du code. D'après la remarque précédente, c'est un code localisateur-dominateur, qui est de taille $\frac{k(k-1)}{2}$.

Pour $n \geq 3$, sur chaque sous-graphe de $S(n,k)$ isomorphe à $S(2,k)$, les sommets correspondant aux éléments du code construit pour $n = 2$ sont mis dans le code (voir la figure 3.3). Il y a k^{n-2} sous-graphes isomorphes à $S(2,k)$ qui sont tous disjoints, le code construit est donc de taille $\frac{k^{n-1}(k-1)}{2}$. □

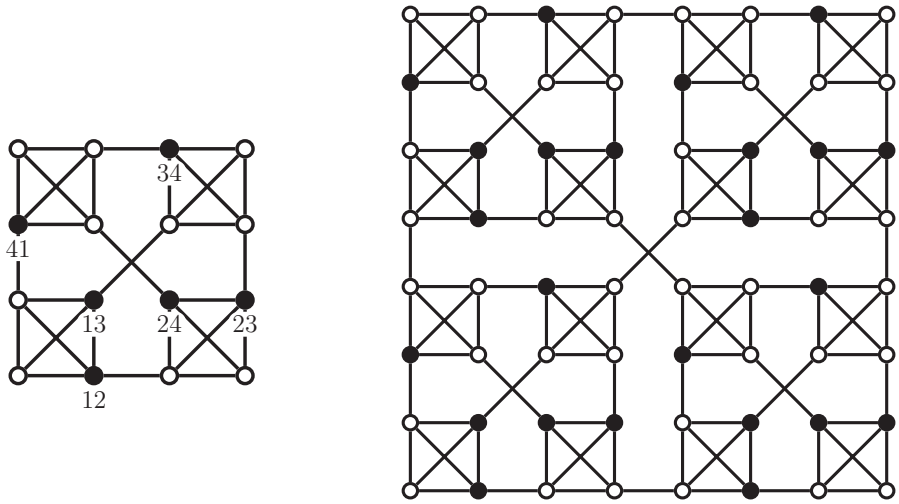


Figure 3.3 – *Des codes localisateur-dominateurs optimaux pour $S(2,4)$ (à gauche) et $S(3,4)$ (à droite). Le code localisateur-dominateur de $S(3,4)$ est construit en répétant sur chaque copie de $S(2,4)$ le code localisateur-dominateur de $S(2,4)$.*

3.1.4 Ensembles résolvants

Pour terminer cette partie, nous calculons la dimension des graphes de Sierpiński. Rappelons que la *dimension* d'un graphe G est la taille du plus petit ensemble C tel que, pour toute paire de sommets de V, il existe un sommet du code qui n'est pas à la même distance des deux sommets. Un tel ensemble est appelé *ensemble résolvant*.

Distances dans $S(n,k)$

Pour étudier les ensembles résolvants, nous avons besoin de connaître la distance entre deux sommets du graphe $S(n,k)$. Nous reprenons les résultats de Klavžar et Milutinović [66] à ce sujet. Posons $\rho_{i,j} = 1$ si $i \neq j$ et $\rho_{i,j} = 0$ si $i = j$. Pour un mot $j_1 \dots j_n$ de $\{0, \dots, k-1\}^n$ et un entier $i \in \{0, \dots, k-1\}$, $\mathcal{P}^i_{j_1 \dots j_n}$ désigne l'entier dont l'écriture binaire est $\rho_{i,j_1} \dots \rho_{i,j_n}$. Par exemple, si tous les j_t sont différents de i, $\mathcal{P}^i_{j_1 \dots j_n} = 2^n - 1$. Il est alors très simple de calculer la distance entre un sommet extrême et un sommet quelconque du graphe.

Lemme 3.3 (Klavžar, Milutinović [66]) *La distance entre le sommet extrême $u = i \dots i$ et un sommet quelconque $v = j_1 \dots j_n$ est $\mathcal{P}^i_{j_1 \dots j_n}$. De plus il y a un unique plus court chemin entre u et v.*

Dans le cas particulier où v est un autre sommet extrême, la distance entre u et v est[2] donc $2^n - 1$. Ce lemme nous donne aussi que le diamètre du graphe de Sierpiński est $2^n - 1$. Lorsqu'aucun des deux sommets n'est un sommet extrême, alors il peut y avoir deux plus courts chemins. La distance entre les deux sommets peut encore s'exprimer à l'aide des nombres $\mathcal{P}^i_{j_1 \dots j_n}$.

Lemme 3.4 (Klavžar, Milutinović [66]) *Soit $u = i_1 \dots i_n$ et $v = j_1 \dots j_n$ deux sommets quelconques de $S(n,k)$. Soit ℓ le premier indice sur lequel u et v diffèrent. Si $\ell = n$, alors $d(u,v) = 1$, sinon :*

$$d(u,v) = \min\{\mathcal{P}^{j_\ell}_{i_{\ell+1} \dots i_n} + 1 + \mathcal{P}^{i_\ell}_{j_{\ell+1} \dots j_n}, \mathcal{P}^h_{i_{\ell+1} \dots i_n} + 1 + 2^{n-\ell} + \mathcal{P}^h_{j_{\ell+1} \dots j_n} | h \neq i_\ell, j_\ell\}.$$

Ce résultat est obtenu en considérant les chemins possibles pour aller entre u et v. Supposons par exemple que $\ell = 1$, c'est-à-dire que les deux sommets sont dans des copies différentes de $S(n-1,k)$. Alors une manière d'aller de u à v est de passer directement de la copie i_1 où u est situé à la copie j_1 où v est situé, en passant par les sommets $i_1 j_1 \dots j_1$ et $j_1 i_1 \dots j_1$. Cela donne un chemin dont la longueur est $\mathcal{P}^{j_1}_{i_2 \dots i_n} + 1 + \mathcal{P}^{i_1}_{j_2 \dots j_n}$, le premier membre du minimum dans le résultat. Une autre manière pour relier le sommet u au sommet v est de passer par une troisième copie h de $S(n-1,k)$, ce qui amène à la deuxième partie du minimum. Le résultat s'obtient de la même manière pour $\ell \neq 1$ en se ramenant au graphe $S(n-\ell+1,k)$ et en remarquant que ce n'est pas efficace d'en sortir.

Dimension métrique de $S(n,k)$

Nous commençons par exhiber un ensemble résolvant de taille $k-1$.

Lemme 3.5 *Un ensemble formé par $k-1$ sommets extrêmes est un ensemble résolvant de $S(n,k)$.*

Preuve : Soit C l'ensemble des sommets extrêmes excepté le sommet extrême n. Soit $u = i_1 \dots i_n$ et $v = j_1 \dots j_n$ deux sommets distincts de $S(n,k)$. Soit ℓ un indice tel que $i_\ell \neq j_\ell$. Un des deux entiers parmi i_ℓ et j_ℓ n'est pas égal à n. Sans perte de

2. Ce résultat correspond au nombre minimum de coups pour gagner aux tours de Hanoï.

généralité, $i_\ell \neq n$. D'après le lemme 3.3, la distance entre u et le sommet extrême i_ℓ vaut $\mathcal{P}_u^{i_\ell}$, et dans son écriture binaire, la coordonnée ℓ est nulle. Cette distance est donc différente de celle entre v et le sommet extrême i_ℓ qui contient un 1 dans son écriture binaire en position ℓ. Comme $i_\ell \neq n$, il y a un sommet de C qui n'est pas à la même distance de u et v. □

La figure 3.4 montre un ensemble résolvant de $S(2,4)$.

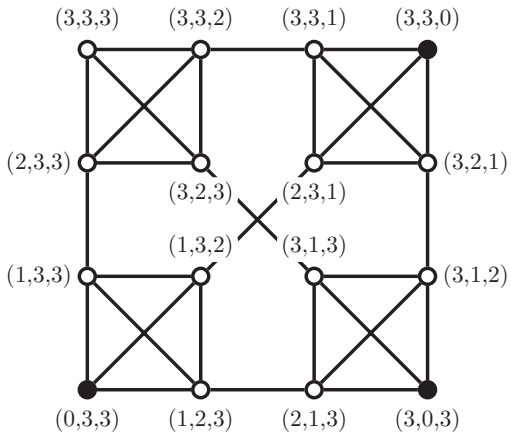


Figure 3.4 – *Un ensemble résolvant de $S(2,4)$ de taille 3. Les distances aux trois sommets de l'ensemble sont indiquées entre parenthèses.*

Théorème 3.6 *La dimension métrique de $S(n,k)$ est $k-1$.*

Preuve : Le lemme 3.5 montre que la dimension de $S(n,k)$ est au plus $k-1$. Supposons qu'il existe un ensemble résolvant C de taille $k-2$. Alors deux copies de $S(n-1,k)$ n'ont pas d'élément de C. Sans perte de généralité, supposons que ce sont les copies 1 et 2. Soient u le sommet $13\ldots31$ et v le sommet $13\ldots32$. Comme C est un ensemble résolvant, il existe un sommet $c = j_1 \ldots j_n$ de l'ensemble C qui n'est pas à la même distance de u et v. Le sommet c n'est ni dans la copie 1 ni dans la copie 2 de $S(n-1,k)$.

Supposons tout d'abord que c est dans la copie 3 de $S(n-1,k)$, c'est-à-dire que $j_1 = 3$. Alors, d'après le lemme 3.4, la distance entre u et c est (ici $\ell = 1$) :

$$d(u,c) = \min\{2 + \mathcal{P}^1_{j_2\ldots j_n}, \mathcal{P}^h_{3\ldots31} + 1 + 2^{n-1} + \mathcal{P}^h_{j_2\ldots j_n} | h \neq 1,3\}.$$

Comme $P^1_{j_2\ldots j_n} \leq 2^{n-1} - 1$, nous avons nécessairement $d(u,c) = 2 + \mathcal{P}^1_{j_2\ldots j_n}$. De la même manière $d(v,c) = 2 + \mathcal{P}^1_{j_2\ldots j_n}$. Les sommets u et v sont à la même distance de c, ce qui est contradictoire.

Nous pouvons donc supposer que c n'est pas dans la copie 3, et sans perte de généralité, c est dans la copie 4 (si $k \leq 3$, la preuve se termine ici). La distance de u à c est alors :

$$d(u,c) = \min\{\mathcal{P}^4_{3\ldots31} + 1 + \mathcal{P}^1_{j_2\ldots j_n}, \mathcal{P}^h_{3\ldots31} + 1 + 2^{n-1} + \mathcal{P}^h_{j_2\ldots j_n} | h \neq 1,4\}.$$

Or $\mathcal{P}^4_{3\ldots31} = 2^{n-1} - 1$ (il y a $n-2$ chiffres 3) et $\mathcal{P}^h_{3\ldots31} = 1$ si $h = 3$ et $2^{n-1} - 1$ si $h \neq 3$ (car $h \neq 1$). Donc

$$d(u,c) = \min\{2^{n-1} + \mathcal{P}^1_{j_2\ldots j_n}, 2 + 2^{n-1} + \mathcal{P}^3_{j_2\ldots j_n}, 2^n - 1 + \mathcal{P}^h_{j_2\ldots j_n} | h \neq 1,3,4\}.$$

Comme $\mathcal{P}^1_{j_2\ldots j_n,} \leq 2^{n-1} - 1$, nous avons en fait :

$$d(u,c) = \min\{2^{n-1} + \mathcal{P}^1_{j_2\ldots j_n}, 2 + 2^{n-1} + \mathcal{P}^3_{j_2\ldots j_n}\}.$$

Le résultat obtenu est exactement le même pour v (u et v jouent un rôle symétrique vis à vis de c), donc $d(u,c) = d(v,c)$, ce qui mène à une contradiction. La figure 3.5 donne un exemple d'ensemble avec $k-2$ éléments qui n'est pas un ensemble résolvant.

Il n'existe donc pas d'ensemble résolvant de taille $k-2$. □

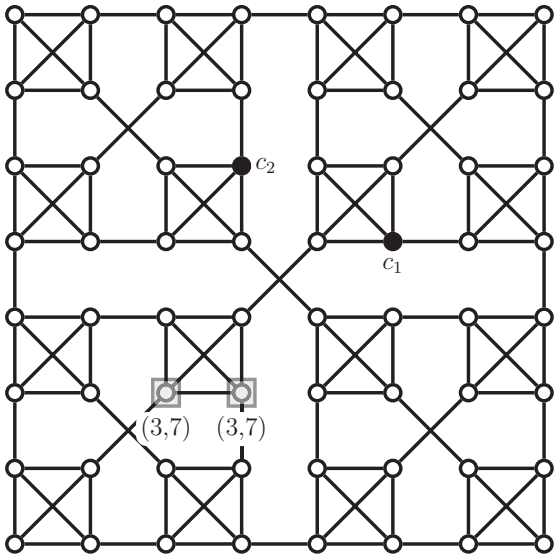


Figure 3.5 – *L'ensemble* $\{c_1, c_2\}$ *ne forme pas un ensemble résolvant de* $S(3,4)$. *Les sommets* 131 *et* 132 *(encadrés) sont à la même distance des deux éléments de* c_1 *et* c_2 *(les distances sont précisées entre parenthèses).*

3.1.5 Conclusion sur les graphes de Sierpiński

Nous avons étudié dans cette partie différents problèmes d'identification sur les graphes de Sierpiński et pu donner les valeurs exactes des ensembles identifiants optimaux dans chacun des cas. De plus, dans le cas des codes identifiants, les graphes de Sierpiński atteignent la borne de la conjecture 1.5.

La structure des graphes de Sierpiński se prête donc à l'étude des problèmes de métrique et de code. Ces graphes peuvent se généraliser en ne débutant pas la construction inductive par un graphe complet mais par un graphe quelconque et en ajoutant les arêtes en fonction du graphe de départ choisi. Nous avons commencé récemment l'étude de cette généralisation (sous le nom de *graphes de Sierpiński généralisés*) et pensons qu'elle peut mener à d'autres exemples extrémaux pour les codes identifiants ou pour d'autres problèmes du même type.

3.2 Graphes d'intervalles

Nous nous intéressons dans cette partie à la classe des graphes d'intervalles. Nous améliorons la borne inférieure $\log(n+1)$ du théorème 1.3 en montrant qu'un code identifiant dans un graphe d'intervalles est de taille au moins $\theta(\sqrt{2n})$. Cette borne est affinée pour les graphes d'intervalles unitaires. Nous terminons cette partie en discutant brièvement la complexité du problème des codes identifiants dans les graphes d'intervalles. Ces travaux de recherche, effectués avec Florent Foucaud, Reza Naserasr et Petru Valicov, sont en cours de rédaction.

3.2.1 Borne inférieure générale

Pour un intervalle I, $a(I)$ désigne le début de l'intervalle I et $b(I)$ sa fin. Nous confondons ici un sommet avec l'intervalle qu'il représente. Les intervalles considérés sont ouverts des deux côtés.

Proposition 3.7 *Soit G un graphe d'intervalles sans jumeaux possédant n sommets. Soit C un code identifiant de G de taille k. Alors*

$$n \leq \frac{k(k+1)}{2}.$$

Preuve : Soit G un graphe d'intervalles sans jumeaux sur n sommets. Soit $\mathcal{I}$ la famille d'intervalles dont G est le graphe d'intersection et soit C un code identifiant de G.

Soient $I_1, ..., I_k$ les intervalles de C ordonnés par *fin croissante*. Pour $i \in \{2, \dots k\}$, nous définissons P_i comme l'intervalle semi ouvert $P_i = [b(I_{i-1}), b(I_i)[$. Nous posons de plus $P_1 =]-\infty, b(I_1)[$. Soit $\mathcal{E}_i$ l'ensemble des intervalles qui commencent durant P_i. Les ensembles $\mathcal{E}_i$ forment une partition des intervalles du graphe. En effet, aucun intervalle ne commencent après $b(I_k)$ ($b(I_k)$ comprise) sinon il ne serait pas dominé.

Soient $i \geq 2$ et I un intervalle de l'ensemble $\mathcal{E}_i$ (I peut être un intervalle du code). L'intervalle I ne peut intersecter que des intervalles de C commençant après $b(I_{i-1})$ donc des intervalles I_j avec $j \geq i$. De plus, les intervalles du code que I intersecte sont consécutifs lorsqu'on considère les intervalles I_j classés par *début croissant*. Enfin, I doit intersecter l'intervalle de C commençant en premier parmi ceux de $\mathcal{E}_i$ pour être dominé. Il y a $k-i+1$ intervalles I_j avec $j \geq i$ et donc $k-i+1$ intersections possibles pour les éléments de $\mathcal{E}_i$. Ainsi, il y a au plus $k-i+1$ éléments dans $\mathcal{E}_i$. De la même manière, un intervalle de $\mathcal{E}_1$ intersecte un ensemble d'intervalles du code consécutifs lorsque les intervalles sont ordonnés par début croissant, et intersecte le premier intervalle pour cet ordre. Il y a donc au plus k éléments dans $\mathcal{E}_1$.

Ainsi, pour tout $i \in \{1, ..., k\}$,

$$|\mathcal{E}_i| \leq k - i + 1.$$

La famille $(\mathcal{E}_i)_{1 \leq i \leq k}$ formant une partition des sommets, nous avons

$$n \leq \frac{k(k+1)}{2}.$$

□

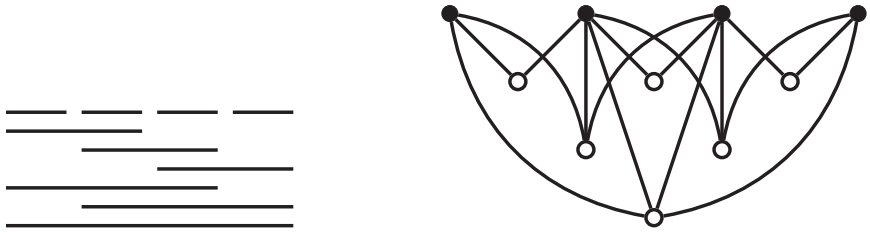


Figure 3.6 – *Une famille d'intervalles (à gauche) dont le graphe d'intervalles (à droite) est G_4. G_4 a 10 sommets et possède un code identifiant de taille 4 représenté par les sommets en noir.*

Corollaire 3.8 *Soit G un graphe d'intervalles sans jumeaux d'ordre n. Alors :*

$$\gamma^{ID}(G) \geq \frac{\sqrt{8n+1}-1}{2} \geq \sqrt{2n} - \frac{1}{2}.$$

La borne obtenue est serrée : pour tout entier k, il existe un graphe d'intervalles sur $\frac{k(k+1)}{2}$ sommets possédant un code identifiant de taille k. Considérons par exemple le graphe G_k représentant l'intersection de la famille d'intervalles

$$\{]i,j[\mid 1 \leq i < j \leq k+1, i,j \in \mathbb{N}\}.$$

La figure 3.6 montre le graphe G_4. Le graphe G_k a $\frac{k(k+1)}{2}$ sommets et le sous-ensemble d'intervalles $\{]i,i+1[\mid 1 \leq i \leq k\}$ est un code identifiant de taille k.

3.2.2 Graphes d'intervalles unitaires

Nous considérons maintenant les graphes d'intersection des familles d'intervalles où tous les intervalles sont de longueur 1. Ces graphes sont appelés *graphes d'intervalles unitaires.*

Proposition 3.9 *Soit G un graphe d'intervalles unitaires. Alors*

$$\gamma^{ID}(G) \geq \frac{|V(G)|+1}{2}.$$

Preuve : Soit $\mathcal{I}$ une famille d'intervalles unitaires dont le graphe d'intersection est G. Nous pouvons supposer que tous les intervalles sont ouverts.

Soit C un code identifiant de taille k. Considérons l'ensemble $T = \{a(I)-1 \mid I \in C\} \cup \{a(I)+1 \mid I \in C\}$ et trions T par ordre croissant : $T = \{t_1 \leq t_2 \leq \cdots \leq t_{2k}\}$. Soient I et J deux intervalles commençant entre deux éléments de T consécutifs : $a(I), a(J) \in]t_i, t_{i+1}]$, pour un certain i. Montrons que I et J ne sont pas séparés par C. Si c'était le cas et que par exemple $a(I) < a(J)$, alors il y aurait un intervalle I_0 de C tel que $a(I) \leq b(I_0) < a(J)$ ou bien tel que $b(I) \leq a(I_0) < b(J)$. Dans le premier cas, $a(I_0)+1 \in]t_i, t_{i+1}[$ et dans le deuxième cas, nous aurions $a(I_0)-1 \in]t_i, t_{i+1}[$, ce qui contredit la définition des t_i.

Aussi, il y a au plus un intervalle commençant dans la période $]t_i, t_{i+1}]$. Il ne peut y avoir un intervalle commençant avant t_1 ou après t_{2k} pour des raisons de

domination. Il y a donc au plus $2k-1$ intervalles dans la famille $\mathcal{I}$, ce qui implique la borne. □

Cette borne est atteinte par exemple pour les chemins de taille impaire qui sont des graphes d'intervalles unitaires.

3.2.3 Discussion sur la complexité

Nous ne connaissons pas la réponse à la question suivante :

Question 3.10 *Quelle est la complexité du problème des codes identifiants sur les graphes d'intervalles ?*

Trouver un ensemble dominant ou un ensemble totalement dominant (tout sommet est à distance 1 de l'ensemble dominant) de taille minimum est un problème polynomial pour la classe des graphes d'intervalles [8, 13, 36, 65]. Par contre, le problème de domination devient $\mathcal{NP}$-complet lorsque la classe est élargie à la classe des graphes 2-intervallaires : les graphes d'intersection de paires d'intervalles. En effet, le problème de domination est $\mathcal{NP}$-complet pour les graphes planaires de degré maximum 3 (Garey et Johnson [43]) qui sont 2-intervallaires. Le problème des codes identifiants dans les graphes d'intervalles se rapproche de ce problème puisque, pour séparer deux intervalles, il faut dominer leur différence symétrique qui est justement une paire d'intervalles. La réponse à la question 3.10 nous semble donc cruciale parmi les questions de complexité sur les codes identifiants.

Le lemme suivant donne une piste pour trouver un algorithme polynomial pour les graphes d'intervalles unitaires. Dans un graphe d'intervalles unitaires, les intervalles représentés par les sommets peuvent être ordonnés de manière naturelle par leur début (ou de manière équivalente, par leur fin). Dans le lemme suivant, les intervalles sont *consécutifs* suivant cet ordre.

Lemme 3.11 *Soit G un graphe d'intervalles unitaires sans jumeaux. Soit C un ensemble dominant qui sépare les paires d'intervalles consécutifs. Alors C est un code identifiant de G.*

Preuve : Soit I_i et I_j deux intervalles avec $i < j$. Si I_i et I_j sont consécutifs, c'est-à-dire, si $j = i+1$, alors I_i et I_j sont séparés par C. Sinon il existe un intervalle $I_k \in C$ qui sépare I_i et I_{i+1}. Si $I_k \leq I_i$ alors I_k intersecte I_i mais pas I_{i+1} et donc pas I_j. L'intervalle I_k sépare donc I_i et I_j. Sinon, $I_k \geq I_{i+1}$ et I_k n'est pas adjacent à I_i. Soit I_k intersecte I_j et sépare I_i et I_j. Soit I_j est dominé par un intervalle $I_\ell \in C$ qui est tel que $I_\ell \geq I_k$ et qui n'est pas donc pas adjacent à I_i. □

Ainsi pour identifier un graphe d'intervalles unitaires, il suffit de dominer les sommets du graphe et de séparer les intervalles consécutifs. De plus, un intervalle peut séparer au plus deux paires d'intervalles consécutifs. Il est alors possible de construire un graphe dont les sommets sont les paires d'intervalles consécutifs et où deux sommets sont reliés si un même intervalle peut séparer les deux paires en même temps, avec comme convention d'ajouter une boucle si un intervalle ne sépare qu'une seule paire. Trouver un ensemble d'arêtes de taille minimum qui couvre tous les sommets de ce graphe est polynomial mais cet ensemble ne donne alors qu'un code séparant les paires d'intervalles consécutifs, qui n'est pas nécessairement dominant.

3.3 Graphes adjoints

Dans cette partie, nous nous intéressons aux codes identifiants dans la classe des graphes adjoints (appelés *line graph* en anglais).

Rappelons que pour tout graphe G, le *graphe adjoint de* G, noté $\mathcal{L}(G)$ est le graphe ayant pour sommets les arêtes de G et où deux sommets sont adjacents si les arêtes correspondantes s'intersectent dans G.

Étudier les codes identifiants dans la classe des graphes adjoints revient à étudier des ensembles d'arêtes identifiant les arêtes dans un graphe quelconque. Nous introduisons donc la notion de *code arête-identifiant*. Nous donnons ensuite des bornes inférieures et supérieures sur les codes arête-identifiants qui impliquent des bornes sur les codes identifiants dans la classe des graphes adjoints. En particulier, comme pour les graphes d'intervalles, la borne inférieure d'ordre $\log(n)$ est améliorée en une borne d'ordre $\sqrt{n}$, avec n le nombre de sommets du graphe. Enfin, nous montrons que le problème des codes identifiants est $\mathcal{NP}$-complet dans la classe des graphes adjoints. Ces travaux sont issus de l'article [37] réalisé avec Florent Foucaud, Sylvain Gravier, Reza Naserasr et Petru Valicov.

3.3.1 Codes arête-identifiants

Définition

Soit G un graphe tel que son graphe adjoint $\mathcal{L}(G)$ soit sans jumeaux. Soit C un code identifiant de $\mathcal{L}(G)$. Alors C peut être vu dans G comme un sous-ensemble d'arêtes C_E tel que, pour toute arête e de G, l'ensemble des arêtes de C_E adjacentes à e soit unique et non vide.

Nous transposons aux arêtes la terminologie des codes identifiants sur les sommets. L'ensemble des arêtes adjacentes à une arête e est noté $N(e)$, si l'on ajoute de plus l'arête e, l'ensemble obtenu est $N[e] = N(e) \cup \{e\}$. Une arête e domine une arête f si $e \in N[f]$ et sépare les arêtes f_1 et f_2 si $e \in N[f_1]\Delta N[f_2]$. Un ensemble d'arêtes C_E est un *code arête-identifiant* si toutes les arêtes sont dominées par un élément de C_E et si toutes les paires d'arêtes sont séparées par un élément de C_E. La taille du plus petit code arête-identifiant de G, s'il existe, est noté $\gamma^{\text{EID}}(G)$.

Un ensemble C_E est un code arête-identifiant de G si et seulement si les sommets correspondant aux arêtes de C_E forment un code identifiant de $\mathcal{L}(G)$ et $\gamma^{\text{EID}}(G) = \gamma^{\text{ID}}(\mathcal{L}(G))$. Il est donc équivalent d'étudier les codes arête-identifiants et d'étudier les codes identifiants (les sommets) dans les graphes adjoints. Nous utiliserons donc surtout les codes arête-identifiants pour étudier les codes identifiants dans les graphes adjoints.

Arêtes pendantes

Comme pour les codes identifiants les sommets, tous les graphes ne possèdent pas de codes arête-identifiants. Pour pouvoir posséder un code arête-identifiant, il ne faut pas que deux arêtes e et f soient telles que $N[e] = N[f]$. Deux telles arêtes sont appelées *arêtes pendantes*. Elles sont adjacentes et aucune autre arête ne peut être adjacente à leurs sommets non commun. La figure 3.7 montre les deux possibilités

Figure 3.7 – *Les arêtes e et f sont des arêtes pendantes, ce sont les deux seules configurations possibles.*

pour des arêtes pendantes. Cette condition est suffisante : un graphe G possède un code arête-identifiant si et seulement si il n'a pas d'arêtes pendantes, c'est-à-dire, si et seulement si $\mathcal{L}(G)$ est sans jumeaux.

Nous utiliserons de nombreuses fois le résultat suivant sur les arbres sans arêtes pendantes.

Lemme 3.12 *Soit T un arbre sans arêtes pendantes. Soit $T = K_2$, soit T a au moins quatre sommets et possède au moins deux sommets de degré 2 adjacent à des sommets de degré 1.*

Preuve : Un arbre à trois sommets est un P_3 et ses deux arêtes sont pendantes. Donc, si $T \neq K_2$, T possède au moins quatre sommets.

Considérons maintenant le plus long chemin de T. Alors chacune de ses extrémités est un sommet de degré 1 dont le voisin est de degré 2 (sinon le chemin peut être agrandi). Les deux sommets de degrés 2 sont nécessairement différents. □

Un exemple de code arête-identifiant

Considérons comme exemple le graphe de Petersen $\mathcal{P}$ représenté sur la figure 3.8. Un couplage parfait forme un code arête-identifiant de ce graphe. Sauf mention contraire, les arêtes appartenant aux codes arête-identifiants seront représentées en gras comme sur la figure 3.8. La borne inférieure $\gamma^{\text{ID}}(\mathcal{L}(\mathcal{P})) \geq \log(|V(\mathcal{L}(\mathcal{P}))| + 1)$ nous donne ici que $\gamma^{\text{EID}}(\mathcal{P}) \geq 4$. Nous verrons que cette borne peut être améliorée et que $\gamma^{\text{EID}}(\mathcal{P}) = 5$.

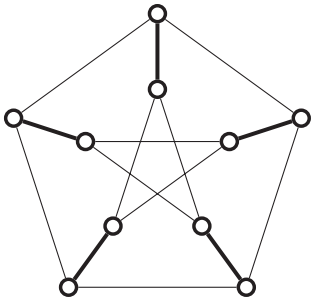


Figure 3.8 – *Un code arête-identifiant (en gras) du graphe de Petersen.*

3.3.2 Premiers résultats

Le lemme suivant est trivial mais sera utilisé de nombreuses fois :

Lemme 3.13 *Soit C_E un code arête-identifiant d'un graphe G. Alors C_E est sans arêtes pendantes.*

Les deux prochains lemmes permettent de construire des codes arête-identifiants à l'aide d'ensembles d'arêtes couvrant les sommets.

Lemme 3.14 *Soit G un graphe de maille au moins 5. Soit C_E un sous-ensemble d'arêtes de G, sans arêtes pendantes, qui couvrent tous les sommets de G. Alors C_E est un code arête-identifiant de G. En particulier, si G a un couplage parfait M, M est un code arête-identifiant.*

Preuve : Comme C_E domine toutes les arêtes de G, nous devons juste montrer que chaque paire d'arêtes est séparée.

Soient e_1 et e_2 deux arêtes de G. Si e_1 et e_2 sont dans le code C_E, alors $C_E \cap N[e_1] \neq C_E \cap N[e_2]$ car le graphe formé par les arêtes de C_E est sans arêtes pendantes.

Sinon, supposons par exemple que $e_2 \notin C_E$. Si $e_1 \in C_E$ et $C_E \cap N[e_1] = C_E \cap N[e_2]$, alors e_2 doit être adjacente à e_1. Soient u leur sommet commun et v l'autre extrémité de e_2. Comme C_E couvre toutes les arêtes, il existe une arête $e_3 \in C_E$ adjacente à e_2 en v. Puisque G est sans triangle, e_3 n'est pas adjacente à e_1. Par conséquent, e_3 sépare e_1 et e_2.

Finalement, supposons que ni e_1 ni e_2 ne sont dans C_E. Alors il y a deux arêtes e_3 et e_4 qui sont adjacentes aux extrémités de e_1. Comme G ne contient ni $\mathcal{C}_3$ ni $\mathcal{C}_4$, e_2 ne peut pas être adjacente à e_3 et e_4 à la fois. Donc, e_1 et e_2 sont séparées. □

Dans la preuve précédente, nous utilisons l'absence de $\mathcal{C}_4$ seulement à la fin lorsque e_1, e_3, e_2 et e_4 devraient formé un $\mathcal{C}_4$ qui n'est adjacent à aucune autre arête de C_E. Nous avons donc le résultat suivant qui est plus fort que le lemme 3.14.

Lemme 3.15 *Soit G un graphe de maille au moins 4. Soit C_E un sous-ensemble d'arêtes de G, sans arêtes pendantes, qui domine tous les sommets de G. Si aucune paire d'arêtes isolées de C_E n'induit un $\mathcal{C}_4$ dans G, alors C_E est un code arête-identifiant de G.*

La proposition suivante donne la valeur de γ^{EID} pour certaines classes de graphes.

Proposition 3.16 *Nous avons :*

$$\gamma^{EID}(K_n) = \begin{cases} 5 & \text{si } n = 4 \text{ ou } 5, \\ n-1 & \text{si } n \geq 6. \end{cases}$$

De plus, si C_E est un code arête-identifiant de K_n de taille $n-1$ avec $n \geq 6$, alors il y a au plus un sommet non couvert par C_E et toutes les composantes connexes du graphe formé par C_E sont isomorphes à des cycles de taille au moins 5.

Preuve : Remarquons tout d'abord que $\mathcal{L}(K_4)$ est isomorphe au graphe complet K_6 privé d'un couplage parfait. D'après le théorème 2.33, le plus petit code identifiant de $\mathcal{L}(K_4)$ est de taille 5, donc $\gamma^{\text{EID}}(K_4) = 5$.

Quatre arêtes ne suffisent pas pour identifier K_5. En effet, les seuls graphes avec quatre arêtes sur cinq sommets sans arêtes pendantes sont $\mathcal{C}_4$ et P_5, et ces graphes n'identifient pas K_5. Par ailleurs les arêtes d'un cycle de taille 5 forment un code arête-identifiant de K_5 et donc $\gamma^{\text{EID}}(K_5) = 5$.

Pour $n \geq 6$, les arêtes d'un cycle de taille $n-1$ identifient toutes les arêtes de K_n. Aussi $\gamma^{\text{EID}}(K_n) \leq n-1$. Pour montrer $\gamma^{\text{EID}}(K_n) \geq n-1$, nous montrons la deuxième partie de la proposition, qui est un résultat plus fort.

Soit C_E un code arête-identifiant de K_n ($n \geq 6$) de taille minimale. D'après ce qui précède, $|C_E| \leq n-1$.

Soit G' le graphe ayant pour sommet $V(K_n)$ et pour arêtes C_E. Soient $G_1, \ldots, G_k$ les composantes connexes de G'. Comme G' a n sommets mais seulement $n-1$ arêtes, au moins une composante de G' est un arbre. Nous allons montrer qu'en fait *au plus* une composante de G' est un arbre et que cet arbre est réduit à un sommet isolé. Soit G_i une composante qui est un arbre. Si G_i a au moins trois sommets, alors d'après le lemme 3.12, il y a un sommet x de degré 1 dans G_i dont le voisin u est de degré 2. Soit v l'autre voisin de u. Les arêtes xv et uv de K_n ne sont alors pas séparées par C_E. Si G_i a seulement deux sommets x et y, alors pour tout autre sommet u, les arêtes ux et uy ne sont pas séparées. Ainsi toute composante qui est un arbre est en fait un sommet isolé, car toutes les paires d'arêtes de K_n sont séparées par C_E. S'il existe deux telles composantes, alors l'arête reliant les deux sommets correspondants n'est pas dominée.

Ainsi, exactement une composante de G', par exemple G_1, est un arbre et c'est un sommet isolé. Cela implique que $\gamma^{\text{EID}}(K_n) \geq n-1$ et donc $\gamma^{\text{EID}}(K_n) = n-1$. De plus, chaque composante de G' différente de G_1 contient exactement un cycle.

Il reste à montrer que chaque G_i pour $i \neq 1$ est isomorphe à un cycle de taille au moins 5. Supposons par exemple que G_2 ne soit pas un cycle. Comme G_2 contient un seul cycle, il y aurait un sommet v de degré 1 dans G_2. Soit t le voisin de v dans G_2 et soit u l'unique sommet de G_1. Les arêtes tu et tv ne seraient pas séparées par C_E, ce qui est contradictoire.

Toutes les composantes sont donc des cycles, de taille au moins cinq car $\mathcal{C}_3$ possède deux arêtes pendantes et que les deux cordes de $\mathcal{C}_4$ ne sont pas séparées par les arêtes de $\mathcal{C}_4$. □

Proposition 3.17 *Nous avons* $\gamma^{EID}(K_{n,n}) = \left\lceil \frac{3n-1}{2} \right\rceil$ *pour* $n \geq 3$.

Preuve : Soient X et Y les deux parties de $K_{n,n}$. Si n est pair, soit C_E un ensemble d'arêtes formé en prenant $\frac{n}{2}$ chemins P_4 disjoints. Si n est impair, C_E est formé avec une arête isolée et $\frac{n-1}{2}$ chemins P_4 disjoints qui sont disjoints de l'arête isolée. Alors l'ensemble C_E est un ensemble d'arêtes couvrant les sommets, avec au plus une arête isolée, et $K_{n,n}$ est sans triangle. D'après le lemme 3.15, C_E est un code arête-identifiant de $K_{n,n}$ de taille $\left\lceil \frac{3n-1}{2} \right\rceil$.

Montrons maintenant qu'il n'existe pas de meilleur code arête-identifiant. Soit C_E un code arête-identifiant de G. Soient $G_1, \ldots, G_k$ les composantes connexes du

graphes $(X \cup Y, C_E)$. Nous montrons que toutes les composantes G_i sauf peut-être une ont au moins quatre sommets. Il n'y a pas de graphe connexe sans arêtes pendantes sur trois sommets. Supposons que G_1 et G_2 ne contiennent toutes deux que deux sommets. Alors les arêtes reliant G_1 et G_2 ne sont pas séparées. Si G_1 est un sommet isolé et si G_2 est une arête isolée, alors l'unique arête reliant G_1 à G_2 n'est pas séparée de l'arête de G_2. Si G_1 et G_2 sont deux sommets isolés, alors les deux sommets doivent être du même côté (autrement l'arête reliant G_1 et G_2 existe et n'est pas dominée). Mais alors pour tout sommet x qui n'est pas dans la partie de G_1 et G_2, les arêtes reliant x à G_1 et x à G_2 ne sont pas séparées.

Sans perte de généralité, G_1 est la composante ayant éventuellement un ou deux sommets et pour tout $i \in \{2, \ldots, k\}$, G_i a au moins quatre sommets. Le nombre k de composantes vérifie donc $1 + 4(k-1) \leq 2n$, c'est-à-dire $k \leq \frac{n+1}{2}$. Chaque composante étant connexe, le nombre d'arêtes m_i de la composante G_i est au moins $|V(G_i)| - 1$. En sommant,

$$|C_E| = \sum_{i=1}^{i=k} m_i \geq 2n - k \geq 2n - \frac{n+1}{2} = \frac{3n-1}{2}.$$

□

3.3.3 Bornes inférieures

Nous donnons tout d'abord une borne inférieure en fonction du nombre de sommets, qui nous permet de calculer γ^{EID} pour l'hypercube. Cette borne donne aussi une borne en fonction du nombre d'arêtes, que l'on peut affiner, et qui, traduit en terme de graphes adjoints, améliore la borne d'ordre $\log(n)$ à une borne d'ordre $\sqrt{n}$.

Borne inférieure pour γ^{EID} en fonction du nombre de sommets

Proposition 3.18 *Soit G un graphe connexe sans arêtes pendantes. Alors,*

$$\gamma^{EID}(G) \geq \frac{|V(G)|}{2}.$$

Preuve : Soit C_E un code arête-identifiant de G. Soit X l'ensemble des sommets qui n'appartiennent à aucune arête de C_E. Soient G' le graphe induit par C_E sur les sommets de $V(G) \setminus X$ et $G_1, \ldots, G_s$ les composantes connexes de G'.

Soient n_i l'ordre de G_i et k_i son nombre d'arêtes, $\sum_{i=1}^{s} k_i = |C_E|$. Soit n'_i le nombre de sommets de X qui sont adjacents à un sommet de G_i dans le graphe initial G. Chaque sommet de X a au moins un voisin (le graphe est connexe) et tous ses voisins sont hors de X (pour des raisons de domination). Aussi, $|X| \leq \sum_{i=1}^{s} n'_i$. De plus, un sommet de G_i a au plus un voisin dans X, donc $n_i \geq n'_i$.

Montrons que $n'_i + n_i \leq 2k_i$. Si $k_i = 1$ (G_i est une arête isolée), $n_i = 2$ et $n'_i = 0$, donc $n'_i + n_i = 2k_i$. Si G_i est un arbre sur au moins trois sommets, alors $n_i = k_i + 1$, et d'après le lemme 3.12, G_i a au moins deux sommets de degré 2 ayant un voisin de degré 1. Aucun sommet de X ne peut alors être adjacent à l'un de ces deux sommets. De plus, tout autre sommet de G_i est adjacent à au plus un sommet dans

X. Donc $n'_i \leq n_i - 2 = k_i - 1$ et finalement $n_i + n'_i \leq 2k_i$. Si G_i n'est pas un arbre, alors $n_i \leq k_i$ et $n'_i \leq n_i$, donc $n'_i + n_i \leq 2k_i$.

Finalement, en comptant les sommets de G,

$$|V(G)| \leq |X| + |V(G) \setminus X| \leq \sum_{i=1}^{s}(n_i + n'_i) \leq 2\sum_{i=1}^{s} k_i \leq 2|C_E|.$$

□

Cette proposition a quelques conséquences sur l'optimalité de certains codes. La proposition 3.18 avec le lemme 3.14 donne le résultat suivant :

Corollaire 3.19 *Soit G un graphe sans arête pendante de maille au moins 5. Si G possède un couplage parfait M, alors M est un code arête-identifiant optimal et $\gamma^{EID}(G) = \frac{|V(G)|}{2}$.*

Cela montre par exemple que le code arête-identifiant du graphe de Petersen donné dans la figure 3.8 est optimal. Ce résultat peut être étendu pour les graphes sans triangle qui ont un couplage parfait dans lequel aucune paire d'arête n'induit un $\mathcal{C}_4$. Nous allons montrer que c'est le cas pour l'hypercube de dimension $d \geq 4$.

Les hypercubes ont une place centrale dans les problématiques de codage. Les codes identifiants ont été étudiés depuis l'introduction des codes identifiants et de nombreuses études ont suivies, beaucoup de questions à ce sujet sont encore ouvertes aujourd'hui. Pour en savoir plus, nous renvoyons à la thèse de Junnila [60] qui donne un aperçu des travaux en cours sur le domaine, ainsi qu'aux récents travaux de Charon *et al.* [17] qui améliorent certaines bornes.

Dans le cas des codes arête-identifiants, le problème est beaucoup plus facile et nous sommes en mesure de donner une valeur exacte.

Proposition 3.20 *Pour $d \geq 4$, $\gamma^{EID}(\mathcal{H}_d) = 2^{d-1}$.*

Preuve : Dans cette preuve, nous voyons l'hypercube $\mathcal{H}_d$ comme deux hypercubes de dimension $d-1$, $\mathcal{H}_0$ et $\mathcal{H}_1$, induit par les sommets dont les premières coordonnées sont respectivement 0 et 1. À une arête de $\mathcal{H}_d$ est associée une *direction* correspondant à la coordonné sur laquelle les deux sommets de l'arête diffèrent. Par exemple, toutes les arêtes entre $\mathcal{H}_0$ et $\mathcal{H}_1$ sont dans la première direction.

D'après la proposition 3.18, $\gamma^{\text{EID}}(\mathcal{H}_d) \geq 2^{d-1}$. Pour $d \geq 4$, nous construisons par induction un couplage parfait M_d de $\mathcal{H}_d$ tel qu'aucune paire d'arêtes n'induise un $\mathcal{C}_4$. D'après le lemme 3.15, M_d sera un code arête-identifiant de $\mathcal{H}_d$, prouvant le résultat.

Deux couplages parfaits disjoints de $\mathcal{H}_4$ vérifiant cette propriété sont montrés sur la figure 3.9.

Le couplage parfait M_5 est construit à partir de ces deux couplages parfaits de $\mathcal{H}_4$, un par copie de $\mathcal{H}_4$. Aucune paires d'arête de M_5 n'induit dans $\mathcal{H}_5$ un cycle $\mathcal{C}_4$ car les deux couplages sont disjoints. De plus, M_5 à la propriété de n'avoir aucune arête dans la première direction.

Nous construisons maintenant un couplage parfait M_d de $\mathcal{H}_d$ pour $d \geq 6$ par induction de telle sorte qu'aucune paire d'arêtes de M_d n'appartienne à un $\mathcal{C}_4$ de $\mathcal{H}_d$

et qu'aucune arête xy de M_d ne soit sur la direction correspondant à la première coordonnée.

Pour cela, soient $\mathcal{H}_0$ et $\mathcal{H}_1$ les deux copies de $\mathcal{H}_{d-1}$ dans $\mathcal{H}_d$ correspondant aux sommets de premières coordonnées respectives 0 et 1. Les arêtes de M_d sont celles de M_{d-1} sur $\mathcal{H}_0$ et celles de $\phi(M_{d-1})$ sur $\mathcal{H}_1$ où ϕ est l'automorphisme de $\mathcal{H}_{d-1}$ qui inverse la première coordonnée : $\phi(x) = x + (1, 0, \ldots, 0)$. Alors M_d ne contient aucune arête dont la direction est la première coordonnée. De plus, le couplage parfait $\phi(M_{d-1})$ est totalement disjoint de M_{d-1} car dans M_{d-1} il n'y a pas d'arêtes sur la première direction. En prenant une arête dans chaque copie de ces deux couplages parfaits, aucun $\mathcal{C}_4$ ne peut être induit dans $\mathcal{H}_d$. □

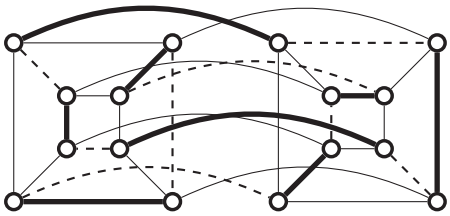


Figure 3.9 – *Deux couplages parfaits disjoints de $\mathcal{H}_4$ (en pointillé et en gras). Ces deux couplages ont la propriété qu'aucune paire d'arêtes au sein d'un même couplage n'induit un $\mathcal{C}_4$. Ce sont des codes arête-identifiant de $\mathcal{H}_4$.*

La formule de la proposition 3.20 n'est pas vrai pour $d = 2$ et $d = 3$. En effet, $\mathcal{H}_2$ correspond à $\mathcal{C}_4$ et donc $\gamma^{\text{EID}}(\mathcal{H}_2) = 3$. Pour $d = 3$, supposons qu'il existe un code arête-identifiant de taille 4. Si ce code est un couplage parfait et aucune paire du couplage de doit induire un cycle de taille 4, une petite étude de cas montre que ce n'est pas possible. Sinon c'est nécessairement un P_4 et une arête isolée (un P_5 laisserait une arête non dominée), et les deux sommets qui ne sont pas dans une des arêtes du code sont tous deux dans le voisinage d'un des sommets de l'arête isolée, ce qui est contradictoire. Donc $\gamma^{\text{EID}}(\mathcal{H}_3) \geq 5$. Un code arête-identifiant de taille 5 est donné sur la figure 3.10, montrant que $\gamma^{\text{EID}}(\mathcal{H}_3) = 5$.

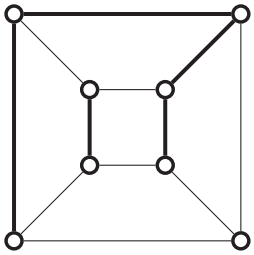


Figure 3.10 – *Un code arête-identifiant optimal de $\mathcal{H}_3$.*

Borne inférieure sur γ^{EID} en fonction du nombre d'arêtes

Le nombre d'arêtes d'un graphe est borné supérieurement par son nombre de sommets, la proposition 3.18 nous donne donc une borne sur γ^{EID} en fonction du nombre d'arêtes du graphe :

Proposition 3.21 *Soit G un graphe sans arêtes pendantes possédant m arêtes. Alors*

$$\gamma^{EID}(G) \geq \frac{\sqrt{8m+1}+1}{4} > \frac{\sqrt{2}}{2}\sqrt{m}.$$

Preuve : Soit n le nombre de sommets de G. On a $2m \leq n^2 - n$ et donc $n \geq \frac{\sqrt{8m+1}+1}{2} > \sqrt{2m}$. La proposition 3.18 nous permet de conclure. □

Nous allons affiner cette borne pour obtenir un coefficient $\frac{3\sqrt{2}}{4}$ à la place de $\frac{\sqrt{2}}{2}$. La borne obtenue sera serrée.

Soit G un graphe sans arêtes pendantes et soit C_E un code arête-identifiant de G. Nous supposons que G est connexe pour éviter les cas triviaux comme les sommets isolés. Cela n'implique pas que le graphe partiel de G dont les arêtes sont C_E soit connexe. En fait, dans la plupart des cas, un code arête-identifiant de taille optimale induit un graphe partiel non connexe. Nous montrons tout d'abord une borne inférieure lorsque le code induit un sous-graphe partiel connexe.

Proposition 3.22 *Soit C_E un code arête-identifiant connexe d'un graphe G, non isomorphe à K_2. Alors G contient au plus $\binom{|C_E|+2}{2} - 4$ arêtes. La borne est serrée uniquement lorsque C_E induit un chemin.*

Preuve : Soit G' le sous-graphe partiel constitué par les arêtes de C_E, les sommets de G' étant uniquement les sommets des arêtes de C_E.

Comme G' est supposé connexe et que G' n'a pas d'arêtes pendantes, G' a au moins quatre sommets et C_E a au moins trois arêtes. Pour chaque sommet x de G', soit C_x l'ensemble des arêtes incidentes à x dans G'. Soit $e = uv$ une arête de G. Alors u ou bien v est dans le graphe G' et e est uniquement déterminée par C_u, C_v ou bien $C_u \cup C_v$. Le nombre total d'ensembles de cette forme est $\binom{|V(G')|+1}{2}$. Si $|V(G')| \leq |C_E|$, il y a donc au plus $\binom{|C_E|+1}{2} \geq \binom{|C_E|+2}{2} - 4$ arêtes (car $|C_E| \geq 3$).

Sinon, comme G' est connexe, $|V(G')| = |C_E| + 1$ et G' est un arbre sans arêtes pendantes sur au moins quatre sommets. Si v est un sommet de degré 1 adjacent à u, alors $C_v = \{uv\}$ mais $uv \in C_u$ donc $C_u = C_u \cup C_v$. Par ailleurs, d'après le lemme 3.12, il y a deux sommets de degré 2 ayant un voisin de degré 1. Soit u un tel sommet, v son voisin de degré 1 et x son autre voisin. Alors $C_v = \{uv\}$ et $C_u = \{uv, ux\}$. Par conséquent, $C_u \cup C_x = C_v \cup C_x$. Finalement, le nombre total d'ensembles différents de la forme C_y ou $C_y \cup C_z$ est au plus $\binom{|C_E|+2}{2} - 4$. S'il y a égalité, alors il y a au plus deux sommets de degré 1 et G' est donc un chemin. □

Cette borne est serrée, par exemple lorsque G' est un chemin. De plus, pour tout chemin P_{k+1}, il est possible de construire un graphe dont P_{k+1} est un code arête-identifiant et qui a $\binom{k+2}{2} - 4$ arêtes. L'ensemble de tous les graphes vérifiant cela est noté $\mathcal{J}_k$. Un exemple d'élément de $\mathcal{J}_k$ est obtenu en prenant le graphe K_{k+2} et en enlevant quatre arêtes comme montré sur la figure 3.11. Tous les autres membres de J_k sont obtenus à partir de cet exemple en dédoublant le sommet qui n'est pas dans le chemin, sans ajouter d'arêtes.

Nous considérons maintenant le cas où le sous-graphe formé par les arêtes de C_E n'est pas nécessairement connexe.

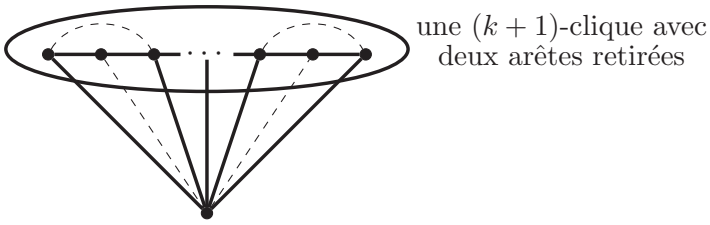


Figure 3.11 – *Un graphe extrémal de $\mathcal{J}_k$ avec son code arête-identifiant connexe.*

Théorème 3.23 *Soit G un graphe sans arêtes pendantes et soit C_E un code arête-identifiant de G de taille k. Alors,*

$$E(G) \leq \begin{cases} \binom{\frac{4}{3}k}{2}, & si\ k \equiv 0 \bmod 3 \\ \binom{\frac{4}{3}(k-1)+1}{2} + 1, & si\ k \equiv 1 \bmod 3 \\ \binom{\frac{4}{3}(k-2)+2}{2} + 2, & si\ k \equiv 2 \bmod 3 \end{cases}$$

Preuve : Soit G un graphe possédant le nombre maximum d'arêtes possible parmi les graphes tels que $\gamma^{\text{EID}}(G) = k$. Lorsque $k = 1, 2$ ou 3, le nombre d'arêtes de G est 1,3 ou 6 respectivement. Pour $k \geq 4$, nous prouvons tout d'abord le résultat suivant :

Étant donné un code arête-identifiant C_E de taille k de G, toutes les composantes connexes de C_E sauf deux sont isomorphes au graphe P_4. Lorsque il y a une seule composante non isomorphe à P_4, alors cette composante est isomorphe à P_2, P_5 ou P_6. S'il y a deux composantes non isomorphes à P_4, alors elles sont isomorphes soit à deux P_2, soit à P_2 et P_5, soit à deux P_5. Cela dépend de la valeur de k mod 3.

Soit G' le graphe dont les arêtes sont C_E et les sommets sont uniquement ceux appartenant à l'une des arêtes de C_E. Pour chaque sommet $u \in V(G) \setminus V(G')$, nous pouvons supposer que u est de degré 1. En effet si u est de degré $d > 1$, ses voisins $v_1, \ldots, v_d$ sont nécessairement dans $V(G')$. En remplaçant u par d sommets de degré 1, $u_1, \ldots, u_d$ et en reliant u_i à v_i, le nombre d'arêtes ne change pas et le code C_E reste un code arête-identifiant de taille k.

Soient $G'_1, G'_2, \ldots, G'_r$ les composantes connexes de G' et soit $n'_i = |V(G'_i)|$. Pour chaque $i \in \{1, \ldots, r\}$, soit G_i le graphe induit dans G par les sommets de G'_i et les sommets connectés seulement à G'_i. À chaque sommet x de G', nous attribuons l'ensemble C_x des arêtes de G' incidentes à x.

Remarquons tout d'abord qu'aucun G'_i n'est d'ordre 3 car il n'y a pas de graphe connexe sans arêtes pendantes sur trois sommets. Si u et v sont des sommets de deux composantes disjointes de G' d'ordre au moins 4, alors la paire $\{u, v\}$ est uniquement déterminée par $C_u \cup C_v$ et par maximalité de G, uv est une arête.

Si une composante de G' est une arête isolée uu' alors pour tout autre sommet v de G' appartenant à une composante de taille au moins 4, il y a exactement une arête parmi uv et $u'v$ qui est un arête de G. Si vv' est une autre arête isolée de G', alors il n'y a qu'une seule arête dans G parmi uv,uv',$u'v$ et $u'v'$.

Nous affirmons maintenant que chaque graphe G'_i avec $n'_i \geq 4$ est un chemin. Par l'absurde, si G'_i n'est pas un chemin, nous remplaçons G_i par un membre $J_{n'_i-1}$ de $\mathcal{J}_{n'_i-1}$ avec $P_{n'_i}$ comme code arête-identifiant. Chaque sommet de $P_{n'_i}$ est alors relié à chaque sommet de chaque G'_j, pour $j \neq i$ et $n'_j \geq 4$, et a exactement un sommet de G'_j lorsque $n'_j = 2$. Le nouveau graphe a toujours un code arête-identifiant de taille k, mais il a plus d'arêtes que G. En effet, le nombre d'arêtes reliant G'_i aux autres G'_j, $j \neq i$ n'a pas changé et le nombre d'arêtes de G_i augmente quand nous remplaçons G_i par $J_{n'_i-1}$, (proposition 3.22 appliquée à G_i).

Nous montrons maintenant que chaque composante G_i n'a pas plus de six sommets. Par l'absurde, supposons que G'_1 est une composante avec $n'_1 \geq 7$. Nous construisons un nouveau graphe G^*_1 à partir de G comme suit. Nous prenons des copies disjointes de $J_3 \in \mathcal{J}_3$ et $J_{n'_1-4} \in \mathcal{J}_{n'_1-4}$ avec P_4 et $P_{n'_1-3}$ comme codes arête-identifiants respectifs.

Soit maintenant $V(G^*_1) = V(J_3) \cup V(J_{n'_1-4}) \cup (V(G) \setminus V(G_1))$. Les arêtes de J_3, $J_{n'_1-4}$ et de $G - G_1$ sont des arêtes de G^*_1. Nous ajoutons des arêtes entre les trois parties comme suit. Chaque sommet de P_4 est relié à tous les sommets de $P_{n'_1-3}$. Pour $i = 2, 3, \ldots, r$ si $n'_i \geq 4$, chaque sommet de G'_i est relié à tous les sommets de $P_4 \cup P_{n'_1-3}$. Si $n'_i = 2$, nous choisissons exactement un sommet de G'_i et le relions à tous les sommets de $P_4 \cup P_{n'_1-3}$. La construction de G^*_1 assure qu'il y a toujours un code arête-identifiant de taille k dans G, mais il y a plus d'arêtes. En effet, le nombre d'arêtes augmente de deux manières. Premièrement, $P_4 \cup P_{n'_1-3}$ a un sommet de plus que G'_1, le nombre d'arêtes connectant $P_4 \cup P_{n'_1-3}$ à $G - G_1$ augmente donc (sauf si $r = 1$). Plus important, le nombre d'arêtes induites par $J_3 \cup J_{n'_1-4}$ est $6 + \binom{n'_1-2}{2} - 4 + 4(n'_1 - 3) = \frac{{n'_1}^2}{2} + \frac{3n'_1}{2} - 7$ ce qui est strictement plus grand que $|E(G'_1)| = \frac{{n'_1}^2}{2} + \frac{n'_1}{2} - 4$ pour $n'_1 \geq 3$. Comme $n'_1 \geq 7$, cela contredit la maximalité de G.

Avec une méthode similaire, les transformations suivantes augmentent le nombre d'arêtes du nouveau graphe qui admet toujours un code arête-identifiant de taille k :

1. Deux composantes de G' de six sommets sont transformées en deux graphes de $\mathcal{J}_3$ et un graphe de $\mathcal{J}_4$.
2. Une composante de G' sur six sommets et une composante sur cinq sommets sont transformées en trois graphes de $\mathcal{J}_3$.
3. Une composante de G' sur six sommets et une composante sur deux sommets sont transformées en deux graphes de $\mathcal{J}_3$.
4. Trois composantes de G' sur cinq sommets sont transformées en quatre graphes de $\mathcal{J}_3$.
5. Deux composantes de G' sur cinq sommets et une sur deux sommets sont transformées en trois graphes de $\mathcal{J}_3$.
6. Une composante de G' sur cinq sommets et deux sur deux sommets sont transformées en deux graphes de $\mathcal{J}_4$.
7. Trois composantes de G' sur deux sommets sont transformées en un graphe de $\mathcal{J}_3$.

Dans la preuve du cas 7, nous observons que le nombre d'arêtes identifiées par trois P_2 sera le même que le nombre d'arêtes identifiée par P_4. Comme $k \geq 4$, il y a une autre composante dans G'. Le nombre d'arêtes reliées à cette composante est plus grande dans le cas d'un P_4 (quatre sommets pour le P_4 et trois pour les trois P_2).

Les cas 1, 2 et 3 impliquent que si une composante de G' est isomorphe à P_6 alors toutes les autres composantes sont des P_4. Les cas 4, 5 et 6 impliquent que si une composante est isomorphe à un P_5 alors au plus une autre composante n'est pas un P_4 et est nécessairement un P_2 ou un P_5. Enfin, le cas 7 montre qu'il y a au plus deux composantes isomorphes à un P_2.

Pour obtenir le résultat du théorème, il faut compter le nombre d'arêtes de G en fonction des valeurs de k mod 3 et de la structure de G. Nous détaillons le cas $k = 0$ mod 3. Lorsque $k = 0$ mod 3, toutes les composantes de C_E sont des chemins P_4. Il y a donc $\frac{k}{3}$ composantes. Dans chaque composante G_i, il y a six arêtes : les trois du code et les trois arêtes hors du code que l'on peut ajouter. Entre deux composantes G_i, il y a 16 arêtes. Il y a $\binom{\frac{k}{3}}{2}$ manières de choisir les deux composantes de G'. Finalement, le nombre d'arêtes de G est

$$6\frac{k}{3} + 16\binom{\frac{k}{3}}{2} = \binom{\frac{4k}{3}}{2}.$$

□

Cette borne est serrée, les exemples extrémaux étant construits dans la preuve. Plus précisément, pour $k \equiv 0$ mod 3, un exemple atteignant la borne est construit en prenant $\frac{k}{3}$ copies disjointes de graphes de $\mathcal{J}_3$ ayant P_4 comme code arête-identifiant et en ajoutant une arête entre chaque sommet des P_4 venant de copies différentes. L'union de tous les P_4 forme alors un code arête-identifiant de taille minimale. Si $k \not\equiv 0$ mod 3, les exemples sont construits de la même manière en utilisant un élément de $\mathcal{J}_4$ ou $\mathcal{J}_5$ en plus.

Conséquence sur les graphes adjoints

Le théorème 3.23 se reformule dans le langage des graphes adjoints de la manière suivante :

Théorème 3.24 *Soit G un graphe adjoint sans jumeaux. Nous avons*

$$\gamma^{ID}(G) \geq \frac{3}{8}(1 + \sqrt{1 + 8|V(G)|}) > \frac{3\sqrt{2}}{4}\sqrt{|V(G)|}.$$

Ainsi, comme pour les graphes d'intervalles (voir le théorème 3.8), la taille minimale d'un code identifiant pour un graphe adjoint est bornée inférieurement par une fonction de l'ordre de $\sqrt{|V(G)|}$. Ce résultat peut-il être étendu à une classe plus large de graphes ? La raison fondamentale derrière ce résultat est l'idée (non formelle) suivante. Pour chacune des deux classes, un sommet peut être caractérisé à l'aide de deux cliques : pour les graphes d'intervalles, ces deux cliques correspondent aux voisinages à gauche et à droite du sommet considéré, pour les graphes adjoints, elles correspondent aux arêtes incidentes aux deux extrémités de l'arête correspondant au

sommet. Un sommet du code correspond donc à deux cliques-codes, et chaque sommet est identifié à l'aide d'une paire de ces cliques. Le nombre de sommets est donc quadratique par rapport à la taille du code identifiant. D'autres graphes vérifiant le même type de propriétés pourraient donc avoir le même type de borne.

Une autre manière d'étendre ce résultat pourrait être d'étudier des sur-classes définies par des sous-graphes induits interdits. En effet, la classe des graphes d'intervalle et celle des graphes adjoints sont, par définition, stables par sous-graphe induit. Pour chacune d'elles, il existe donc une caractérisation par sous-graphes induits interdits. Par exemple pour les graphes adjoints, Beineke [5] a donné une caractérisation des graphes adjoints avec neuf sous-graphes induits interdits. Une question naturelle est alors de savoir si la borne d'ordre $\sqrt{|V(G)|}$ est toujours vérifiée en interdisant moins de sous-graphes. Nous savons par exemple que ce n'est pas le cas en interdisant uniquement la griffe $K_{1,3}$ qui fait partie de la liste des neuf sous-graphes interdits pour les graphes adjoints. En effet, considérons le graphe G construit sur deux ensembles de sommets A et B de taille k et $2^k - 1$ respectivement. Les ensembles de sommets A et B induisent des graphes complets. Chaque sommet de B correspond à un sous-ensemble non trivial de A et est relié aux sommets correspondants. Le graphe G ne possède pas de griffes induites et a un code identifiant de taille $2k$ (en prenant dans le code les sommets de A et les sommets de B correspondant aux singletons). Il existe donc des familles infinies de graphes G sans griffes avec des codes identifiants dont la taille est de l'ordre de $\log(|V(G)|)$.

3.3.4 Bornes supérieures

Dans cette partie nous donnons des bornes supérieures sur la taille d'un code identifiant dans un graphe adjoint.

Pour les graphes adjoints, avec le nombre de sommets

Nous avons classifier au théorème 1.4 du chapitre 2 tous les graphes G connexes tels que $\gamma^{\text{ID}}(G) = |V(G)| - 1$. Seulement six de ces graphes sont des graphes adjoints, ce qui nous amène au résultat suivant :

Corollaire 3.25 *Soit G un graphe adjoint connexe qui ne fait pas partie des graphes suivants $P_3, P_4, C_4, P_4 \bowtie K_1, C_4 \bowtie K_1, \mathcal{L}(K_4)$, alors $\gamma^{ID}(G) \leq |V(G)| - 2$.*

Pour les codes arête-identifiants, avec le nombre d'arêtes

Un graphe est *k-dégénéré* si dans tout sous-graphe, il y a un sommet de degré inférieur ou égal à k. De manière équivalente, un graphe est k-dégénéré s'il existe un ordre des sommets $v_1, \ldots, v_n$ tel que v_i soit de degré au plus k dans le sous-graphe induit par $\{v_1, ..., v_i\}$.

Un code arête-identifiant de taille minimale est 2-dégénéré, et de ce fait, sa taille est limitée par le nombre de sommets du graphe.

Proposition 3.26 *Soit G un graphe sans arêtes pendantes et soit C_E un code arête-identifiant minimal de G pour l'inclusion. Alors le sous-graphe formé par les arêtes de C_E est 2-dégénéré.*

Preuve : Soit G' le graphe formé par les arêtes de C_E dont les sommets sont ceux de G. Soit uv une arête de G' avec $d_{G'}(u), d_{G'}(v) \geq 3$. Par minimalité de C_E, le sous-ensemble $C' = C_E \setminus \{uv\}$ n'est pas un code arête-identifiant de G. Or C' est un ensemble dominant les arêtes. Il existe donc deux arêtes, e_1 et e_2, qui ne sont pas séparées par C'. Au moins l'une d'entre elles, disons e_1, est incidente à u ou v et l'autre non. Nous considérons deux cas selon si $e_1 = uv$ ou si e_1 est incidente à seulement un des sommets u ou v.

Dans le premier cas, e_2 est adjacente à toutes les arêtes de C' auxquelles e_1 est adjacente. Comme chaque sommet de uv est incident à deux autres arêtes de C', le sous graphe induit par u, v et les sommets de e_2 est une clique de taille 4 et il n'y a pas d'autres arêtes de C' incidentes aux sommets de ce K_4 (voir la figure 3.12a).

Dans l'autre cas, supposons que e_1 soit incidente à u. Soient x et y deux voisins de u dans G' différents de v. Alors $e_2 = xy$ car e_2 doit être adjacente aux arêtes xu et yu mais pas incidente à u. Aussi le degré dans G' de u est exactement 3. Soit z l'autre extrémité de e_1. Nous considérons deux sous-cas selon si $z \notin \{x, y\}$ ou, sans perte de généralité, si $z = x$. Supposons tout d'abord que $z \notin \{x, y\}$. L'arête uv est la seule arête séparant e_1 et e_2. Mais e_1 doit être séparée de ux et uy. Aussi $zy, zx \in C_E$. De plus, les sommets x, y et z sont tous de degré 2 dans G' (voir la figure 3.12b). Supposons maintenant que $e_1 = ux$. Comme uv est la seul arête séparant e_1 et e_2, uy et xy sont les seules arêtes de G' incidentes à y. Le sommet y est donc de degré au plus 2 dans G', et u est de degré 3 (voir les figures 3.12c et 3.12d).

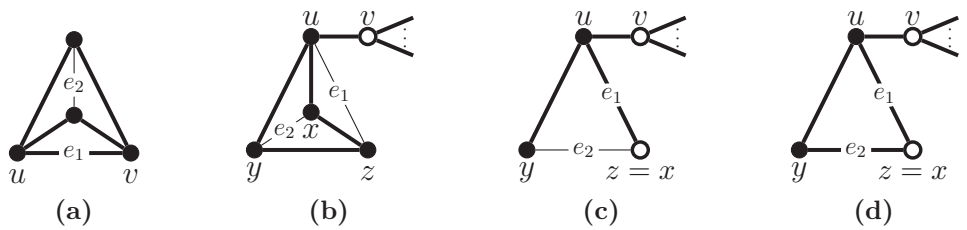


Figure 3.12 – *Distinctions de cas dans la preuve de la proposition 3.26. Les sommets noirs ont un degré fixé dans G'. Les arêtes en gras sont celles du code.*

Pour résumer, nous avons montré qu'étant donné une arête uv d'un code minimal arête-identifiant C_E, nous avons l'un des cas suivant (ou un cas symétrique en inversant le rôle de u et v) :

- L'un des sommets u ou v est de degré au plus 2 dans G'.
- L'arête uv est une arête d'une composante connexe de G' isomorphe à K_4^- (K_4 privé d'une arête), voir la figure 3.12a.
- Le degré de u est 3 et u est incident à un cycle C_4 dont les sommets sont de degrés 2 dans G' (figure 3.12b) ou a un sommet de degré 1 dans G' (figure 3.12c) ou encore à un triangle dont un des sommets est de degré 2 dans G' (figure 3.12d).

Dans chacun des cas, il existe un sommet de degré au plus 2 dans G' et lorsque ce sommet est supprimé du graphe u ou v devient de degré au plus 2 dans le graphe

restant. Cela définit donc un ordre d'élimination des sommets de G' montrant que G' est 2-dégénéré. □

Un graphe 2-dégénéré sur n sommets a un nombre d'arêtes limité. La proposition précédente nous donne ainsi une borne sur $\gamma^{\text{EID}}(G)$ en fonction de $|V(G)|$.

Corollaire 3.27 *Soit G un graphe sans arêtes pendantes. Alors :*

$$\gamma^{EID}(G) \leq 2|V(G)| - 3.$$

De plus, les seuls graphes à atteindre la borne sont K_2 et K_4.

Preuve : Notons $n = |V(G)|$. Soit C_E un code arête-identifiant minimal de G et soit G' le graphe dont les arêtes sont celles de C_E et les sommets sont ceux de G, $G' = (V(G), C_E)$. D'après la proposition 3.26, G' est 2-dégénéré. Soit $v_n, v_{n-1}, \ldots, v_1$ un ordre des sommets de G' obtenu en éliminant l'un après l'autre des sommets de degré au plus 2. Comme v_1 et v_2 peuvent induire au plus le graphe K_2, et qu'à chaque étape au plus deux arêtes sont ajoutées, il y a au plus $2n-3$ arêtes dans G' et donc $|C_E| \leq 2n-3$.

S'il y a exactement $2n-3$ arêtes dans C_E, cela signifie que les sommets enlevés sont toujours de degré exactement 2 et qu'il n'y a pas de sommet isolé. Si G n'est pas le graphe K_2, G a au moins quatre sommets. Le sous-graphe induit par v_1, v_2, v_3 et v_4 dans G' est isomorphe à K_4^- (le graphe complet privé d'une arête). Si G n'a que quatre sommets, alors G est soit le graphe K_4, soit le graphe K_4^-. Mais ce dernier a un code identifiant de taille 4. Si G n'est pas le graphe K_4, G a donc au moins cinq sommets. Le sommet v_5 est de degré 2 dans le graphe induit par les sommets de v_1 à v_5 dans G'. Par symétrie, il n'y a que trois sous-graphes possibles pour G' qui sont représentés sur la figure 3.13. Dans les trois cas, l'arête uv a ses deux extrémités de degré au moins 3 mais ne correspond à aucune des configurations décrites dans la preuve de la proposition 3.26 (voir la figure 3.12). Cela n'est pas possible car C_E est un code minimal, les seuls graphes ayant un code arête-identifiant minimal avec $2n-3$ arêtes sont donc les graphes K_2 et K_4.

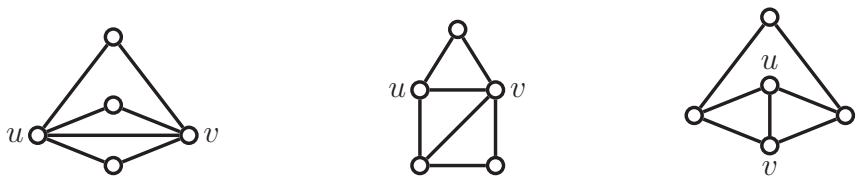


Figure 3.13 – *Les trois possibilités pour un graphe 2-dégénéré maximal sur cinq sommets.*

□

En poussant l'analyse plus loin, on peut montrer que le seul graphe ayant un code arête-identifiant de taille $2|V(G)|-4$ est le graphe K_4^-. Il existe par contre une infinité de graphes G tels que $\gamma^{\text{EID}}(G) = 2|V(G)|-6$, les graphes $K_{2,n}$ vérifiant cela. La borne précédente ne peut donc pas être beaucoup plus affinée.

Borne pour les graphes adjoints avec le degré maximum

Pour des graphes adjoints de graphes de degré moyen au moins 5, la conjecture 1.5 est vérifiée.

Corollaire 3.28 *Soit G un graphe à n sommets sans arêtes pendantes de degré moyen $\bar{d}(G) \geq 5$. Alors :*

$$\gamma^{ID}(\mathcal{L}(G)) \leq n - \frac{n}{\Delta(\mathcal{L}(G))}.$$

Preuve : Soit u un sommet de degré $d(u) \geq \bar{d}(G) \geq 5$. Comme G est sans arêtes pendantes, il y a au moins un voisin v de u qui est de degré au moins 2. L'arête uv dans G vérifie donc $d(u) + d(v) \geq \bar{d}(G) + 2$. Donc $\Delta(\mathcal{L}(G)) \geq \bar{d}(G)$. D'après le corollaire 3.27, il suffit donc de montrer que $2|V(G)| - 4 \leq |E(G)| - \frac{|E(G)|}{\bar{d}(G)}$.

Pour cela, comme $\bar{d}(G) \geq 5$, nous avons $4|V(G)| \leq (\bar{d}(G) - 1)|V(G)|$. Donc

$$4|V(G)| - 8 \leq (\bar{d}(G) - 1)|V(G)|.$$

En multipliant les deux côtés par $\frac{\bar{d}(G)}{2}$ nous obtenons :

$$(2|V(G)| - 4)\bar{d}(G) \leq (\bar{d}(G) - 1)\frac{\bar{d}(G)}{2}|V(G)| = (\bar{d}(G) - 1)|E(G)|.$$

□

La prochaine proposition montre que la borne supérieure de la conjecture 1.5 est serrée dans la classe des graphes adjoints. Ces exemples ont d'abord été introduits dans [40], nous les reprenons ici en terme de code arête-identifiant.

Proposition 3.29 (Foucaud et al., [40]) *Soit G un graphe k-régulier ($k \geq 3$). Soit G_1 le graphe obtenu à partir de G en subdivisant chaque arête de G exactement une fois. Alors*

$$\gamma^{EID}(G_1) = (k-1)|V(G)| = |E(G_1)| - \frac{|E(G_1)|}{2k-2} = |V(\mathcal{L}(G_1))| - \frac{|V(\mathcal{L}(G_1))|}{\Delta(\mathcal{L}(G_1))}.$$

Preuve : Soit x un sommet de G_1 de degré au moins 3, qui est donc un sommet du graphe de départ. Pour chaque arête e_i^x incidente à x, soit $e_i'^x$ l'arête adjacente à e_i^x mais pas incidente à x, et soit $A_x = \{e_i'^x\}_{i=1}^k$.

Alors $\{A_x \mid x \in V(G)\}$ est une partition des arêtes de $E(G_1)$. Pour un code arête-identifiant C_E de G_1, si deux éléments de A_x, $e_1'^x$ et $e_2'^x$ ne sont pas dans C_E alors e_1^x et e_2^x ne sont pas séparées. Aussi $|C_E \cap A_x| \geq k - 1$. Cela montre que $|C_E| \geq (k-1)|V(G)|$.

Pour trouver un code arête-identifiant de cette taille, nous choisissons exactement une arête dans chaque ensemble A_x, de telle sorte que que pour chaque sommet du graphe de départ, au plus une arête incidente à ce sommet soit choisie. L'ensemble des arêtes non choisies forment alors un code arête-identifiant. Pour choisir ces arêtes, nous considérons le graphe biparti B construit à partir de G en prenant dans chaque partie l'ensemble des sommets de G et en reliant deux sommets de

deux parties différentes s'ils sont adjacent dans G. Soient V_1 et V_2 les deux parties. Le graphe biparti H est k-régulier et, d'après le lemme des mariages de Hall (voir par exemple [11]), admet donc un couplage parfait M. Pour chaque sommet x, nous choisissons alors l'arête e_x de A_x qui est adjacente au sommet y tel que $xy \in M$, avec $x \in V_1$ et $y \in V_2$. Aucune paire d'arêtes sélectionnée ne sont incidentes à un même sommet de G. L'ensemble des arêtes restantes $C = E(G_1) \setminus \{e_x\}_{x \in V(G)}$ forme un code arête-identifiant de G_1. □

3.3.5 Complexité

$\mathcal{NP}$-complétude

Dans cette partie nous étudions la complexité du problème de décision associé au concept des codes arête-identifiants. Le problème EDGE-IDCODE est défini ainsi :

EDGE-IDCODE
INSTANCE : Un graphe G sans arêtes pendantes et un entier k.
QUESTION : G possède-t-il un code arête-identifiant de taille au plus k ?

Nous allons montrer que EDGE-IDCODE est $\mathcal{NP}$-complet en réduisant à partir du problème PLANAR $(\leq 3, 3)$-SAT qui est une variante du problème SAT définie dans [32].

PLANAR $(\leq 3, 3)$-SAT
INSTANCE : Une collection $\mathcal{Q}$ de clauses sur un ensemble X de variables booléennes, où chaque clause contient au moins deux et au plus trois littéraux distincts (une variable x ou sa négation $\overline{x}$). De plus, chaque variable apparaît exactement dans trois clauses, deux fois sous forme positive et une fois sous forme négative. Finalement, le graphe biparti d'incidence de $\mathcal{Q}$, noté $B(\mathcal{Q})$ est planaire ($B(\mathcal{Q})$ a comme ensemble de sommets $\mathcal{Q} \cup X$ et $Q \in \mathcal{Q}$ est adjacent à $x \in X$ si x ou $\overline{x}$ apparaît dans Q).
QUESTION : $\mathcal{Q}$ est-elle satisfiable ?

PLANAR $(\leq 3, 3)$-SAT est un problème $\mathcal{NP}$-complet [32]. Nous allons maintenant montrer le théorème suivant en réduisant PLANAR $(\leq 3, 3)$-SAT à EDGE-IDCODE.

Théorème 3.30 *EDGE-IDCODE est $\mathcal{NP}$-complet, même restreint aux graphes bipartis planaires de degré maximum* 3 *et de maille arbitrairement large.*

Le problème est clairement dans $\mathcal{NP}$: étant donné un sous-ensemble C d'arêtes de G, on peut vérifier en temps polynomial si C est un code arête-identifiant de G en calculant les ensembles $C \cap N[e]$ pour chaque arête e et en les comparant.

Nous montrons le résultat pour la maille 8 et expliquons ensuite comment étendre ce résultat à des mailles plus grande.

Pour la réduction, nous aurons besoin du gadget représenté par la figure 3.14 que nous appelons *P-gadget*. Pour avoir des figures plus compactes nous utiliserons la représentation du gadget dessinée sur la figure 3.14. Un P-gadget est *attaché* à

un sommet x si x est incident à l'arête a dessinée sur la figure. Dans toute la suite, les sommets du P-gadget, excepté le sommet x, ne seront adjacents à aucun autre sommet du graphe.

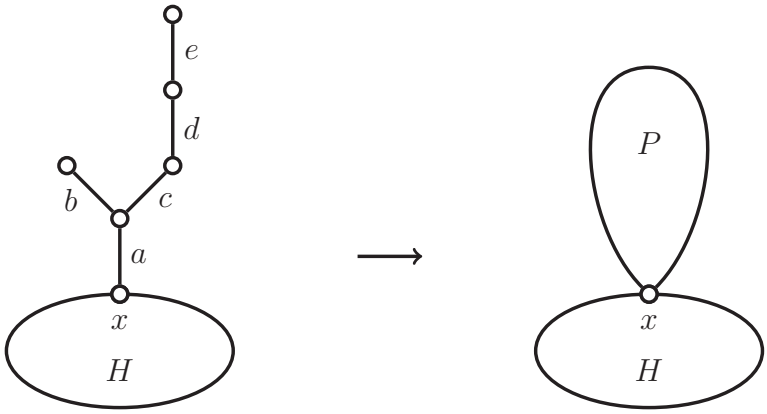


Figure 3.14 – *Le P-gadget et sa représentation.*

Lemme 3.31 *Dans tout graphe contenant un P-gadget, au moins trois arêtes de ce gadget doivent appartenir à tout code arête-identifiant.*

Preuve : L'arête d est la seule arête qui sépare b et c. De la même manière, c est la seule arête qui sépare d et e. Finalement, pour séparer d et c, il faut prendre au moins une arête parmi a, b et e. Il faut donc au moins trois arêtes. □

De plus, l'arête a doit être séparée de l'arête b :

Lemme 3.32 *Si G est un graphe sans arêtes pendantes obtenu à partir de H avec un P-gadget attaché au sommet x de H, alors tout code arête-identifiant de G doit contenir une arête de H incidente à x.*

Nous sommes maintenant armés pour réaliser la réduction. Étant donné une instance $\mathcal{Q} = \{Q_1, \ldots, Q_m\}$ de PLANAR $(\leq 3, 3)$-SAT sur les variables booléennes $X = \{x_1, \ldots, x_n\}$ avec une représentation planaire de son graphe d'incidence $B(\mathcal{Q})$, nous construisons le graphe $G_{\mathcal{Q}}$ comme suit.

Pour chaque variable x_j et chaque clause Q_i, nous construisons les sous-graphes G_{x_j} et G_{Q_i} respectivement, comme illustrés sur les figures 3.15 et 3.16. Rappelons qu'une variable donnée x_j apparaît dans sa forme positive dans exactement deux clauses, Q_p et Q_q et dans sa forme négative dans exactement une clause Q_r.

Nous unifions alors le sommet x_j^1 de G_{x_j} avec le sommet l_{p_k} de G_{Q_p} qui correspond à x_j. De la même manière, nous unifions les sommets x_j^2 et $\overline{x_j}^1$ avec les sommets correspondants dans G_{Q_q}et G_{Q_r}. Cela peut être fait en assurant la planarité de $G_{\mathcal{Q}}$, grâce à la représentation planaire de $B(\mathcal{Q})$. De plus $G_{\mathcal{Q}}$ est biparti car $B(\mathcal{Q})$ est biparti et qu'il n'y a pas de cycles impairs dans les gadgets des clauses. Finalement, $G_{\mathcal{Q}}$ est de degré maximum 3 et de maille 8. Comme le gadget d'une clause a 45 sommets et que le gadget d'une variable en compte 42, $G_{\mathcal{Q}}$ a $45m + 42n$ sommets, la construction est donc polynomiale.

Pour compléter la preuve nous avons besoin des deux résultats suivant.

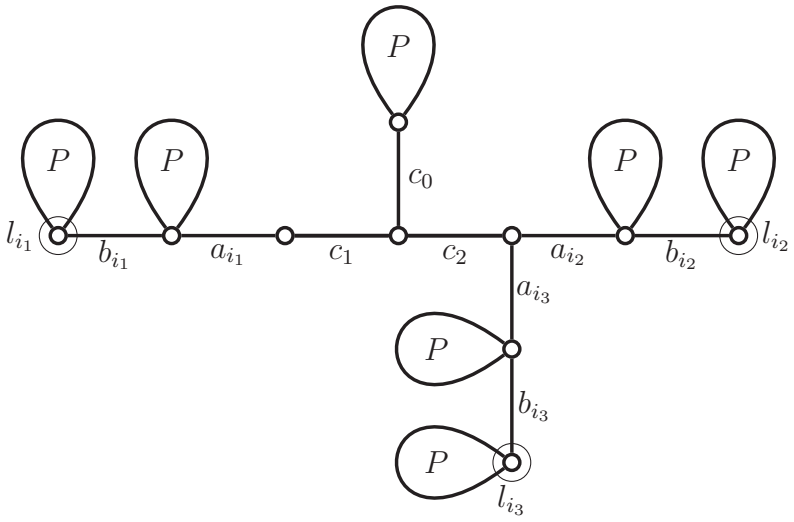


Figure 3.15 – *Gadget G_{Q_i} pour la clause $Q_i = l_{i_1} \cup l_{i_2} \cup l_{i_3}$.*

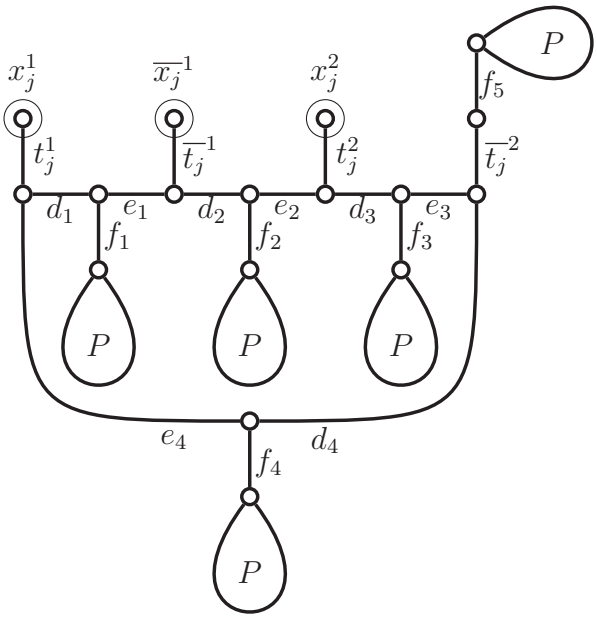


Figure 3.16 – *Gadget G_{x_j} pour la variable x_j.*

Lemme 3.33 *Dans le gadget G_{x_j} de la variable x_j, pour séparer les quatre paires d'arêtes $\{d_i, e_i\}$ pour $1 \leq i \leq 4$, au moins deux arêtes de $A = \{d_i, e_i \mid 1 \leq i \leq 4\} \cup \{t_j^1, \overline{t_j}^1, t_j^2, \overline{t_j}^2\}$ doivent appartenir à tout code arête-identifiant C. De plus, si $|C \cap A| = 2$, alors soit $C \cap A = \{t_j^1, t_j^2\}$, soit $C \cap A = \{\overline{t_j}^1, \overline{t_j}^2\}$.*

La première partie de ce lemme provient du fait que les deux arêtes de chacune des paires $\{d_1, e_1\}$ et $\{d_3, e_3\}$ doivent être séparées. La deuxième partie se montre avec une étude de cas.

Le résultat suivant est une conséquence du lemme 3.32 :

Lemme 3.34 *Soit $v_1v_2v_3v_4$ un chemin sur quatre sommets où chacun des sommets v_2 et v_3 sont de degré 3 avec un P-gadget attaché . Alors au moins une des trois arêtes du chemin est dans un code arête-identifiant du graphe. Si exactement une seule est dans le code, c'est v_2v_3.*

Lemme 3.35 *Une instance $\mathcal{Q}$ est satisfiable si et seulement si $G_{\mathcal{Q}}$ a un code arête identifiant de taille $k = 25m + 22n$.*

Preuve : Supposons dans un premier temps qu'il existe une affectation des variables qui satisfait $\mathcal{Q}$ et construisons un code arête-identifiant C de $G_{\mathcal{Q}}$ de taille $k = 25m + 22n$.

Pour chaque P-gadget, les arêtes a, c, d sont dans C. Pour chaque gadget de clauses, l'arête c_0 appartient à C. Pour chaque littéral l_{i_k} de Q_i, $1 \leq k \leq 3$, si l_{i_k} est vraie, l'arête a_{i_k} appartient à C ; sinon, l'arête b_{i_k} est dans C. Si Q_i a seulement deux littéraux et que le sommet l_{i_k} est le sommet qui ne correspond pas à un littéral de Q_i, alors l'arête b_{i_k} est dans C. Toutes les arêtes de G_{Q_i} sont alors dominées.

De plus, avec une exception pour la paire d'arêtes $\{c_1, c_2\}$, chaque arête de G_{Q_i} est séparée des autres arêtes. Dans chaque gadget de variable G_{x_j}, si x_j est vrai les arêtes t_j^1 et t_j^2 appartiennent à C. Sinon, les arêtes $\overline{t_j}^1$ et $\overline{t_j}^2$ sont dans C. Les arêtes f_1, f_2, f_3, f_4 et f_5 sont aussi dans C. Les arêtes de $G_{x_j} \setminus \{t_j^1, t_j^2, \overline{t_j}^1\}$ sont alors dominées. Comme chacune des trois arêtes $t_j^1, t_j^2, \overline{t_j}^1$ est incidente à un sommet d'un P-gadget d'un gadget de clause, elles sont toutes les trois dominées. De plus, toutes les paires d'arêtes contenant au moins une arête de $G_{x_j} \setminus \{t_j^1, t_j^2, \overline{t_j}^1\}$ sont clairement séparées. Maintenant, comme pour chaque P-gadget d'un gadget de clause, l'arête a est dans C, les arêtes $t_j^1, t_j^2, \overline{t_j}^1$ sont séparées de toutes les arêtes de $G_{\mathcal{Q}}$. Il reste à montrer que C sépare les paires $\{c_1, c_2\}$ de chaque gadget de clause. Comme nous avons considéré une affectation qui satisfait $\mathcal{Q}$, dans chaque clause Q_i de $\mathcal{Q}$ il existe un littéral vrai. D'où, pour chaque clause Q_i, au moins une arête a_{i_j} avec $1 \leq j \leq 3$ doit être dans le code et la paire $\{c_1, c_2\}$ est donc séparée. Nous pouvons donc conclure que C est un code arête-identifiant de taille k.

Pour montrer l'autre implication, soit C' un code arête-identifiant de $G_{\mathcal{Q}}$ avec $|C'| \leq k$. D'après le lemme 3.31, trois arêtes de chacun des sept P-gadgets d'un gadget de clause G_{Q_i} doit appartenir à C'. De plus, d'après le lemme 3.32, l'arête c_0 est obligatoirement dans C'. Finalement, avec le lemme 3.32, pour chaque sommet l_{i_k}, $1 \leq k \leq 3$ de G_{Q_i} une des arêtes a_{i_k} ou b_{i_k} est dans C'. Il y a donc au moins 25 arêtes dans le code pour chaque gadget de clause.

De la même manière, le lemme 3.31 implique que dans chaque gadget de variable G_{x_j}, au moins quinze arêtes de C' sont contenues dans les P-gadgets de G_{x_j}. D'après le lemme 3.32, toutes les arêtes f_i, $1 \leq i \leq 5$, sont dans C', ce qui fait un total d'au moins vingt arêtes dans chaque gadget de variable. Nous avons donc déjà $25m+20n$ arêtes de C' qui sont localisées. Pour les $2n$ arêtes restantes, le lemme 3.33 implique que pour chaque gadget de variable, deux arêtes additionelles sont dans C' (pour séparer les paires $\{d_i, e_i\}$). Aussi, dans chaque gadget de variable, exactement deux arêtes additionelles sont dans C'. D'après la deuxième partie du lemme 3.33, soit $\{t_j^1, t_j^2\}$, soit $\{\overline{t_j}^1, \overline{t_j}^2\}$ est un sous-ensemble de C'.

Nous avons maintenant retrouvé les $k = 25m + 22n$ arêtes de C'. Dans chaque gadget de clause G_{Q_i}, il y a donc *exactement* une des arêtes a_{i_k} et b_{i_k} de G_{Q_i} qui est dans C'.

Nous construisons maintenant un affectation des variables. Pour chaque gadget de variable, si $\{t_j^1, t_j^2\}$ est un sous-ensemble de C', x_j est mis à VRAI. Sinon $\{\overline{t_j}^1, \overline{t_j}^2\}$ est un sous-ensemble de C' et x_j est mis à FAUX. Montrons que cette affectation satisfait $\mathcal{Q}$.

Dans chaque gadget de clause G_{Q_i}, les arêtes c_1 et c_2 doivent être séparées par C', cela signifie qu'une arête a_{i_k} parmi $\{a_{i_1}, a_{i_2}, a_{i_3}\}$ est dans C'. Or, comme noté précédemment, $b_{i_k} \notin C'$ et d'après le lemme 3.34, dans le chemin formé par les arêtes $\{a_{i_k}, b_{i_k}, t_j^1\}$, t_j^1 est dans C' (sans perte de généralité, nous supposons que $l_{i_k} = x_j$ et que t_j^1 est l'arête de G_{x_j} incidente au sommet l_{i_k} de G_{Q_i}). Ainsi le litéral l_{i_k} est VRAI et la clause est satisfaite. En répétant ces arguments pour chaque clause, $\mathcal{Q}$ est satisfaite. □

Il reste à montrer que des arguments similaires peuvent être utilisés pour étendre la construction à des graphes de plus grande maille. Soient $\lambda \geq 1$ et $\mu \geq 2$ des entiers. Nous construisons le graphe $G_{\mathcal{Q}}(\lambda, \mu)$ en utilisant le gadget modifié de variable $G_{x_j}(\mu)$ et le gadget modifié de clause $G_{Q_i}(\lambda)$, qui sont représentés sur les figures 3.17 et 3.18. La construction est la même que précédemment et $G_{\mathcal{Q}}(\lambda, \mu)$ a $(36\lambda+9)m + (30\mu - 18)n$ sommets. La maille de $G_{\mathcal{Q}}(\lambda, \mu)$ est d'au moins $\min\{4\mu, 8(\lambda+1)\}$. En effet, $G_{x_j}(\mu)$ a un cycle de taille 4μ et comme la maille de $B(\mathcal{Q})$ est d'au moins 4, la taille minimale d'un cycle entre les gadgets de clause (au moins 2) et les gadgets de variables (au moins 2) est d'au moins $4(2\lambda+1)+2+2 = 8(\lambda+1)$.

En utilisant une preuve similaire à la précédente, nous pouvons montrer que $\mathcal{Q}$ est satisfiable si et seulement si $G_{\mathcal{Q}}(\lambda, \mu)$ a un code arête-identifiant de taille au plus $k = (21\lambda+4)m + (17\mu - 12)n$, ce qui termine la preuve du théorème 3.30.

Un graphe adjoint $\mathcal{L}(G)$ est parfait si et seulement si G n'a pas de cycles induits de taille impaire strictement plus grande que 3, voir [93] par exemple. Les graphes adjoints des graphes construits dans la preuve précédente sont parfaits, de degré maximum 4 et de nombre de clique 3 (et donc 3-coloriable). Cela implique le corollaire suivant :

Corollaire 3.36 *IDCODE est $\mathcal{NP}$-complet même restreint aux graphes parfaits adjoints 3-coloriables de degré maximum 4.*

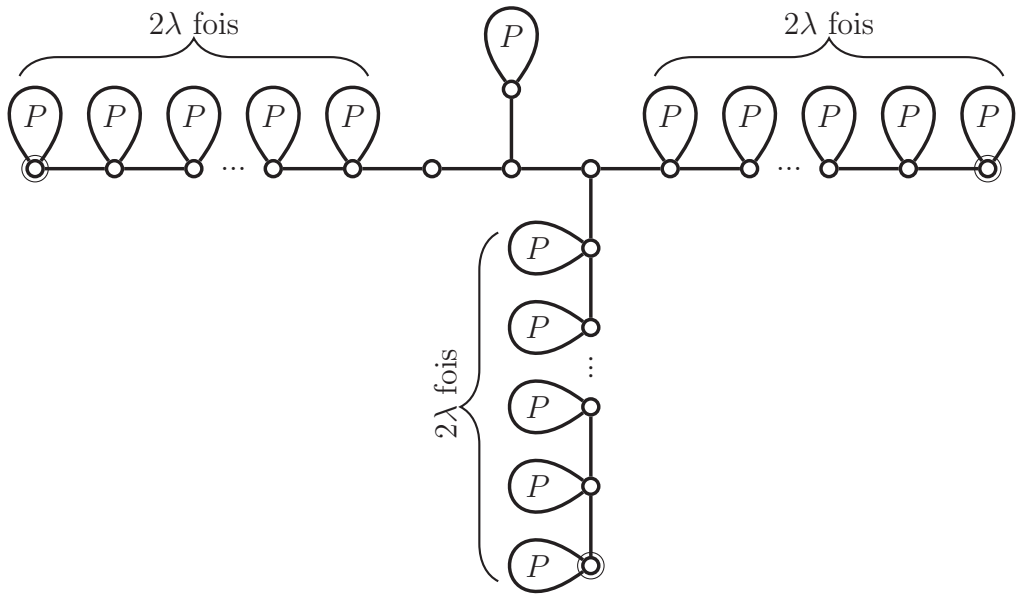


Figure 3.17 – *Gadget de clause $G_{Q_i}(\lambda)$ pour les grandes mailles.*

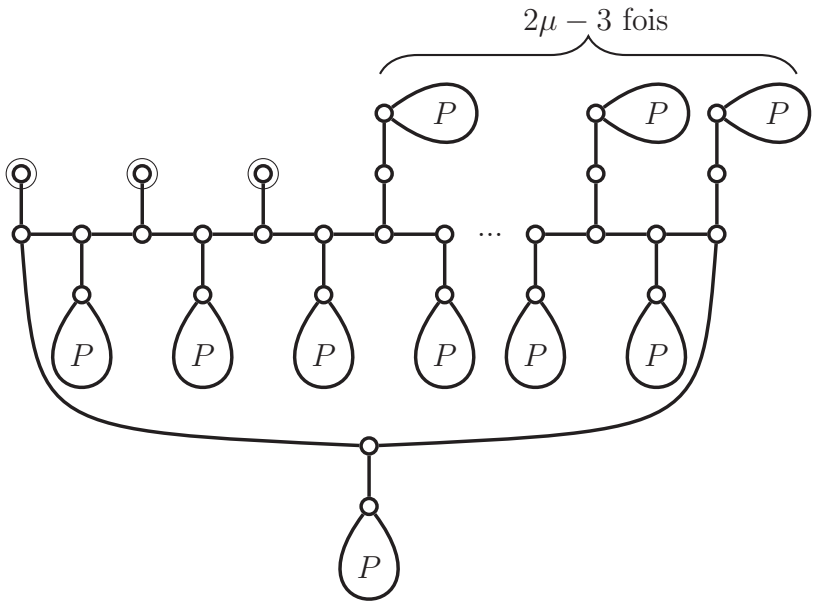


Figure 3.18 – *Gadget de variable $G_{x_j}(\mu)$ pour les grandes mailles.*

4-approximabilité

Au problème de décision EDGE-IDCODE est naturellement associé un problème d'optimisation MIN-EIDCODE :

MIN-EIDCODE
INSTANCE : Un graphe G sans arêtes pendantes.
QUESTION : Quelle est la taille du plus petit code arête-identifiant de G ?

Le problème EDGE-IDCODE étant lui même $\mathcal{NP}$-complet, nous ne pouvons pas espérer avoir un algorithme polynomial qui résout le problème MIN-EIDCODE. Néanmoins, nous pouvons obtenir facilement un algorithme donnant un code arête-identifiant de G dont la taille est au plus un facteur 4 de l'optimal.

Proposition 3.37 *Il existe une 4-approximation polynomiale au problème MIN-EIDCODE.*

Preuve : D'après la proposition 3.18, pour tout graphe G, $\gamma^{\text{EID}}(G) \geq \frac{|V(G)|}{2}$. Par ailleurs, le corollaire 3.27 dit que $\gamma^{\text{EID}}(G) \leq 2|V(G)| - 3$. Aussi, en approchant $\gamma^{\text{EID}}(G)$ par $2|V(G)| - 3$, l'erreur commise est au plus un facteur 4 de la valeur optimale.

De plus, il est effectivement possible de construire un code de taille $2|V(G)| - 3$ en un temps polynomial, il suffit de prendre un code arête-identifiant minimal pour l'inclusion car la proposition 3.26 est valable pour tout code arête-identifiant minimal pour l'inclusion. De manière gloutonne, un tel code peut être construit en partant de l'ensemble de toutes les arêtes, puis en retirant des arêtes les unes après les autres, tant qu'il est possible d'en retirer une et que le code reste arête-identifiant (ce qui se teste en temps polynomial). □

Ainsi, le problème des codes identifiants est 4-approximable dans la classe des graphes adjoints. Pour les graphes quelconques, il existe une $O(\log(n))$-approximation polynomiale pour le problème des codes identifiants, n étant le nombre de sommets du graphe (résultat montré indépendamment dans [45, 69, 91]). De plus, il n'existe pas de $(1 - \epsilon)\log(n)$-approximation polynomiale pour le problème général (Laifenbeld *et al.* [69]).

Utilisation de la logique monadique du second ordre

Nous allons maintenant voir que le problème des codes arête-identifiants devient linéaire pour des classes restreintes de graphes. Pour cela nous utilisons le théorème de Courcelle et la logique monadique du second ordre. Nous rappelons tout d'abord les bases nécessaires (voir [29] pour plus de détails).

Une propriété $\mathcal{P}$ sur les graphes est exprimable en *logique monadique du second ordre avec comptage* (CMSOL), si $\mathcal{P}$ peut être définie en utilisant :

- Les sommets, arêtes, sous-ensembles de sommets ou d'arêtes du graphe.
- La relation binaire d'adjacence *adj* avec $adj(u, v)$ vraie si et seulement si u et v sont des sommets adjacents.

- La relation binaire d'incidence *inc* avec $inc(u, e)$ vraie si et seulement si e est incidente à u.
- L'opérateur $=$ pour les sommets et les arêtes.
- La relation d'appartenance $\in$ pour vérifier si un élément appartient à un ensemble.
- L'opérateur *card* de cardinalité pour les ensembles de sommets.
- Les opérateurs logiques OR, AND et NOT, dénotés par $\vee$, $\wedge$, $\neg$.
- Les quantificateurs logiques $\exists$, $\forall$ sur les sommets, arêtes, sous-ensembles de sommets et d'arêtes.

La logique monadique du second ordre est particulièrement utile pour obtenir des résultats de complexité sur les graphes de largeur d'arborescence bornée. Un graphe G est de *largeur d'arborescence* k s'il existe des ensembles $X_1, ..., X_s$ et un arbre T dont les nœuds sont les ensembles X_i tels que :

- chaque ensemble X_i est de taille au plus k,
- pour toute arête uv de G, il existe un ensemble X_i contenant les deux sommets u et v,
- pour chaque sommet u, l'ensemble des X_i contenant u est non vide et forme un sous-arbre de l'arbre T.

De manière alternative ce sont aussi des sous-graphes partiels des k-arbres, que nous utiliserons au chapitre 5. Les arbres sont exactement les graphes connexes de largeur d'arborescence 1, les graphes séries-parallèles ainsi que les graphes planaires extérieurs sont de largeur d'arborescence 2.

Toute propriété exprimable en CMSOL est vérifiable en temps linéaire sur les graphes de largeur d'arborescence bornée :

Théorème 3.38 (Courcelle [29]) *Soit $\mathcal{P}$ une propriété sur les graphes exprimables en CMSOL et soit c une constante. Alors pour tout graphe G de largeur d'arborescence au plus c, il peut être vérifié en temps linéaire si G a la propriété $\mathcal{P}$.*

Être un code identifiant de taille k est exprimable en CMSOL (voir par exemple la thèse de Moncel [76]), ce qui implique par exemple qu'il existe un algorithme linéaire pour les arbres. C'est aussi le cas pour les codes arête-identifiants :

Proposition 3.39 *Soient $G = (V, E)$ un graphe et k un entier. Soit $\mathcal{EID}(G, k)$ la propriété $\gamma^{EID}(G) \leq k$. La propriété $\mathcal{EID}(G, k)$ est exprimable en CMSOL.*

Preuve : Nous définissons tout d'abord en CMSOL la relation $dom(e, f)$ qui est vraie si et seulement si e et f sont des arêtes de E et si e et f sont adjacentes, c'est-à-dire si e et f ont un sommet en commun :

$$dom(e, f) := \exists x \in V, (inc(x, e) \wedge inc(x, f)).$$

Nous définissons alors $\mathcal{EID}(G, k)$ de la manière suivante :

$$\mathcal{EID}(G, k) := \exists C, C \subseteq E, card(C) \leq k, \big(\forall e \in E, \exists f \in C, dom(e, f)\big) \wedge$$
$$\forall e \in E, \forall f \in E, e \neq f, \exists g \in C, \big((dom(e, g) \wedge \neg dom(f, g)) \vee (dom(f, g) \wedge \neg dom(e, g))\big)$$

□

Corollaire 3.40 *EDGE-IDCODE est un problème résoluble en temps linéaire pour toutes les classes de graphes dont la largeur d'arborescence est bornée par une constante.*

En particulier, il existe un algorithme linéaire pour résoudre le problème des codes arête-identifiants dans les arbres. La preuve du théorème 3.38 est constructive et donne donc cet algorithme, mais il est très technique. Il serait donc intéressant d'avoir un algorithme linéaire simple pour ce problème sur les arbres, comme cela a été fait par Auger[1] pour les codes identifiants classiques.

3.4 Grille du roi

Pour terminer ce chapitre, nous nous intéressons aux codes $(1, \leq 2)$-identifiants dans la grille du roi. La grille du roi est le graphe dont les sommets sont les éléments de $\mathbb{Z}^2$ et où deux sommets sont voisins si leur distance euclidienne est au plus $\sqrt{2}$ (voir la figure 3.19). Un code $C \subseteq \mathbb{Z}^2$ est $(r, \leq \ell)$-identifiant si pour tout ensemble X de sommets de taille au plus ℓ les ensembles $(\bigcup_{v \in X} N[v]) \cap C$ sont uniques.

La mesure utilisée pour étudier des codes dans les grilles infinies est la *densité* plutôt que le cardinal du code qui est infini. Soit Q_n l'ensemble de sommets de $\mathbb{Z}^2$ dont les coordonnées sont comprises entre $-n$ et n :

$$Q_n = \{(u, v) \in \mathbb{Z}^2 \mid |u| \leq n, |v| \leq n\}.$$

Soit $C \subseteq \mathbb{Z}^2$ un code de $\mathbb{Z}^2$. La *densité* de C, notée $D(C)$, est définie par :

$$D(C) = \limsup_{n \to +\infty} \frac{|C \cap Q_n|}{|Q_n|}.$$

Honkala et Laihonen ont donné dans [55] un code $(1, \leq 2)$-identifiant de la grille du roi de densité $\frac{3}{7} = 0.4286$ qui est représenté sur la figure 3.19. C'est le meilleur code connu à ce jour. Par ailleurs, Pelto [81] a montré qu'un code $(1, \leq 2)$-identifiant dans la grille du roi est de densité au moins $\frac{5}{12} = 0.41\overline{6}$.

Théorème 3.41 (Pelto [81], Honkala, Laihonen [55]) *Soit d la densité optimale d'un code $(1, \leq 2)$-identifiant de la grille du roi. Alors*

$$0,41\overline{6} = \tfrac{5}{12} \leq d \leq \tfrac{3}{7} = 0,4286...$$

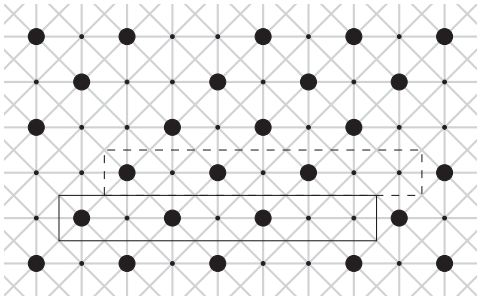


Figure 3.19 – *La grille du roi avec un code $(1, \leq 2)$-identifiant de densité $\frac{3}{7}$.*

Nous améliorons la borne inférieure $\frac{5}{12} = 0,41\overline{6}$ connue sur la densité optimale d'un code $(1, \leq 2)$-identifiant dans la grille du roi en montrant le résultat suivant :

Théorème 3.42 *La densité optimale d'un code $(1, \leq 2)$-identifiant de la grille du roi est d'au moins $\frac{47}{111} = 0,\overline{423}$.*

L'écart entre la borne inférieure et la borne supérieure est donc réduit de moitié. Nous exposons tout d'abord le principe général de la méthode utilisée. Nous appliquons cette méthode pour montrer de manière simple la borne inférieure de Pelto [81]. Nous donnons ensuite les détails de notre analyse plus fine qui permet d'améliorer la borne inférieure. Ces travaux ont été réalisés durant mon séjour à Turku avec Florent Foucaud et Tero Laihonen [41].

3.4.1 Méthode

Principe général de la démonstration

Pour établir des bornes inférieures sur la densité d'un code dans une grille, une technique courante est d'utiliser une *fenêtre*, c'est-à-dire un sous-ensemble fini de sommets, que l'on fait glisser sur la grille. Plus précisément, on suit le procédé suivant :

1. Fixer un code C quelconque.
2. Choisir un sous-ensemble fini de sommets X.
3. Compter le nombre N_{min} d'éléments du code qu'il y a au minimum dans une translation de X.
4. En déduire que la densité du code est d'au moins $\frac{N_{min}}{|X|}$.

Souvent, si dans une translation de X il y a N_{min} éléments du code, alors il y a une translation de X proche qui contient plus d'éléments du code. Nous utilisons alors des règles locales pour déterminer le *nombre moyen d'éléments d'un code C dans une translation de X*, noté $n(X, C)$. La quantité $n(X, C)$ est définie formellement par :

$$n(X, C) = \limsup_{n\to\infty} \frac{\sum_{v\in Q_n} |(v + X) \cap C|}{|Q_n|}.$$

La méthode se transforme ainsi :

1. Fixer un code C quelconque.
2. Choisir un sous-ensemble X de sommets finis.
3. Trouver une borne inférieure N_{moy} sur $n(X, C)$ en utilisant des règles locales de déchargement :
   (a) Donner une charge à chaque translation de X correspondant au nombre de sommets dans X.
   (b) Redistribuer *localement* les charges pour que chaque translation de X est une charge d'au moins N_{moy}.
   (c) En déduire que $n(X, C) \geq N_{moy}$.
4. En déduire que la densité du code est d'au moins $\frac{N_{moy}}{|X|}$.

Pour la dernière étape, la proposition suivante relie la quantité $n(X, C)$ avec la densité de C.

Proposition 3.43 *Soient C un code de $\mathbb{Z}^2$ et X un ensemble non vide fini de $\mathbb{Z}^2$. Alors :*

$$D(C) = \frac{n(X, C)}{|X|}.$$

Preuve : Soient h un entier tel que $X \subseteq Q_h$ et n un entier positif. Un élément de $(v+X)\cap C$ pour $v \in Q_{n+h}$ est soit dans $C\cap Q_n$, soit dans $Q_{n+2h} \setminus Q_n$. Parmi tous les ensembles $v + X$ avec $v \in Q_{n+h}$, chaque élément de $C \cap Q_n$ est compté exactement $|X|$ fois et chaque élément de $Q_{n+2h} \setminus Q_n$ est compté au plus $|X|$ fois. Nous avons donc :

$$|X||C \cap Q_n| \leq \sum_{v \in Q_{n+h}} |(v+X) \cap C| \leq |X|(|C \cap Q_n| + |Q_{n+2h} \setminus Q_n|).$$

En divisant chaque terme par $|Q_n|$, nous obtenons :

$$|X| \cdot \frac{|C \cap Q_n|}{|Q_n|} \leq \frac{\sum_{v \in Q_{n+h}} |(v+X) \cap C|}{|Q_n|} \leq |X| \cdot \frac{|C \cap Q_n|}{|Q_n|} + |X| \cdot \frac{|Q_{n+2h} \setminus Q_n|}{|Q_n|}.$$

Comme h est un entier fixé, $\limsup_{n \to \infty} \frac{|Q_{n+2h} \setminus Q_n|}{|Q_n|} = 0$. Aussi lorsque $n \to \infty$,

$$\limsup_{n \to \infty} \frac{\sum_{v \in Q_{n+h}} |(v+X) \cap C|}{|Q_n|} = |X| \cdot D(C).$$

Pour finir la preuve, il suffit de vérifier que

$$\limsup_{n \to \infty} \frac{\sum_{v \in Q_{n+h}} |(v+X) \cap C|}{|Q_n|} = n(X, C).$$

□

L'essentiel de l'étude consiste donc à choisir un bon ensemble X et à étudier finement la disposition des éléments du code dans X pour pouvoir distribuer les charges (étape 3*b*). L'étape 3*c* se justifie de la même manière que dans la preuve de la proposition 3.43 à condition que :

- Il n'y ait pas de perte de charges.
- La quantité de charge qui transite soit bornée pour chaque translation de X.
- Chaque translation de X reçoive ou donne des charges à d'autres translations dans un rayon borné.

Simplification du problème en un problème de couverture d'ensembles

Les codes $(1, \leq 2)$-identifiants dans la grille du roi, comme tous les codes identifiants, peuvent être vu comme des problèmes de couverture. Il s'agit de couvrir, pour toute paire d'ensembles Y, Y' avec $|Y|, |Y'| \leq 2$ et $Y \neq Y'$, la différence symétrique $N[Y] \Delta N[Y']$. Par exemple, si $Y = \{(0,0)\}$ et $Y' = \{(0,0),(0,1)\}$, l'ensemble à couvrir consiste en trois sommets alignés (voir la figure 3.20a). Lorsque $Y = \{(0,0),(1,1)\}$ et $Y' = \{(0,1),(1,0)\}$, il s'agit de quatre sommets formant les coins d'un carré 4×4 (voir la figure 3.20b). Beaucoup d'autres exemples pourraient être donnés mais Honkala et Laihonen [55] ont montré qu'ils suffisaient de couvrir ces deux types d'ensembles (modulo symétrie) pour avoir un code $(1, \leq 2)$-identifiant, réduisant considérablement la complexité du problème.

Théorème 3.44 (Honkala, Laihonen [55]) *Un code C est un code $(1, \leq 2)$-identifiant dans la grille du roi si et seulement si, pour tout $(x,y) \in \mathbb{Z}^2$, les ensembles*

- $\{(x,y),(x+3,y),(x,y+3),(x+3,y+3)\}$,
- $\{(x,y),(x+1,y),(x+2,y)\}$,
- $\{(x,y),(x,y+1),(x,y+2)\}$

(a) $Y = \{(0,0)\}$ *et* $Y' = \{(0,0),(0,1)\}$ **(b)** $Y = \{(0,0),(1,1)\}$ *et* $Y' = \{(0,1),(1,0)\}$

Figure 3.20 – *Les ensembles* $N[Y]\Delta N[Y']$ *pour différents ensembles* Y *et* Y'*. Ce sont les seuls ensembles à couvrir (modulo symétrie) pour obtenir un code* $(1,\leq 2)$*-identifiant de la grille du roi.*

contiennent chacun au moins un élément de C*.*

Ce théorème, avec la proposition 3.43, implique que la densité d'un code $(1,\leq\ 2)$-identifiant dans la grille du roi est de densité au moins $\frac{1}{3}$ en prenant comme ensemble X trois sommets consécutifs de la grille. Pour améliorer cette borne, nous allons utiliser un ensemble X plus grand.

Fenêtres

Nous allons utiliser pour l'ensemble X l'ensemble représenté sur la figure 3.21. Dans la suite nous appelons *fenêtre* toute translation de cet ensemble. Il y a une bijection naturelle entre toutes les fenêtres et l'ensemble $\mathbb{Z}^2$. Cela nous permet de considérer le *graphe des fenêtres* qui a comme sommets l'ensemble de toutes les fenêtres, vu comme l'ensemble $\mathbb{Z}^2$, et où les voisinages considérés sont les mêmes que ceux de la grille du roi. Par exemple, la boule de rayon 2 d'une fenêtre F est l'ensemble des fenêtres $F + (x,y)$ avec $(x,y) \in \mathbb{Z}^2, \max\{|x|,|y|\} \leq 2$.

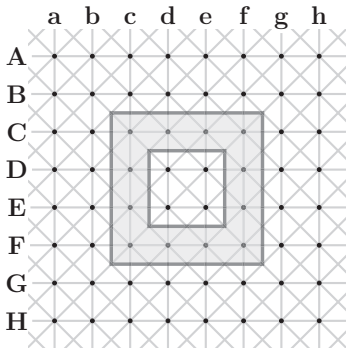


Figure 3.21 – *Notation pour les sommets dans le voisinage d'une fenêtre.*

Dans notre étude, nous allons considérer les sommets à distance 2 d'une fenêtre donnée. Nous utiliserons la notation **Xy** avec **X** une lettre majuscule comprise entre **A** et **H** et **y** une lettre minuscule comprise entre **a** et **h**, pour un sommet sur la

ligne **X** et la colonne **y**, comme représenté sur la figure 3.21. Les quatre sommets de la fenêtre en positions **Cc**, **Cf**, **Fc** et **Ff** sont les *coins* de la fenêtre. Un *côté* de la fenêtre est un ensemble de quatre sommets situés sur une même ligne ou colonne.

Les conditions du théorème 3.44 se reformulent de la manière suivante. Un code C est un $(1, \leq 2)$-code identifiant dans la grille du roi si et seulement si, pour chaque fenêtre F :

Condition 1 : Au moins un coin de F est un élément de C.

Condition 2 : Chaque ensemble de trois sommets consécutifs sur un côté de F contient au moins un élément de C.

Dans tout ce qui suit nous supposerons que C est un code $(1, \leq 2)$-identifiant de la grille du roi. Pour montrer le théorème principal, il suffit de montrer qu'il y a en moyenne $5 + \frac{3}{111}$ éléments du code par fenêtre, la proposition 3.43 donnant alors directement le théorème. Nous commençons par montrer qu'il y a en moyenne au moins cinq éléments du code par fenêtre, menant de manière relativement simple à la borne inférieure de $\frac{5}{12}$ donnée par Pelto [81]. Pour alléger l'écriture des résultats et preuves qui suivent, nous utiliserons le terme *k-fenêtre* (resp. *k^+-fenêtre*) pour désigner une fenêtre possédant k éléments du code (resp. au moins k éléments du code).

Premiers résultats sur les fenêtres

Nous montrons tout d'abord qu'il y a toujours au moins quatre éléments du code dans une fenêtre et précisons le cas où cette borne est atteinte.

Lemme 3.45 *Si F est une fenêtre, alors F est une 4^+-fenêtre. Si F est une 4-fenêtre, alors les quatre éléments du code sont situés aux quatre coins de F.*

Preuve : Chaque fenêtre peut être partitionnée en quatre ensembles de trois sommets consécutifs sur la même ligne ou colonne. D'après la condition 2, F contient donc au moins quatre éléments du code.

Supposons maintenant que F contienne exactement quatre éléments du code. D'après la condition 1, un des éléments du code doit être dans un coin. Sans perte de généralité, nous pouvons donc supposer qu'il y a un élément du code en position **Cc**. D'après la condition 2, il doit y avoir un autre élément du code sur la ligne **C**. Si celui-ci était situé sur la colonne **d** or **e**, nous pourrions partitionner le reste des sommets en trois ensembles de trois sommets consécutifs situés sur des côtés de F qui contiendrait donc cinq éléments du code ce qui mène à une contradiction.

Il y a donc un élément du code dans le coin **Cf**. De la même manière, les deux autres éléments du code sont en positions **Fc** et **Ff**. □

Les deux observations suivantes seront utilisées à de nombreuses reprises dans la suite.

Lemme 3.46 *Si une fenêtre a trois éléments du code qui ne sont pas sur des coins et qui sont situés sur seulement deux côtés, alors c'est une 6^+-fenêtre.*

Preuve : Soit F une fenêtre ayant trois éléments du code qui ne sont pas sur des coins et qui sont situés sur seulement deux côtés. Supposons tout d'abord que les deux côtés concernés sont des côtés opposés de la fenêtre. Sans perte de généralité, les éléments du code sont en position **Cd**, **Ce** et **Fd**.

D'après la condition 1, il y a au moins un élément du code dans un coin. Supposons que ce soit dans la colonne **c**. Alors, d'après la condition 2, il doit y avoir un autre élément du code dans la colonne **c** et un dans la colonne **f**, ce qui fait au moins six éléments. S'il y a un élément du code dans un coin de la colonne **f**, alors il y aura deux éléments du code dans la colonne **f** et un dans la colonne **c**, ce qui fait encore six éléments ou plus du code.

Supposons maintenant que les trois éléments du code qui ne sont pas dans les coins sont situés sur des côtés adjacents. Sans perte de généralité, nous pouvons supposer qu'ils ne sont pas dans l'ensemble S composé de la ligne **C** et de la colonne **f**. D'après la condition 2, il y a au moins deux éléments du code dans S. D'après la condition 1, il y a un élément du code dans un des coins. S'il y a un élément du code en position **Fc**, alors F contient au moins six éléments du code. Sinon, il y a un élément du code qui est dans un coin faisant partie de l'ensemble S et il y a donc au moins trois éléments du code dans S, ce qui fait en tout six éléments du code dans F. □

Lemme 3.47 *Si une fenêtre F a un élément c du code dans un de ses coins et deux éléments du code qui ne sont pas dans des coins et situés sur un même côté qui ne contient pas c, alors F est une 6^+-fenêtre.*

Preuve : Sans perte de généralité, il y a trois éléments du code en position **Cc**, **Fd** et **Fe**. D'après la condition 2, il y a un autre élément du code dans la colonne **c**, un dans la ligne **C** et un parmi les positions **Df**, **Ef** et **Ff**, ce qui fait donc au moins six éléments du code dans F. □

Nous montrons maintenant qu'aux alentours d'une 4-fenêtre, il y a plusieurs 6^+-fenêtres.

Lemme 3.48 *Soit F une 4-fenêtre. Alors les fenêtres $F + (0, 2)$, $F + (0, -2)$, $F + (2, 0)$, $F + (-2, 0)$ sont des 6^+-fenêtres.*

Preuve : D'après le lemme 3.45, les quatre éléments du code qui sont dans la fenêtre F sont les quatre coins. D'après la condition 1 appliquée sur $F + (1, 0)$, il y a un élément du code en position **Cg** ou **Fg**. D'après le lemme 3.46, $F + (2, 0)$ a donc au moins six éléments du code. Le résultat est obtenu pour les autres fenêtres par symétrie. □

Preuve simple de la borne $\frac{5}{12}$

Nous illustrons la méthode utilisée pour montrer la borne inférieure connue.

Théorème 3.49 (Pelto [81]) *La densité d'un code $(1, \leq 2)$-identifiant dans la grille du roi est au moins $\frac{5}{12}$.*

Preuve : Soit C un code $(1, \leq 2)$-identifiant de la grille du roi. Nous utilisons une méthode de déchargement pour montrer que le nombre moyen d'éléments du code est au moins cinq. Comme chaque fenêtre contient douze éléments, le résultat suit alors avec la proposition 3.43.

Au départ chaque fenêtre reçoit comme charge le nombre d'éléments du code qu'elle contient. Puis chaque fenêtre contenant au moins six éléments du code donne une charge $\frac{1}{4}$ aux fenêtres contenant quatre éléments du code parmi les fenêtres $F + (0, 2)$, $F + (0, -2)$, $F + (2, 0)$ et $F + (-2, 0)$.

D'après le lemme 3.45, chaque fenêtre a au moins une charge 4 au début. Chaque fenêtre ayant exactement une charge 4 reçoit une charge 1 et chaque fenêtre ayant une charge supérieure ou égale à 6 donne au plus une charge 1. À la fin, chaque fenêtre a donc une charge d'au moins 5. Un fenêtre donne une quantité de charge bornée (au plus 1) à distance au plus 2, les charges sont conservées donc le nombre moyen d'éléments du code dans un fenêtre est au moins 5. □

3.4.2 Détails de la preuve

Pour montrer notre borne inférieure, nous utilisons une procédure de déchargement pour obtenir à la fin une charge strictement plus grande que 5 sur chaque fenêtre. Une charge de 5 est vue comme la charge de référence, et l'*excès de charge* d'une k-fenêtre est $k - 5$. L'excès de charge d'un ensemble de fenêtres est la somme de tous les excès de charge de chaque fenêtre de l'ensemble.

Analyse fine du voisinage des fenêtres

Nous étudions tout d'abord le voisinage des 4-fenêtres et des 6^+-fenêtres à proximité. Une 6^+-fenêtre F est appelée 4-*bienfaitrice* s'il y a une 4-fenêtre parmi les fenêtres $F + \{(0, 2), (0, -2), (2, 0), (-2, 0)\}$.

Lemme 3.50 *Soit F une 4-bienfaitrice. Alors il y a une 6^+-fenêtre dans chaque coin de la boule de rayon 2 de F à distance 2 de chaque 4-fenêtre parmi $F + \{(0, 2), (0, -2), (2, 0), (-2, 0)\}$. De plus, si F est une 6-fenêtre, il y a une unique 4-fenêtre parmi les fenêtres $F + \{(0, 2), (0, -2), (2, 0), (-2, 0)\}$. Si F est une 7^+-fenêtre, il y a au plus deux 4-fenêtres parmi les fenêtres $F + \{(0, 2), (0, -2), (2, 0), (-2, 0)\}$.*

Preuve : La première partie du lemme est une conséquence directe du lemme 3.48. Il y a au moins une 4-fenêtre parmi $F + \{(0, 2), (0, -2), (2, 0), (-2, 0)\}$, sans perte de généralités, supposons que $F + (-2, 0)$ est une 4-fenêtre. Alors $F + (2, 0)$ ne peut pas être une 4-fenêtre, sinon aucun coin de F n'est un élément du code, ce qui contredit la condition 1 sur F.

Supposons maintenant qu'il y a une autre 4-fenêtre parmi $F + \{(0, 2), (0, -2)\}$, par exemple $F + (0, 2)$. Comme précédemment, $F + (0, -2)$ n'est pas une 4-fenêtre. D'après la condition 1 appliquée sur F, il y a un élément du code en position **Ff**. D'après la condition 1 appliquée sur $F + (-1, 0)$ et $F + (0, 1)$, il y a deux élément du code parmi les positions **Ce, Fe, Ec** et **Ef**. Ainsi F est une 7^+-fenêtre, ce qui termine la preuve. □

6^+	X	X	X	X
X	Y	Z	Z	X
4	X	**F**	Z	X
X	Y	Z	Z	X
6^+	X	X	X	X

Figure 3.22 – *La boule de rayon* 2 *d'une* 6*-fenêtre qui est une* 4*-bienfaitrice.*

Les 6^+-fenêtres dans les coins d'une 4-bienfaitrice décrites dans le lemme précédent sont appelées les *co-bienfaitrices* de F. Le prochain lemme est exprimé pour une configuration précise mais est valide pour les autres cas par symétrie. Sa preuve qui est longue et technique est donnée en fin de partie (partie 3.4.3).

Lemme 3.51 *Soit F une* 4*-bienfaitrice ayant exactement six éléments du code, orientée comme dans la figure 3.22. Alors :*
- *La fenêtre F a au moins deux 6^+-fenêtres dans sa boule de rayon* 2*, en plus de ses co-bienfaitrices.*
- *S'il n'y a pas de 6^+-fenêtres en position Y ou Z (voir la figure 3.22), alors il y a au moins trois 6^+-fenêtres dans la boule de rayon* 2 *de F (en plus des co-bienfaitrices) ou bien l'un des co-bienfaitrices de F est une 7^+-fenêtre et $F + (-1, 0)$ n'est pas une 6^+-fenêtre.*

Nous étudions maintenant les 5-fenêtres et montrons qu'il y a quelques 6^+-fenêtres dans leur voisinage. Certaines de ces 5-fenêtres n'ont pas beaucoup de 6^+-fenêtres et sont appelées pour cette raison *pauvres*. Plus précisément, une fenêtre est 1*-pauvre* (resp. *2-pauvre*) si c'est une 5-fenêtre qui a au plus une 6-fenêtre dans sa boule de rayon 2 (resp. aucune 6^+-fenêtres à distance 1 et au plus deux 6-fenêtres à distance 2). Le lemme suivant donne les différentes possibilités pour une 5-fenêtre. La preuve est donnée en fin de partie (sous-partie 3.4.3) pour plus de lisibilité.

Lemme 3.52 *Soit F une* 5*-fenêtre. F a au moins une 6^+-fenêtre dans sa boule de rayon* 2*. De plus, une des propriétés suivantes est vérifiée :*
- *F a une* 6*-fenêtre à distance 1 et une autre 6^+-fenêtre dans sa boule de rayon 2,*
- *F a un excès de charge d'au moins 3 dans sa boule de rayon* 2 *avec au moins deux 6^+-fenêtres,*
- *F est* 1*-pauvre, et, à symétrie près, le voisinage de rayon* 2 *de F est décrit par la figure 3.23,*
- *F est* 2*-pauvre, et, à symétrie près, le voisinage de rayon* 2 *de F est décrit par la figure 3.24.*

Finalement, nous étudions les 6^+-fenêtres à proximité des fenêtres pauvres.

Lemme 3.53 *Soit F une fenêtre* 1*-pauvre. Alors F a une* 6*-fenêtre F_6 à distance 1 telle que :*

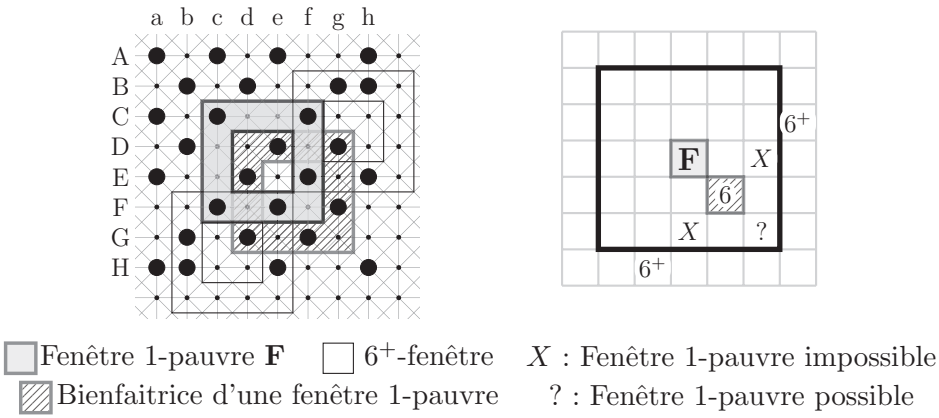


Figure 3.23 – *Voisinage d'une fenêtre* 1*-pauvre dans le graphe des sommets (à gauche) et dans le graphe des fenêtres (à droite).*

- *F_6 a au plus deux fenêtres* 1*-pauvres dans sa boule de rayon* 2*. Elles sont à distance 1 de F_6,*
- *F_6 a au moins deux 6^+-fenêtres à distance 2,*
- *F_6 n'est pas une* 4*-bienfaitrice.*

Preuve : D'après le lemme 3.52, si F est une fenêtre 1-pauvre, le voisinage de F est fixé et décrit par la figure 3.23 (à symétrie près). La seule 6^+-fenêtre dans le voisinage de F est $F_6 = F + (1, -1)$ et contient exactement six éléments du code. La première partie du lemme vient du fait qu'une fenêtre 1-pauvre a nécessairement une 6-fenêtre dans un coin de sa boule de rayon 1. En utilisant la partie gauche de la figure 3.23, les positions marquées par un "X" sur la partie droite de la figure ne peuvent pas être des fenêtres 1-pauvres.

La seconde partie vient en utilisant le lemme 3.46 sur la partie gauche de la figure 3.23.

Enfin, le dernier point est une conséquence du fait qu'une 4-bienfaitrice doit avoir deux coins sans élément du code dans la même ligne ou colonne, ce qui n'est pas le cas ici. □

Une fenêtre correspondant à la fenêtre F_6 du lemme précédent est appelée *bienfaitrice d'une fenêtre* 1*-pauvre.*

Lemme 3.54 *Soit F une fenêtre* 2*-pauvre. Alors F a deux* 6*-fenêtres à distance* 2*. L'une d'elle, notée F_6, vérifie les propriétés suivantes :*
- *F_6 n'est ni une* 4*-bienfaitrice, ni une bienfaitrice d'une fenêtre* 1*-pauvre,*
- *F_6 a au moins quatre 6^+-fenêtres dans sa boule de rayon* 2*,*
- *F_6 a seulement une fenêtre* 2*-pauvre dans sa boule de rayon* 2*.*

Preuve : D'après le lemme 3.52, si F est une fenêtre 2-pauvre, le voisinage de F est fixé et décrit par la figure 3.24, à symétrie près. Soit $F_- = F + (-2, -1)$.

Le premier point du lemme vient du fait qu'une fenêtre 1-pauvre a exactement une 6-fenêtre dans sa boule de rayon 2 (lemme 3.52) et qu'une 4-bienfaitrice à deux

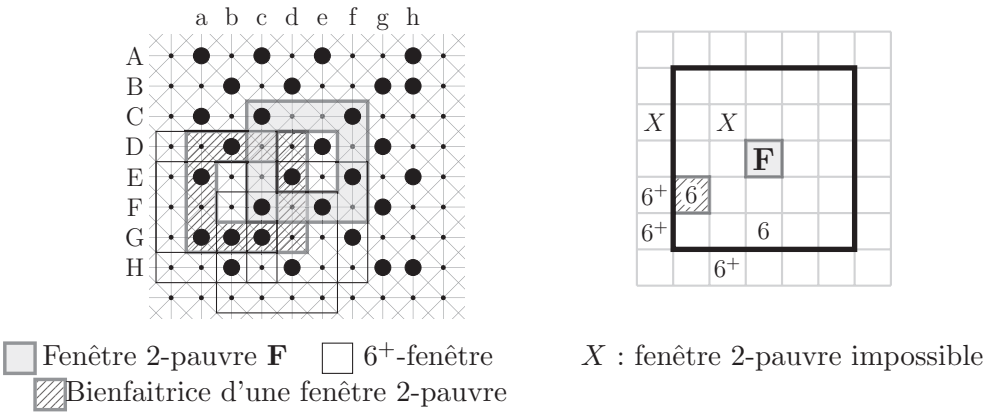


Figure 3.24 – *Voisinage d'une fenêtre* 2*-pauvre dans le graphe des sommets (à gauche) et dans le graphe des fenêtres (à droite).*

coins sans élément du code (la seule possibilité serait donc que $F + (-2, 1)$ soit une 4-fenêtre ce qui n'est pas vrai).

Le deuxième point est obtenu en utilisant le lemme 3.46 sur la partie gauche de la figure 3.24.

Pour le troisième point, les seuls endroits où il pourrait y avoir une autre fenêtre 2-pauvre dans la boule de rayon 2 de F_6 sont ceux marqués par un "X" dans la partie droite de la figure 3.24 : $F + (-3, 1)$ et $F + (-1, 1)$. Cependant, ces deux fenêtres ont deux 6^+-fenêtres dans leur boule de rayon 2 qui ne correspondent pas à la configuration des éléments du code d'une fenêtre 2-pauvre donnée par la partie gauche de la figure 3.24. □

Une fenêtre jouant le rôle de F_6 dans le lemme précédent est appelée *bienfaitrice d'une fenêtre* 2*-pauvre*.

Règles de déchargement

Nous pouvons maintenant décrire la procédure de déchargement qui mène à notre résultat. Au début une k-fenêtre a la charge k.

Soit $\alpha = \frac{1}{37}$. Nous appliquons les règles suivantes :

1. Soit F une 4-bienfaitrice qui contient exactement six éléments du code, sans 7^+-fenêtres parmi ses co-bienfaitrices. On suppose que F est orientée comme dans la figure 3.22 (les autres cas suivent par symétrie). On considère deux sous-règles :

   (a) S'il y a une 6^+-fenêtre sur une position Y de la figure 3.22, par exemple en position $F + (-1, 1)$, et pas d'autres 6^+-fenêtres sur les positions Y ou Z de la figure 3.22, alors F donne une charge $\frac{3\alpha+1}{4}$ à la 4-fenêtre $F + (-2, 0)$, une charge α à la fenêtre $F + (0, 1)$, une charge 2α aux autres 5-fenêtres sur les positions Z de la figure 3.22, une charge $\frac{3\alpha}{2}$ aux autres 5-fenêtres sur les positions Y de la figure 3.22 et une charge α à toutes les autres 5-fenêtres de la boule de rayon 2 de F.

(b) Sinon, F donne une charge $\frac{3\alpha+1}{4}$ à la 4-fenêtre $F+(-2,0)$, une charge 2α aux 5-fenêtres sur les positions Z de la figure 3.22, une charge $\frac{3\alpha}{2}$ aux 5-fenêtres sur les positions Y de la figure 3.22 et une charge α à toutes les autres 5-fenêtres de la boule de rayon 2 de F.

2. Soit F une 4-bienfaitrice qui contient exactement six éléments du code et qui a une 7^+-fenêtre parmi ses co-bienfaitrices. On suppose que F est orientée comme dans la figure 3.22, et que $F+(-2,2)$ est une 7^+-fenêtre (les autres cas se déduisent par symétrie). Alors F donne les mêmes charges que dans la règle 1 sauf pour la fenêtre $F+(-1,1)$ qui ne reçoit rien de F.
3. Une 7-fenêtre qui est une 4-bienfaitrice donne une charge $\frac{3\alpha+1}{4}$ aux 4-fenêtres parmi les fenêtres $F+\{(0,2),(0,-2),(2,0),(-2,0)\}$, une charge 3α aux fenêtres à distance 1 et une charge 2α aux fenêtres à distance 2.
4. Une bienfaitrice d'une fenêtre 1-pauvre, qui est nécessairement une 6-fenêtre, donne une charge 3α aux fenêtres 1-pauvre de sa boule de rayon 2, une charge 2α aux autres 5-fenêtres à distance 1 et une charge α aux 5-fenêtres à distance 2.
5. Une bienfaitrice d'une fenêtre 2-pauvre, qui est nécessairement une 6-fenêtre, donne une charge 2α à la seule fenêtre 2-pauvre dans sa boule de rayon 2, une charge 2α aux autres 5-fenêtres à distance 1 et une charge α aux autres 5-fenêtres à distance 2.
6. Les autres 6-fenêtres donnent une charge 2α aux 5-fenêtres à distance 1 et une charge α aux 5-fenêtres à distance 2.
7. Les autres 7-fenêtres donnent une charge 3α aux 5-fenêtres à distance 1 et une charge 2α aux 5-fenêtres à distance 2.

Validité des règles

Les règles énoncées précédemment ne sont pas ambigües. En effet, d'après les lemmes 3.53 et 3.54, une fenêtre F est soit une 4-bienfaitrice, soit une bienfaitrice d'une fenêtre 1-pauvre ou 2-pauvre, soit ce n'est pas une bienfaitrice, mais aucun de ces choix-là en même temps. De plus, il y a conservation de la charge.

Les lemmes suivant montrent que chaque fenêtre, quelque soit le nombre d'éléments du code qu'elle contient, a une charge d'au moins $5+3\alpha$ à la fin de la procédure.

Lemme 3.55 *Après l'application des règles de déchargement, chaque 7^+-fenêtre a une charge d'au moins $5+3\alpha$.*

Preuve : Soit F une 7^+-fenêtre. Il suffit de vérifier que F donne au plus une charge de $2-3\alpha=\frac{71}{37}$.

Si F est une 4-bienfaitrice, alors d'après le lemme 3.50, il y a au plus deux 4-fenêtres parmi $F+\{(0,2),(0,-2),(2,0),(-2,0)\}$.

Si F a deux 4-fenêtres parmi $F+\{(0,2),(0,-2),(2,0),(-2,0)\}$ alors F a au moins trois co-bienfaitrices. Il y a donc au plus onze 5-fenêtres à distance 2 de F. D'après la règle 3, F donne donc au plus une charge de $2\cdot\frac{3\alpha+1}{4}+8\cdot 3\alpha+11\cdot 2\alpha=\frac{66}{37}$.

Si F a seulement une 4-fenêtre parmi $F + \{(0,2),(0,-2),(2,0),(-2,0)\}$, alors il a deux co-bienfaitrices et donc au plus treize 5-fenêtre à distance 2. D'après la règle 3, F donne au plus une charge de $\frac{3\alpha+1}{4} + 8 \cdot 3\alpha + 13 \cdot 2\alpha = \frac{60}{37}$.

Finalement, si F n'est pas une 4-bienfaitrice, d'après la règle 7, F donne au plus une charge de $8 \cdot 3\alpha + 16 \cdot 2\alpha = \frac{56}{37}$. □

Lemme 3.56 *Après l'application des règles de déchargement, chaque* 6*-fenêtre a une charge d'au moins* $5 + 3\alpha$.

Preuve : Soit F une 6-fenêtre. Il suffit de montrer que F donne au plus une charge de $1 - 3\alpha = \frac{34}{37}$.

Supposons tout d'abord que F est une 4-bienfaitrice sans 7^+-fenêtre parmi ses co-bienfaitrices. S'il n'y avait pas d'autres 6^+-fenêtres que les co-bienfaitrices de F dans sa boule de rayon 2, F donnerait, selon la règle 1, une charge de $\frac{3\alpha+1}{4} + 5 \cdot 2\alpha + 2 \cdot \frac{3\alpha}{2} + 14\alpha = \frac{37}{37}$. Cependant, d'après le lemme 3.51, il y a au moins deux 6^+-fenêtres qui ne sont pas des co-bienfaitrices dans la boule de rayon 2 de F. De plus, toujours d'après le lemme 3.51, l'une d'entre elles est sur une position Y ou Z ou bien il y a une troisième 6^+-fenêtre dans la boule de rayon de F, hors co-bienfaitrices. Nous distinguons deux cas. S'il y a une 6^+-fenêtre sur une position Y et une sur une position X, la règle 1a s'applique et nous sauvons une charge d'au moins $\frac{3\alpha}{2} + \alpha$ des 6^+-fenêtres, et une charge α de la fenêtre $F + (0,1)$. Sinon, la règle 1b s'applique, et la charge totale sauvée grâce aux 6^+-fenêtres est d'au moins 3α. En effet, nous sauvons au moins $2 \cdot \frac{3\alpha}{2}$ si les deux 6^+-fenêtres supplémentaires sont toutes les deux sur des positions Y, $2 \cdot 2\alpha$ si elles sont toutes les deux sur des positions Z, $2\alpha + \frac{3\alpha}{2}$ s'il y en a une sur une position Z et l'autre sur une position Y, $2\alpha + \alpha$ s'il y en a une sur une position Z et l'autre sur une position X et enfin $3 \cdot \alpha$ s'il y a trois 6^+-fenêtres en plus. Dans tous les cas, F donne une charge d'au plus $\frac{34}{37}$.

Supposons maintenant que F est une 4-bienfaitrice avec une 7^+-fenêtre parmi ses co-bienfaitrices. D'après le lemme 3.51, il y a deux 6^+ fenêtres F_1 et F_2 dans la boule de rayon 2 de F en plus de ses co-bienfaitrices. D'après la règle 2, nous appliquons le même déchargement que pour la règle, sauf que la fenêtre $F + (-1,1)$ ne reçoit aucune charge. D'après le lemme 3.51, soit F_1 ou F_2 est sur une position Y ou Z, soit il y a une troisième 6^+-fenêtre qui n'est pas un co-bienfaitrice de F, soit $F + (-1,0)$ n'est pas une 6^+-fenêtre. Si l'un des deux premiers cas est vérifié, les mêmes arguments que ceux utilisés au paragraphe précédent s'appliquent. Sinon, F_1 et F_2 ne sont pas sur une position Y ou Z et $F + (-1,0)$ a au plus cinq éléments du code. Les fenêtres F_1 et F_2 sont donc à distance 2 de F. Il y a donc au plus douze 5-fenêtres sur des positions X dans la boule de rayon 2 de F et F donne au plus $\frac{3\alpha+1}{4} + 5 \cdot 2\alpha + \frac{3\alpha}{2} + 12\alpha = \frac{33.5}{37}$.

Si F est une bienfaitrice d'une fenêtre 1-pauvre, d'après le lemme 3.53, F a au plus deux fenêtres 1-pauvres dans sa boule de rayon 2, et chacune d'elles est à distance 1 de F. De plus, il y a au moins deux 6^+-fenêtres à distance 2 de F. Nous pouvons supposer que F a deux fenêtres 1-pauvres dans son voisinage puisque selon la règle 4, cela correspond au cas où F donne le plus de charges, c'est-à-dire au plus $2 \cdot 3\alpha + 6 \cdot 2\alpha + 14 \cdot \alpha = \frac{32}{37}$.

Si F est une bienfaitrice d'une fenêtre 2-pauvre, d'après le lemme 3.54, F a seulement une fenêtre 2-pauvre dans son voisinage et au moins quatre 6^+-fenêtres

dans sa boule de rayon 2. Nous pouvons supposer que ces fenêtres sont à distance 2 de F puisque selon la règle 5, c'est le pire cas, et alors F donne au plus une charge de $2\alpha + 8 \cdot 2\alpha + 11 \cdot \alpha = \frac{29}{37}$.

Finalement, si F n'est pas une bienfaitrice, d'après la règle 6, F donne une charge d'au plus $8 \cdot 2\alpha + 16 \cdot \alpha = \frac{32}{37}$. □

Lemme 3.57 *Après l'application des règles de déchargement, chaque* 5*-fenêtre a une charge d'au moins* $5 + 3\alpha$.

Preuve : Il suffit de montrer que chaque 5-fenêtre reçoit une charge d'au moins 3α.

Remarquons tout d'abord que d'après les règles de déchargement, chaque 6-fenêtre F_6 donne une charge d'au moins α à chaque 5-fenêtre dans sa boule de rayon 2, excepté avec la règle 2 où F_6 ne donne rien à une 5-fenêtre F' à distance 1 de F_6. Dans ce cas, F' a une 4-bienfaitrice qui a sept élément du code à distance 1 qui, d'après la règle 3, donne une charge 3α à F', ce qui termine ce cas là.

Soit F une 5-fenêtre. D'après le paragraphe précédent, on peut considérer que F reçoit une charge d'au moins α de chaque 6-fenêtre dans sa boule de rayon 2. D'après les règles 3 et 7, F reçoit une charge d'au moins 2α de chaque 7^+-fenêtre de sa boule de rayon 2. Aussi si F a un excès de charge d'au moins 3 dans sa boule de rayon 2 avec au moins deux 6^+-fenêtres, F reçoit au moins une charge 3α.

Sinon, d'après le lemme 3.52, F est soit 1-pauvre, soit 2-pauvre, ou bien F a une 6-fenêtre à distance 1 et une autre 6-fenêtre dans sa boule de rayon 2.

Si F est 1-pauvre, d'après le lemme 3.53, il y a une bienfaitrice d'une fenêtre 1-pauvre dans la boule de rayon 2 de F qui donne, d'après la règle 4, une charge 3α à F.

Si F est 2-pauvre, d'après le lemme 3.54, il y a une bienfaitrice d'une fenêtre 2-pauvre dans la boule de rayon 2 de F qui donne, d'après la règle 5, une charge 2α à F. L'autre 6^+-fenêtre présente dans la boule de rayon 2 de F donne une charge d'au moins α à F.

Finalement, supposons que F a une 6-fenêtre F_6 à distance 1 et une autre F_6' dans sa boule de rayon 2. Si F reçoit une charge 2α de F_6, c'est fini. Sinon, d'après nos règles de déchargement, F_6 est nécessairement une 4-bienfaitrice. Sans perte de généralité, supposons que F_6 est orientée comme dans la figure 3.22. Alors F_6 donne toujours une charge 2α aux 5-fenêtres sur les positions Z, sauf dans la règle $1a$ où $F_6 + (0,1)$ reçoit seulement une charge α, mais dans ce cas, $F_6 + (0,1)$ a trois 6^+-fenêtres dans sa boule de rayon 2 et donc $F \neq F_6 + (0,1)$. La fenêtre $F_6 + (-1,0)$ a aussi trois 6^+-fenêtres dans sa boule de rayon 2, donc $F \neq F_6 + (-1,0)$. Nous pouvons donc supposer que F est sur une position Y. Alors F_6' est aussi une 4-bienfaitrice. Dans ce cas, d'après les règles 1 et 2, F_6 et F_6' donnent toutes deux une charge $\frac{3\alpha}{2}$ à F, ce qui termine la preuve. □

Lemme 3.58 *Après l'application des règles de déchargement, chaque* 4*-fenêtre a une charge d'au moins* $5 + 3\alpha$.

Preuve : D'après le lemme 3.48, une 4-fenêtre F a quatre 4-bienfaitrices. D'après les règles 1 à 3, chaque 4-bienfaitrice donne une charge $\frac{3\alpha+1}{4}$ à F. Donc, F a reçu une charge totale de $1 + 3\alpha$ et termine avec une charge de $5 + 3\alpha$. □

Fin de la preuve

Après l'application des règles de déchargement, d'après les lemmes 3.55, 3.56, 3.57 et 3.58, chaque fenêtre a une charge d'au moins $5 + 3\alpha = \frac{188}{37}$. Par conséquent, le nombre moyen d'éléments du code dans chaque fenêtre est d'au moins $\frac{188}{37}$ (cela est dû au fait que chaque fenêtre donne des charges à des fenêtres à distance au plus 2 et que la quantité de charge donnée est bornée). Une fenêtre compte douze sommets, en appliquant la proposition 3.43, nous obtenons le théorème 3.42.

3.4.3 Preuves des lemmes techniques

Preuve du lemme 3.51

Lemme 1 *Soit F une 4-bienfaitrice ayant exactement six éléments du code, orientée comme dans la figure 3.22. Alors :*

- *La fenêtre F a au moins deux 6^+-fenêtres dans sa boule de rayon 2, en plus de ses co-bienfaitrices.*
- *S'il n'y a pas de 6^+-fenêtres en position Y ou Z (voir la figure 3.22), alors il y a au moins trois 6^+-fenêtres dans la boule de rayon 2 de F (en plus des co-bienfaitrices) ou bien l'un des co-bienfaitrices de F est une 7^+-fenêtre et $F + (-1, 0)$ n'est pas une 6^+-fenêtre.*

Preuve : Dans cette preuve, nous utilisons souvent les conditions 1 et 2 sans y faire référence explicitement. Soit F une 4-bienfaitrice avec six éléments du code. Nous construisons pas à pas le voisinage de F en éliminant au fur et à mesure les cas amenant dans l'une des deux possibilités du lemme.

D'après la condition 1 appliquée sur $F + (-1, 0)$, il y a un élément du code en position **Ce** ou **Fe**. Sans perte de généralité, supposons qu'il y ait un élément du code en position **Ce** (l'autre cas est symétrique).

Comme F n'est pas une 7^+-fenêtre, il n'y a pas d'élément du code en position **Fe** et seulement un parmi les positions **Dc** et **Ec**.

Supposons qu'il y ait un élément du code en position **Ec** et aucun en position **Dc**. D'après la condition 2, il y a un élément du code en position **De**. D'où, en utilisant le lemme 3.46, $F + (-1, 1)$ est une 6^+-fenêtre. Si $F + (-1, 0)$ est aussi une 6^+-fenêtre, c'est bon. Sinon $F + (2, 1)$ est une 6^+-fenêtre.

Supposons maintenant qu'il n'y a pas d'élément du code en position **Ec**. D'après la condition 2, il y a un élément du code en position **Dc**.

S'il y a un élément du code en position **De**, $F + (-1, 0)$ est une 6^+-fenêtre, et, d'après le lemme 3.47, $F + (-1, 1)$ est une 6^+-fenêtre. Aussi nous pouvons supposer qu'il n'y a pas d'élément du code en position **De**. Cela implique qu'il y a un élément du code en position **Df**, et pas d'élément du code en position **Ef**.

S'il y a un élément du code en position **Cf**, d'après le lemme 3.46, $F + (0, 1)$ est une 6^+-fenêtre. De plus, d'après le lemme 3.47, $F + (1, 0)$ est une 6^+-fenêtre. Donc, nous pouvons supposer qu'il n'y a pas d'élément du code en position **Cf**, ce qui implique qu'il y a un élément du code en position **Ff**.

S'il y a un élément du code en position **Gd**, d'après le lemme 3.46, $F + (-1, -1)$ et $F + (1, -2)$ sont des 6^+-fenêtres. Aussi nous pouvons supposer qu'il n'y a pas d'élément du code en position **Gd**.

S'il y a un élément du code en position **Fg**, $F + (1, 0)$ est une 6^+-fenêtre et d'après le lemme 3.46, $F + (2, 0)$ est une 6^+-fenêtre. Donc nous pouvons supposer qu'il n'y a pas d'élément du code en position **Fg**.

S'il y a un élément du code en position **Db**, $F + (-1, 0)$ est une 6-fenêtre et d'après le lemme 3.46, $F + (-2, -1)$ est une 6^+-fenêtre. Si $F + (-1, -1)$ est une 6^+-fenêtre, c'est fini. Sinon, d'après le lemme 3.46, $F + (0, -2)$ est une 6^+-fenêtre. Aussi nous pouvons supposer qu'il n'y a pas d'élément du code en position **Db**.

S'il n'y a pas d'élément du code en position **Bf**, il y a un élément du code en position **Bc**, et d'après le lemme 3.46, $F + (-2, 1)$ est une 6^+-fenêtre. Alors, il doit y avoir un élément du code parmi les positions **Bd** et **Be**, ce qui implique que $F + (-1, 1)$ est une 6^+-fenêtre. Aussi nous pouvons supposer qu'il y a un élément du code en position **Bf**.

S'il y a un élément du code en position **Be**, d'après le lemme 3.46, $F + (1, 1)$ et $F + (2, 2)$ sont des 6^+-fenêtres. Aussi nous supposons qu'il n'y a pas d'élément du code en position **Be**.

S'il n'y a pas d'élément du code en position **Eg**, d'après le lemme 3.46, $F+(2, -1)$ et $F + (2, 2)$ sont des 6^+-fenêtres. Maintenant, s'il y a un élément du code in **Bd**, d'après le lemme 3.46, $F+(1, 2)$ est une 6^+-fenêtre. Sinon, il y a un élément du code en positions **Bc** et **Ba**, et $F + (-2, 1)$ est une 6-fenêtre. Supposons donc qu'il y a un élément du code en position **Eg**.

S'il y a un élément du code en position **Dg**, alors $F + (1, 0)$ est une 6^+-fenêtre et d'après le lemme 3.46, $F + (2, -1)$ est une 6^+-fenêtre. Donc nous pouvons supposer qu'il n'y a pas d'élément du code en position **Dg**, ce qui implique qu'il y a un élément du code en position **Gg** d'après la condition 1.

S'il y a un élément du code en position **Gf**, d'après le lemme 3.46, $F + (0, -2)$ et $F + (2, -1)$ sont des 6^+-fenêtres. Maintenant, s'il y a un élément du code en position **Ge**, $F + (1, -1)$ est une 6^+-fenêtre. Sinon, il y a un élément du code en position **Gb** et $F + (-2, -1)$ est une 6^+-fenêtre. Nous supposons donc qu'il n'y a pas d'élément du code en position **Gf**, ce qui implique qu'il y a un élément du code en position **Ge**.

S'il y a un élément du code en position **Gb**, $F + (-1, -1)$ et $F + (-2, -1)$ sont des 6-fenêtres. Nous supposons donc qu'il n'y a pas d'élément du code en position **Gb**.

S'il y a un élément du code en position **Bd**, d'après le lemme 3.46, $F + (1, 2)$ est une 6^+-fenêtre. De plus $F + (1, 1)$ est une 6^+-fenêtre. Supposons donc qu'il n'y a pas d'élément du code en position **Bd**. Alors il y a un élément du code aux deux positions **Ba** et **Bc**, ce qui implique que $F + (-2, 1)$ est une 6-fenêtre.

S'il y a un élément du code en position **Cg**, $F+(1, 0)$ est une 6-fenêtre. Supposons donc qu'il n'y a pas d'élément du code en position **Cg**. Alors, d'après le lemme 3.46, $F + (2, 1)$ est une 6-fenêtre.

Maintenant, si $F + (1, -2)$ est une 6^+-fenêtre, c'est fini. Sinon, il n'y a pas d'élément du code en position **Hd** et il y a exactement un élément du code parmi les positions **He** et **Hf**. Comme $F + (0, -2)$ est une 5^+-fenêtre, il y a un élément du code en position **Hc**. D'après la condition 1, il y a un élément du code en position **Ha**. Cela implique que $F + (-2, -2)$ est une 7-fenêtre et finit l'analyse des cas. □

Preuve du lemme 3.52

Lemme 2 *Soit F une 5-fenêtre. F a au moins une 6^+-fenêtre dans sa boule de rayon 2. De plus, une des propriétés suivantes est vérifiée :*

- *F a une 6-fenêtre à distance 1 et une autre 6^+-fenêtre dans sa boule de rayon 2,*
- *F a un excès de charge d'au moins 3 dans sa boule de rayon 2 avec au moins deux 6^+-fenêtres,*
- *F est 1-pauvre, et, à symétrie près, le voisinage de rayon 2 de F est décrit par la figure 3.23,*
- *F est 2-pauvre, et, à symétrie près, le voisinage de rayon 2 de F est décrit par la figure 3.24.*

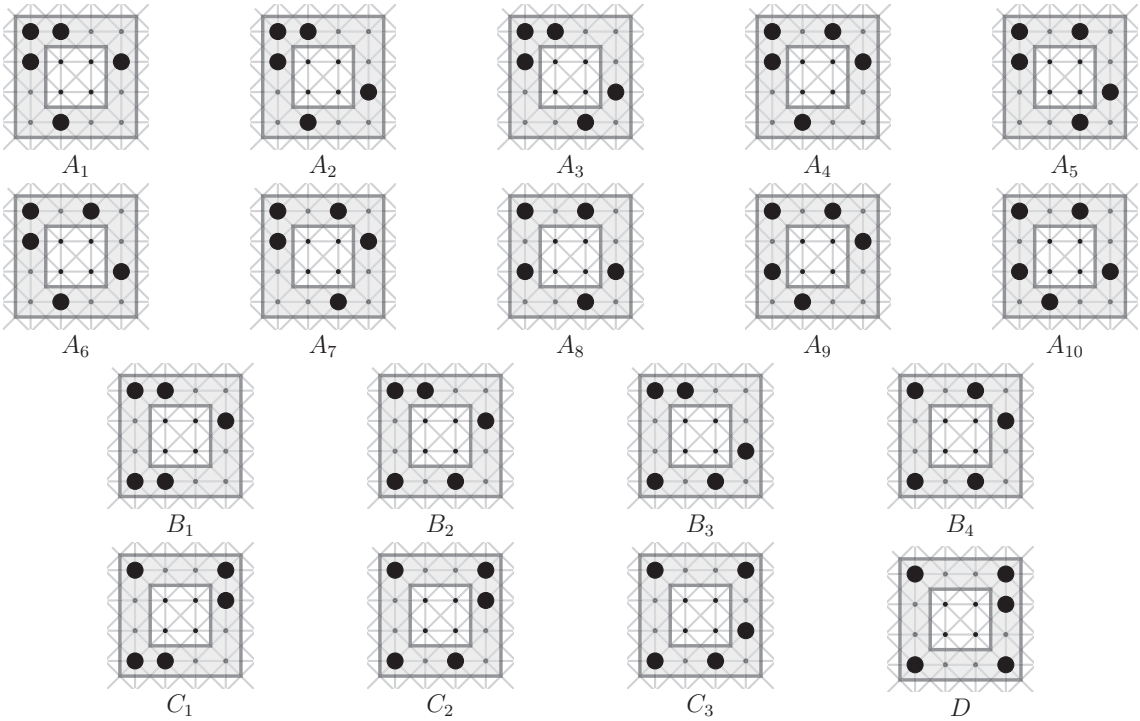


Figure 3.25 – *Toutes les 5-fenêtres possibles.*

Preuve : Soit F une 5-fenêtre. Nous procédons par une analyse de cas selon les configurations possibles pour F décrites par la figure 3.25. Les cas $A1$ à A_{10} sont ceux où F a seulement un élément du code dans un coin. Les cas B_1 à B_4 correspondent à ceux où il y a exactement deux éléments du code dans des coins se trouvant sur un même côté. Les cas C_1 à C_3 sont ceux où F a trois éléments du code dans des coins et enfin le cas D est celui où les quatre coins de F sont des éléments du code. Les cas symétriques à ceux de la figure 3.25 se démontrent avec les mêmes arguments. Nous pouvons remarquer qu'il n'y a aucun cas où F a exactement deux éléments du code situés dans des coins, qui sont opposés. En effet, dans ce cas, avec la condition 2, chaque côté de F contiendrait aussi un élément du code, menant à une 6^+ fenêtre. Pour chaque cas, nous construisons le voisinage de F pas à pas en s'arrêtant lorsque F est dans un des cas du lemme.

Cas A_1**.** D'après le lemme 3.46, $F + (-1, 0)$ et $F + (0, 1)$ sont toutes deux des 6^+-fenêtres.

Cas A_2**.** D'après le lemme 3.46, $F + (-1, 0)$ est une 6^+-fenêtre. D'après le lemme 3.47, $F + (0, 1)$ est aussi une 6^+-fenêtre.

Cas A_3**.** D'après le lemme 3.47, $F + (-1, 0)$ et $F + (0, 1)$ sont toutes deux des 6^+-fenêtres.

Cas A_4**.** D'après le lemme 3.46, $F+(0, 1)$ est une 6^+-fenêtre. Maintenant, d'après la condition 2, il y a un élément du code en position **Dd** ou **Ed**. Dans les deux cas, d'après le lemme 3.46, $F + (-1, 2)$ ou $F + (1, -1)$ est une 6^+-fenêtre.

Cas A_5**.** D'après le lemme 3.47, $F+(0, 1)$ est une 6^+-fenêtre. Maintenant, d'après la condition 2, il y a un élément du code en position **Dd** ou **De**. Dans les deux cas, d'après le lemme 3.46, $F + (-1, -1)$ ou $F + (2, 1)$ est une 6^+-fenêtre.

Cas A_6**.** D'après le lemme 3.47, $F+(0, 1)$ est une 6^+-fenêtre. Maintenant, d'après la condition 2, il y a un élément du code en position **Dd** ou **Ed**. Dans les deux cas, d'après le lemme 3.46, $F + (-1, -1)$ ou $F + (1, -1)$ est une 6^+-fenêtre.

Cas A_7**.** D'après le lemme 3.46, $F+(0, 1)$ est une 6^+-fenêtre. Si $F+(-1, -1)$ est une 6^+-fenêtre, c'est fini. Sinon, par hypothèse, il y a déjà quatre éléments du code qui ne sont pas des coins dans $F + (-1, -1)$, donc $F + (-1, -1)$ est une 5-fenêtre et, d'après la condition 1, toutes les autres positions qui ne sont pas des coins ne contiennent pas d'élément du code. En particulier, il n'y a pas d'élément du code aux positions **Dd** et **Gd**. D'après la condition 1, il y a un élément du code sur au moins une des positions **Dg** et **Gg**. Donc, d'après le lemme 3.46, $F+(2, -1)$ est une 6^+-fenêtre.

Cas A_8**.** Supposons qu'il n'y a pas d'élément du code en position **Dd**. Alors, $F + (-1, 0)$ et $F + (2, 1)$ sont des 6^+-fenêtres. Supposons donc qu'il y a un élément du code en position **Dd**.

S'il y a un élément du code en position **Ed**, d'après le lemme 3.46, $F + (-2, 0)$ et $F+(1, 0)$ sont des 6^+-fenêtres. Aussi nous pouvons supposer qu'il n'y a pas d'élément du code en position **Ed**, et par symétrie, pas d'élément du code en position **De** non plus.

S'il y a un élément du code en position **Ee**, d'après le lemme 3.46, $F + (1, 1)$ et $F + (-1, -1)$ sont des 6^+-fenêtres. Aussi nous pouvons supposer qu'il n'y a pas d'élément du code en position **Ee**.

S'il y a un élément du code en position **Be**, d'après le lemme 3.46, $F + (1, 1)$ et $F + (2, 2)$ sont des 6^+-fenêtres. Aussi nous pouvons supposer qu'il n'y a pas d'élément du code en position **Be**, et par symétrie, pas d'élément du code en position **Eb**. D'après la condition 1, il n'y a donc pas d'élément du code en position **Bb**.

S'il y a un élément du code en position **Eg**, d'après le lemme 3.46, $F + (1, 0)$ et $F + (2, 1)$ sont des 6^+-fenêtres. Aussi nous pouvons supposer qu' il n'y a pas d'élément du code en position **Eg**, et par symétrie, pas d'élément du code en position **Ge**.

S'il n'y a pas d'élément du code en position **Bd**, d'après le lemme 3.46, $F+(0,2)$ et $F+(-2,1)$ sont des 6^+-fenêtres. Si $F+(1,1)$ est une 6^+-fenêtre, c'est fini. Sinon, d'après le lemme 3.46, $F+(2,1)$ est une 6^+-fenêtre. Nous pouvons donc supposer qu'il y a un élément du code en position **Bd**, et par symétrie, un autre élément du code en position **Db**.

S'il y a un élément du code en position **Bc**, $F+(0,1)$ est une 6-fenêtre et d'après le lemme 3.46, $F+(-1,1)$ est une 6^+-fenêtre. Aussi nous pouvons supposer qu' il n'y a pas d'élément du code en position **Bc**, et par symétrie, pas d'élément du code en position **Cb**.

S'il y a un élément du code en position **Gb**, $F+(-1,-1)$ et $F+(-1,-2)$ sont des 6^+-fenêtres. Aussi nous pouvons supposer qu' il n'y a pas d'élément du code en position **Gb**, et par symétrie, pas d'élément du code en position **Bg**.

S'il n'y a pas d'élément du code en position **Gf**, il y a, d'après la condition 1, un élément du code en position **Gc** et alors $F+(-1,-1)$ est une 6-fenêtre. Maintenant, s'il n'y a pas d'élément du code en position **Fg**, $F+(1,-1)$ est une 4-fenêtre et d'après le lemme 3.48, $F+(1,1)$ est une 6^+-fenêtre. Sinon, d'après le lemme 3.46, $F+(1,-2)$ est une 6^+-fenêtre. Aussi nous pouvons supposer qu'il y a un élément du code en position **Gf**, et par symétrie, un autre élément du code en position **Fg**.

S'il y a un élément du code en position **Gg**, $F+(1,-1)$ est une 6-fenêtre et d'après le lemme 3.46, $F+(1,-2)$ est une 6^+-fenêtre. Aussi nous pouvons supposer qu'il n'y a pas d'élément du code en position **Gg**.

S'il y a un élément du code en position **Gc**, $F+(0,-1)$ et $F+(-1,-1)$ sont des 6^+-fenêtres. Nous pouvons donc supposer qu'il n'y a pas d'élément du code en position **Gc** et **Cg**.

Si $F+(-1,2)$ est une 5-fenêtre, d'après la condition 2, il n'y a pas d'élément du code en position **Ae** et un élément du code sur exactement une des positions **Ac** ou **Ad**. Alors il doit y avoir un élément du code en position **Af** car $F+(0,2)$ doit avoir un cinquième élément du code. Donc, $F+(2,2)$ est une 6^+-fenêtre. Ainsi, il y a au moins une 6^+-fenêtre parmi $F+(-1,2)$ et $F+(2,2)$. Par symétrie, il y a aussi une 6^+-fenêtre parmi $F+(-2,1)$ et $F+(-2,-2)$.

Si $F+(-1,-2)$ est une 6^+-fenêtre, c'est fini. Sinon, il n'y a pas d'élément du code en position **Hd**, ce qui implique qu'il y a un élément du code en position **Hg**, et un élément du code parmi les positions **He** et **Hf**. Par symétrie, si $F+(2,1)$ est une 5-fenêtre, il y a un élément du code en position **Gh**, et un élément du code parmi les positions **Eh** et **Fh**. Alors, $F+(2,-2)$ est une 6^+-fenêtre.

Cas A_9. D'après la condition 2, il y a au moins deux éléments du code parmi les positions **Dd**, **De**, **Ed** et **Ee**. S'il y a deux éléments du code sur la même ligne ou colonne, (resp. **d**, **D**, **e**, **E**), alors d'après le lemme 3.46, resp. $F+(1,0)$, $F+(0,-1)$, $F+(-1,0)$, $F+(0,1)$, est une 6^+-fenêtre. Ainsi, s'il y a trois éléments du code parmi les positions **Dd**, **De**, **Ed** et **Ee**, il y a deux 6^+-fenêtres à distance 1 de F et c'est fini.

Aussi, nous pouvons supposer qu'il y a exactement deux élément du code parmi les positions **Dd**, **De**, **Ed** et **Ee**. Alors d'après la condition 2, ils ne peuvent pas être situés sur la même ligne ou colonne. S'il y a des éléments du code en position **De** et **Ed**, d'après le lemme 3.46, $F+(1,-1)$ et $F+(-1,1)$ sont des 6^+-fenêtres et c'est fini. Nous supposons donc qu'il y a des éléments du code en position **Dd** et **Ee**.

S'il y a un élément du code en position **Eb**, d'après le lemme 3.46, $F + (-2, 1)$ est une 6^+-fenêtre, et d'après la condition 2, $F + (-1, 1)$ est une 6^+-fenêtre aussi. Nous pouvons donc supposer qu'il n'y a pas d'élément du code en position **Eb**.

S'il y a un élément du code en position **Cb**, d'après le lemme 3.46, $F + (-2, 0)$ est une 6^+-fenêtre, et d'après la condition 2, $F + (-1, 0)$ est aussi une 6^+-fenêtre. Nous pouvons donc supposer qu' il n'y a pas d'élément du code en position **Cb**.

S'il y a un élément du code en position **Gd**, $F + (1, -1)$ est nécessairement une 6^+-fenêtre, et d'après le lemme 3.46, $F + (1, -2)$ est une 6^+-fenêtre. Aussi nous pouvons supposer qu'il n'y a pas d'élément du code en position **Gd**.

Par symétrie, nous pouvons supposer qu'il n'y a pas d'élément du code parmi les positions **Be**, **Bc** et **Dg**. En utilisant la condition 2, cela implique qu'il y a deux éléments du code en position **Db** et **Bd**.

S'il y a un élément du code en position **Gg**, $F + (1, -1)$ est nécessairement une 6^+-fenêtre, et d'après le lemme 3.46, $F + (2, -1)$ est une 6^+-fenêtre. Aussi nous pouvons supposer qu'il n'y a pas d'élément du code en position **Gg**.

S'il y a un élément du code en position **Ge**, $F+(1, -1)$ est nécessairement une 6^+-fenêtre. S'il y a un élément du code en position **Fb**, $F+(-1, -1)$ est nécessairement une 6-fenêtre aussi. Sinon, $F + (-2, -1)$, $F + (-1, 1)$ ou $F + (-2, 2)$ est une 6^+-fenêtre. Nous pouvons donc supposer qu'il n'y a pas d'élément du code en position **Ge** et par symétrie, pas d'élément du code en position **Eg**.

S'il y a un élément du code en position **Fb** (resp. en position **Bf**), $F + (-1, 0)$ (resp. $F + (0, 1)$) est nécessairement une 6-fenêtre. S'il y a un élément du code dans les deux positions, c'est fini. Par symétrie, nous pouvons donc supposer qu'il n'y a pas d'élément du code en position **Fb**. Cela implique que $F + (-2, -1)$ est une 6^+-fenêtre. S'il y a une 6^+-fenêtre à distance 1 de F, c'est fini. Sinon, $F + (1, 2)$ est une 6^+-fenêtre. Si $F + (-2, 2)$ est une 6^+-fenêtre, c'est fini. Sinon, $F + (-2, -1)$ et $F + (1, 2)$ sont toutes deux des 7^+-fenêtres.

Cas A_{10}. Supposons qu'il n'y a pas d'élément du code parmi les positions **Ed** et **Ee**. Alors, d'après la condition 2, il y a des éléments du code en position **Dd** et **De**, et d'après le lemme 3.46, $F + (0, -1)$ et $F + (2, 1)$ sont des 6^+-fenêtres.

Supposons qu'il y ait un élément du code en position **Ed** ou **Ee**. Alors, d'après la condition 2, $F + (0, 1)$ est une 6^+-fenêtre. S'il y a un élément du code en position **Ed**, d'après le lemme 3.46, $F + (1, -1)$ est une 6^+-fenêtre. S'il y a un élément du code en position **Ee**, d'après le lemme 3.46, $F + (1, 1)$ est une 6^+-fenêtre.

Cas B_1. D'après le lemme 3.46, $F + (-1, 0)$ est une 6^+-fenêtre. D'après la condition 1, il y a un élément du code parmi les positions **Cb** et **Fb**. Alors, d'après le lemme 3.46, $F + (-2, 0)$ est une 6^+-fenêtre.

Cas B_2. D'après le lemme 3.46, $F + (-1, 0)$ est une 6^+-fenêtre. D'après la condition 2, il y a un élément du code parmi les positions **Dd** et **Ed**. S'il y a un élément du code en position **Dd** (resp. **Ed** et pas en **Dd**), d'après le lemme 3.46, $F + (1, 1)$ (resp. $F + (1, -1)$) est une 6^+-fenêtre.

Cas B_3. D'après le lemme 3.46, $F + (-1, 0)$ est une 6^+-fenêtre. Si $F + (1, 1)$ est une 6^+-fenêtre, c'est fini. Sinon, $F + (1, 1)$ est une 5-fenêtre et il y a un unique élément du code dans un coin, en position **Ed**. D'après la condition 2, cela implique

qu'il y a un élément du code en position **Bc**, et d'après le lemme 3.46, $F + (0, 2)$ est une 6^+-fenêtre.

Cas B_4**.** Supposons tout d'abord qu'il n'y a pas d'élément du code parmi les positions **De** et **Ee**. Alors d'après la condition 2, il y a des éléments du code aux positions **Dd** et **Ed**. Donc, d'après le lemme 3.46, $F + (1, 0)$ et $F + (-2, 0)$ sont des 6^+-fenêtres.

S'il y a un élément du code en position **De** ou **Ee**, alors $F + (-1, 0)$ est une 6^+-fenêtre. S'il y a un élément du code en position **De** (resp. **Ee**), d'après le lemme 3.46, $F + (1, -1)$ (resp. $F + (2, -1)$) est une 6^+-fenêtre.

Cas C_1**.** D'après le lemme 3.46, $F + (-1, 0)$ et $F + (0, 1)$ sont toutes deux des 6^+-fenêtres.

Cas C_2**.** D'après le lemme 3.46, $F + (0, 1)$ est une 6^+-fenêtre. D'après la condition 1, il y a un élément du code parmi les positions **Bc** et **Bf**. Alors, d'après le lemme 3.46, $F + (0, 2)$ est une 6^+-fenêtre.

Cas C_3**.** Supposons tout d'abord qu'il n'y a pas d'élément du code en position **De**. Alors, il y a des éléments du code en positions **Dd** et **Ee**. Alors, d'après le lemme 3.46, $F + (1, 1)$ et $F + (-1, -1)$ sont des 6^+-fenêtres. Supposons donc qu'il y a un élément du code en position **De**, et, par symétrie en position **Ed**.

S'il y a un élément du code en position **Dd**, d'après le lemme 3.46, $F + (0, 2)$ et $F + (0, -1)$ sont des 6^+-fenêtres. Aussi nous pouvons supposer qu' il n'y a pas d'élément du code en position **Dd**.

S'il y a un élément du code en position **Ee**, d'après le lemme 3.46, $F + (0, 1)$ et $F + (2, 0)$ sont des 6^+-fenêtres. Aussi nous pouvons supposer qu' il n'y a pas d'élément du code en position **Ee**.

S'il y a un élément du code en position **Bc** ou **Bf**, d'après le lemme 3.46, $F+(0, 2)$ est une 6^+-fenêtre et $F + (0, 1)$ doit être une 6^+-fenêtre aussi. Aussi nous pouvons supposer qu' il n'y a pas d'élément du code en position **Bc** et **Bf** et, par symétrie, en position **Cb** et **Fb**.

S'il y a un élément du code en position **Be**, alors $F + (0, 1)$ est une 6^+-fenêtre. Si $F + (-1, 1)$ ou $F + (-2, 2)$ est une 6^+-fenêtre. Sinon, il n'y a pas d'élément du code en position **Ca** et il y a donc un élément du code en position **Fa** d'après la condition 1. Cela implique que $F + (-2, -1)$ est une 6^+-fenêtre. Nous pouvons donc supposer qu'il n'y a pas d'élément du code en position **Be** et, par symétrie, pas d'élément du code en position **Eb**.

S'il y a un élément du code en position **Eg**, d'après le lemme 3.46, $F + (2, 1)$ et $F+(2, -2)$ sont des 6^+-fenêtres. Si $F+(1, 0)$ est une 6^+-fenêtre, c'est fini. Sinon, il n'y a pas d'élément du code en position **Dg**, et un élément du code parmi les positions **Cg** et **Fg**. S'il y a un élément du code en position **Fg**, d'après le lemme 3.46, $F + (1, -1)$ est une 6^+-fenêtre. S'il y a un élément du code en position **Cg**, d'après le lemme 3.46, $F + (2, 0)$ est une 6^+-fenêtre. Aussi nous pouvons supposer qu' il n'y a pas d'élément du code en position **Eg** et, par symétrie, en position **Ge**.

S'il y a un élément du code en position **Gc**, d'après le lemme 3.46, $F + (0, -2)$ et $F + (-2, -1)$ sont des 6^+-fenêtres. Si ce sont deux 6-fenêtres et que ce sont les seules dans la boule de rayon 2 de F, F est 2-pauvre. Alors, en utilisant les mêmes

techniques que précédemment, les éléments du code autour de F sont fixés comme sur la figure 3.24 et c'est fini. Sinon, il y a un excès de charge d'au moins 3 dans la boule de rayon 2 de F avec au moins de 6^+-fenêtres. Nous pouvons donc supposer qu'il n'y a pas d'élément du code en position **Gc** et, par symétrie, en position **Cg**.

Remarquons qu'à cette étape $F + (1, -1)$ est une 6^+-fenêtre. Si $F + (1, -1)$ est une 6-fenêtre et que c'est la seule dans la boule de rayon 2 de F alors F est 1-pauvre et tous les éléments du code dans le voisinage à distance 2 de F sont fixées comme sur la figure 3.23. Sinon, soit $F + (1, -1)$ est une 7^+-fenêtre et alors $F + (1, -2)$ est une 6^+-fenêtre, ou bien $F + (1, -1)$ est une 6-fenêtre et il existe une autre 6^+-fenêtre dans la boule de rayon 2 de F.

Cas D**.** D'après le lemme 3.46, $F + (0, 1)$ est une 6^+-fenêtre. D'après la condition 1, il y a un élément du code parmi les positions **Bc** et **Bf**. Alors, d'après le lemme 3.46, $F + (0, 2)$ est une 6^+-fenêtre. □

3.4.4 Conclusion sur la grille du roi

	$r = 1$	$r = 2$	$r \geq 3$
$\ell = 1$	$\frac{2}{9}$[20, 27]	$\frac{1}{4r}$[19]	
$\ell = 2$	$\frac{47}{111}$, $\frac{3}{7}$[55]	$\frac{31}{120}$[55], $\frac{2}{7}$[81]	$\frac{1}{4}$[55]
$\ell \geq 3$	pas de code [55]		

Table 3.1 – *Bornes ou valeurs exactes connues sur la densité optimale d'un code* $(r, \leq \ell)$*-identifiant dans la grille du roi.*

Le tableau 3.1 donne les bornes connues à ce jour sur la densité optimale des codes $(r, \leq \ell)$-identifiants pour différentes valeurs de r et ℓ dans la grille du roi. La densité optimale d'un code $(r, \leq \ell)$-identifiant est connue pour le cas $\ell = 1$ et r quelconque et le cas $\ell = 2$ et $r \geq 3$. Pour $\ell \geq 3$, il n'existe pas de code identifiant dans la grille du roi car les ensembles $\{(0, -1), (0, 0), (0, 1)\}$ et $\{(-1, 0), (0, 1)\}$ ont le même voisinage. Les deux cas où la valeur optimale n'est pas connue sont celui considéré dans cette partie ($r = 1$ et $\ell = 2$) et le cas $r = 2$, $\ell = 2$. Nous pensons que la méthode utilisée ici pourrait être reprise pour améliorer la borne inférieure de ce dernier cas. L'enjeu est de choisir le bon ensemble comme fenêtre. Nous ne pensons pas qu'avec celle utilisée ici nous puissions encore améliorer la borne inférieure pour le cas $r = 1$, $\ell = 2$, ou en tout cas pas de manière significative, l'analyse donnée ici étant déjà très poussée.

3.5 Conclusion du chapitre

Dans ce chapitre, nous avons étudié les codes identifiants pour quatre classes de graphes. Ces classes peuvent être regroupées en deux catégories : les classes de graphes dépendantes d'une métrique et pour ces raisons étudiées en théorie des codes, et les classes de graphes héréditaire étudiées en théorie des graphes et particulièrement en optimisation combinatoire.

D'un côté, nous avons étudié les graphes de Sierpiński et la grille du roi (métrique ℓ_∞), qui sont des classes de graphes définis par une métrique. Ces graphes sont au cœur des études sur les codes. Pour les graphes de Sierpiński, il est souvent possible de déterminer les valeurs exactes des tailles optimales des codes considérés. Est-il possible d'étendre cette famille de graphes tout en gardant ces bonnes propriétés pour les problèmes de codage ? Les graphes de Sierpiński généralisés sont une extension potentielle qui nous semble pertinent d'étudier. Pour la grille du roi, bien que sa structure soit très régulière, nous ne connaissons pas la valeur exacte de la densité optimale d'un code $(1, \leq 2)$-identifiant. La technique utilisée ici pour améliorer la borne inférieure est une technique assez classique de théorie des codes, qui pourrait amener à la valeur optimale en choisissant un ensemble de départ différent de celui utilisé ici.

D'un autre côté, nous avons étudié les graphes d'intervalles et les graphes adjoints, qui sont des classes de graphes héréditaires pouvant être vus comme graphes d'intersections d'objets. Ces familles peuvent être définies en interdisant une famille (infinie ou finie) de sous-graphes induits, et sont, par conséquent, d'un intérêt majeur en théorie des graphes. Pour le problème des codes identifiants, nous ne savons pas déterminer la taille optimale d'un code identifiant, le problème étant par exemple $\mathcal{NP}$-difficile pour la classe des graphes adjoints. Nous avons par contre augmenté la borne inférieure sur cette valeur, montrant que les codes identifiants sont d'ordre au moins $\theta(\sqrt{n})$. Il serait intéressant d'étudier d'autres classes de graphes du même type, comme les graphes (fortement) triangulés ou les graphes parfaits, d'y étendre les résultats obtenus et d'affiner les résultats de complexité obtenus.

Chapitre 4

Modifier les voisinages des sommets

Sommaire

4.1 Codes identifiants évolutifs 118
4.1.1 Définitions et premières remarques 118
4.1.2 Sans mémoire dans les cycles 123
4.1.3 Avec mémoire dans les cycles 128
4.1.4 Limitons le nombre de rayons par sommet 131
4.2 Codes identifiants tolérants 134
4.2.1 Cadre : le plan euclidien 134
4.2.2 Codes identifiants tolérants 135
4.2.3 Premier exemple 137
4.2.4 Existence d'un code identifiant tolérant 137
4.2.5 Étude du motif horizontal 139
4.2.6 Bornes générales 141
4.2.7 Amélioration des bornes pour des valeurs fixées 144
4.2.8 Étude des petites valeurs de r 154
4.2.9 Conclusion sur les codes identifiants tolérants 159
4.3 Conclusions et perspectives du chapitre 160

Jusqu'à présent, nous avons identifié les sommets du graphe uniquement avec le voisinage fermé des sommets, autrement dit, avec des capteurs capables de détecter une erreur que dans l'ensemble entier de leur voisinage fermé. Nous considérons dans ce chapitre des procédés d'identification donnant plus de flexibilité aux capteurs.

Parmi les problèmes d'identification rentrant dans ce cadre, nous pouvons citer les *systèmes de surveillance* (*watching systems*). Dans les systèmes de surveillance, introduits en 2010 par Auger, Charon, Hudry et Lobstein [3], un capteur est toujours situé sur un sommet mais n'a qu'une partie de son voisinage fermé à surveiller. Plus formellement, un système de surveillance est un ensemble de couples, appelés *contrôleurs*, (u, Z) avec u un sommet du graphe et $Z \subseteq N[u]$, tels que pour chaque sommet v, l'ensemble des contrôleurs surveillant v soit non vide et unique. Il est possible d'avoir plusieurs contrôleurs sur le même sommet. Tout code identifiant peut être vu comme un système de contrôle. À l'inverse, pour certains graphes, il existe des systèmes de contrôles beaucoup plus efficaces (en terme de nombre de contrôleurs) que les codes identifiants. C'est le cas par exemple de l'étoile $K_{1,n}$ pour laquelle il existe un système de contrôle avec $\log(n+2)$ contrôleurs (tous placés sur le sommet central), alors que le plus petit code identifiant est de taille n.

Nous nous intéressons ici à deux autres variantes motivées par des considérations pratiques. La première utilise des capteurs qui peuvent tester plusieurs rayons de voisinage. Le processus d'identification s'effectue en plusieurs étapes : les capteurs commencent par tester les sommets à distance 1, puis 2, ainsi de suite jusqu'à atteindre leur rayon limite de détection. Nous considérons deux cas selon si le processus a une mémoire ou non. Le cas avec mémoire revient simplement à considérer des ensembles résolvants avec un rayon limité. Nous donnons les valeurs exactes de la taille optimale de ces codes pour la famille des cycles. La deuxième variante considérée utilise des capteurs qui détectent des erreurs jusqu'à un rayon d'environ r, des incertitudes pouvant avoir lieu aux limites du voisinage. Pour pouvoir quantifier l'incertitude sur les capteurs, nous nous plaçons dans le plan muni de la distance euclidienne.

4.1 Codes identifiants évolutifs

Nous commençons par définir les codes identifiants évolutifs et les relions à certains problèmes d'identification. Nous étudions ensuite ces deux problèmes dans la famille des cycles puis donnons une définition générale qui englobe ces problèmes ainsi que celui des codes identifiants. Ces travaux sont issus d'une collaboration avec Olivier Delmas, Mickaël Montassier et Sylvain Gravier [33].

4.1.1 Définitions et premières remarques

Définitions

Jusqu'à présent, le rayon des voisinages considérés était 1 ou sous-entendu par le contexte. Dans cette partie, les valeurs des rayons vont être cruciales, nous intro-

duisons donc une terminologie prenant en compte les rayons. Le *r-voisinage* (sous-entendu, *fermé*) du sommet x est l'ensemble des sommets à distance au plus r de x, noté $\mathcal{B}_r(x)$. Deux sommets x et y sont dit *r-séparés* (ou encore *séparés pour le rayon* r) par un code C si $\mathcal{B}_r(x) \cap C \neq \mathcal{B}_r(y) \cap C$. Un sommet x d'un graphe G est *r-séparé* (ou encore, *séparé pour le rayon* r) par C si pour tout autre sommet y, x et y sont r-séparés. Un sous-ensemble de sommets est *r-dominant* si l'union des r-voisinages de l'ensemble couvre l'ensemble des sommets. Nous utiliserons de nombreuses fois des ensembles d'entiers consécutifs. Pour plus de clarté, nous utiliserons la notation $[\![a, b]\!]$ pour désigner tous les entiers compris entre a et b.

Code r-identifiant évolutif sans mémoire : Un *code r-identifiant évolutif sans mémoire* d'un graphe G est un sous-ensemble de sommets r-dominant $C \subseteq V(G)$ tel que chaque sommet x est séparé par C pour un rayon r_x dans $[\![0, r]\!]$:

$$\forall x \in V(G), \exists r_x \in [\![0, r]\!],\ \forall y \neq x \in V(G), \mathcal{B}_{r_x}(x) \cap C \neq \mathcal{B}_{r_x}(y) \cap C$$

Le rayon r_x est appelé *rayon de séparation* de x.

Code r-identifiant évolutif avec mémoire : Un *code r-identifiant évolutif avec mémoire* de G est un ensemble r-dominant $C \subseteq V(G)$ tel que chaque paire de sommets (x, y) est séparée par C pour un rayon r_{xy} dans $[\![0, r]\!]$:

$$\forall x \in V(G), \forall y \neq x \in V(G), \exists r_{xy} \in [\![0, r]\!],\ \mathcal{B}_{r_{xy}}(x) \cap C \neq \mathcal{B}_{r_{xy}}(y) \cap C$$

Comme pour les codes identifiants, nous étudions la taille minimale d'un code identifiant évolutif de G. Nous noterons $\gamma_r^{\text{IDD}}(G)$ (resp. $\gamma_r^{\text{IDDM}}(G)$) la cardinalité minimum d'un code r-identifiant évolutif sans mémoire (resp. code r-identifiant évolutif avec mémoire) de G. Tout code r-identifiant est un code r-identifiant évolutif sans mémoire et tout code r-identifiant évolutif sans mémoire est un code r-identifiant évolutif avec mémoire. Aussi nous avons la chaîne d'inégalités suivante :

$$\gamma_r^{\text{ID}}(G) \geq \gamma_r^{\text{IDD}}(G) \geq \gamma_r^{\text{IDDM}}(G).$$

L'inverse n'est pas vrai comme nous allons le voir dans la suite. Notons aussi que contrairement aux codes identifiants, il y a toujours un code identifiant évolutif avec ou sans mémoire dans G (en prenant comme code l'ensemble des sommets).

Nous pouvons aussi remarquer qu'un code 1-identifiant évolutif avec mémoire est un code localisateur-dominateur.

Premiers exemples

La figure 4.1 montre un code 2-identifiant évolutif sans mémoire du graphe P_5. Les sommets v_3 et v_4 sont chacun d'eux séparés pour le rayon 0, les sommets v_2 et v_5 pour le rayon 1 et le sommet v_1 pour le rayon 2. Ce code n'est pas un code 2-identifiant car les sommets v_2,v_3,v_4 et v_5 ont tous le même voisinage sous C à rayon 2.

La figure 4.2 montre un code 2-identifiant évolutif avec mémoire de P_5 qui n'est pas un code 2-identifiant évolutif sans mémoire : le sommet v_2 n'est séparé du sommet v_1 que pour le rayon 0, et pour ce rayon, les sommets v_2 et v_3 ne sont pas séparés.

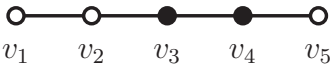


Figure 4.1 – *Un code 2-identifiant évolutif sans mémoire qui n'est pas 2-identifiant.*

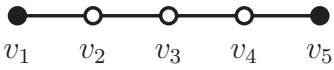


Figure 4.2 – *Un code 2-identifiant évolutif avec mémoire mais pas évolutif sans mémoire.*

Mise en œuvre

Expliquons maintenant pourquoi ces codes permettent d'identifier un sommet défectueux avec un procédé en plusieurs étapes. Considérons un graphe G muni d'un code r-identifiant évolutif sans mémoire C dans lequel il y a potentiellement un sommet défectueux. Nous supposons que les rayons r_x de séparation de chaque sommet sont connus, ainsi que les ensembles $\mathcal{B}_{r_x}(x) \cap C$ correspondant.

Pour localiser le sommet défectueux potentiel, nous interrogeons tout d'abord le 0-voisinage de chaque sommet de C, ce qui revient à interroger si un des sommets du code est défectueux. Si une des réponses est positive, nous pouvons localiser l'erreur. Sinon nous interrogeons les 1-voisinages de chaque sommet du code. Si l'ensemble des réponses positives est non vide et correspond à l'un des ensembles $\mathcal{B}_1(x) \cap C$ pour lequel $r_x = 1$, alors nous arrêtons le processus, l'élément défectueux est le sommet correspondant. Sinon, nous continuons le processus en interrogeant les 2-voisinages de chaque sommet. De la même manière, si l'ensemble des réponses positives est non vide et correspond à l'un des ensembles $\mathcal{B}_2(x) \cap C$ pour lequel $r_x = 2$, alors nous arrêtons le processus, l'élément défectueux est le sommet correspondant. Si l'ensemble des réponses positives est non vide, qu'il était vide à l'étape 1 et qu'un sommet x vérifie $\mathcal{B}_1(x) \cap C = \emptyset$ et $r_x = 1$, nous arrêtons le processus, x est l'élément défectueux. Sinon nous continuons ce processus jusqu'au rayon r.

Si aucune réponse positive n'a eu lieu jusqu'au rayon r, nous pouvons conclure qu'il n'y a pas de sommets défectueux. Sinon nous avons deux cas : soit le sommet défectueux x vérifie $\mathcal{B}_{r_x}(x) \cap C \neq \emptyset$ et dans ce cas, le processus s'arrêtera au rayon r_x. Soit $\mathcal{B}_{r_x}(x) \cap C = \emptyset$ et le processus s'arrêtera au rayon $r_x + 1$ et $\mathcal{B}_{r_x+1}(x) \cap C \neq \emptyset$. En effet, à l'étape r_x, nous sommes seulement capable de dire que soit il n'y a pas d'erreurs, soit le sommet défectueux est le sommet x (puisque tous les autres sommets vérifient $\mathcal{B}_{r_x}(y) \cap C \neq \emptyset$). Il faudra attendre la première réponse positive pour pouvoir distinguer les deux cas, si x est défectueux, cela arrivera au rayon $r_x + 1$.

Une manière de pallier ce problème de l'ensemble vide identifiant serait d'interdire cette possibilité. Nous avons choisi de la garder par souci d'homogénéité.

Dans le cas d'un code r-identifiant évolutif avec mémoire, nous effectuons exactement le même processus, en enregistrant à chaque étape l'ensemble des réponses positives obtenues. Nous pouvons alors arrêter le processus lorsque les réponses

obtenues jusqu'alors correspondent à un unique sommet. C'est une localisation en plusieurs passes, avec mémoire, puisque l'on doit retenir tous les réponses obtenues auparavant.

Lien avec les ensembles résolvants

Rappelons qu'un *ensemble résolvant* d'un graphe G est un sous-ensemble de sommets $\{v_1, ..., v_k\}$ tel que les k-uplets $(d(x, v_1), ..., d(x, v_k))$ soient différents pour chaque sommet x du graphe. Autrement dit, un ensemble est résolvant si et seulement si toute paire de sommets est séparée pour un certain rayon, sans borne sur le rayon. De ce point de vue, un code r-identifiant évolutif avec mémoire est toujours un ensemble résolvant du graphe et un ensemble résolvant est un code δ-identifiant évolutif avec mémoire où δ est le diamètre du graphe (la plus grande distance entre deux sommets du graphe). Nous avons donc :

$$\gamma_r^{\text{IDDM}}(G) \geq \dim(G) = \gamma_\delta^{\text{IDDM}}(G)$$

où $\dim(G)$ désigne la taille du plus petit ensemble résolvant de G ($\dim(G)$ est appelé *dimension* du graphe G).

Les codes r-identifiants évolutifs avec mémoire sont donc une généralisation des ensembles résolvants où le rayon de détection des capteurs est borné, ce qui semble pertinent pour les applications des ensembles résolvants, un capteur n'ayant rarement en pratique un rayon d'action infini. Les codes r-identifiants évolutifs avec mémoire peuvent ainsi être vus comme un intermédiaire entre les ensembles résolvants et les codes identifiants.

Borne inférieure pour les codes r-identifiants évolutifs sans mémoire

Le résultat suivant est l'analogue pour les codes r-identifiants évolutifs sans mémoire de la borne inférieure $\log(|V(G)| + 1)$ des codes identifiants (théorème 1.3).

Proposition 4.1 *Soient r et k deux entiers et soit $w_r(k)$ le nombre maximal de sommets d'un graphe G qui possède un code r-identifiant évolutif sans mémoire de taille k. Nous avons :*

$$w_r(k) = k + r(2^k - 2)$$

Preuve : Nous commençons par construire un graphe H_r^k ayant $k + r(2^k - 2)$ sommets et possédant un code r-identifiant évolutif sans mémoire de taille k. La figure 4.3 montre la construction dans le cas $r = 4$ et $k = 3$.

Le graphe H_r^k a pour ensemble de sommets l'ensemble $C \cup I_1 \cup ... \cup I_r$ où $C = \{1, ..., k\}$ et I_j est de taille $2^k - 2$ pour $1 \leq j \leq r$; chaque sommet de I_j correspondant à un ensemble non trivial de $\{1, ..., k\}$. Les arêtes de H_r^k sont les suivantes. Chaque sommet x de I_1 est adjacent aux sommets de C qui forment le sous-ensemble de $\{1, ..., k\}$ correspondant à x. Chaque sommet de I_j pour $j > 1$ est relié au sommet de I_{j-1} qui correspond au même sous ensemble. De plus, tous les sommets de C sont connectés.

Le graphe H_r^k a $k+r(2^k-2)$ sommets et l'ensemble C forme un code r-identifiant évolutif sans mémoire de H_r^k. Nous venons donc de montrer que $w_r(k) \geq k+r(2^k-2)$.

Pour montrer l'autre sens, prenons un graphe G d'ordre maximum qui possède un code r-identifiant évolutif sans mémoire C de taille k. Comptons pour chaque rayon $i \leq r$ combien de sommets au maximum peuvent être i-identifiés.

Pour le rayon 0, les k sommets du code sont identifiés et seulement ceux-ci (aucun sommet ne peut-être identifié par l'ensemble vide à cette étape, car il y a strictement plus de $k+1$ sommets dans G).

Pour le rayon 1, au plus 2^k sommets supplémentaires peuvent être identifiés (un par sous-ensemble de $\{1, ..., k\}$). Cependant, un sommet du code c donne un ensemble $\mathcal{B}_1(c) \cap C$ qui ne peut pas être l'ensemble identifiant d'un sommet hors de C. Nous avons donc au plus 2^k-1 sommets supplémentaires qui peuvent être identifiés. Si les 2^k-1 sommets sont réalisés, alors nécessairement, pour tous les éléments c du code, l'ensemble $\mathcal{B}_1(c) \cap C$ est le même ensemble, et c'est donc l'ensemble C tout entier. De plus, l'ensemble vide est aussi réalisé par un de sommets qui n'est pas dans C. Donc C n'est pas un ensemble 1-dominant, et nous avons donc $r \geq 2$. Le graphe G ne peut alors pas contenir d'autres sommets : en effet, si G contient un sommet x qui n'est pas encore identifié, alors $x \notin C$ et $\mathcal{B}_1(x) \cap C = C$ (sinon x n'est pas séparé d'un sommet qui serait identifié au rayon 1). Mais x ne pourra jamais être identifié car pour tout $r_x \geq 1$, $\mathcal{B}_{r_x}(x) \cap C = C$. Ainsi, si 2^k-1 sommets sont identifiés au rayon 1, alors G a $k+2^k-1$ sommets et $r=2$. Or G a au moins $k+2(2^k-2)$ sommets, ce qui est contradictoire. Il s'ensuit qu'au plus 2^k-2 sommets sont identifiés pour le rayon 1.

Pour un rayon $2 \leq i \leq r$, nous procédons de la même manière pour montrer qu'au plus 2^k-2 sommets sont identifiés pour le rayon i.

En additionnant le nombre de sommets identifiés à chaque rayon, nous pouvons conclure que l'ordre de G est au plus $k+r(2^k-2)$, ce qui termine la preuve. □

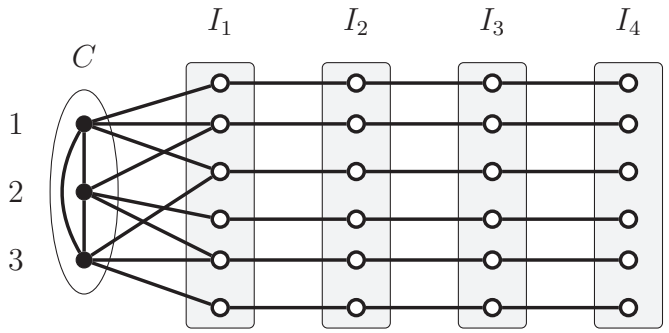


Figure 4.3 – *Le graphe H_4^3 qui est un exemple de graphe extrémal contenant un code 4-identifiant évolutif sans mémoire de taille 3.*

4.1.2 Codes identifiants évolutifs sans mémoire dans la famille des cycles

L'étude des codes r-identifiants classiques dans la famille des cycles, qui est pourtant une famille simple de graphe, n'a pas été évidente. Bertrand *et al.* [10] et Gravier *et al.* [49] ont posé les premières bases, résolu le cas $r = 1$, le cas où le nombre de sommets est pair ainsi qu'un certain nombre d'autres cas. D'autres articles ont suivi résolvant encore d'autres petits cas [25, 44, 85, 95]. Finalement, Junnila et Laihonen [62] ont complété les cas restants.

Dans cette partie et la suivante, nous étudions les codes identifiants évolutifs dans les cycles. L'étude est plus aisée que pour les codes identifiants classiques et nous sommes en mesure de donner la valeur exacte des tailles optimales de ces codes pour tous les cas.

Nous notons $\mathcal{C}_n$ le cycle de taille n et $\{v_0, v_1, \ldots, v_{n-1}\}$ l'ensemble de ses sommets, les indices des sommets seront considérés modulo n. Nous supposons dans un premier temps que $n \geq 2r + 2$.

Les sommets d'un code r-identifiant évolutif sans mémoire ne peuvent pas être trop éloignés :

Lemme 4.2 *Soit S un ensemble de $2r + 2$ sommets consécutifs de $\mathcal{C}_n$. Tout code r-identifiant évolutif sans mémoire de $\mathcal{C}_n$ doit contenir au moins deux éléments de C.*

Preuve : Supposons par l'absurde qu'il existe un code r-identifiant évolutif sans mémoire C de $\mathcal{C}_n$ qui ne contient qu'un élément a de S. Sans perte de généralité, $S = \{v_0, v_1, \ldots, v_{2r+1}\}$ et $a = v_i$ avec $i \leq r$. La figure 4.4 illustre la situation.

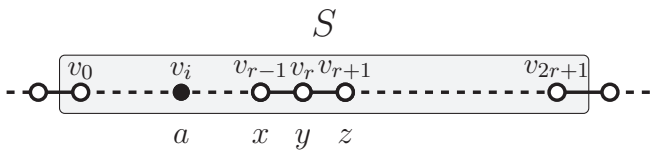


Figure 4.4 – *Notations pour la preuve du lemme 4.2*

Nous étudions les trois sommets $x = v_{r-1}$, $y = v_r$ et $z = v_{r+1}$. Les r-voisinages de y et z sont inclus dans S. Pour un rayon t, les ensembles $\mathcal{B}_t(y) \cap C$ et $\mathcal{B}_t(z) \cap C$ sont donc soit vides soit contiennent seulement l'élément a. Ces deux ensembles diffèrent uniquement pour le rayon $t = d(a, y) = r - i$ et nous avons donc $r_y = r_z = t$ (r_y et r_z désignent les rayons de séparation du code C). L'ensemble $\mathcal{B}_t(y) \cap C$, qui est exactement $\{a\}$, doit être différent de l'ensemble $\mathcal{B}_t(x) \cap C$. Cet ensemble doit donc contenir un élément b, différent de a. Nécessairement, $b \notin S$. Comme b est dans le r-voisinage de x mais pas dans S, $b = v_{-1}$ et donc, $t = r$ et $i = 0$. Le sommet z n'est pas r-dominé ce qui contredit le fait que C soit un code r-identifiant évolutif sans mémoire. □

Le lemme précédent nous donne une première borne qui va s'avérer être serrée dans un grand nombre de cas.

Corollaire 4.3 *Soit C un code r-identifiant évolutif sans mémoire de $\mathcal{C}_n$. Alors :*

$$|C| \geq \left\lceil \frac{n}{r+1} \right\rceil.$$

Preuve : Dans $\mathcal{C}_n$, il y a n ensembles différents de $2r+2$ sommets consécutifs. Si C est un code r-identifiant évolutif sans mémoire, d'après le lemme 4.2, il y a au moins deux sommets du code dans chacun de ces ensembles. Chaque sommet du code est compté exactement $2r+2$ fois, ce qui donne :

$$|C| \geq \left\lceil \frac{2n}{2r+2} \right\rceil = \left\lceil \frac{n}{r+1} \right\rceil.$$

□

L'étude des codes r-identifiants évolutifs sans mémoire dans les cycles va donc dépendre des congruences de n modulo $2r+2$. Notons donc n sous la forme $n = (2r+2)p + R$ avec $0 \leq R \leq 2r+1$ et $p \geq 1$ (car $n \geq 2r+2$).

Le corollaire 4.3 peut alors se reformuler ainsi : si C est code r-identifiant évolutif sans mémoire de $\mathcal{C}_n$, nous avons :

- si $R = 0$, $|C| \geq 2p$;
- si $1 \leq R \leq r+1$, $|C| \geq 2p+1$;
- si $r+2 \leq R \leq 2r+1$, $|C| \geq 2p+2$.

Les prochains lemmes donnent des constructions de codes qui mèneront à des bornes supérieures. Certaines sont serrées, pour d'autres, nous aurons besoin de raffiner la borne inférieure.

Lemme 4.4 *Lorsque $n = (2r+2)p$, $\mathcal{C}_n$ possède un code r-identifiant évolutif sans mémoire de cardinalité $2p = \frac{n}{r+1}$.*

Preuve : Nous construisons un code périodique avec comme motif de base le motif représenté dans la figure 4.5. Plus précisément, nous posons

$$C = \{v_i | i \equiv r \mod 2r+2 \text{ ou } i \equiv r+1 \mod 2r+2\}.$$

L'ensemble C est un ensemble r-dominant de cardinalité $2p$. Nous posons $r_k = r - k$ si $k \in [\![0, r]\!]$, $r_k = k - (r+1)$ si $k \in [\![r+1, 2r+1]\!]$, et pour $k > 2r+1$, $r_k = r_{k'}$ avec $k \equiv k' \mod 2r+2$. Alors le sommet v_k est r_k-séparé par C : ce sommet est le seul à avoir comme unique sommet dans son r_k-voisinage le sommet du code est le plus proche de v_k. Ainsi C est un code r-identifiant évolutif sans mémoire, qui est optimal d'après le corollaire 4.3. La figure 4.6 donne un exemple de tel code pour le cycle de taille 12. □

Pour le cas général, nous utilisons la construction précédente en ajoutant un ou deux sommets dans la partie restante.

Lemme 4.5 *Lorsque $R = 1$, $\mathcal{C}_n$ possède un code r-identifiant évolutif sans mémoire de cardinalité $2p+1$. Lorsque $2 \leq R \leq 2r+1$, $\mathcal{C}_n$ possède un code r-identifiant évolutif sans mémoire de cardinalité $2p+2$.*

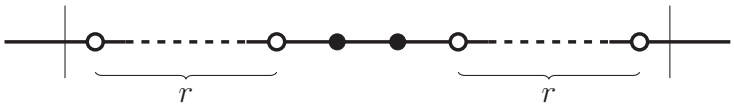


Figure 4.5 – *Le motif utilisé pour construire un code r-identifiant évolutif sans mémoire pour les cycles $\mathcal{C}_{(2r+2)p}$ avec $p \geq 1$.*

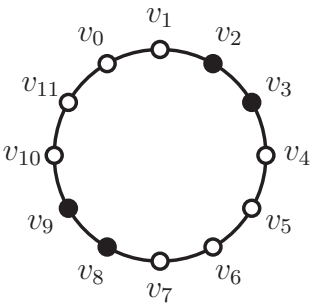


Figure 4.6 – *Un code 2-identifiant évolutif sans mémoire optimal pour le cycle $\mathcal{C}_{12}$*

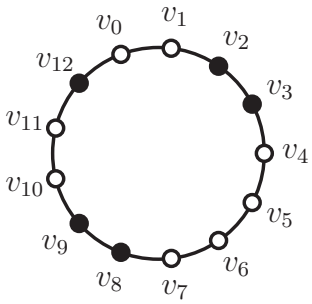


Figure 4.7 – *Un code 2-identifiant évolutif sans mémoire optimal pour le cycle $\mathcal{C}_{13}$*

Preuve : Supposons tout d'abord que $R = 1$. Le code

$$C = \{v_i | i \equiv r \mod 2r+2 \text{ ou } i \equiv r+1 \mod 2r+2\} \cup \{v_{n-1}\}$$

est un code r-identifiant évolutif sans mémoire de $\mathcal{C}_n$ et $|C| = 2p+1$. La figure 4.7 donne un exemple de ce code pour $\mathcal{C}_{13}$.

Pour $R \geq 2$, si $R \geq r+2$, nous posons

$$C = \{v_i | i \equiv r \mod 2r+2 \text{ ou } i \equiv r+1 \mod 2r+2\}.$$

Sinon $R \geq r+2$, nous posons

$$C = \{v_i | i \equiv r \bmod 2r+2 \text{ ou } i \equiv r+1 \bmod 2r+2\} \cup \{v_{n-2}, v_{n-1}\}.$$

Alors C est un code r-identifiant évolutif sans mémoire de $\mathcal{C}_n$ de cardinalité $2p+2$. □

Ces codes ne sont pas toujours optimaux :

Lemme 4.6 *Pour $(r, R) = (1, 2)$, $\mathcal{C}_n$ a un code 1-identifiant évolutif sans mémoire optimal de cardinalité $2p+1$.*

Pour $(r, R) = (2, 2)$, $\mathcal{C}_n$ a un code 2-identifiant évolutif sans mémoire optimal de cardinalité $2p+1$.

Preuve : Pour $(r, R) = (1, 2)$, l'ensemble $C = \{v_i | i \equiv 0 \bmod 2\}$ est un code 1-identifiant évolutif sans mémoire : chaque sommet x dans l'ensemble C est 0-séparé par C et chaque sommet x hors de C est 1-séparé par C.

De la même manière, pour $(r, R) = (2, 2)$, l'ensemble $C = \{v_i | i \equiv 0 \mod 6$ ou $i \equiv 2 \mod 6\}$ est un code 2-identifiant évolutif sans mémoire.

D'après le corollaire 4.3, ces deux codes sont optimaux. □

Le prochain lemme montre que la borne inférieure issue du corollaire 4.3 n'est pas serrée lorsque $2 \leq R \leq r + 1$ et $(r, R) \notin \{(1, 2), (2, 2)\}$. Cela implique que les codes du lemme 4.5 sont optimaux.

Lemme 4.7 *Lorsque $2 \leq R \leq r + 1$ et $(r, R) \notin \{(1, 2), (2, 2)\}$, $\mathcal{C}_n$ n'a pas de code r-identifiant évolutif sans mémoire de cardinalité $2p + 1$.*

Preuve : Supposons par l'absurde qu'il existe un code r-identifiant évolutif sans mémoire C de $\mathcal{C}_n$ de cardinalité $2p + 1$.

Observons tout d'abord les deux faits suivant :

(O.1) *Parmi R sommets consécutifs du cycle, il y a au plus un sommet du code C. En effet, si ce n'est pas le cas, dans le reste de $\mathcal{C}_n$ il y aurait au plus $2p-1$ sommets du code parmi $(2r + 2)p$ sommets consécutifs, ce qui contredit le lemme 4.2. En particulier, comme $R > 1$, il n'y a pas deux sommets consécutifs dans C.*

(O.2) *Pour des raisons similaires, parmi $2r + 2 + R$ sommets consécutifs de $\mathcal{C}_n$, il y a au plus trois sommets du code C.*

Soit M le nombre maximum de sommets consécutifs du cycle qui ne sont pas dans C et soit S_M un ensemble de M sommets consécutifs non éléments de C. D'après l'observation O.1, nous avons $M \geq R - 1$. Nous avons en fait dans tous les cas $M > 1$: si $M = 1$, nécessairement $R = 2$ et le code contient exactement un sommet sur deux, donc $|C| = \frac{n}{2} = 2p + 1$, $n = 4p + 2$ et $(r, R) = (1, 2)$ qui est un cas hors du lemme.

Notons c_1 et c_2 les deux éléments du code qui entoure l'ensemble S_M, et notons S_1 et S_2 les deux ensembles maximaux de sommets consécutifs hors de C situés avant c_1 et après c_2 (voir la figure 4.8 pour les notations). Enfin, nous notons c_0 et c_3 les deux sommets du code qui sont avant S_1 et après S_2.

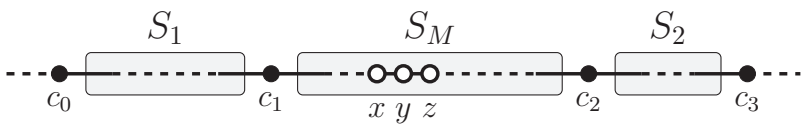


Figure 4.8 – *Notation pour la preuve du lemme 4.7. Les ensembles S_1, S_2 et S_M ne contiennent pas d'éléments du code.*

Nous avons $M \geq 1$, donc C est de cardinalité au moins 2, et par l'observation O.1, les ensembles S_1 et S_2 ne sont pas vide. Les éléments c_1, c_2 et c_3 sont distincts deux à deux ainsi que les éléments c_0, c_1 et c_2. Par contre, c_0 et c_3 peuvent désigner le même sommet.

Toujours d'après l'observation O.1, les ensembles S_1, S_2 et S_M ont au moins $R - 1$ éléments. Nous notons S l'ensemble $S_1 \cup \{c_1\} \cup S_M \cup \{c_2\} \cup S_2$.

Remarquons que $|S| \geq 2r + 3$. En effet, si c_0 et c_3 sont des sommets différents, alors $\{c_0\} \cup S \cup \{c_3\}$ est ensemble contenant quatre sommets du code, donc d'après

l'observation O.2, $|S|+2 > 2r+2+R \geq 2r+4$. Si c_0 et c_3 désigne le même sommet, alors $S \cup \{c_3\}$ correspond à l'ensemble des sommets du cycle. Alors nécessairement, $p = 1$ et $|S| = n - 1 = 2r + 1 + R \geq 2r + 3$.

Nous pouvons donc trouver trois sommets consécutifs x,y,z de S tels que chacun de leurs r-voisinages soit inclus dans S. Autrement dit :

$$\{\mathcal{B}_r(x) \cup \mathcal{B}_r(y) \cup \mathcal{B}_r(z)\} \cap C \subseteq \{c_1, c_2\}.$$

Nous pouvons de plus imposer que y soit dans S_M.

Pour séparer y de x, le rayon de séparation de y, r_y, ne peut être que $d(x, c_1)$ ou $d(y, c_2)$. Pour séparer y de z, r_y ne peut être que $d(y, c_1)$ ou $d(z, c_2)$. Ainsi nous avons soit $r_y = d(x, c_1) = d(z, c_2)$, soit $r_y = d(y, c_2) = d(y, c_1)$. Dans tous les cas, M doit être impair et y doit être l'élément médian de S_M, ce qui signifie que $d(y, c_2) = d(y, c_1)$. Comme $M \neq 1$, nous avons en fait $M \geq 3$ et $x, z \in S_M$.

Notons d_y la distance $d(y, c_1)$ et w le sommet situé juste avant x. Alors $\mathcal{B}_r(w) \cap C \subseteq \{c_0, c_1, c_2\}$. Pour séparer x de y, r_x doit être égal à $d(y, c_2) = d_y$ ou bien à $d(x, c_1) = d_y - 1$. Pour séparer x de w, r_x doit être égal à $d(w, c_1) = d_y - 2$, à $d(x, c_2) = d_y + 1$ ou bien à $d(w, c_0)$. Pour satisfaire les deux conditions nous avons donc nécessairement $r_x = d(w, c_0)$. Cela implique que $d(w, c_0) = r$ car $d(w, c_0) = d(x, c_0) - 1 \geq r$ et $r_x \leq r$. Comme $d_y \leq r$ et $r_x = d_y$ ou $r_x = d_y - 1$, $r_x = d_y = r$. Finalement, $M = 2r - 1$, $|S_1| = 1$, et $R = 2$. En utilisant les mêmes arguments pour z, nous obtenons la situation décrite par la figure 4.9.

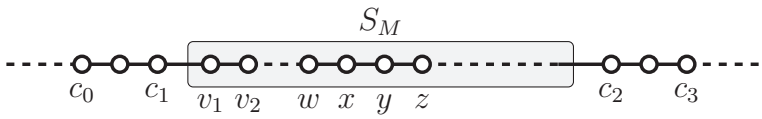


Figure 4.9 – *Les ensembles S_1, S_2 et S_M après quelques déductions.*

Comme $(r, R) \notin \{(1, 2), (2, 2)\}$ et que $R = 2$, nous avons $r \geq 3$ et donc $M \geq 5$. Soient v_1 et v_2 les deux sommets consécutifs de S_M situés après c_1 (voir la figure 4.9). Nous avons $d(v_2, c_2) = M - 1 > r$ et $d(v_1, c_2) > r$ donc v_1 et v_2 ne peuvent être séparés que par des éléments du code situés à la gauche de v_1 et v_2. Considérons r_{v_1} le rayon qui sépare v_1. Il doit y avoir un élément du code à gauche et à distance r_{v_1} de v_1 pour séparer v_1 et v_2. Pour des raison similaires, il doit y avoir un élément du code à distance $r_{v_1} + 1$ de v_1 pour séparer v_1 de c_1. Nous avons donc deux éléments consécutifs qui sont dans le code, ce qui contredit l'observation O.1 et termine la preuve. □

Nous avons maintenant tous les éléments pour conclure l'étude des codes r-identifiants évolutifs sans mémoire sur les cycles, lorsque $p \geq 1$:

Théorème 4.8 *Soit r un entier et $n = (2r+2)p+R$, avec $0 \leq R \leq 2r+1$ et $p \geq 1$.*

(i) Si $R = 0$, $\gamma_r^{IDD}(\mathcal{C}_n) = 2p$;

(ii) Si $R = 1$ ou si $r \leq 2$ et $R = 2$, $\gamma_r^{IDD}(\mathcal{C}_n) = 2p + 1$;

(iii) Sinon, $R \geq 2$ et $(r, R) \neq (1, 2)$ et $(2, 2)$, et : $\gamma_r^{IDD}(\mathcal{C}_n) = 2p + 2$.

Enfin, ce dernier lemme termine l'étude avec les petits cas.

Lemme 4.9 *Soient r et n deux entiers tels que $3 \leq n \leq 2r+1$. Alors $\gamma_r^{IDD}(\mathcal{C}_n) = 2$.*

Preuve : D'une part, un singleton ne peut pas être un code r-identifiant évolutif sans mémoire car ses deux voisins ne sont séparés pour aucun rayon.

D'autre part, deux sommets adjacents forment un code r-identifiant évolutif sans mémoire pour tout r. Il est à noter que si n est impair, le sommet antipodal aux deux sommets du code est identifié par l'ensemble vide. □

4.1.3 Codes identifiants évolutifs avec mémoire dans la famille des cycles

Nous étudions dans cette partie les codes r-identifiants évolutifs avec mémoire dans les cycle. Dans un premier temps, nous supposons que $n \geq 3r+2$. L'étude suit globalement le même schéma que pour les codes r-identifiants évolutifs sans mémoire. Le lemme suivant permettra de simplifier certains raisonnements.

Lemme 4.10 *Soient C un code r-identifiant évolutif avec mémoire de $\mathcal{C}_n$ et c un élément de C. Il existe un autre élément de C à distance au plus $r+1$ de c.*

Preuve : Soient x et y les deux voisins de c. Comme C est un code r-identifiant évolutif avec mémoire, il existe un entier r_{xy} compris entre 0 et r qui sépare x de y. Il existe donc un sommet $c' \in C$ qui, sans perte de généralité, se situe dans $\mathcal{B}_{r_{xy}}(x)$ mais pas dans $\mathcal{B}_{r_{xy}}(y)$. Nécessairement, $c' \neq c$ car $d(x,c) = d(c,y) = 1$, et la distance entre c et c' est d'au plus $r+1$. □

Le lemme suivant est l'analogue pour les codes r-identifiants évolutifs avec mémoire du lemme 4.2.

Lemme 4.11 *Soit S un ensemble de $3r+2$ sommets consécutifs de $\mathcal{C}_n$. Tout code r-identifiant évolutif avec mémoire de $\mathcal{C}_n$ possède au moins deux éléments dans S.*

Preuve : Soit C un code r-identifiant évolutif avec mémoire de $\mathcal{C}_n$. Supposons par l'absurde que S ne contienne qu'un seul élément de C (S contient au moins un élément du code pour des raisons de domination). Notons cet élément c. D'après le lemme 4.10, il existe un élément c' à distance au plus $r+1$ de c. Par hypothèse, $c' \notin S$. Sans perte de généralité, nous notons $v_0, ...v_{3r+1}$ les éléments de S et posons $c = v_i$ avec $i < 2r$. Alors c' est nécessairement un des sommets v_j avec $-(r+1) \leq j \leq -1$ et donc $i \leq r$. Mais alors le sommet v_{2r+1} n'est pas r-dominé par un élément de C, menant à une contradiction. □

Le lemme 4.11 nous donne directement une borne inférieure sur la taille des codes r-identifiants évolutifs avec mémoire de $\mathcal{C}_n$:

Corollaire 4.12 *Soit C un code r-identifiant évolutif avec mémoire de $\mathcal{C}_n$. Alors*

$$|C| \geq \left\lceil \frac{2n}{3r+2} \right\rceil.$$

L'étude des codes r-identifiants évolutifs avec mémoire va dépendre de la valeur de n modulo $3r+2$. Aussi nous écrivons dans la suite $n = (3r+2)p+R$ avec $0 \leq R \leq 3r+1$ et $p \geq 1$ (par hypothèse, $n \geq 3r+2$). Le corollaire 4.12 peut alors se reformuler de la manière suivante. Soit C un code r-identifiant évolutif avec mémoire de $\mathcal{C}_n$, alors :

- Si $R = 0$, $|C| \geq 2p$;
- Si $0 < 2R \leq 3r+2$, $|C| \geq 2p+1$;
- Sinon, $2R > 3r+2$, et $|C| \geq 2p+2$.

Le prochain lemme donne une construction de code optimal.

Lemme 4.13 *Soit $n = (3r+2)p$, alors $\mathcal{C}_n$ possède un code r-identifiant évolutif avec mémoire de cardinalité $2p$.*

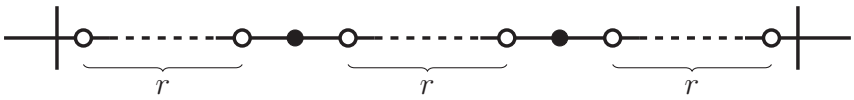


Figure 4.10 – *Le motif utilisé pour construction des codes r-identifiants évolutifs avec mémoire du cycle $\mathcal{C}_{(3r+2)p}$.*

Preuve : Nous considérons le code obtenu en répétant le motif représenté sur la figure 4.10. Plus formellement, posons

$$C = \{v_i | i \equiv r \mod 3r+2 \text{ ou } i \equiv 2r+1 \mod 3r+2\}.$$

L'ensemble C est un ensemble r-dominant de cardinalité $2p$ et nous devons vérifier que toute paire de sommets est séparée par C pour un rayon dans $[\![0, r]\!]$.

Il est suffisant de vérifier cela pour une paire de sommets située dans une même répétition du motif. Aussi, prenons v_i, v_j deux sommets tels que $0 \leq i < j \leq 3r+1$. Nous définissons r_{ij} de la manière suivante :

- Si $j \leq r$, $r_{ij} = r - j$;
- Si $i \leq r < j$, $r_{ij} = |(2r+1) - j|$;
- Si $r < i \leq 2r$, $r_{ij} = i - r$;
- Si $i \geq 2r+1$, $r_{ij} = i - (2r+1)$.

Nous avons bien dans tous les cas $0 \leq r_{ij} \leq r$ et la paire de sommets (v_i, v_j) est r_{ij}-séparée par C (dans les cas 1 et 3 par l'élément v_r et dans les cas 2 et 4 par l'élément v_{2r+1}). □

Cette construction peut être généralisée lorsque n n'est pas un multiple de $3r+2$ en ajoutant un ou deux sommets au code.

Lemme 4.14 *Si $1 \leq R \leq r+1$, alors $\mathcal{C}_n$ a un code r-identifiant évolutif avec mémoire de cardinalité $2p+1$. Si $R > r+1$, alors $\mathcal{C}_n$ a un code r-identifiant évolutif avec mémoire de cardinalité $2p+2$.*

Preuve : Considérons les trois cas suivants : (1) $R \in [\![1, r+1]\!]$, (2) $R \in [\![r+2, 2r+2]\!]$, et (3) $R \in [\![2r+3, 3r+1]\!]$. Pour chacun de ces cas, nous définissons le code C comme suit :

(1) $C = \{v_i | i < (3r+2)p, i \equiv r \mod 3r+2 \text{ ou } i \equiv 2r+1 \mod 3r+2\} \cup \{v_{(3r+2)p}\}$
(2) $C = \{v_i | i < (3r+2)p, i \equiv r \mod 3r+2 \text{ ou } i \equiv 2r+1 \mod 3r+2\} \cup \{v_{(3r+2)p}, v_{(3r+2)p+r}\}$
(3) $C = \{v_i | i < (3r+2)p, i \equiv r \mod 3r+2 \text{ ou } i \equiv 2r+1 \mod 3r+2\} \cup \{v_{(3r+2)p+r}, v_{(3r+2)p+2r}\}$

De la même manière que dans la preuve du lemme 4.13, nous montrons que ces ensembles sont des codes r-identifiants évolutifs avec mémoire de cardinalité respective $2p+1$, $2p+2$ et $2p+2$. □

Le lemme suivant permet de terminer l'étude en montrant que la borne inférieure du corollaire 4.12 n'est pas atteinte lorsque $R > r+1$. Les codes construits auparavant sont donc optimaux dans ce cas.

Lemme 4.15 *Si $R > r+1$, alors $\mathcal{C}_n$ n'a pas de code r-identifiant évolutif avec mémoire de cardinalité $2p+1$.*

Preuve : Supposons par l'absurde que $R > r+1$ et qu'il existe un code r-identifiant évolutif avec mémoire de $\mathcal{C}_n$ de cardinalité $2p+1$. Observons tout d'abord que dans dans un ensemble de R sommets consécutifs, il y a au plus un élément de C. Sinon il y aurait seulement $2p-1$ éléments du code dans le reste du cycle dont les sommets peuvent être séparés en p ensembles disjoints de taille $3r+2$. Un de ces ensembles n'aurait qu'un seul sommet du code, contredisant le lemme 4.11.
Choisissons maintenant un élément c du code. D'après le lemme 4.10, il y a un sommet c' du code à distance $d \leq r+1$ de c. Alors l'ensemble de sommets S comprenant tous les sommets situés entre c et c', c et c' inclus, est de taille au plus $r+2 \leq R$ et contient deux sommets de C, menant à une contradiction. □

Nous résumons les résultats obtenus dans le théorème suivant :

Théorème 4.16 *Soient r et b des entiers avec $n = (3r+2)p+R$, où $0 \leq R < 3r+2$ et $p \geq 1$. Nous avons :*

(i) Si $R = 0$, $\gamma_r^{IDDM}(\mathcal{C}_n) = 2p$;
(ii) Si $R \leq r+1$, $\gamma_r^{IDDM}(\mathcal{C}_n) = 2p+1$;
(iii) Sinon, $R > r+1$ et $\gamma_r^{IDDM}(\mathcal{C}_n) = 2p+2$.

Le théorème 4.16-*i* (*resp.* 4.16-*ii*, 4.16-*iii*) découle du corollaire 4.12 et du lemme 4.13 (*resp.* du corollaire 4.12 et du lemme 4.14, et des Lemmes 4.14 et 4.15).

Le prochain lemme complète l'étude pour les petites valeurs de n :

Lemme 4.17 *Soient r et n des entiers, avec $3 \leq n \leq 3r+1$. Alors $\gamma_r^{IDDM}(\mathcal{C}_n) = 2$.*

Preuve : Un code r-identifiant évolutif avec mémoire dans un cycle doit contenir au moins deux sommets, sinon les deux sommets adjacents au sommet du code ne sont pas séparé. Lorsque $n \leq 2r+2$, deux sommets adjacents forment un code r-identifiant évolutif avec mémoire. Pour $n > 2r+2$, deux sommets à distance $r+1$ forment un code r-identifiant évolutif avec mémoire. □

4.1.4 Limitons le nombre de rayons par sommet

Lorsque nous construisons un code r-identifiant évolutif avec mémoire, nous « donnons » à un sommet jusqu'à $r+1$ rayons pour le séparer des autres sommets. Dans le cas des cycles, nous pouvons en fait limiter ce nombre à 3.

Proposition 4.18 *Soit C un code r-identifiant évolutif avec mémoire de $\mathcal{C}_n$ et soit x un sommet du cycle. Supposons que $n > 2r+1$. Alors il existe un sous-ensemble R_x de $[\![0,r]\!]$ de taille au plus 3 tel que pour tout autre sommet y de $\mathcal{C}_n$, il existe un rayon r_{xy} dans R_x qui sépare x et y.*

Preuve : Sans perte de généralité, supposons que $x = v_0$.

Supposons tout d'abord qu'il y a deux sommets du code, $a = v_i$ et $b = v_j$, tels que $-r \leq i \leq 0 \leq j \leq r$ (si $x \in C$, alors nous avons $a = b = x$). Alors les deux rayons $d(x,a)$ et $d(x,b)$ permettent de séparer x de tous les autres sommets : les sommets v_k avec $0 < k < \frac{n}{2}$ sont séparés de x pour le rayon $d(x,a)$ et les autres pour le rayon $d(x,b)$. L'ensemble $R_x = \{d(x,a), d(x,b)\}$ convient donc.

Supposons maintenant qu'il n'y a des sommets du code que d'un seul côté du r-voisinage de x. Notons $a = v_i$ le sommet du code le plus proche de x. Sans perte de généralité, $0 < i \leq r$ et il n'y a pas de sommets du code parmi les sommets v_j avec $-r \leq j \leq 0$. D'après le lemme 4.10, il existe un autre élément du code $b = v_j$ tel que $i < j$ et $j - i \leq r+1$. Alors le sommet x est séparé de tous les sommets hors de $\mathcal{B}_i(a)$ pour le rayon i et de tous les sommets de $\mathcal{B}_{i-1}(a)$ pour le rayon $i-1$. Il reste alors un seul sommet qui n'est pas séparé de x : le sommet v_{2i} qui est séparé de x pour le rayon $d(v_{2i}, b) \leq r$. Finalement, l'ensemble $R_x = \{i-1, i, d(v_{2i}, b)\}$ est suffisant. □

Cette proposition nous amène naturellement à nous poser la question suivante : quelle est la cardinalité minimale d'un code r-identifiant évolutif avec mémoire sur $\mathcal{C}_n$ pour lequel deux rayons suffisent à séparer un sommet des autres ?

Pour répondre à cette question, nous appelons $(2, [\![0,r]\!])$-code d'un graphe G un sous-ensemble de sommets de G qui est r-dominant et tel que pour chaque sommet x, il existe deux rayons dans $[\![0,r]\!]$ qui séparent x de tous les autres sommets du graphe.

Lemme 4.19 *Soient $r \geq 1$, $k = \left\lfloor \frac{r+1}{3} \right\rfloor$ et $s = 3r - k + 2$. Si n est un multiple de s, alors $\mathcal{C}_n$ possède un $(2, [\![0,r]\!])$-code de cardinalité $\frac{n}{s}$.*

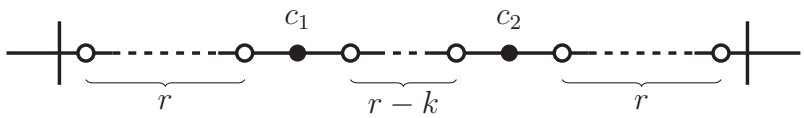


Figure 4.11 – *Le motif utilisé pour construire un $(2, [\![0,r]\!])$-code du cycle lorsque n est un multiple de s (voir le lemme 4.19).*

Preuve : Nous étudions le code construit en répétant le motif S de la figure 4.11. Notons c_1 et c_2 les deux sommets du code qui sont dans le motif S et supposons que $c_1 = v_0$. Alors $c_2 = v_{r-k+1}$ et les sommets de S sont les sommets situés entre v_{-r} et v_{2r-k+1}. Nous divisons alors les sommets de S en cinq sous-ensembles : $A_1 = \{v_{-r}, \ldots, v_{-k-1}\}$, $A_2 = \{v_{-k}, \ldots, v_{-1}\}$, $A_3 = \{v_0, \ldots, v_{r-k+1}\}$, $A_4 = \{v_{r-k+2}, \ldots, v_{r+1}\}$ et $A_5 = \{v_{r+2}, \ldots, v_{2r-k+1}\}$. Si $r = 1$, alors A_2 et A_4 sont vides.

Soit x un sommet de S. Suivant la position de x, nous définissons l'ensemble R_x :
- Si $x \in A_1$, $R_x = \{d(x, c_1), d(x, c_1) - 1\}$;
- Si $x \in A_2$, $R_x = \{d(x, c_1), d(x, c_2) - 1\}$;
- Si $x \in A_3$, $R_x = \{d(x, c_1), d(x, c_2)\}$;
- Si $x \in A_4$, $R_x = \{d(x, c_1) - 1, d(x, c_2)\}$;
- Si $x \in A_5$, $R_x = \{d(x, c_2), d(x, c_2) - 1\}$.

Nous avons alors dans tous les cas $R_x \subseteq [\![0, r]\!]$. Par symétrie, nous vérifions seulement qu'un sommet x de $A_1 \cup A_2 \cup A_3$ est séparé des autres sommets par un rayon dans R_x.

Si $x \in A_1$, alors x est séparé des sommets hors de $\mathcal{B}_{d(x,c_1)}(c_1)$ pour le rayon $d(x, c_1)$ et des sommets de $\mathcal{B}_{d(x,c_1)-1}(c_1)$ pour le rayon $d(x, c_1) - 1$. Il reste le sommet y à distance $d(x, c_1)$ de c_1. Notons $x = v_{-i}$ avec $k + 1 \leq i \leq d(x, c_1)$. Alors $y = v_i$ et $d(y, c_2) = r - k + 1 - i \leq r - 2k \leq k + 1 \leq d(x, c_1)$, par définition de k. Or, $d(x, c_2) > d(x, c_1)$ donc x et y sont séparés pour le rayon $d(x, c_1)$.

Si $x \in A_2$, alors x est séparé des sommets hors de $\mathcal{B}_{d(x,c_1)}(c_1)$ pour le rayon $d(x, c_1)$ et des sommets de $\mathcal{B}_{d(x,c_2)-1}(c_1)$ pour le rayon $d(x, c_2) - 1$. Cela couvre tous les sommets du cycle.

Avec des arguments similaires, $x \in A_3$ est aussi séparé de tous les sommets du cycle pour le rayon $d(x, c_1)$ ou $d(x, c_2)$. □

Lemme 4.20 *Soit C un $(2, [\![0, r]\!])$-code de $\mathcal{C}_n$. Soit S un ensemble de $s = 3r - k + 2$ sommets avec $k = \lfloor (r + 1)/3 \rfloor$. Alors S contient au moins deux sommets de C.*

Preuve : Pour $r = 1$, un $(2, [\![0, 1]\!])$-code est un code 1-identifiant évolutif avec mémoire, aussi le lemme est vrai et nous pouvons supposer que $r \geq 2$.

Remarquons que $3r - k + 2 > 2r$, donc S contient au moins un sommet de C. Supposons par l'absurde que S contient un seul sommet du code, c. Sans perte de généralité, $c = v_0$.

Soient v_{-a} le premier sommet de S et v_b le dernier. Nous avons $a + b = 3r - k + 1$, et nous pouvons supposer que $a \leq b$. Le code C est toujours un code r-identifiant évolutif avec mémoire donc le lemme 4.10 est valide et $a \leq r$, $b \geq 2r - k + 1$. Le code C est r-dominant, donc $b \leq 2r$ et $a \geq r - k + 1$. Les sommets v_{k-1}, v_k et v_{k+1} n'ont alors dans leur r-voisinage que le sommet c parmi les sommets du code. Ainsi, v_k et v_{k-1} sont uniquement séparés pour le rayon $k - 1$, et v_k et v_{k+1} uniquement pour le rayon k. Le sommet v_k est donc séparé de tous les sommets soit pour le rayon k, soit pour le rayon $k - 1$. Cela est en particulier vrai pour le sommet v_{-k}. Il y a donc un sommet du code $c' \notin S$ à distance au plus k de v_{-k}. Mais alors $d(c', v_{-k}) = d(c', v_{-a}) + d(v_{-a}, v_{-k}) \geq 1 + a - k \geq r - 2k + 2 \geq k + 1$ (par définition de k), menant à une contradiction. □

Ce lemme implique que le code du lemme 4.19 est optimal, et nous avons les bornes inférieures suivantes :

Corollaire 4.21 *Soit C un $(2, [\![0, r]\!])$-code de $\mathcal{C}_n$. Alors*

$$|C| \geq \left\lceil \frac{2n}{s} \right\rceil$$

avec $s = 3r - \lfloor (r+1)/3 \rfloor + 2$.

Il reste le cas où s ne divise pas n. En utilisant des arguments similaires à ceux utilisés pour les codes r-identifiants évolutifs avec mémoire et les codes r-identifiants évolutifs sans mémoire, nous obtenons le théorème suivant :

Théorème 4.22 *Soient n, r des entiers. Posons $k = \lfloor (r+1)/3 \rfloor$, $s = 3r - k + 2$ et $n = sp + R$, avec $0 \leq R < s$. La cardinalité d'un $(2, [\![0, r]\!])$-code optimal de $\mathcal{C}_n$ est :*

(i) $2p$ *si* $R = 0$;

(ii) $2p + 1$ *si* $R \leq r + 1$;

(iii) $2p + 2$ *sinon.*

4.2 Codes identifiants tolérants dans le plan euclidien

Nous étudions dans cette partie des codes identifiants qui sont tolérants pour des variations légères du voisinage des sommets. Pour cela nous nous plaçons dans le plan $\mathbb{Z}^2$ avec la distance euclidienne. Nous pouvons ainsi considérer des capteurs qui détectent jusqu'à une distance euclidienne r. D'un point de vue pratique, il est naturel de considérer que les capteurs ne sont pas parfaits et qu'aux abords de la distance limite de détection, il puisse y avoir des petites variations. En se plaçant dans le plan euclidien, nous pouvons considérer des rayons réels et donc des variations minimes. Ces travaux ont étés réalisés avec Ville Junnila et Tero Laihonen [63] lors de mon séjour à Turku en Finlande au printemps 2011.

4.2.1 Cadre : le plan euclidien

Dans toute cette partie, nous nous plaçons dans la grille infinie $\mathbb{Z}^2$. Au lieu d'utiliser la distance habituelle sur les graphes nous utilisons la distance *euclidienne* $d_{\mathcal{E}}$: pour deux sommets $u = (x_u, y_u)$ et $v = (x_v, y_v)$ de $\mathbb{Z}^2$:

$$d_{\mathcal{E}}(u, v) = \sqrt{(x - x_u)^2 + (y - y_u)^2}.$$

La boule de rayon r centrée $u \in \mathbb{Z}^2$, $\mathcal{B}_r(u)$, devient naturellement :

$$\mathcal{B}_r(u) = \{v \in \mathbb{Z}^2 | d_{\mathcal{E}}(u, v) \leq r\} = \{(x, y) \in \mathbb{Z}^2 | (x - x_u)^2 + (y - y_u)^2 \leq r^2\}.$$

Notons que les rayons pris ici sont réels mais que les boules considérées ne varient pas continument. En effet, si r et r' sont assez *proches*, c'est-à-dire, lorsque $\lfloor r^2 \rfloor = \lfloor r'^2 \rfloor$, alors les boules $\mathcal{B}_r(u)$ et $\mathcal{B}_{r'}(u)$ sont les mêmes et les codes considérés aussi. Ainsi dans cette partie nous ne préciserons pas toujours que les résultats sont vrais pour des rayons *proches*. La figure 4.12 montre les boules euclidiennes pour $r \leq 3$.

Le graphe $\mathbb{Z}^2$ où le voisinage d'un sommet est la boule euclidienne de rayon r centrée sur ce sommet est appelé *plan euclidien de rayon* r. Les codes r-identifiants dans le plan euclidien sont alors définis de manière usuelle. Comme dans la partie 3.4, nous étudions la densité minimale de ces codes. Nous noterons $D(r)$ la densité optimale d'un code identifiant dans le plan euclidien de rayon r.

Le plan euclidien de rayon 1 correspond à la grille carrée classique. La densité optimale d'un code identifiant pour ce graphe est de $\frac{7}{20}$ (Charon *et al.* [28] pour la borne supérieure et Ben-Haim *et al.* [6] pour la borne inférieure). Le plan euclidien de rayon 2 correspond à la grille carrée de rayon 2. Les plans euclidiens de rayon $\sqrt{2}$ et $2\sqrt{2}$ correspondent à la grille du roi de rayon 1 et 2 que nous avons étudiée dans la partie 3.4. Les autres rayons ne correspondent pas à des grilles « classiques ».

Junnila et Laihonen [61] ont étudié les codes identifiants dans le plan euclidien pour des rayons quelconques. Ils montrent le résultat suivant :

Théorème 4.23 (Junnila, Laihonen [61]) *La densité minimale d'un code identifiant dans le plan euclidien de rayon r vérifie :*

$$D(r) \geq \frac{3}{4\lfloor r \rfloor + 4\lfloor b \rfloor + 4\lfloor \sqrt{r^2 - (\lfloor b \rfloor + 1)^2} \rfloor + 8} \geq \frac{3}{4(\sqrt{2} + 1)r + 12}$$

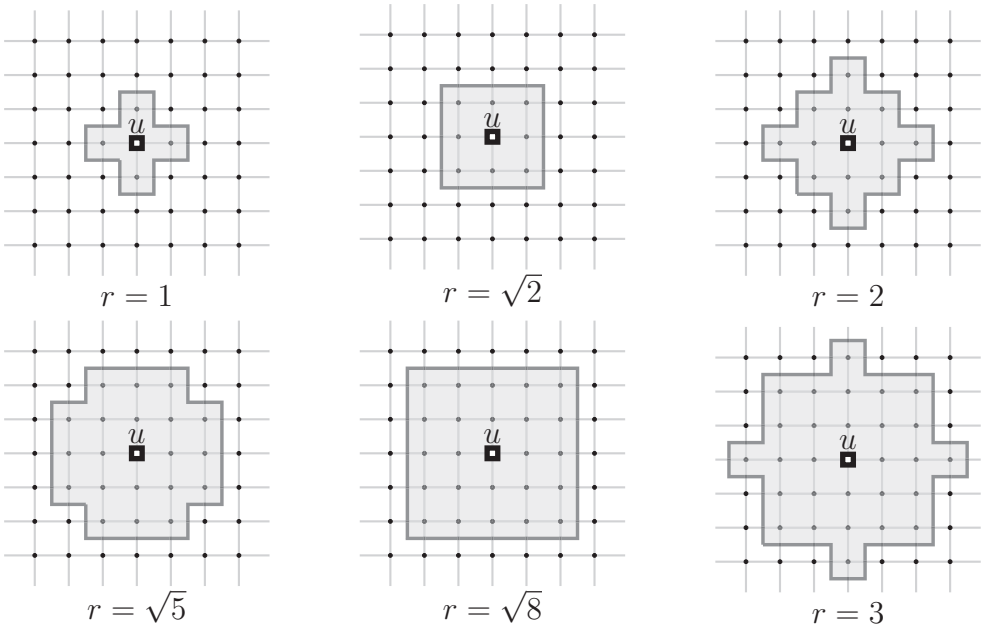


Figure 4.12 – *Les boules euclidiennes $\mathcal{B}_r(u)$ pour $r \leq 3$.*

avec $b = \frac{\sqrt{2r^2-1}-1}{2}$.

La figure 4.13 donne un exemple issu de [61] de code identifiant, optimal, dans le plan euclidien de rayon $\sqrt{5}$. Nous pouvons remarquer que ce code n'est plus identifiant si le rayon de la boule diminue faiblement. L'objet de cette partie est de construire des codes identifiants du plan euclidien qui restent identifiant lorsque les rayons des boules varient. Le fait de se placer dans le plan euclidien permet de considérer des variations en fonction du rayon r des boules, rayon qui n'est pas obligatoirement entier et permet des variations infinitésimales.

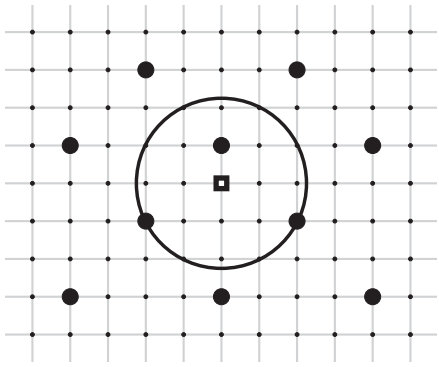


Figure 4.13 – *Un code identifiant optimal pour le plan euclidien de rayon $\sqrt{5}$.*

4.2.2 Codes identifiants tolérants

Soient un *rayon* r (pas nécessairement entier) et une *variation* Δ, réel strictement positif. Dans la suite, nous considérons que le voisinage d'un élément c du code est

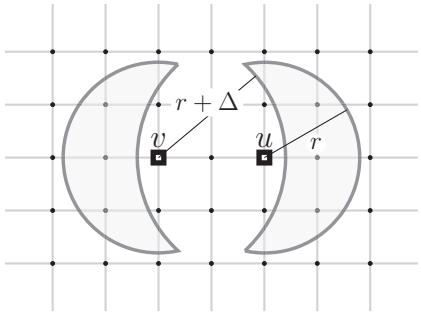


Figure 4.14 – *L'ensemble $\mathcal{S}_{r,\Delta}(u,v)$ qui permet de séparer les sommets u et v.*

un ensemble contenant $\mathcal{B}_r(c)$ et inclus dans $\mathcal{B}_{r+\Delta}(c)$ où Δ représente la tolérance sur le rayon. Cet ensemble ne correspond pas nécessairement à un ensemble $\mathcal{B}_{r'}(c)$ avec r' compris entre r et $r+\Delta$.

Pour construire un code identifiant C qui permette une tolérance de valeur Δ sur le rayon r, il faut tout d'abord que les éléments du code dominent tous les sommets de $\mathbb{Z}^2$. Comme le voisinage d'un élément c du code contient $\mathcal{B}_r(c)$, la domination est assurée en choisissant un ensemble r-dominant. Ensuite, pour que c sépare deux sommets u et v, il faut par exemple que c domine u mais pas v. Pour assurer que c domine u, il faut que $u \in \mathcal{B}_r(c)$, ou de manière équivalente, que $c \in \mathcal{B}_r(u)$. Pour assurer que c ne domine pas v, il faut que $v \notin \mathcal{B}_{r+\Delta}(c)$, ou encore que $c \notin \mathcal{B}_{r+\Delta}(v)$. Ainsi, si $c \in \mathcal{B}_r(u) \setminus \mathcal{B}_{r+\Delta}(v)$, c sépare u et v. De la même manière, si $c \in \mathcal{B}_r(v) \setminus \mathcal{B}_{r+\Delta}(u)$, c sépare encore u et v.

Pour deux sommets u et v de $\mathbb{Z}^2$, nous posons donc :

$$\mathcal{S}_{r,\Delta}(u,v) := (\mathcal{B}_r(u) \setminus \mathcal{B}_{r+\Delta}(v)) \cup (\mathcal{B}_r(v) \setminus \mathcal{B}_{r+\Delta}(u)).$$

La figure 4.14 donne un exemple d'ensemble $\mathcal{S}_{r,\Delta}(u,v)$. Un *code (r,Δ)-identifiant tolérant* est un sous-ensemble r-dominant de sommets S du plan euclidien $\mathbb{Z}^2$ tel que pour toute paire de sommets distincts u et v,

$$\mathcal{S}_{r,\Delta}(u,v) \cap C \neq \emptyset.$$

Comme pour les codes identifiants, nous cherchons la densité optimale $D(r,\Delta)$ d'un code (r,Δ)-identifiant tolérant. Un code $(r,0)$-identifiant tolérant est simplement un code identifiant dans le plan euclidien de rayon r, donc $D(r,0) = D(r)$. De plus, si $\Delta \leq \Delta'$, tout code (r,Δ')-identifiant tolérant est un code (r,Δ)-identifiant tolérant : $D(r,\Delta) \leq D(r,\Delta')$.

Dans la suite, nous utiliserons simplement le terme *code (r,Δ)-identifiant* pour se référer à ces codes. Trouver un code (r,Δ)-identifiant correspond à résoudre un problème de couverture de sous-ensembles. Aussi nous utiliserons la proposition 3.43 montrée au chapitre 3 dont nous rappelons l'énoncé ici :

Proposition 4.24 *Soient C un code de $\mathbb{Z}^2$ et X un ensemble fini non vide de $\mathbb{Z}^2$. Alors :*

$$D(C) = \frac{n(X,C)}{|X|}$$

où $n(X,C)$ désigne le nombre moyen de sommets de C dans une translation de l'ensemble X.

4.2.3 Premier exemple

Afin d'illustrer la notion de code (r, Δ)-identifiant tolérant, nous traitons dans cette sous-partie un exemple. Pour que la définition ait un sens et soit différente de celle des codes identifiants classiques, il faut que les boules de rayon r et $r+\Delta$ soient distinctes, et que les ensembles $\mathcal{S}_{r,\Delta}(u, v)$ soient non vides. Les premières valeurs que nous pouvons étudier sont donc $r = 1$ et $r + \Delta = \sqrt{2}$.

Un code $(1, \sqrt{2} - 1)$-identifiant doit couvrir tous les ensembles $\mathcal{S}_{1,\sqrt{2}-1}(u, v)$. La figure 4.15 montre ces ensembles pour $u - v = (1, 0)$, $u - v = (0, 1)$, $u - v = (1, -1)$ et $u - v = (1, 1)$. Notons que ces ensembles sont invariants par translation et symétriques en fonction de u et v. Ce qui nous importe est donc uniquement la différence vectorielle « absolue » entre u et v.

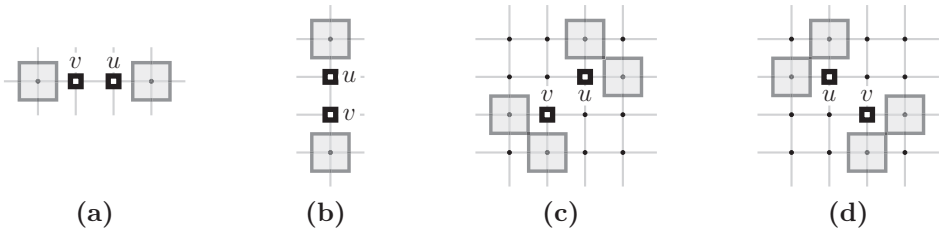


Figure 4.15 – *Les ensembles* $\mathcal{S}_{1,\sqrt{2}-1}(u, v)$ *pour (a)* $u - v = (1, 0)$*, (b)* $u - v = (0, 1)$*, (c)* $u - v = (1, 1)$ *et (d)* $u - v = (1, -1)$*.*

L'ensemble $\mathcal{S}_{1,\sqrt{2}-1}((0, 0), (-1, 0))$ contient seulement deux éléments et doit être couvert pour toute translation par un code $(1, \sqrt{2} - 1)$-identifiant. D'après la proposition 4.24, nous avons donc :

$$D(1, \sqrt{2} - 1) \geq \frac{1}{2}.$$

Par ailleurs, la figure 4.16 donne un exemple de code 1-dominant de densité $\frac{1}{2}$ couvrant tous les ensembles $\mathcal{S}_{1,\sqrt{2}-1}(u, v)$ pour u et v tels que $d_{\mathcal{E}}(u, v) \leq \sqrt{2}$. Nous pouvons vérifier que ce code est $(1, \sqrt{2} - 1)$-identifiant en remarquant que tout ensemble $\mathcal{S}_{1,\sqrt{2}-1}(u, v)$ avec $d_{\mathcal{E}}(u, v) > \sqrt{2}$ contient trois sommets consécutifs en forme de L : $\begin{smallmatrix} \bullet & \\ \bullet & \bullet \end{smallmatrix}$. Comme ce code intersecte toutes les rotations et translations de ce motif, il intersecte tous les ensembles $\mathcal{S}_{1,\sqrt{2}-1}(u, v)$. Nous avons donc :

$$D(1, \sqrt{2} - 1) = \frac{1}{2}.$$

4.2.4 Existence d'un code identifiant tolérant

Nous ne pouvons pas construire de codes identifiants tolérants pour n'importe quelles valeurs de r et de Δ. Par exemple, si $\Delta \geq 1$, nous avons $\mathcal{B}_r((0, 0)) \subseteq \mathcal{B}_{r+\Delta}((-1, 0))$ et donc $\mathcal{S}_{r,\Delta}((0, 0), (-1, 0)) = \emptyset$. Il n'existe alors pas de code (r, Δ)-identifiant. Avoir $\mathcal{S}_{r,\Delta}((0, 0), (-1, 0)) \neq \emptyset$ est en fait une condition nécessaire et suffisante pour qu'un code (r, Δ)-identifiant existe :

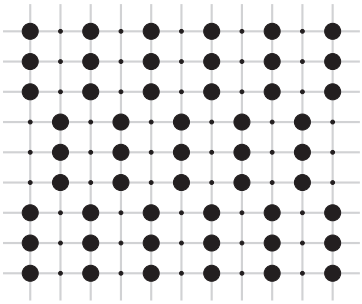


Figure 4.16 – *Un code* $(1, \sqrt{2}-1)$*-identifiant de densité optimale* $\frac{1}{2}$.

Proposition 4.25 *Soient* r *et* Δ *des réels positifs ou nuls. Il existe un code* (r, Δ)*-identifiant si et seulement si* $\mathcal{S}_{r,\Delta}((0,0),(-1,0)) \neq \emptyset$.

Preuve : S'il existe un code (r, Δ)-identifiant, par définition, $\mathcal{S}_{r,\Delta}((0,0),(-1,0))$ contient au moins un élément.

Supposons que $\mathcal{S}_{r,\Delta}((0,0),(-1,0)) \neq \emptyset$. Par symétrie, $\mathcal{S}_{r,\Delta}((0,0),(0,-1)) \neq \emptyset$ et par translation tous les ensembles $\mathcal{S}_{r,\Delta}(u,v)$ pour $d_{\mathcal{E}}(u,v) = 1$ sont non vides. Soient u et v deux sommets de $\mathbb{Z}^2$ avec $u - v = (x, y)$. Par symétrie, supposons que $x > 0$ et $y \geq 0$. Comme $\mathcal{S}_{r,\Delta}((0,0),(-1,0)) \neq \emptyset$, il contient un sommet (x', y') avec $x' > 0$ et $y' \geq 0$. Alors $u + (x', y') \in \mathcal{S}_{r,\Delta}(u,v)$. Ainsi tous les ensembles $\mathcal{S}_{r,\Delta}(u,v)$ sont non vides et l'ensemble de tous les sommets de $\mathbb{Z}^2$ est un code (r, Δ)-identifiant. □

Notons que s'il existe un code (r, Δ)-identifiant pour un certain couple (r, Δ), alors il y aura un code (r, Δ')-identifiant pour tout Δ' tel que $0 \leq \Delta' \leq \Delta$ (le même code reste identifiant lorsque la variation Δ diminue). Il est donc naturel de définir la valeur limite de Δ pour laquelle il existe un code (r, Δ)-identifiant, que l'on note $\Delta_m(r)$:

$$\Delta_m(r) := \sup\{\Delta | \text{ il existe un code } (r, \Delta)\text{-identifiant}\}.$$

D'après les remarques précédentes, $\Delta_m(r) \leq 1$ et il existe un code (r, Δ)-identifiant si et seulement si $\Delta < \Delta_m(r)$. De plus, $\Delta_m(r) = 1$ si et seulement si r est un entier (ou est r est suffisamment *proche* d'un entier, au sens défini précédemment). Lorsque r n'est pas entier, la valeur de $\Delta_m(r)$ varie, mais tend vers 1 lorsque $r \to +\infty$:

Proposition 4.26 *Lorsque* $r \to +\infty$ *:*

$$\Delta_m(r) \geq 1 - \sqrt{\frac{2}{r}} + O\left(\frac{1}{r}\right).$$

Preuve : Considérons la plus grande abscisse $\lfloor r \rfloor$ d'un sommet de $\mathcal{B}_r((0,0))$, et u le sommet de $\mathcal{B}_r((0,0))$ qui a la plus grande ordonnée dans la colonne $\lfloor r \rfloor$. Les coordonnées de u sont alors $(\lfloor r \rfloor, \lfloor \sqrt{r^2 - \lfloor r \rfloor} \rfloor)$. Posons $\alpha = r - \lfloor r \rfloor$, alors $u = (r - \alpha, \lfloor \sqrt{\alpha(2r - \alpha)} \rfloor)$. Soit $\Delta_0 = d_{\mathcal{E}}((-1,0), u) - r$. Nécessairement, $\Delta_0 > 0$ et pour $\Delta < \Delta_0$, $u \in \mathcal{S}_{r,\Delta}((0,0),(-1,0))$. Donc l'ensemble $\mathcal{S}_{r,\Delta}((0,0),(-1,0))$ est non vide et d'après la proposition 4.25, il existe un code (r, Δ_0)-identifiant, ce qui revient à dire que $\Delta_m(r) \geq \Delta_0$.

Les calculs qui suivent donnent une borne inférieure sur la valeur de Δ_0, utilisant le fait que $\alpha \in [0, 1[$.

$$\begin{aligned} r + \Delta_0 &= \left((r-\alpha+1)^2 + (\lfloor\sqrt{\alpha(2r-\alpha)}\rfloor)^2\right)^{\frac{1}{2}} \\ &\geq \left((r-\alpha+1)^2 + (\sqrt{\alpha(2r-\alpha)}-1)^2\right)^{\frac{1}{2}} \\ &\geq (r^2+2r-2\sqrt{\alpha(2r-\alpha)}+2-2\alpha)^{\frac{1}{2}} \\ &\geq (r^2+2r-2\sqrt{2r})^{\frac{1}{2}} \\ &= r\left(1+\frac{2}{r}-\frac{2\sqrt{2}}{r^{3/2}}\right)^{\frac{1}{2}} \end{aligned}$$

Lorsque $r \to +\infty$, en utilisant les séries de Taylor, nous obtenons :

$$r + \Delta_0 \geq r\left(1+\frac{1}{r}-\frac{\sqrt{2}}{r\sqrt{r}}+O\left(\frac{1}{r^2}\right)\right).$$

Et donc :

$$\Delta_m(r) \geq \Delta_0 \geq 1-\sqrt{\frac{2}{r}}+O\left(\frac{1}{r}\right).$$

□

Dans ce qui suit nous considérons implicitement des valeurs de r et Δ pour lesquelles il existe un code (r, Δ)-identifiant, de telles valeurs sont dites *admissibles*.

4.2.5 Étude du motif horizontal

Les sommets les plus difficiles à séparer sont les sommets qui sont à distance 1, dans le sens où l'ensemble $S_{r,\Delta}(u, v)$ contient moins d'éléments lorsque les sommets u et v sont à distance 1 que lorsqu'ils sont plus éloignés. Pour étudier les codes (r, Δ)-identifiants nous avons donc besoin d'une bonne compréhension des ensembles $S_{r,\Delta}(u, v)$ avec $d_{\mathcal{E}}(u, v) = 1$. Par symétrie, nous avons seulement besoin d'étudier l'ensemble $\mathcal{S}_{r,\Delta}((0, 0), (-1, 0))$ que nous appelons *motif horizontal*. Nous appelons de la même manière *motif vertical* l'ensemble $\mathcal{S}_{r,\Delta}((0, 0), (0, -1))$ et *motif diagonal* l'ensemble $\mathcal{S}_{r,\Delta}((0, 0), (-1, -1))$.

Dans ce qui suit, nous utiliserons les notations suivantes :

- $C_r(u)$, où u est un sommet de $\mathbb{Z}^2$ et r est un réel positif, désigne le cercle (géométrique) de rayon r centré en u.

- $h_{r,\Delta}(x) = \sqrt{r^2 - x^2} - \sqrt{(r + \Delta)^2 - (x + 1)^2}$, défini pour un entier $x \in [0, \lfloor r + \Delta - 1\rfloor]$, est la distance verticale signée à l'abscisse x entre les cercles $C_r((0, 0))$ et $C_{r+\Delta}((-1, 0))$. Lorsque $h_{r,\Delta}(x) \leq 0$, il ne peut y avoir de sommet d'abscisse

x dans $\mathcal{S}_{r,\Delta}((0,0),(-1,0))$. Quand $h_{r,\Delta}(x) > 1$, il y a toujours un sommet d'abscisse x dans $\mathcal{S}_{r,\Delta}((0,0),(-1,0))$. Enfin, si $h_{r,\Delta}(x) \in]0,1]$, il y a au plus un sommet d'abscisse x dans $\mathcal{S}_{r,\Delta}((0,0),(-1,0))$.

– $x_0(r,\Delta)$ (ou simplement x_0 lorsque le contexte est clair) est la plus petite abscisse positive ou nulle d'un élément de $\mathcal{S}_{r,\Delta}((0,0),(-1,0))$. C'est au moins la partie entière supérieure de la solution positive de $h_{r,\Delta}(x) = 0$. Nous avons donc :

$$x_0(r,\Delta) \geq \left\lceil \frac{\Delta(2r+\Delta)-1}{2} \right\rceil = r\Delta + O(1) \text{ lorsque } r \to +\infty. \tag{4.1}$$

– $x_1(r,\Delta)$ (ou simplement x_1 lorsque le contexte est clair) est la partie entière de la solution positive de $h_{r,\Delta}(x) = 1$. La valeur exacte de x_1 est :

$$x_1(r,\Delta) = \left\lfloor \frac{1}{4}\Big(-2+\Delta^2+2r\Delta+ \sqrt{-4+4\Delta^2-\Delta^4+8r\Delta-4r\Delta^3+8r^2-4r^2\Delta^2}\Big)\right\rfloor. \tag{4.2}$$

Lorsque $r \to +\infty$:

$$x_1(r,\Delta) = r\left(\frac{\Delta}{2} + \frac{\sqrt{2-\Delta^2}}{2}\right) + O(1). \tag{4.3}$$

Nous aurons aussi besoin de la valeur de x_1 lorsque Δ est proche de 1. Supposons que $\Delta = 1 - \epsilon(r)$ avec $\epsilon(r)$ de l'ordre de $\frac{1}{r^\alpha}$ lorsque $r \to +\infty$. D'après la proposition 4.26, nous avons toujours un code (r,Δ)-identifiant lorsque $\epsilon(r) \geq \sqrt{\frac{2}{r}} + O(\frac{1}{r})$. Aussi nous supposons que $0 < \alpha \leq \frac{1}{2}$. Alors, lorsque $r \to +\infty$:

$$x_1(r,\Delta) = r - \frac{r\epsilon(r)^2}{2} + O(r\epsilon(r)^3) + O(1). \tag{4.4}$$

Dans le cas spécial où $\alpha = \frac{1}{2}$, nous obtenons que la quantité $r - x_1$ est bornée.

– $m(r,\Delta)$ (ou simplement m lorsque le contexte est clair) est la partie entière supérieure de la valeur de $h_{r,\Delta}$ à l'abscisse $\lfloor r+\Delta-1 \rfloor$. La valeur exacte de m est :

$$m(r,\Delta) = \left\lceil \sqrt{r^2 - \lfloor r+\Delta-1\rfloor^2} - \sqrt{(r+\Delta)^2 - \lfloor r+\Delta\rfloor^2} \right\rceil. \tag{4.5}$$

Lorsque $r \to +\infty$:

$$m(r,\Delta) \leq 2\sqrt{r} + O(1). \tag{4.6}$$

Comme $h_{r,\Delta}$ est une fonction strictement croissante sur $[0, \lfloor r+\Delta-1 \rfloor]$, $m(r,\Delta)$ est une borne supérieure de $h_{r,\Delta}$.

Dans $\mathcal{S}_{r,\Delta}((0,0),(-1,0))$, pour un entier $x \in [0, \lfloor r+\Delta-1 \rfloor]$, il y a :
– pas de sommets d'abscisse x si $x < x_0$,
– au plus un sommet d'abscisse x si $x_0 \leq x \leq x_1$,
– au moins un sommet et au plus m sommets d'abscisse x si $x > x_1$.

Remarquons que dans la plupart des cas, $x_0 \geq 0$. En fait, $x_0 = 0$ seulement dans le cas où $\Delta = 0$ et r est entier.

Nous finissons cette partie par un résultat qui sera utilisé dans la suite et qui montre l'importance du motif horizontal :

Lemme 4.27 *Soient u et v deux sommets de $\mathbb{Z}^2$ situés sur la même ligne horizontale ou verticale. Si $d_{\mathcal{E}}(u,v) \leq 4x_0(r,\Delta)+1$, alors il existe deux sommets u' et v' à distance 1 tels que $\mathcal{S}_{r,\Delta}(u',v') \subseteq \mathcal{S}_{r,\Delta}(u,v)$.*

Preuve : Supposons tout d'abord que $d_{\mathcal{E}}(u,v)$ est impaire : $d_{\mathcal{E}}(u,v) = 2k+1$ avec $k \leq 2x_0$. Sans perte de généralité, nous pouvons supposer que $u = (k,0)$ et $v = (-k-1,0)$. Soient $u' = (0,0)$ et $v' = (-1,0)$. Montrons que $\mathcal{S}_{r,\Delta}(u',v') \subseteq \mathcal{S}_{r,\Delta}(u,v)$. Soit $w = (x,y) \in \mathcal{S}_{r,\Delta}(u',v')$. Par symétrie, nous pouvons supposer que $w \in \mathcal{B}_r(u') \setminus \mathcal{B}_{r+\Delta}(v')$ avec $x \geq 0$ et $y \geq 0$. Nous avons alors :

$$d_{\mathcal{E}}(w,v)^2 = (x+k+1)^2 + y^2 \geq (x+1)^2 + y^2 = d_{\mathcal{E}}(w,v')^2.$$

Donc $d_{\mathcal{E}}(w,v) \geq d_{\mathcal{E}}(w,v') > r+\Delta$ et $w \notin \mathcal{B}_{r+\Delta}(v)$. Par ailleurs, si $x \geq k$, alors $d_{\mathcal{E}}(w,u) \leq d_{\mathcal{E}}(w,u') \leq r$. Sinon, nous avons de toute façon, $x \geq x_0$ et comme $k \leq 2x_0$, nous avons $0 \leq k-x \leq x$. Alors nous avons aussi $d_{\mathcal{E}}(w,u) \leq d_{\mathcal{E}}(w,u') \leq r$. Dans tous les cas $w \in \mathcal{B}_r(u)$ et donc $w \in \mathcal{B}_r(u) \setminus \mathcal{B}_{r+\Delta}(v) \subseteq \mathcal{S}_{r,\Delta}(u,v)$. Le cas où la distance $d_{\mathcal{E}}(u,v)$ est paire se montre de la même manière. □

4.2.6 Bornes générales

Dans cette partie, nous donnons des bornes inférieures et supérieures sur la densité optimale $D(r,\Delta)$ d'un code (r,Δ)-identifiant. Nous considérons des valeurs générales de r et Δ ainsi que la tendance asymptotique de $D(r,\Delta)$ lorsque $r \to \infty$. Pour ce dernier point, nous distinguons deux cas : soit Δ reste fixe, soit Δ varie avec r en restant très proche de sa valeur maximale $\Delta_m(r)$.

Borne inférieure

La borne inférieure pour $\Delta = 0$ (correspondant aux codes identifiants classiques) issue de [61] et rappelée au théorème 4.23, est toujours valable pour une valeur générale de Δ :

$$D(r,\Delta) \geq D(r,0) \geq \frac{3}{4(\sqrt{2}+1)r+12}. \tag{4.7}$$

Cette borne a été obtenue en considérant l'union des différences symétriques pour toutes les paires possibles de sommets parmi quatre sommets en carré et en utilisant une variante pondérée de la proposition 4.24.

Nous n'avons pas pu améliorer cette borne lorsque Δ est fixé et que r est grand. Par contre, lorsque Δ approche 1 alors que r tend vers l'infini, nous obtenons une meilleur borne. Pour cela nous utilisons la proposition suivante qui est une conséquence directe de la proposition 4.24 :

Proposition 4.28 *Pour tout r et Δ admissibles, nous avons :*

$$D(r,\Delta) \geq \frac{1}{|\mathcal{S}_{r,\Delta}((0,0),(-1,0))|}.$$

Pour trouver une borne inférieure sur $D(r,\Delta)$, il faut donc majorer la quantité $|\mathcal{S}_{r,\Delta}((0,0),(-1,0))|$. Pour compter le nombre d'éléments contenu dans l'ensemble $\mathcal{S}_{r,\Delta}((0,0),(-1,0))$, nous comptons le nombre d'éléments d'abscisse $x \geq x_0$ et nous multiplions le résultat par 2. A l'abscisse entière $x \leq r$, il y a $\lfloor\sqrt{r^2-x^2}\rfloor$ éléments dans $\mathcal{B}_r((0,0))$ d'ordonnée positive, et parmi eux, si $x+1 \leq r+\Delta$, $\lfloor\sqrt{(r+\Delta)^2-(x+1)^2}\rfloor$ éléments de $\mathcal{B}_{r+\Delta}((-1,0))$. Suivant si pour la dernière abscisse $\lfloor r\rfloor$ il y a des éléments de $\mathcal{B}_{r+\Delta}((-1,0))$ ou non, nous obtenons la formule exacte suivante pour $|\mathcal{S}_{r,\Delta}((0,0),(-1,0))|$:

$$|\mathcal{S}_{r,\Delta}((0,0),(-1,0))| = 4\sum_{x=x_0(r,\Delta)}^{\lfloor r+\Delta-1\rfloor}\left(\lfloor\sqrt{r^2-x^2}\rfloor - \lfloor\sqrt{(r+\Delta)^2-(x+1)^2}\rfloor\right) + \delta\cdot(4\lfloor\sqrt{r^2-\lfloor r\rfloor^2}\rfloor+2)$$

où $\delta = 1$ si $\lfloor r\rfloor = \lfloor r+\Delta\rfloor$ et 0 sinon.

Pour avoir une borne inférieure concrète, nous utilisons les notations définies précédemment en remarquant que $\lfloor\sqrt{r^2-x^2}\rfloor - \lfloor\sqrt{(r+\Delta)^2-(x+1)^2}\rfloor$ est compris entre $\lfloor h_{r,\Delta}(x)\rfloor$ et $\lceil h_{r,\Delta}(x)\rceil$:

$$\begin{aligned}|\mathcal{S}_{r,\Delta}((0,0),(-1,0))| \quad \leq \quad & 4((x_1-x_0+1)+m\cdot(\lfloor r+\Delta-1\rfloor - x_1)) + \\ & \delta\cdot(4\lfloor\sqrt{r^2-\lfloor r\rfloor^2}\rfloor+2).\end{aligned}$$

Supposons maintenant que $\Delta = 1-\epsilon(r)$ avec $\epsilon(r)$ de l'ordre de $\frac{1}{r^\alpha}$ lorsque $r\to\infty$, avec $\alpha \leq \frac{1}{2}$. En utilisant les équations 4.4 et 4.6, nous pouvons montrer que l'ordre de $|\mathcal{S}_{r,\Delta}((0,0),(-1,0))|$ est au plus $r\sqrt{r}\epsilon(r)^2$ et donc que la densité de $D(r,\Delta)$ est d'ordre au moins $\frac{1}{r\sqrt{r}\epsilon(r)^2}$.

Dans le cas particulier où $\alpha = \frac{1}{2}$, c'est-à-dire que Δ est très proche de Δ_m, nous obtenons que $D(r,\Delta)$ est d'ordre au moins $\frac{1}{\sqrt{r}}$ ce qui améliore l'ordre de la borne donnée dans le cas général.

Borne supérieure

Pour obtenir une borne supérieure générale, nous construisons des codes en utilisant les ensembles de base suivants :

– lignes verticales modulo k : $L_k^v = \{(x,y) \in \mathbb{Z}^2 |\ x \equiv 0 \bmod k\}$,
– lignes horizontales modulo k : $L_k^h = \{(x,y) \in \mathbb{Z}^2 |\ y \equiv 0 \bmod k\}$.

Si nous pouvons garantir que tout ensemble $\mathcal{S}_{r,\Delta}(u,v)$ intersecte k lignes verticales ou horizontales consécutives, alors l'union de L_k^v et L_k^h sera un code (r,Δ)-identifiant. C'est la méthode utilisée pour montrer le résultat suivant :

Proposition 4.29 *Soit $k = \lfloor r \rfloor - x_1(r,\Delta)$. Alors l'ensemble $C = L_k^v \cup L_k^h$ est un code (r,Δ)-identifiant, et donc :*

$$D(r,\Delta) \leq \frac{2}{k} - \frac{1}{k^2}.$$

Preuve : L'ensemble C est un ensemble r-dominant car il y a plus de k sommets d'abscisses consécutives dans $\mathcal{B}_r((0,0))$. Pour montrer que C est un code (r,Δ)-identifiant, nous considérons les sommets de $\mathcal{S}_{r,\Delta}((0,0),(-1,0))$ quadrants par quadrants.

Soient X_1 les sommets de $\mathcal{S}_{r,\Delta}((0,0),(-1,0))$ situés dans le premier quadrant :

$$X_1 = \mathcal{S}_{r,\Delta}((0,0),(-1,0)) \cap \{(x,y) \in \mathbb{Z}^2 |\ x \geq 0, y \geq 0\}$$

et X_2 ceux situés dans le deuxième quadrant :

$$X_2 = \mathcal{S}_{r,\Delta}((0,0),(-1,0)) \cap \{(x,y) \in \mathbb{Z}^2 |\ x \geq 0, y \leq 0\}.$$

Pour le motif vertical, nous définissons de la même manière,

$$X_3 = \mathcal{S}_{r,\Delta}((0,0),(0,-1)) \cap \{(x,y) \in \mathbb{Z}^2 |\ x \geq 0, y \geq 0\}$$

et

$$X_4 = \mathcal{S}_{r,\Delta}((0,0),(0,-1)) \cap \{(x,y) \in \mathbb{Z}^2 |\ x \leq 0, y \geq 0\}.$$

Pour toute paire de sommets u et v de $\mathbb{Z}^2$, avec $u \neq v$, il existe un indice $i \in \{1,2,3,4\}$ tel que $u + X_i \subseteq \mathcal{S}_{r,\Delta}(u,v)$ ou bien $v + X_i \subseteq \mathcal{S}_{r,\Delta}(u,v)$. Pour montrer que C est (r,Δ)-identifiant, il suffit donc de montrer que tous les ensembles X_i contiennent k sommets d'abscisses ou d'ordonnées consécutives, ce qui découle de la définition de $x_1(r,\Delta)$ et de k.

Ainsi C est un code (r,Δ)-identifiant dont la densité est $\frac{2}{k} - \frac{1}{k^2}$ (la partie $\frac{1}{k^2}$ vient du fait que certains éléments sont comptés deux fois). □

Lorsque r tend vers l'infini, cela mène à la borne supérieure suivante :

$$D(r,\Delta) \leq \frac{4}{r(2-\Delta-\sqrt{2-\Delta^2})+K} \text{ où } K \text{ est une constante .} \tag{4.8}$$

En combinant (4.7) et (4.8), nous obtenons que la densité optimale d'un code (r,Δ)-identifiant, pour Δ fixé, est d'ordre $\frac{1}{r}$ lorsque $r \to +\infty$.

Dans le cas où Δ varie avec r en étant proche de Δ_m, c'est-à-dire lorsque $\Delta = 1 - \epsilon(r)$, avec $\epsilon(r)$ d'ordre $\frac{1}{r^\alpha}$ lorsque $r \to \infty$, et $\alpha \leq \frac{1}{2}$, nous obtenons que $D(r,\Delta)$

est d'ordre au plus $\frac{1}{r\epsilon(r)^2}$. Dans le cas particulier où Δ est très proche de Δ_m, c'est-à-dire, lorsque $\alpha = \frac{1}{2}$, le résultat obtenu est trivial. Comme nous le verrons dans la suite, il existe des familles de valeurs (r, Δ), avec r non borné, pour lesquelles la densité optimale est $\frac{1}{2}$, aussi nous ne pouvions pas espérer obtenir une borne générale d'ordre meilleur qu'une constante.

4.2.7 Amélioration des bornes pour des valeurs fixées

Construction générale

La construction de la proposition 4.29 ne prend pas en compte les fortes symétries que contient l'ensemble $\mathcal{S}_{r,\Delta}((0,0),(-1,0))$. Dans beaucoup de cas, nous pouvons construire de meilleurs codes en utilisant des lignes diagonales qui « cassent » la symétrie. Pour cela nous avons besoin d'informations sur $\mathcal{S}_{r,\Delta}((0,0),(-1,0))$ que nous ne sommes pas capables de calculer pour des valeurs générales. La construction qui sera donnée dans la proposition 4.32 doit être vue comme une méthode de construction de code (r, Δ)-identifiant pour des valeurs données de r et Δ.

Un ensemble U de $\mathbb{Z}^2$ *intersecte toutes les lignes diagonales (resp. anti-diagonales, horizontales et verticales) modulo* k si pour tout $i \in \{0, 1, \ldots, k-1\}$, il y a un élément (x, y) de U tel que $y - x \equiv i \bmod k$ (resp. $x + y \equiv i \bmod k$, $y \equiv i \bmod k$ et $x \equiv i \bmod k$).

Par exemple, le motif horizontal $\mathcal{S}_{3,\sqrt{10}-3}((0,0),(-1,0))$ (voir la figure 4.17) intersecte toutes les lignes diagonales et anti-diagonales modulo 6, alors que le motif diagonal $\mathcal{S}_{3,\sqrt{10}-3}((0,0),(-1,-1))$ intersecte toutes les lignes horizontales et verticales modulo 8.

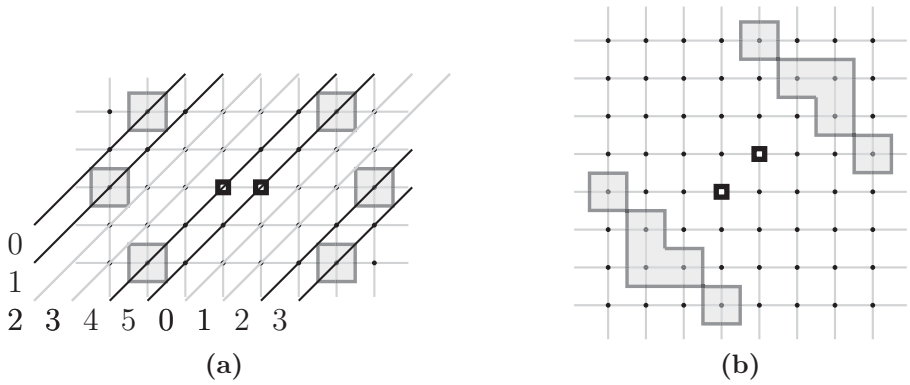


Figure 4.17 – *Le motif horizontal* $\mathcal{S}_{3,\sqrt{10}-3}((0,0),(-1,0))$ *(figure (a)) intersecte toutes les lignes diagonales et anti-diagonales modulo 6. Le motif diagonal* $\mathcal{S}_{3,\sqrt{10}-3}((0,0),(-1,-1))$ *(figure (b)) intersecte toutes les lignes horizontales et verticales modulo 8.*

Nous notons $diag(U)$ le plus grand entier k tel que l'ensemble U intersecte toutes les lignes diagonales modulo k.

Lemme 4.30 *Soit u un sommet de $\mathbb{Z}^2$. Alors :*

$$diag(\mathcal{B}_r(u)) = 4\left\lfloor \frac{r}{\sqrt{2}} \right\rfloor + 2\delta + 1,$$

où $\delta = 0$ *si* $\left\lfloor \frac{r}{\sqrt{2}} \right\rfloor^2 + \left(\left\lfloor \frac{r}{\sqrt{2}} \right\rfloor + 1 \right)^2 \leq r^2$ *et* $\delta = 1$ *sinon.*

Preuve : Par symétrie, il suffit de montrer le résultat pour $u = (0,0)$. Soient $v = \left(-\left\lfloor \frac{r}{\sqrt{2}} \right\rfloor, \left\lfloor \frac{r}{\sqrt{2}} \right\rfloor \right)$ et $w = \left(\left\lfloor \frac{r}{\sqrt{2}} \right\rfloor, -\left\lfloor \frac{r}{\sqrt{2}} \right\rfloor \right)$ les deux sommets de $\mathbb{Z}^2$ sur la droite d'équation $y = -x$ les plus éloignés de u. Toutes les diagonales se situant entre celles contenant v et w (incluses) intersectent $\mathcal{B}_r(u)$. Cela fait $4 \left\lfloor \frac{r}{\sqrt{2}} \right\rfloor + 1$ lignes diagonales.

Il peut y avoir une diagonale de plus avant v si le sommet $v + (-1,0)$ est dans $\mathcal{B}_r(u)$. Cela correspond à la condition

$$\left\lfloor \frac{r}{\sqrt{2}} \right\rfloor^2 + \left(\left\lfloor \frac{r}{\sqrt{2}} \right\rfloor + 1 \right)^2 \leq r^2$$

et dans ce cas il y a aussi une ligne diagonale de plus après w. □

Remarquons que l'ensemble $\mathcal{B}_r((0,0))$ intersecte toutes les lignes diagonales modulo s pour n'importe quel $s \leq diag(\mathcal{B}_r((0,0)))$. Cela n'est pas toujours vrai, par exemple, l'ensemble $\mathcal{S}_{3,\sqrt{10}-3}((0,0),(-1,0))$ n'intersecte pas toutes les lignes diagonales modulo 4 alors que $diag(\mathcal{S}_{3,\sqrt{10}-3}((0,0),(-1,0)) = 6$ (voir la figure 4.17).

Lemme 4.31 *Si* $\mathcal{S}_{r,\Delta}((0,0),(-1,0))$ *intersecte toutes les lignes diagonales modulo* s*, avec* $s \leq diag(\mathcal{B}_r((0,0)))$*, alors pour toute paire de sommets* u, v *de* $\mathbb{Z}^2$ *situés sur la même ligne horizontale ou verticale,* $\mathcal{S}_{r,\Delta}(u,v)$ *intersecte toutes les lignes diagonales modulo* s*.*

Preuve : Remarquons tout d'abord que si u et v sont situés sur la même ligne horizontale et que $\mathcal{S}_{r,\Delta}(u,v)$ intersecte toutes les lignes diagonales modulo s, alors $\mathcal{S}_{r,\Delta}(u,v)$ intersecte aussi toutes les lignes anti-diagonales modulo s, du fait de la symétrie verticale de $\mathcal{S}_{r,\Delta}(u,v)$. Ainsi, l'ensemble $\mathcal{S}_{r,\Delta}(u',v')$ où u' et v' sont les images de u et v par une rotation d'angle $\frac{\pi}{2}$ centrée en $(0,0)$, intersecte aussi toutes les lignes diagonales et anti-diagonales modulo s. C'est pourquoi il est suffisant de montrer le lemme seulement pour des sommets situés sur la même ligne horizontale.

Par hypothèse, $\mathcal{S}_{r,\Delta}((0,0),(-1,0))$ intersecte toutes les lignes diagonales modulo s, avec $s \leq diag(\mathcal{B}_r((0,0)))$. Soient u et v situés sur la même ligne horizontale. Si $d_{\mathcal{E}}(u,v) \leq 4x_0 + 1$, d'après le lemme 4.27, $\mathcal{S}_{r,\Delta}(u,v)$ contient une translation de l'ensemble $\mathcal{S}_{r,\Delta}((0,0),(-1,0))$, et intersecte donc toutes les lignes diagonales modulo s. Si $d_{\mathcal{E}}(u,v) > \lfloor r \rfloor + \lfloor r + \Delta \rfloor$, alors $\mathcal{B}_r(v) \setminus \mathcal{B}_{r+\Delta}(u) = \mathcal{B}_r(v)$. L'ensemble $\mathcal{B}_r(v)$ intersecte toutes les lignes diagonales modulo $s \leq diag(\mathcal{B}_r(v)) = diag(\mathcal{B}_r((0,0)))$ et donc l'ensemble $\mathcal{S}_{r,\Delta}(u,v)$ aussi.

Nous pouvons donc supposer que $4x_0 + 2 \leq d_{\mathcal{E}}(u,v) \leq \lfloor r \rfloor + \lfloor r + \Delta \rfloor$. Sans perte de généralité, supposons que $u = (-k',0)$ et que $v = (k,0)$ avec $k = \left\lfloor \frac{d_{\mathcal{E}}(u,v)}{2} \right\rfloor$, $k + k' = d_{\mathcal{E}}(u,v)$ et $k' \geq k$.

Nous avons, d'après l'équation 4.1, $\Delta(2r + \Delta) \leq 2x_0 + 1 \leq k \leq \lfloor r \rfloor$. Soient w_u le sommet $u + \left(-\left\lfloor \frac{r}{\sqrt{2}} \right\rfloor, \left\lfloor \frac{r}{\sqrt{2}} \right\rfloor \right)$ et w_v le sommet $v + \left(\left\lfloor \frac{r}{\sqrt{2}} \right\rfloor, -\left\lfloor \frac{r}{\sqrt{2}} \right\rfloor \right)$. Nous allons montrer que $\mathcal{S}_{r,\Delta}(u,v)$ intersecte toutes les diagonales comprises entre celle contenant w_u et celle contenant w_v, ce qui fera plus de s diagonales consécutives.

Considérons la ligne diagonale D définie par l'équation $y = x$. Nous montrons tout d'abord que toutes les lignes diagonales situées entre D et celle contenant w_v intersectent $\mathcal{B}_r(v) \setminus \mathcal{B}_{r+\Delta}(u)$. Comme $k \leq \lfloor r \rfloor$, toutes ces diagonales intersectent $\mathcal{B}_r(v)$ en un sommet de $\mathbb{Z}^2$ d'abscisse strictement positive. Si D n'intersecte pas $\mathcal{B}_{r+\Delta}(u)$ en une abscisse positive, alors c'est fini. Supposons donc que D intersecte le cercle $C_{r+\Delta}(u)$ en un point d'abscisse x_u positive. (x_u n'est pas nécessairement entier). Soit x_v l'abscisse du point d'intersection de D et du cercle $C_r(v)$ d'abscisse positive. Alors si la distance entre les deux points (x_u, x_u) et (x_v, x_v) est supérieure à $\sqrt{2}$, il y a obligatoirement un sommet de $\mathbb{Z}^2$ situé sur D dans l'ensemble $\mathcal{B}_r(v) \setminus \mathcal{B}_{r+\Delta}(u)$. Comme la distance entre les deux cercles sur chaque diagonale située à la droite de D sera aussi supérieure à $\sqrt{2}$, il y aura un sommet de $\mathbb{Z}^2$ sur chaque ligne diagonale située entre D et w_v dans $\mathcal{B}_r(v) \setminus \mathcal{B}_{r+\Delta}(u)$.

La distance entre les deux points (x_u, x_u) et (x_v, x_v) est égale à $\sqrt{2}(x_v - x_u)$. Il suffit donc de montrer que $x_v - x_u \geq 1$. Nous avons :

$$x_v - x_u = \frac{k+k'}{2} + \frac{1}{2}(\sqrt{2r^2 - k^2} - \sqrt{2r^2 - k'^2 + 4r\Delta + 2\Delta^2})$$

Il y a alors deux cas : soit $x_0(r, \Delta) = 0$ et alors $\Delta = 0$,$k = k' = 1$ et c'est fini, soit $x_0(r, \Delta) \geq 1$ et alors $k \geq 3$ et :

$$\begin{aligned} x_v - x_u &\geq k + \frac{1}{2}(\sqrt{2r^2 - k^2} - \sqrt{2r^2 - k^2 + 2k}) \\ &= k + \frac{2r^2 - k^2 - (2r^2 - k^2 + 2k)}{2(\sqrt{2r^2 - k^2} + \sqrt{2r^2 - k^2 + 2k})} \\ &\geq k - \frac{k}{2\sqrt{2r^2 - k^2}} \\ &\geq \frac{k}{2} > 1. \end{aligned}$$

Nous pouvons faire la même étude pour $\mathcal{B}_r(u) \setminus \mathcal{B}_{r+\Delta}(v)$ pour montrer qu'il existe un sommet dans $\mathcal{B}_r(u) \setminus \mathcal{B}_{r+\Delta}(v)$ situé sur chaque diagonale comprise entre w_u et la diagonale définie par $y = x + 1$, ce qui termine la preuve. □

La prochaine proposition donne, pour des valeurs données de r et Δ, une méthode pour construire des meilleurs codes (r, Δ)-identifiant que la méthode de la proposition 4.29. Lorsque r et Δ sont fixés, il est assez facile de vérifier que le code construit dans la prochaine proposition est bien (r, Δ)-identifiant. Cependant, la preuve générale est technique, même en se restreignant à certains valeurs de r. Nous pensons que cette méthode fonctionne pour toutes valeurs de r et de Δ mais ne sommes pas en mesure de le prouver. La figure 4.18 donne un exemple d'application de cette méthode.

Proposition 4.32 *Soient $s \leq diag(\mathcal{B}_r((0,0)))$ et t des entiers. Soit*

$$L_s^d = \{(x, y) | y - x \equiv 0 \bmod s\}.$$

Si les conditions suivantes sont vérifiées :

(a) $\mathcal{S}_{r,\Delta}((0,0),(-1,0))$ *intersecte toutes les lignes diagonales modulo* s *;*
(b) $\mathcal{S}_{r,\Delta}((0,0),(-1,-1))$ *intersecte toutes les lignes horizontales modulo* t *;*
(c) r *n'est pas trop proche d'un entier :* $\sqrt{\lfloor r\rfloor^2+4} \leq r < \lfloor r\rfloor + 1$ *;*
alors le code $C = L_s^d \cup L_t^h$ *est un code* (r,Δ)*-identifiant de densité* $\frac{1}{s}+\frac{1}{t}-\frac{1}{st}$.

Preuve : Supposons que les conditions de la proposition sont remplies.

Nous remarquons tout d'abord qu'il y a au plus $2\lfloor r\rfloor + 2$ lignes horizontales qui intersectent $\mathcal{S}_{r,\Delta}((0,0),(-1,-1))$, donc nécessairement, $t \leq 2r+2$.

L'ensemble L_s^d est un code contenant des lignes diagonales espacées de s, il est donc de densité $\frac{1}{s}$. Le code $C = L_s^d \cup L_t^h$ est alors de densité $\frac{1}{s}+\frac{1}{t}-\frac{1}{st}$. Nous allons montrer que c'est un code (r,Δ)-identifiant.

Comme $s \leq diag(\mathcal{B}_r((0,0)))$, C est un ensemble r-dominant. Soient u et v deux sommets distincts. Sans perte de généralité, nous pouvons écrire $v = u + (x,y)$ avec $x \geq 0$. D'après le lemme 4.31, $C \cap \mathcal{S}_{r,\Delta}(u,v)$ est non vide lorsque u et v sont situés sur la même ligne horizontale ou verticale. Supposons donc que $x \geq 1$ et $|y| \geq 1$.

Soit k la plus grande ordonnée positive d'un sommet de $\mathcal{B}_r(0,0)$ d'abscisse $\lfloor r\rfloor$. La condition (c) implique que $k \geq 2$. Soit $E = \{(x_0,y_0),\ldots,(x_{t-1},y_{t-1})\}$ un ensemble de sommets de $\mathcal{S}_{r,\Delta}((0,0),(-1,-1))$ tel que $y_i \equiv i \bmod t$. Nous pouvons supposer que $x_iy_i \geq 0$ pour tout $i \in \{0,...,t-1\}$. En effet, supposons par exemple que $x_i < 0$ et $y_i > 0$. Si $(x_i,y_i) \in \mathcal{B}_r((0,0)) \setminus \mathcal{B}_{r+\Delta}((-1,-1))$, alors $(-x_i,y_i) \in \mathcal{B}_r((0,0)) \setminus \mathcal{B}_{r+\Delta}((-1,-1))$ et nous changeons le couple (x_i,y_i) par le couple $(-x_i,y_i)$. Sinon $(x_i,y_i) \in \mathcal{B}_r((-1,-1)) \setminus \mathcal{B}_{r+\Delta}((0,0))$, et $(-x_i+1,y_i) \in \mathcal{B}_r((0,0)) \setminus \mathcal{B}_{r+\Delta}((-1,-1))$. Alors nous changeons (x_i,y_i) par $(-x_i+1,y_i)$.

Pour $0 \leq i \leq k$, l'ensemble $E + (0,-i)$ est alors inclus dans les ensembles $\mathcal{S}_{r,\Delta}((0,0),(-1,-2i))$ et $\mathcal{S}_{r,\Delta}((0,0),(-1,-2i-1))$. Aussi, par translation, $\mathcal{S}_{r,\Delta}(u,v)$ intersecte toutes les lignes horizontales modulo y pour $|y| \leq 2k+1$ et $x = 1$. Par symétrie, cela est aussi vrai pour tout $x > 0$ et $|y| \leq 2k+1$. Si $x > 0$ et $y > 2k+1$, alors $\mathcal{S}_{r,\Delta}(u,v)$ intersecte $diag(B_r(0,0))$ lignes diagonales consécutives, donc $\mathcal{S}_{r,\Delta}(u,v) \cap C \neq \emptyset$.

Le cas restant est le cas $x > 0$ et $y < -(2k+1)$. Nous traitons tout d'abord le cas $x = 1$. Sans perte de généralité, supposons que $u = (0,h)$ et $v = (1,-h')$ avec $h'+1 \geq h \geq h' \geq k+1$. Si $h \geq r$, la ligne diagonale d'équation $y = x$ n'intersecte pas le cercle $C_{r+\Delta}(v)$ en une abscisse positive, donc $\mathcal{S}_{r,\Delta}(u,v)$ intersecte au moins $diag(B_r((0,0)))$ lignes diagonales consécutives. Dans l'autre cas $(h < r)$, nous montrons, en utilisant la même méthode que pour la preuve du lemme 4.31, que a distance sur la ligne diagonale $y = x$ entre les deux cercles $C_r(u)$ et $C_{r+\Delta}(v)$ est au moins $\sqrt{2}$. Cela impliquera que $\mathcal{S}_{r,\Delta}(u,v)$ intersecte au moins $diag(B_r((0,0)))$ lignes diagonales consécutives et terminera le cas $x = 1$. Pour cela, notons (x_u,x_u) et (x_v,x_v) les points du plan (non-éléments de $\mathbb{Z}^2$) à l'intersection de la ligne diagonale $y = x$ et des cercles $C_r(u)$ et $C_{r+\Delta}(v)$ respectivement. Alors $x_u = \frac{h}{2} + \frac{\sqrt{2r^2-h^2}}{2}$ et $x_v = 1 - \frac{h'}{2} + \frac{\sqrt{2r^2-h'^2+4r\Delta-2h'+2\Delta^2-1}}{2}$. D'où :

$$x_v - x_u = \frac{h+h'}{2} - 1 + \frac{\sqrt{2r^2-h^2} - \sqrt{2r^2-h'^2+4r\Delta-2h'+2\Delta^2-1}}{2} \geq h'-2 \geq 1.$$

La distance entre (x_u,x_u) et (x_v,x_v) est donc d'au moins $\sqrt{2}$, ce qui termine le cas $x = 1$.

Avec la méthode, nous pouvons montrer que l'ensemble $\mathcal{S}_{r,\Delta}(u,v)$ intersecte aussi $diag(B_r((0,0)))$ lignes diagonales consécutives lorsque $x \geq 2$. □

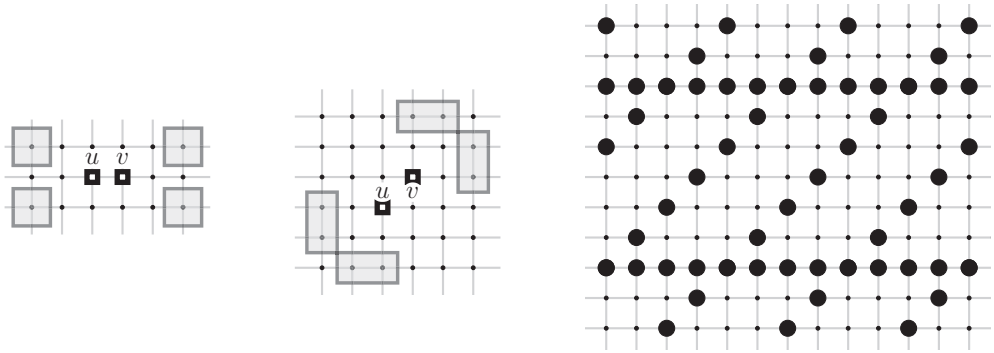


Figure 4.18 – *Code obtenu avec la proposition 4.32 dans le cas* $(r,\Delta) = (\sqrt{5}, 3-\sqrt{5})$. *Le motif horizontal (figure (a)) intersecte toutes les lignes diagonales modulo* 4. *Le motif diagonal (figure (b)) intersecte toutes les lignes horizontales modulo* 6. *Le code* $L_4^d \cup L_6^h$ *(figure (c)) est un code* $(\sqrt{5}, 3-\sqrt{5})$*-identifiant de densité* $\frac{5}{12}$. *Les sommets aux intersections des lignes diagonales et horizontales peuvent être retirés, menant à un code de densité* $\frac{1}{3}$.

Dans certains cas, le code construit dans la proposition 4.32 peut être légèrement amélioré en enlevant les sommets qui sont à l'intersection des lignes horizontales et diagonales. En effet, pour tout sommet du motif horizontal, il y a un autre sommet situé sur la même ligne horizontale. Ainsi, si la distance entre deux sommets du motif horizontal situés sur la même ligne horizontale n'est jamais égale à s, alors lorsque le code de la proposition 4.32 intersecte le motif horizontal au niveau de l'intersection des lignes horizontales et diagonales du code, il l'intersecte aussi à un autre point de la ligne horizontale, qui n'est pas sur une des diagonales. Pour le motif diagonal, nous pouvons faire le même raisonnement avec les lignes diagonales.

C'est par exemple le cas lorsque $(r,\Delta) = (\sqrt{5}, 3-\sqrt{5})$ (voir la figure 4.18). Le code $C = (L_4^d \cup L_6^h) \setminus (L_4^d \cap L_6^h)$ est un code $(\sqrt{5}, 3-\sqrt{5})$-identifiant de densité $\frac{1}{3}$. Cela ne fonctionne pas toujours, comme par exemple dans le cas $(r,\Delta) = (\sqrt{41}, 5\sqrt{2} - \sqrt{41})$.

Construction optimale

En pratique, le motif diagonal contient généralement beaucoup plus de sommets que le motif horizontal et les lignes diagonales du code de la proposition 4.32 vont suffire. Nous obtenons alors un code densité $\frac{1}{s}$. C'est en particulier le cas dès que le rayon r devient assez grand et que s reste petit (ou fixé).

Proposition 4.33 *Soit* s *un entier fixé.*

(i) *S'il existe* $r_0 \in \mathbb{N}$ *tel que pour tout* $r \geq r_0$ *et tout* $\Delta \in [0, \Delta_m(r)]$, *l'ensemble* $\mathcal{S}_{r,\Delta}((0,0),(-1,-1))$ *interesecte toutes les lignes diagonales et anti-diagonales modulo* s *dans le premier quadrant.*

(ii) *Si de plus, l'ensemble* $\mathcal{S}_{r,\Delta}((0,0),(-1,0))$ *intersecte toutes les lignes diagonales modulo* s, *il existe un code* (r,Δ)*-identifiant de densité* $\frac{1}{s}$.

Preuve :

(i) Comme $\Delta_m(r) \leq 1$, l'ensemble $\mathcal{S}_{r,1}((0,0),(-1,-1))$ est contenu dans l'ensemble $\mathcal{S}_{r,\Delta}((0,0),(-1,-1))$. Aussi pour montrer la première partie de la proposition, nous supposons que $\Delta = 1$.

Considérons l'ensemble des sommets $\mathbb{Z}^2$ comme une partition en deux ensembles de sommets : les *sommets pairs* $\mathbb{Z}^2_e = \{(i,j) \in \mathbb{Z}^2 \mid i+j \equiv 0 \bmod 2\}$ et les *sommets impairs* $\mathbb{Z}^2_o = \{(i,j) \in \mathbb{Z}^2 \mid i+j \equiv 1 \bmod 2\}$ (voir la figure 4.19).

L'ensemble des sommets pairs $\mathbb{Z}^2_e$ peut être vu comme la rotation du réseau $\mathbb{Z}^2$ d'angle $\pi/4$ où la distance unitaire entre deux sommets les plus proches est $\sqrt{2}$ (à la place de 1). Les lignes diagonales (resp. anti-diagonales) du réseau d'origine $\mathbb{Z}^2$ correspondent aux lignes horizontales (resp. verticales) du nouveau réseau $\mathbb{Z}^2_e$.

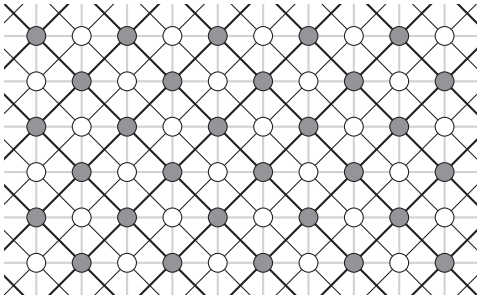


Figure 4.19 – *Partition de $\mathbb{Z}^2$ en sommets pairs et impairs. Les sommets pairs (resp. impairs) forment un réseau similaire à $\mathbb{Z}^2$ avec une rotation $\pi/4$ et une dilatation de $\sqrt{2}$.*

Considérons tout d'abord les lignes anti-diagonales du premier quadrant. L'ensemble $\mathcal{S}_{r,1}((0,0),(-1,-1))$ intersecte $r - x_1(r,1)$ lignes verticales consécutives. En passant au réseau $\mathbb{Z}^2_e$, cela fait donc $\frac{r}{\sqrt{2}} - x_1(\frac{r}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ lignes verticales intersectées, qui correspondent à des lignes anti-diagonales contenant des sommets pairs qui intersectent $\mathcal{S}_{r,1}((0,0),(-1,-1))$. Avec le même raisonnement sur $\mathbb{Z}^2_o$, il y a, entre chacune de ces lignes, une ligne anti-diagonale de sommets impairs qui intersecte $\mathcal{S}_{r,1}((0,0),(-1,-1))$. Le nombre de lignes anti-diagonales consécutives intersectant $\mathcal{S}_{r,1}((0,0),(-1,-1))$ est de l'ordre de $\frac{r}{\sqrt{2}} - x_1(\frac{r}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ qui, d'après l'équation 4.3, tend vers l'infini lorsque $r \to +\infty$. Pour r assez grand, il y a donc s lignes anti-diagonales consécutives qui intersectent $\mathcal{S}_{r,1}((0,0),(-1,-1))$ dans le premier quadrant.

Considérons maintenant les lignes diagonales (du premier quadrant). En utilisant le résultat sur $m(r,\Delta)$ de l'équation (4.5), nous pouvons remarquer qu'il y a une ligne verticale de $\mathbb{Z}^2_e$ (qui est une ligne anti-diagonale dans $\mathbb{Z}^2$) telle que le nombre de sommets de cette ligne qui sont des éléments de $\mathcal{S}_{r,1}((0,0),(-1,-1))$ tend vers l'infini quand r augmente. De la même manière, il y a une ligne verticale de $\mathbb{Z}^2_o$ qui contient un nombre croissant de sommets de $\mathcal{S}_{r,1}((0,0),(-1,-1))$. Ainsi il y a s lignes diagonales consécutives qui intersectent $\mathcal{S}_{r,1}((0,0),(-1,-1))$ dans le premier quadrant, lorsque r est assez grand.

Par symétrie et translation, l'ensemble $\mathcal{S}_{r,\Delta}(u,v)$ intersecte toutes les lignes diagonales modulo s du premier quadrant (resp. du second quadrant) pour tous u et v dans $\mathbb{Z}^2$ avec $v = u + (1,1)$ (resp. $v = u + (1,-1)$).

(ii) : Supposons maintenant que $\mathcal{S}_{r,\Delta}((0,0),(-1,0))$ intersecte toutes les lignes diagonales modulo s. Soit $C = L^d_s$ (voir la proposition 4.32). Montrons que C est

un code (r, Δ)-identifiant. Soient $u = (x_u, y_u) \in \mathbb{Z}^2$ et $v = (x_v, y_v) \in \mathbb{Z}^2$ deux sommets différents. Sans perte de généralité, il est suffisant de considérer les deux cas suivants :

(a) Si $|x_u - x_v| \geq 1$ et $|y_v - y_u| \geq 1$, alors d'après la première partie de la proposition il y a une ligne diagonale intersectant $\mathcal{S}_{r,\Delta}(u, v)$.

(b) Si $y_v = y_u$ ou $x_v = x_u$, c'est-à-dire si u et v sont situés sur la même ligne verticale ou horizontale, alors d'après le lemme 4.31, $\mathcal{S}_{r,\Delta}(u, v)$ intersecte toutes les diagonales modulo s et intersecte donc C. □

Constructions particulières et codes optimaux pour Δ proche de 1

Les résultats obtenus précédemment ne sont pas souvent serrés. Cela est dû en grande partie au fait que nous n'ayons pas une connaissance précise et générale de l'ensemble $\mathcal{S}_{r,\Delta}((0, 0), (-1, 0))$. Cependant, lorsque Δ est proche de 1, l'ensemble $\mathcal{S}_{r,\Delta}((0, 0), (-1, 0))$ devient de plus en plus petit. Lorsqu'il ne contient que quelques éléments (moins de dix), nous pouvons alors l'étudier en détail. C'est l'objet de cette sous-partie : nous étudions les code (r, Δ)-identifiants lorsque $|\mathcal{S}_{r,\Delta}((0, 0), (-1, 0))| \leq 8$. En particulier, nous pouvons extraire des familles infinies de valeurs (r, Δ) pour lesquelles nous connaissons la valeur de la densité optimale $D(r, \Delta)$.

Dans la suite, s désigne $|\mathcal{S}_{r,\Delta}((0, 0), (-1, 0))|$. Notons que s est nécessairement pair. De plus, si s n'est pas divisible par 4 alors r est un entier (ou très proche d'un entier). Nous commençons avec $s = 2$ (et donc r entier). Dans ce cas là, nous avons toujours $D(r, \Delta) = \frac{1}{2}$, quelque soit la valeur de r.

Proposition 4.34 *Soient r un entier et Δ tel que $\sqrt{r^2 + 2r - 1} \leq r + \Delta < 1$. Alors $|\mathcal{S}_{r,\Delta}((0, 0), (-1, 0))| = 2$, et $D(r, \Delta) = \frac{1}{2}$.*

Preuve : La première partie de la proposition n'est pas difficile : r est un entier donc pour tout $\Delta < 1$, $(-r - 1, 0)$ et $(r, 0)$ sont des éléments de $\mathcal{S}_{r,\Delta}((0, 0), (-1, 0))$, et sont les seuls éléments de $\mathcal{S}_{r,\Delta}((0, 0), (-1, 0))$ d'ordonnée 0. Comme $r + \Delta \geq \sqrt{r^2 + 2r - 1}$, si un sommet (x, y) est dans $\mathcal{B}_r((0, 0))$, avec $u \neq 0$, alors $x \leq r - 1$ et $(x + 1)^2 + y^2 \leq r^2 + 2r - 1 \leq (r + \Delta)^2$ donc $(x, y) \in \mathcal{B}_{r+\Delta}((-1, 0))$.

Pour la deuxième partie, remarquons tout d'abord que $\mathcal{S}_{r,\Delta}((0, 0), (-1, 0))$ intersecte toutes les lignes diagonales modulo 2 car r et $-r - 1$ n'ont pas la même parité. Ainsi, lorsque r est assez grand, la proposition 4.33 nous assure que $D(r, \Delta) = \frac{1}{2}$. Nous allons construire explicitement un code de densité $\frac{1}{2}$ pour des petites valeurs de r. La figure 4.20 montre la construction dans le cas $r = 4$.

Nous construisons le code en utilisant des lignes verticales contenant un sommet sur deux : $L_{0.5}^{(o)} = \{(0, y) \in \mathbb{Z}^2 | y \text{ impair}\}$ et $L_{0.5}^{(e)} = \{(0, y) \in \mathbb{Z}^2 | y \text{ pair}\}$. Si r est impair, posons $X = \{0, 1, \ldots, 2r\}$. Si r est pair, posons

$$X = \{0, 2, 4, \ldots, r - 2\} \cup \{r - 1\} \cup \{r + 1, r + 3, \ldots, 2r - 3\} \cup \{2r\}$$

. Dans le cas $r = 4$ représenté sur la figure 4.20, nous obtenons $X = \{0, 2, 3, 5, 8\}$.

Nous construisons ensuite un sous-ensemble de $\mathbb{Z}^2$ inclus dans la ligne d'ordonnée 0 qui sera la base du motif :

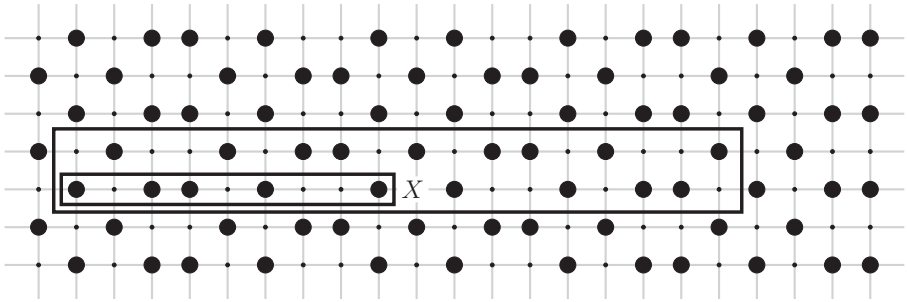


Figure 4.20 – *Un code optimal de densité* $\frac{1}{2}$ *pour* $r = 4$ *avec* $\sqrt{23} - 4 \leq \Delta < 1$.

$$U = \{(x,0) | x \equiv i \bmod 4r+2, 0 \leq i \leq 2r \text{ et } x \in X\} \cup \\ \{(x,0) | x \equiv i \bmod 4r+2, 2r+1 \leq i \leq 4r+1 \text{ et } i-(2r+1) \notin X\}. \quad (4.9)$$

Nous pouvons maintenant définir le code C par :

$$C = \bigcup_{u \in U} (u + L_{0.5}^{(o)}) \bigcup_{u \notin U} (u + L_{0.5}^{(e)}).$$

Pour montrer que ce code est (r, Δ)-identifiant, nous avons uniquement besoin de montrer qu'il intersecte le motif diagonal. En effet, le code intersecte tous les motifs en L : et chaque ensemble $\mathcal{S}_{r,\Delta}(u,v)$ contient un motif en L dès lors que $d_{\mathcal{E}}(u,v) > \sqrt{2}$. De plus, C intersecte clairement le motif vertical (les deux sommets ne sont pas de même parité et chaque colonne est alternée) et le motif horizontal : dans les ensembles U et $\overline{U}$, il n'y a pas deux sommets à distance $2r+1$. Ainsi il reste uniquement le cas $d_{\mathcal{E}}(u,v) = \sqrt{2}$ à traiter.

Nous montrons le résultat seulement pour le cas $v = u + (1,1)$ et r pair, les autres cas étant similaires ou plus faciles. Dans $\mathcal{S}_{r,\Delta}(u,v)$, il y a nécessairement les quatre sommets suivant : $A = u + (-r, 0)$, $B = u + (0, -r)$, $C = v + (0, r)$ et $D = v + (r, 0)$. Comme les lignes verticales alternent une fois sur deux les sommets du code, B (resp. C) est dans le code si et seulement si u (resp. v) est dans le code. Soit (x,y) les coordonnées de A. Si ni A ni u ne sont dans le code, alors soit y est pair et $x \equiv 2r-1 \bmod 4r+2$, ou bien y est impair et $x \equiv 4r \bmod 4r+2$. Dans tous les cas, nous pouvons vérifier que v est dans le code, ce qui termine la preuve. □

Nous considérons maintenant le cas où $\mathcal{S}_{r,\Delta}((0,0),(-1,0))$ contient quatre éléments ($s = 4$). Nous pouvons tout d'abord assurer l'existence dans tous les cas d'un code de densité $\frac{3}{8}$:

Proposition 4.35 *Soient* r *et* Δ *tels que* $|\mathcal{S}_{r,\Delta}((0,0),(-1,0))| = 4$. *Alors*

$$D(r, \Delta) \leq \frac{3}{8}.$$

Preuve : L'ensemble $\mathcal{S}_{r,\Delta}((0,0),(-1,0))$ contient nécessairement un sommet de coordonnées (a,b) avec $a \geq b > 0$. Les autres sommets de $\mathcal{S}_{r,\Delta}((0,0),(-1,0))$ sont alors $(a,-b)$, $(-a-1,b)$ et $(-a-1,-b)$.

Écrivons $b = 2^k b'$ avec b' impair. Nous construisons un code C en deux parties C_1 et C_2 : $C_1 = \{(x,y) \in \mathbb{Z}^2 \mid x \text{ et } y \text{ sont pairs}\}$ et

$$C_2 = \{(x,y) \in \mathbb{Z}^2 | (y-x) = i \bmod 2^{k+2} \text{ avec } i \in \{2,4,6,\ldots,2^{k+1}\}\}.$$

Alors le code $C = C_1 \cup C_2$ est de densité $\frac{3}{8}$. Nous allons montrer que c'est un code (r,Δ)-identifiant.

Dans le motif diagonal, $\mathcal{S}_{r,\Delta}((0,0),(-1,-1))$, il y a au moins les quatre sommets (a,b), (b,a), $(a,b-1)$, $(b-1,a)$ et leurs images par la rotation centrale de centre $(-0.5,-0.5)$. Certains de ces sommets peuvent être égaux (si $a = b$ par exemple) mais dans toute translation du motif diagonal il y aura un sommet donc les deux coordonnées sont paires, qui est obligatoirement dans C_1. Par symétrie, C_1 intersecte aussi tous les motifs anti-diagonaux.

Pour le motif horizontal, remarquons que pour tout $c \in \mathbb{Z}$, parmi les quatre valeurs $c+b-a$, $c-b-a$, $c+b+a+1$ et $c-b+a+1$ modulo 2^{k+2}, il y en a exactement une d'entre elle qui est dans l'ensemble $\{2,4,6,\ldots,2^{k+1}\}$. Cela implique que C_2 intersecte tous les motifs horizontaux. Nous raisonnons de même pour le motif vertical.

Enfin, prenons u et v deux sommets de $\mathbb{Z}^2$ avec $d_{\mathcal{E}}(u,v) > \sqrt{2}$. Si u et v sont situés sur la même ligne et que $d_{\mathcal{E}}(u,v) \leq 5$, avec le lemme 4.27, $\mathcal{S}_{r,\Delta}(u,v)$ contient un ensemble isomorphe au motif vertical ou horizontal et intersecte C_2. Sinon, $\mathcal{S}_{r,\Delta}(u,v)$ contient un carré de quatre sommets et intersecte donc C_1, ce qui termine la preuve. □

Pour certaines familles (r,Δ) avec $|\mathcal{S}_{r,\Delta}((0,0),(-1,0))| = 4$, nous pouvons assurer que $\mathcal{S}_{r,\Delta}((0,0),(-1,0))$ intersecte toutes les diagonales modulo 4. En utilisant la proposition 4.33, nous obtenons alors des familles dont la densité optimale d'un code (r,Δ)-identifiant est $\frac{1}{4}$ si r est assez grand.

Proposition 4.36 *Soient k et i deux entiers tels que i soit impair et $i^2 < 2k+1$. Soient $r = \sqrt{k^2+i^2}$ et Δ tel que $r+\Delta = \sqrt{r^2+2k}$. Alors $|\mathcal{S}_{r,\Delta}((0,0),(-1,0))| = 4$ et $\mathcal{S}_{r,\Delta}((0,0),(-1,0))$ intersecte toutes les lignes diagonales modulo 4. Si k est choisi assez grand, la densité optimale d'un code (r,Δ)-identifiant est $\frac{1}{4}$.*

Preuve : Pour montrer que $|\mathcal{S}_{r,\Delta}((0,0),(-1,0))| = 4$ nous montrons que l'ensemble

$$X_1 = \mathcal{S}_{r,\Delta}((0,0),(-1,0)) \cap \{(x,y) | x > 0, y \geq 0\}$$

ne contient qu'un seul sommet et que l'ordonnée de ce sommet est strictement positive.

Soit $(x,y) \in X_1$. Nous avons $x \leq \lfloor r \rfloor = k$. Si $x = k$ alors $y \leq i$. Si $y < i$ alors

$$(x+1)^2 + y^2 \leq (k+1)^2 + i^2 - 1 \leq r^2 + 2k = (r+\Delta)^2,$$

et $(x,y) \notin X_1$. Si $y = i$ alors

$$(x+1)^2 + y^2 = r^2 + 2k + 1 > (r+\Delta)^2,$$

et donc $(k,i) \in X_1$. Si $x < k$, alors

$$(x+1)^2 + y^2 \leq r^2 + 2k - 1 \leq (r+\Delta)^2,$$

et nécessairement $(x,y) \notin X_1$. Ainsi, le seul sommet de X_1 est (k,i), avec $i > 0$. Cela implique que $|\mathcal{S}_{r,\Delta}((0,0),(-1,0))| = 4$ et que les quatre sommets de l'ensemble $\mathcal{S}_{r,\Delta}((0,0),(-1,0))$ sont (k,i), $(k,-i)$, $(-k-1,i)$ et $(-k-1,-i)$.

Nous devons maintenant montrer que les quatre valeurs $i-k$, $-i-k$, $i+k+1$ et $-i+k+1$ sont différentes modulo 4. En ajoutant $i+k$, il est équivalent de montrer que les quatre valeurs 0, $2i$, $2k+1$ et $2k+1+2i$ sont différentes modulo 4, ce qui découle du fait que i est impair. La dernière partie de la proposition est une conséquence directe de la proposition 4.33 car $r \to +\infty$ lorsque $k \to +\infty$. □

Passons maintenant au cas $s = 6$. Nous donnons une famille de valeurs pour lesquelles l'ensemble $\mathcal{S}_{r,\Delta}((0,0),(-1,0))$ est de cardinalité 6 et intersecte toutes les lignes diagonales modulo 6. Nous avons construit ces familles par tâtonnements. Puisque 6 n'est pas divisible par 4, le rayon r doit être entier. La valeur de Δ ne doit pas être trop proche de 1 sinon il n'y aura que deux sommets dans l'ensemble $\mathcal{S}_{r,\Delta}((0,0),(-1,0))$, ni trop petite pour n'avoir que six sommets. Enfin parmi toutes les valeurs possibles où nous avons six sommets, il ne faut garder que celles pour lesquelles les six sommets sont sur six diagonales différentes. L'idée est de n'avoir les sommets que sur les deux dernières colonnes.

Proposition 4.37 *Soit k un entier impair non multiple de 3 et soient $r = 2k^2+1$, et Δ tel que $r + \Delta = \sqrt{r^2+2r-3}$.*

Alors l'ensemble $\mathcal{S}_{r,\Delta}((0,0),(-1,0))$ contient six éléments et intersecte toutes les lignes diagonales modulo 6.

Si k est choisi est assez grand, la densité optimale d'un code (r,Δ)-identifiant est $\frac{1}{6}$.

Preuve : Comme précédemment, nous notons

$$X_1 = \mathcal{S}_{r,\Delta}((0,0),(-1,0)) \cap \{(x,y) | x > 0, y \geq 0\}.$$

Par le calcul, nous pouvons vérifier que les seuls sommets de X_1 sont $(r,0)$ et $(r-1,2k)$. Cela implique que $|\mathcal{S}_{r,\Delta}((0,0),(-1,0))| = 6$ et que les six sommets de $\mathcal{S}_{r,\Delta}((0,0),(-1,0))$ sont $(r,0)$, $(-r-1,0)$, $(r-1,2k)$, $(r-1,-2k)$, $(-r,2k)$, $(-r,-2k)$. Les six valeurs $-r$, $r+1$, $2k-r+1$, $-2k-r+1$, $2k+r$, $-2k+r$ sont toutes différentes modulo 6 donc $\mathcal{S}_{r,\Delta}((0,0),(-1,0))$ intersecte toutes les diagonales modulo 6. Nous pouvons conclure en utilisant la proposition 4.33. □

Nous procédons de la même manière pour huit sommets. Ici il est plus compliqué de vérifier que notre famille vérifie les conditions demandées aussi nous détaillons la preuve.

Proposition 4.38 *Soit $k \geq 18$ tel que $k = 2 \bmod 16$ et soit $L = \left(\frac{k}{2}\right)^2 - 1$. Posons $r = \sqrt{L^2+8}$ et $r + \Delta = \sqrt{L^2+2L+4}$.*

Alors l'ensemble $\mathcal{S}_{r,\Delta}((0,0),(-1,0))$ contient huit éléments et intersecte toutes les lignes diagonales modulo 8.

Si k est choisi assez grand, la densité optimale d'un code (r,Δ)-identifiant est $\frac{1}{8}$.

Preuve : Nous montrons tout d'abord que les seuls sommets de

$$X_1 = \mathcal{S}_{r,\Delta}((0,0),(-1,0)) \cap \{(x,y)|x > 0, y \geq 0\}$$

sont $(L,2)$ et $(L-2,k)$. Par le calcul, ces deux sommets sont des éléments de X_1, et ce sont les seuls d'abscisse L et d'abscisse plus petite que $L-2$ (comprise).

Il reste à montrer qu'il n'y a pas de sommets d'abscisse $L-1$. Supposons pour cela qu'il y a un sommet $(L-1,y)$ dans X_1. Nous avons nécessairement $y^2 \leq 2L+7$. Mais $2L \equiv 0 \bmod 16$, et les restes $5,6,7$ modulo 16 ne peuvent pas être des carrés. Donc $2L+5$, $2L+6$ et $2L+7$ ne sont pas carrés d'un entier, et nous avons donc $y^2 \leq 2L+4$. Alors $(x+1)^2+y^2 \leq (r+\Delta)^2$, et $(L-1,y)$ n'est pas dans X_1, menant à une contradiction.

Un calcul simple permet de voir que les huit sommets de $\mathcal{S}_{r,\Delta}((0,0),(-1,0))$ sont sur des lignes diagonales différentes modulo 8 et nous finissons la preuve en utilisant encore une fois le proposition 4.33. □

Comme nous venons de le voir, plus $\mathcal{S}_{r,\Delta}((0,0),(-1,0))$ contient d'éléments, plus il va être difficile de construire des familles explicites qui satisfont toutes les contraintes. Par exemple pour le cas de huit éléments, il n'existe pas de famille où les deux éléments de X_1 sont sur les deux dernières colonnes : cela crée automatiquement deux sommets sur la même diagonale modulo 8. Nous devons prendre des sommets sur l'avant-avant-dernière colonne, mais il faut alors assurer qu'il n'y ait pas de sommets sur les colonnes intermédiaires. Il doit sûrement être possible de créer d'autres familles de ce type pour plus de sommets dans $\mathcal{S}_{r,\Delta}((0,0),(-1,0))$ mais cela se complique à chaque fois et nous n'avons pas de méthode globale pour construire ce type de famille.

4.2.8 Étude des petites valeurs de r

Nous finissons ce chapitre par l'étude des petites valeurs de r. Le tableau 4.1 résume les résultats obtenus. Pour trouver des bornes supérieures, nous exhibons des codes (r,Δ)-identifiants. Pour obtenir des bornes inférieures, nous utilisons la proposition 4.28 qui nous donne une borne triviale en comptant le nombre de sommets du motif horizontal. Nous pouvons dans certains cas améliorer cette borne en utilisant une fenêtre X et des techniques de déchargement comme nous l'avons fait dans la partie 3.4 au chapitre 3.

Cas $(r,\Delta) = (\sqrt{2}, 2-\sqrt{2})$

Nous avons représenté dans les figures 4.21a et 4.21b, les motifs horizontal et diagonal pour les valeurs $(r,\Delta) = (\sqrt{2}, 2-\sqrt{2})$. D'après la proposition 4.28, nous avons donc $D(\sqrt{2}, 2-\sqrt{2}) \geq \frac{1}{4}$. Nous pouvons améliorer cette borne en utilisant des techniques de déchargement sur une fenêtre particulière. Nous donnons d'abord une amélioration faible qui illustre simplement la méthode utilisée.

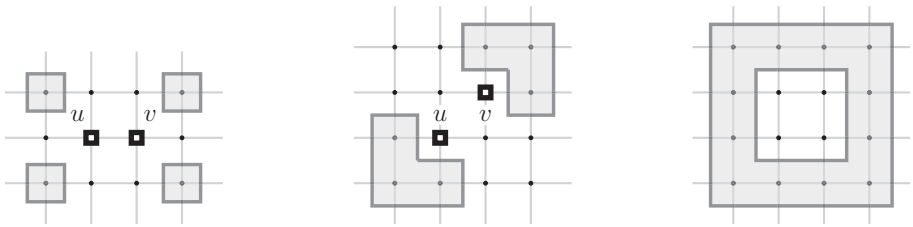
Proposition 4.39

$$D(\sqrt{2}, 2-\sqrt{2}) \geq \frac{4}{15}.$$

$r \backslash r+\Delta$	$\sqrt{2}$	2	$\sqrt{5}$	$\sqrt{8}$	3	$\sqrt{10}$
1	$0.5^{a,b}$	X	X	X	X	X
$\sqrt{2}$	–	$[\frac{16}{57}^{c}, \frac{1}{3}^{d}]$	X	X	X	X
2	–	–	$0.5^{a,b}$	$0.5^{a,b}$	X	X
$\sqrt{5}$	–	–	–	$[0.17^{e}, \frac{2}{9}^{f}]$	$[0.25^{a}, \frac{1}{3}^{g}]$	X
$\sqrt{8}$	–	–	–	–	$[\frac{1}{7}^{h}, \frac{4}{21}^{i}]$	$[\frac{1}{4}^{a}, \frac{3}{8}^{j}]$

- : Pas de sens car $\Delta \leq 0$
X : Pas de code (r, Δ)-identifiant
a : Proposition 4.28
b : Code de la proposition 4.34
c : Proposition 4.40
d : Code de la figure 4.23
e : Proposition 4.41
f : Code de la figure 4.25
g : Code de la proposition 4.32
h : Proposition 4.42
i : Code de la figure 4.27
j : Code de la proposition 4.35

Table 4.1 – *Bornes sur $D(r, \Delta)$ pour $r < 3$.*



(a) **(b)** **(c)**

Figure 4.21 – *En (a), l'ensemble $\mathcal{S}_{\sqrt{2},2-\sqrt{2}}((0,0),(-1,0))$, en (b), l'ensemble $\mathcal{S}_{\sqrt{2},2-\sqrt{2}}((0,0),(-1,-1))$ et en (c) la fenêtre utilisée pour augmenter la borne inférieure, dans le cas $(r, \Delta) = (\sqrt{2}, 2-\sqrt{2})$.*

Preuve : Pour montrer cette borne inférieure, nous utilisons l'ensemble de la figure 4.21c [1]. Toute translation de cet ensemble est appelé *fenêtre* dans cette preuve. Soit C un code $(\sqrt{2}, 2-\sqrt{2})$-identifiant. Une k-fenêtre (resp. k^+-fenêtre) est une fenêtre contenant exactement k (resp. au moins k) éléments du code.

Chaque fenêtre est en fait une 3^+-fenêtre. En effet, une fenêtre contient deux motifs horizontaux disjoints, donc doit contenir au moins deux éléments du code. Si une fenêtre contenait exactement deux éléments du code, alors l'un d'eux serait dans un coin, sinon un motif vertical ne sera pas intersecté. Le deuxième élément du code sera obligatoirement dans le coin opposé et un motif diagonal ne sera pas intersecté.

Il y a, à symétrie près, seulement cinq configurations possibles pour la position des trois éléments du code dans une 3-fenêtre, ces configurations sont décrites dans la figure 4.22. Pour chaque 3-fenêtre F, une des fenêtres voisines de F, c'est-à-dire

1. Notons que c'est la même fenêtre qui est utilisée pour la grille du roi dans la partie 3.4. C'est d'ailleurs l'étude des codes identifiants tolérants et de ce cas précis qui a donné l'idée à Tero Laihonen d'utiliser cette fenêtre pour la grille du roi.

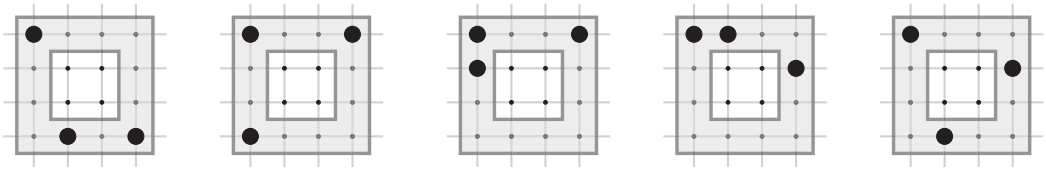


Figure 4.22 – *Dans un code $(\sqrt{2}, 2-\sqrt{2})$-identifiant, il y a seulement cinq configurations possibles lorsque la fenêtre de la figure 4.21c ne contient que trois éléments du code. Pour chacune de ces configurations, il y a une fenêtre voisine qui contient au moins quatre éléments du code.*

une fenêtre $F+u$ avec $d_{\mathcal{E}}((0,0),u) = 1$, contient au moins quatre éléments du code : celle au dessus dans le cas (c) et celle à gauche dans les autres cas.

Ainsi il ne peut y avoir uniquement des 3-fenêtres et nous avons régulièrement des fenêtres avec quatre sommets. Pour pouvoir augmenter la borne $\frac{1}{4}$, nous utilisons une procédure de déchargement.

Au début, chaque k-fenêtre a la charge k. Nous appliquons ensuite la règle suivante : chaque 4^+-fenêtre donne la charge $\frac{1}{5}$ à ses 3-fenêtres voisines.

Chaque 3-fenêtre reçoit donc au moins une charge $\frac{1}{5}$ et chaque 4^+-fenêtre perd au plus la charge $\frac{4}{5}$. Après l'application de la règle, chaque fenêtre a une charge d'au moins $3 + \frac{1}{5} = \frac{16}{5}$.

Comme il n'y a pas de perte de charge et que les charges transmises restent dans un rayon borné, le nombre moyen de sommets du code dans une fenêtre $u + F$ est donc au moins $\frac{16}{5}$. D'après la proposition 4.24, comme une fenêtre F contient 12 sommets, la densité de C est au moins $\frac{16}{5} \times \frac{1}{|F|} = \frac{4}{15}$. □

En analysant plus finement le voisinage des 3-fenêtres et en utilisant des règles de déchargement plus élaborées, nous pouvons améliorer le résultat précédent. La preuve détaillée est écrite dans notre article [63].

Proposition 4.40

$$D(\sqrt{2}, 2-\sqrt{2}) \geq \frac{16}{57}.$$

Pour la borne supérieure, nous utilisons le code C de la figure 4.23 qui est de densité $\frac{1}{3}$. Pour montrer que C est un code $(\sqrt{2}, 2-\sqrt{2})$-identifiant nous n'avons besoin, dans ce cas particulier, de vérifier que les ensembles $\mathcal{S}_{\sqrt{2},2-\sqrt{2}}(u,v)$ avec $d_{\mathcal{E}}(u,v) \leq \sqrt{2}$, ainsi que le fait que ce soit un ensemble $\sqrt{2}$-dominant. En effet, pour les autres paires de sommets, $\mathcal{S}_{\sqrt{2},2-\sqrt{2}}(u,v)$ contient soit un ensemble $\mathcal{S}_{\sqrt{2},2-\sqrt{2}}(u',v')$ avec $d_{\mathcal{E}}(u',v') \leq \sqrt{2}$ soit l'ensemble $\mathcal{B}_r(u)$. Cela n'est pas vrai dans le cas général.

Cas $(r, \Delta) = (\sqrt{5}, \sqrt{8} - \sqrt{5})$

Sur les figures 4.24b et 4.24c, nous avons représenté les motifs horizontal et diagonal pour le cas $(r, \Delta) = (\sqrt{5}, \sqrt{8} - \sqrt{5})$. Comme auparavant, nous pouvons améliorer la borne triviale de $\frac{1}{6}$ en utilisant la fenêtre de la figure 4.24d.

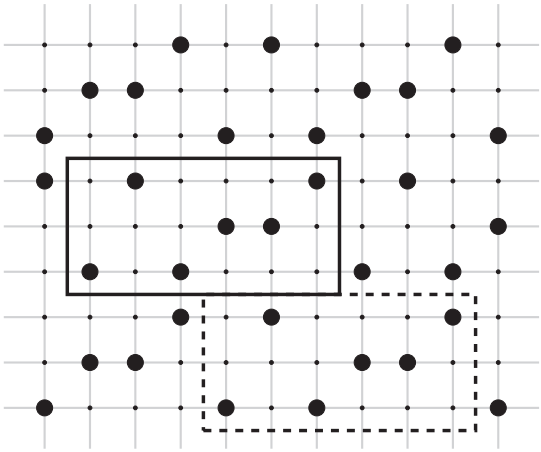


Figure 4.23 – *Un code* $(\sqrt{2}, 2-\sqrt{2})$*-identifiant de densité* $1/3$.

Proposition 4.41 *Nous avons :*

$$D(\sqrt{5}, \sqrt{8}-\sqrt{5}) \geq 0.17.$$

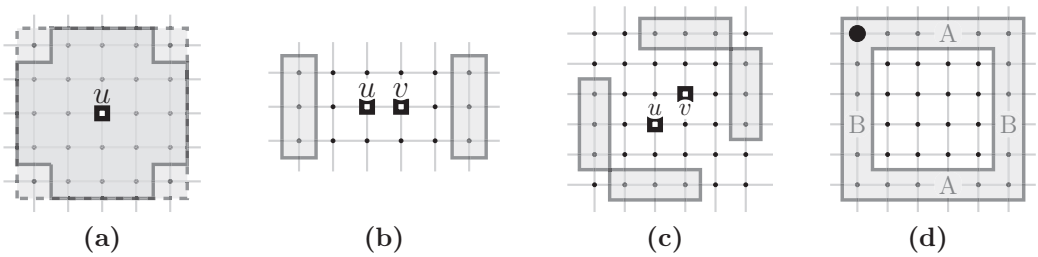


Figure 4.24 – *Sur la figure (a), les ensembles* $\mathcal{B}_r(u)$ *pour* $r = \sqrt{5}$ *et* $r = \sqrt{8}$*. Sur les figures (b) et (c), les ensembles* $\mathcal{S}_{\sqrt{5},\sqrt{8}-\sqrt{5}}(u, v)$ *lorsque* $v - u = (1, 0)$ *(b) et* $v - u = (1, 1)$*(c). Sur la figure (d), la fenêtre utilisée pour améliorer la borne inférieure dans le cas* $(r, \Delta) = (\sqrt{5}, \sqrt{8}-\sqrt{5})$.

Preuve : Nous utilisons la fenêtre F de la figure 4.24d. Nous reprenons le même vocabulaire que dans la preuve de la proposition 4.39. Dans un code $(\sqrt{5}, \sqrt{8}-\sqrt{5})$-identifiant C, il y a au moins trois éléments du code dans chaque fenêtre. Si F est une 3-fenêtre, alors nécessairement, un deséléments du code, c, est dans un coin de F. Sans perte de généralité, nous pouvons supposer que c est situé dans le coin supérieur gauche. Alors, il doit y avoir un élément du code parmi les positions $c + (3, 0)$ et $c + (3, -6)$, représentées par la lettre A sur la figure 4.24d, et de la même manière, il y a un élément du code en position $c + (0, -3)$ ou $c + (6, -3)$ (positions représentées par la lettre B sur la figure). Alors nécessairement, les fenêtres $F + (-1, 0)$ (à gauche) et $F + (0, 1)$ (en haut) contiennent sont des 4^+-fenêtres.

Cette première remarque ne permet pas d'améliorer la borne $\frac{1}{6}$. Nous allons analyser plus finement cette fenêtre pour obtenir le résultat annoncé. Soit C un code $(\sqrt{5}, \sqrt{8}-\sqrt{5})$-identifiant. Soit $\mathcal{F}_{3A}$ (resp. $\mathcal{F}_{3B}$) l'ensemble des 3-fenêtres qui ont exactement deux (resp. au moins trois) 4^+-fenêtres voisines (i.et. à distance 1).

Nous avons les faits suivant :

1. Une 4-fenêtre ne peut pas avoir ses quatre voisines dans $\mathcal{F}_{3A}$,
2. Une 4-fenêtre ne peut pas avoir trois voisines dans $\mathcal{F}_{3A}$ et une dans $\mathcal{F}_{3B}$.
3. Une 3-fenêtre de $\mathcal{F}_{3B}$ a, soit une 4^+-fenêtre voisine qui elle-même a une 4^+-fenêtre voisine, soit ses quatre voisines sont des 4^+-fenêtres.

Montrons par exemple le premier fait. Supposons qu'il existe une 4-fenêtre F avec ses quatre voisines dans $\mathcal{F}_{3A}$. Parmi les quatre fenêtres $F + v$ avec $d_{\mathcal{E}}(0, v) = \sqrt{2}$ (fenêtres en diagonal), exactement deux d'entre elles sont des 4^+-fenêtres, et sont nécessairement en diagonal. Sans perte de généralité, supposons que $F + (1, 1)$ et $F + (-1, -1)$ sont des 4^+-fenêtres. Alors les quatre éléments du code qui sont dans F sont fixés. En effet, chaque 3-fenêtre voisine de F contient un élément du code dans un de ses coins, qui est déterminé par la position de ses deux 4^+-fenêtres voisines. Cet élément du code est nécessairement dans $u + F$. Un motif diagonal n'est alors pas capturé par le code.

Les deux autres faits sont montrés de la même manière.

Nous effectuons maintenant une procédure de déchargement. Chaque k-fenêtre reçoit au départ la charge k. Puis chaque 4^+-fenêtre donne la charge $\frac{1}{5}$ à chaque 3-fenêtre voisine dans $\mathcal{F}_{3A}$ et pour chaque 3-fenêtre voisine dans $\mathcal{F}_{3B}$, elle donne :

– $\frac{1}{10}$ si F n'a pas de 4^+-fenêtre voisine,
– $\frac{1}{5}$ sinon.

A la fin de cette procédure, chaque 4^+-fenêtre aura donné au plus $\frac{3}{5}$ si elle contient quatre éléments du code et au plus $\frac{4}{5}$ sinon. Chaque 3-fenêtre aura reçu au moins la charge $\frac{2}{5}$. Ainsi, chaque fenêtre aura une charge d'au moins $3 + \frac{2}{5} = \frac{17}{5}$. Dans la fenêtre, il y a donc en moyenne au moins $\frac{17}{5}$ éléments du code. Comme une fenêtre contient 20 sommets, la proposition 4.24 nous permet de conclure que $D(C) \geq \frac{17}{100} = 0.17$. □

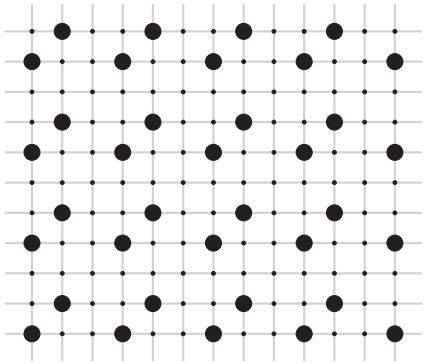


Figure 4.25 – *Un code* $(\sqrt{5}, \sqrt{8} - \sqrt{5})$*-identifiant de densité* $\frac{2}{9}$.

Nous pensons que cette borne inférieure n'est pas optimale. En poussant plus loin l'analyse effectuée lors de la précédente preuve, nous pensons qu'il est possible d'améliorer cette borne, mais le gain sera sans doute minime.

Pour la borne supérieure, nous utilisons le code de la figure 4.25 de densité $\frac{2}{9}$. Ce code est clairement un ensemble $\sqrt{5}$-dominant, et même un ensemble 2-dominant.

Il intersecte les motifs horizontaux et diagonaux ainsi que leurs rotations. De plus, si u et v sont deux sommets distincts de $\mathbb{Z}^2$, nous pouvons observer que l'ensemble $\mathcal{S}_{\sqrt{5},\sqrt{8}-\sqrt{5}}(u,v)$ contient soit une boule de rayon 2, soit un motif horizontal ou diagonal, ce qui montre que le code de la figure est un code $(\sqrt{5},\sqrt{8}-\sqrt{5})$-identifiant.

Cas $(r,\Delta)=(\sqrt{8},3-\sqrt{8})$

Les figures 4.26b et 4.26c représentent les motifs diagonal et horizontal pour le cas $(r,\Delta)=(\sqrt{8},3-\sqrt{8})$. Nous pouvons améliorer la borne inférieure $\frac{1}{8}$ en utilisant la fenêtre de la figure 4.26d.

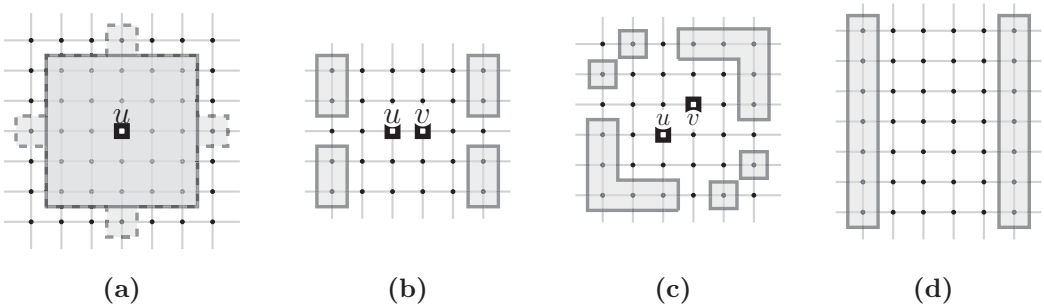


Figure 4.26 – *Sur la figure (a), les ensembles $\mathcal{B}_r(u)$ pour $r=\sqrt{8}$ et $r+\Delta=3$. Sur les figures (b) et (c), l'ensemble et $\mathcal{S}_{\sqrt{8},3-\sqrt{8}}(u,v)$ lorsque (b) $v-u=(1,0)$ et (c) $v-u=(1,1)$. Sur la figure (d), la fenêtre utilisée pour améliorer la borne inférieure dans le cas $(r,\Delta)=(\sqrt{8},3-\sqrt{8})$.*

Proposition 4.42 *Nous avons :*

$$D(\sqrt{8},3-\sqrt{8})\geq\frac{1}{7}.$$

Preuve : Soit F la fenêtre de la figure 4.26d et soit C un code $(\sqrt{8},3-\sqrt{8})$-identifiant. Toute fenêtre contient au moins deux éléments du code. La densité de C est donc d'au moins $\frac{1}{7}$. □

Le code de la figure 4.27 est un code $(\sqrt{8},3-\sqrt{8})$-identifiant de densité $\frac{4}{21}$. Dans ce cas-ci, vérifier si les motifs horizontaux et diagonaux intersectent le code n'est pas suffisant. En effet, lorsque (u,v) vérifie $v=u+(1,2)$, l'ensemble $\mathcal{S}_{\sqrt{8},3-\sqrt{8}}(u,v)$ ne contient aucun motif diagonal, horizontal, vertical ni aucune boule de rayon $\sqrt{8}$. La vérification se fait à la main.

4.2.9 Conclusion sur les codes identifiants tolérants

Dans cette partie, nous avons étudié des codes identifiants dans le plan euclidien qui reste identifiant lorsque le voisinage des éléments du code peut varier légèrement sur les bords. Nous rappelons les principaux résultats obtenus :

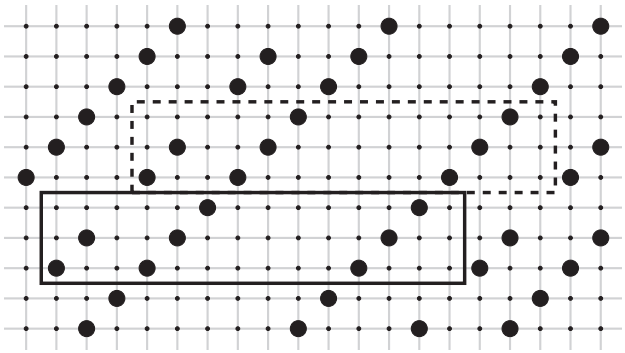


Figure 4.27 – *Code de densité* $4/21$ *pour* $(r, \Delta) = (\sqrt{8}, 3 - \sqrt{8})$.

- La densité d'un code (r, Δ)-identifiant tolérant optimal est d'ordre $\frac{1}{r}$ lorsque $r \to +\infty$ et que Δ est fixé.
- Lorsque Δ varie avec r en tendant vers 1 ($\Delta = 1 - \frac{1}{\sqrt{r}}$), la densité optimale d'un code (r, Δ)-identifiant est au moins d'ordre $\frac{1}{\sqrt{r}}$.
- Pour des valeurs fixées de r et Δ, nous donnons des méthodes pour construire de meilleurs codes identifiants que ceux donnés par la borne générale.
- Il existe des familles infinies de valeurs (r, Δ) pour lesquelles nous connaissons des constructions optimales.
- Les cas avec $r \leq \sqrt{8}$ ont été étudiés plus en détail, en utilisant d'autres techniques, notamment des techniques de déchargement. Les résultats sont plus précis. Le tableau 4.1 (page 155) résume les résultats obtenus pour les petites valeurs.

4.3 Conclusions et perspectives du chapitre

Nous avons étudié dans ce chapitre des problèmes d'identification où les capteurs ont plus de flexibilité. La première variante introduite dans ce sens est une généralisation des ensembles résolvants. Elle peut être formalisée dans le cadre suivant qui englobe plusieurs problèmes d'identification.

Soient p un entier et $\mathcal{R}$ un ensemble d'entiers positifs. Un *code* $(p, \mathcal{R})$*-identifiant* d'un graphe G est un sous-ensemble C de sommets tels que :

$$\text{(domination)} \quad \forall x \in V(G), \exists r \in \mathcal{R}, \mathcal{B}_r(x) \cap C \neq \emptyset$$

$$\text{(séparation)} \quad \begin{cases} \forall x \in V(G), \exists R_x \subset \mathcal{R}, |R_x| \leq p, \forall y \in V(G), y \neq x, \exists r_{xy} \in R_x \text{ t.q.} \\ \mathcal{B}_{r_{xy}}(x) \cap C \neq \mathcal{B}_{r_{xy}}(y) \cap C \end{cases}$$

Intuitivement, l'entier p correspond au nombre de rayons que l'on peut donner à chaque sommet, et l'ensemble $\mathcal{R}$ correspond aux rayons disponibles. Cette définition unifie quelques-unes des notions sur les codes identifiants : un code r-identifiant est un code $(1, \{r\})$-identifiant, un code r-identifiant évolutif sans mémoire est un code $(1, [\![0, r]\!])$-identifiant, un code r-identifiant évolutif avec mémoire est un code $(r + 1, [\![0, r]\!])$-identifiant, un code r-localisateur dominateur est un code $(2, \{0, r\})$-identifiant, un ensemble résolvant est un code $(\infty, \mathbb{N})$-identifiant.

Nous avons déterminé la taille optimale des codes $(p, [\![0, r]\!])$-identifiants sur les cycles. En effet, la proposition 4.18 revient à dire que tout code $(p, [\![0, r]\!])$-identifiant d'un cycle avec $p \geq 3$ est en fait un code $(3, [\![0, r]\!])$-identifiant. Le problème lorsque $\mathcal{R}$ est quelconque reste ouvert.

Par ailleurs, d'autres questions nous semblent intéressantes à étudier sur ces problèmes :

- Étudier la complexité du problème pour des restrictions sur p ou $\mathcal{R}$.
- Étudier le problème sur d'autres classes de graphes. L'étude sur les cycles des codes $(p, [\![0, r]\!])$-identifiants semble plus facile que pour les codes identifiants classiques. Il serait intéressant d'étudier ces notions sur des classes de graphe pour lesquelles le problème des codes identifiants n'est pas complètement résolu, comme l'hypercube.

La deuxième variante utilise fortement une métrique continue donnée par la distance euclidienne du plan. Cela nous permet de considérer des voisinages avec un rayon approximatif. Avec une métrique discrète comme celle de Hamming, le problème n'aurait pas de sens car deux sommets à distance 1 ne pourraient être séparés. La plupart des résultats obtenus sont partiels et nous ne connaissons en général pas les densités optimales des codes étudiés, ce qui laisse des perspectives de recherche. Par ailleurs, ce problème pourrait être étendu sur d'autres réseaux qui se plongent dans le plan, comme le réseau hexagonal, ou bien dans d'autres dimensions.

Chapitre 5

Coloration identifiante

Sommaire

5.1 Coloration globalement identifiante 164
5.1.1 Comment identifier avec des couleurs ? 164
5.1.2 Application pratique 165
5.1.3 Lien avec d'autres colorations 165
5.1.4 Lien avec les codes identifiants 165
5.1.5 Bornes générales 167
5.1.6 Cycles et lien avec les cycles universels 171
5.1.7 Conclusion sur les colorations identifiantes globales 174
5.2 Du global au local 174
5.3 Bornes générales 176
5.3.1 En fonction du nombre de sommets 176
5.3.2 En fonction du nombre chromatique 177
5.3.3 En fonction du degré maximum 180
5.4 Étude dans les graphes bipartis et complexité 186
5.4.1 Motivations 186
5.4.2 Coloration en quatre couleurs des graphes bipartis 187
5.4.3 Lien avec la coloration d'hypergraphe et complexité . . . 188
5.4.4 Cas des arbres 190
5.4.5 Produit cartésien de graphes bipartis 191
5.5 Étude dans les graphes parfaits 191
5.5.1 k-arbres 192
5.5.2 Graphes d'intervalles 193
5.5.3 Graphes scindés 195
5.5.4 Cographes 197
5.5.5 Bilan sur les graphes parfaits et conjecture 199
5.6 Étude dans les graphes planaires 201
5.6.1 Graphes planaires extérieurs 201
5.6.2 Graphes planaires de grande maille 204
5.7 Conclusion sur les colorations identifiantes 205

Dans ce chapitre, nous nous intéressons aux manières d'identifier un graphe en coloriant ses sommets. Nous avons considéré les deux cas suivant : tous les sommets doivent être séparés (coloration globalement identifiante) ou bien seulement les sommets adjacents (coloration localement identifiante).

Nous nous sommes principalement penchés sur le second cas, qui a fait l'objet de deux articles [35, 39] avec respectivement Louis Esperet, Sylvain Gravier, Mickaël Montassier, Pascal Ochem et Florent Foucaud, Iiro Honkala, Tero Laihonen, Guillem Perarnau. Les résultats présentés dans le premier cas sont en cours de rédaction et ont été obtenus en grande partie lors de la semaine discrète de l'Institut Fourier en 2009.

5.1 Coloration globalement identifiante

5.1.1 Comment identifier avec des couleurs ?

Lors de la semaine discrète de l'Institut Fourier de 2009, Eric Duchêne puis Julien Moncel ont posé la question suivante : que se passe-t-il si nous essayons d'identifier un graphe non pas avec un sous-ensemble de sommets mais avec des couleurs ?

Une manière naturelle d'identifier un graphe avec des couleurs est de donner des couleurs à chaque sommet et de demander à ce que chaque sommet ait un ensemble unique de couleurs dans son voisinage fermé.

Plus formellement, donner des couleurs aux sommets d'un graphe G revient à définir une application $c : V(G) \to \mathbb{N}$ appelée *coloration*. Pour un sous-ensemble X de sommets, nous notons alors $c(X)$ l'ensemble des couleurs apparaissant dans X : $c(X) = \{c(u)|u \in X\}$. Une coloration c est dite *identifiante* si pour toute paire de sommets u et v, nous avons $c(N[u]) \neq c(N[v])$. La figure 5.1 donne un exemple de coloration identifiante.

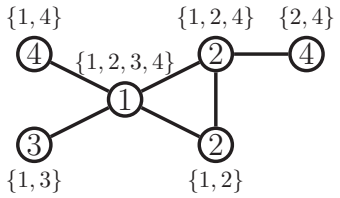


Figure 5.1 – *Une coloration identifiante d'un graphe en quatre couleurs. L'ensemble $c(N[v])$ est indiqué pour chaque sommet v.*

Remarquons que nous ne demandons pas que la coloration soit propre, c'est-à-dire à ce que deux sommets adjacents reçoivent des couleurs différentes.

Pour chaque sommet u, l'ensemble $c(N[u])$ est unique. Chaque sommet u est donc *identifié* par l'ensemble $c(N[u])$. Nous dirons aussi que u *réalise* l'ensemble de couleurs $c(N[u])$. Une couleur *sépare* deux sommets u et v si elle appartient à l'ensemble $c(N[u])\Delta c(N[v])$.

Comme pour les codes identifiants où nous cherchons à minimiser la taille du code, ici nous voulons minimiser le nombre de couleurs nécessaire pour avoir une coloration identifiante. Le *nombre chromatique identifiant* d'un graphe G, noté $\chi_{id}(G)$,

est le nombre minimum de couleurs nécessaire dans une coloration identifiante de G. Pour le graphe G de la figure 5.1, la coloration identifiante n'utilise que quatre couleurs et nous avons donc $\chi_{id}(G) \leq 4$. En étudiant le graphe, nous pouvons voir qu'il n'existe pas de coloration identifiante en trois couleurs, nous avons donc $\chi_{id}(G) = 4$.

5.1.2 Application pratique

Les colorations identifiantes correspondent à une identification par une partition des sommets du graphe. Elles se généralisent dans le cadre des graphes bipartis. Soit $G = (S \cup T, A)$ un graphe biparti où les sommets à identifier sont les sommets de S et les éléments permettant d'identifier sont ceux de T. Alors une coloration identifiante devient une coloration des sommets de T tels que chaque sommet de S ait dans son voisinage un ensemble de couleurs uniques.

Cela prend un sens pratique lorsque les sommets de S correspondent à des maladies et les sommets de T à des syndromes. Chercher un code identifiant permettait de choisir un sous-ensemble de symptômes efficaces pour différencier chaque maladie. Une coloration identifiante permet de regrouper certains symptômes ensemble. Par exemple, on pourrait ainsi regrouper « avoir de la fièvre ou ressentir de la fatigue ». Notons tout de même qu'avec une coloration identifiante, qui est une partition des sommets, nous ne permettons pas de tester un même symptôme dans plusieurs groupes. Ce serait une autre définition à étudier correspondant à un procédé d'identification avec des sous-ensembles.

5.1.3 Lien avec d'autres colorations

Des colorations du même type ont déjà été étudiées auparavant. Harary et Plantholt [51] ont par exemple considéré des colorations propres d'arêtes telles que pour toutes paires de sommets u et v, l'ensemble des couleurs des arêtes incidentes à u soit différent de l'ensemble des couleurs des arêtes incidentes à v. Une telle coloration est appelée *coloration propre des arêtes distinguant les sommets* (vertex-distinguishing edge-coloring en anglais). Le nombre minimum de couleurs nécessaire dans une telle coloration est l'*observabilité* du graphe et a été étudié largement.

Cette notion a été étendue par Zhang *et al.* [97] en une variante locale où seules les paires de sommets adjacents doivent être distingués. Zhang *et al.* [96] ont introduit une variante où la coloration utilisée est une coloration totale (sommets et arêtes). De manière étonnante, le cas où la coloration est uniquement une coloration des sommets n'a pas encore été étudié. Dans les parties 5.2 et suivante, nous considérerons une variante locale des colorations identifiantes.

5.1.4 Lien avec les codes identifiants

Tous les graphes n'admettent pas de coloration identifiante. Comme pour les codes identifiants, il est nécessaire d'avoir un graphe sans jumeaux. Cela est aussi suffisant car si le graphe est sans jumeaux et que tous les sommets du graphe ont des couleurs différentes alors nous obtenons une coloration identifiante. Une manière d'éviter cela serait de définir des colorations identifiantes pour tout graphe G est de

demander à ce que $c(N[u]) \neq c(N[v])$ seulement si $N[u] \neq N[v]$. Nous utiliserons cette astuce pour la variante locale.

Par ailleurs, nous pouvons construire à partir d'un code identifiant d'un graphe G une coloration identifiante et obtenir une borne sur le nombre chromatique identifiant de G :

Proposition 5.1 *Soit G un graphe sans jumeaux, alors :*

$$\chi_{id}(G) \leq \gamma^{ID}(G) + 1.$$

Preuve : Soit C un code identifiant de G de taille minimale $\gamma^{\text{ID}}(G)$. Nous construisons une coloration identifiante de G en donnant des couleurs différentes allant de 1 à $\gamma^{\text{ID}}(G)$ aux sommets de C et une nouvelle couleur $\gamma^{\text{ID}}(G)+1$ aux autres sommets. La figure 5.2 donne un exemple de construction. Cette coloration est identifiante car le sous-ensemble de couleurs dans le voisinage fermé d'un sommet donne directement le sous-ensemble identifiant du code. □

Figure 5.2 – *Construction d'une coloration identifiante avec $k+1$ couleurs à partir d'un code identifiant de taille k.*

Remarquons que le fait de ne pas imposer une coloration propre nous permet d'avoir cette borne. Si l'on voulait une coloration identifiante propre de G, il faudrait prendre une coloration propre des sommets qui ne sont pas dans le code identifiant et la borne deviendrait $\chi_{id}(G) \leq \gamma^{\text{ID}}(G) + \chi(G)$.

Nous verrons dans le paragraphe suivant que la borne de la proposition 5.1 est serrée, notamment par les graphes qui nécessitent autant de couleurs que de sommets dans une coloration identifiante.

Nous donnons maintenant une manière – moins naturelle – pour obtenir un code identifiant à partir d'une coloration identifiante. Supposons que nous ayons une coloration identifiante c d'un graphe G avec k couleurs (numérotées de 1 à k). Nous construisons un graphe G' obtenu à partir de G en ajoutant $k+1$ sommets $v_0, v_1, ...v_k$ de la manière suivante :

- v_0 est adjacent à tous les sommets de G,
- v_i est adjacent à tous les sommets de G de couleur i.

Alors l'ensemble $\{v_0, ...v_k\}$ est un code identifiant de G' de taille $k+1$. Notons que si nous n'ajoutons pas le sommet v_0, nous obtenons un code localisateur-dominateur.

Cette construction ne donne pas de borne inférieure générale de $\chi_{id}(G)$ par $\gamma^{\text{ID}}(G)$. Nous verrons par la suite qu'il existe, pour tout entier k, des graphes (les cycles par exemple) dont le nombre chromatique identifiant est de l'ordre de $\gamma^{\text{ID}}(G)^{1/k}$. Le plus grand écart possible entre $\chi_{id}(G)$ et $\gamma^{\text{ID}}(G)$ est théoriquement un facteur logarithmique, nous ne savons pas si des graphes réalisant cet écart existent.

Question 5.2 *Existe-t-il des familles infinies de graphes G pour lesquelles $\chi_{id}(G) = O(\log(\gamma^{ID}(G)))$?*

5.1.5 Bornes générales

Proposition 5.3 *Soit G un graphe sans jumeaux avec n sommets, alors :*

$$\lceil \log(n+1) \rceil \leq \chi_{id}(G) \leq n.$$

Preuve : Soit c une coloration identifiante utilisant k couleurs. Les ensembles identifiants $c(N[x])$ pour chaque sommet x de G sont des sous-ensembles non vide et distincts de l'ensemble $\{1, ..., k\}$. Nous avons donc $n \leq 2^k - 1$ ce qui induit la borne inférieure de la proposition. La borne supérieure est triviale. □

Les deux lemmes suivants montrent que les bornes de la proposition précédente sont serrées.

Lemme 5.4 *Pour tout entier $k \geq 4$, il existe un graphe G_k avec $2^k - 1$ sommets tel que $\chi_{id}(G_k) = k$.*

Preuve : Nous raisonnons par récurrence sur l'entier $k \geq 4$ en montrant le résultat plus précis suivant :

Pour tout entier $k \geq 4$, il existe un graphe G_k avec $2^k - 1$ sommets qui possède une coloration identifiante c_k avec k couleurs qui vérifie la propriété suivante : soit x_k l'unique sommet tel que $c_k(N[x_k]) = \{1, ..., k\}$, alors il existe un autre sommet de couleur $c_k(x_k)$.

Pour $k = 4$, nous considérons le graphe G_4 de la figure 5.3. Sur la figure, une coloration identifiante en quatre couleurs est indiquée. Le sommet x_4 est l'unique sommet dont toutes les couleurs sont présentes dans son voisinage fermé. Remarquons qu'il y a un autre sommet de couleur 1.

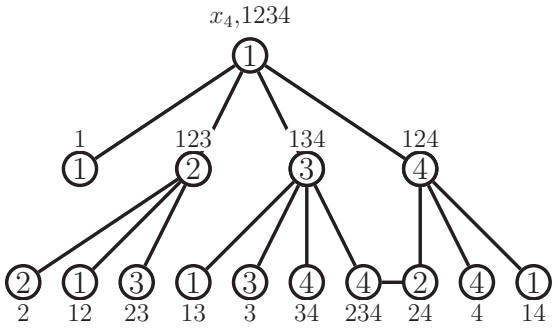


Figure 5.3 – *Le graphe G_4 avec 15 sommets et $\chi_{id}(G_4) = 4$. Les ensembles $c(N[u])$ sont indiqués pour chaque sommet, sous forme d'un mot pour plus de lisibilité.*

Soit maintenant $k > 4$. Supposons que le graphe G_{k-1} est construit et que nous disposons d'une coloration c_{k-1} identifiante en $k-1$ couleurs de G_{k-1}. Un des sommets, notons le x_{k-1}, a nécessairement toutes les couleurs dans son voisinage.

Le graphe G_k est construit à partir du graphe G_{k-1} de la manière suivante (voir la figure 5.4).

- On prend deux copies du graphe G_{k-1} dont les sommets sont notés V_1 et V_2 dont on unifie les deux sommets x_{k-1}.
- On crée un nouveau sommet x_k relié à tous les sommets de V_1 sauf x_{k-1}.
- On ajoute un nouveau sommet u relié seulement à x_k.

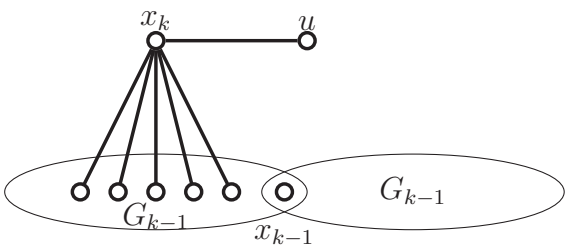


Figure 5.4 – *Construction du graphe G_k à partir du graphe G_{k-1}.*

Clairement, le graphe G_k ainsi créé possède $2^k - 1$ sommets. Nous définissons une coloration c_k de G_k de la manière suivante :

- Dans chaque copie de G_{k-1}, les sommets sont coloriés avec les couleurs 1 à $k-1$ selon la coloration identifiante c_{k-1} de G_{k-1}.
- Les sommets u et x_k sont coloriés avec la couleur k.

Comme G_k contient $2^k - 1$ sommets, pour montrer que c_k est une coloration identifiante, il suffit de montrer que tous les sous-ensembles non vides de $\{1, ..., k\}$ sont réalisés. Soit S un tel sous-ensemble. Supposons d'abord que $k \notin S$. Alors il existe un sommet v dans G_{k-1} tel que $c_{k-1}(N[v]) = S$ (dans G_{k-1}). Soit v' le sommet correspondant de V_2. Le sommet v' n'est pas relié à x_k donc $c_k(N[v']) = S$ (dans G_k).

Supposons maintenant que $k \in S$. Si $S = \{k\}$ alors S est réalisé par le sommet u. Si $S = \{1, ..., k\}$, S est réalisé par x_k car toutes les couleurs $\{1, ..., k-1\}$ sont présentes dans la coloration c_{k-1} restreinte aux sommets différents de x_{k-1} (par hypothèse de récurrence, la couleur de x_{k-1} est aussi présente sur un autre sommet de V_1).

Sinon, posons $S' = S - \{k\}$. Alors il existe un sommet v dans G_{k-1} tel que $c_{k-1}(N[v]) = S'$ (dans G_{k-1}). Soit v' le sommet correspondant dans V_1. Le sommet v' n'est pas le sommet x_{k-1} car $S' \neq \{1, ..., k-1\}$, il est donc adjacent à v_k et donc $c_k(N[v']) = S' \cup \{k\} = S$, ce qui termine la preuve. □

Notons qu'il n'existe pas de tel graphe pour $k \leq 3$. Supposons par exemple qu'un tel graphe G_3 existe pour $k = 3$ et soit c_3 une coloration identifiante avec trois couleurs de G_3. Alors un sommet, appelons-le x, a les trois couleurs dans son voisinage. Sans perte de généralité, supposons que sa couleur soit 1. L'ensemble $\{2, 3\}$ doit être réalisé, supposons qu'il le soit par le sommet y et que la couleur de y soit 2. Le sommet y a un voisin z de couleur 3. Mais alors $z \neq x$ et $\{2, 3\} \subseteq c(N[z])$, donc z ne peut être séparé de y et x à la fois.

Nous montrons maintenant que la borne supérieure de la proposition 5.3 est serrée.

Lemme 5.5 *Soit G un graphe complet privé d'un couplage maximal. Alors $\chi_{id}(G) = |V(G)|$.*

Preuve : Notons n l'ordre de G. Alors le couplage M est de taille $\lfloor \frac{n}{2} \rfloor$.

Supposons qu'il existe une coloration identifiante c de G avec $n-1$ couleurs. Deux sommets u et v ont donc la même couleur. Nécessairement, uv est une arête du graphe, car autrement v serait le seul sommet qui n'est pas adjacent à u, et réciproquement. Les ensembles $c(N[u])$ et $c(N[v])$ seraient tous deux égaux à l'ensemble des couleurs utilisées et les sommets u et v ne seraient pas séparés. Sans perte de généralité, nous pouvons supposer que u n'est pas un sommet universel et notons u' le sommet qui n'est pas adjacent à u. Le sommet u' est alors adjacent à tous les autres sommets, en particulier à v, et donc $c(N[u'])$ contient toutes les couleurs utilisées. Le graphe G n'a donc pas de sommet universel (autrement il ne serait pas séparé de u'), ce qui signifie que n est pair. Entre autre, le sommet v n'est pas universel, et l'unique sommet v' qui n'est pas adjacent à v a toutes les couleurs dans son voisinage. Le sommet v' n'est donc pas séparé de u', menant à une contradiction. □

Nous allons maintenant utiliser la proposition 5.1 pour caractériser tous les graphes sans jumeaux dont le nombre de couleurs nécessaire dans une coloration identifiante est le nombre de sommets du graphe. Nous venons de voir un exemple, les graphes complets privés d'un couplage maximal, qui font aussi partie des exemples extrémaux pour les codes identifiants. Cela est en fait une conséquence directe de la proposition 5.1 et du fait que pour tout graphe G, $\gamma^{\text{ID}}(G) \leq |V(G)| - 1$ (théorème 1.4). En utilisant la caractérisation des graphes G tels que $\gamma^{\text{ID}}(G) = |V(G)| - 1$ donnée dans le chapitre 2, nous n'avons donc qu'à rechercher parmi les graphes de cette famille quels sont les graphes concernés.

Proposition 5.6 *Soit G un graphe sans jumeaux connexe. Il y a équivalence entre :*

1. *$\chi_{id}(G) = |V(G)|$;*
2. *G est un graphe complet privé d'un couplage maximum ou bien G s'écrit $G = K_1 \bowtie G_1 \bowtie ... \bowtie G_t$ avec G_i de la forme P_{2k}^{k-1}, $k \geq 2$ ou $G_i = \overline{K_2}$.*

Preuve :

(2)⇒(1) : D'après le lemme 5.5, si G est un graphe complet privé d'un couplage, alors $\chi_{id}(G) = |V(G)|$. Ce cas inclus aussi le cas $G = K_1 \bowtie \overline{K_2}$ pour lequel nous pouvons remarquer que pour chaque couleur d'un des sommets de $\overline{K_2}$, il existe une paire de sommets qui est séparée uniquement par cette couleur.

Soit $k \geq 2$ un entier et H_k le graphe $P_{2k}^{k-1} \bowtie K_1$. Soit c une coloration identifiante de H_k. Nous montrons que c utilise $2k+1 = |V(H_k)|$ couleurs. Soient $x_1, ...x_{2k}$ les sommets du chemin P_{2k} et x le sommet universel. Nous montrons par récurrence que x_i, pour $i \leq k$, a une couleur qui n'est pas utilisée par un autre sommet.

Le sommet x_1 est l'unique sommet dans la différence symétrique de $N[x_{k+1}]$ et $N[x]$. Pour que x et x_{k+1} soient séparés, il faut donc que la couleur de x_1 n'apparaissent pas dans le voisinage fermé de x_{k+1}, qui contient exactement tous les autres sommets du graphe. Donc la couleur de x_1 n'est pas utilisée par les autres sommets.

Supposons que cela soit vrai pour les sommets x_1,...,x_{i-1}, $1 < i \leq k$. Alors x_i ne peut pas prendre une des couleurs d'un des x_j, $j < i$. Or x_i est l'unique sommet dans la différence symétrique de $N[x_{i+k-1}]$ et $N[x_{i+k}]$. La couleur de x_i n'apparaît donc pas dans le voisinage de $N[x_{i+k}]$ qui comprend tous les sommets différents de x_j pour $j > i$ ainsi que le sommet x. La propriété est donc vraie pour x_i, par récurrence, elle est vraie pour tous les sommets x_1 jusqu'à x_k, et par symétrie, pour tous les sommets du chemin.

Enfin, x n'a pas une couleur d'un des sommets x_i, ce qui termine la preuve. Nous avons de plus montré que pour chaque couleur autre que celle de x, il existe une paire de sommets qui est séparée uniquement par cette couleur.

Prenons maintenant G la forme $G = K_1 \bowtie G_1 \bowtie ... \bowtie G_t$ avec G_i de la forme P_{2k}^{k-1}, $k \geq 2$ ou $G_i = \overline{K_2}$ et prenons c une coloration identifiante de G. Alors c induit une coloration identifiante sur chaque facteur du joint car deux sommets dans un même facteur ont le même voisinage hors de ce facteur. Comme l'opération $\bowtie$ est associative et commutative, c induit une coloration identifiante sur $K_1 \bowtie G_i$ pour tout i. D'après ce qui précède, au sein de chaque facteur G_i, toutes les couleurs des sommets sont donc distinctes et différentes de la couleur du sommet correspondant à K_1. Supposons qu'une couleur c_0 soit utilisée dans deux facteurs différents G_1 et G_2 par exemple. Alors il existe deux sommets u_1 et v_1 de $K_1 \bowtie G_1$ qui sont séparés uniquement par la couleur c_0 dans le graphe $K_1 \bowtie G_1$, et donc aussi dans le graphe G. Une contradiction car u_1 et v_1 sont tous les deux adjacents au sommet colorié avec c_0 dans G_2. Ainsi toutes les couleurs sont différentes et c utilise $|V(G)|$ couleurs.

(1)⇒(2) : Supposons par l'absurde qu'il existe un graphe G vérifiant $\chi_{id}(G) = |V(G)|$ qui ne soit pas dans la famille décrite par la proposition. D'après la proposition 5.1, G vérifie $\gamma^{\text{ID}}(G) = V(G) - 1$, et d'après le théorème 2.33 du chapitre 2, G est donc soit :

1. une étoile $K_{1,n}, n \geq 3$;
2. un graphe P_{2k}^{k-1}, $k \geq 2$;
3. le joint de deux graphes $P_{2k}^{k-1} \bowtie H$, avec $k \geq 2$ et H n'ayant pas de sommet universel.

Dans chacun de ces cas, nous donnons une coloration identifiante de G avec $|V(G)| - 1$ couleurs.

Supposons que $G = K_{1,n}, n \geq 3$. Soient $v_1, ...v_n$ les feuilles de G et v le sommet universel. Nous posons $c(v) = 1$ et $c(v_i) = i$ pour tout $i \in \{1, ..., n\}$. Alors $c(N[v]) = \{1, \ldots, n\}$, $c(N[v_i]) = \{1, i\}$ si $i > 1$ et $c(N[v_1]) = \{1\}$. La coloration c est donc identifiante et n'utilise que $|V(G)| - 1$ couleurs.

Supposons maintenant que $G = P_{2k}^{k-1}$ avec $k \geq 2$ et notons v_1,...v_{2k} les sommets de G dans l'ordre naturel du chemin P_{2k}. Posons $c(v_i) = i - 1$ si $i > 1$ et $c(v_1) = 1$. Alors $c(N[v_i]) = \{1, ..., i + k - 2\}$ pour $i \leq k + 1$ et $c(N[v_i]) = \{i - k, ..., 2k - 1\}$ pour $i \geq k + 2$, et c est un coloration identifiante n'utilisant que $2k - 1$ couleurs.

Enfin supposons que G s'écrit $G = P_{2k}^{k-1} \bowtie H$, avec $k \geq 2$ et H sans sommet universel. Nous colorions les sommets de P_{2k}^{k-1} sont coloriés comme au cas précédent, et les sommets de H avec des nouvelles couleurs en donnant une couleur différente par sommet. Notons c la coloration obtenue. Soient u et v deux sommets. Si u et v

sont dans P_{2k}^{k-1}, ils sont bien séparés car les couleurs ajoutées dans H ne sont pas des couleurs qui sont dans P_{2k}^{k-1}. Si u est dans P_2^{k-1} et v est dans H, alors v n'est pas universel et $c(N[v])$ ne contient pas toutes les couleurs de H contrairement à u. Enfin si u et v sont dans H, comme $N[u] \cap H \neq N[v] \cap H$ et que les couleurs de H sont toutes distinctes et non utilisées dans P_{2k}^{k-1}, les sommets sont séparés. Finalement, c est une coloration identifiante de G n'utilisant que $|V(G)|-1$ couleurs. □

5.1.6 Étude dans la classe des cycles et lien avec les cycles universels

Dans cette partie, nous étudions le problème de coloration identifiante dans la classe des cycles. Plus précisément, nous fixons le nombre de couleurs L autorisé et cherchons la taille $n(L)$ maximale d'un cycle qui possède une coloration identifiante avec L couleurs. Étant donné qu'un sommet d'un cycle n'a que deux voisins, $n(L)$ ne peut être plus grand que le nombre de sous-ensembles à moins de trois éléments parmi un ensemble à L éléments :

$$\begin{aligned} n(L) &\leq \binom{L}{3} + \binom{L}{2} + \binom{L}{1} \\ &\leq \frac{L^3 + 5L}{6} \end{aligned}$$

Nous allons montrer que l'ordre de cette borne est atteint en utilisant les cycles universels.

Cycle universel des sous-ensembles à trois éléments

Un *cycle universel des sous-ensembles à trois éléments de* $\{1, ..., L\}$ est une séquence d'entiers $(x_1, ..., x_{\binom{L}{3}})$, où chaque x_i est un entier compris entre 1 et L et où chaque sous-ensemble à trois éléments de $\{1, ..., L\}$ est représenté par un unique bloc de taille 3 : (x_i, x_{i+1}, x_{i+2}) (pris cycliquement). La figure 5.5 donne un exemple (issu de [59]) lorsque $L = 8$.

Cela correspondant exactement à une coloration identifiante c avec L couleurs du cycle de taille $\binom{L}{3}$ où pour chaque sommet u, $c(N[u])$ est de taille 3. Pour certains entiers L, il n'existe pas de cycle universel. En effet, chaque occurrence d'un entier dans le cycle correspond à trois sous-ensembles à trois éléments de $\{1, ..., L\}$, donc le nombre de sous-ensembles à trois éléments contenant un entier donné doit être divisible par 3. Ce qui revient à dire que $\binom{L-1}{2} \equiv 0 \bmod 3$, ou encore que $\text{pgcd}(L, 3) = 1$. Jackson [59] a montré dans que c'était une condition suffisante pour l'existence des cycles universels :

Théorème 5.7 (Jackson [59]) *Pour tout entier $L \geq 8$ non divisible par 3, il existe un cycle universel des sous-ensembles à trois éléments de $\{1, ..., L\}$.*

8	1	2	4	6	7	2	5	6	7	1	3	4	7	2	3	4	6	8	1	4	7
5																					8
2																					1
1																					3
7																					5
5																					6
4																					1
3	8	5	4	2	8	7	6	3	8	7	5	3	2	1	6	3	2	8	6	5	4

Figure 5.5 – *Un cycle universel des sous-ensembles à trois éléments de* $\{1, \ldots, 8\}$. *Chaque sous-ensemble à trois éléments de triplet* $\{1, \ldots, 8\}$ *correspond à une unique suite de trois chiffres (les chiffres de l'ensemble).*

En terme de colorations identifiantes, cela implique :

Corollaire 5.8 *Soit* $L \geq 8$ *un entier non divisible par* 3. *Le cycle de taille* $\binom{L}{3}$ *admet une coloration identifiante avec* L *couleurs. Par conséquent.*

$$n(L) \geq \binom{L}{3}.$$

Ainsi pour un cycle de taille n, le nombre de couleurs nécessaire est d'ordre $n^{\frac{1}{3}}$. En effet, même si la taille du cycle ne correspond pas exactement à la taille du cycle universel, il est possible de construire une coloration identifiante en prenant le premier cycle universel de taille supérieure et de n'en utiliser qu'une partie, quitte à ajouter trois couleurs pour le recollement. Par ailleurs, un code identifiant du cycle de taille n est de taille environ $\frac{n}{2}$ (voir par exemple Junnila et Laihonen [62]). Les cycles sont donc des exemples de graphes pour lesquels une coloration identifiante demande beaucoup moins de couleurs qu'un code identifiant n'a besoin de sommets. Plus précisément, pour des familles infinies de cycles, $\chi_{id}(\mathcal{C})$ est de l'ordre de $\gamma^{\text{ID}}(\mathcal{C})^{\frac{1}{3}}$.

Notons que $n(L)$ est très certainement plus grand que $\binom{L}{3}$. En effet, dans la coloration identifiante obtenue, les ensembles de taille 1 et 2 ne sont pas réalisés. Or il est assez facile d'ajouter des sommets pour obtenir ces ensembles. Par exemple, si dans le cycle universel la séquence $(1, 2, 3)$ est présente et que nous ajoutons deux sommets entre le 1 et le 2, coloriés avec la couleur 2, alors nous réaliserons les ensembles $\{1, 2\}$, $\{2\}$ et $\{2, 3\}$ en plus, mais nous ne réaliserons plus l'ensemble $\{1, 2, 3\}$. Si nous ajoutons seulement un sommet colorié 2, alors nous ne créons pas l'ensemble $\{2\}$. Il faut alors choisir où intégrer ces entiers pour réaliser toutes les paires et singletons, sans créer de doublons, et recréer les sous-ensembles à trois éléments manquants. Nous ne savons pas pour l'instant si cela est possible ou non :

Question 5.9 *Existe-t-il des entiers* L *tels que* $n(L) = \frac{L^3+5L}{6}$ *?*

Une piste pour répondre à cette question est d'utiliser la construction donnée dans la preuve du théorème 5.7. En effet, l'auteur utilise des parcours eulériens dans

des graphes orientés bien choisis, chaque arête correspondant à un sous-ensemble à trois éléments. Les cycles construit ont la propriété suivante. Soit $\{i, j, k\}$ un sous-ensemble de $\{1, ..., L\}$ tel que $i < j < k$. Alors cet ensemble apparaît dans l'ordre (i, j, k), (j, k, i) ou bien (k, i, j) de telle sorte que les différences d_1 et d_2 entre les deux premiers éléments et les deux derniers éléments (prises modulo L) vérifient soit $d_1 = d_2$ (et $d_1 \leq \frac{L-1}{2}$) soit $d_1, d_2 \leq \frac{L}{2} - 1$. Cette représentation est unique. À titre d'exemple, le cycle de la figure 5.5 est construit de cette manière-là.

Ainsi nous savons dans quel ordre les éléments de chaque sous-ensemble apparaissent. Cela permet de savoir quelles paires de $\{1, ..., L\}$ peuvent être ajoutées ensemble en transformant un sous-ensemble à trois éléments avec l'opération mentionnée précédemment. Par exemple, pour $L = 8$, nous sommes en mesure d'ajouter toutes les paires en transformant les triplets suivant :

123	345	567	781
135	571	246	682
258	836	614	472

Comme chaque élément de $\{1, ..., 8\}$ est le centre d'une paire, les singletons peuvent aussi être réalisés. Ce tableau a été construit en suivant un chemin eulérien du graphe complet en effectuant tout d'abord des sauts de taille 1 (première ligne) puis de taille 2 (deuxième ligne) et enfin de taille 3 (troisième ligne), et peut-être généralisé pour tout entier pair. Nous ajoutons alors $\frac{1}{2}\binom{L}{2} + L$ sommets dans le cycle.

Notons que cela revient à chercher un cycle universel des sous-ensembles à moins de trois ou trois éléments de $\{1, \ldots, L\}$. Dans la même idée, les cycles universels des multi-ensembles ont déjà été considérés par Jackson ([59]).

Cycle universel des sous-ensembles à k-éléments et puissance de cycles

Les cycles universels se généralisent aux sous-ensembles à k éléments et font partie des problèmes de type code de Gray. Un *cycle universel des sous-ensembles à k éléments de* $\{1, ..., L\}$ est une séquence d'entiers $(x_1, ..., x_{\binom{L}{k}})$, où chaque x_i est un entier compris entre 1 et L et où chaque sous-ensemble de taille k de $\{1, ..., L\}$ est représenté par un unique bloc de taille k : $(x_i, ..., x_{i+k})$ (pris cycliquement). Par exemple, la séquence d'entiers $(1, 2, 3, 4, 5, 1, 3, 5, 2, 4)$ est un cycle universel des paires de $\{1, ..., 5\}$.

Comme auparavant, il faut que $\binom{L-1}{k-1}$ (le nombre de sous-ensembles de taille k de $\{1, ..., L\}$ contenant un entier donné) soit divisible par k (le nombre d'ensembles auquel appartient une occurrence d'un entier). Chung *et al.* [26] ont proposé en 1992 la conjecture suivante[1] :

Conjecture 5.10 (Chung, Diaconis, Graham [26]) *Pour tout $k \geq 2$, il existe un entier $L_0(k)$ tel que pour tout $L \geq L_0(k)$ vérifiant $\binom{L-1}{k-1} \equiv 0 \bmod k$, il existe un cycle universel des sous-ensembles à k éléments de $\{1, ..., L\}$.*

1. Un prix de 100$ est promis à celui qui résoudra cette conjecture

Le théorème 5.7 confirme la conjecture pour $k = 3$. La conjecture est aussi vérifiée pour $k = 2$, $k = 4$ et 5, et partiellement pour $k = 6$ (voir [26, 59, 58, 57]).

Il a aussi été montré dans Stevens *et al.* [90] qu'il n'existe pas de cycle universel pour $k = L - 2$ et $L \geq 3$.

Lorsque k est impair, un cycle universel nous donne une coloration identifiante en L couleurs du graphe $\mathcal{C}_{\frac{k-1}{2}}$. Le fait que la conjecture soit vraie pour $k = 5$ implique, de la même manière que précédemment que pour les carrés des cycles, il existe des colorations identifiantes avec de l'ordre de $n^{\frac{1}{5}}$ couleurs, où n est le nombre de sommets. Or pour ces graphes, les meilleurs codes identifiants sont de l'ordre de $\frac{n}{2}$. Il existe donc des familles infinies de graphes pour lesquels $\chi_{id}(G) = O(\gamma^{\text{ID}}(G)^{\frac{1}{5}})$.

La conjecture reste ouverte pour $k \geq 6$. Cependant, Curtis *et al.* [31] ont montré que l'on peut toujours atteindre presque tous les sous-ensembles : c'est-à-dire construire des cycles qui contiennent tous les sous-ensembles une seule fois, sauf un nombre négligeable d'entre eux (par rapport au nombre total de sous-ensembles). Cela implique le résultat suivant, en utilisant les puissances $\frac{k-1}{2}$-ème de cycles.

Proposition 5.11 *Pour tout entier k impair, il existe une famille infinie de graphes G tels que*

$$\chi_{id}(G) = O(\gamma^{ID}(G)^{\frac{1}{k}}).$$

5.1.7 Conclusion sur les colorations identifiantes globales

Le problème des colorations identifiantes n'a été pour l'instant que peu étudié. Nous n'avons donné ici que quelques bornes et constructions générales. Le lien établi avec les cycles universels nous permet de considérer ces colorations comme une généralisation de ce problème. Il serait par exemple intéressant d'étudier ce problème sur d'autres structures régulières, comme les grilles.

5.2 Du global au local

Afin de mieux comprendre les colorations identifiantes, nous avons étudié une variante locale. Cela va nous permettre de comparer nos colorations avec les colorations propres des graphes et d'étudier des problèmes similaires.

Soit $G = (V, E)$ un graphe. Une coloration c des sommets de G est *localement identifiante* si les deux conditions suivantes sont vérifiées :

(i) c est une coloration *propre* : pour toute arête xy de G, $c(x) \neq c(y)$;

(ii) pour toute arête xy de G telle que $N[x] \neq N[y]$, $c(N[x]) \neq c(N[y])$.

Un graphe G est dit *k-lid-coloriable* s'il existe une coloration localement identifiante de G en k couleurs. La figure 5.6 montre un exemple de graphe 3-lid-coloriable.

La première condition nous permet de comparer ces colorations aux colorations propres des graphes. La deuxième condition traduit le caractère local et identifiant de la coloration. Nous avons ajouté la condition « $N[x] \neq N[y]$ » afin de pouvoir colorier n'importe quel graphe. Ainsi nous tolérons les jumeaux dans nos graphes. Cela est utile dans la suite pour certaines preuves où nous raisonnons par induction. Imposer cette condition ne change pas fondamentalement le problème.

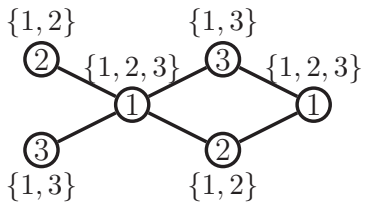


Figure 5.6 – *Une coloration localement identifiante d'un graphe.*

Comme auparavant, nous définissons le *nombre chromatique localement identifiant* ou *nombre lid-chromatique* de G, $\chi_{lid}(G)$, comme le plus petit nombre de couleurs nécessaire dans une coloration localement identifiante de G. Le graphe G de la figure 5.6 n'est pas 2-lid-coloriable, donc $\chi_{lid}(G) = 3$. En fait, les seuls graphes connexes 2-lid-coloriables sont K_1 et K_2 :

Lemme 5.12 *Soit G un graphe connexe 2-lid-coloriable, alors G a au plus deux sommets.*

Preuve : Soit un graphe G connexe ayant au moins trois sommets. Si G contient un triangle, G a besoin d'au moins trois couleurs pour avoir une coloration propre. Sinon G a au moins deux sommets adjacents u et v qui ne sont pas jumeaux. Soit c une coloration de G en deux couleurs. Les sommets u et v étant adjacent, leurs couleurs sont différentes. Les ensembles $c(N[u])$ et $c(N[v])$ contiennent les deux couleurs $c(u)$ et $c(v)$, et il n'existe pas d'autres couleurs. Les sommets u et v ne sont donc pas séparés. □

Dans la suite, nous étudions divers problèmes inspirés de ceux étudiés dans le cadre des colorations propres. Dans la partie 5.3 nous donnons plusieurs types de borne pour χ_{lid} : avec le nombre de sommets, avec le degré maximum, avec le nombre chromatique. Dans la partie 5.4 nous étudions les colorations localement identifiantes dans les graphes bipartis : si le nombre lid-chromatique est borné pour cette classe, il n'en est pas moins difficile de le calculer exactement. Dans la partie 5.5 nous étudions le problème pour plusieurs sous-classes de graphes parfaits et enfin dans la partie 5.6 nous nous intéressons aux graphes planaires.

Contrairement aux colorations propres, la propriété d'être localement identifiante n'est pas une propriété héréditaire. En effet, si c est une coloration localement identifiante d'un graphe G, alors en enlevant un sommet u, la coloration n'est pas toujours localement identifiante. De fait, nous verrons (proposition 5.19) qu'il existe des graphes G de nombre lid-chromatique borné, tels qu'en enlevant un sommet, le nombre lid-chromatique soit aussi grand que souhaité. Nous ne pouvons donc pas utiliser les techniques classiques de coloration propre qui font appel à cette hérédité.

5.3 Bornes générales pour les colorations localement identifiantes

5.3.1 En fonction du nombre de sommets

Contrairement aux codes identifiants ou aux colorations identifiantes, il n'y a pas de borne inférieure fonction du nombre de sommets pour les colorations localement identifiantes. En effet, il existe des graphes infiniment grand qui sont 3-lid-coloriable. C'est le cas par exemple de toutes les grilles qui se colorient avec trois couleurs, comme le montre la figure 5.7.

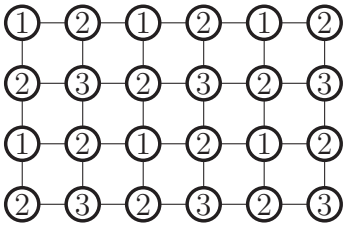


Figure 5.7 – *Coloration localement identifiante de la grille avec trois couleurs.*

Quant à une borne supérieure, nous avons la borne triviale $\chi_{lid}(G) \leq |V(G)|$, qui est atteinte pour les graphe complets. Il existe des exemples non triviaux de graphes atteignant aussi cette borne :

Proposition 5.13 $\chi_{lid}(P_{2k}^{k-1}) = 2k$ *et* $\chi_{lid}(P_{2k}^{k-1} \bowtie K_1) = 2k+1$

Preuve : Considérons les sommets x_1,...,x_{2k} du chemin P_{2k} et c une coloration localement identifiante de P_{2k}^{k-1}. Les k sommets x_1 à x_k de P_{2k}^{k-1} forment une clique et sont donc de couleurs différentes, car c est une coloration propre. Montrons par récurrence sur i, $k \leq i \leq 2k$, que les sommets de x_1 à x_i sont tous de couleurs différentes dans une coloration localement identifiante de P_{2k}^{k-1}. Supposons que cela soit vrai pour $i < 2k-1$, x_{i+1} est l'unique sommet dans la différence symétrique de x_{i-1-k} et x_{i-k}, il doit donc être colorié avec une couleur qui n'apparaît pas parmi les sommets x_1 à x_i car ces sommets sont voisins de x_{i-1-k}. Pour colorier x_{2k}, si nous utilisons une couleur déjà prise auparavant, deux sommets auront la même différence symétrique, x_{2k} ne peut donc pas prendre une couleur déjà utilisée auparavant et tous les sommets sont de couleurs différentes.

Pour $P_{2k}^{k} \bowtie K_1$, nous raisonnons de la même manière pour les sommets du chemin qui induisent le graphe P_{2k}^{k-1} (le sommet universel ne change donc rien au raisonnement). Enfin le sommet universel est adjacent à tous les autres sommets et a donc une couleur non utilisée par les autres sommets. □

Il est intéressant de remarquer que ces exemples de graphes sont des graphes extrêmes aussi pour les codes identifiants. Cependant, comme dans le cas des colorations identifiantes globales, tous les graphes extrêmes pour le problème des codes identifiants ne le sont pas pour les colorations localement identifiantes :

Lemme 5.14 *Pour* $n \geq 3$, $\chi_{lid}(K_{1,n}) = 3$.

Preuve : Le sommet central reçoit la couleur 1, une des feuilles est de couleur 3 et les autres sont de couleur 2. Cette coloration est localement identifiante. L'égalité découle du lemme 5.12. □

Contrairement aux colorations identifiantes globales, il existe des graphes sans jumeaux à n sommets qui ont besoin de n couleurs pour être localement identifiés mais qui ne sont pas extrêmes pour le problème des codes identifiants :

Lemme 5.15 *Soit k un entier. Soit G_k le graphe obtenu en prenant une clique de taille k et en ajoutant des sommets pendants à tous les sommets de la clique sauf un. Alors $\chi_{lid}(G_k) = |V(G_k)| = 2k - 1$.*

Une construction du graphe G_k est donnée dans la figure 5.8.

Preuve : Pour colorier la clique toutes les couleurs doivent être différentes (coloration propre). Soient x un des sommets pendant, x' son unique voisin et y le sommet de la clique qui n'a pas de sommets pendant. Le seul sommet dont la couleur peut séparer x' de y est le sommet x, donc la couleur de x ne doit pas apparaître dans le voisinage de y, c'est-à-dire, dans la clique. De plus si deux sommets pendants sont de même couleurs, leurs uniques voisins dans la clique ne seront pas séparés. Par conséquent, dans toute coloration localement identifiante de G_k les sommets reçoivent tous des couleurs différentes. □

Figure 5.8 – *Exemple de graphe extrémal pour les codes identifiants mais pas pour les colorations localement identifiantes (à gauche) et vice versa (à droite).*

La question suivante est ouverte :

Question 5.16 *Quels sont les graphes G tels que $\chi_{lid}(G) = |V(G)|$?*

5.3.2 En fonction du nombre chromatique

Une coloration localement identifiante est une coloration propre, nous avons donc :

$$\chi(G) \leq \chi_{lid}(G).$$

Cette borne est atteinte pour le graphe complet où nous n'avons pas d'autre possibilité que de colorier tous les sommets avec des couleurs différentes. Lorsque $\chi(G) \leq 2$ et que G est connexe, le lemme 5.12 nous dit que seuls les graphes complets atteignent la borne. C'est aussi vrai pour $\chi(G) = 3$ comme nous allons le voir dans la partie 5.4.

Pour des valeurs plus grandes de $\chi(G)$, nous ne connaissons pas à ce jour d'autres exemples pour lesquels la borne est serrée.

Question 5.17 *Existe-t-il des graphes G connexes non complets tels que $\chi(G) = \chi_{lid}(G)$?*

Nous pensons que de tels graphes existent. Ces graphes ne sont pas parfaits. En effet, si un graphe connexe n'est pas complet et possède une clique de taille k, alors ce graphe nécessite au moins $k+1$ couleurs pour être localement identifié car il faut k couleurs pour colorier la clique et un des sommets reliés à la clique doit avoir une couleur différente de celles de la clique pour séparer les sommets de la clique entre eux.

Une question naturelle est de savoir si l'on peut borner supérieurement χ_{lid} par une fonction de χ. La réponse est non : le résultat suivant nous permet de construire des familles de graphes dont le nombre lid-chromatique est très grand mais qui sont 3-coloriables : il suffit de prendre un graphe G et de subdiviser deux fois les arêtes :

Lemme 5.18 *Soit G un graphe. Soit G' le graphe obtenu en subdivisant les arêtes de G deux fois. Alors $\chi_{lid}(G') \geq \chi(G)$ et $\chi(G') \leq 3$.*

Preuve : Pour obtenir une coloration propre de G', les sommets initiaux de G sont coloriés avec la couleur 1. Nous utilisons les couleurs 2 et 3 pour colorier les deux sommets ajoutés sur chaque arête. La coloration obtenue est propre et donc $\chi(G') \leq 3$.

Prenons maintenant une coloration localement identifiante c de G'. Pour chaque arête $e = xy$ de G, nous notons x_e et y_e les deux sommets ajoutés sur l'arête xy, avec x_e adjacent à x et y_e adjacent à y. Si $c(x) = c(y)$ dans la coloration de G', alors les sommets x_e et y_e ont dans leur voisinage l'ensemble de couleurs $\{c(x), c(x_e), c(y_e)\}$ et ne sont donc pas séparés, ce qui est contradictoire. Nous avons donc nécessairement $c(x) \neq c(y)$ pour toute arête xy de G. Ainsi la coloration c induit une coloration propre des sommets de G et $\chi_{lid}(G') \geq \chi(G)$. □

La figure 5.9 donne la construction pour le graphe K_4. Les graphes complets subdivisés deux fois nous donnent directement des exemples de graphes avec un nombre lid-chromatique très grand mais un nombre chromatique borné.

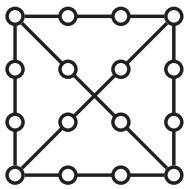


Figure 5.9 – *Le graphe K_4 dont les arêtes ont été subdivisées deux fois.*

Cette construction nous permet aussi de construire des graphes ayant un nombre lid-chromatique borné mais dont ce nombre augmente lorsque un sommet est retiré.

Proposition 5.19 *Pour tout entier $k \geq 4$, il existe un graphe G d'ordre $O(k^2)$ et un sommet u de G tel que $\chi_{lid}(G) \leq 5$ et $\chi_{lid}(G - u) \geq k$.*

Preuve : Soit k un entier. Considérons le graphe G obtenu à partir d'une clique de taille k subdivisée deux fois à laquelle est ajouté un sommet u relié pour chaque arête du graphe complet à exactement un des nouveaux sommets. Ce graphe est représenté sur la figure 5.10 lorsque $k = 6$.

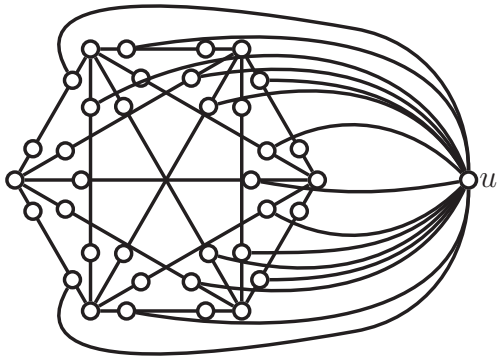


Figure 5.10 – *Un graphe* 5*-lid-coloriable* G *tel qu'en enlevant le sommet* u*, le nombre lid-chromatique de* $G - u$ *augmente d'un. En prenant une clique de départ plus grande, le nombre lid-chromatique de* $G - u$ *peut devenir arbitrairement grand alors que* G *reste* 5*-lid-coloriable.*

D'après la proposition 5.18, $\chi_{lid}(G - u) \geq k$. Nous donnons maintenant une coloration localement identifiante de G avec cinq couleurs.

Le sommet u est colorié avec la couleur 5, c'est le seul sommet qui aura cette couleur. Chaque sommet de la clique initiale est colorié avec des couleurs allant de 1 à 4 de telle sorte que chaque couleur apparaisse au moins une fois. Soient $e = xy$ une arête initiale de la clique et x_e, y_e les sommets ajoutés sur l'arête, x_e étant adjacent à x. Si x et y ont reçu des couleurs différentes, alors x_e est colorié avec la couleur de y et y_e avec la couleur de x (en reprenant les notations de la preuve du lemme 5.18). Sinon, x_e et y_e reçoivent deux couleurs parmi les couleurs non utilisées par x et y. Nous pouvons toujours effectuer cette coloration de telle sorte que le sommet u soit adjacent à des sommets de toutes les couleurs (en prenant par exemple tous les voisins d'un des sommets initiaux de la clique dans le voisinage de u sauf un qui est à distance 2).

La coloration obtenue c est bien une coloration propre. Si x est un sommet initial de la clique, alors $c(N[x]) = \{1, 2, 3, 4\}$, car chaque couleur d'un voisin initial de la clique apparaît sur l'un des nouveaux voisins. Si x est un sommet sur une arête subdivisée, alors soit x est relié à u et alors x est séparé de ces deux voisins car il a la couleur 5 dans son voisinage. Soit x n'est pas relié à u et a moins de trois couleurs dans son voisinage, son voisin initialement dans la clique en a quatre et son autre voisin est relié à u. Enfin u est séparé de tous les sommets car il a les cinq couleurs dans son voisinage. □

5.3.3 En fonction du degré maximum

Concernant les colorations propres classiques, il existe une borne simple utilisant le degré maximum du graphe, obtenue avec un algorithme glouton : pour tout graphe G de degré maximum Δ,

$$\chi(G) \leq \Delta + 1.$$

Cette partie a pour but de trouver une borne similaire pour les colorations localement identifiante.

Une première borne supérieure

Une façon de procéder est de noter le fait suivant :

Lemme 5.20 *Soient G un graphe et c une coloration propre de G^3. Alors la coloration c induite sur les sommets de G est localement identifiante :*

$$\chi_{lid}(G) \leq \chi(G^3)$$

Preuve : Soient x et y deux sommets adjacents de G vérifiant $N[x] \neq N[y]$. Il existe un sommet z qui, sans perte de généralité, est adjacent à x mais pas à y. Soit c un coloration propre de G^3. Dans le graphe G, la couleur de z ne peut apparaître dans $N[y]$ car tous les sommets de $N[y]$ sont à distance inférieure ou égale à 3 de z. Donc $c(N[x]) \neq c(N[y])$ et la coloration est localement identifiante. □

Corollaire 5.21 *Soit G un graphe de degré maximum Δ,*

$$\chi_{lid}(G) \leq \Delta^3 - \Delta^2 + \Delta + 1.$$

Preuve : Le degré maximum dans G^3 est $\Delta^3 - \Delta^2 + \Delta$ où Δ est le degré maximum dans G. Le résultat s'obtient en utilisant l'inégalité $\chi(H) \leq \Delta(H) + 1$ avec $H = G^3$ et le résultat du lemme 5.20. □

Des graphes nécessitant Δ^2 couleurs

Nous construisons maintenant une famille de graphes pour lesquels $\chi_{lid}(G) \geq \Delta^2 - \Delta + 1$. Ces graphes sont construits à partir des plans projectifs discrets. Un plan projectif P d'ordre n est composé de $n^2 + n + 1$ points et $n^2 + n + 1$ droites (vues comme ensembles de points) de telle sorte que :

- chaque droite contient exactement $n + 1$ points,
- chaque point est dans $n + 1$ droites,
- chaque paire de points est contenue dans une unique droite,
- toute paire de droites a un point en commun.

La figure 5.11 donne un exemple de plan projectif d'ordre 2. Les plans projectifs discrets correspondent aux plans projectifs sur des corps finis d'ordre n. Lorsque n est une puissance d'un nombre premier, il existe des constructions (voir par exemple le livre de Hughes et Piper [56]).

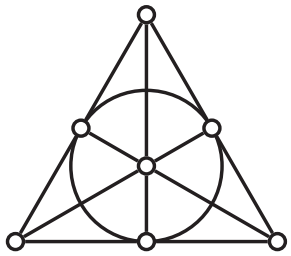


Figure 5.11 – *Le plan de Fano est un plan projectif d'ordre* 2*. Le cercle central correspond à une droite.*

Proposition 5.22 *Pour tout entier n puissance d'un nombre premier, il existe un graphe G de degré maximum $n+1$ tel que*

$$\chi_{lid}(G) \geq n^2 + n + 1 = \Delta^2 - \Delta + 1.$$

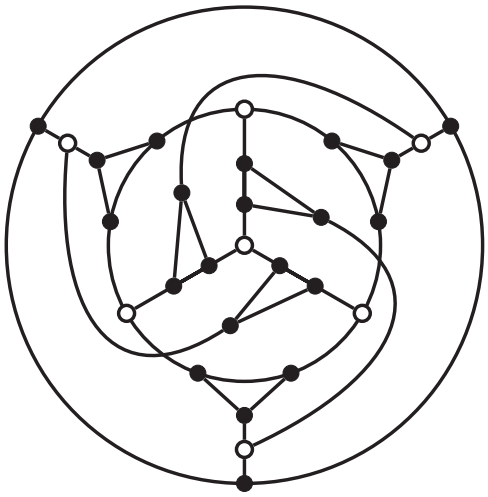


Figure 5.12 – *Le graphe H_3, 3-régulier. Dans toute coloration localement identifiante, les sommets blancs doivent avoir des couleurs différentes.*

Preuve : Soit P un plan projectif d'ordre n. Soit G_{n+1} le graphe obtenu à partir d'une clique de taille $n+1$ en ajoutant pour chaque sommet v de la clique un sommet v' adjacent uniquement à v. Dans une coloration localement identifiante, tous les sommets v' doivent avoir des couleurs différentes. Pour chaque droite D de P, nous créons une copie G^D_{n+1} de G_{n+1} dans laquelle les sommets v' sont indexés par les points de la droite D. Pour chaque point p de P, tous les sommets indexés par p sont confondus en un unique sommet p^*. Le graphe obtenu, H_{n+1} est $(n+1)$-régulier et possède $(n^2+n+1)\times(n+1)$ sommets. La figure 5.12 donne la construction pour $n=3$. Par construction, tous les sommets p^* doivent avoir des couleurs différentes : pour chaque paire de tels sommets, il existe une droite D qui contient les deux points du plan correspondant, et le graphe G_{n+1} correspondant impose d'avoir des couleurs différentes. Il y a n^2+n+1 points p^*, donc $\chi_{lid}(H_{n+1}) \geq n^2+n+1 = \Delta^2-\Delta+1$. □

Borne supérieure d'ordre 2

Dans ce paragraphe, nous montrons que tout graphe de degré maximum Δ possède une coloration localement identifiante avec $O(\Delta^2)$ couleurs :

Théorème 5.23 *Pour tout graphe G de degré maximum $\Delta \geq 3$:*

$$\chi_{lid}(G) \leq 2\Delta^2 - 3\Delta + 3.$$

Pour montrer ce théorème, nous utilisons un résultat qui nous permet, dans un graphe déjà bien colorié, de recolorier un sommet en évitant un nombre fixé de couleurs.

Lemme 5.24 *(Lemme de recoloration) Soit G un graphe de degré maximum $\Delta \geq 3$. Soient v un sommet et d son degré. Supposons qu'il existe une coloration localement identifiante c de G utilisant strictement plus de $2d(\Delta - 1)$ couleurs. Alors il existe une liste de couleurs L de taille au plus $2d(\Delta - 1)$ telle que si le sommet v est recolorié avec n'importe quelle couleur n'appartenant pas à la liste L, la coloration reste localement identifiante.*

Notons que dans l'énoncé de ce lemme, la couleur initiale du sommet v n'apparaît pas obligatoirement dans la liste de couleurs interdites L. Ainsi, la couleur du sommet v fait partie des couleurs autorisées pour le « recolorier ».

Preuve : L'idée de la preuve est d'ajouter dans L des couleurs *interdites* pour v. Soient v_1,...v_d les voisins de v. Pour chaque sommet v_i, soit $u_{i,1}$,...,u_{i,s_i} les voisins de v_i qui ne sont pas voisins de v (voir la figure 5.13 pour les notations). Remarquons que deux sommets $u_{i,j}$ et $u_{i',j'}$ peuvent correspondre au même sommet. Pour chaque $1 \leq i \leq d$, nous construisons une liste de couleurs L_i de taille au plus $2(\Delta - 1)$.

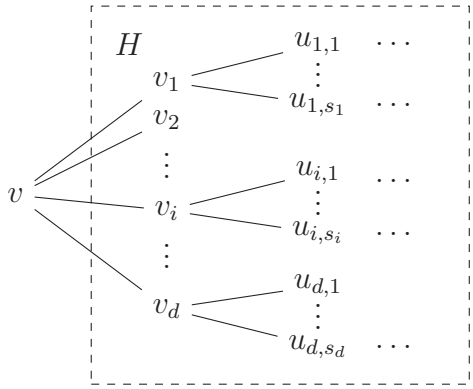


Figure 5.13 – *Voisinage du sommet v.*

Tout d'abord, la couleur $c(v_i)$ est placée dans L_i. Nous distinguons ensuite deux cas :

S'il existe un sommet $u_{i,j}$ tel que $c(N[u_{i,j}]) = c(N[v_i] \setminus \{v\})$, v_i est dit de *type A* et nous ajoutons à L_i toutes les couleurs de $c(N[v_i] \setminus \{v\})$. À cette étape, nous

avons ajouté au plus $\Delta - 1$ couleurs car $c(v_i)$ était déjà dans L_i. Finalement, pour tous les sommets $u_{i,j'}$ tels que $c(N[u_{i,j'}]) \setminus c(N[v_i] \setminus \{v\})$ est non vide, nous ajoutons une couleur arbitraire de $c(N[u_{i,j'}]) \setminus c(N[v_i] \setminus \{v\})$ à L_i. Nous ajoutons alors au plus $\Delta - 2$ couleurs car $j' \neq j$. Au final : $|L_i| \leq 2(\Delta - 1)$.

Sinon, v_i est dit de *type B* et pour chaque voisin $u \neq v$ de v_i, si $c(N[u] \setminus \{v\}) \setminus c(N[v_i] \setminus \{v\})$ est non vide nous ajoutons une couleur de cet ensemble à L_i. Le sommet u peut être l'un des v_j ou bien l'un des $u_{i,j}$ mais il y a au plus $\Delta - 1$ sommets u possibles et nous ajoutons donc $\Delta - 1$ couleurs à L_i. Enfin, si $c(N[v_i]) \setminus c(N[v])$ est non vide, nous ajoutons aussi une des couleurs de cet ensemble à L_i. Au final $|L_i| \leq \Delta + 1 \leq 2(\Delta - 1)$ car $\Delta \geq 3$.

Soit maintenant $L = \bigcup_{i=1}^{d} L_i$. Par construction, $L \leq 2d(\Delta - 1)$. Soit c' une nouvelle coloration de G obtenue en recoloriant v avec une couleur qui n'est pas dans L. Montrons que c' est localement identifiante. La liste L contient toutes les couleurs de v_1,...,v_d donc c' est une coloration propre. Soient x et y deux sommets adjacents tels que $N[x] \neq N[y]$. Si ni x ni y ne sont dans dans $N[v]$ alors les couleurs dans leur voisinage fermé n'ont pas changé et $c'(N[x]) \neq c'(N[y])$. Sans perte de généralité, supposons que $x = v_1$.

Supposons d'abord que $y = v$ et qu'il existe une couleur $c_0 \in c(N[v_1]) \setminus c(N[v])$. Nécessairement, $c_0 \neq c(v)$. Si v_1 est de type A, $c_0 \in L$ et nous avons donc toujours $c_0 \in c'(N[v_1]) \setminus c'(N[v])$, c_0 sépare v et v_1. Si v_1 est de type B, une couleur de $c(N[v_1]) \setminus c(N[v])$ est dans L (qui n'est pas obligatoirement c_0) et sépare toujours v_1 de v. Supposons maintenant que $c(N[v_1]) \setminus c(N[v])$ est vide. Comme $c(N[v_1]) \neq c(N[v])$, il existe une couleur $c_0 \in c(N[v]) \setminus c(N[v_1])$. Comme $c_0 \neq c(v)$, $c_0 \in L$ et sépare toujours v de v_1. Ainsi, dans tous les cas $c'(N[v_1]) \neq c'(N[v])$.

Supposons maintenant que $y = v_j$ avec $j \neq 1$ et v_1, v_j adjacents. Sans perte de généralité, supposons qu'il existe une couleur $c_0 \in c(N[v_1]) \setminus c(N[v_j])$. Nécessairement, $c_0 \neq c(v)$. Si v_1 est de type A, $c_0 \in L$ donc sépare toujours v_1 de v_j. Sinon, v_1 est de type B. Si $c(N[v_j] \setminus \{v\}) \setminus c(N[v_1] \setminus \{v\})$ est non vide, une couleur de cette ensemble est dans L et sépare v_j de v_1. Sinon, si v_j est de type A, nous pouvons conclure en raisonnant de la même manière que pour v_1. Si v_j est aussi de type B, alors une couleur de $c(N[v_j] \setminus \{v\}) \setminus c(N[v_1] \setminus \{v\})$ est dans L et sépare v_1 de v_j. Dans tous les cas, $c'(N[v_1]) \neq c'(N[v_j])$.

Sinon, nous pouvons supposer que $y = u_{1,1}$. Si $c(N[u_{1,1}]) = c(N[v_1] \setminus \{v\})$ alors v_1 est de type A et donc $c(N[u_{1,1}]) \subseteq L$. La nouvelle couleur de v sépare v_1 de $u_{1,1}$. Si $c(N[v_1] \setminus \{v\}) \setminus c(N[u_{1,1}])$ est non vide, n'importe quelle couleur de cet ensemble séparera toujours v_1 de $u_{1,1}$. Sinon $c(N[u_{1,1}]) \setminus c(N[v_1] \setminus \{v\})$ est non vide et une couleur de cet ensemble est dans L et sépare toujours v_1 de $u_{1,1}$. Dans tous les cas $c'(N[v_1]) \neq c'(N[u_{1,1}])$, ce qui termine la preuve. □

Ce lemme nous permet de montrer assez rapidement que $\chi_{lid}(G) \leq 2\Delta(\Delta-1)+1$ pour n'importe quel graphe G de degré maximum $\Delta \geq 3$. En effet, nous pouvons partir d'une coloration localement identifiante triviale de G (tous les sommets de couleurs différentes par exemple) et enlever au fur et à mesure des couleurs en recoloriant un par un les sommets jusqu'à obtenir le bon nombre de couleurs.

Nous obtenons un résultat plus fort sur les graphes d-dégénérés [2] avec le résultat qui suit, et qui induira le théorème 5.23.

Proposition 5.25 *Soit G un graphe d-dégénéré de degré maximum $\Delta \geq 3$ avec $d < \Delta$, alors :*

$$\chi_{lid}(G) \leq 2(\Delta - 1)^2 + d.$$

Preuve : Fixons Δ pour toute la preuve. Nous montrons la proposition par induction sur le nombre de sommets de G. Cela est clairement vrai pour un graphe trivial avec un sommet.

Soit $n > 1$ un entier. Supposons que le résultat soit vrai pour tout graphe d-dégénéré avec $n-1$ sommets et de degré maximum au plus Δ. Soit G un graphe d-dégénéré avec n sommets et de degré maximum au plus Δ.

Soit v un sommet de degré minimum. Soit t son degré. G étant d-dégénéré, $t \leq d$. Doit $H = G - v$. Le graphe H est d-dégénéré avec $n-1$ sommets et de degré maximum au plus Δ, donc par induction il existe une coloration localement identifiante c de H utilisant $2(\Delta-1)^2 + d$ couleurs.

Nous utilisons les mêmes notations que dans la preuve du lemme 5.24 : v_1,...,v_t désignent les voisins de v et $u_{i,j}$ les voisins de v_i qui ne sont pas voisins de v. Nous construisons une liste L' de taille au plus d contenant pour chaque $i \in \{1, ..., t\}$ la couleur $c(u_{i,1})$ si $u_{i,1}$ existe. Chaque sommet v_i est de degré au plus $\Delta - 1$ dans H. D'après le lemme 5.24, nous pouvons recolorier chaque sommet v_i avec une couleur qui n'est pas dans L' en maintenant une coloration localement identifiante avec au plus $2(\Delta-1)^2+d$ couleurs. En effet, il y a $2(\Delta-1)^2$ couleurs interdites par le lemme appliqué à v_i et d couleurs interdites par L', mais la couleur de $u_{i,1}$ (si ce sommet existe) est comptée deux fois. Il y a donc $2(\Delta-1)^2 + d - 1$ couleurs interdites et $2(\Delta-1)^2 + d$ couleurs disponibles, donc une des couleurs est libre pour v_i.

Supposons donc maintenant que c est une coloration localement identifiante de H telle qu'aucun sommet v_i n'a une couleur de L'. La coloration c est étendue à G en donnant à v une nouvelle couleur qui n'est pas utilisée dans H. Montrons que la coloration c' ainsi obtenue de G est une coloration localement identifiante. La coloration est clairement propre et toute paire de sommets qui était déjà séparée dans H par c l'est encore par c'. Les seules paires de sommets adjacents qui peuvent poser problème sont en fait les paires $\{v, v_i\}$ pour $1 \leq i \leq t$. Si $u_{i,1}$ existe, alors sa couleur est dans L' et donc v ne l'a pas dans son voisinage, elle sépare donc la paire $\{v, v_i\}$. Sinon, $N[v_i] \subseteq N[v]$. Comme v a été pris de degré minimum, cela signifie que $N[v_i] = N[v]$. Les deux sommets sont jumeaux et ne peuvent donc pas être séparés.

Finalement, c' est une coloration localement identifiante de G avec $2(\Delta-1)^2 + d + 1 \geq 2(\Delta-1)^2 + 2$ couleurs, mais la couleur de v n'est utilisée qu'une seule fois. D'après le lemme 5.24, nous pouvons utiliser au moins deux couleurs pour v, l'une d'elle est donc différente de celle de v. On peut donc changer la couleur de v pour une couleur déjà utilisée et l'on obtient une coloration localement identifiante de G avec $2(\Delta-1)^2 + d$ couleurs. □

2. Déjà rencontrés dans la partie 3.3 du chapitre 3 où l'on a rappelé la définition.

Preuve du théorème 5.23 : Si G n'est pas connexe, chaque composante peut être coloriée indépendamment. Supposons donc que G est connexe. Si G n'est pas Δ-régulier et connexe, alors G est $(\Delta-1)$-dégénéré et la proposition 5.25 s'applique.

Supposons donc que G est connexe et Δ-régulier. Soit v un sommet de G. Chaque composante connexe de $G-v$ est $(\Delta-1)$-dégénérée et donc $G-v$ est $(\Delta-1)$-dégénéré. D'après la proposition 5.25, $G-v$ a une coloration localement identifiante avec $2\Delta^2-3\Delta+1$ couleurs. Comme auparavant, nous recolorions chaque voisin de v de telle sorte que la coloration reste localement identifiante et qu'un voisin de v n'ayant pas le même voisinage que v ait dans son voisinage une couleur qui n'apparaît pas dans les voisins de v. Pour faire cela, il y a $2(\Delta-1)^2$ couleurs interdites d'après le lemme 5.24 et Δ couleurs interdites par les couleurs des voisins, dont une comptée en double. Cela fait donc $2\Delta^2-3\Delta+1$ couleurs interdites. En ajoutant éventuellement une couleur, nous obtenons une coloration localement identifiante de $G-v$ avec $2\Delta^2-3\Delta+2$ couleurs. On peut maintenant donner une couleur complètement nouvelle à v et obtenir une coloration localement identifiante de G avec $2\Delta^2-3\Delta+3$. □

Nous avons donc montré que $\chi_{lid}(G)=O(\Delta^2)$ et qu'il existe des graphes ayant besoin d'un nombre de couleurs de l'ordre de Δ^2 couleurs. Cependant, nous ne connaissons pas le coefficient optimal pour cette borne :

Question 5.26 *A-t-on pour tout graphe G de degré maximum Δ, $\chi_{lid}(G)\leq\Delta^2+O(\Delta)$?*

Le cas des graphes de degré maximum 2

Le théorème 5.23 ne s'applique que pour les graphes de degré maximum $\Delta\geq 3$. La preuve pourrait s'adapter afin d'obtenir une borne pour les graphes de degré maximum 2, c'est-à-dire des cycles et des chemins. Un calcul direct nous donne une borne exacte.

Proposition 5.27 *Soit $n\geq 3$ un entier. Soit P_n le chemin d'ordre n. Alors*
- $\chi_{lid}(P_n)=3$ *si* $n\equiv 1 \bmod 2$,
- $\chi_{lid}(P_n)=4$ *sinon.*

Preuve : Soient $v_0,\ldots,v_{n-1}$ les sommets de P_n. D'après le lemme 5.12, $\chi_{lid}(P_n)\geq 3$. En coloriant le sommet v_i avec la couleur i modulo 4, nous obtenons une coloration localement identifiante avec quatre couleurs : $\chi_{lid}(P_n)\leq 4$.

Supposons maintenant qu'il existe une coloration localement identifiante c avec trois couleurs. Alors sans perte de généralité, $c(v_0)=1$ et $c(v_1)=2$. Donc $c(N[v_0])=\{1,2\}$ et $\{1,2\}\subseteq c(N[v_1])$. Pour séparer v_1 de v_0, nécessairement $c(v_2)=3$. De la même manière, nous avons nécessairement $c(v_3)=2$. Ainsi de suite, toutes les couleurs sont forcées : $c(v_i)=1$ si $i\equiv 0 \bmod 4$, $c(v_i)=2$ si $i\equiv 1 \bmod 2$ et $c(v_i)=3$ si $i\equiv 2 \bmod 4$. Si $n\equiv 0 \bmod 2$, alors $c(v_{n-1})=2$ et $c(N[v_{n-1}])=c(N[v_{n-2}])$. La coloration n'est donc pas localement identifiante. Par contre si $n\equiv 1 \bmod 2$, la coloration obtenue est localement identifiante, ce qui termine la preuve. □

Proposition 5.28 *Soit $n\geq 4$ un entier. Soit $\mathcal{C}_n$ le cycle d'ordre n. Alors*

- $\chi_{lid}(\mathcal{C}_n) = 3$ *si* $n \equiv 0 \bmod 4$,
- $\chi_{lid}(\mathcal{C}_n) = 5$ *si* $n = 5$ *ou* $n = 7$,
- $\chi_{lid}(\mathcal{C}_n) = 4$ *sinon.*

Preuve : Soient v_0,...,v_{n-1} les sommets de $\mathcal{C}_n$. D'après le lemme 5.12, $\chi_{lid}(\mathcal{C}_n) \geq 3$.

Pour $n \geq 4$, $n \neq 5, 7$, nous colorions $\mathcal{C}_n$ avec quatre couleurs en utilisant le motif suivant :

$$[124341232][42](1232)^*$$

où un motif entre crochet, $[M]$ signifie que l'on peut prendre ou ne pas prendre le motif M, et où $(M)^*$ signifie que l'on peut prendre le motif M autant de fois que l'on veut, ou ne pas l'utiliser. La coloration obtenue est toujours localement identifiante, et donne une coloration en trois couleurs si $n \equiv 0 \bmod 4$ et quatre couleurs sinon.

Si $n \not\equiv 0 \bmod 4$, il n'y a pas de coloration localement identifiante avec trois couleurs. En effet, essayons de construire une telle coloration. Nous n'avons pas le choix : comme pour le chemin dans la preuve de la proposition 5.27, nous aurons, sans perte de généralité, le sommet v_i colorié avec la couleur 1 si $i \equiv 0 \bmod 4$, 2 si $i \equiv 1, 3 \bmod 4$ et 3 si $i \equiv 2 \bmod 4$. Mais alors v_{n-1} doit avoir la couleur 2, et v_{n-2} doit avoir la couleur 3. Cela implique que $n - 2 \equiv 2 \bmod 4$ et donc $n \equiv 0 \bmod 4$, ce qui est contradictoire.

Enfin, supposons que $n = 5$ ou $n = 7$. S'il existait une coloration localement identifiante avec quatre couleurs, deux sommets auraient la même couleur. Ils ne peuvent pas être adjacent ni à deux sommets d'écart. Cela n'est pas possible pour $n = 5$. Pour $n = 7$ la seule possibilité est de prendre des sommets à distance 2. Pour une couleur nous ne pouvons alors avoir que deux sommets de cette couleur, et il faut donc avoir trois couples de sommets à distance 2 de la même couleur. Alors nécessairement deux de ces paires sont imbriquées l'une dans l'autre et les deux sommets au centre ne sont pas séparés. Ainsi $\chi_{lid}(\mathcal{C}_n) \geq 5$ pour $n = 5$ et $n = 7$. Enfin, il existe des colorations en cinq couleurs pour ces graphes. □

En coloriant chaque composante connexe d'un graphe de degré maximum 2 séparément, nous obtenons le résultat suivant :

Corollaire 5.29 *Si G est de degré maximum 2, $\chi_{lid}(G) \leq 5$.*

5.4 Étude dans les graphes bipartis et complexité

Cette partie est dédiée à l'étude des graphes bipartis. Nous donnons tout d'abord quelques résultats montrant la pertinence des graphes bipartis dans l'étude des colorations identifiantes.

5.4.1 Motivations

Proposition 5.30 *Si G est un graphe connexe 3-lid-coloriable, alors G est soit un triangle soit un graphe biparti.*

Preuve : Considérons une coloration localement identifiante c de G avec trois couleurs 1,2,3.

Supposons dans un premier temps que G est sans jumeaux. La coloration c' définie par $c'(x) = |c(N[x])|$ pour chaque sommet x est une coloration propre de G avec deux couleurs 2 et 3 (en supposant que G ne soit pas un sommet isolé). En effet, si deux sommets u et v satisfont $c'(u) = c'(v) = 3$ alors $c(N[u]) = c(N[v]) = \{1, 2, 3\}$ donc u et v ne peuvent pas être adjacents. Si $c'(u) = c'(v) = 2$ et que u et v sont adjacents, alors sans perte de généralité, $c(N[u]) = \{1, 2\}$, $c(N[v]) = \{1, 3\}$. Mais alors $c(u) = c(v) = 1$ ce qui est contradictoire. Donc G est un graphe biparti.

Supposons maintenant que G ait au moins une paire u,v de jumeaux : $N[u] = N[v]$ et que G n'est pas seulement constitué de l'arête uv. Le graphe G étant connexe, il y a un sommet w qui est relié à u et v. Alors $c(N[u]) = c(N[v]) = c(N[w]) = \{1, 2, 3\}$. Donc nécessairement, $N[w] = N[u] = N[v]$. Si G n'est pas un triangle, il existe un quatrième sommet z relié aux trois sommets u,v,w. Mais alors nous avons besoin d'au moins quatre couleurs. □

Cette proposition a pour conséquence que le seul graphe G tel que $\chi(G) = \chi_{lid}(G) = 3$ est le triangle, comme annoncé dans la partie 5.3.2.

La proposition suivante précise les classes de couleurs d'un graphe biparti 3-lid-coloriable :

Proposition 5.31 *Soit $G = (U \cup V, E)$ un graphe biparti connexe 3-lid-coloriable. Soit c une coloration localement identifiante en trois couleurs 1,2,3. Si pour un sommet $u \in U$, $c(N[u]) = \{1, 2, 3\}$, alors :*

- *Les sommets de U sont tous de la même couleur : $c(U) = \{c(u)\}$.*
- *Les sommets de V sont des deux autres couleurs : $c(V) = \{1, 2, 3\} \setminus \{c(u)\}$.*
- *Chaque sommet u' de U vérifie $c(N[u']) = \{1, 2, 3\}$.*

Preuve : Comme $c(N[u]) = \{1, 2, 3\}$, G a au moins trois sommets. Le graphe G est biparti et connexe, donc ne possède pas de jumeaux. Par conséquent, toute paire de sommets adjacents est séparée par la coloration c.

Sans perte de généralité, supposons que $c(u) = 1$. Tous les voisins de u sont alors coloriés avec les couleurs 2 et 3, et ne doivent avoir que deux couleurs dans leur voisinage pour être séparés de u. Donc tous les sommets à distance 2 de u sont aussi de couleur 1. En itérant ce procédé, tous les sommets à distance paire de u doivent avoir la couleur 1 et tous ceux à distance impaire doivent être de couleur 2 ou 3. Cela montre les deux premiers points de la proposition.

Pour le dernier point : supposons qu'un sommet u' de U soit tel que $c(N[u']) = \{1, 2\}$. Alors il a un voisin $v \in V$ de couleur 2. Nécessairement, $c(N[v]) = \{1, 2\}$ et u' et v ne sont pas séparés. C'est impossible donc tous les sommets de U ont les trois couleurs dans leur voisinage. □

5.4.2 Coloration en quatre couleurs des graphes bipartis

Pour colorier les graphes bipartis, nous utilisons le même type de coloration que pour les chemins dans la proposition 5.27 :

Proposition 5.32 *Tout graphe biparti G est 4-lid-coloriable.*

Preuve : Supposons G connexe (sinon nous colorions chaque composante de manière indépendante). Si G a quatre sommets ou moins le résultat est clairement vrai. Supposons donc que G a au moins cinq sommets. Alors il existe un sommet u qui n'est pas adjacent a un sommet de degré 1 (u lui même pouvant être de degré 1).

Pour chaque sommet v de G , nous donnons à v la couleur parmi $\{0, 1, 2, 3\}$ qui est congruente à $d(u, v)$ modulo 4. Ainsi u a la couleur 0, les sommets à distance 1 de u sont coloriés avec la couleur 1, ceux à distance 2 avec la couleur 2,... Montrons que la coloration c ainsi obtenue est localement identifiante. Le graphe étant biparti, deux sommets à la même distance de u ne sont pas adjacents. Plus précisément, il n'y a d'arête entre deux sommets v et w que si leur distance à u diffère d'exactement 1. La coloration c est donc une coloration propre.

Soient maintenant v et w deux sommets adjacents tels que $N[v] \neq N[w]$. Sans perte de généralité, v est à distance k de u et w à distance $k + 1$. Si $k = 0$, $v = u$ et w n'est pas de degré 1. Donc $c(N[v]) = \{0, 1\}$ et $c(N[w]) = \{0, 1, 2\}$, les sommets sont séparés. Sinon la couleur $(k - 1)$ modulo 4 est dans le voisinage de v mais pas de w et les sommets sont aussi séparés. □

5.4.3 Lien avec la coloration d'hypergraphe et complexité

D'après ce qui précède, tout graphe biparti connexe se colorie avec quatre couleurs. Nous allons voir dans cette partie qu'il est « difficile » de décider si un graphe biparti peut se colorier avec trois couleurs, et ce même pour des graphes bipartis. Pour cela nous utilisons la 2-coloration d'hypergraphes. Un hypergraphe $\mathcal{H}$ est 2-coloriable si ses sommets peuvent être coloriés avec deux couleurs de manière à ce que chaque hyperarête contienne les deux couleurs. Le problème de savoir si un hypergraphe est 2-coloriable est $\mathcal{NP}$-complet, même lorsque toutes les hyperarêtes sont de taille 3 (Lovász [73]).

Comme vu au chapitre 1, deux hypergraphes sont associés à un graphe biparti $G = (U \cup V, E)$: l'hypergraphe dont les sommets sont l'ensemble U, que nous noterons dans la suite $\mathcal{H}_U$, et celui dont les sommets sont l'ensemble V, noté $\mathcal{H}_V$.

La proposition 5.31 se traduit de la manière suivante en terme de 2-coloration d'hypergraphes :

Proposition 5.33 *Soit $G = (U \cup V, E)$ un graphe biparti connexe avec au moins trois sommets. G est 3-lid-coloriable si et seulement si $\mathcal{H}_U$ ou $\mathcal{H}_V$ est 2-coloriable.*

Preuve : Supposons que $\mathcal{H}_U$ ait une 2-coloration c utilisant les couleurs 2 et 3. Une coloration localement identifiante de G avec trois couleurs peut être obtenue en donnant la couleur 1 aux sommets de V et à chaque sommet u de U la couleur $c(u)$. Alors chaque sommet de V aura les trois couleurs dans son voisinage et chaque sommet de U n'aura que deux couleurs. La coloration est donc localement identifiante.

L'autre sens est une traduction directe de la proposition 5.31. □

De manière réciproque à partir d'un hypergraphe $\mathcal{H} = (V, \mathcal{E})$, nous pouvons associer le graphe biparti d'incidence $G_{\mathcal{H}}$. Alors si $\mathcal{H}$ est 2-coloriable, $G_{\mathcal{H}}$ est 3-lid-coloriable. Cependant le contraire n'est pas vrai car cela peut très bien être

l'hypergraphe dual de $\mathcal{H}$ qui est 2-coloriable (le rôle des arêtes et des sommets est échangé). Pour pouvoir réduire à partir de la 2-coloration d'hypergraphe nous avons besoin de forcer une des parties de $G_{\mathcal{H}}$ à être unicolore (celle correspondant aux hyperarêtes de l'hypergraphe). Pour cela, nous ajoutons deux sommets u et v au graphe $G_{\mathcal{H}}$: u est adjacent à un des sommets correspondant à un sommet de l'hypergraphe, et v est seulement adjacent à u. Notons $G'_{\mathcal{H}}$ le graphe ainsi obtenu. La figure 5.14 donne un exemple de construction de $G'_{\mathcal{H}}$ à partir d'un hypergraphe.

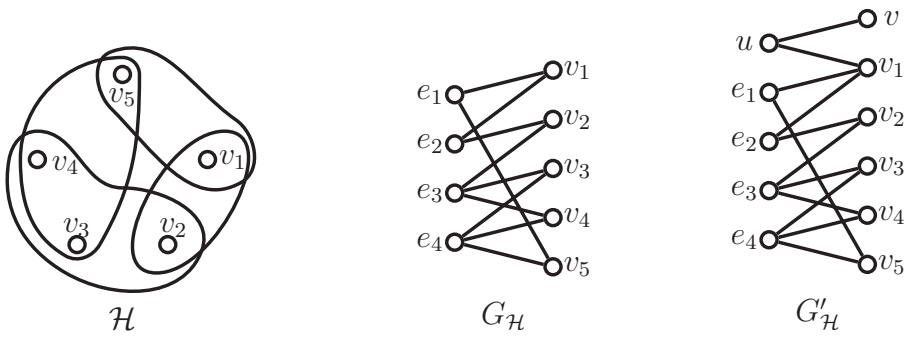


Figure 5.14 – *Réduction du problème de 2-coloration d'hypergraphe au problème de 3-lid-coloration dans les graphes bipartis.*

Lemme 5.34 *Soit $\mathcal{H}$ un hypergraphe. $\mathcal{H}$ est 2-coloriable si et seulement si $G'_{\mathcal{H}}$ est 3-lid-coloriable.*

Preuve : Supposons que $\mathcal{H}$ soit 2-coloriable. D'après la proposition 5.33, $G_{\mathcal{H}}$ est alors 3-lid-coloriable et nous pouvons le colorier avec les couleurs 1,2,3 de manière à ce que tous les sommets correspondant aux hyperarêtes aient la même couleur, 1, et aient trois couleurs dans leur voisinage. Soit w le sommet de $G_{\mathcal{H}}$ relié à u, supposons qu'il est de couleur 2. Nous complétons alors la coloration dans $G'_{\mathcal{H}}$ en donnant la couleur 1 à u et la couleur 3 à v. La coloration reste localement identifiante.

Réciproquement, supposons que c soit une 3-lid-coloration de $G'_{\mathcal{H}}$. Le sommet v étant de degré 1, il a nécessairement seulement deux couleurs dans son voisinage. Le sommet u a donc les trois couleurs dans voisinage. D'après la proposition 5.31, tous les sommets de la bipartition correspondant à u sont donc coloriés comme u et ont les trois couleurs dans leur voisinage. Les sommets de cette bipartition correspondent aux hyperarêtes de $\mathcal{H}$. Les autres sommets sont coloriés avec les deux autres couleurs. Nous définissons une coloration de $\mathcal{H}$ en donnant à chaque sommet la couleur du sommet correspondant dans $G'_{\mathcal{H}}$. Chaque hyperarête contient les deux couleurs et nous avons donc une 2-coloration de $\mathcal{H}$. □

Théorème 5.35 *Décider si un graphe biparti est 3-lid-coloriable est $\mathcal{NP}$-complet.*

Preuve : Le problème est clairement dans $\mathcal{NP}$. La 2-coloration d'hypergraphe se réduit à ce problème en utilisant le lemme 5.34, la transformation de $\mathcal{H}$ à $G'_{\mathcal{H}}$ étant clairement polynomiale. □

Notons que la classe de graphes pour laquelle ce problème est $\mathcal{NP}$-complet peut être réduite. Par exemple, pour un entier g fixé, le problème de décider si un graphe est 3-lid-coloriable reste $\mathcal{NP}$-complet pour la classe des graphes bipartis de maille au moins g. Pour cela, nous modifions le graphe $G'_{\mathcal{H}}$ en remplaçant chaque sommet de la partie "sommets" par un long chemin. Le "sommet" est alors relié à son arête correspondante tous les quatre sommets (ou plus, mais un multiple de 4), en partant du premier sommet. Comme un chemin a une coloration localement identifiante unique en trois couleurs (lorsqu'elle existe), nous assurerons que tous les sommets du chemin qui sont connectés au reste du graphe auront la même couleur.

Ce lien fort avec les hypergraphes nous donne des résultats encore plus précis pour certaines classes de graphes pour lesquelles nous connaissons des résultats pour la 2-coloration des hypergraphes correspondants. Nous donnons trois exemples de ce type. Le premier concerne la coloration de graphes bipartis planaires de degré maximum 3. Moret [78] a montré que l'on peut vérifier en temps polynomial si un hypergraphe planaire (i.e., son graphe biparti d'incidence est planaire) de degré maximum 3 est 2-coloriable. En utilisant la proposition 5.33, nous avons donc :

Proposition 5.36 *Vérifier si un graphe planaire biparti de degré maximum* 3 *est* 3*-lid-coloriable se fait en temps polynomial.*

Par ailleurs, Burstein [14] et Penaud [82] ont montré que tout hypergraphe planaire dont les hyperarêtes sont de taille au moins 3 sont 2-coloriables. Cela implique directement le résultat suivant :

Proposition 5.37 *Soit $G = (U \cup V, E)$ un graphe planaire biparti tel que tous les sommets de U ou tous les sommets de V soient de degré au moins 3. Alors G est* 3*-lid-coloriable.*

Et enfin, Thomassen [92] a montré que tous les hypergraphes k-réguliers k-uniformes, pour $k \geq 4$ sont 2-coloriables. Nous avons donc :

Proposition 5.38 *Soit G un graphe k-régulier biparti, avec $k \geq 4$. Alors G est* 3*-lid-coloriable.*

5.4.4 Cas des arbres

Décider en temps polynomial si un arbre est 3-lid-coloriable est un problème polynomial. Nous pouvons donner une condition nécessaire et suffisante simple pour qu'un arbre soit 3-lid-coloriable. La coloration construite lorsqu'il existe une coloration en trois couleurs suit le même schéma que pour le chemin : les sommets sont coloriés en fonction de leur distance à un sommet de référence.

Proposition 5.39 *Un arbre T est* 3*-lid-coloriable si pour toute paire de feuilles u,v de T, la distance entre u et v est paire.*

Preuve : Pour chaque feuille u de T nous avons dans toute coloration c localement identifiante $|c(N[u])| = 2$.

Supposons que T soit 3-lid-coloriable. D'après la proposition 5.31, toutes les feuilles de T sont du même côté de la bipartition lorsque T est vu comme un graphe biparti. Cela signifie exactement que la distance entre toute paire de feuilles est paire.

Supposons maintenant que la distance entre toute paire de feuilles de T soit paire. Considérons une feuille u de T et colorions les sommets de T en fonction de leur distance à u. Soit v un sommet de T. Si $d(u,v) \equiv 1 \bmod 2$ nous donnons la couleur 1 à v, si $d(u,v) \equiv 0 \bmod 4$ nous donnons la couleur 2 à v et sinon nous donnons la couleur 3 à v. Alors tous les sommets à distance impaire k de u ne sont pas des feuilles et ont donc un voisin à distance $k-1$ de u et un voisin à distance $k+1$. Ils ont donc les trois couleurs dans leur voisinage et par ailleurs forment un stable. Les sommets à distance paire de u ont tous deux couleurs dans leur voisinage et sont aussi un stable. La coloration obtenue est donc localement identifiante et utilise trois couleurs. □

Cette proposition appliquée dans le cas particulier des chemins nous redonne la proposition 5.27.

5.4.5 Produit cartésien de graphes bipartis

Une classe particulière de graphes bipartis 3-lid-coloriables sont les produits cartésiens de deux graphes bipartis.

Proposition 5.40 *Soient G_1 et G_2 deux graphes bipartis sans sommets isolés. Alors $G_1 \square G_2$ est 3-lid-coloriable.*

Preuve : Soient $\{U_1, V_1\}$ et $\{U_2, V_2\}$ les bipartitions respectives de G_1 et G_2. Alors $G_1 \square G_2$ est un graphe biparti de bipartition $\{(U_1 \times U_2) \cup (V_1 \times V_2), (U_1 \times V_2) \cup (V_1 \times U_2)\}$. Comme G_1 et G_2 n'ont pas de sommets isolés, chaque sommet de $(U_1 \times U_2) \cup (V_1 \times V_2)$ a un voisin dans $U_1 \times V_2$ et un voisin dans $V_1 \times U_2$.

Les sommets de $G_1 \square G_2$ sont coloriés de la manière suivante : les sommets dans $(U_1 \times U_2) \cup (V_1 \times V_2)$ sont coloriés avec la couleur 1, les sommets dans $U_1 \times V_2$ avec la couleur 2 et les sommets dans $V_1 \times U_2$ avec la couleur 3. La coloration c ainsi obtenue est localement identifiante : $c(N[u]) = \{1,2,3\}$ pour les sommets dans $(U_1 \times U_2) \cup (V_1 \times V_2)$, $c(N[u]) = \{1,2\}$ pour les sommets dans $U_1 \times V_2$ et $c(N[u]) = \{1,3\}$ pour les sommets dans $U_2 \times V_1$. □

Comme corollaire, toutes les grilles et hypercubes de toute dimension sont 3-lid-coloriables. La coloration obtenue pour la grille bidimensionnelle est celle donnée dans la figure 5.7.

5.5 Étude dans les graphes parfaits

Nous nous intéressons maintenant à la classe des graphes parfaits (définie dans le chapitre 1), qui généralise celle des graphes bipartis. Nous étudions différentes sous-classes de graphes parfaits. Les définitions de chaque classe sont rappelées au début de chaque étude. Les trois premières classes étudiées sont des sous-classes de graphes

triangulés (ou encore graphes *chordaux*). Un graphe est triangulé si tout cycle d'ordre au moins 4 possède une corde. La figure 5.16 (page 200) montre l'imbrication des classes entre elles et recense les résultas obtenus.

5.5.1 k-arbres

Un *k-arbre* G est construit de manière inductive à partir d'une $(k+1)$-clique en ajoutant à chaque étape un sommet dont le voisinage ouvert est une k-clique dans le graphe précédent (à la différence des graphes triangulés qui peuvent aussi être construit inductivement en ajoutant à chaque étape un sommet dont le voisinage est une clique, sans contrainte de taille). Cela donne un ordre naturel des sommets de G, $v_0,...,v_{n-1}$, où le voisinage de v_i (pour $i > k$) dans le graphe G_i induit par les sommets de v_0 à v_i est une k-clique. Nous utiliserons cet ordre des sommets dans la suite. La plus grande clique d'un k-arbre G est de taille $\omega(G) = k+1$.

Les 1-arbres correspondent à la définition classique des arbres, pour lesquels la proposition 5.39 nous dit qu'ils sont tous 4-lid-coloriables. Ce résultat se généralise de la manière suivante :

Théorème 5.41 *Si G est un k-arbre, $\chi_{lid}(G) \leq 2k+2 = 2\omega(G)$.*

Pour montrer ce résultat, nous avons besoin du lemme suivant :

Lemme 5.42 *Toute k-clique d'un k-arbre est contenu dans une $(k+1)$-clique.*

Preuve : Soit C une k-clique d'un k-arbre G. Considérons les sommets de G dans l'ordre naturel dû à la construction inductive de G : $v_0,...,v_{n-1}$. Soit v_j le dernier sommet dans cet ordre qui soit dans C. La clique C est incluse dans le voisinage fermé de v_j dans G_j, qui par construction est une $(k+1)$-clique. □

Pour montrer le théorème 5.41, nous construisons de manière inductive une coloration utilisant $2k+2$ couleurs dans laquelle les couleurs sont associées par deux de sorte qu'il n'y ait aucune arête dont les extrémités ont deux couleurs associées. Cela était déjà le cas dans la coloration en quatre couleurs des chemins (ainsi que des graphes bipartis) où la couleur 1 et la couleur 3 (resp. la couleur 0 et la couleur 2) ne sont jamais données à deux sommets adjacents.

Preuve du théorème 5.41: Soit G un k-arbre. Dans cette preuve, nous construisons une coloration localement identifiante de G avec les couleurs de 0 à $2k+1$. Les couleurs sont considérées modulo $2k+2$. En particulier, la fonction $x \to x+k+1$ sur les couleurs est une involution.

Considérons les sommets de G dans l'ordre naturel : $v_0,...v_{n-1}$. Le sous-graphe G_i induit par les sommets de v_0 à v_i est aussi un k-arbre. La coloration c de G est construite de manière itérative. Pour $i \leq k$, $c(v_i) = i$. Soit maintenant $i \geq k+1$. Soit C le voisinage ouvert de v_i dans G_i. C est une k-clique dans le k-arbre G_{i-1}. D'après le lemme 5.42, il existe un sommet v_j de G_{i-1} tel que $C \cup \{v_j\}$ soit une $(k+1)$-clique dans G_{i-1}. Nous posons alors $c(v_i) = c(v_j) + k + 1$ (il peut y avoir plusieurs possibilités pour le choix de v_j et donc plusieurs colorations possibles).

Montrons que c est une coloration localement identifiante. Pour cela, nous montrons que les deux faits suivant sont satisfait tout au long de la procédure :

(i) c est une coloration propre.

(ii) Aucun sommet colorié avec i n'est adjacent à un sommet colorié $i+k+1$.

Les propriétés (i) et (ii) sont vraies lorsque l'on colorie les $k+1$ premiers sommets. Lorsque v_i est colorié avec la couleur $c(v_j)+k+1$, nous ne créons pas d'arêtes $c(v_i)(c(v_i)+k+1)$. En effet, aucun sommet de la clique C n'a la couleur $c(v_i)+k+1 = c(v_j)$ car la coloration c est propre dans G_{i-1} (fait (i)). De plus, tous les sommets de C sont adjacent à v_j, et d'après (ii) n'ont donc pas la couleur $c(v_j)+k+1 = c(v_i)$. Donc la coloration reste propre.

Pour montrer que la coloration est localement identifiante, prenons deux sommets v_i et v_j adjacents, tels que $N[v_i] \neq N[v_j]$. Supposons tout d'abord que $i, j \leq k$. Sans perte de généralité, supposons qu'il existe un sommet adjacent à v_i mais pas à v_j. Considérons alors le plus petit indice ℓ tel que v_ℓ soit un voisin de v_i mais pas de v_j. Par définition de c, nous avons $c(v_\ell) = c(v_j)+k+1$ (v_j est le seul sommet que l'on peut prendre ici pour v_ℓ lors de la définition de c).

Supposons maintenant que l'un des i, j est strictement plus grand que k. Sans perte de généralité, supposons que $i < j$ et que $j > k$. Soit C le voisinage ouvert de v_j dans G_j. Par définition de c, il y a un sommet v_ℓ dans G_{j-1} tel que $C \cup \{v_\ell\}$ soit une $(k+1)$-clique et tel que $c(v_j) = c(v_\ell)+k+1$.

Dans tous les cas, $c(v_j) = c(v_\ell)+k+1$, et la couleur $c(v_\ell)$ est dans le voisinage de v_i mais pas de v_j d'après le fait (ii), et donc $c(N[v_i]) \neq c(N[v_j])$. □

Cette construction donne dans le cas des arbres la même coloration que dans la preuve de la propriété 5.39. De plus cette borne est serrée : le graphe P^k_{2k+2} est un k-arbre et d'après la proposition 5.13 $\chi_{lid}(P^k_{2k+2}) = 2k+2$.

Étant donné un entier ℓ fixé, la propriété « G est ℓ-lid-coloriable » peut s'exprimer en CMSOL (voir le chapitre 3 page 93). En effet, tester si ℓ ensembles E_1,...,E_ℓ forment une partition s'exprime facilement en CMSOL, ce qui nous une coloration propre. Pour tester si elle est localement identifiante, on peut tester pour chaque paire de sommets $\{x, y\}$, s'ils sont jumeaux et si il existe une couleur qui les sépare. Pour cette dernière partie, il suffit de tester pour chaque ensemble E_i, s'il existe un sommet $z \in E_i$ qui soit adjacent à x et aucun sommet de E_i n'est adjacent à y, ou le contraire. Si le résultat est positif pour l'un des E_i alors les sommets sont séparés.

D'après le théorème 3.38, il peut être vérifié en temps linéaire si un graphe est ℓ-lid-coloriable dans la classe des graphes de largeur d'arborescence fixée. Les k-arbres étant de largeur d'arborescence $k+1$, le problème des colorations localement identifiante est linéaire dans cette classe.

5.5.2 Graphes d'intervalles

Nous nous intéressons maintenant à la classe des graphes d'intervalles (définie au chapitre 1). Les graphes d'intervalles sont des graphes triangulés. L'objectif de cette partie est de montrer le résultat suivant :

Théorème 5.43 *Pour tout graphe d'intervalles G, $\chi_{lid}(G) \leq 2\omega(G)$.*

Pour cela, nous utilisons une technique similaire à celle utilisée pour les k-arbres : une coloration localement identifiante est construite de manière inductive de sorte

que pour chaque couleur, une autre couleur lui est associée dans le sens où il n'y a pas d'arêtes dont les extrémités sont coloriées avec les deux couleurs associées.

Preuve : Soit G un graphe d'intervalles connexe d'ordre n. Dans cette preuve, les couleurs sont des entiers pris modulo $2k$ où $k = \omega(G)$. Les sommets v_1,...,v_n de G sont identifiés avec la famille d'intervalles $\mathcal{I} = (I_i = [a_i, b_i])_{1 \leq i \leq n}$ de telle sorte que $v_i v_j$ soit une arête de G lorsque I_i et I_j s'intersectent. Supposons que $a_1 \leq a_2 \leq ... \leq a_n$. Sans perte de généralité, supposons que si $a_i < a_j$ et que $I_i \cap I_j = \emptyset$, alors il existe un intervalle I_ℓ tel que $a_i \leq b_\ell < a_j$; sinon, nous pouvons changer l'intervalle I_j pour l'intervalle $[a_i, b_j]$ et le graphe d'intersection G reste le même. Par un argument similaire, nous pouvons aussi supposer que si $b_j < b_i$ et $I_i \cap I_j \neq \emptyset$, alors il existe un intervalle I_ℓ tel que $b_j < a_\ell \leq b_i$.

Soit $\{a_{t_1} = a_1 < a_{t_2} < ... < a_{t_s} = a_n\}$ l'ensemble des points de départs des intervalles de la famille $\mathcal{I}$. Notons $\mathcal{I}(t_i)$, $1 \leq i \leq s$, l'ensemble des intervalles commençant à a_{t_i}. Remarquons qu'il y a au plus k intervalles dans $\mathcal{I}(t_i)$. Nous construisons une coloration c de G en s étapes : à l'étape i, $1 \leq i \leq s$, tous les intervalles de $\mathcal{I}(t_i)$ sont coloriés. Nous commençons par colorier à l'étape 1 les intervalles de $\mathcal{I}(t_1)$ avec des couleurs comprises entre 0 et $k-1$. Supposons maintenant que tous les intervalles commençant strictement avant a_{t_i}, $1 < i \leq s$, sont coloriés. Nous allons colorier les sommets de $\mathcal{I}(t_i)$. Pour cela, nous avons besoin de définir les sous-ensembles d'intervalles suivant :

- $\mathcal{V}(t_i)$: intervalles I_j commençant strictement avant a_{t_i} et finissant après : $a_j < a_{t_i} \leq b_j$;
- $\mathcal{U}(t_i)$: intervalles I_j finissant entre $a_{t_{i-1}}$ et a_{t_i}, a_{t_i} pris strictement : $a_{t_{i-1}} \leq b_j < a_{t_i}$;
- $\mathcal{T}(t_i)$: intervalles I_j de $\mathcal{U}(t_i)$ pour lesquels il y a un intervalle I_ℓ de $\mathcal{V}(t_i)$ commençant à la même date.

Notons que $\mathcal{V}(t_i)$ est exactement l'ensemble des intervalles déjà coloriés qui intersectent $\mathcal{I}(t_i)$. Les intervalles de $\mathcal{U}(t_i)$ sont déjà coloriés et intersectent tous les intervalles de $\mathcal{V}(t_i)$. Cet ensemble est non vide : l'intervalle qui finit le plus tard mais avant a_{t_i} est dans cet ensemble. Remarquons que tous les intervalles de $\mathcal{U}(t_i)$ finissent en même temps car aucun intervalle ne commencent entre $a_{t_{i-1}}$ et a_{t_i}.

Si $\mathcal{T}(t_i)$ n'est pas vide, posons I_0 un des intervalles de $\mathcal{T}(t_i)$ commençant le plus tôt, sinon prenons pour I_0 n'importe quel intervalle de $\mathcal{U}(t_i)$. Soit c_0 la couleur de I_0. Tous les intervalles de $\mathcal{V}(t_i)$ et $\mathcal{U}(t_i)$ intersectent I_0 et ont donc la couleur c_0 dans leur voisinage.

Nous pouvons maintenant colorier les sommets de $\mathcal{I}(t_i)$. La couleur $c_0 + k$ est donnée à l'un des intervalles terminant en dernier. Les autres intervalles sont coloriés avec des couleurs allant de 0 à $2k-1$ de sorte que la coloration reste propre et qu'il n'y ait pas d'arêtes dont les couleurs aux extrémités diffèrent d'exactement k. Cela est toujours possible car $\mathcal{V}(t_i) \cup \mathcal{I}(t_i)$ est de taille au plus k.

La coloration c ainsi obtenue est une coloration propre de G utilisant $2k$ couleurs, sans arêtes dont les couleurs aux extrémités diffèrent d'exactement k. Montrons qu'elle est localement identifiante. Soient I_i et I_j deux intervalles s'intersectant tels que $N[I_i] \neq N[I_j]$. Nous pouvons supposer que $a_i \leq a_j$.

Traitons tout d'abord le cas où $a_i < a_j$. Pendant la procédure, lorsque I_j est colorié, un intervalle I_ℓ, commençant à la date a_j est colorié avec une couleur $c_0 + k$ (cela peut éventuellement être I_j), tel que c_0 soit une couleur du voisinage de I_i. Nécessairement $I_j \subseteq I_\ell$ car I_ℓ termine en dernier parmi les intervalles commençant à la date a_j, et donc $c(N[I_j]) \subseteq c(N[I_\ell])$. Or $c_0 \notin c(N[I_\ell])$, donc $c_0 \notin c(N[I_j])$, et les intervalles I_i et I_j sont séparés.

Supposons maintenant que $a_i = a_j$. Nous pouvons alors supposer que $b_j < b_i$ et donc que $I_j \subset I_i$. Soit a_{t_ℓ} la date de début d'intervalles la plus petite telle que $b_j < a_{t_\ell} \leq b_i$ (cette date existe car $N[I_i] \neq N[I_j]$). Alors $I_i \in \mathcal{V}(t_\ell)$ et $I_j \in \mathcal{T}(t_\ell)$. Par construction, un des intervalles de $\mathcal{I}(t_\ell)$, est colorié avec une couleur $c_0 + k$ où c_0 est la couleur d'un intervalle I_0 de $\mathcal{T}(t_\ell)$. Nécessairement, $I_j \subseteq I_0$ et $c_0 + k \notin c(N[I_0])$, donc $c_0 + k \notin c(N[I_j])$. Or $c_0 + k \in c(N[I_i])$, ce qui termine la preuve. □

Notons que cette borne est serrée car les graphes P_{2k}^{k-1} sont des graphes d'intervalles.

5.5.3 Graphes scindés

Un *graphe scindé* (ou graphe *split*) est un graphe $G = (V, E)$ où l'ensemble des sommets peut être partionné en une clique K et un ensemble indépendant S. Le but de cette partie est de montrer le théorème suivant :

Théorème 5.44 *Soit $G = (K \cup S, E)$ un graphe scindé. Si $|K| \neq 2$ ou si $G = K_{1,n}$, alors $\chi_{lid}(G) \leq 2\omega(G) - 1$. Sinon G est un arbre et $\chi_{lid}(G) \leq 2\omega(G)$.*

Dans la suite, nous considérerons à chaque fois des partitions $K \cup S$ de V avec K de taille maximum (dans une partition quelconque un sommet de S peut être adjacent à toute la clique K qui n'est alors pas de taille maximum). Un graphe scindé est un graphe triangulé dont le nombre de clique et le nombre chromatique sont égaux à $|K|$ (avec la convention précédente sur K). Pour montrer le théorème 5.44 nous avons besoin d'un résultat intermédiaire, conséquence directe du théorème de Bondy [12] dont nous avons rappelé l'énoncé au chapitre 2 (page 32).

Dans un graphe scindé, nous étendons la définition d'un ensemble séparant en disant qu'un ensemble $S' \subseteq S$ *sépare* un ensemble $K' \subseteq K$ si pour tous éléments distincts u et v de K' avec $N[u] \neq N[v]$, $N[u] \cap S' \neq N[v] \cap S'$. Le lemme suivant est similaire au corollaire 2.3 du chapitre 2.

Lemme 5.45 *Soit $G = (K \cup S, E)$ un graphe scindé. Pour tout $K' \subset K$, il existe un sous-ensemble S' de S de taille au plus $|K'| - 1$ tel que S' sépare K'.*

Preuve : Nous appliquons le théorème de Bondy [12] (voir page 32) avec pour l'ensemble $\mathcal{A}$ l'ensemble S. Pour les ensembles A_i, nous prenons les ensembles distincts $N(v_i) \cap S$ pour $v_i \in K'$. D'après le théorème de Bondy, il existe un ensemble $S' \subseteq S$ de taille au plus $|K'| - 1$ tels que les ensembles $A_i \cap S'$ soient distincts. Or $A_i \cap S' = N(v_i) \cap S'$ et les sommets de K' sont tous adjacents deux à deux. Donc les ensembles $N[v_i] \cap S'$ sont tous distincts, ce qui finit la preuve. □

À partir du lemme 5.45, nous pouvons montrer que tout graphe scindé peut se colorier avec $2|K|$ couleurs en donnant les couleurs $1, ..., |K|$ aux sommets de K, les

couleurs $|K|+1$,...,$|K|+k$ à un ensemble S' de taille $k < |K|$ séparant K et la couleur $|K|+k+1$ aux sommets de $S \setminus S'$. Nous pouvons économiser une couleur et obtenir le théorème 5.44.

Preuve du théorème 5.44: Lorsque $G = K_{1,n}$, nous avons vu précédemment que G est 3-lid-coloriable (voir la proposition 5.14). De plus, tout arbre est 4-lid-coloriable. Nous supposons donc que $|K| \neq 2$.

Soient $k = |K|$ et v_1,...,v_k les sommets de K. Nous supposons que $k \geq 3$, le cas $k = 1$ étant trivial. Si $|S| \leq k-1$ ou si S contient un ensemble qui sépare K de taille $k-2$, le résultat se déduit en utilisant la méthode précédente. Supposons donc $|S| \geq k$ et soit S_1 un ensemble séparant K de taille minimale. Nécessairement $|S_1| = k-1$ et pour tous u, v dans K, $N[u] \neq N[v]$ (sinon il y avait seulement $k-1$ sommets à séparer et nous aurions pu le faire avec $k-2$ sommets). Considérons deux cas.

Cas 1. Il y a un sommet $x \in S \setminus S_1$ de degré $k-1$ et un voisin $v_i \in K$ de x tel que $N[v_i] \cap S_1 = \emptyset$. Sans perte de généralité, supposons que $v_i = v_{k-1}$ et que $K \setminus N(x) = \{v_k\}$. Soit $S_x = \{y \in S, N(y) = N(x) = K \setminus \{v_k\}\}$. Nous avons $S_x \cap S_1 = \emptyset$ car v_{k-1} est adjacent à tout S_x et à aucun sommet de S_1. Par définition de S_1, pour chaque sommet $v_i \neq v_{k-1}$, $N[v_i] \cap S_1 \neq \emptyset$ car v_i doit être séparé de v_{k-1}.

Soit $K_1 = K \setminus \{v_{k-1}, v_k\}$. Soit S_2 un sous-ensemble de S_1 de taille au plus $|K_1| - 1 = k-3$ qui sépare K_1 et soit $S' = S \setminus (S_1 \cup S_x)$. La coloration c des sommets est définie de la manière suivante :

- pour $1 \leq i \leq k$, $c(v_i) = i$;
- les sommets de S_2 sont coloriés avec les couleurs $k+1$ à $2k-3$, chaque sommet ayant sa couleur propre ;
- pour $u \in S_1 \setminus S_2$, $c(u) = 2k-2$;
- pour $u \in S_x$, $c(u) = 2k-1$;
- pour $u \in S'$, soit $v_i \in K \setminus N(u)$ (v_i existe toujours par maximalité de K) et $c(u) = c(v_i)$.

Par construction, c est une coloration propre de G. Montrons que c'est une coloration localement identifiante. Pour cela, remarquons tout d'abord que pour tout sommet v_i de K, $c(N[v_i])$ contient une couleur de $\{k+1, ..., 2k-1\}$. En effet, $2k-1 \in c(N[v_k - 1])$ et si $v_i \neq v_{k-1}$, alors $N[v_i] \cap S_1 \neq \emptyset$, et donc $c(N[v_i]) \cap \{k+1, ..., 2k-2\} \neq \emptyset$. Cela implique que pour tout $v_i \in K$, $c(N[v_i])$ est distinct de tous les ensembles $c(N[y])$ avec $y \in S$: soit $c(y) \in c(K)$ et alors $c(N[y]) \subseteq c(K)$, ou bien $c(y) \notin c(K)$ mais alors il y a une couleur de $c(K)$ que $c(N[y])$ ne contient pas.

Il reste donc à vérifier que pour tous i et j, $c(N[v_i]) \neq c(N[v_j])$. L'ensemble $c(N[v_k])$ est différent de tous les ensembles $c(N[v_i])$, $i \neq k$, car $2k-1 \in c(N[v_i])$ mais $2k-1 \notin c(N[v_k])$. L'ensemble $c(N[v_{k_1}])$ est différent de tous les ensembles $c(N[v_i])$, $i \neq k-1$, car $c(N[v_{k_1}])$ ne contient aucune couleur de $c(S_1)$ alors que $c(N[v_i])$ en contient au moins une. Finalement, $c(N[v_i]) \neq c(N[v_j])$ pour $i < j < k-1$ car il y a un sommet de S_2 qui les sépare et que sa couleur n'est utilisée qu'une seule fois.

Cas 2. Pour chaque sommet x de $S \setminus S_1$, soit x est de degré au plus $k-2$, soit x est de degré $k-1$ et chaque sommet de $N(x)$ a un sommet dans S_1. La coloration c est définie de la manière suivante :

- pour $1 \leq i \leq k$, $c(v_i) = i$;
- les sommets de S_1 ont les couleurs $k+1$ à $2k-1$, chaque sommet est colorié de manière différente ;
- pour $u \in S \setminus S_1$, soit v_i un sommet dans $K \setminus N(u)$ et alors $c(u) = c(v_i)$.

Montrons que c est une coloration localement identifiante de G. Par construction, c est une coloration propre. Soit uv une arête de G avec $N[u] \neq N[v]$. Si u et v sont dans K, sans perte de généralité, il y a un sommet w de S_1 tel que $w \in N[u]$ et $w \notin N[v]$. Alors, $c(w) \in c(N[u])$ et $c(w) \notin c(N[v])$. Sinon, sans perte de généralité, $u \in K$ et $v \in S$. Si $v \in S_1$, v n'est pas adjacent à tout l'ensemble K et $c(N[v])$ ne contient pas $c(K)$, contrairement à $c(N[u])$. Supposons maintenant que $v \notin S_1$. Si v est de degré $k-1$, alors u a nécessairement un voisin w dans S_1, et $c(w) \in c(N[u]) \setminus c(N[v])$. Si v est de degré strictement inférieur à $k-1$, il y a une couleur i entre 1 et k, telle que $i \in c(N[u]) \setminus c(N[v])$. Dans tous les cas, $c(N[u]) \neq c(N[v])$, ce qui termine la preuve. □

Notons que cette borne est serrée avec l'exemple de la figure 5.8 lorsque $|K| > 2$ et avec le graphe P_4 pour $|K| = 2$.

5.5.4 Cographes

Un *cographe* est un graphe qui ne contient pas le graphe P_4 comme sous-graphe induit. Les cographes sont une sous-classe de graphes de permutation et sont donc des graphes parfaits. Ce ne sont pas en général des graphes triangulés. La classe des cographes est close par union disjointe et passage au complémentaire (car P_4 est auto-complémentaire).

La caractérisation suivante des cographes nous sera utile. Tout cographe G est de l'un des trois types suivant :

(S) G est un sommet isolé.

(U) $G = \bigsqcup_{i=1}^{k} G_i$ avec $k \geq 2$ et chaque G_i est un cographe de type S ou J.

(J) $G = \bowtie_{i=1}^{k} G_i$ avec $k \geq 2$ et chaque G_i est un cographe de type S ou U.

Théorème 5.46 *Soit G un cographe, alors $\chi_{lid}(G) \leq 2\omega(G) - 1$.*

Preuve : Dans cette preuve, nous considérons des colorations localement identifiantes qui ont de plus la propriété qu'un sommet qui n'est pas universel ne contient pas toutes les couleurs dans son voisinage. Une telle coloration est appelée dans cette preuve *coloration localement identifiante forte*. Nous notons $\widetilde{\chi_{lid}}(G)$ le nombre minimum de couleurs nécessaires pour avoir une telle coloration dans G. Clairement $\chi_{lid}(G) \leq \widetilde{\chi_{lid}}(G)$. Remarquons aussi que si G a un sommet universel, alors G est de type S ou J.

Nous montrons par induction le résultat suivant, qui implique le résultat du théorème :

Pour tout cographe G, $\chi_{lid}(G) \leq 2\omega(G) - 1$ et $\widetilde{\chi_{lid}}(G) \leq 2\omega(G)$.

Si G est de type S, c'est un sommet isolé, qui est universel et $\chi_{lid}(G) = \widetilde{\chi_{lid}}(G) = 1$, le résultat est vrai.

Supposons maintenant que G est de type J. Alors G s'écrit $G = \bowtie_{i=1}^{k} G_i$ où $k \geq 2$ et chaque G_i est un cographe de type S ou U. Sans perte de généralité, supposons que les graphes G_1 à G_s $(0 \leq s \leq k)$ sont de type S et que les graphes G_{s+1} à G_k sont de type U. Soit c_1 une coloration localement identifiante de G_1 et c_i une coloration localement identifiante forte de G_i pour $2 \leq i \leq k$ tels que les ensembles de couleurs utilisées pour chaque graphe soient disjoints. Nous définissons une coloration c de G par $c(v) = c_i(v)$ si v est un sommet de G_i. La coloration c est clairement une coloration propre. Montrons que c est une coloration localement identifiante de G. Supposons par l'absurde qu'il existe deux sommets adjacents u et v tels que $N[u] \neq N[v]$ et $c(N[u]) = c(N[v])$. Comme chaque c_i est une coloration localement identifiante de G_i, les sommets u et v sont nécessairement dans des graphes G_i différents. Sans perte de généralité, supposons que u est dans G_i et v est dans G_j avec $i < j$. Pour avoir $c(N[u]) = c(N[v])$, les sommets u et v doivent avoir toutes les couleurs de G_i et G_j dans leur voisinage. Or c_j est une coloration localement identifiante forte, donc v est un sommet universel dans G_j. Donc nécessairement G_j (et donc G_i) est de type S (il n'y a pas de sommet universel dans un graphe de type U). Ainsi u et v sont universels dans G et $N[u] = N[v]$, menant à une contradiction.

Si de plus c_1 est une coloration localement identifiante forte, alors c est aussi localement identifiante forte. En effet, si un sommet v dans G_i contient toutes les couleurs dans son voisinage, alors il les contient aussi toutes dans la coloration c_i et par conséquent est universel dans G_i, et donc dans G.

Nous avons donc $\chi_{lid}(G) \leq \chi_{lid}(G_1) + \sum_{i=2}^{k} \widetilde{\chi_{lid}}(G_i)$ et $\widetilde{\chi_{lid}}(G) \leq \sum_{i=1}^{k} \widetilde{\chi_{lid}}(G_i)$. De plus, $\omega(G) = \sum_{i=1}^{k} \omega(G_i)$.

Par induction, nous avons donc :

$$\chi_{lid}(G) \leq 2\omega(G_1) - 1 + \sum_{i=2}^{k} 2\omega(G_i)) = 2 \times \sum_{i=1}^{k} \omega(G_i) - 1 = 2\omega(G) - 1$$

et

$$\widetilde{\chi_{lid}}(G) \leq \sum_{i=1}^{k} 2\omega(G_i) = 2\omega(G).$$

Supposons maintenant que G soit de type U. Alors G s'écrit $G = \bigsqcup_{i=1}^{k} G_i$ où $k \geq 2$ et chaque G_i est un cographe de type S ou U. Sans perte de généralité, supposons que $\chi_{lid}(G_1) = \max_{i=1}^{k}\{\chi_{lid}(G_i)\}$. Pour chaque graphe G_i, soit c_i une coloration localement identifiante utilisant les couleurs de 1 à $\chi_{lid}(G_i)$. La coloration c de G obtenue en posant $c(v) = c_i(v)$ lorsque v est un sommet de G_i est clairement localement identifiante, et donc $\chi_{lid}(G) = \max_{i=1}^{k}\{\chi_{lid}(G_i)\}$. Pour obtenir une coloration qui soit de plus forte, nous changeons la couleur des sommets de G_1 coloriés avec la couleur 1 par la couleur $\chi_{lid}(G_1) + 1$ et les autres sommets gardent la même couleur. La coloration c' ainsi obtenue est toujours localement identifiante, et aucun sommet de G n'a toutes les couleurs dans son voisinage. C'est donc une coloration localement identifiante forte. Ainsi $\widetilde{\chi_{lid}}(G) \leq \max_{i=1}^{k}\{\chi_{lid}(G_i)\} + 1$. Comme $\omega(G) = \max_{i=1}^{k}\{\omega(G_i)\}$, par induction :

$$\chi_{lid}(G) \leq \max_{i=1}^{k}\{2\omega(G_i) - 1\} = 2\omega(G) - 1$$

et

$$\widetilde{\chi_{lid}}(G) \leq \max_{i=1}^{k}\{2\omega(G_i) - 1\} + 1 = 2\omega(G).$$

□

La borne du théorème 5.46 est serrée en prenant l'exemple suivant. Pour tout $k \geq 1$, nous prenons le graphe complet sur les sommets v_1,...,v_k et ajoutons pour tout $2 \leq i \leq k$ un sommet u_i adjacent aux sommets v_i à v_k. Ce graphe est un cographe de nombre de clique k où les sommets u_2 à u_k forment un ensemble indépendant U. Pour tout sommet v_i, $N[v_i] \cap U = \{u_2, ..., u_i\}$. Dans une coloration localement identifiante c de G, le sommet u_i, pour $2 \leq i \leq k$ doit avoir une couleur différente des couleurs utilisées par les sommets v_1 à v_k ainsi qu'une couleur différente de celles utilisées par les sommets u_j pour $j < i$, sinon les sommets v_i et v_{i-1} ne sont pas séparés. Ainsi, tous les sommets sont de couleurs différentes et $\chi_{lid}(G) = 2k + 1$. La figure 5.15 montre la construction pour $k = 5$.

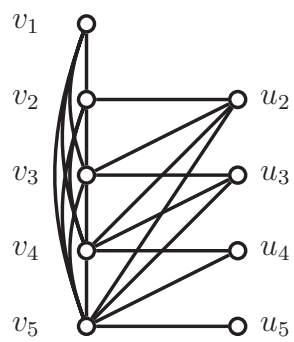


Figure 5.15 – *Exemple de cographe G vérifiant $\chi_{lid}(G) = |V(G)| = 2\omega(G) + 1$.*

5.5.5 Bilan sur les graphes parfaits et conjecture pour les graphes triangulés

Nous venons d'étudier quelques sous-classes de graphes parfaits. Les résultats obtenus sont résumés sur la figure 5.16. Ils nous incitent à penser que tout graphe parfait serait lid-coloriable avec $2\omega(G) = 2\chi(G)$ couleurs.

Cela n'est en fait pas vrai comme nous le montre l'exemple suivant. Considérons le graphe G_k ($k \geq 2$) dont les sommets sont l'union de trois ensembles indépendants, S_1, S_2 et S_3, chacun de taille k, et dont les arêtes sont :

- toutes les arêtes possibles entre S_1 et S_2 ;
- un couplage parfait entre S_1 et S_3 ;
- le complémentaire d'un couplage parfait entre S_2 et S_3.

Ce graphe est dessiné pour le cas $k = 3$ dans la figure 5.17.

Montrons que G_k est parfait. G_k est 3-coloriable car les ensembles S_1, S_2 et S_3 induisent des stables, et $\omega(G_k) = 3$. Soit H un sous-graphe induit de G_k. Si

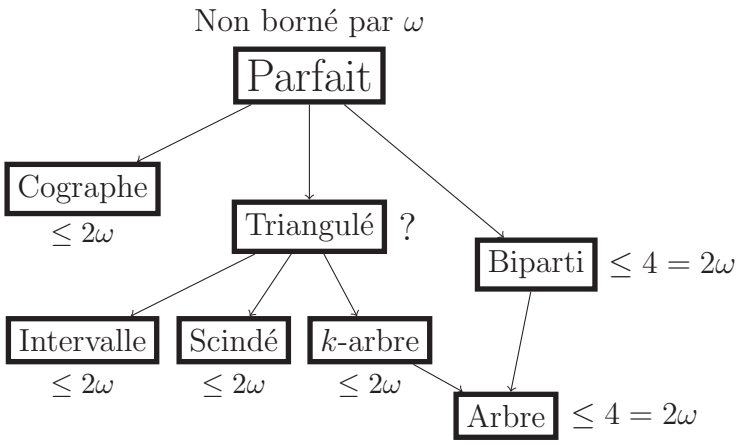


Figure 5.16 – *Résumé des résultats obtenus pour des sous-classes des graphes parfaits*

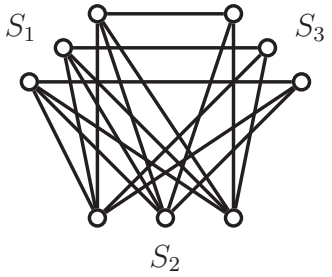


Figure 5.17 – *Construction de graphes parfaits dont le nombre lid-chromatique n'est pas borné par ω.*

$\chi(H) = 2$ ou si $\omega(H) = 3$, alors $\chi(H) = \omega(H)$. Supposons donc que H n'a pas de triangle et n'est pas biparti. Alors H contient un cycle $\mathcal{C}$ impair induit de taille au moins 5. Le cycle $\mathcal{C}$ contient nécessairement un sommet u de S_3. Les deux voisins v_1 et v_2 dans u doivent être dans la même composante S_i (sinon un triangle apparaît). Nécessairement, cette composante est S_2. Ainsi, dans $\mathcal{C}$, il n'y a aucune arête entre S_1 et S_3, et donc $\mathcal{C}$ est d'ordre pair, ce qui est contradictoire. Donc H est soit biparti, soit contient un triangle et dans tous les cas $\chi(H) = \omega(H)$.

Montrons maintenant qu'au moins $k + 2$ couleurs sont nécessaires pour colorier G_k de manière localement identifiante. Soit c une lid-coloration de G_k. Soient x_2 un sommet de S_2, x_3 le seul sommet de S_3 qui n'est pas adjacent à x_2 et x_1 l'unique voisin de x_3 dans S_1. On a $N[x_1] = N[x_3] \cup \{x_2\}$ et $c(N[x_1]) \neq c(N[x_3])$, donc la couleur de x_2 n'apparaît qu'une seule fois dans S_2. Ainsi, tous les sommets de S_2 sont de couleurs différentes. Les sommets de S_1 et S_3 doivent avoir aussi une couleur qui n'apparaît pas dans S_2. Au final $\chi_{lid}(G_k) \geq k + 2$ alors que $\chi(G_k) = \omega(G_k) = 3$.

Cependant, nous pensons que le résultat est vrai si l'on se restreint à la classe des graphes triangulés :

Conjecture 5.47 *Pour tout graphe triangulé G, $\chi_{lid}(G) \leq 2\omega(G)$.*

En effet, les techniques de preuves utilisées pour les k-arbres et pour les graphes d'intervalles pourraient peut-être se généraliser aux graphes triangulés. Dans les

deux cas, nous construisons une coloration de manière itérative en coloriant les sommets selon un ordre simplicial (le voisinage du sommet ajouté est une clique). Dans le cas des k-arbres, nous ajoutons un sommet à une clique toujours de même taille, et dans le cas des graphes d'intervalles, la taille des cliques peut varier mais nous utilisons fortement le fait que les cliques soient ordonnées. Une première piste pour cette conjecture serait de montrer le résultat pour les graphes *fortement triangulés*, pour lesquels il existe un ordre simplicial des sommets où de plus les voisinages de chacun des sommets de la clique sont inclus les uns dans les autres.

5.6 Étude dans les graphes planaires

Cette dernière partie est consacrée à l'étude des graphes planaires. Nous étudions tout d'abord les graphes *planaires extérieurs*, c'est-à-dire les graphes planaires dont il existe une représentation où tous les sommets sont sur la face extérieure.

5.6.1 Graphes planaires extérieurs

Les graphes planaires extérieurs maximaux (en terme d'arêtes) font partie de la classe des 2-arbres. En effet, tout graphe planaire extérieur à un sommet de degré 2, et s'il est de plus maximal, les deux voisins de ce sommet sont nécessairement adjacent. En enlevant ce sommet, le graphe reste planaire extérieur maximal. Ainsi, les graphes planaires extérieurs maximaux peuvent se construire en ajoutant à chaque étape un sommet adjacent à une arête située sur la face extérieure. D'après le théorème 5.41, tout graphe planaire extérieur maximal est donc 6-lid-coloriable. Cependant, la propriété d'être 6-lid-coloriable n'est pas héréditaire, ce résultat ne nous apporte donc rien sur les graphes planaires extérieurs en général. Nous allons utiliser une méthode différente pour colorier les graphes planaires extérieurs, basée sur celle utilisée pour colorier les graphes bipartis : nous allons colorier les sommets en fonction de leur distance à un sommet particulier. Cela fonctionne bien car pour les graphes planaires extérieurs, l'ensemble des sommets à même distance d'un sommet particulier est une union de chemins.

Théorème 5.48 *Tout graphe planaire extérieur connexe est* 20*-lid-coloriable.*

Preuve : Soit G un graphe planaire extérieur connexe. Soit H un graphe planaire extérieur maximal contenant G (H est obtenu en ajoutant des arêtes à G). Le graphe H est 2-connexe et de degré minimum 2. Considérons une représentation de H dans le plan telle que tous les sommets soient situés sur la face extérieure. Soit x_1 un sommet de degré 2. En suivant la face extérieure dans le sens horaire, nous obtenons un ordre $x_1, \ldots, x_n$ des sommets de H, et donc de G. Cet ordre implique les propriétés suivantes sur les sommets de G :

- Pour quatre entiers $i, j, k, \ell \in \{1, \ldots, n\}$ tels que $i < j < k < \ell$, au plus une des paires $\{x_i, x_k\}$ et $\{x_j, x_\ell\}$ correspond à une arête de G.
- Soit x_{i_0} un sommet et $x_{i_1}, \ldots, x_{i_k}$ ses voisins dans G de telle sorte que les sommets $x_{i_0}, x_{i_1}, \ldots, x_{i_k}$ apparaissent dans cet ordre sur la face extérieure de H (en suivant le sens horaire). La propriété précédente implique que, pour

$1 \leq j \leq k$, les voisins de x_{i_j} différents de x_{i_0} apparaissent (dans l'ordre horaire sur la face extérieure de H) entre $x_{i_{j-1}}$ et $x_{i_{j+1}}$ (si $j \neq k$) et entre $x_{i_{k-1}}$ et x_{i_0} (si $j = k$). De plus, deux sommets distincts x_{i_j} et x_{i_ℓ} ont au plus un voisin en commun en dehors de $N[x_{i_0}]$. Si un tel voisin commun existe, alors $|j - \ell| = 1$.

Pour $i \geq 1$, nous définissons $L_i = \{x_{i_1}, \ldots, x_{i_{k_i}}\}$ comme l'ensemble des sommets à distance i de x_1 dans G, avec $i_1 < \cdots < i_{k_i}$. Soit L_s le dernier ensemble L_i non vide. Pour plus de clarté, nous noterons $x_1^i, \ldots, x_{k_i}^i$ les sommets de L_i à la place de $x_{i_1}, \ldots, x_{i_{k_i}}$. Deux sommets x_j^i et x_{j+1}^i de L_i sont dits *consécutifs*. Nous avons les faits suivants :

- Un sommet de L_{i+1} a au plus deux voisins dans L_i.
- Deux sommets de L_i ont au plus un voisin commun dans L_{i+1}.
- Si deux sommets de L_i ont un voisin commun dans L_{i+1}, ils sont consécutifs dans L_i.
- Si deux sommets de L_i sont adjacents, ils sont consécutifs dans L_i. Ainsi le graphe induit par les sommets de L_i est une union disjointe de chemins.

Ces faits sont dûs aux propriétés énoncées précédemment, impliquées par l'ordre des sommets sur la représentation planaire de H et donc de G. Nous pouvons maintenant oublier le graphe H que nous avons construit pour établir proprement l'ordre sur les sommets de G. Avec les faits précédents, nous pouvons remarquer que dans l'ordre des sommets de L_{i+1}, nous trouvons tout d'abord les voisins de x_1^i, puis ceux de x_2^i et ainsi de suite.

Nous allons colorier les sommets de G avec vingt couleurs, partitionnées en quatre classes de cinq couleurs C_0, C_1, C_2 et C_3, avec $C_i = \{5j, \ldots, 5j+4\}$. Les sommets de L_i seront coloriés avec les couleurs de $C_{i \bmod 4}$, reprenant ainsi l'idée de la coloration des graphes bipartis (voir la partie 5.4.2). Nous utiliserons de plus des sommets *marqués*.

Nous commençons par colorier x_1 avec la couleur 0 et par marquer le dernier sommet ${x_1}^{k_1}$ de L_1. Puis nous appliquons l'algorithme 1.

Décrivons tout d'abord cet algorithme. Les ensembles de sommets L_i sont coloriés les uns après les autres (ligne 4). Lorsque l'ensemble L_i va être colorié, nous marquons tout d'abord certains sommets de L_{i+1} : les derniers voisins dans L_{i+1} des sommets de L_i (lignes 5 et 6). Puis nous colorions les sommets de L_i dans l'ordre dans lequel ils apparaissent. Il y a quatre couleurs courantes de $C_{i \bmod 4}$, c_0 à c_3, et une couleur interdite, $c_\varnothing$. Au départ, ces couleurs sont respectivement initialisées avec les valeurs $5 \times (i \bmod 4)$, $1+5\times(i \bmod 4)$, $2+5\times(i \bmod 4)$, $3+5\times(i \bmod 4)$, et $4+5\times(i \bmod 4)$. Les sommets de L_i sont coloriés suivant le motif $c_1c_2c_3c_0c_1$... (ligne 11), mais chaque fois qu'un sommet marqué v_j^i est rencontré, nous changeons de manière cyclique les valeurs de $c_{(j+1) \bmod 4}$, $c_\varnothing$, et $c_{(j-1) \bmod 4}$ (lignes 12 à 16). Cela implique les faits suivants :

- La coloration est propre.
- Quatre sommets consécutifs de L_i reçoivent quatre couleurs différentes.
- Deux sommets consécutifs de L_{i-1} n'ont pas le même ensemble de couleurs dans leur voisinage dans L_i, lorsque ces voisinages sont différents.

Ainsi, cet algorithme renvoie une coloration propre c de G avec vingt couleurs, telle que, pour tout i, $c(L_i) \subseteq C_{i \bmod 4}$. Montrons maintenant que la coloration c est

Entrée : Un graphe planaire extérieur G
Sortie : Une coloration localement identifiante c de G en 20 couleurs

```
1  Algorithme
2      c(x_1) = 1;
3      Marquer le sommet x_1^{k_1};
4      pour i = 1 à s faire
5          pour j = 1 à k_i faire
6              Marquer, s'il existe, le dernier voisin de x_j^i dans L_{i+1};
7          pour k = 0 à 3 faire
8              c_k ← k + 5 × (i mod 4);
9          c_∅ ← 4 + 5 × (i mod 4);
10         pour j = 1 à k_i faire
11             c(v_j^i) = c_{j mod 4};
12             si v_j^i est marqué alors
13                 tmp ← c_{(j+1) mod 4};
14                 c_{(j+1) mod 4} ← c_∅;
15                 c_∅ ← c_{(j−1) mod 4};
16                 c_{(j−1) mod 4} ← tmp;
17 retourner c
```

Algorithme 1: Coloration des graphes planaires extérieurs

localement identifiante. Soit uv une arête de G telle que $N[u] \neq N[v]$.

Supposons tout d'abord que uv n'est pas une arête incluse dans un niveau L_i, nous pouvons alors supposer que $u \in L_i$ et $v \in L_{i+1}$. Si $u \neq x_1$, u a un voisin t dans L_{i-1} et $c(t) \notin c(N[v])$. Supposons donc que $u = x_1$. Si le sommet v est de degré 1, alors u est de degré 2 et a un autre voisin, t, et $c(t) \notin c(N[v])$. Sinon, v a un voisin $t \neq u$. Si $t \in L_1$, comme $N[u] \neq N[v]$, il y a un autre voisin t' de v, dans L_2. Nous pouvons donc supposer que $t \in L_2$ et alors $c(t) \notin c(N[u])$. Dans tous les cas $c(N[u]) \neq c(N[v])$.

Supposons maintenant que uv est incluse dans un niveau L_i. Sans perte de généralité, $u = x_i^j$ et $v = x_i^{j+1}$, pour un certain j, et nous pouvons supposer qu'il existe un sommet t adjacent à exactement un sommet parmi $\{u, v\}$. Si $t \in L_i$, alors $c(N[u]) \neq c(N[v])$ car quatre sommets consécutifs ont des couleurs différentes dans L_i. Supposons maintenant que $t \in L_{i-1}$ et, sans perte de généralités, que t est un voisin de u. Le sommet v a au plus deux voisins dans L_{i-1}. Ces voisins (s'ils existent) suivent obligatoirement t dans le niveau L_{i-1} et donc ne peuvent pas avoir la même couleur que t. Sinon, $t \in L_{i+1}$. Les sommets u et v sont des sommets consécutifs de L_i ayant des voisinages différents dans L_{i+1}, les ensembles de couleurs dans leur voisinage sont donc différents (d'après la propriété énoncée précédemment). □

Nous pensons que cette borne n'est pas serrée. En effet, tous les graphes planaires extérieurs que nous avons étudiés sont 6-lid-coloriables.

Question 5.49 *Tout graphe planaire extérieur est-il* 6*-lid-coloriable ?*

Nous savons que ce résultat est vrai pour des graphes planaires extérieurs maximaux. Dans la partie suivante, nous étudions une classe de graphes planaires qui est en quelque sorte à l'opposé de ces graphes en terme de densité : les graphes planaires de grande maille.

5.6.2 Graphes planaires de grande maille

Nous montrons dans cette partie qu'un graphe planaire de maille suffisamment grande est 5-lid-coloriable. Pour cela nous utilisons des techniques de déchargement. Lorsqu'un graphe planaire a une maille suffisamment grande, $5k+1$ alors il contient soit un sommet de degré 1, soit un chemin de k sommets consécutifs de degré 2. Nous colorions alors le graphe de manière inductive en supprimant les sommets de degré 1 et les longs chemins.

Théorème 5.50 *Tout graphe planaire de maille au moins 36 est* 5*-lid-coloriable.*

Preuve : Pour cette preuve, nous avons besoin d'une contrainte en plus sur les colorations. Nous dirons (dans cette preuve) qu'une coloration localement identifiante est *régulière* si tout sommet de degré au moins 2 satisfait $|c(N[u])| = 3$. Nous allons montrer que tout graphe planaire de maille au moins 36 admet une lid-coloration régulière avec cinq couleurs.

Tout d'abord, les cycles de taille au moins 12 ont une lid-coloration régulière avec cinq couleurs en utilisant le motif $(1234)^*[12345][12345][12345]$ (voir la partie 5.3.3 pour ces notations).

Supposons par l'absurde qu'il existe un graphe G planaire de maille au moins 36 qui n'admet pas de 5-lid-coloration régulière. Nous choisissons G avec le plus petit nombre de sommets possible.

Montrons tout d'abord que G est connexe et n'a pas de sommets de degré 1. Si G n'est pas connexe, chacune de ses composantes est plus petite et 5-lid-coloriable de manière régulière, G l'est donc aussi. Supposons maintenant que G contienne un sommet u de degré 1, adjacent à un sommet v. Par minimalité de G, $G' = G - u$ admet une 5-lid-coloration régulière c. Nous considérons trois cas selon le degré de v dans G', et dans chacun des cas, nous étendons la coloration c a une 5-lid-coloration régulière de G, ce qui nous donne une contradiction.

Si v est de degré au moins 2 dans G', alors nous donnons à u une couleur de $c(N[v]) \setminus \{c(v)\}$. Alors $c(N[v])$ ne change pas et $c(N[u]) \neq c(N[v])$ car $|c(N[u])| = 2$ et $|c(N[u])| = 3$. Nous avons donc une lid-coloration régulière de G. Si v est de degré 1 dans G', alors v est adjacent à un sommet w de G' et nous donnons à u une couleur qui n'est pas dans $c(N[w])$. Une telle couleur existe car $|c(N[w])| \leq 3$. La coloration obtenue est régulière : $|c(N[v])| = 3$ et $c(N[v]) \neq c(N[w])$. Enfin, v ne peut pas être de degré 0 dans G' car G est connexe et K_2 a une 2-lid-coloration régulière.

Le graphe G est donc de degré minimum 2 et de maille au moins 36. G possède donc un chemin de sept sommets consécutifs de degré 2. Il existe donc un chemin $P = x_1x_2...x_9$ tel que x_1 soit de degré au moins 3 (car G n'est pas un cycle) ; x_2,...,x_8 soient de degré 2 et x_9 de degré au moins 2. Par minimalité de G, le graphe $G' = G \setminus \{x_2, ..., x_8\}$ admet une 5-lid-coloration régulière. Sans perte de généralité, supposons que $c(x_1) = 1$ et que $c(N[x_1]) = \{1, 2, 3\}$ (x_1 est de degré au moins 2 dans G'). Soit a la couleur de x_9. Si le degré de x_9 dans G' est au moins 2, soit $\{b, c\} = c(N(x_9))$. Si le degré de x_9 dans G' est 1, alors x_9 est adjacent à un sommet x_{10} de degré au moins 2 et nous posons $b = c(x_{10})$ et $c \in \{1, 2, 3, 4, 5\} \setminus c(N[x_{10}])$. Le tableau suivant donne les valeurs de $x_2, x_3, ...x_8$ en fonction des valeurs possibles de $(a; b, c)$. Notons que $c(x_2) \in \{1, 2, 3\}$, $c(x_3) \notin \{1, 2, 3\}$, $c(x_6) \neq a$, $c(x_7) \notin \{a, b, c\}$, $c(x_8) = c$ et quatre sommets consécutifs ont des couleurs distinctes. La coloration obtenue est donc une lid-coloration régulière de G, ce qui est contradictoire.

$2431243(1; 2, 3)$	$2431543(2; 1, 3)$	$2431542(3; 1, 2)$	$2534152(4; 1, 2)$	$2435142(5; 1, 2)$
$2431254(1; 2, 4)$	$2541354(2; 1, 4)$	$2431254(3; 1, 4)$	$2431253(4; 1, 3)$	$2431243(5; 1, 3)$
$2431245(1; 2, 5)$	$2451345(2; 1, 5)$	$2431245(3; 1, 5)$	$2451235(4; 1, 5)$	$2435124(5; 1, 4)$
$2431254(1; 3, 4)$	$3512354(2; 3, 4)$	$2431254(3; 2, 4)$	$2431253(4; 2, 3)$	$2431243(5; 2, 3)$
$2431245(1; 3, 5)$	$3412345(2; 3, 5)$	$2431245(3; 2, 5)$	$2451235(4; 2, 5)$	$2435214(5; 2, 4)$
$2531425(1; 4, 5)$	$3521435(2; 4, 5)$	$2531425(3; 4, 5)$	$2534125(4; 3, 5)$	$2435124(5; 3, 4)$

□

Si nous nous restreignons comme dans la preuve précédente aux colorations régulières, alors cette borne ne peut être améliorée, même pour des graphes de très grande maille. En effet, pour un cycle $\mathcal{C}_n$ avec $n \not\equiv 0 \bmod 4$, nous avons besoin d'au minimum cinq couleurs pour assurer que tous les sommets du cycle aient trois couleurs différentes dans leur voisinage.

A l'inverse, nous pouvons obtenir des résultats similaires en utilisant plus de couleurs pour des graphes planaires de maille plus petite.

Dans [35], nous conjecturions que les graphes planaires peuvent se colorier de manière localement identifiante avec un nombre borné de couleurs. Il semblerait que cette conjecture soit vraie d'après des remarques de Daniel Gonçalves et Alexandre Pinlou qui sont en cours de rédaction. Cependant la borne trouvée semble très grande, alors que le graphe planaire connu ayant besoin du plus grand nombre de couleurs est P_8^3 qui a besoin de seulement huit couleurs.

Question 5.51 *Tout graphe planaire est-il 8-lid-coloriable ?*

5.7 Conclusion sur les colorations identifiantes

Nous avons étudié dans ce chapitre un procédé d'identification utilisant des couleurs. Nous avons tout d'abord considéré une identification globale des sommets. Le nombre minimum de couleurs utilisées pour identifier un graphe est au plus la taille d'un code identifiant optimal, plus un. Ces colorations dans les cycles correspondent

à un problème de théorie des nombres, les cycles universels. Cela nous a permis d'exhiber des familles infinies de graphes G pour lesquelles il est possible d'identifier les sommets de G avec $\gamma^{\text{ID}}(G)^{\frac{1}{k}}$ couleurs. Une poursuite envisagée pour ce travail est d'étudier les colorations identifiants sur d'autres structures régulières comme les grilles ou les hypercubes.

Nous avons ensuite considéré une identification locale des sommets avec des couleurs. Les colorations localement identifiantes ont été comparées avec les colorations propres, et des résultats similaires ont été montrés. Par exemple, tout graphe possède une coloration localement identifiante avec $O(\Delta^2)$ couleurs, où Δ est le degré maximum du graphe. Ces colorations ont ensuite été étudiées dans des sous-classes de graphes parfaits, dans chaque cas, deux fois le nombre chromatique suffisent comme couleurs pour identifier localement le graphe. Cela n'est pas vrai pour des graphes parfaits quelconque. Nous conjecturons par contre que cela est vrai pour les graphes triangulés. Enfin, nous avons étudié ces colorations dans les graphes planaires, montrant des bornes supérieures sur le nombre de couleurs nécessaire dans le cas des graphes planaires extérieurs ou de grande maille. Une borne existe pour les graphes planaires quelconques, mais qui semble pour l'instant loin d'être serrée. Par ailleurs, les colorations localement identifiantes peuvent être vues comme un renforcement des colorations propres : dans une coloration propre, pour toute paire de voisins u et v, $\mathcal{B}_0(u) \neq \mathcal{B}_0(v)$. Dans les colorations localement identifiantes, nous demandons en plus que $\mathcal{B}_1(u) \neq \mathcal{B}_1(v)$. Une manière d'étendre les colorations localement identifiantes serait de fixer un rayon r et de demander que pour tout entier $i \leq r$, $\mathcal{B}_i(u) \neq \mathcal{B}_i(v)$.

Conclusion

Dans cette thèse, nous avons considéré plusieurs aspects de l'identification des sommets d'un graphe. Les graphes dont la taille minimale d'un code identifiant est la plus grande ont été caractérisés, et ce pour les cas finis, infinis, orientés ou non orientés (chapitre 2). Nous avons ensuite étudié les codes identifiants dans plusieurs classes de graphes, d'une part pour des graphes dont la structure est dépendante d'une métrique, graphes par conséquent souvent étudiés en théorie des codes, et d'autre part pour des familles héréditaires de graphes, définies comme graphes d'intersections d'objet. Ces dernières familles sont très étudiées en théorie des graphes. Suivant les cas, nous avons donné des valeurs exactes de codes optimaux, amélioré des bornes existantes ou montré des résultats de complexité (chapitre 3). Nous avons ensuite introduit plus de flexibilité dans les codes identifiants en introduisant des codes identifiants évolutifs et en construisant des codes identifiants tolérants dans le plan euclidien (chapitre 4). Enfin, nous avons étudié une identification des sommets avec des couleurs. Cette notion dans sa variante locale a été comparée avec la coloration propre des graphes, notamment dans des classes de graphes parfaits (chapitre 5).

Certaines questions abordées dans cette thèse ne sont pas complètement résolues ou amènent d'autres questions qui sont autant de pistes de recherche comme poursuite de cette thèse. Nous recensons ici celles qui d'après nous sont les plus intéressantes.

- Étudier les codes identifiants dans les graphes de Sierpiński généralisés (partie 3.1).
- Étudier les codes identifiants dans des classes de graphes héréditaires, en généralisant par exemple la classe de graphes comprenant les graphes adjoints et les graphes d'intervalles pour lesquelles tout code identifiant est de taille au moins $\Theta(\sqrt{n})$, et affiner les résultats de complexité et d'approximabilité sur ces classes.
- Étudier la notion de code $(p, \mathcal{R})$-identifiant, introduite dans la partie 4.1, en commençant par des structures de type cycles, hypercubes, grilles.
- Étudier le problème des colorations identifiantes sur d'autres graphes de classes correspondant à des métriques comme les grilles ou les hypercubes.
- Améliorer les bornes sur les colorations localement identifiantes pour des classes de graphes héréditaires comme les graphes triangulés (en étudiant la conjecture 5.47) ou sur les graphes planaires (question 5.51).

À titre plus personnel, j'aimerais étudier des problèmes d'identification sur des structures discrètes plus générales, comme par exemple les hypergraphes. Le problème

des colorations identifiantes, en particulier ses liens avec la combinatoire des mots et les colorations propres des graphes, sont un autre thème que je souhaiterais approfondir. Enfin, les notions étudiées dans cette thèse sont généralement abordables pour des non-initiés. J'aimerais les utiliser pour construire des situations de recherche ludiques pour les élèves ou le grand public.

Bibliographie

[1] D. Auger : Minimal identifying codes in trees and planar graphs with large girth. *European Journal of Combinatorics*, 31:1372–1384, 2010.

[2] D. Auger, I. Charon, O. Hudry et A. Lobstein : Complexity results for identifying codes in planar graphs. *International Transactions in Operational Research*, 17(6):691–710, 2010.

[3] D. Auger, I. Charon, O. Hudry et A. Lobstein : Watching systems in graphs : an extension of identifying codes. *À paraître dans Discrete Applied Mathematics*, 2012.

[4] L. Beaudou, S. Gravier, S. Klavžar, M. Kovše et M. Mollard : Covering codes in Sierpiński graphs. *Discrete Mathematics & Theoretical Computer Science*, 12(3):63–74, 2010.

[5] L. W. Beineke : Characterizations of derived graphs. *Journal of Combinatorial Theory*, 9(2):129 – 135, 1970.

[6] Y. Ben-Haim et S. Litsyn : Exact minimum density of codes identifying vertices in the square grid. *SIAM Journal on Discrete Mathematics*, 19(1):69–82, 2005.

[7] C. Berge : *Graphes et hypergraphes*. Dunod, 1970.

[8] A. A. Bertossi : Total domination in interval graphs. *Information Processing Letters*, 23(3):131 – 134, 1986.

[9] N. Bertrand : Codes identifiants et codes localisateurs-dominateurs sur certains graphes. Mémoire de D.E.A., ENST Paris, 2001.

[10] N. Bertrand, I. Charon, O. Hudry et A. Lobstein : Identifying and locating-dominating codes on chains and cycles. *European Jouranl of Combinatorics*, 25(7):969–987, 2004.

[11] B. Bollobás : *Combinatorics : Set Systems, Hypergraphs, Families of Vectors, and Combinatorial Probability*. Cambridge University Press, 1986.

[12] J. A. Bondy : Induced subsets. *Journal of Combinatoral Theory Serie B*, 12(2):201–202, 1972.

[13] K. S. Booth et J. H. Johnson : Dominating sets in chordal graphs. *SIAM Journal on Computing*, 11(1):191–199, 1982.

[14] M. Burstein : An upper bound for the chromatic number of hypergraphs. *Sakharth. SSR Mecn. Akad. Moambe*, 1974.

[15] E. Charbit, I. Charon, G. D. Cohen et O. Hudry : Discriminating codes in bipartite graphs. *Electronic Notes in Discrete Mathematics*, 26:29–35, 2006.

[16] E. Charbit, I. Charon, G. D. Cohen, O. Hudry et A. Lobstein : Discriminating codes in bipartite graphs : bounds, extremal cardinalities, complexity. *Advances in Mathematics of Communications*, 4:403–420, 2008.

[17] I. Charon, G. D. Cohen, O. Hudry et A. Lobstein : New identifying codes in the binary hamming space. *European Journal of Combinatorics*, 31(2):491–501, 2010.

[18] I. Charon, S. Gravier, O. Hudry, A. Lobstein, M. Mollard et J. Moncel : A linear algorithm for minimum 1-identifying codes in oriented trees. *Discrete Applied Mathematics*, 154(8):1246–1253, 2006.

[19] I. Charon, I. Honkala, O. Hudry et A. Lobstein : The minimum density of an identifying code in the king lattice. *Discrete Mathematics*, 276(1-3):95–109, 2004.

[20] I. Charon, O. Hudry et A. Lobstein : Identifying codes with small radius in some infinite regular graphs. *Electronic Journal of Combinatorics*, 8, 2002.

[21] I. Charon, O. Hudry et A. Lobstein : Minimizing the size of an identifying or locating-dominating code in a graph is NP-hard. *Theoretical Computer Science*, 290(3):2109–2120, 2003.

[22] I. Charon, O. Hudry et A. Lobstein : Extremal cardinalities for identifying and locating-dominating codes in graphs. *Discrete Mathematics*, 307(3-5):356–366, 2007.

[23] I. Charon, O. Hudry et A. Lobstein : Minimum sizes of identifying codes in graphs differing by one edge or one vertex. *Soumis*, 2012+.

[24] G. Chartrand, L. Eroh, M. Johnson et O. Oellermann : Resolvability in graphs and the metric dimension of a graph. *Discrete Applied Mathematics*, 105(1–3):99–113, 2000.

[25] C. Chen, C. Lu et Z. Miao : Identifying codes and locating-dominating sets on paths and cycles. *Discrete Appl. Math.*, 159(15):1540–1547, sept. 2011.

[26] F. Chung, P. Diaconis et R. Graham : Universal cycles for combinatorial structures. *Discrete Mathematics*, 110(1-3):43–59, 1992.

[27] G. Cohen, I. Honkala, A. Lobstein et G. Zémor : On codes identifying vertices in the two-dimensional square lattice with diagonals. *IEEE Transactions on Computers*, 50:174–176, 2001.

[28] G. Cohen, I. Honkala, M. Mollard, S. Gravier, A. Lobstein, C. Payan et G. Zémor : Improved identifying codes for the grid. *Electronic Journal of Combinatorics*, 6, 1999.

[29] B. Courcelle : The monadic second-order logic of graphs. i. recognizable sets of finite graphs. *Information and Computation*, 85(1):12–75, mars 1990.

[30] P. Cull et I. Nelson : Error-correcting codes on the towers of hanoi graphs. *Discrete Mathematics*, 208-209:157–175, 1999.

[31] D. Curtis, T. Hines, G. Hurlbert et T. Moyer : Near-universal cycles for subsets exist. *SIAM Journal on Discrete Mathematics*, 23(3):1441–1449, 2009.

[32] E. Dahlhaus, D. S. Johnson, C. H. Papadimitriou, P. D. Seymour et M. Yannakakis : The complexity of multiterminal cuts. *SIAM Journal on Computing*, 23:864–894, 1994.

[33] O. Delmas, S. Gravier, M. Montassier et A. Parreau : On two variations of identifying codes. *Discrete Mathematics*, 311(17):1948–1956, 2011.

[34] R. Diestel : *Graph Theory*. Graduate Texts in Mathematics, volume 173. Springer-Verlag, Heidelberg, troisième édn, 2005.

[35] L. Esperet, S. Gravier, M. Montassier, P. Ochem et A. Parreau : Locally identifying coloring of graphs. *Electronic Journal of Combinatorics*, 19, 2012.

[36] M. Farber : Domination, independent domination, and duality in strongly chordal graphs. *Discrete Applied Mathematics*, 7(2):115–130, 1984.

[37] F. Foucaud, S. Gravier, R. Nasersar, A. Parreau et P. Valicov : Identifying codes in line graphs. *Accepté dans Journal of Graph Theory*, 2012.

[38] F. Foucaud, E. Guerrini, M. Kovše, R. Naserasr, A. Parreau et P. Valicov : Extremal graphs for the identifying code problem. *European Journal of Combinatorics*, 32(4):628–638, 2011.

[39] F. Foucaud, I. Honkala, T. Laihonen, A. Parreau et G. Perarnau : Locally identifying colouring of graphs with given maximum degree. *Discrete Mathematics*, 312(10), 2012.

[40] F. Foucaud, R. Klasing, A. Kosowski et A. Raspaud : Bounds on the size of identifying codes in triangle-free graphs. *Discrete Applied Mathematics*, 160(10-11):1532–1546, 2012.

[41] F. Foucaud, T. Laihonen et A. Parreau : An improved lower bound for $(1, \leq 2)$-identifying codes in the king grid. *Soumis*, 2011+.

[42] F. Foucaud, R. Naserasr et A. Parreau : Extremal problems on identifying codes in digraphs and bondy's theorem on induced subsets. *À paraître dans Graphs and Combinatorics*, 2012.

[43] M. R. Garey et D. S. Johnson : *Computers and Intractability : A Guide to the Theory of NP-Completeness*. W. H. Freeman, 1979.

[44] J. Gimbel, B. Van Gorden, M. Nocolescu, C. Umstead et N. Vaiana : Location with dominating sets. *Congressus Numerantium*, 151:129–144, 2001.

[45] S. Gravier, R. Klasing et J. Moncel : Hardness results and approximation algorithms for identifying codes and locating-dominating codes in graphs. *Algorithmic Operations Research*, 3(1), 2008.

[46] S. Gravier, S. Klavžar et M. Mollard : Codes and $l(2,1)$-labelings in Sierpiński graphs. *Taiwanese Journal of Mathematics*, 4:671–681, 2004.

[47] S. Gravier, M. Kovse, M. Mollard, J. Moncel et A. Parreau : New results on variants of covering codes in Sierpiński graphs. *À paraître dans Designs, Codes and Cryptography*, 2012.

[48] S. Gravier et J. Moncel : On graphs having a $V \setminus \{x\}$ set as an identifying code. *Discrete Mathematics*, 307(3-5):432–434, 2007.

[49] S. Gravier, J. Moncel et A. Semri : Identifying codes of cycles. *European Journal of Combinatorics*, 27(5):767–776, 2006.

[50] F. Harary et R. A. Melter : On the metric dimension of a graph. *Ars Combinatoria*, 2:191–195, 1976.

[51] F. Harary et M. Plantholt : The point-distinguishing chromatic index., 1985.

[52] T. W. Haynes, S. T. Hedetniemi et P. J. Slater : *Fundamentals of domination in graphs*. Marcel Dekker, 1998.

[53] I. Honkala et T. Laihonen : On a new class of identifying codes in graphs. *Information Processing Letters*, 102(2–3):92–98, 2007.

[54] I. Honkala, T. Laihonen et S. Ranto : On codes identifying sets of vertices in Hamming spaces. *Designs, Codes and Cryptography*, 24:193–204, 2001. 10.1023/A :1011256721935.

[55] I. S. Honkala et T. Laihonen : Codes for identification in the king lattice. *Graphs and Combinatorics*, 19(4):505–516, 2003.

[56] D. Hughes et F. Piper : *Projective planes*. Graduate texts in mathematics. Springer-Verlag, 1973.

[57] G. Hurlbert : On universal cycles for k-subsets of an n-set. *SIAM Journal on Discrete Mathematics*, 7:598–604, 1994.

[58] B. W. Jackson : Universal cycles of 4-subsets and 5-subsets. *Unpublished*.

[59] B. W. Jackson : Universal cycles of k-subsets and k-permutations. *Discrete Mathematics*, 117:141–150, 1993.

[60] V. Junnila : *On identifying and locating-dominating codes*. Thèse de doctorat, Turku Centre for computer science, 2011.

[61] V. Junnila et T. Laihonen : Identification in $\mathbb{Z}^2$ using euclidean balls. *Discrete Applied Mathematics*, 159(5):335–343, 2011.

[62] V. Junnila et T. Laihonen : Optimal identifying codes in cycles and paths. *À paraître dans Graphs and Combinatorics*, p. 1–13, 2012.

[63] V. Junnila, T. Laihonen et A. Parreau : Tolerant identification with euclidean balls. *À paraître dans Networks*, 2012.

[64] M. G. Karpovsky, K. Chakrabarty et L. B. Levitin : On a new class of codes for identifying vertices in graphs. *IEEE Transactions on Information Theory*, 44(2):599–611, 1998.

[65] J. M. Keil : Total domination in interval graphs. *Information Processing Letters*, 22(4):171–174, 1986.

[66] S. Klavžar et U. Milutinović : Graphs $s(n, k)$ and a variant of the tower of Hanoi problem. *Czechoslovak Mathematical Journal*, 47(122):95–104, 1997.

[67] S. Klavžar, U. Milutinović et C. Petr : 1-perfect codes in sierpiński graphs. *Bulletin of the Australian Mathematical Society*, 66:369–384, 2002.

[68] A. Kogan : On the essential test sets of discrete matrices. *Discrete Applied Mathematics*, 60(13):249–255, 1995.

[69] M. Laifenfeld, A. Trachtenberg et T. Y. Berger-wolf : Identifying codes and the set cover problem. *Proceedings of the 44th annual Allerton Conference on Communication, Control, and Computing*, 2006.

[70] C.-K. Li et I. Nelson : Perfect codes on the towers of hanoi graphs. *Bulletin of the Australian Mathematical Society.*, 57:367–376, 1998.

[71] S. L. Lipscomb et J. C. Perry : Lipscomb's $l(a)$ space fractalized in hilbert's $l^2(a)$ space. *Proceedings of the American Mathematical Society*, 115:1157–1165, 1992.

[72] A. Lobstein : Codes identifiants et codes localisateurs-dominants, une bibliographie. Publiée en ligne : www.infres.enst.fr/∼lobstein/bibLOCDOMetID.html.

[73] L. Lovász : Coverings and colorings of hypergraphs. *Proceedings of the fourth south-eastern conference on combinatorics, graph theory, and computing, Boca Raton, Florida*, 1973.

[74] L. Lovász : *Combinatorial problems and exercises.* North-Holland Pub. Co., 1979.

[75] U. Milutinović : Completeness of the lipscomb space. *Glasnik Matematički Serija III*, 27(47):343–364, 1992.

[76] J. Moncel : *Codes identifiants dans les graphes.* Thèse de doctorat, Université Joseph Fourier, Grenoble, 2005.

[77] J. Moncel : On graphs on n vertices having an identifying code of cardinality $\log_2(n+1)$. *Discrete Applied Mathematics*, 154(14):2032–2039, 2006.

[78] B. Moret : Planar NAE-3-SAT is in $\mathcal{P}$. *ACM SIGACT News*, 19:51–54, 1988.

[79] R. Motwani et M. Sudan : Computing roots of graphs is hard. *Discrete Applied Mathematics*, 54(1):81–88, 1994.

[80] R. W. Payne et D. A. Preece : Identification keys and diagnostic tables : A review. *Journal of the Royal Statistical Society. Series A (General)*, 143(3):253–292, 1980.

[81] M. Pelto : New bounds for $(r, \leq 2)$-identifying codes in the infinite king grid. *Cryptography and Communications*, 2:41–47, 2010.

[82] J. Penaud : Une propriété de bicoloration des hypergraphes planaires. *Cahiers du Centre d'Études de Recherche Opérationnelle*, 17:345–349, 1975.

[83] S. Ray, D. Starobinski, A. Trachtenberg et R. Ungrangsi : Robust location detection with sensor networks. *IEEE Journal on Selected Areas in Communications*, 22(6):1016–1025, 2004.

[84] S. Ray, R. Ungrangsi, F. D. Pellegrini, A. Trachtenberg et D. Starobinski : Robust location detection in emergency sensor networks. *INFOCOM*, 2003.

[85] D. L. Roberts et F. S. Roberts : Locating sensors in paths and cycles : The case of 2-identifying codes. *European Journal of Combinatorics*, 29(1):72–82, 2008.

[86] R. D. Skaggs. : *Identifying vertices in graphs and digraphs.* Thèse de doctorat, University of South Africa, South Africa, 2007.

[87] P. J. Slater : Leaves of tree. *Congressus Numerantium*, 14:549–559, 1975.

[88] P. J. Slater : Dominating and reference sets in a graph. *Journal of Mathematical and Physical Sciences*, 22(4):445–455, 1988.

[89] P. J. Slater et D. F. Rall : On location-domination numbers for certain classes of graphs. *Congressus Numerantium*, 45:97–106, 1984.

[90] B. Stevens, P. Buskell, P. Ecimovic, C. Ivanescu, A. M. Malik, A. Savu, T. S. Vassilev, H. Verrall, B. Yang et Z. Zhao : Solution of an outstanding conjecture : the non-existence of universal cycles with k=n-2. *Discrete Mathematics*, 258(1-3):193–204, 2002.

[91] J. Suomela : Approximability of identifying codes and locating-dominating codes. *Information Processing Letters*, 103(1):28 – 33, 2007.

[92] C. Thomassen : The even cycle problem for directed graphs. *Journal of American Mathematical Society*, 5:217–219, 1992.

[93] L. E. Trotter : Line perfect graphs. *Mathematical Programming*, 12:255–259, 1977. 10.1007/BF01593791.

[94] A. Winter : Another algebraic proof of bondy's theorem on induced subsets. *Journal of Combinatorial Theory, Series A*, 89(1):145–147, 2000.

[95] M. Xu, K. Thulasiraman et X. Hu : Identifying codes of cycles with odd orders. *European Journal of Combinatorics*, 29(7):1717–1720, 2008.

[96] Z. Zhang, X. Chen, J. Li, B. Yao, X. Lu et J. Wang : On adjacent-vertex-distinguishing total coloring of graphs. *Science in China Series A : Mathematics*, 48:289–299, 2005.

[97] Z. Zhang, L. Liu et J. Wang : Adjacent strong edge coloring of graphs. *Applied Mathematics Letters*, 15(5):623–626, 2002.

Résumé : Dans cette thèse, nous étudions des problèmes d'identification des sommets dans les graphes. *Identifier* les sommets d'un graphe consiste à attribuer à chaque sommet un objet qui rend le sommet unique par rapport aux autres. Nous nous intéressons particulièrement aux *codes identifiants* : sous-ensembles de sommets d'un graphe, dominants, tels que le voisinage fermé de chaque sommet du graphe a une intersection unique avec l'ensemble. Les sommets du code identifiant peuvent être considérés comme des capteurs et chaque sommet du graphe comme un lieu possible pour une défaillance.

Nous caractérisons tout d'abord l'ensemble des graphes pour lesquels tous les sommets sauf un sont nécessaires dans tout code identifiant. Le problème consistant à trouver un code identifiant optimal, c'est-à-dire de taille minimale, étant $\mathcal{NP}$-difficile, nous l'étudions sur quatre classes restreintes de graphes. Suivant les cas, nous pouvons résoudre complètement le problème (pour les graphes de Sierpiński), améliorer les bornes générales (pour les graphes d'intervalles, les graphes adjoints, la grille du roi) ou montrer que le problème reste difficile même restreint (pour les graphes adjoints).

Nous considérons ensuite des variations autour des codes identifiants permettant plus de flexibilité pour les capteurs. Nous étudions par exemple des capteurs du plan capables de détecter des défaillances à un rayon connu avec une erreur tolérée. Nous donnons des constructions de tels codes et bornons leur taille pour des valeurs de rayons et d'erreurs fixés ou asymptotiques. Nous introduisons enfin la notion de *coloration identifiante* d'un graphe, permettant d'identifier les sommets d'un graphe avec les couleurs présentes dans son voisinage. Nous comparons cette coloration avec la coloration propre des graphes et donnons des bornes sur le nombre de couleurs nécessaires pour identifier un graphe, pour plusieurs classes de graphes.

Mots clés : Codes identifiants, détection de défaillance, domination, coloration, théorie des graphes, combinatoire.

Abstract : In this thesis, we study problems on vertices identification of graphs. *To identify* the vertices of a graph consists in giving to each vertex of the graph an object that makes it unique. We are specially interested in the problem of identifying codes : dominating sets of vertices for which the closed neighborhood of each vertex has a unique intersection with the set. The vertices of the identifying code can be seen as *sensors* and each vertex of the graph as the location of a potential *fault*.

We first classify all finite graphs for which all but one of the vertices are needed in any identifying code. Finding an optimal identifying code, *i.e*, an identifying code of minimum size, is a $\mathcal{NP}$-hard problem. Therefore, we study this problem in some restricted classes of graphes. Depending on the class considered, we are able to solve this problem (for Sierpiński graphs), to give better bounds on the size of an identifying code than the general one (for interval graphs, line graphs and the king grid) or to prove that the problem remains $\mathcal{NP}$-hard even in the restricted class (for line graphs).

Then, we consider some variations of identifing codes that give flexibility to the sensors. For example, we study codes sensors able to detect faults within a radius around a fixed value. We give constructions of such codes and bounds on their size for general and asymptotic values of the radius and the tolerance on it. Finally, we introduce *identifying colourings* of graph ; verex-colouring of graph such that each vertex is identified by the set of colours in its closed neighbourhood. We compare this colouring of graphs with proper vertex-colouring and give bounds on the number of colours required to identify a graph, for several class of graphs.

Key words : Identifying codes, fault detection, domination, colouring, graph theory, combinatorics.