Majoration de la taille minimale d'une décomposition d'un graphe régulier

Guillaume Aubian

23 août 2015

Résumé

Voici le rapport de mon stage effectué à l'École Normale Supérieure de Lyon sous la direction de Stéphan Thomassé. L'objectif de ce stage était de majorer la taille d'une décomposition minimale d'un graphe k-régulier en chemins disjoints.

Table des matières

1 Préambule

1.1 Le LIP

Il s'agit du Laboratoire de l'Informatique du Parallélisme, situé au centre de Lyon. C'est le laboratoire d'informatique de l'École Normale Supérieure de Lyon, et la plupart des chercheurs y enseignent. J'étais dans l'équipe MC2 - pour Modèles de calcul, Complexité, Combinatoire - qui s'intéresse à la théorie de la complexité, à l'algorithmique sur les structures discrètes et à la combinatoire de manière générale.

Le chercheur qui m'a encadré était Stéphan Thomassé, dont le domaine de recherche est plutôt axé sur la théorie des graphes, même si j'ai aussi beaucoup travaillé et communiqué avec d'autres chercheurs ou doctorants que je remercie chaleureusement, parmi lesquels Matthieu Rosenfeld, Aurélie Lagoutte et Éric Thierry.

1.2 Le problème étudié

Le problème que m'a proposé Stéphan Thomassé concerne donc la théorie des graphes, et possède l'énorme avantage d'être très facile d'accès, ce qui m'a permis de m'y attaquer directement.

Tout d'abord, nous allons donner quelques précisions, notamment sur le vocabulaire qui sera utilisé.

Nous ne nous intéresserons qu'à des graphes simples non-orientés i.e. avec au plus une arête entre chaque paire de sommets et aucune arête d'un sommet vers lui-même. Un tel graphe est dit k-régulier si tous ses sommets sont de degré k, i.e. ont k voisins.

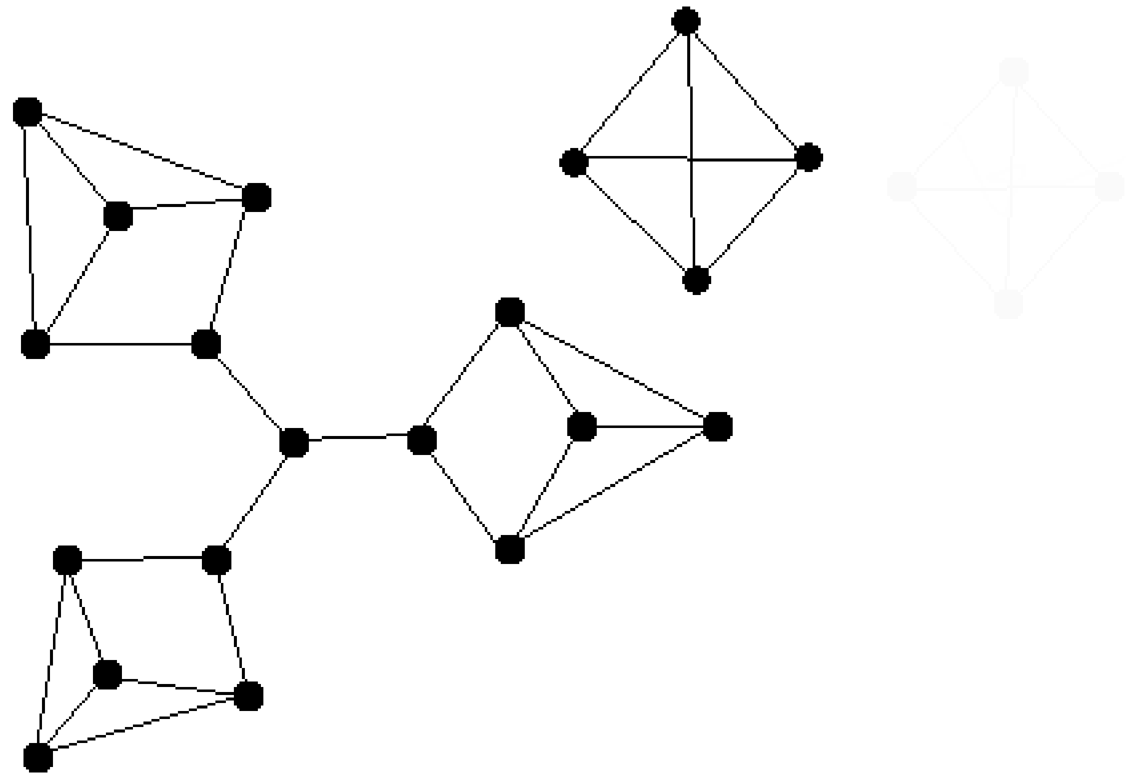

Fig 1: un graphe 3-régulier

Pour un tel graphe G, on nomme décomposition en chemins disjoints de G un ensemble C de chemins disjoints de G tel que l'ensemble des sommets des chemins de C recouvre l'ensemble des sommets de G. On note

$\mu(G){=}min\{|C|,\ C$ est une décomposition en chemins disjoints de $G\}$.

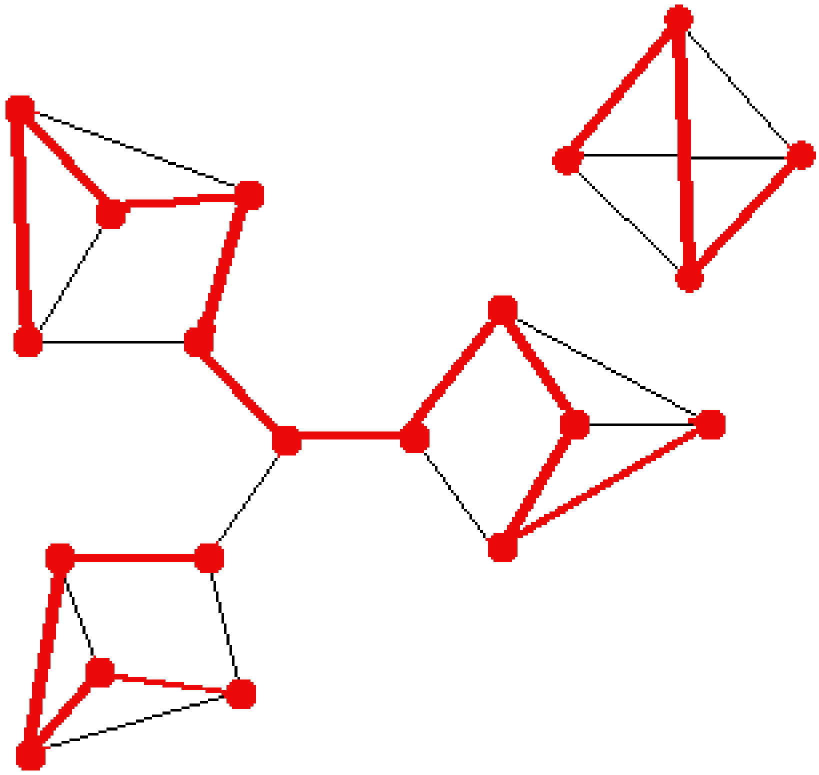

Fig. 2: Une décomposition minimale de Fig. 1

Enfin, un «tour» dans un graphe est un chemin qui parcourt tous les sommets du graphe au moins une fois, et revient à son point de départ. Trouver un «tour» de taille minimale est un problème NP-complet, puisque cela implique de trouver un cycle hamiltonien dans le cas des graphes qui en admettent un, et on se contente donc d'approximations, la meilleure étant en $(1 + O(1))n$ à l'heure actuelle.

La conjecture à démontrer est la suivante :

Conjecture 1.1. Si G est un graphe k-régulier à n sommets, $\mu(G) \leq \frac{n}{k+1}$

C'est Magnant et Martin qui l'ont établie, en 2009 dans [1].

En 2014, dans [2], R. Ravi a trouvé un algorithme qui calcule en temps polynomial un «tour» d'un graphe G avec une longueur de seulement $(1 + O(\frac{1}{k}))n$, dans le cas des graphes k-réguliers, à condition que notre conjecture soit vérifiée.

Quelques semaines avant le début de mon stage, Paul Seymour a fait remarquer en conférence que dans le cas de graphes denses - typiquement $k = \epsilon n$ - cela donnait une $1 + O(\frac{1}{\epsilon n})$-approximation, et que ça semblait donc être un angle d'attaque intéressant. Mon maitre de stage s'étant alors rendu compte que le problème était beaucoup plus difficile que prévu et que c'était le plus intéressant, nous avons décidé de nous attaquer en priorité aux graphes denses.

Notre plus grand résultat a été de le démontrer pour $k \geq \frac{n}{3}$, même si nous avons pu aborder plusieurs autres stratégies, qui n'ont malheureusement pas abouti, mais pourraient être approfondies.

Il faut aussi noter que l'on démontre quand même le résultat de Ravi si on montre seulement $\mu(G) = O(\frac{n}{k})$.

2 Les résultats obtenus

2.1 Les résultats directs

Tout d'abord, on peut énoncer un certain nombre de résultats directs.

Dans un premier temps, on peut remarquer que la conjecture est immédiate si G admet un chemin hamiltonien : en effet, on a alors $\mu(G) = 1$, ce qui vérifie forcément la conjecture vu que $k + 1 \leq n$.

Or d'après le théorème de Dirac :

Théorème 1. *Si le degré minimal d'un graphe G à n noeuds est supérieur ou égal à $\frac{n}{2}$, alors G est hamiltonien*

Donc, si $k \geq \frac{n}{2}$, la conjecture est vérifiée. À fortiori, la conjecture est vérifiée sur les graphes complets (vu que $k = n - 1 \geq \frac{n}{2}$).

De plus, il est important de voir qu'on peut se limiter aux graphes connexes. En effet, supposons qu'on puisse partager un graphe G en deux composantes G_0 et G_1 telles qu'il n'y ait aucune arête entre G_0 et G_1. On remarque que tout chemin de G est soit dans G_0, soit dans G_1. Donc $\mu(G) \leq \mu(G_0) + \mu(G_1)$. On a donc notre relation de récurrence, et il suffit de résoudre le cas initial, qui est celui des graphes connexes.

En s'inspirant de ceci, il est évident que la conjecture est «sharp», puisqu'il suffit de considérer l'union de $\frac{n}{k+1}$ graphes complets à k sommets : chaque composante connexe admet un chemin hamiltonien d'après le théorème de Dirac, et donc $\mu(G) = \frac{n}{k+1}$.

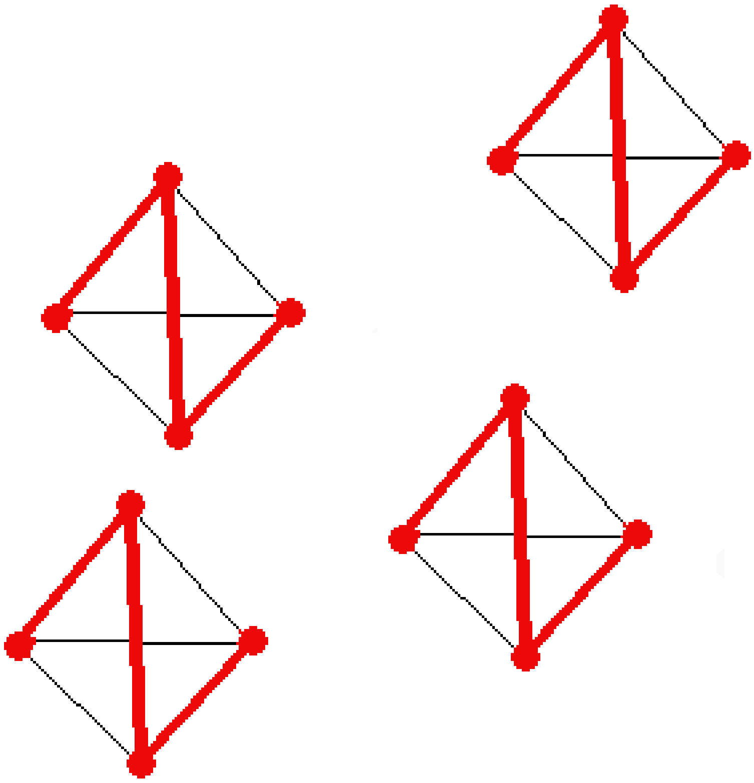

Fig. 3: Cas limite pour k=3 et n = 16

Que se passe-t'il pour des petites valeurs de k ? On se restreint comme décrit plus haut à des graphes connexes.

Si $k = 0$, alors le graphe est réduit à un sommet i.e $\mu(G) = 1 = \frac{1}{0+1} = \frac{n}{k+1}$

Si $k = 1$, le graphe n'est constitué que de deux sommets reliés par une arête, et donc possède un chemin hamiltonien i.e. $\mu(G) = \frac{2}{1+1} = \frac{n}{k+1}$.

Si $k = 2$, G est réduit à un cycle i.e. $\mu(G) = 1 \leq \frac{n}{3}$ car $n \geq k + 1 = 3$.

Il y a aussi des preuves pour k=3, 4 et 5 dans [1] mais on les redémontre facilement avec notre méthode du maximum lexicographique.

2.2 La méthode de l'arbre couvrant

Nous avons tout d'abord découvert deux théorèmes qui nous semblaient prometteurs pour la suite.

On notera $nb_{leaves}(T)$ le nombre de feuilles d'un arbre T. Le premier des deux théorèmes stipule :

Théorème 2. *Soit ST un arbre couvrant d'un graphe connexe G, alors $\mu(G) \leq nb_{leaves}(ST)$.*

Démonstration. Il est évident que $\mu(G) \leq \mu(ST)$, vu qu'un chemin de ST est un chemin de G.

On va alors noter $PW(T)$ l'ensemble des chemins pointés d'un arbre T que l'on définit par récurrence de la manière ainsi :

Si T est réduit à un sommet r (qui est aussi sa racine), et en notant c le chemin réduit à r, $PW(T) = \{c\}$. On remarque que la racine de T (qui n'est autre que r), est à l'extrémité d'un chemin de $PW(T)$.

Sinon, si T est réduit à une racine r avec s fils $T_1, T_2...T_s$. On choisit c le chemin de $PW(T_1)$ tel que la racine de T_1 est à l'extrémité de c. On définit alors $PW(T)$ comme étant l'union des $PW(T_i)$ (pour i allant de 1 à s), si ce n'est qu'on rajoute r à l'extrémité de c. On remarque alors qu'un des chemins de $PW(T)$ a pour extrémité la racine de T.

On démontre facilement par induction structurelle que $|PW(T)| = nb_{leaves}(T)$.

Et comme $PW(T)$ est un cas particulier de décomposition en chemins disjoints, $\mu(ST) \leq |PW(T)|$.

Donc $\mu(G) \leq \mu(ST) \leq |PW(T)| = nb_{leaves}(T)$ □

On aimerait alors pouvoir démontrer qu'un graphe k-régulier à n noeuds G admet toujours $O(\frac{n}{k})$ feuilles. Et le théorème 2 nous permet de montrer que si notre conjecture est vraie, c'est le cas :

Théorème 3. *Soit G un graphe connexe, alors :*
$\exists ST$ un arbre couvrant de G, $nb_{leaves}(ST) \leq 2\mu(G)$.

Démonstration. Soit G un graphe connexe. On note $W = \{w_1; w_2; ...; w_{\mu(G)}\}$ une décomposition minimale de G en chemins disjoints.

On note G_{folded} le graphe replié de G i.e. les sommets de G_{folded} sont en bijection avec les chemins de W et il y a une arête entre deux sommets de G_{folded} si et seulement si il y a une arête entre les deux chemins correspondants de W.

Par construction de G_{folded}, et G étant connexe, il est connexe lui-aussi. On considère alors ST_{folded}, un arbre couvrant de G_{folded}. En «dépliant» ST_{folded}, on obtient un arbre couvrant ST de G (ceci se voit aisément, mais se démontre en remarquant que ST_{folded} possède $\mu(G) - 1$ arêtes, et que les chemins de W utilisent $n - \mu(G)$ arêtes. Donc ST utilise $n - \mu(G) + \mu(G) - 1 = n - 1$ arêtes. Comme W partitionne les sommets de G, ST est un arbre couvrant de G). □

Donc :
$\forall G, \exists ST$, arbre couvrant de G, $\mu(G) \leq nb_{leaves}(ST) \leq 2\mu(G)$

Il est à noter que ceci ne se restreint pas qu'aux graphes réguliers.

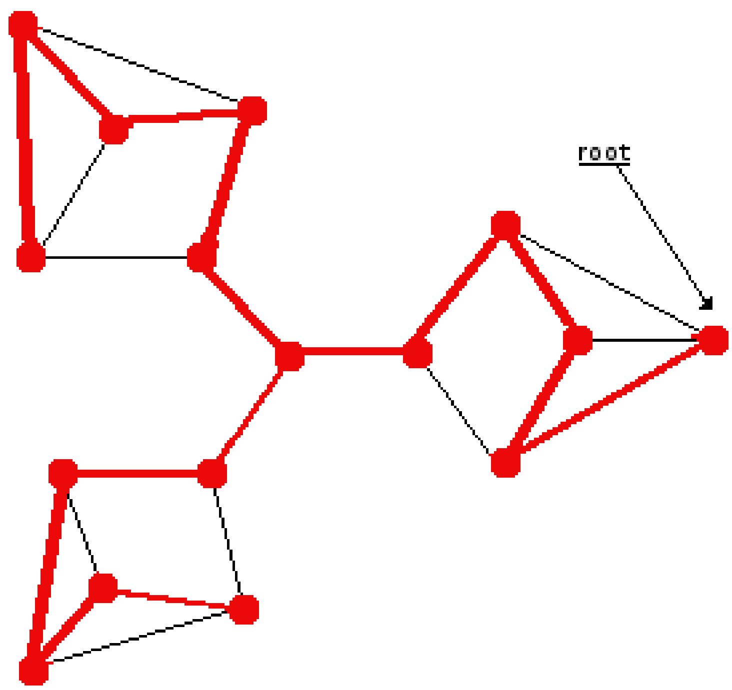

Fig. 4: Un arbre couvrant minimal d'une composante connexe de Fig. 1

On doit alors démontrer que les graphes k-réguliers à n noeuds admettent un arbre couvrant à $O(\frac{n}{k})$ feuilles. On peut partir d'un arbre couvrant avec un nombre minimal de feuilles, et montrer qu'il a forcément moins de $\frac{n}{k+1}$ feuilles, par exemple par l'absurde avec l'aide du théorème suivant.

Théorème 4. *Soit G un graphe connexe k-régulier, et ST_{min} un arbre couvrant qui minimise $nb_{leaves}(ST_{min})$. Si $nb_{leaves} > \frac{n}{k+1}$, alors deux feuilles de ST_{min} ont un voisin commun dans G*

Démonstration. Si, à chaque feuille, on associe ses k voisins et elle-même - i.e. $k+1$ éléments - alors comme il y a plus de $\frac{n}{k+1}$ feuilles et n sommets, on obtient presque notre résultat en utilisant le théorème des tiroirs. En effet, il reste encore le cas où deux feuilles sont voisines : ceci est impossible, car si c'était le cas, on pourrait rajouter l'arête entre les deux, et supprimer une arête adjacente à un noeud du cycle nouvellement créé de degré ≥ 3 dans ST_{min}. Ainsi, on aurait supprimé deux feuilles pour en créer une, ce qui contredit l'argument de minimalité. □

Dans la pratique, nous ne sommes pas arrivés à conclure malgré ce théorème. Néanmoins, on aimerait bien garder une structure qui permette de conclure avec un théorème semblable.

2.3 La méthode du maximum lexicographique

C'est le cas avec cet angle d'attaque : ici, on considère une partition minimale de G en chemins disjoints $W = \{w_1; w_2; ...; w_{\mu(G)}\}$ (on demande donc que $|W|$ soit minimal), qui soit un maximum lexicographique parmi les partitions minimales en chemins disjoints de G. Pour rappel, cela signifie que pour toute partition minimale de G en chemins disjoints $W' = \{w'_1; w'_2; ...; w'_{\mu(G)}\}$ différente de W, $w_1 > w'_1$ ou ($w_1 = w'_1$ et $w_2 > w'_2$) ou ($w_1 = w'_1$ et $w_2 = w'_2$ et $w_3 > w'_3$) etc...

i.e. $\exists i \in [1; ...; \mu(G)] \forall i' < i, (w_i > w'_i$ et $w_{i'} = w'_{i'})$

Théorème 5. *Soit x un sommet à l'extrémité d'un w_i, et y un voisin de x situé sur w_j, pour $i \neq j$. Alors, y est à une distance supérieure ou égale à $|w_i|$ d'une extrémité de w_j.*

Démonstration. Par l'absurde, on pourrait rajouter l'arête entre x et y, et supprimer l'arête partant de y vers l'extrémité de w_j la plus proche. On obtient alors une contradiction puisque le nouveau chemin obtenu est plus grand que w_i et que w_j. □

A fortiori, cela implique que les extrémités des w_i ont tous leurs voisins situés sur des $w_{i'}$ pour $i' \leq i$, comme on peut le voir sur la figure suivante :

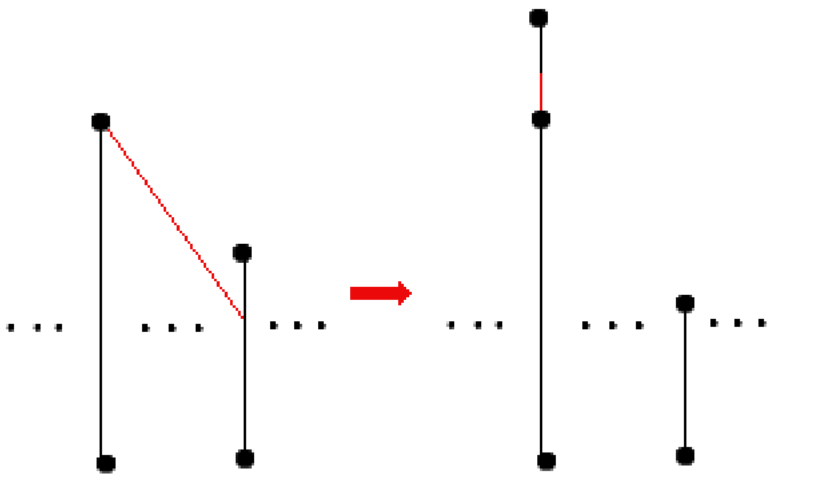

Fig. 5: ceci est absurde car le nouveau chemin est trop grand

Une autre conséquence est que si une extrémité x de w_i a un voisin sur w_j pour $i \neq j$, alors $|w_j| \geq 2|w_i|$. On voudrait alors faire une chaîne $w_{i_0} \rightarrow w_{i_1} \rightarrow w_{i_2} \rightarrow$ ... tel que chaque élément soit deux fois plus grand que le précédent, d'autant qu'on démontre facilement un théorème, semblable à celui sur les arbres, qui nous garantit que si il y a plus de $\frac{n}{k+1}$ chemins, l'un d'entre eux a ses extrémités qui ont chacune au moins un voisin à l'extérieur de ce chemin.

Même si on arrive à quelques résultats, par exemple on peut prouver que le plus grand chemin est au moins 8 fois plus long que le plus petit, le cas de figure suivant semble être une barrière infranchissable :

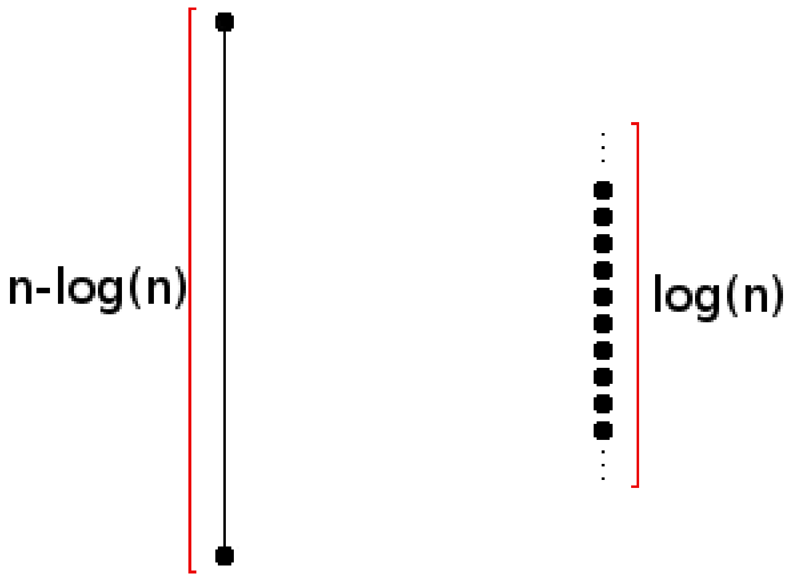

Fig. 6: le cas qui pose problème; un grand chemin et beaucoup de petits

Supposons qu'on ne puisse pas trouver deux extrémités voisines (par exemple si les k voisins de chaque sommet isolé sont vers le milieu du grand chemin), alors on semble bloqué avec $log(n)$ chemins, alors que si k est très grand (par exemple $k \geq \frac{n}{3}$), il faudrait trouver un nombre «constant» de chemins (2 ou 3 par exemple)

Dans le cas présent, une solution est de remarquer que, comme k est très grand, alors on pourra forcément trouver dans le long chemin deux sommets séparés par un seul sommet x et ayant un voisin commun parmi les sommets isolés : auquel cas, on peut échanger ce voisin commun avec x. Ainsi, cela nous permet d'augmenter le nombre de sommets isolés éventuels. En continuant ainsi tant que possible, il est probable que deux sommets isolés soient voisins, d'où l'absurdité.

Il faudrait que l'on puisse transformer les chemins afin d'augmenter le nombre éventuel d'extrémités, vu que l'on ne peut travailler qu'à partir des extrémités : par exemple, si les deux extrémités d'un segment sont voisines i.e. s'il forme en fait un cycle, alors on peut choisir comme extrémité de notre chemin le sommet de notre choix. Afin de faire ceci, il nous faut une structure adaptée.

2.4 Le Dominating Cycle

Soit G un graphe, on définit le dominating cycle C de ce graphe comme étant un cycle de longueur maximale, et comme étant parmi ceux-ci un qui minimise le nombre de composantes connexes de $G - C$.

Soit $C = c_0 c_1 ... c_{l-1} c_0$ un dominating cycle de G, supposons qu'au moins une composante connexe de $G - C$ soit réduite à un sommet v. Un théorème très visuel que nous avons démontré avec mon maître de stage, sans savoir qu'il était déjà mentionné dans [3] sous le nom de «Hopping Lemma», est le suivant :

Lemme 6. *On définit une suite Y_i de sous-ensembles de sommets de G de la manière suivante :*

$Y_0 = \{v\}$

$Y_i = \{v_j | v_{j-1} et v_{j+1}$ *voisins de* $x, x \in Y_{i-1}\}$

Si on pose $Y = \cup_{i=0}^{+\infty} Y_i \cup G - C$, alors il n'y a aucune arête entre éléments de Y, et ils n'ont pas de voisins successifs sur C.

Démonstration. La preuve complète est très visuelle, puisqu'elle consiste simplement à remarquer que Y_1 ne contient que des éléments que l'on peut échanger avec $Y_0 = \{v\}$, c'est à dire sortir du cycle. De même on peut échanger les éléments de Y_2 avec Y_1, et donc les sortir eux aussi, et ainsi de suite. Et comme évidemment, il n'y a pas d'arêtes entre élèments de $G - C$, et que les éléments de $G - C$ n'ont pas deux voisins consécutifs sur C, on conclut. □

J'ai dans un premier temps essayé d'utiliser des versions affaiblies du Hopping Lemma pour chercher des «patterns interdits». Par exemple, on ne peut pas avoir ceci :

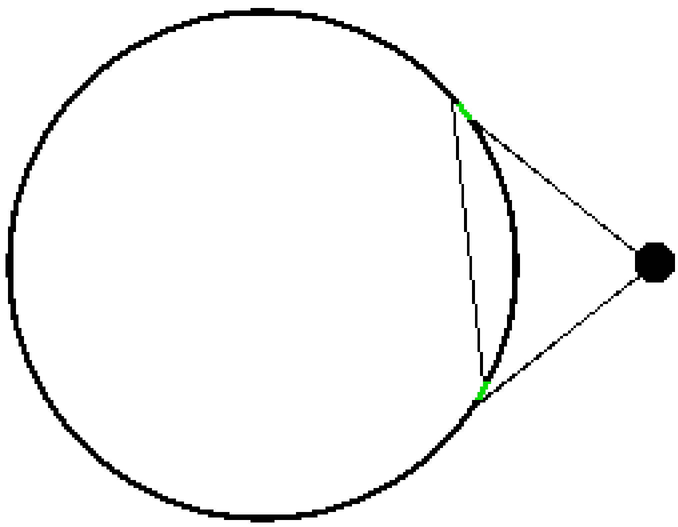

Fig. 7: Un pattern interdit

où le grand cercle représente le dominating cycle initial. En effet il suffirait alors de supprimer les arêtes vertes pour obtenir un cycle strictement plus grand, ce qui est absurde.

Ceci nous a ainsi permis de démontrer que :
$\forall \epsilon > 0, \exists N, \forall n > N, \forall k \geq \frac{n}{2} - \epsilon$ la conjecture est démontrée pour k et n

Néanmoins, le hopping lemma permet surtout de démontrer notre résultat principal :

Théorème 7. *Soit G un graphe k-régulier à n sommets, avec $k \geq \frac{n}{3}$, alors $\mu(G) \leq \frac{n}{k+1}$*

Démonstration. Comme expliqué plus haut, on peut supposer G connexe, quitte à raisonner par récurrence sur le nombre de composantes connexes.

Pour rappel, on dit qu'un graphe est c-connexe, si quels que soient les c sommets que l'on choisit, leur suppression laisse G connexe

Il y a alors deux cas de figure possibles, selon que G est 2-connexe ou pas.

Si G est 2-connexe, Bill Jackson démontre dans [4] que G est hamiltonien (en utilisant le Hopping Lemma).

Sinon, si G n'est pas 2-connexe, il existe un sommet, que nous noterons $v_{critical}$ dont la disparition a pour résultat d'éclater G en c composantes connexes $C_1, C_2, ..., C_c$. On remarque que quel que soit i, $|\{v_{critical}\} \cup C_i| \geq 1 + \frac{n}{3}$, car si on considère $v_i \in C_i$, v_i a tous ses voisins dans $|\{v_{critical}\} \cup C_i|$.

Or n $= v_{critical} + C_1 + ... + C_c = 1 + \frac{cn}{3}$ ce qui entraîne $c \leq 2$. Et comme G n'est pas 2-connexe, $c = 2$.

Donc il existe un sommet $v_{critical}$ dont la disparition entraîne l'apparition de deux uniques composantes connexes C_1 et C_2. Or $|C_1| \geq \frac{n}{3}$, donc $|C_2| < \frac{2n}{3}$. De la même manière $|C_1| < \frac{2n}{3}$. On aimerait alors conclure en appliquant le théorème de Dirac à $C_1 \cup \{v_{critical}\}$, en remarquant que $k \geq \frac{|C_1 \cup \{v_{critical}\}|}{2}$.

Malheureusement, on ne peut pas, vu que $v_{critical}$ envoie une partie de son degré dans C_2. Il suffit simplement d'appliquer le théorème de Posa, qui est une version généralisée du théorème de Dirac, et qui peut s'appliquer ici.

Théorème 8. *Soit G un graphe simple à n sommets tel que :*

— *$\forall k, 1 \leq k < \frac{n-1}{2}$, le nombre de sommets de degré inférieur ou égal à k est strictement inférieur à k*

— *le nombre de sommets de degré inférieur ou égal à $\frac{n-1}{2}$ est inférieur ou égal à $\frac{n-1}{2}$*

Alors G est hamiltonien.

Comme $v_{critical}$ est le seul sommet de degré inférieur à $\frac{n-1}{2}$, on peut effectivement utiliser le théorème de Posa.

Et donc, $C_1 \cup \{v_{critical}\}$ et $C_2 \cup \{v_{critical}\}$ admettent chacun un cycle hamiltonien. En supprimant dans les deux une arête adjacente à $v_{critical}$ et en les fusionnant, on trouve donc un chemin hamiltonien de G. $\square$

On remarque de plus qu'on a prouvé bien mieux, puisqu'on a démontré que tout graphe connexe et k-régulier avec moins de $3k$ sommets admet un chemin hamiltonien.

3 Conclusion

Au final, je n'ai pas réussi à résoudre le problème, mais je suis quand même satisfait des résultats trouvés, même s'il y a une sensation d'inachevé. En effet, nous avons passé les deux premières semaines à nous engouffrer dans des impasses stériles (les lollipops, les décompositions en anse, découper le graphe pour trouver une relation de récurrence etc...), et je serais bien resté quelques semaines de plus à approfondir dans la direction finale.

En particulier, on remarque dans notre théorème final que l'on pourrait surement pousser jusqu'à $k \geq \frac{n}{4}$. En effet, le cas «connexe mais pas 2-connexe» se résout de la même manière. Reste alors le cas «2-connexe», que l'on résolvait avec le théorème de Bill Jackson. Si on regarde la démonstration de ce théorème, on se rend compte que Bill Jackson démontre dans un premier temps que si $k \geq \frac{n}{3}$, et C est un dominating cycle, alors $G - C$ n'est composé que de composantes réduites à un sommet.

Pourtant, dans le reste de sa démonstration, on a besoin d'un résultat beaucoup moins fort, qui est que $G - C$ possède *au moins* une composante connexe réduite à un sommet, ce qui pourrait être toujours le cas lorsque $k \geq \frac{n}{4}$.

De plus, il se pourrait que la conjecture soit encore vraie si on la généralise aux graphes quelconques (i.e. pas forcément régulier) et qu'on appelle k le degré minimal de G, mais cela me semble hors d'atteinte, sachant que nous n'avons même pas pu démontré $\mu(G) = O(\frac{n}{k})$.

4 Bibliographie

Références

[1] C. MAGNANT, D. M. MARTIN. *A note on the path cover number of regular graphs.* Australasian Journal of Combinatorics 43, 211-217 (2009)

[2] U. FEIGE, R. RAVI, M. SINGH, *Short Tours through large Linear Forests.* Springer International Publishing Switzerland, 273-284 (2014)

[3] D. R. WOODALL, *The binding number of a graph and its Anderson number.* J. Combinatorial Theory Ser. B 15, 225-255 (1973)

[4] B. JACKSON, *Hamilton cycles in regular 2-connected graphs.* J. Combinatorial Theory Ser. B (1978)