

90.05M

SESSION 2009

Filière MP (groupe I)

Épreuve commune aux ENS de Paris, Lyon et Cachan

MATHÉMATIQUES – INFORMATIQUE

Durée : 4 heures

Les calculatrices ne sont pas autorisées.

Le sujet porte sur l'étude de structures combinatoires et algébriques associées aux graphes. La première partie traite de quelques propriétés de la matrice Laplacienne. La seconde partie étudie une structure algébrique naturellement associée à un graphe, son groupe critique. La troisième partie porte sur l'étude d'un système de réécriture sur le graphe qui conduit à une représentation combinatoire du groupe critique. La quatrième et dernière partie est consacrée à des questions de complexité du calcul dans cette représentation.

Le sujet progresse au fil des parties et il est conseillé de les aborder dans l'ordre. Par contre il est bien entendu permis d'utiliser les résultats des questions précédentes sans y avoir répondu.

PARTIE 1 : Matrice d'incidence et matrice Laplacienne d'un graphe

Un graphe G est formé d'un ensemble fini X de sommets, qu'on prendra, sauf mention explicite du contraire, égal à $X_n = \{1, \dots, n\}$, et d'un ensemble E de paires non ordonnées de sommets appelées les arêtes. On représente graphiquement un graphe en associant les sommets à des points distincts du plan et en traçant des arcs joignant les paires de sommets qui forment des arêtes. Le graphe $G_0 = (X, E)$ avec $X = \{1, 2, 3, 4, 5\}$ et $E = \{\{1, 2\}, \{1, 4\}, \{2, 3\}, \{2, 4\}, \{3, 4\}, \{3, 5\}\}$ est représenté à la figure 1, ainsi que le graphe cycle C_6 et le graphe roue W_5 .

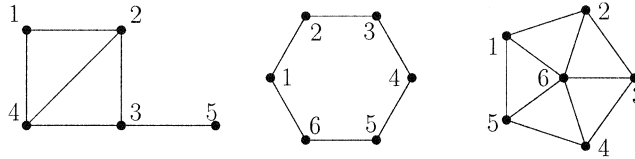


FIG. 1 – Trois graphes : G_0 , C_6 et W_5

L'arête $\{i, j\}$ est dite incidente aux sommets i et j et dans ce cas les sommets i et j sont dits adjacents. Le degré d_i du i ème sommet est le nombre d'arêtes qui lui sont incidentes. La matrice d'incidence d'un graphe $G = (X, E)$ à n sommets et m arêtes est la matrice $L_G = (\ell_{i,j})$ de taille $n \times m$ avec

$$\ell_{i,k} = \begin{cases} 1 & \text{si la } k\text{-ème arête est } \{i, j\} \text{ avec } i < j, \\ -1 & \text{si la } k\text{-ème arête est } \{i, j\} \text{ avec } i > j, \\ 0 & \text{sinon.} \end{cases}$$

où on a ordonné les arêtes arbitrairement, par exemple en écrivant chaque arête (i, j) avec $i < j$ et en utilisant l'ordre lexicographique : $(i, j) < (k, \ell)$ si $i < k$ ou si $(i = k \text{ et } j < \ell)$.

La matrice d'adjacence complétée ou matrice Laplacienne du graphe G est la matrice symétrique $\Delta_G = (e_{i,j})$ de taille $n \times n$ avec

$$e_{i,j} = \begin{cases} d_i & \text{si } j = i, \\ -1 & \text{si } j \neq i \text{ et les sommets } i \text{ et } j \text{ sont adjacents,} \\ 0 & \text{sinon.} \end{cases}$$

Pour le graphe G_0 de la figure 1 on a

$$L_{G_0} = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 1 & 1 \\ 0 & -1 & 0 & -1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{pmatrix}, \quad \Delta_{G_0} = \begin{pmatrix} 2 & -1 & 0 & -1 & 0 \\ -1 & 3 & -1 & -1 & 0 \\ 0 & -1 & 3 & -1 & -1 \\ -1 & -1 & -1 & 3 & 0 \\ 0 & 0 & -1 & 0 & 1 \end{pmatrix}$$

Question 1.1. Montrer que pour tout G , $\Delta_G = L_G^t L_G$ où ${}^t L_G$ désigne la matrice transposée de L_G .

Question 1.2. Montrer que si un réel λ et un vecteur colonne $v \neq 0$ satisfont $\Delta_G v = \lambda v$ alors $\lambda = \frac{\|{}^t L_G v\|^2}{\|v\|^2}$ et en déduire $\lambda = \frac{\sum_{\{i,j\} \in E} (v_i - v_j)^2}{\sum_i v_i^2}$ où $v = (v_i)_{i=1}^n$. Commentez le fait que les valeurs propres de Δ_G soient réelles positives.

Un chemin de longueur k du sommet x au sommet y du graphe G est une suite u_0, u_1, \dots, u_k de sommets avec $u_0 = x$, $u_k = y$ et telle que $\{u_{i-1}, u_i\} \in E$ pour tout $i = 1, \dots, k$. S'il existe un chemin de x à y dans G on dit que x et y sont connectés dans G . La relation être connectés est

une relation d'équivalence sur les sommets de G et on appelle composantes connexes de G classes d'équivalences. Le graphe G est connexe s'il n'a qu'une composante connexe, autrement dit s'il existe un chemin entre toute paire de sommets de G .

Question 1.3. Montrer que si le graphe G est connexe alors $\ker(\Delta_G) = \{v \mid \Delta_G v = 0\}$ est le sous-espace vectoriel engendré par $e = {}^t(1, \dots, 1)$. En déduire dans ce cas que le rang de Δ_G est $n - 1$. Plus généralement donner une formule qui lie le rang de Δ_G au nombre de composantes connexes de G .

On va montrer que les résultats précédents restent vrais si on remplace L_G par $L_{G,k}$ la matrice $n \times (m + 1)$ dont les lignes $i \neq k$ sont les lignes de L_G augmentée d'un 0 en colonne $m + 1$ et la ligne k est une ligne de zéro terminée par un 1 en colonne $m + 1$, et Δ_G par $\Delta_{G,k}$ la matrice $n \times n$ formée à partir de Δ_G en remplaçant la k ème ligne et la k ème colonne par les k ème ligne et k ème colonne d'une matrice identité $n \times n$.

Question 1.4. Donner une relation analogue à celle de la question 1 entre $L_{G,k}$ et $\Delta_{G,k}$, et une relation analogue à celle de la question 2 entre une valeur propre λ de $\Delta_{G,k}$ et un éventuel vecteur propre v non nul. En déduire que si le graphe G est connexe alors $\Delta_{G,k}$ est de rang n .

On note x_1, \dots, x_n les n vecteurs de la base canonique de l'espace vectoriel \mathbb{R}^n : x_i a toutes ses coordonnées nulles sauf la i ème égale à 1. On pose $\Delta_i = d_i x_i + \sum_{j \neq i} e_{i,j} x_j$, le vecteur dont les coordonnées dans la base canonique sont données par la i ème ligne de Δ_G .

Question 1.5. Déduire des résultats précédents que la famille $\{\Delta_1, \dots, \Delta_n\} \setminus \{\Delta_k\} \cup \{x_k\}$ est une famille libre de vecteurs de \mathbb{R}^n , quel que soit k .

La question suivante est indépendante des précédentes et termine cette partie préliminaire.

Question 1.6. Soit A une matrice carrée $n \times n$ inversible et a un vecteur, tout deux à coefficients entiers. Notons v l'unique solution du système d'équations linéaires $Av = a$. Montrer que le vecteur v s'écrit sous la forme $(v'_1/\delta + v''_1, \dots, v'_n/\delta + v''_n)$ avec les v'_i , les v''_i et δ entiers et $0 \leq v'_i < \delta$ pour tout i .

PARTIE 2 : Le groupe critique d'un graphe

On considère dans tout le reste du sujet un graphe connexe $G = (X_n, E)$ de matrice Laplacienne $\Delta_G = (e_{i,j})$ de taille $n \times n$, telle que définie à la partie précédente. En particulier pour $i \neq j$, $e_{i,j} = -1$ si $\{i, j\} \in E$ et 0 sinon, et $d_i = e_{i,i}$ est le degré du sommet i .

Un groupe abélien K est un ensemble muni d'une loi interne, notée $+$, associative, commutative, admettant un élément neutre et telle que chaque élément de K ait un inverse. Un sous-groupe H de K est un sous-ensemble de K stable pour la loi interne $+$ et qui forme lui-même un groupe. Étant donnés des éléments g_1, \dots, g_k de K , le sous-groupe de K engendré par les g_i , noté $\langle g_1, \dots, g_k \rangle$, est le plus petit sous-groupe de K les contenant.

Question 2.1. Montrer que $\langle g_1, \dots, g_k \rangle$ est formé de tous les éléments de K qui s'écrivent $\sum_{i=1}^k a_i g_i$ avec $a_i \in \mathbb{Z}$.

On sera attentif à ne pas confondre la notion de sous-groupe engendré par des éléments d'un groupe abélien avec la notion d'espace vectoriel réel engendré par des vecteurs d'un espace vectoriel : dans le premier cas on forme des sommes et différences de vecteurs, qui se réécrivent comme combinaisons linéaires mais à coefficients entiers, alors que le second cas on forme toutes les combinaisons linéaires à coefficients réels.

Par exemple, l'ensemble \mathbb{R}^n peut-être vu comme un groupe pour la loi usuelle d'addition des vecteurs, et le sous-ensemble \mathbb{Z}^n des vecteurs à coordonnées entières en est le sous-groupe $\langle x_1, \dots, x_n \rangle$ engendré par les vecteurs de la base canonique. Ce sous-groupe \mathbb{Z}^n est bien différent du sous-espace vectoriel engendré par les mêmes vecteurs x_1, \dots, x_n (qui est \mathbb{R}^n tout entier).

On note $\Delta(G, k)$ le sous-groupe de \mathbb{Z}^n engendré par le vecteur x_k et les Δ_i pour $i = 1, \dots, n$, définis comme à la partie précédente par $\Delta_i = d_i x_i + \sum_{j \neq i} e_{i,j} x_j$.

Question 2.2. Montrer que $\Delta(G, k)$ est engendré par x_k et n'importe quel sous-ensemble de $n - 1$ des Δ_i .

Deux éléments x et y d'un groupe abélien K sont dits équivalents relativement au sous-groupe H si $x - y \in H$.

Question 2.3. Montrer que la relation d'équivalence relativement à un sous-groupe est bien une relation d'équivalence.

L'ensemble des classes d'équivalences pour la relation d'équivalence relativement à un sous-groupe est noté K/H et appelé le quotient de K par H . La classe d'équivalence de x relativement à H est notée x/H , ou plus simplement \bar{x} lorsque H est clair. Dans chaque classe d'équivalence de K relativement à H on choisit arbitrairement un élément qu'on appelle le représentant de la classe et on définit la somme de deux classes d'équivalence comme la classe d'équivalence de la somme de leurs représentants.

Question 2.4. Montrer que l'opération de somme décrite ci-dessus munit K/H d'une structure de groupe abélien.

On définit le groupe critique $C(G, k)$ du graphe G enraciné en k comme le quotient $\mathbb{Z}^n / \Delta(G, k)$.

Question 2.5. Montrer que $C(G, k)$ est un groupe de cardinal fini.

Deux groupes K, K' sont isomorphes (on écrit $K \sim K'$) s'il existe un isomorphisme de groupe entre K et K' , c'est-à-dire une bijection ϕ de K sur K' telle que $\phi(x + y) = \phi(x) + \phi(y)$.

Question 2.6. Montrer que si ϕ est un isomorphisme de groupe entre K et K' , et H est un sous-groupe de K alors $K/H \sim K'/\phi(H)$.

Question 2.7. Montrer que $C(G, k) \sim C(G, \ell)$ pour tout k, ℓ . Indication : on pourra considérer les $\Delta'_i = \Delta_i - e_{i,k} x_\ell$ pour tout i , les vecteurs $y_i = x_i - x_k$ pour $i \neq k$ et $y_k = -x_\ell$, et l'application $\phi(x_i) = y_i$ pour tout i .

Le groupe critique du graphe G étant, à isomorphisme près, indépendant de la racine on le note $C(G)$.

PARTIE 3 : Tas de sable sur un graphe et configurations récurrentes

On considère toujours un graphe connexe $G = (X_n, E)$ de matrice Laplacienne Δ_G comme à la partie précédente. On va maintenant étudier un système de réécriture qu'il est commode d'interpréter comme un modèle de tas de sable : des grains de sable sont répartis sur les sommets d'un graphe et se déplacent le long des arêtes selon des règles d'éboulements. L'un des sommets du graphe est appelé le puits, les grains qui y tombent s'y accumulent.

Formellement on définit une configuration du graphe comme un élément de \mathbb{Z}^n , la i ème coordonnée s'interprétant comme le nombre de grains au sommet i , le n ème sommet représentant le puits. Une configuration $u = (u_i)_{1 \leq i \leq n}$ est positive si $u_i \geq 0$ pour tout $i = 1, \dots, n - 1$: le nombre de grains de sable doit être positif en chaque sommet, sauf peut-être dans le puits.

L'éboulement d'un sommet consiste à lui retirer un grain par arête incidente et à distribuer ces grains aux voisins. Plus précisément, étant données deux configurations positives u et v on écrit $u \rightarrow v$ s'il existe un $i \leq n - 1$ tel que $v = u - \Delta_i$ (où Δ_i est défini comme dans la partie précédente), et on dit que v est obtenu à partir de u par éboulement du sommet i . On note $\xrightarrow{*}$ la clôture transitive de la relation $\rightarrow : u \xrightarrow{*} v$ si et seulement si il existe des configurations u^0, \dots, u^p telles que $u^0 = u$, $u^p = v$ et $u^i \rightarrow u^{i+1}$ pour tout $0 \leq i < p$. Remarquons encore une fois qu'on éboule jamais le puits.

Question 3.1. Montrer que si $u \xrightarrow{*} v$ alors u et v ont la même image dans le groupe $C(G, n)$.

On note S_k l'ensemble des sommets à distance k du sommet n dans le graphe $G : S_0 = \{n\}$, S_1 contient les voisins du sommet n , S_2 les voisins de ces voisins qui ne sont pas déjà dans $S_0 \cup S_1$, etc. On pose $\ell = \max\{i \mid S_i \neq \emptyset\}$ et pour $k \leq \ell$, $\mu_k(u) = \sum_{i \in S_k} u_i$ le nombre de grains à distance k du puits dans la configuration u . Le vecteur $\mu(u) = (\mu_0(u), \dots, \mu_\ell(u))$ est appelé potentiel de la configuration u . Une configuration positive u est dite stable si aucun sommet ne peut s'ébouler : $u_i \leq d_i$ pour tout $i \leq n - 1$.

Question 3.2. Montrer que (a) pour toute configuration positive u il existe une configuration stable v telle que $u \xrightarrow{*} v$ et que (b) cette configuration est unique. Indication : pour (a) on pourra s'appuyer sur la notion de potentiel introduite plus haut.

On appelle avalanche une suite d'éboulements qui se termine par une configuration stable. Une configuration u est dite récurrente si elle est stable et s'il existe une configuration positive $v \neq 0$ telle que $u + v \xrightarrow{*} u$. Soit δ la configuration avec $\delta_i = d_i$ pour tous les sommets, on remarque que si u est stable alors $\delta - u$ est positive.

Question 3.3. Montrer les points suivants :

- (a). Si $u \xrightarrow{*} u'$ et $v \xrightarrow{*} v'$ alors $u + v \xrightarrow{*} u' + v'$.
- (b). Pour toute configuration positive $v \neq 0$ il existe un entier k et une configuration w (non nécessairement stable) telle que $kv \xrightarrow{*} w$ et $w_i > 0$ pour tout $1 \leq i \leq n - 1$.
- (c). Une configuration stable u est récurrente si et seulement s'il existe une configuration positive u' telle que $u' + \delta \xrightarrow{*} u$.

Pour toute paire (u, v) de configurations positives on note $u \oplus v$ l'unique configuration stable telle que $u + v \xrightarrow{*} u \oplus v$. Enfin pour toute paire de configurations (non nécessairement positives) on écrit $u \xRightarrow{*} v$ s'il existe un sommet $i \leq n - 1$ tel que $v = u - \Delta_i$ et $\xRightarrow{*}$ la clôture transitive de cette relation.

Question 3.4. Montrer que si u et v sont deux configurations telles que $u - v \in \langle \Delta_1, \dots, \Delta_n \rangle$ alors il existe une configuration w telle que $w \xRightarrow{*} u$ et $w \xRightarrow{*} v$.

Soit $\varepsilon = 2\delta - (\delta \oplus \delta)$ où δ est la configuration définie précédemment.

Question 3.5. Montrer que la configuration ε est positive et que $\delta + \varepsilon \xrightarrow{*} \delta$.

Question 3.6. Montrer qu'une configuration u est récurrente si et seulement si $u + \varepsilon \xrightarrow{*} u$.

Question 3.7. Montrer que pour toute configuration u il existe une unique configuration récurrente v telle que $u - v \in \langle \Delta_1, \Delta_2, \dots, \Delta_n \rangle$.

Question 3.8. Montrer que l'opération \oplus munit l'ensemble $R(G)$ des configurations récurrentes de G d'une structure de groupe et que le groupe ainsi obtenu est isomorphe à $C(G)$.

Nous avons donc construit une représentation du groupe critique d'un graphe en terme d'éboulements.

PARTIE 4 : Configurations récurrentes et arbres couvrants

On suppose dans cette partie que le graphe connexe G sur lequel on étudie le tas de sable a n sommets et m arêtes. On désigne comme précédemment par S_i l'ensemble des sommets à distance i du puits, on pose $\ell = \max(i \mid S_i \neq \emptyset)$ et on note $\mu_i(u)$ le nombre de grains à distance i du puits dans la configuration u . On pose de plus $S_{>i} = \bigcup_{j>i} S_j$ (l'ensemble des sommets à distance au moins $i + 1$ du puits) et $\mu_{\geq i}(u) = \sum_{j \geq i} \mu_j(u)$ (le nombre de grains à distance au moins i du puits).

Question 4.1. Montrer que le nombre total de grains hors du puits dans une configuration stable est majoré par $2m$.

Question 4.2. Montrer que deux séquences d'éboulements qui conduisent d'une configuration u à une configuration stable v font ébouler le même nombre de fois le sommet i pour tout i : autrement dit, deux séquences d'éboulements complètes ne diffèrent que par l'ordre des éboulements.

Il est donc possible de parler du nombre de sommets éboulés pour passer d'une configuration positive à la configuration stable associée. Remarquons pour simuler efficacement un éboulement sur ordinateur il faudrait, entre autres choses, gérer dynamiquement l'ensemble des sommets éboulables, pour ne pas perdre de temps à les chercher. On se concentre ici sur la complexité intrinsèque du modèle en prenant comme définition de complexité d'une avalanche le nombre d'éboulements qui la compose.

Question 4.3. On considère une configuration u ayant p grains hors du puits et on l'éboule dans l'ordre suivant :

- Pour i décroissant de ℓ jusqu'à 1, répéter les opérations suivantes :
 - (a). ébouler simultanément tous les sommets instables de S_i ;
 - (b). tant qu'il y a des sommets instables dans $S_{>i}$, les ébouler ;
 - (c). s'il y a à nouveau des sommets instables dans S_i reprendre à l'étape (a).

Montrer que le nombre d'éboulements effectués à l'étape i est au plus $\mu_{\geq i} \cdot |S_{>i}|$ et en déduire une borne sur la complexité d'une avalanche partant de u en fonction de ℓ , n et p (autrement dit majorer le nombre d'éboulements de $u \xrightarrow{*} v$ avec v stable).

Question 4.4. Appliquer votre borne pour majorer en fonction de ℓ , n et m la complexité du test $u + \varepsilon \xrightarrow{*} u$ pour déterminer si une configuration stable est récurrente ? Même question pour le calcul de $u \oplus v$ pour u et v récurrentes ?

Le test $u + \varepsilon \xrightarrow{*} u$ fait intervenir la configuration ε qui contient au moins $2m$ grains et qu'il faut avoir calculé au préalable à partir de sa définition. On donne maintenant une caractérisation un peu plus « efficace » des configurations récurrentes. Soit $\beta = -\Delta_n$ la configuration obtenue par éboulement du puits : $\beta_i = 1$ pour tous les voisins du puits et $\beta_n = -d_n$.

Question 4.5. Montrer qu'une configuration est récurrente si et seulement si $u + \beta \xrightarrow{*} u$. Montrer de plus que si u est récurrent, dans une séquence d'éboulements de $u + \beta$ à u , chaque sommet s'éboule exactement une fois. En déduire un test de récurrence de complexité linéaire.

Nous allons maintenant voir que le test $u + \beta \xrightarrow{*} u$ permet d'associer les configurations récurrentes à d'autres structures naturelles sur le graphe.

L'algorithme suivant, appelé algorithme thermique, prend en entrée une configuration stable u et effectue les opérations suivantes :

-
- $u^{(0)} = u + \beta$; $R_{-1} = \{n\}$; $i := 0$
 - répéter tant que $u^{(i)}$ n'est pas stable :
 - soit $R_i = \{k \mid k < n, u_k^{(i)} \geq d_k\}$ l'ensemble des sommets instables de $u^{(i)}$,
 - soit $u^{(i+1)} = u^{(i)} - \sum_{j \in R_i} \Delta_j$,
 - soit $A_i = \{\{k, \ell\} \mid k \in R_i, \ell = \text{select}(V_k \cap R_{i-1}, u_k^{(i)} - d_k)\}$
 - $i := i+1$
 - $A = \bigcup A_i$

où la fonction $\text{select}(I, j)$ renvoie le $(j+1)$ ème plus petit élément d'un ensemble fini I d'entiers (ou $+\infty$ si $|I| \leq j$), et où V_k désigne l'ensemble des sommets voisins de k dans G .

Question 4.6. Montrer que l'algorithme thermique termine et que si la configuration de départ est récurrente les R_i forment une partition des sommets de G et les A_i sont des ensembles d'arêtes bien formées (pas de $+\infty$ renvoyé par select).

Un sous-graphe d'un graphe $G = (X, E)$ est un graphe $G = (X, E')$ avec $E' \subset E$. Un chemin u_0, \dots, u_k de longueur k est un cycle si $u_0 = u_k$ et $k \geq 2$. Un cycle est simple si pour tout $0 \leq i < j \leq k-1$, $u_i \neq u_j$. Un graphe est une forêt s'il ne contient pas de cycle simple, c'est un arbre s'il est de plus connexe. Un arbre couvrant d'un graphe $G = (X, E)$ est un sous-graphe $T = (X, A)$ de G qui est un arbre.

Question 4.7. Montrer que l'algorithme thermique appliqué à une configuration récurrente construit un ensemble d'arêtes A tel que (X, A) soit un arbre couvrant du graphe.

Question 4.8. Montrer que deux configurations récurrentes différentes donnent par l'algorithme thermique deux arbres couvrants distincts.

Question 4.9. Montrer réciproquement qu'à tout arbre couvrant (X, A) est associée une configuration récurrente qui le redonne par l'algorithme thermique.

Question 4.10. Dédurre des questions précédentes que le nombre de configurations récurrentes d'un graphe est égal au nombre de ses arbres couvrants.

Fin de l'épreuve