



Markov processes spectrum

Yiyan Pan-Yu

► **To cite this version:**

Yiyan Pan-Yu. Markov processes spectrum. Modélisation et simulation. Université Joseph-Fourier - Grenoble I, 1997. Français. tel-00004959

HAL Id: tel-00004959

<https://tel.archives-ouvertes.fr/tel-00004959>

Submitted on 20 Feb 2004

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.

Année 1997

THESE

Présentée par

Yiyan PAN-YU

Pour obtenir le titre de

Docteur en Mathématiques Appliquées

DE L'UNIVERSITE JOSEPH FOURIER-GRENOBLE 1

(arrêtés ministériels du 5 juillet 1984 et du 30 mars 1992)

SPECTRES DE PROCESSUS DE MARKOV

DATE DE SOUTENANCE : 18 juin 1997

COMPOSITION DU JURY :

Yves	COLIN DE VERDIERE	Président
Laurent	MICLO	Rapporteur
Laurent	SALOFF-COSTE	Rapporteur
Olivier	FRANCOIS	Examineur
Bernard	YCART	Directeur de thèse

Thèse préparée au sein du Laboratoire LMC de l'Institut IMAG de Grenoble

A mes parents

Remerciements

Il est vain d'espérer remercier en quelques mots une personne qui a changé ma vie en m'amenant sur la voie de recherches mathématiques et en me disant "Je vous fais confiance", une personne qui n'a jamais hésité à me permettre d'enrichir mes expériences, une personne qui m'encourage en cas d'hésitations et de difficultés, qui m'applaudit aux moindres progrès, une personne qui avec sa clairvoyance a dirigé ma thèse toujours dans la bonne direction, une personne qui pour moi a été un modèle d'enthousiasme, de conscience professionnelle, de générosité et de compréhension. Devant une personne que je respecte tant, la seule crainte que j'avais le long de ma thèse était de la décevoir. Je remercie du fond de mon cœur Bernard Ycart.

Je suis très reconnaissante à Yves Colin de Verdière de m'avoir inspirée par ses idées exceptionnelles le théorème A que vous aller lire. Je suis très honorée qu'il ait accepté de présider le jury. Qu'il soit remercié chaleureusement.

Je tiens à remercier Laurent Miclo et Laurent Saloff-Coste pour avoir minutieusement lu ce texte en y consacrant beaucoup de temps et pour avoir fait des remarques pertinentes.

Je voudrais remercier également Olivier François pour les discussions que j'ai pu avoir avec lui et pour sa participation au jury.

Enfin merci à Serge Dégirine, Carole Desprez et Bernard Van Cutsem qui m'ont aidée dans mes préparations des cours T.D.

Table des matières

1	Introduction	9
1.1	Théorème A	9
1.2	Théorème B	13
1.3	Plan du texte	14
2	Gaps des processus de Markov	17
2.1	Générateurs de Markov et leurs gaps	17
2.1.1	Réversibilité, stationnarité et gap	18
2.1.2	Processus et chaînes	24
2.2	Exemples d'applications du gap	27
2.2.1	Algorithme de Metropolis et échantillonnage de Gibbs	28
2.2.2	Modèle d'Ising	30
2.3	Méthodes principales d'étude du gap	34
2.3.1	Techniques de comparaison	34
2.3.2	Différentes versions de l'inégalité de Cheeger	35
2.3.3	Etude de comportement asymptotique spectral	37
2.3.4	Méthode de couplage	38
3	Spectres des générateurs quasi réductibles	43
3.1	Enoncé et démonstration du théorème A	44
3.1.1	Structure hiérarchique	44
3.1.2	Décomposition hiérarchique du générateur	67

3.1.3	Propositions préparatoires	71
3.1.4	Démonstration du théorème A	93
3.1.5	Discussion	114
3.2	Applications	116
3.2.1	Processus quasi décomposables	116
3.2.2	Processus de naissance et de mort à homothétie interne .	118
3.2.3	Recuit simulé à basse température	121
3.2.4	Processus d'Ising en dimension un	134
4	Moyenne harmonique du spectre	147
4.1	Démonstration du théorème B	147
4.2	Applications	158
4.2.1	Gap géométrique	158
4.2.2	Processus d'Ising	163

Chapitre 1

Introduction

Le sujet central de cette thèse est l'étude des spectres des processus de Markov finis irréductibles et réversibles. Dans le spectre, le trou spectral qu'on appellera *gap* est surtout notre point focal d'intérêt. C'est un sujet étudié dans de nombreux domaines: en probabilités ([1, 20]), en statistiques ([32, 72]), en géométrie ([6, 10, 17, 63]), en physique mathématiques ([55, 64]) etc. Les deux résultats principaux de la thèse sont le théorème A et le théorème B. Le premier donne un développement asymptotique du spectre d'un processus quasi réductible. Le second donne une minoration du *gap* d'un processus, basée sur le calcul explicite de la somme des inverses des valeurs propres non nulles d'un processus dont le graphe de transition est arborescent. Nous commençons par décrire et énoncer ces deux résultats.

1.1 Théorème A

Dans ce premier résultat, on considère des générateurs de Markov Λ sur un espace d'états fini E qui sont fonctions de ε , strictement positif et destiné à tendre vers 0, et qui sont analytiques en $\varepsilon = 0$ au sens où il existe un prolongement analytique de Λ sur un voisinage de 0. Les générateurs que nous étudions sont réversibles et presque réductibles, au sens où la limite de Λ , noté $\Lambda(0)$, est

réductible. Pour $k \geq 0$, notons $\theta(\varepsilon^k)$ une grandeur telle que

$$0 < \left| \lim_{\varepsilon \searrow 0} \frac{\theta(\varepsilon^k)}{\varepsilon^k} \right| < +\infty$$

et $O(\varepsilon^k)$ telle que

$$0 \leq \left| \lim_{\varepsilon \searrow 0} \frac{O(\varepsilon^k)}{\varepsilon^k} \right| < +\infty .$$

D'après la théorie des perturbations des opérateurs linéaires (voir [44]), les valeurs propres de Λ , ainsi que les sous-espaces propres associés convergent vers les valeurs propres et les sous-espaces propres correspondants de $\Lambda(0)$, autrement dit, le spectre de $\Lambda(0)$ donne l'équivalent des valeurs propres d'ordre $\theta(1)$ de Λ . Mais $\Lambda(0)$ étant réductible, la valeur propre 0 a en général une multiplicité strictement supérieure à 1. Ce qui signifie que certaines valeurs propres de Λ sont d'ordre $O(\varepsilon)$. Or ce sont précisément ces valeurs propres qui contrôlent la vitesse d'accès à l'équilibre du processus Λ . Il est important d'avoir une estimation précise, c'est-à-dire un équivalent de ces "petites" valeurs propres.

Il est difficile voire impossible de calculer les valeurs propres, même si la taille de l'espace d'états est modérée. Nous proposons une méthode générale de calcul explicite des équivalents des valeurs propres d'ordre $\theta(\varepsilon^k)$ pour tout $k \geq 1$ qui vise à calculer les valeurs propres sur un espace d'états largement réduit. Précisément, notons $S_k(\Lambda)$ le vecteur composé des valeurs propres d'ordre $O(\varepsilon^k)$ du générateur Λ par ordre croissant avec multiplicité éventuelle. Notre méthode donne explicitement un générateur Λ_k sur un espace d'états de taille $\#S_k(\Lambda)$, la dimension de $S_k(\Lambda)$, tel que $S_0(\Lambda_k)$ et $S_k(\Lambda)/\varepsilon^k$ ont la même limite. Les équivalents des valeurs propres d'ordre $\theta(\varepsilon^k)$ de Λ sont ainsi donnés par les valeurs propres non nulles de $\lim_{\varepsilon \searrow 0} \Lambda_k$ multipliées par ε^k . Supposons que le gap de Λ est d'ordre $\theta(\varepsilon^{K-1})$. S'il n'y a que deux valeurs propres de Λ d'ordre $O(\varepsilon^{K-1})$ (0 et le gap), alors l'équivalent du gap est donné par la valeur propre non nulle multipliée par ε^{K-1} de $\lim_{\varepsilon \searrow 0} \Lambda_{K-1}$ qui est une matrice d'ordre 2.

Expliquons l'idée de cette méthode intuitivement. Notons \mathcal{E}_1 l'ensemble des classes irréductibles, transientes et récurrentes, par rapport à $\Lambda(0)$ et \mathcal{F}_1 l'ensemble des classes récurrentes de $\Lambda(0)$ qui est donc un sous-ensemble de

\mathcal{E}_1 . En fait, on peut voir les classes de \mathcal{E}_1 d'une autre façon en considérant le processus Λ . Pour ε suffisamment petit, partant d'une classe transiente élément de $\mathcal{E}_1 \setminus \mathcal{F}_1$, le processus en sort et va vers une classe récurrente dans un temps en moyenne d'ordre $\theta(1)$, et dans une même échelle de temps, partant d'une classe récurrente de \mathcal{F}_1 , le processus peut atteindre n'importe quel état de cette classe. Or la sortie des classes récurrentes se produira dans un temps en moyenne d'ordre $\theta(1/\varepsilon)$. Donc si on se place sur une échelle de temps de temps d'ordre $\theta(1)$, on voit que le processus atteint son équilibre à l'intérieur de chaque classe récurrente de \mathcal{F}_1 , mais on ne voit pas de communications entre les classes de \mathcal{F}_1 . Si on souhaite observer des transitions entre les classes de \mathcal{F}_1 , il faut se placer sur des intervalles de temps beaucoup plus longs d'ordre $\theta(1/\varepsilon)$, c'est-à-dire changer l'échelle de temps d'un facteur ε . Cette fois ci, le processus atteint à son équilibre à l'intérieur de chaque classe de \mathcal{F}_1 presque immédiatement. On peut alors ignorer la façon dont le processus atteint son équilibre à l'intérieur des classes et identifier les états d'une même classe récurrente en les considérant comme un seul état. Par contre ce sont les mouvements entre les classes récurrentes de \mathcal{F}_1 qui sont liés aux équivalents des valeurs propres d'ordre $\theta(\varepsilon)$. Ces mouvements correspondent à un processus de Markov sur \mathcal{F}_1 . Notant Λ_1 le générateur de ce processus qui est d'ordre $\#\mathcal{F}_1$, on montre que une fois multipliées par ε , les valeurs propres non nulles de $\lim_{\varepsilon \searrow 0} \Lambda_1$ ont les même équivalents que les valeurs propres d'ordre $\theta(\varepsilon)$ du générateur initial Λ . En itérant cette procédure, nous définirons une partition \mathcal{E}_k de l'espace d'états E et un sous-ensemble \mathcal{F}_k de \mathcal{E}_k ayant les propriétés suivantes. Dans un temps d'ordre $\theta(1/\varepsilon^{k-1})$ en moyenne, partant d'une classe de \mathcal{E}_k le processus peut atteindre n'importe quel état de cette classe et en sortir en allant vers une classe de \mathcal{F}_k si la classe de départ est une classe de $\mathcal{E}_k/\mathcal{F}_k$. Par contre il faut un temps d'ordre $\theta(1/\varepsilon^k)$ en moyenne pour sortir d'une classe de \mathcal{F}_k . Alors on trouve que non seulement le nombre des valeurs propres d'ordre $\theta(\varepsilon^k)$ est égal au cardinal de \mathcal{F}_k mais aussi que l'équivalent de ces valeurs propres correspond à l'évolution du processus à l'échelle de temps d'ordre $\theta(1/\varepsilon^k)$. Sur cette échelle de temps, le processus atteint son équilibre à l'intérieur de chaque classe de \mathcal{F}_k tellement rapidement

qu'on peut ignorer les mouvement dans ces classes en considérant chacune de ces classes comme un "état". Or c'est exactement les mouvements entre ces "états" qui sont liés à l'équivalent des valeurs propres d'ordre $\theta(\varepsilon^k)$. En décrivant ces mouvements par un générateur Λ_k défini sur l'espace d'états \mathcal{F}_k , on montre que le nombre $\#\mathcal{F}_k$ est égale au nombre des valeurs propres d'ordre $O(\varepsilon^k)$ et que multipliées par ε^k , les valeurs propres non nulles de $\lim_{\varepsilon \searrow 0} \Lambda_k$ donnent les équivalents des valeurs propres d'ordre $\theta(\varepsilon^k)$ de Λ . En fait $\#\mathcal{F}_k$ est décroissant. Soit K le plus petit entier tel que

$$\#\mathcal{F}_K = 1 .$$

Les valeurs propres d'ordre ε^{K-1} sont les plus proches de zéro de tout le spectre, et ε^{K-1} est l'ordre de grandeur du gap. La construction récursive de $\{\mathcal{E}_k\}_{0 \leq k \leq K}$ et $\{\mathcal{F}_k\}_{0 \leq k \leq K}$ est assez longue et sera donnée explicitement dans le paragraphe 3.1.1.

Posons

$$U_{k+1} = E \Leftrightarrow \bigcup_{\alpha_k \in \mathcal{F}_k} \alpha_k$$

et notons $\Lambda_{U_{k+1}}$ le générateur restreint de Λ sur U_{k+1} . Pour toute classe α_k de \mathcal{F}_k , notons Π_{α_k} la mesure de probabilité réversible du générateur restreint de Λ sur α_k . Pour tous sous-ensembles S_1, S_2 de E avec $S_1 \cap S_2 = \emptyset$, notons

$$G_{S_1}^{S_2} = (\lambda_{ab})_{a \in S_1, b \in S_2}$$

la matrice composée des taux de transition de S_1 à S_2 . Posons $D_{U_{k+1}}$ la matrice diagonale telle que

$$D_{U_{k+1}} \mathbb{1}_{U_{k+1}} = \sum_{\alpha_k \in \mathcal{F}_k} G_{U_{k+1}}^{\alpha_k} \mathbb{1}_{\alpha_k}$$

où $\mathbb{1}_{U_{k+1}}$ et $\mathbb{1}_{\alpha_k}$ désignent les vecteurs constants de valeur 1 sur U_{k+1} et α_k respectivement.

Notre résultat principal est le suivant.

Théorème A *Soit $1 \leq k \leq K$.*

- (1) Pour tous $\alpha_k \neq \beta_k \in \mathcal{F}_k$, définissons le taux de transition de α_k à β_k comme suit,

$$\lambda_{\alpha_k \beta_k} = \frac{1}{\varepsilon^k} \Pi_{\alpha_k} \left[G_{\alpha_k}^{\beta_k} \mathbb{1}_{\beta_k} \Leftrightarrow G_{\alpha_k}^{U_{k+1}} \begin{array}{c} N_{U_{k+1}} \\ \swarrow \\ D_{U_{k+1}} \end{array} \Leftrightarrow D_{U_{k+1}} \right]^{-1} G_{U_{k+1}}^{\beta_k} \mathbb{1}_{\beta_k} \Big]$$

Alors $\lim_{\varepsilon \searrow 0} \lambda_{\alpha_k \beta_k}$ existe. Soit $l \geq k+1$ l'entier le plus petit tel que α_k soit contenu dans un élément de \mathcal{F}_l noté α_l . Alors si

$$\beta_k \not\subseteq \alpha_l$$

on a en plus

$$\lim_{\varepsilon \searrow 0} \lambda_{\alpha_k \beta_k} = 0 .$$

- (2) Pour tout a de U_{k+1} , notons $p_a^{\beta_k}$ la probabilité que le processus Λ , partant de a , atteigne β_k pour la première fois sans visiter les autres classes de \mathcal{F}_k . Soit $P_{U_{k+1}}^{\beta_k}$ le vecteur colonne composé des $p_a^{\beta_k}$ pour a de U_{k+1} . Alors

$$\lambda_{\alpha_k \beta_k} = \frac{1}{\varepsilon^k} \Pi_{\alpha_k} \left(G_{\alpha_k}^{\beta_k} \mathbb{1}_{\beta_k} + G_{\alpha_k}^{U_{k+1}} P_{U_{k+1}}^{\beta_k} \right) .$$

- (3) Soit Λ_k le générateur sur \mathcal{F}_k composé des taux de transition $\lambda_{\alpha_k \beta_k}$. Alors

$$\lim_{\varepsilon \searrow 0} S_0(\Lambda_k) = \lim_{\varepsilon \searrow 0} \frac{1}{\varepsilon^k} S_k(\Lambda) .$$

1.2 Théorème B

Dans le deuxième résultat, on donne un moyen d'estimer le gap d'un processus de Markov irréductible et réversible au travers de la moyenne harmonique du spectre du processus. L'idée est la suivante. On sait que le générateur a une valeur propre nulle, les autres étant négatives. Nous noterons SIVP la somme des inverses des valeurs propres non nulles de l'opposé du générateur. Leur moyenne harmonique est n/SIVP si la taille de l'espace d'états est $n+1$. Du fait que le gap est la valeur absolue de la valeur propre la plus proche de 0, il est compris dans l'intervalle $[1/\text{SIVP}, n/\text{SIVP}]$. C'est la minoration qui est la plus intéressante. Il n'est pas facile de calculer SIVP en général. Mais dans le

cas où le graphe réduit du générateur est un arbre, on donne une expression explicite de cette somme. On en déduit une borne inférieure du gap dans ce cas là. De plus, on constate que plus le gap est décalé des autres valeurs propres, plus important est le rôle joué par le terme de l'inverse du gap dans la somme, et donc meilleure est l'estimation. Revenons au cas général où le graphe réduit du générateur est quelconque. Intuitivement, moins le graphe possède d'arêtes, moins vite le processus parcourt toutes les arêtes selon la mesure stationnaire et plus faible sera la vitesse de convergence. De sorte que, si on choisit un arbre maximal du graphe réduit du générateur initial, le gap du générateur restreint sur l'arbre est plus petit que celui du générateur initial. On obtient ainsi une borne inférieure du gap d'un générateur quelconque.

Plus précisément, soit Λ un générateur irréductible et π -réversible sur l'espace d'états E de cardinal n . Soit G' un arbre maximal du graphe réduit du générateur Λ . Si on fixe un sommet $a_0 \in E$ comme racine, G' est alors orienté: pour tous $a \neq b \in E$, on note $b > a$ s'il existe un chemin sans répétition de b à a_0 passant par a . On note B_a le sous-arbre issu de a

$$B_a = \{b; a \leq b\} \quad .$$

Pour un sommet a différent de a_0 , on note μ_a le taux de transition de a vers son père, le sommet unique b tel que $b < a$ et (a, b) est une arête de G' . On note

$$\phi^*(B_a) = \frac{\pi(a)\mu_a}{\pi(B_a) [1 \leftrightarrow \pi(B_a)]} \quad .$$

On a le résultat suivant.

Théorème B

$$\text{gap}(\Lambda) \geq \left[\sum_{\substack{a \in E \\ a \neq a_0}} \frac{1}{\phi^*(B_a)} \right]^{-1} \quad .$$

1.3 Plan du texte

Le chapitre 2 est un chapitre de préliminaires, qui contient une description des motivations et des objets utilisés dans le reste de la thèse ainsi que des résultats existants dans ce domaine.

Le chapitre 3 est consacré à la démonstration et aux applications du théorème A. On commence (paragraphe 3.1.1) par donner une structure hiérarchique de l'espace d'états, déterminée par le générateur de Markov, et les propriétés qui lui sont attachées. Puis on donne les propriétés du générateur décomposé en blocs selon la structure hiérarchique de l'espace d'états (3.1.2). Ensuite, certaines propositions techniques qui sont indispensables pour la démonstration finale du théorème A sont rassemblées dans le paragraphe 3.1.3. En fin la démonstration du théorème A sera donnée dans le paragraphe 3.1.4. Dans de nombreux cas d'importance pratique, le calcul de la structure hiérarchique et des générateurs Λ_k est très simplifié par rapport au cas général. Nous traiterons en particulier le cas des processus quasi décomposables ([18]), avec comme illustration un processus de naissance et de mort à homothétie interne (paragraphe 3.2.1 et 3.2.2). Nous appliquons également le théorème A au recuit simulé ([46]) à basse température (3.2.3). Comme cas particulier, nous étudierons le gap de différentes dynamiques de processus d'Ising en dimension un, à basse température (3.2.4).

Le chapitre 4 est consacré à la démonstration et aux applications du théorème B. Nous donnons d'abord l'expression explicite de la somme des inverses des valeurs propres d'un générateur dont le graphe réduit est un arbre et ensuite la démonstration du théorème B (4.1). Dans les applications nous examinons d'abord le cas où tous les taux de transition non nuls sont égaux (4.2.1). A une transformation simple près, le générateur Λ est alors le laplacien de son graphe réduit ([17]). Le théorème généralise des résultats connus sur le spectre des graphes non orientés. Nous reprenons ensuite l'exemple du processus d'Ising en dimension un (4.2.2). Nous montrerons que la minoration déduite du théorème B, si elle est évidemment moins précise que l'équivalent calculé par le théorème A, fournit néanmoins le bon ordre de grandeur pour le gap.

Chapitre 2

Gaps des processus de Markov

On consacre ce chapitre à introduire les motivations, le contexte et les connaissances de base de ce travail. Notre objectif principal est d'étudier la deuxième valeur propre (trou spectral ou gap) des processus de Markov. Les notions et propositions concernées sont rappelées à la première partie. La deuxième partie consiste à présenter quelques applications des calculs de gap. Certaines méthodes d'étude du gap ainsi que du spectre sont introduites dans la troisième partie.

2.1 Générateurs de Markov et leurs gaps

On ne s'intéresse ici qu'aux processus de Markov irréductibles et réversibles dont les valeurs propres sont réelles et négatives ou nulles. Le gap du générateur qui mesure en quelque sorte la vitesse de convergence du processus est la valeur absolue de la valeur propre la plus proche de la valeur propre zéro. Le premier paragraphe 2.1.1 de cette partie contient un bref rappel des notions élémentaires, des propriétés nécessaires des processus de Markov. Au lieu des chaînes de

Markov sur lesquelles portent beaucoup de travaux étudiant le gap [20, 28, 38, 40], le point de vue qu'on adopte ici est celui des processus de Markov. Les raisons de cette préférence sont exposées au paragraphe 2.1.2.

2.1.1 Réversibilité, stationnarité et gap

Fixons d'abord quelques notations. Soit $M = (m_{ij})$ une matrice quelconque. Sa transposée sera notée M^* . On dit que M est *positive*, si tous les m_{ij} sont positifs ou nuls et il existe au moins un m_{ij} strictement positif. Et M est dite strictement positive, si tous les m_{ij} sont strictement positifs. Notons I la matrice identité dont l'ordre pourra changer selon le cas. Un vecteur $v = (v(i))_{1 \leq i \leq n}$ désignera un vecteur ligne d'ordre n . On note $\#E$ le cardinal d'un ensemble E quelconque.

Pour les définitions, propositions évoquées dans cette partie, nous donnons [14, 42, 43, 45, 62] comme références de base.

On s'intéresse à une famille de variables aléatoires $\{X(t); t \geq 0\}$ qui prennent leurs valeurs dans l'ensemble fini $E = \{i, j, \dots\}$ appelé l'*espace d'états*. On se restreint au cas où $\{X(t)\}$ est un processus de Markov avec des probabilités de transition homogènes. Ainsi, la probabilité de transition pour $t \geq 0$

$$p_{ij}(t) = \Pr \{ X(t+u) = j \mid X(u) = i \}, \forall i, j \in E$$

est indépendante de $u \geq 0$.

Il est en général plus naturel de donner à l'avance les taux de transition et de dériver à partir d'eux une expression explicite des probabilités de transition.

Les taux de transition sont les constantes λ_{ij} définies pour tous $i \neq j \in E$ tels que quand $h \searrow 0$,

$$(1) \quad p_{ij}(h) = \lambda_{ij}h + o(h),$$

$$(2) \quad p_{ii}(h) = 1 \Leftrightarrow \sum_{j \neq i} \lambda_{ij}h + o(h).$$

Le symbole $o(h)$ signifie que si on divise ce terme par h sa valeur tend vers zéro lorsque h tend vers zéro. L'ensemble des variables aléatoires $\{X(t); t \geq 0\}$ est

un *processus de Markov fini et homogène*. Appelons

$$\{P(t) = (p_{ij}(t)); t \geq 0\}$$

son *semi-groupe*. La constante λ_{ij} est le *taux de transition de i à j* qui mesure l'intensité du passage de i à j dans l'évolution aléatoire du processus.

Définition 2.1.1 Soit $\Lambda = (\lambda_{ij})_{i,j \in E}$ une matrice indicée par E . On dit que Λ est un *générateur de Markov sur E* si et seulement si

$$\lambda_{ij} \begin{cases} \geq 0 & \text{si } i \neq j \\ = \Leftrightarrow \sum_{j \neq i} \lambda_{ij} & \text{si } i = j . \end{cases}$$

Pour simplifier, on appelle un générateur de Markov aussi un générateur dans cette thèse. Soit $\Lambda = (\lambda_{ij})_{i,j \in E}$ un générateur de Markov. Un processus de Markov homogène est dit *de générateur Λ* s'il a λ_{ij} comme taux de transition. La proposition suivante nous dit qu'il existe un seul semi-groupe $\{P(t); t \geq 0\}$ tel que les deux postulats ci-dessus soient vérifiés (voir [42]).

Proposition 2.1.1 Soit $\Lambda = (\lambda_{ij})$ un *générateur de Markov*. Alors il n'existe qu'un *semi-groupe* $\{P(t); t \geq 0\}$ tel que

$$\frac{d}{dt}P(t) = \Lambda P(t) \quad \text{et} \quad P(0) = I$$

où I est la *matrice identité* et on a

$$P(t) = \exp(\Lambda t) .$$

En pratique, il est plus facile de connaître le générateur que le semi-groupe, et nous allons travailler directement sur les générateurs de Markov. On définit l'*horloge interne* λ_h du générateur Λ comme suit

$$\lambda_h = \Leftrightarrow \min_{i \in E} \lambda_{ii} .$$

Les générateurs qui nous intéressent sont irréductibles et réversibles.

Définition 2.1.2 Le *générateur de Markov* $\Lambda = (\lambda_{ij})_{i,j \in E}$ est dit *irréductible* si et seulement si $\forall i, j \in E, \exists l_1, \dots, l_n$ tels que

$$\lambda_{il_1} \lambda_{l_1 l_2} \cdots \lambda_{l_{n-1} l_n} \lambda_{l_n j} > 0 .$$

Définition 2.1.3 Soit π une mesure sur E . On dit que π est une mesure stationnaire du générateur de Markov Λ si

$$\pi\Lambda = 0 .$$

Nous avons la proposition suivante (voir [42]).

Proposition 2.1.2 Si Λ est un générateur irréductible, alors la mesure stationnaire de Λ est unique et strictement positive.

Définition 2.1.4 Soit π une mesure strictement positive sur E . On dit que Λ admet π comme mesure réversible, ou plus simplement que Λ est π -réversible, si la condition suivante, dite de bilan détaillé, est vérifiée.

$$\forall i, j \in E , \pi(i)\lambda_{ij} = \pi(j)\lambda_{ji} .$$

En sommant les équations ci-dessus par rapport à l'indice j , on voit immédiatement qu'une mesure réversible est nécessairement stationnaire. Elle est donc unique si le processus est irréductible. Remarquons que pour un processus réversible,

$$\forall i, j \quad \lambda_{ij} > 0 \iff \lambda_{ji} > 0 .$$

Il est légitime d'associer à un tel processus un graphe non orienté dont l'ensemble des sommets est E et l'ensemble des arêtes est

$$A = \{(i, j) ; \lambda_{ij} > 0\} .$$

Ce graphe sera appelé dans la suite *graphe réduit* du générateur Λ et noté $G = (E, A)$. Le processus étant supposé irréductible, son graphe réduit est nécessairement connexe. Dans le cas où $G = (E, A)$ est un arbre, si on fixe un sommet $a_0 \in E$ comme *racine*, un ordre partiel est défini comme suit. Quels que soient a et b dans E avec a différent de b , on dit que b est supérieur à a ou bien que a est antérieur à b et on note $b > a$ s'il existe un chemin sans répétition de b à a_0 passant par a . Dans ce cas là, on dit que G est orienté par rapport à la racine a_0 . Pour un sommet a différent de a_0 , il existe un sommet unique

b tel que $b < a$ et $(a, b) \in A$. On appelle b le *père* de a et a un *fil* de b . Un sommet est appelé *feuille de l'arbre*, si il n'a pas de fils.

Soit D la matrice diagonale

$$D = \text{diag} \left(\sqrt{\pi(i)} \right)_{i \in E} ,$$

et D^{-1} son inverse. Dire que Λ est π -réversible équivaut à dire que la matrice $D\Lambda D^{-1}$ est symétrique. Cela entraîne alors que Λ , comme $D\Lambda D^{-1}$ est diagonalisable, ses valeurs propres étant toutes réelles.

Munissons l'ensemble des fonctions de E dans \mathbb{R} du produit scalaire

$$\forall f, g \in \mathbb{R}^E, \langle f, g \rangle_\pi = \sum_{i \in E} f(i)g(i)\pi(i) .$$

Le générateur Λ est π -réversible si et seulement si il est auto-adjoint dans $L^2(E, \pi)$ muni de ce produit scalaire. Pour tout vecteur f de \mathbb{R}^E , nous noterons $D(f)$ la forme Dirichlet associée à Λ en f .

$$D(f) = \langle \Lambda f, f \rangle_\pi = \frac{1}{2} \sum_{i, j \in E} [f(i) - f(j)]^2 \pi(i)\lambda_{ij} .$$

Le spectre d'un générateur de Markov irréductible et réversible possède les propriétés suivantes (voir [62]).

Proposition 2.1.3 *Soit Λ un générateur irréductible et réversible. Alors les valeurs propres de Λ sont réelles, négatives ou nulles, et 0 est valeur propre simple.*

L'un de nos objectifs principaux est d'étudier le taux de convergence au sens L^2 des processus de Markov. Soit $\{P(t)\}$ le semi-groupe d'un processus de Markov de générateur Λ et de mesure stationnaire π . On dit que $P(t)$ *converge exponentiellement par rapport à la norme de l'espace $L^2(\pi)$* $\|\cdot\|_\pi$, s'il existe une valeur positive λ telle que pour tout f de \mathbb{R}^E ,

$$\|P(t)f - \pi(f)\|_\pi \leq \exp(-\lambda t) \|f - \pi(f)\|_\pi ,$$

où $\pi(f) = \sum_{i \in E} f(i)\pi(i)$. On appelle le maximum des λ vérifiant la condition ci-dessus le *taux de convergence exponentielle* de Λ qu'on va noter λ^* .

Dans le spectre d'un générateur de Markov irréductible et réversible, ce qui nous intéresse le plus est la valeur propre correspondant au taux de convergence exponentielle.

Définition 2.1.5 *Soit Λ un générateur de Markov irréductible et réversible. Nous appelons trou spectral ou gap de Λ l'opposé de la deuxième valeur propre de Λ par ordre décroissant.*

Le théorème suivant montre que le taux de convergence au sens L^2 n'est rien d'autre que le gap du générateur et donne une formule variationnelle du gap au travers de la forme de Dirichlet. Par conséquent, contrôler la vitesse de convergence revient à estimer le gap. Pour la démonstration de la proposition suivante, voir [9].

Proposition 2.1.4 *Si Λ est un générateur de Markov irréductible et de la mesure π réversible, alors*

$$\lambda^* = \text{gap} = \inf \{ D(f) ; \pi(f) = 0, \|f\|_\pi^2 = 1 \} . \quad (2.1)$$

C'est-à-dire, on a

$$\|P(t)f \ominus \pi(f)\|_\pi \leq \exp(\ominus \text{gap} \cdot t) \|f \ominus \pi(f)\|_\pi .$$

La convergence est d'autant plus rapide que le gap est grand.

Bien que la réversibilité de Λ soit une hypothèse très restrictive par rapport au cas général des mesures stationnaires non réversibles, c'est sous cette condition que se placent la plupart des références du domaine [20, 23, 28, 70]. La raison en est double. D'une part le cas réversible est beaucoup plus facile que le cas général, du fait de la possibilité de se ramener à des matrices symétriques. La seconde raison tient au fait que l'on peut contrôler la vitesse d'accès à l'équilibre d'un processus de Markov quelconque par le gap d'un générateur réversible (cf. [25, 61]).

Proposition 2.1.5 *Soit $\{X_t; t \geq 0\}$ un processus de Markov irréductible de générateur Λ sur E . Soit π la mesure stationnaire de Λ , D^2 la matrice diagonale*

$$D^2 = \text{diag}(\pi(i))_{i \in E} ,$$

et D^{-2} son inverse. Soit M le générateur

$$M = \frac{1}{2}(\Lambda + D^{-2} \Lambda^* D^2) \quad .$$

Alors pour toute fonction f de E dans \mathbb{R} , il existe une constante C , telle que

$$|\mathbf{E}[f(X_t)] - \pi(f)| \leq C \exp[-\text{gap}(M)t] \quad .$$

L'interprétation du générateur M est la suivante. Supposons que $\{X_t; t \geq 0\}$ soit stationnaire (pour tout t , X_t est de loi π). Fixons $t_0 > 0$. Le processus $\{X_{t_0-t}; 0 \leq t \leq t_0\}$ est markovien et homogène, de générateur $D^{-2} \Lambda^* D^2$. La mesure stationnaire de ce nouveau générateur est encore π . Passer de Λ à $D^{-2} \Lambda^* D^2$ revient à retourner l'échelle de temps. Un processus de générateur M évolue comme suit. Chaque fois qu'il atteint un nouvel état, il choisit en tirant à pile ou face si le pas suivant se fera vers le futur ou vers le passé d'un processus de générateur Λ qui serait dans le même état. Il n'est donc pas surprenant que M soit π -réversible, ce que l'on vérifie immédiatement en constatant que DMD^{-1} est symétrique.

Un autre avantage de la réversibilité, que nous serons amenés à utiliser, est de se conserver par troncature du générateur.

Définition 2.1.6 Soit $\Lambda = (\lambda_{ij})_{i,j \in E}$ un générateur de Markov sur E . Soit E_1 un sous-ensemble de E . Le générateur restreint de Λ sur E_1 qu'on note $\Lambda_1 = (\lambda_{ij}^1)_{i,j \in E_1}$ est un générateur de Markov défini sur E_1 tel que pour tous $i \neq j \notin E_1$,

$$\lambda_{ij}^1 = \lambda_{ij} \quad .$$

Proposition 2.1.6 Soit Λ un générateur de Markov irréductible et π -réversible sur E . Soit E_1 un sous-ensemble de E . Soit π_{E_1} la mesure de probabilité sur E_1 telle que pour tout i de E_1 ,

$$\pi_{E_1}(i) = \frac{\pi(i)}{\pi(E_1)}$$

où $\pi(E_1)$ désigne la probabilité de l'ensemble E_1 . Posons Λ_1 le générateur restreint de Λ sur E_1 . Alors Λ_1 est π_{E_1} -réversible.

Ce résultat est un cas particulier du corollaire 1.10 p. 26 de Kelly [45].

2.1.2 Processus et chaînes

Une des motivations importantes dans l'étude du gap vient des méthodes de Monte Carlo par chaînes de Markov (MCMC) qui visent à simuler une chaîne de Markov admettant une distribution cible comme mesure stationnaire sur un espace d'états de grande taille (cf. 2.2). La vitesse de convergence de la chaîne de Markov simulée est liée à la valeur propre dont le module est plus proche de la valeur propre triviale 1 de la matrice de transition de la chaîne.

D'une façon précise, une chaîne de Markov irréductible, π -réversible et apériodique de matrice de transition P admet une valeur propre simple 1, les autres étant réelles et comprises strictement entre $\Leftrightarrow 1$ et 1. Notons $\rho(P)$ le maximum des valeurs propres en valeur absolue différentes de 1,

$$\rho(P) = \max\{|\alpha| ; \alpha \neq 1 \text{ valeur propre de } P\} \quad .$$

Il est connu que $\rho(P)$ contrôle la vitesse de convergence de la chaîne au sens suivant ([62])

$$\|P^k(i, \cdot) \Leftrightarrow \pi(\cdot)\|_{\text{TV}} \leq \frac{1}{2} \sqrt{\frac{1}{\pi(i)}} \rho(P)^k$$

où $P^k(i, \cdot)$ est la distribution de la chaîne à l'étape k en partant de l'état i et $\|\cdot\|_{\text{TV}}$ est la norme de variation totale.

Tout algorithme itératif est une chaîne de Markov. Le temps continu est une approximation commode mais a peu d'intérêt apparemment pour des méthodes algorithmiques. D'autre part, l'étude des valeurs propres d'une matrice de transition est de difficulté comparable à ce qu'elle est pour un générateur. Pourquoi donc préférer les processus aux chaînes? Les raisons sont les suivantes.

En fait, le passage des matrices de transition aux générateurs de Markov, ou des chaînes à temps discret aux processus à temps continu est purement formel. Même si un algorithme de simulation est nécessairement présenté comme une chaîne de Markov (temps discret), il est souvent commode de considérer un processus de Markov (temps continu) qui en est la version harmonisée au sens suivant. Considérons une chaîne de Markov $\{X_n ; n \in \mathbb{N}\}$ de matrice de transition P et un processus de Poisson $\{N_t ; t \geq 0\}$ d'intensité λ . Posons, pour

tout $t \geq 0$,

$$Z_t = X_{N_t} \quad .$$

Ceci revient à remplacer des intervalles de temps fixes (de longueur 1) par des intervalles de temps aléatoires indépendants, dont la longueur suit la loi exponentielle de paramètre λ . Le processus $\{Z_t ; t \geq 0\}$ est un processus de Markov de générateur

$$\Lambda = \lambda(P \Leftrightarrow I) \quad .$$

Réciproquement, si Λ est le générateur d'un processus de Markov, il peut toujours être vu comme une chaîne harmonisée de matrice de transition

$$P = I + \frac{1}{\lambda} \Lambda$$

où λ est supérieur à l'horloge interne λ_h de Λ .

On constate que Λ et P définis ainsi ont les mêmes vecteurs propres et le spectre de l'un se déduit aisément de celui de l'autre. On suppose que l'espace d'états est de taille $n + 1$. En ordonnant les spectres de Λ et P , on pose que leurs valeurs propres sont respectivement

$$0 = \lambda_0 > \lambda_1 \geq \dots \geq \lambda_n$$

et

$$1 = \alpha_0 > \alpha_1 \geq \dots \geq \alpha_n > \Leftrightarrow 1 \quad .$$

On a donc

$$\text{gap}(\Lambda) = \lambda_1 \quad \text{et} \quad \rho(P) = |\alpha_n| \vee \alpha_1 \quad .$$

D'un point de vue technique, l'étude du gap d'un processus, étant tout simplement la deuxième valeur propre, est plus simple que celle de $\rho(P)$ d'une chaîne qui nécessite la connaissance de tout le spectre. Mais si toutes les valeurs propres de P sont positives, $\rho(P)$ est donc la deuxième valeur propre α_1 de P , $\text{gap}(\Lambda)$ et $\rho(P)$ sont liés par:

$$\text{gap}(\Lambda) = \lambda [1 \Leftrightarrow \rho(P)] \quad . \tag{2.2}$$

En fait, les valeurs propres négatives d'une chaîne de Markov correspondent à l'oscillation ou bien au comportement quasi périodique qui ne peut pas avoir lieu, si tout état i a une probabilité de retour p_{ii} assez grande. En particulier, toutes les valeurs propres de la chaîne sont positives dès que

$$\min_i p_{ii} \geq \frac{1}{2} .$$

Pour voir cela, on considère la matrice de transition $2P \Leftrightarrow I$ dont les valeurs propres sont $\beta_i = 2\alpha_i \Leftrightarrow 1$. Le fait $\beta_i \geq \Leftrightarrow 1$ implique $\alpha_i \geq 0$ pour tout $0 \leq i \leq n$.

En fait, les valeurs négatives ne sont pas vraiment gênantes. Pour une matrice de transition P quelconque, on peut modifier la chaîne de façon simple de sorte que la condition ci-dessus soit vérifiée sans ralentir trop la convergence. Considérons la nouvelle chaîne

$$P' = \frac{1}{2}(I + P) .$$

Alors P' , irréductible et π -réversible, a toutes ses valeurs propres positives. De toutes façons, tous les résultats portant sur le gap des chaînes s'adaptent à celui des processus et inversement si on se ramène à (2.2).

Par ailleurs, nous étudions les processus au lieu des chaînes, compte tenu de la plus grande facilité de manipulation d'un processus. Libérés de la contrainte $\sum_j p_{ij} = 1$, les générateurs sont plus faciles à modifier. Si on fait varier un taux de transition, deux termes du générateur changent. Mais il faut modifier toute une ligne de la matrice de transition de la chaîne harmonisée. Mais pour les processus qui n'ont souvent pas la même horloge interne, comparer les vitesses de convergence entre deux générateurs Λ_1 et Λ_2 ne revient pas à comparer $\text{gap}(\Lambda_1)$ et $\text{gap}(\Lambda_2)$ comme dans le cas des chaînes dont l'horloge de référence est toujours 1. Afin de comparer de façon significative, définissons une référence par le rapport

$$\text{gap}(\Lambda)/\lambda_h .$$

Ceci explique pourquoi λ_h est nommé l'horloge interne. On pourra dire que le processus de générateur Λ_1 est meilleur que celui de générateur Λ_2 au sens où

$\text{gap}(\Lambda_1)/\lambda_h(\Lambda_1) \geq \text{gap}(\Lambda_2)/\lambda_h(\Lambda_2)$ (la vitesse relative d'accès à l'équilibre du premier est supérieure à celle du second).

2.2 Exemples d'applications du gap

De nombreux travaux ont portés ces dernières années sur l'estimation du gap. Les motivations de ces études sont des domaines multiples: en probabilités ([1, 20, 39, 47]), en statistiques ([4, 32, 72]), en physique mathématique ([55, 64]), en géométrie ([6, 10, 49, 63]), en théorie des graphes ([13, 17]), en informatique ([40, 68, 41]) etc.

Une application fréquente porte sur les méthodes de Monte-Carlo par chaînes de Markov (méthodes MCMC). Ces dernières années sont apparues, par exemple en analyse d'image [31, 34] et en statistique bayésienne [30], de nombreuses méthodes nécessitant des échantillonnages aléatoires sur des espaces discrets de grande taille. Mais la difficulté voire l'impossibilité d'échantillonner directement en calculant la probabilité de chaque état se rencontre à cause de la taille de l'espace d'états. On fait alors appel aux méthodes MCMC (Monte-Carlo Markov Chains) qui visent à simuler une chaîne de Markov admettant la distribution désirée comme mesure stationnaire sur l'espace de configurations. Or la vitesse de convergence vers le régime stationnaire est contrôlée par $\rho(P)$ de la chaîne ayant P comme matrice de transition.

Un autre sujet important est celui des transitions de phase en physique statistique. Considérons en particulier le modèle d'Ising sans champ magnétique externe dont le générateur est fonction de la température T . Le système est L^2 exponentiellement ergodique i.e. $\text{gap}(\Lambda_T) > 0$ si T est supérieure à la température critique T_c . Par contre, la transition de phase a lieu si T est inférieure à T_c , dans ce cas là, $\text{gap}(\Lambda_T) = 0$. Ainsi, l'étude du gap fournit une description de la transition de phase.

2.2.1 Algorithme de Metropolis et échantillonnage de Gibbs

Les algorithmes le plus souvent utilisés par les méthodes MCMC sont l'algorithme de Metropolis introduit par Metropolis et *al.* [56] (1953) et l'échantillonnage de Gibbs par Glauber [33] (1963). Ces deux algorithmes sont largement utilisés en pratique en restauration d'image [3, 31, 32], en optimisation stochastique (recuit simulé) [2, 46] et en mécanique statistique [50, 71]. Ici, on va les présenter dans le cadre suivant.

Soit S un ensemble fini de sites muni d'une structure de graphe G . Chaque site i de S correspond à un ensemble fini Ω_i . L'espace de configurations Ω est le produit des Ω_i , $\otimes_{i \in S} \Omega_i$. Le système d'intérêt est constitué de l'espace d'états Ω , la fonction d'énergie H sur Ω et le paramètre de température T . La fonction d'énergie H peut souvent s'écrire comme une somme de potentiels qui ne dépendent que des cliques correspondant à la structure de voisinage déduite du graphe G . La mesure dont on voudrait des échantillonnages est la mesure de Gibbs de paramètre T , notée π_T , déduite de la fonction d'énergie H par la loi de Boltzmann,

$$\forall \eta \in \Omega, \pi_T(\eta) = \frac{1}{Z_T} \exp \left[\frac{H(\eta)}{T} \right]$$

où Z_T est la constante de normalisation définie par

$$Z_T = \sum_{\eta \in \Omega} \exp \left[\frac{H(\eta)}{T} \right].$$

Quand la température tend vers 0, la famille de mesures ainsi définie converge vers la loi uniforme sur l'ensemble des minimas globaux de la fonction H . (voir [76]).

Proposition 2.2.1 *Quand la température T tend vers 0, la mesure de Gibbs π_T tend vers la mesure uniforme sur le sous-ensemble des configurations d'énergie minimale.*

Démonstration. Soit Ω' l'ensemble des configurations d'énergie minimale.

Evidemment, pour toute configuration η de $\Omega \Leftrightarrow \Omega'$,

$$\lim_{T \rightarrow 0} \pi_T(\eta) = 0 .$$

Or pour toute configuration η de Ω' ,

$$\begin{aligned} \lim_{T \rightarrow 0} \pi_T(\eta) &= \lim_{T \rightarrow 0} \frac{\exp \left[\frac{H(\eta)}{T} \right]}{\sum_{\xi \in \Omega} \exp \left[\frac{H(\xi)}{T} \right]} \\ &= \lim_{T \rightarrow 0} \left\{ \frac{1}{\#\Omega'} + \sum_{\xi \in \Omega \setminus \Omega'} \exp \left[\frac{H(\xi) - H(\eta)}{T} \right] \right\}^{-1} \\ &= \frac{1}{\#\Omega'} . \end{aligned}$$

□

L'algorithme de Metropolis consiste à simuler une chaîne de Markov ergodique tendant vers π_T dont la probabilité de transition ne dépend que de la différence d'énergie entre deux états successifs. Quand la probabilité de transition est remplacée par la caractéristique locale de la mesure de Gibbs (la probabilité conditionnelle d'un site sachant les autres), on obtient l'échantillonnage de Gibbs. Pour une description plus générale, voir [60]. Une adaptation se fera des chaînes aux processus. Si la chaîne de Markov simulée est de matrice de transition P , ce qu'on va donner ici sera le processus de Markov de générateur $P \Leftrightarrow I$.

L'algorithme de Metropolis repose sur l'utilisation d'un générateur $\Lambda' = (\lambda'_{\eta\xi})$ irréductible tel que

$$\lambda'_{\eta\xi} > 0 \iff \lambda'_{\xi\eta} > 0 .$$

Il produit un processus de Markov de générateur $\Lambda = (\lambda_{\eta\xi})$ tel que, pour $\eta \neq \xi$,

$$\lambda_{\eta\xi} = \begin{cases} 1 & \text{si } \frac{\pi(\xi)\lambda'_{\xi\eta}}{\pi(\eta)\lambda'_{\eta\xi}} \geq 1 \\ \frac{\pi(\xi)\lambda'_{\xi\eta}}{\pi(\eta)} & \text{sinon.} \end{cases}$$

L'irréductibilité de Λ se déduit de celle de Λ' et la réversibilité de Λ par rapport à π_T se vérifie facilement. Le processus de Markov de générateur Λ converge donc en loi vers π_T .

En pratique, la fonction d'énergie H est facile à calculer, par contre la constante de normalisation Z_T est difficile à obtenir à cause de la grande taille de l'espace de configurations. L'avantage principal de l'algorithme est d'éviter le calcul de la constante Z_T en ne faisant intervenir que les rapports $\frac{\pi_T(\xi)}{\pi_T(\eta)}$ qui dépendent seulement de la différence entre $H(\eta)$ et $H(\xi)$. Un autre avantage est de ne requérir qu'une connaissance limitée de la loi π_T en proposant une infinité de possibilités de générateurs Λ .

Un algorithme de Metropolis n'est pas nécessairement multidimensionnel comme présenté ici, par contre l'échantillonnage de Gibbs est toujours sur un espace de configurations. Pour un site x fixé, soit $\Omega(\eta, x)$ l'ensemble des configurations ξ qui ne sont différentes de η qu'au site x , c'est-à-dire

$$\forall y \neq x, \eta(y) = \xi(y) \text{ et } \eta(x) \neq \xi(x) .$$

Pour tous $\eta \neq \xi$, posons

$$\lambda_{\eta\xi} = \begin{cases} \exp \left[\frac{H(\xi) - H(\eta)}{T} \right] / \sum_{\zeta \in \Omega(\eta, x)} \exp \left[\frac{H(\zeta) - H(\eta)}{T} \right] & \text{si } \exists x \in S \text{ t.q. } \xi \in \Omega(\eta, x) \\ 0 & \text{sinon .} \end{cases}$$

Alors le générateur déduit de l'échantillonnage de Gibbs est $\Lambda = (\lambda_{\eta\xi})$. On vérifie que Λ est irréductible et $\pi_T \leftrightarrow$ réversible.

2.2.2 Modèle d'Ising

En physique statistique, en statistique bayésiennes, le modèle d'Ising apparaît souvent. En adoptant les notations du paragraphe précédent, pour tout site i de S , on identifie l'ensemble Ω_i à $\{0, 1\}$. Chaque configuration η de l'espace $\Omega = \{0, 1\}^S$ est une application de S dans $\{0, 1\}$ (à chaque site x , on associe un état binaire $\eta(x) \in \{0, 1\}$). Un *système de spin* est un processus de Markov sur l'espace d'états Ω , qui à chaque transition modifie la configuration courante en

au plus un site. Pour toute configuration η de Ω , et tout site x de S , on note η_x la configuration obtenue en modifiant η au site x

$$\eta_x(y) = \begin{cases} \eta(y) & \text{si } y \neq x \\ 1 \Leftrightarrow \eta(x) & \text{si } y = x . \end{cases}$$

Les taux d'un système de spin sont donc, pour $\eta \neq \xi \in \Omega$,

$$\lambda_{\eta\xi} = \begin{cases} c(x, \eta) & \text{si } \xi = \eta_x \\ 0 & \text{sinon} . \end{cases}$$

On appelle *processus d'Ising* tout système de spin qui admet la mesure de Gibbs π_T comme mesure réversible. Il suffit qu'il vérifie la condition de bilan détaillé, pour toute configuration η de Ω et tout site x de S .

$$c(x, \eta) \exp \left[\frac{H(\eta)}{T} \right] = c(x, \eta_x) \exp \left[\frac{H(\eta_x)}{T} \right] .$$

De nombreux générateurs vérifient cette équation. Voici quelques exemples.

Exemple 2.2.1

$$c(x, \eta) = \exp \left[\frac{H(\eta) \Leftrightarrow H(\eta_x)}{2T} \right]$$

Les rôles de 0 et 1 sont symétriques. C'est le choix le plus naturel du point de vue mathématique [52].

Exemple 2.2.2

$$c(x, \eta) = \begin{cases} 1 & \text{si } \eta(x) = 1 \\ \exp \left[\frac{H(\eta_x) \Leftrightarrow H(\eta)}{T} \right] & \text{sinon} . \end{cases}$$

Il s'agit d'un modèle de déposition de particules [77].

Exemple 2.2.3

$$c(x, \eta) = \begin{cases} 1 & \text{si } H(\eta_x) \leq H(\eta) \\ \exp \left[\frac{H(\eta_x) \Leftrightarrow H(\eta)}{T} \right] & \text{sinon} . \end{cases}$$

C'est le générateur de Metropolis en choisissant

$$\lambda'_{\eta\xi} = \begin{cases} 1 & \text{si } \exists x \in S \text{ t.q. } \xi = \eta_x \\ 0 & \text{sinon} \end{cases}$$

pour le générateur Λ' .

Exemple 2.2.4

$$c(x, \eta) = \left\{ 1 + \exp \left[\frac{H(\eta_x) \Leftrightarrow H(\eta)}{T} \right] \right\}^{-1}.$$

C'est le générateur de l'échantillonnage de Gibbs ou dynamique de Glauber.

Ici, on s'intéresse au cas où l'ensemble de sites S est une partie finie de \mathbb{Z}^d . Pour tout $x \in \mathbb{Z}^d$, notons $x(i)$ la i -ème coordonnée de x avec $1 \leq i \leq d$. La norme de x est

$$\|x\| = \sum_{i=1}^d |x(i)|.$$

On munit \mathbb{Z}^d de la structure de graphe dont les arêtes sont les paires (x, y) telles que $\|x \Leftrightarrow y\| = 1$. On pose

$$S = \{x; 0 \leq x(i) \leq n \Leftrightarrow 1, 1 \leq i \leq d\}.$$

La fonction d'énergie H dépend de la structure de graphe ainsi que la condition de bord de S . Nous allons considérer trois cas de conditions de bord : périodique, ouvert et unitaire. Ils seront distingués par les indices p , o et u respectivement. Nous identifions le graphe G sur S au graphe restreint de \mathbb{Z}^d à S , c'est-à-dire

$$\forall x, y \in S, (x, y) \in G \text{ si } \|x \Leftrightarrow y\| = 1.$$

Définissons le bord de S par:

$$\partial S = \{x \in S; \exists y \notin S \text{ t.q. } \|x \Leftrightarrow y\| = 1\}.$$

Pour la condition de bord périodique, définissons la structure de graphe G^p sur S par

$$G^p = G \cup \{(x, y); x, y \in \partial S, \|(x \Leftrightarrow y) \text{ modulo } n\| = 1\}.$$

Définissons la fonction d'énergie par

$$H^p(\eta) = \Leftrightarrow \sum_{(x, y) \in G^p} [2\eta(x) \Leftrightarrow 1] [2\eta(y) \Leftrightarrow 1]$$

qui mesure la différence entre le nombre de paires de voisins de valeur différente et le nombre de paires de voisins de même valeur.

Pour les deux autres conditions de bord, utilisons la notation τ pour désigner un élément générique de l'ensemble $\Omega_\infty = \{0, 1\}^{\mathbb{Z}}$. Quelle que soit la configuration η de Ω , définissons l'énergie $H^{U,\tau}(\eta)$ avec conditions de bord τ dehors de S et couplage de bord $0 \leq U(x, y) \leq 1$ avec $x \in \partial S$, $y \notin S$ et $\|x \leftrightarrow y\| = 1$, par

$$\begin{aligned} H^{U,\tau}(\eta) &= \Leftrightarrow \sum_{(x,y) \in G} [2\eta(x) \leftrightarrow 1] [2\eta(y) \leftrightarrow 1] \\ &\Leftrightarrow \sum_{\substack{x \in \partial S, y \notin S \\ \|x-y\|=1}} U(x, y) [2\eta(x) \leftrightarrow 1] [2\tau(y) \leftrightarrow 1] \end{aligned}$$

On dit que la condition de bord est ouverte si $U \equiv 0$ et unitaire si $U \equiv 1$ et τ est la configuration constante valant 1 en tout site. C'est-à-dire

$$H^o(\eta) = \Leftrightarrow \sum_{(x,y) \in G} [2\eta(x) \leftrightarrow 1] [2\eta(y) \leftrightarrow 1]$$

et

$$H^u(\eta) = \Leftrightarrow \sum_{(x,y) \in G} [2\eta(x) \leftrightarrow 1] [2\eta(y) \leftrightarrow 1] \Leftrightarrow \sum_{x \in \partial S} [2\eta(x) \leftrightarrow 1] \quad .$$

Quand l'ensemble de sites est fini, un processus d'Ising a toujours une mesure stationnaire unique. Si on remplace l'ensemble de sites S par \mathbb{Z}^d , on obtient le modèle d'Ising bien connu en physique statistique. Maintenant, un phénomène tout à fait différent apparaît. Pour $d = 2$, pour une température

$$T > T_c^{(2)} = \frac{2}{\log \frac{1}{1+\sqrt{2}}} \approx 2,27,$$

la mesure stationnaire est unique. Si $T < T_c^{(2)}$, il y a au moins deux mesures stationnaires. Pour $d \geq 3$, nous avons la conclusion analogue pour une température critique $T_c^{(d)} > 0$ dont la valeur exacte reste inconnue. Ce phénomène s'appelle *transition de phase*. Mais pour $d = 1$, il n'y a pas de transition de phase, la mesure stationnaire est unique.

Une autre description de la transition de phase se fait par le gap. Si la transition de phase a lieu, $\text{gap} = 0$, sinon, $\text{gap} > 0$. Il est naturel de demander comment cette nullité du gap se reflète si on considère le processus d'Ising sur un carré S fini mais très grand dont le gap est toujours strictement positif. De nombreux travaux portent sur cette question: [53, 64, 65, 66].

2.3 Méthodes principales d'étude du gap

Les outils pour l'étude du gap ainsi que du spectre sont essentiellement de trois types: analytiques, géométriques et probabilistes. Regardant le générateur Λ et le semi-groupe $\{P(t); t \geq 0\}$ d'un processus de Markov de mesure stationnaire π comme opérateurs sur l'espace $L^2(\pi)$, associées à la forme de Dirichlet, les inégalités analytiques comme les inégalités de Poincaré, de Sobolev, de Nash et de Sobolev logarithmique ont été utilisées pour l'étude des chaînes ou des processus de Markov. D'ailleurs, beaucoup de résultats ont été obtenus récemment à partir des quantités géométriques et la combinaison des chemins du graphe déduit du générateur du processus de Markov. Une présentation précise de ces méthodes est donnée dans le cours de Saloff-Coste [62]. D'un point de vue probabiliste qui est assez différent des précédents, la méthode de couplage est aussi utilisée avec succès dans les études de gap. Donnons quelques références récentes de cette méthode: [50], [9], [11]. Nous nous contenterons de décrire brièvement quelques unes des méthodes évoquées ci-dessus.

2.3.1 Techniques de comparaison

Dans [20] et [23], est développée une méthode pour estimer toutes les valeurs propres d'une chaîne de Markov en comparant sa forme de Dirichlet avec celle d'une autre dont on suppose connaître le spectre. Cette méthode est aussi valable pour les processus de Markov. De façon plus précise, soit Λ et $\tilde{\Lambda}$ deux processus irréductibles sur un espace d'états de taille $n + 1$. Soit π et $\tilde{\pi}$ les mesures réversibles respectivement de Λ et $\tilde{\Lambda}$. Notons D et \tilde{D} les formes de Dirichlet associées. Posons que les valeurs propres de Λ et $\tilde{\Lambda}$ sont respectivement

$$0 = \lambda_0 > \lambda_1 \geq \dots \geq \lambda_n$$

et

$$0 = \tilde{\lambda}_0 > \tilde{\lambda}_1 \geq \dots \geq \tilde{\lambda}_n .$$

La caractérisation minimax des valeurs propres donne les bornes des valeurs

propres de Λ en terme de celles de $\tilde{\Lambda}$ (voir [36]). Pour $0 \leq i \leq n$, si

$$\tilde{D} \leq AD \text{ et } \tilde{\pi} \geq a\pi ,$$

alors

$$\lambda_i \geq \frac{a}{A} \tilde{\lambda}_i .$$

Notons G, \tilde{G} les graphes réduits de Λ et $\tilde{\Lambda}$ respectivement. Soit $\tilde{\gamma}_{kl}$ un chemin dans \tilde{G} liant k et l . Pour toute paire (i, j) de G , $\tilde{G}(i, j)$ désigne l'ensemble des chemins $\tilde{\gamma}_{kl}$ dans \tilde{G} contenant l'arête (i, j) . On note $|\tilde{\gamma}_{kl}|$ la longueur du chemin $\tilde{\gamma}_{kl}$. Alors une version de la constante A est donnée dans [20].

Théorème 2.3.1 (P. Diaconis et L. Saloff-Coste, 1993) *Posons*

$$A = \max_{(i,j) \in G} \left\{ \frac{1}{\pi(i)\lambda_{ij}} \sum_{\tilde{G}(i,j)} |\tilde{\gamma}_{kl}| \tilde{\pi}(k) \tilde{\lambda}_{kl} \right\} .$$

Alors

$$\tilde{D} \leq AD .$$

De nombreuses applications du théorème ci-dessus sont étudiées par Diaconis et Saloff-Coste: [20, 21, 62].

2.3.2 Différentes versions de l'inégalité de Cheeger

En montrant une borne inférieure de la deuxième valeur propre du laplacien sur une variété Riemannienne lisse et compacte, Cheeger [7] a introduit une constante isopérimétrique. Cette quantité géométrique a inspiré beaucoup de bornes inférieures ou supérieures du gap voire de tout le spectre dans notre cadre. On va présenter ici certains résultats qui se servent de cette constante en les adaptant aux processus de Markov si nécessaire.

Soit Λ un générateur de Markov irréductible et π -réversible défini sur l'espace d'états E . Adoptons les notions présentées dans [26]. Pour tout sous-ensemble E_1 de E , définissons selon Sinclair (voir [69]) la *conductance* de E_1 *normalisée*

par

$$\phi(E_1) = \frac{1}{\pi(E_1)} \sum_{\substack{i \in E_1 \\ j \in E - E_1}} \pi(i)\lambda_{ij} .$$

On l'interprète comme le flux probabiliste de sortie de E_1 dans le régime stationnaire. La *conductance symétrique de E_1 normalisée* est définie selon Lawler et Sokal (voir [47]) par

$$\phi^*(E_1) = \frac{\phi(E_1)}{1 \leftrightarrow \pi(E_1)}$$

qui du fait de la réversibilité est égal à $\phi^*(E \setminus E_1)$. On l'interprète ainsi comme le flux probabiliste d'échange entre E_1 et son complémentaire dans le régime stationnaire. Définissons enfin la *conductance globale* ou bien *constante de Cheeger* par

$$\Phi = \min_{\substack{E_1 \subset E \\ \pi(E_1) \leq 1/2}} \phi(E_1)$$

et la *conductance globale symétrique* par

$$\Phi^* = \min_{E_1 \subset E} \phi^*(E_1) .$$

Intuitivement, si le processus sort rapidement de tous les “petits” ensembles $E_1 \subset E$ avec $\pi(E_1) \leq 1/2$, autrement dit si la conductance globale est grande, la convergence doit nécessairement être rapide. Cette intuition est formalisée par Sinclair et Jerrum (cf. [69]):

Théorème 2.3.2 (A. Sinclair et M. Jerrum, 1989) .

$$\frac{\Phi^2}{2\lambda_h} \leq \text{gap}(\Lambda) .$$

Par une intuition analogue, si les flux d'échanges entre tous les sous-ensembles complémentaires sont rapides, ce qui correspond à une grande conductance globale symétrique, Lawler et Sokal (cf. [47]) affirment que le taux de convergence est nécessairement grand :

Théorème 2.3.3 (G.F. Lawler et A.D. Sokal, 1988) .

$$\frac{\Phi^{*2}}{8\lambda_h} \leq \text{gap}(\Lambda) .$$

Inversement, plus le flux d'échange entre E_1 et son complémentaire est faible, plus le processus s'attardera longtemps dans un des deux sous-ensembles avant de visiter l'autre. L'accès à l'équilibre en sera ralenti d'autant et le gap du processus sera plus faible. Dans [23], Diaconis et Stroock ont montré le résultat suivant

$$\text{gap}(\Lambda) \leq \Phi^* .$$

L'estimation du gap de Lawler et Sokal est récemment étendue à toutes les valeurs propres du générateur par O.François [26] dans le cas où la mesure réversible du générateur est assez concentrée sur un état (le cas du recuit simulé à basse température par exemple). On adapte le résultat de François des chaînes aux processus de la manière suivante.

Théorème 2.3.4 (O. François, 1997) *Soit π la mesure réversible du générateur Λ . Supposons qu'il existe un état i tel que*

$$\pi(i) > v^2$$

avec

$$v = \left(\frac{19}{27} + \frac{\sqrt{33}}{9} \right)^{\frac{1}{3}} + \frac{4}{9} \left(\frac{19}{27} + \frac{\sqrt{33}}{9} \right)^{-\frac{1}{3}} \Leftrightarrow \frac{2}{3} \approx 0,701 .$$

Alors pour toute valeur propre $\lambda \neq 0$ de Λ , on a

$$\Leftrightarrow 1 \Leftrightarrow \sqrt{1 \Leftrightarrow x^2} \leq \lambda \leq \Leftrightarrow 1 + \sqrt{1 \Leftrightarrow x^2}$$

où

$$x = \left\{ 1 \Leftrightarrow \sqrt{1 \Leftrightarrow \pi(i)} \left[1 + \sqrt{\pi(i)} \right] \right\} \Phi^* .$$

2.3.3 Etude de comportement asymptotique spectral

Dans de nombreux cas, le générateur de Markov irréductible d'intérêt peut être vu comme une perturbation d'un générateur réductible. C'est le cas par exemple pour les processus d'Ising et le recuit simulé à basse température. C'est aussi le cadre où on se place pour le théorème A qui porte sur le comportement

asymptotique des valeurs propres. Les travaux sur le comportement asymptotique spectral des générateurs de ce type sont nombreux: [44, 27, 5, 57]. L'idée d'étudier le comportement asymptotique au travers de la structure hiérarchique de l'espace d'états liée naturellement au générateur est introduite par Freidlin-Wentzell [27]. Leur technique a inspiré par suite beaucoup de travaux dans de nombreux domaines: [5, 54, 57, 8]. Avec la même idée de base, Miclo [57] a considéré le recuit simulé à basse température et étendu le résultat de Holley, Kusuoka et Stroock ([35]) sur le gap en obtenant le comportement asymptotique de toutes les valeurs propres ainsi que les projections spectrales. De façon précise, il considère sur un graphe fini G , le générateur Λ , à température T , du recuit simulé associé à une fonction d'énergie H . Soit

$$0 = \lambda_0 > \lambda_1 \geq \dots \geq \lambda_{n-1}$$

ses valeurs propres qui sont fonctions de T . Il montre que pour tout $1 \leq i \leq n-1$, il existe des constantes $b_i \geq a_i > 0$ (ne dépendant que de G) et $c_i(G, H) \geq 0$, telles que, pour tout $T \geq 0$,

$$a_i \exp \left[\frac{c_i(G, H)}{T} \right] \leq \lambda_i \leq b_i \exp \left[\frac{c_i(G, H)}{T} \right],$$

et il a donné une description géométrique de $c_i(G, H)$. De plus, il a obtenu la convergence de la somme des projections spectrales de

$$\text{diag} \left(\exp \frac{H(i)}{2T} \right)_{0 \leq i \leq n-1} \Lambda \text{diag} \left(\exp \frac{H(i)}{2T} \right)_{0 \leq i \leq n-1},$$

l'opérateur de Schrödinger associé à Λ , relatives aux valeurs propres λ_i . La description des $c_i(G, H)$ est longue, voir [57].

2.3.4 Méthode de couplage

La méthode de couplage est beaucoup moins utilisée pour estimer le gap que les autres méthodes évoquées ci-dessus. L'objectif de ce paragraphe est d'illustrer l'efficacité et la puissance potentielle de cette méthode en présentant certains travaux récents de Chen ([9, 11, 12]) sur cette méthode et le gap spectral. Au

travers de la méthode de couplage, Chen a obtenu des bornes fines voire des formules variationnelles du gap, dans le cas du laplacien sur une variété lisse, des opérateurs elliptiques, et des processus de Markov. Présentons essentiellement une borne inférieure du gap des processus de Markov infinis. Pour la théorie fondamentale des processus de Markov infinis, voir [9].

Soit $\Lambda = (\lambda_{ij})$ un générateur de Markov irréductible et π -réversible défini sur $E = \{0, 1, 2, \dots\}$. Un *couplage markovien* de Λ est un générateur de Markov Λ^{coup} sur l'espace produit $E \times E$ vérifiant la condition de marginalité, pour tout couple (i, j) de $E \times E$,

$$\Lambda^{\text{coup}} f(i, j) = \Lambda f(i) \quad (\text{resp. } \Lambda^{\text{coup}} f(i, j) = \Lambda f(j))$$

pour toute fonction f bornée de $E \times E$ dans \mathbb{R} , dépendant seulement de la première variable (resp. la deuxième variable). Le couplage markovien appelé *couplage classique* ou *basique* est défini de la façon suivante.

$$\Lambda^{\text{coup}} f(i_1, i_2) = \begin{cases} (\Lambda f(\cdot, i_2))(i_1) + (\Lambda f(i_1, \cdot))(i_2) & \text{si } i_1 \neq i_2 \\ \Lambda \bar{f}(i_1) & \text{si } i_1 = i_2 \end{cases}$$

où $\bar{f}(i) = f(i, i)$.

Soit $\bar{\Lambda} = (\bar{\lambda}_{ij})$ un générateur π -réversible tel que $\forall j < i, \lambda_{ij} \geq \bar{\lambda}_{ij}$. Pour une nouvelle distribution $\tilde{\pi}$ telle que

$$0 < \inf_i \frac{\tilde{\pi}(i)}{\pi(i)} \leq \sup_i \frac{\tilde{\pi}(i)}{\pi(i)} < \infty,$$

définissons un générateur $\tilde{\Lambda} = (\tilde{\lambda}_{ij})$ tel que

$$\tilde{\lambda}_{ij} = \begin{cases} \bar{\lambda}_{ij} & \text{si } i > j \\ \frac{\tilde{\pi}(j)\bar{\lambda}_{ij}}{\tilde{\pi}(i)} & \text{si } i < j. \end{cases}$$

On a aussi besoin d'un générateur $\hat{\Lambda} = (\hat{\lambda}_{ij})$ sur $E_n = \{0, 1, \dots, n\}$ tel que

$$\hat{\lambda}_{ij} = \begin{cases} \bar{\lambda}_{ij} & \text{si } i, j \leq n \Leftrightarrow 1 \\ \sum_{k \geq n} \bar{\lambda}_{ij} & \text{si } i \leq n \Leftrightarrow 1, j = n \\ \frac{\tilde{\pi}(j) \sum_{k \geq n} \tilde{\lambda}_{jk}}{\sum_{k \geq n} \tilde{\pi}(k)} & \text{si } i = n, j \leq n \Leftrightarrow 1. \end{cases}$$

Un générateur de Markov Λ est dit *régulier*, s'il existe un seul semi-groupe admettant Λ comme générateur. Une formule générale de minoration est donnée dans [11].

Théorème 2.3.5 (M.F. Chen, 1995) *Supposons que les générateurs Λ , $\bar{\Lambda}$ et $\tilde{\Lambda}$ définis ci-dessus sont réguliers. Pour tout $n \geq 1$, soit Λ_n^{coup} un couplage markovien de $\hat{\Lambda}_n$. Soit ρ une distance sur E . Si pour tout n , il existe α_n telle que pour tous $i_1 \neq i_2$,*

$$\Lambda_n^{\text{coup}} \rho(i_1, i_2) \leq \Leftrightarrow \alpha_n \rho(i_1, i_2) ,$$

alors

$$\text{gap}(\Lambda) \geq \left[\inf_i \frac{\tilde{\pi}(i)}{\pi(i)} \Big/ \sup_i \frac{\tilde{\pi}(i)}{\pi(i)} \right] \overline{\lim}_{n \rightarrow \infty} \alpha_n . \quad (2.3)$$

Pour transformer ce résultat en une estimation efficace, il s'agit de trouver un bon couplage Λ_n^{coup} et une bonne distance ρ . Comme qu'il y a une infinité de couplages vérifiant (2.3) pour une distance ρ donnée, il est naturel de se poser les questions suivantes. En existe-t-il un optimal? En quel sens un couplage est-il optimal? Selon Chen, un générateur de couplage $\bar{\Lambda}^{\text{coup}}$ est dit ρ -optimal si pour tout générateur de couplage Λ^{coup} , on a

$$\bar{\Lambda}^{\text{coup}} \rho(i_1, i_2) \leq \Lambda^{\text{coup}} \rho(i_1, i_2) , \quad \forall i_1 \neq i_2 .$$

Ainsi un couplage optimal dépend fortement de la distance. De plus même pour un couplage optimal fixé, Chen a montré qu'il y a encore une grande classe de distances ρ pouvant être choisies. Donc la construction de la distance joue un rôle crucial dans l'application du théorème . Une distance convenable conduit à une estimation fine. Dans le cas des processus de naissance et de mort, en choisissant une distance ρ fortement liée au vecteur propre du gap et le couplage classique qui est ρ -optimal, Chen a obtenu une formule variationnelle du gap différente de (2.1).

Considérons un processus de naissance et de mort sur $E = \{0, 1, 2, \dots\}$ de paramètres $\{\nu_i, \mu_i\}_{i \geq 0}$. C'est un processus de Markov sur E dont les taux sont

définis par

$$\lambda_{ij} = \begin{cases} \nu_i & \text{si } i \geq 0, j = i + 1 \\ \mu_i & \text{si } i > 0, j = i \Leftrightarrow 1 \\ 0 & \text{sinon .} \end{cases}$$

Soit $\pi = \{\pi(i)\}$ la mesure réversible du processus avec $\pi(0) = 1$. Soit \mathcal{V} l'ensemble des suites positives $\{v_i\}_{i \geq 0}$. Définissons

$$R_i(v) = \mu_{i+1} + \nu_i \Leftrightarrow \mu_i / v_{i-1} \Leftrightarrow \nu_{i+1} v_i, \quad \mu_0 = 0, v_{-1} = 1, i \geq 0 .$$

Soit \mathcal{W} l'ensemble des suites strictement croissantes $\{w_i\}_{i \geq 1}$ avec $\sum_{i \geq 1} \pi(i)w_i > 0$. Définissons

$$I_i(w) = \nu_i \pi(i) (w_{i+1} \Leftrightarrow w_i) \left/ \sum_{j=i+1}^{\infty} \pi(j)w_j \right., \quad i \geq 1$$

$$I_0(w) = b_0 \left(1 + w_1 \left/ \sum_{j=1}^{\infty} \pi(j)w_j \right. \right) .$$

Théorème 2.3.6 (M.F. Chen, 1995) *Pour le processus de naissance et de mort défini ci-dessus, on a*

$$\text{gap}(\Lambda) = \sup_{v \in \mathcal{V}} \inf_{i \geq 0} R_i(v) .$$

$$\text{gap}(\Lambda) = \sup_{w \in \mathcal{W}} \inf_{i \geq 0} I_i(w) .$$

De plus, le sup peut être atteint dans les deux égalités ci-dessus.

Il est clair que pour toute suite $\{v_i\}$ ou $\{w_i\}$, on obtient une borne inférieure du gap. Ces deux formules variationnelles de gap sont plus utiles que (2.1) pour minorer le gap. En choisissant des suites $\{v_i\}$ ou $\{w_i\}$ convenables, Chen a montré (cf. [11]) qu'il est facile de retrouver les estimations existantes du gap des processus de naissance et de mort.

Chapitre 3

Spectres des générateurs quasi réductibles

Dans ce chapitre, toutes les fonctions f de ε sont définies sur un intervalle $(0, x]$ avec x positif et supposé petit. Par abus de l'angage, on va cependant parler de l'analyticité de f en $\varepsilon = 0$ au sens où il existe un prolongement f' de f à un voisinage de 0 tel que f' est analytique dans ce voisinage. Par conséquent $\lim_{\varepsilon \searrow 0} f$ existe. Notons

$$d(f) = \lim_{\varepsilon \searrow 0} \frac{\log |f|}{\log \varepsilon}$$

si la limite existe et

$$d(f) = +\infty$$

si $f \equiv 0$. On dit que f est d'ordre θ si $d(f) = i$, d'ordre O si $d(f) \geq i$ et d'ordre o si $d(f) > i$. Soit S un ensemble, $O(\varepsilon_s^i)$ et $o(\varepsilon_s^i)$ vont désigner des vecteur indicé par S dont chaque terme est d'ordre $O(\varepsilon^i)$ et $o(\varepsilon^i)$ respectivement.

3.1 Enoncé et démonstration du théorème A

On considère un générateur Λ fonction de ε et analytique en $\varepsilon = 0$, irréductible et réversible. En notant $S_k(\Lambda)$ le vecteur constitué des valeurs propres d'ordre $O(\varepsilon^k)$ du générateur Λ par ordre croissant avec multiplicité éventuelle, on va donner explicitement dans le théorème A (paragraphe 3.1.4) un générateur Λ_k d'ordre $\#S_k(\Lambda)$ tel que $1/\varepsilon^k S_k(\Lambda)$ et $S_0(\Lambda_k)$ ont la même limite. Ainsi les équivalents des valeurs propres d'ordre $\theta \varepsilon^k$ sont donnés par les valeurs propres non nulles multipliées par ε^k . L'énoncé et la démonstration du théorème A se fait en plusieurs étapes. L'idée principale de la démonstration est de structurer hiérarchiquement l'espace d'états à partir du générateur Λ . On va commencer par donner cette structure hiérarchique naturellement liée à Λ et certaines propriétés associées (3.1.1). Le générateur Λ est ainsi agrégé en blocs lui-même hiérarchiquement, et on montre dans le paragraphe 3.1.2 certaines propositions concernant cette hiérarchisation. Ensuite, quelques propriétés d'inversibilité et de continuité de certaines matrices qui sont indispensables techniquement pour la démonstration du théorème A sont rassemblées dans le paragraphe 3.1.3. Enfin, on montre le théorème A dans le paragraphe 3.1.4. Une comparaison entre le théorème A et les résultats existants se fait dans le paragraphe 3.1.5.

3.1.1 Structure hiérarchique

Soit $\Lambda = (\lambda_{ab})_{a,b \in E}$ un générateur de Markov analytique en $\varepsilon = 0$ sur l'espace d'états E . On le suppose irréductible et Π -réversible. La propriété suivante est sur l'analyticité de la mesure réversible Π de Λ .

Proposition 3.1.1 *La mesure réversible Π est analytique en $\varepsilon = 0$.*

Démonstration. Soit $a \in E$. Comme Λ est irréductible, pour tout b de E , il existe un entier l et des états c_1, \dots, c_l avec $c_1 = a$ et $c_l = b$ tels que

$$\lambda_{c_1 c_2} \cdots \lambda_{c_{l-1} c_l} > 0.$$

Par anliticité, l finira par ne plus dependre de ε pour $\varepsilon > 0$ assez petit. Posons

$$\Pi'(b) = \frac{\lambda_{c_1 c_2} \cdots \lambda_{c_{l-1} c_l}}{\lambda_{c_l c_{l-1}} \cdots \lambda_{c_2 c_1}} ,$$

on a

$$\Pi(b) = \frac{\Pi'(b)}{\sum_{c \in E} \Pi'(c)} .$$

Du fait que $\Pi(b)$ est borné et méromorphe en $\varepsilon = 0$, on en déduit que $\Pi(b)$ est analytique en $\varepsilon = 0$.

□

Introduisons d'abord quelques notations. De la réversibilité, pour tous $a, b \in E$, on déduit immédiatement que

$$d(\lambda_{ab}) + d(\Pi(a)) = d(\lambda_{ba}) + d(\Pi(b)) .$$

Posons

$$d(a, b) = d(\lambda_{ab}) + d(\Pi(a)) = d(\lambda_{ba}) + d(\Pi(b)) ,$$

qui est donc entier au vu de la proposition 3.1.1.

On définit la fonction de base B sur l'ensemble des sous-ensembles de E de la façon suivante

$$\forall S \subseteq E , B(S) = \min_{a \in S} d(\Pi(a)) .$$

Soit $\mathcal{E} = \{\alpha, \beta, \dots\}$ une partition de E . On utilise les notations suivantes par rapport à la partition \mathcal{E} .

- Pour tout a de E , on note $\mathcal{E}(a)$ la classe de \mathcal{E} contenant a .
- Soit $a, b \in E$. En notant $\alpha = \mathcal{E}(a)$, on pose

$$\delta_{\mathcal{E}}[a, b] = d(a, b) \Leftrightarrow B(\alpha) ,$$

qui mesure, en ordre de grandeur de ε , la difficulté de communication de l'état a vers l'état b , relative à la partition \mathcal{E} .

- Pour toutes classes α, β de \mathcal{E} , notons encore $\delta_{\mathcal{E}}[\alpha, \beta]$ la quantité suivante:

$$\delta_{\mathcal{E}}[\alpha, \beta] = \min \{ \delta_{\mathcal{E}}[a, b]; a \in \alpha, b \in \beta \}$$

qui mesure en quelque sorte la difficulté de passage de α à β relative à la partition \mathcal{E} .

Définissons les suites de fonctions d'énergie $\{H_k\}_{k \geq 1}$ sur E , de partitions $\{\mathcal{E}_k\}_{k \geq 0}$ de E , et de sous-ensembles des \mathcal{E}_k , $\{\mathcal{F}_k\}_{k \geq 0}$ par récurrence.

Définissons \mathcal{E}_0 comme la partition en singletons de E et identifions \mathcal{F}_0 à \mathcal{E}_0 . La fonction H_1 est définie par

$$\forall a \in E, \quad H_1(a) = d(\Pi(a)) .$$

Par référence au cas particulier du recuit simulé (paragraphe 3.2.3), H_1 sera appelée la fonction d'énergie. Cette fonction H_1 prend des valeurs positives ou nulles. Définissons la relation \mathcal{R}_1 sur E comme suit. On note $a\mathcal{R}_1b$ si et seulement si il existe un entier l et une suite $\{c_i\}_{1 \leq i \leq l}$ dans E avec $a = c_1$ et $b = c_l$ telles que pour tout i compris entre 1 et $l \Leftrightarrow 1$

$$H_1(c_i) = H_1(c_{i+1}) \quad \text{et} \quad \delta_{\mathcal{E}_0}[c_i, c_{i+1}] = 0 .$$

En fait, sous les hypothèses si-dessus,

$$\delta_{\mathcal{E}_0}[c_i, c_{i+1}] = 0 \Leftrightarrow \delta_{\mathcal{E}_0}[c_{i+1}, c_i] = 0 \Leftrightarrow d(\lambda_{c_i c_{i+1}}) = 0 \Leftrightarrow d(\lambda_{c_{i+1} c_i}) = 0 ,$$

La relation \mathcal{R}_1 est donc une relation d'équivalence. En d'autres termes, $a\mathcal{R}_1b$ s'il existe un chemin liant a, b tel que H_1 est constante sur ce chemin et tous les taux sur ce chemin ont une limite strictement positive quand ε tend vers 0. Définissons \mathcal{E}_1 l'ensemble des classes d'équivalence par rapport à \mathcal{R}_1 . Evidemment, H_1 est constante sur toute classe α_1 de \mathcal{E}_1 . Notons $h(\alpha_1)$ cette constante d'énergie sur α_1 . On dit que α_1 est une classe minimale d'ordre 1 si

$$\forall \beta_1 \in \mathcal{E}_1, \quad h(\alpha_1) > h(\beta_1) \implies \delta_{\mathcal{E}_1}[\alpha_1, \beta_1] \geq 1 ,$$

et non-minimale, sinon. Notons \mathcal{F}_1 l'ensemble des classes minimales d'ordre 1 qui est donc un sous-ensemble de \mathcal{E}_1 .

Pour $k \geq 1$, supposons définies la fonction d'énergie H_k et la partition \mathcal{E}_k qui sont telles que

- H_k prend des valeurs $\geq k \Leftrightarrow 1$,
- H_k est constante sur chaque élément de \mathcal{E}_k ,
- le générateur restreint de Λ sur chaque élément de \mathcal{E}_k est irréductible.

Un élément de \mathcal{E}_k sera appelé une classe d'ordre k . Pour toute classe α_k de \mathcal{E}_k , on note $h(\alpha_k)$ la valeur commune des $H_k(a)$, pour $a \in \alpha_k$. Afin d'alléger les écritures, nous notons désormais $\delta_k[a, b]$ la difficulté de la communication de a vers b relative à la partition \mathcal{E}_k i.e.

$$\delta_k[a, b] = \delta_{\mathcal{E}_k}[a, b] \quad .$$

Et pour toutes classes α_k, β_k de \mathcal{E}_k , notons $\delta[\alpha_k, \beta_k]$ la difficulté de la communication de α_k vers β_k relative à la partition \mathcal{E}_k , i.e.

$$\delta[\alpha_k, \beta_k] = \delta_{\mathcal{E}_k}[\alpha_k, \beta_k] \quad .$$

On dit que α_k est une classe minimale d'ordre k si

$$\forall \beta_k \in \mathcal{E}_k, h(\alpha_k) > h(\beta_k) \implies \delta[\alpha_k, \beta_k] \geq k$$

et non-minimale sinon. Notons \mathcal{F}_k l'ensemble des classes minimales d'ordre k . Définissons H_{k+1} de la façon suivante. Pour tout a de E

$$H_{k+1}(a) = \begin{cases} H_k(a) + 1 & \text{si } \mathcal{E}_k(a) \in \mathcal{F}_k \\ H_k(a) & \text{sinon} \end{cases}$$

où $\mathcal{E}_k(a)$ est la classe de la partition \mathcal{E}_k à laquelle a appartient. On constate que $H_{k+1} \geq H_k$ et que H_{k+1} prend des valeurs $\geq k$. De plus, H_{k+1} est constante sur toutes les classes α_k de \mathcal{E}_k . On va noter $H_{k+1}(\alpha_k)$ cette constante d'énergie.

La partition \mathcal{E}_{k+1} de E est définie à partir de \mathcal{E}_k et H_{k+1} en regroupant certaines classes de \mathcal{E}_k . Pour toute classe α_k de \mathcal{E}_k , nous notons $\overline{\alpha_k}$ la classe regroupée autour de α_k de la façon suivante:

$$\overline{\alpha_k} = \bigcup \{ \beta_k \in \mathcal{E}_k ; H_{k+1}(\beta_k) = H_{k+1}(\alpha_k) \text{ et } \delta[\alpha_k, \beta_k] \leq k \}$$

si α_k est minimale, et

$$\overline{\alpha_k} = \alpha_k$$

sinon. En d'autres termes, nous regroupons autour de chaque classe minimale d'ordre k , celles qui sont de même énergie par rapport à H_{k+1} et communiquent avec elle à un ordre inférieur ou égal à k . Définissons la partition \mathcal{E}_{k+1} de E comme la partition la plus fine telle que pour toutes classes α_k, β_k de \mathcal{E}_k , les classes $\overline{\alpha_k}, \overline{\beta_k}$ d'intersection non vide sont dans une même classe de \mathcal{E}_{k+1} . De façon équivalente, \mathcal{E}_{k+1} est l'ensemble des classes d'équivalence par rapport à la relation d'équivalence \mathcal{R}_{k+1} définie comme suit. Pour toutes $\alpha_k, \beta_k \in \mathcal{E}_k$, on note $\overline{\alpha_k} \mathcal{R}_{k+1} \overline{\beta_k}$ si et seulement si il existe un entier m et une suite $\{\gamma_k^i\}_{1 \leq i \leq m}$ de \mathcal{E}_k avec $\gamma_k^1 = \alpha_k$ et $\gamma_k^m = \beta_k$ telles que

$$\overline{\gamma_k^i} \cap \overline{\gamma_k^{i+1}} \neq \emptyset, \quad 1 \leq i \leq m \Leftrightarrow 1.$$

On constate que le générateur restreint de Λ sur chaque classe de \mathcal{E}_{k+1} est irréductible.

La construction définie ci-dessus est achevée dès que l'ensemble des classes minimales se trouve réduit à un seul élément. Notons K l'entier le plus petit tel que

$$\#\mathcal{F}_K = 1.$$

On remarque que pour $1 \leq i \leq k \leq K$, toute classe de \mathcal{E}_k est une union de classes de \mathcal{E}_i .

Définition 3.1.1 *Soit $\alpha_k \in \mathcal{E}_k$, on dit que α_k est une classe de type 1 si elle contient au moins une classe minimale d'ordre $k \Leftrightarrow 1$ et de type 2 sinon.*

On constate que si α_k est de type 2, forcément elle est aussi une classe de \mathcal{E}_{k-1} , autrement dit, il existe une classe non-minimale d'ordre $k \Leftrightarrow 1$ notée α_{k-1} telle que

$$\alpha_{k-1} = \alpha_k.$$

Soit $\alpha_i \in \mathcal{E}_i$, on va noter $\mathcal{E}_k(\alpha_i)$ la classe de \mathcal{E}_k contenant α_i . Et on note $F_i(\alpha_k)$ un sous-ensemble des classes minimales d'ordre i incluses dans α_k tel que

$$F_i(\alpha_k) = \{\alpha_i \in \mathcal{F}_i; \alpha_i \subseteq \alpha_k, \mathcal{E}_l(\alpha_i) \in \mathcal{F}_l, i \leq l \leq k\}.$$

Une autre définition équivalente de $F_i(\alpha_k)$ est la suivante.

$$F_i(\alpha_k) = \{ \alpha_i \in \mathcal{F}_i; \alpha_i \subseteq \alpha_k, h(\alpha_k) \Leftrightarrow h(\alpha_i) = k \Leftrightarrow i \} .$$

Evidemment, pour tout $i \leq j \leq k$,

$$F_i(\alpha_k) = \bigcup_{\alpha_j \in F_j(\alpha_k)} F_i(\alpha_j) .$$

Nous illustrons les notions définies ci-dessus par un exemple explicite.

Exemple 3.1.1

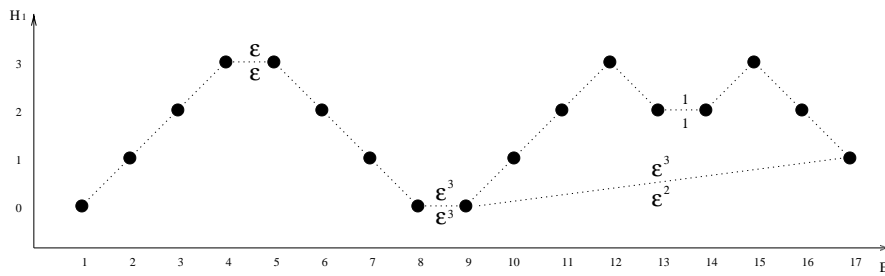


Figure 3.1: Diagramme des taux de transition.

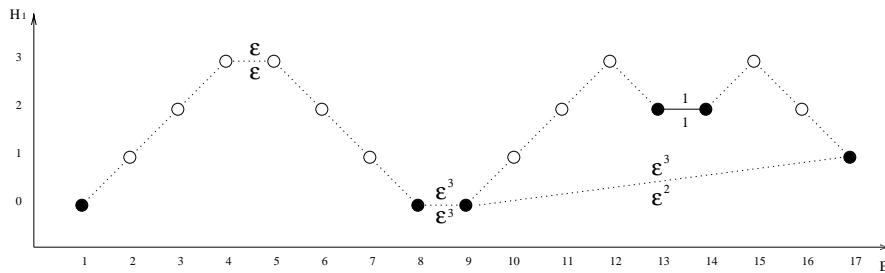
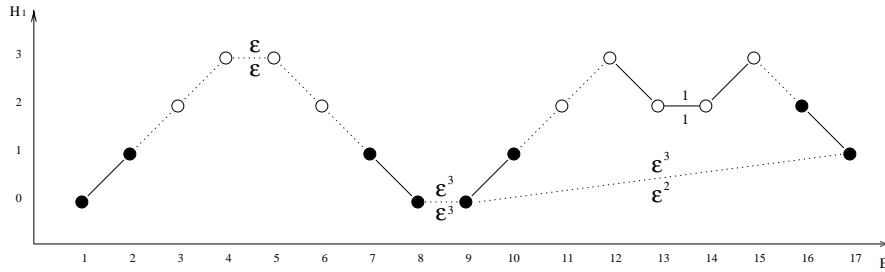
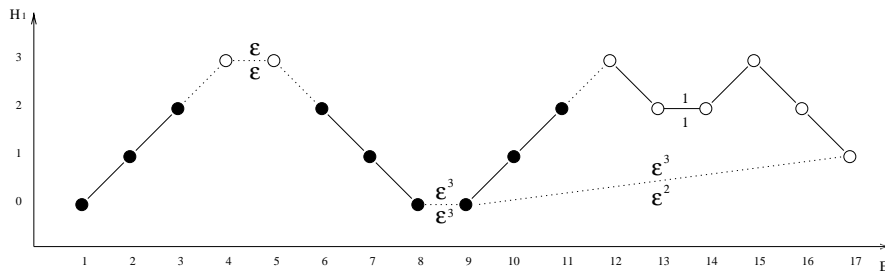


Figure 3.2: H_1 , \mathcal{E}_1 et \mathcal{F}_1 .

Soit $E = \{1, \dots, 17\}$. Donnons-nous un générateur Λ sur E par son diagramme de transition et la fonction d'énergie H_1 dans la figure 3.1. Dans cette figure, il n'y a des taux de transition non nuls qu'entre deux états liés par un trait pointillé. Pour les taux explicitement indiqués dans la figure, le taux au

Figure 3.3: H_2 , \mathcal{E}_2 et \mathcal{F}_2 .Figure 3.4: H_3 , \mathcal{E}_3 et \mathcal{F}_3 .

dessus du trait est celui de la transition de l'état de gauche vers l'état de droite. Le taux au-dessous du trait est celui de la transition de droite à gauche. Pour les taux qui ne sont pas indiqués sur la figure, ils valent ε pour les transitions qui augmentent la valeur de H_1 , ils valent 1 pour les transitions qui diminuent la valeur de H_1 .

Pour $1 \leq k \leq 5$, dans la figure nommée " H_k , \mathcal{E}_k , et \mathcal{F}_k ", une classe de \mathcal{E}_k est un ensemble de points connectés par des traits pleins, si les points d'une classe sont noirs (resp. blancs), il s'agit alors d'une classe minimale (resp. non-minimale). Dans cet exemple, la valeur $h(\alpha_k)$ de la fonction d'énergie se trouve être égale au maximum de H_1 sur cette classe ce qui n'est pas vrai dans le cas général.

Ici, on a $K = 5$. Posons $\alpha_4 = \{5, \dots, 17\}$ qui est une classe minimale d'ordre 4, alors

$$F_3(\alpha_4) = \{\{6, 7, 8\}, \{9, 10, 11\}\},$$

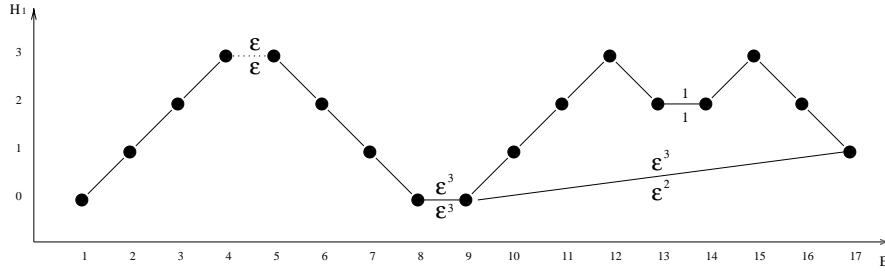


Figure 3.5: H_4 , \mathcal{E}_4 et \mathcal{F}_4 .

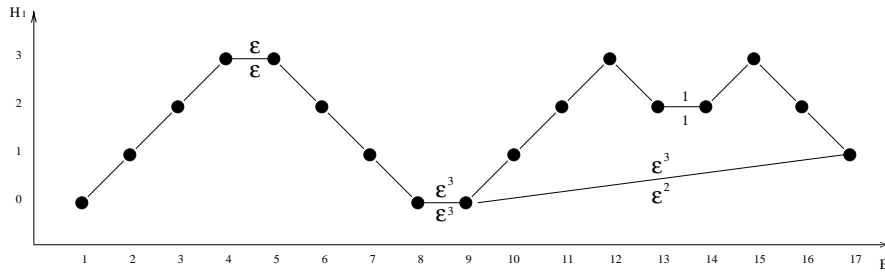


Figure 3.6: H_5 , \mathcal{E}_5 et \mathcal{F}_5 .

$$F_2(\alpha_4) = \{\{7, 8\}, \{9, 10\}\},$$

et

$$F_1(\alpha_4) = \{\{8\}, \{9\}\}.$$

Les suites de partitions $\{\mathcal{E}_k\}_{1 \leq k \leq K}$, d'ensembles de classes minimales $\{\mathcal{F}_k\}_{1 \leq k \leq K}$ et de fonctions d'énergie $\{H_k\}_{1 \leq k \leq K}$ que nous avons construites possèdent un certain nombre de propriétés qui seront cruciales pour la démonstration du théorème A. Ces propriétés sont rassemblées dans la proposition ci-dessous.

Proposition 3.1.2 *Les propriétés ci-dessous sont vraies pour tout k compris entre 1 et K .*

- (1) Soient α_k une classe de \mathcal{E}_k et β_k une classe de $\mathcal{E}_k \setminus \mathcal{F}_k$. S'il existe a de α_k

et b de β_k tels que

$$\delta_k[b, a] = \min \{ \delta_k[b', a']; b' \in \beta_k, a' \notin \beta_k \} ,$$

alors on a

$$h(\alpha_k) < h(\beta_k)$$

et

$$\delta_k[a, b] = h(\beta_k) \Leftrightarrow B(\alpha_k) = \delta[\alpha_k, \beta_k] .$$

(2) Si α_k est une classe minimale de \mathcal{F}_k , alors α_k est de type 1. Donc $F_i(\alpha_k)$ est non vide, pour $1 \leq i \leq k$.

(3) Soit $\alpha_k \in \mathcal{E}_k$. Si α_k est de type 1, alors

$$h(\alpha_k) \Leftrightarrow B(\alpha_k) = k \Leftrightarrow 1 .$$

Si α_k est de type 2, alors

$$h(\alpha_k) \Leftrightarrow B(\alpha_k) < k \Leftrightarrow 1 .$$

(4) Soit $\alpha_k \neq \beta_k \in \mathcal{E}_k$ telles que $h(\beta_k) \geq h(\alpha_k)$. On a

$$\delta[\alpha_k, \beta_k] \geq h(\beta_k) \Leftrightarrow B(\alpha_k) .$$

(5) Soit $\alpha_k \neq \beta_k \in \mathcal{E}_k$. Si α_k est de type 1 et $h(\alpha_k) = h(\beta_k)$, alors

$$\delta[\alpha_k, \beta_k] \geq k .$$

On dira que $P_i(k)$ est vraie si pour tout $j \leq k$, la propriété (i) ci-dessus est vraie, pour $1 \leq i \leq 5$. On va démontrer les 5 propriétés ci-dessus par récurrence selon les étapes suivantes. Dans l'étape I, on montrera que les $P_i(1)$ sont vraies avec $1 \leq i \leq 5$ et l'étape I.i consiste à montrer $P_i(1)$. Ensuite, nous supposons que $P_i(k \Leftrightarrow 1)$ sont vraies avec $1 \leq i \leq 5$. Dans l'étape II, nous montrerons

étape II.1: $P_1(k \Leftrightarrow 1)$ et $P_3(k \Leftrightarrow 1) \implies P_2(k)$;

étape II.2: $P_2(k \Leftrightarrow 1)$ et $P_3(k \Leftrightarrow 1) \implies P_3(k)$;

étape II.3: $P_3(k \Leftrightarrow 1)$ et $P_4(k \Leftrightarrow 1)$ et $P_5(k \Leftrightarrow 1) \implies P_4(k)$;

étape II.4: $P3(k)$ et $P4(k \Leftrightarrow 1)$ et $P5(k \Leftrightarrow 1) \implies P5(k)$;

étape II.5: $P3(k)$ et $P5(k \Leftrightarrow 1)$ et $P1(k \Leftrightarrow 1) \implies P1(k)$.

Démonstration. Constatons que pour toute classe α_1 d'ordre 1,

$$B(\alpha_1) = h(\alpha_1) \tag{3.1}$$

où B est la fonction de base.

Etape I.1: montrons que $P1(1)$ est vraie.

Comme β_1 est non-minimale, alors

$$\delta_1[b, a] \leq 0,$$

d'où

$$\delta_1[b, a] = 0 .$$

Alors

$$d(\lambda_{ab}) = \delta_1[b, a] + B(\alpha_1) \Leftrightarrow d(\Pi(a)) = 0 .$$

Si $d(\lambda_{ba}) = 0$, on a

$$H_1(a) = d(\Pi(a)) = d(\Pi(b)) = H_1(b) .$$

Donc a et b sont dans une même classe d'ordre 1 ce qui contredit l'hypothèse.

On a donc $d(\lambda_{ba}) > 0$ ce qui implique

$$h(\alpha_1) = d(\Pi(a)) < d(\Pi(b)) = h(\beta_1) .$$

En plus

$$\begin{aligned} \delta_1[a, b] &= \delta_1[b, a] + B(\beta_1) \Leftrightarrow B(\alpha_1) \\ &= h(\beta_1) \Leftrightarrow B(\alpha_1) . \end{aligned}$$

Pour le fait que

$$\delta[\alpha_1, \beta_1] = h(\beta_1) \Leftrightarrow B(\alpha_1) ,$$

c'est un résultat immédiat de $P4(1)$ qu'on va montrer après.

Etape I.2: $P2(1)$ est vraie car toute classe d'ordre 0 est minimale.

Etape I.3: On déduit que P3(1) est vraie du fait que toute classe d'ordre 1 est de type 1 et en utilisant l'égalité (3.1).

$$\begin{aligned}\delta[\alpha_1, \beta_1] &= \delta[\beta_1, \alpha_1] + B(\beta_1) \Leftrightarrow B(\alpha_1) \\ &= B(\beta_1) \Leftrightarrow B(\alpha_1) .\end{aligned}$$

Etape I.4: montrons que P4(1) est vraie.

Pour tous $a \in \alpha_1, b \in \beta_1$,

$$\begin{aligned}\delta_1[a, b] &= d(a, b) \Leftrightarrow B(\alpha_1) \\ &= d(\Pi(b)) + d(\lambda_{ba}) \Leftrightarrow B(\alpha_1) \\ &= h(\beta_1) + d(\lambda_{ba}) \Leftrightarrow B(\alpha_1) \\ &\geq h(\beta_1) \Leftrightarrow B(\alpha_1) .\end{aligned}$$

Etape I.5: montrons que P5(1) est vraie.

D'après P4(1), $\forall a \in \alpha_1, b \in \beta_1$

$$\delta_1[a, b] \geq h(\beta_1) \Leftrightarrow B(\alpha_1) = 0 .$$

Or si $\delta_1[a, b] = 0$, alors

$$d(\lambda_{ab}) = \delta_1[a, b] \Leftrightarrow d(\Pi(a)) + B(\alpha_1) = 0 .$$

On en déduit que a et b sont dans une même classe d'ordre 1 ce qui contredit le fait que $\alpha_1 \neq \beta_1$. D'où

$$\delta_1[a, b] \geq 1 .$$

Supposons maintenant que les $P_i(k \Leftrightarrow 1)$ sont vraies avec $1 \leq i \leq 5$.

Etape II.1: $P1(k \Leftrightarrow 1)$ et $P3(k \Leftrightarrow 1) \implies P2(k)$.

Soit α_k une classe minimale d'ordre k . Supposons que α_k est de type 2. Notons α_{k-1} la classe non-minimale d'ordre $k \Leftrightarrow 1$ telle que $\alpha_{k-1} = \alpha_k$. Posons

$$d_1 = \min\{\delta_k[a, b]; a \in \alpha_k, b \notin \alpha_k, H_k(b) < H_k(a)\} .$$

et

$$G_1 = \{(a, b); a \in \alpha_k, b \notin \alpha_k, H_k(b) < H_k(a), \delta_k[a, b] = d_1\} .$$

Comme α_k est minimale, forcément $d_1 \geq k$. Posons

$$d_2 = \min\{\delta_{k-1}[a, b]; a \in \alpha_{k-1}, b \notin \alpha_{k-1}, H_{k-1}(b) < H_{k-1}(a)\} .$$

et

$$G_2 = \{(a, b); a \in \alpha_{k-1}, b \notin \alpha_{k-1}, H_{k-1}(b) < H_{k-1}(a), \delta_{k-1}[a, b] = d_2\} .$$

Du fait que $\alpha_k = \alpha_{k-1}$ et $h(\alpha_k) = h(\alpha_{k-1})$, on déduit

$$G_2 = \{(a, b); a \in \alpha_k, b \notin \alpha_k, H_{k-1}(b) < H_k(a), \delta_k[a, b] = d_2\} .$$

Comme α_{k-1} est non-minimale, forcément $d_2 \leq k \Leftrightarrow 2$. Donc

$$G_1 \cap G_2 = \emptyset .$$

Soit $(a, b) \in G_2$, on en déduit

$$H_k(b) = H_{k-1}(b) + 1 = H_k(a) .$$

Posons $\beta_{k-1} = \mathcal{E}_{k-1}(b)$, alors β_{k-1} est une classe minimale. On a donc

$$\begin{aligned} \delta_{k-1}[b, a] &= h(\alpha_{k-1}) \Leftrightarrow B(\beta_{k-1}) \text{ (d'après P1}(k \Leftrightarrow 1)) \\ &= h(\alpha_{k-1}) \Leftrightarrow [h(\beta_{k-1}) \Leftrightarrow k + 2] \text{ (P3}(k \Leftrightarrow 1)) \\ &= k \Leftrightarrow 1 . \end{aligned}$$

Alors $\alpha_{k-1} \subseteq \overline{\beta_{k-1}}$ et donc α_{k-1} et β_{k-1} sont dans une même classe d'ordre k , ce qui contredit l'hypothèse.

Etape II.2: P2($k \Leftrightarrow 1$) et P3($k \Leftrightarrow 1$) \implies P3(k).

Si α_k est de type 2, soit α_{k-1} la classe non-minimale d'ordre $k \Leftrightarrow 1$ telle que $\alpha_{k-1} = \alpha_k$. Alors

$$h(\alpha_{k-1}) = h(\alpha_k)$$

et donc

$$\begin{aligned} h(\alpha_k) \Leftrightarrow B(\alpha_k) &= h(\alpha_{k-1}) \Leftrightarrow B(\alpha_{k-1}) \\ &\leq k \Leftrightarrow 2 \text{ (P3}(k \Leftrightarrow 1)) . \end{aligned}$$

Si α_k est de type 1, pour toute classe non-minimale β_{k-1} d'ordre $k \Leftrightarrow 1$ incluse dans α_k , on a

$$\begin{aligned} B(\beta_{k-1}) &\geq h(\beta_{k-1}) \Leftrightarrow k + 2 \quad (\text{P3}(k \Leftrightarrow 1)) \\ &= h(\alpha_k) \Leftrightarrow k + 2 . \end{aligned}$$

Or pour toute classe minimale α_{k-1} d'ordre $k \Leftrightarrow 1$ incluse dans α_k , d'après P2($k \Leftrightarrow 1$), α_{k-1} est de type 1, on a

$$\begin{aligned} B(\alpha_{k-1}) &= h(\alpha_{k-1}) \Leftrightarrow k + 2 \quad (\text{P3}(k \Leftrightarrow 1)) \\ &= h(\alpha_k) \Leftrightarrow k + 1 . \end{aligned}$$

On en déduit

$$B(\alpha_{k-1}) = B(\alpha_k) .$$

On a alors

$$\begin{aligned} h(\alpha_k) \Leftrightarrow B(\alpha_k) &= h(\alpha_{k-1}) + 1 \Leftrightarrow B(\alpha_{k-1}) \\ &= k \Leftrightarrow 1 \quad (\text{P3}(k \Leftrightarrow 1)) . \end{aligned}$$

Etape II.3: P3($k \Leftrightarrow 1$) et P4($k \Leftrightarrow 1$) et P5($k \Leftrightarrow 1$) \implies P4(k).

Soit $a \in \alpha_k$, $b \in \beta_k$ tels que

$$\delta_k[a, b] = \delta[\alpha_k, \beta_k] .$$

Posons

$$\alpha_{k-1} = \mathcal{E}_{k-1}(a) \quad \text{et} \quad \beta_{k-1} = \mathcal{E}_{k-1}(b) .$$

Evidemment

$$\delta_{k-1}[a, b] = \delta[\alpha_{k-1}, \beta_{k-1}] .$$

Si $h(\beta_k) = h(\beta_{k-1})$, on a

$$h(\alpha_{k-1}) \leq h(\beta_{k-1}) .$$

Alors

$$\begin{aligned} \delta_k[a, b] &= \delta_{k-1}[a, b] + B(\alpha_{k-1}) \Leftrightarrow B(\alpha_k) \\ &\geq h(\beta_{k-1}) \Leftrightarrow B(\alpha_k) \quad (\text{P4}(k \Leftrightarrow 1)) \\ &= h(\beta_k) \Leftrightarrow B(\alpha_k) . \end{aligned}$$

D'où le résultat.

Supposons maintenant $h(\beta_k) > h(\beta_{k-1})$, alors β_{k-1} est une classe minimale.

Considérons les trois cas suivants.

1) Si $h(\alpha_{k-1}) < h(\beta_{k-1})$, on a

$$\delta_{k-1}[b, a] \geq k \Leftrightarrow 1$$

d'après la définition de classe minimale. Donc

$$\begin{aligned} \delta_k[a, b] &= \delta_{k-1}[b, a] + B(\beta_{k-1}) \Leftrightarrow B(\alpha_k) \\ &\geq k \Leftrightarrow 1 + B(\beta_{k-1}) \Leftrightarrow B(\alpha_k) \\ &= h(\beta_{k-1}) + 1 \Leftrightarrow B(\alpha_k) \quad (\text{P3}(k \Leftrightarrow 1)) \\ &= h(\beta_k) \Leftrightarrow B(\alpha_k) . \end{aligned}$$

2) Si $h(\alpha_{k-1}) = h(\beta_{k-1})$, on a

$$\begin{aligned} \delta_k[a, b] &= \delta_{k-1}[b, a] + B(\beta_{k-1}) \Leftrightarrow B(\alpha_k) \\ &\geq k \Leftrightarrow 1 + B(\beta_{k-1}) \Leftrightarrow B(\alpha_k) \quad (\text{P5}(k \Leftrightarrow 1)) \\ &= h(\beta_{k-1}) + 1 \Leftrightarrow B(\alpha_k) \quad (\text{P3}(k \Leftrightarrow 1)) \\ &= h(\beta_k) \Leftrightarrow B(\alpha_k) . \end{aligned}$$

3) Si $h(\alpha_{k-1}) > h(\beta_{k-1})$, forcément $h(\alpha_{k-1}) = h(\alpha_k) = h(\beta_k)$. On a

$$\begin{aligned} \delta_k[a, b] &= \delta_{k-1}[b, a] + B(\beta_{k-1}) \Leftrightarrow B(\alpha_k) \\ &\geq h(\alpha_{k-1}) \Leftrightarrow B(\alpha_k) \quad (\text{P4}(k \Leftrightarrow 1)) \\ &= h(\beta_k) \Leftrightarrow B(\alpha_k) . \end{aligned}$$

D'où le résultat.

Etape II.4: $\text{P3}(k)$ et $\text{P4}(k \Leftrightarrow 1)$ et $\text{P5}(k \Leftrightarrow 1) \implies \text{P5}(k)$.

Soit $a \in \alpha_k$, $b \in \beta_k$ tels que

$$\delta_k[a, b] = \delta[\alpha_k, \beta_k] .$$

Posons pour tout i compris entre 1 et k ,

$$\alpha_i = \mathcal{E}_i(a) \text{ et } \beta_i = \mathcal{E}_i(b) .$$

Remarquons que

$$\delta_i[a, b] = \delta[\alpha_i, \beta_i] .$$

Si

$$h(\alpha_1) = h(\alpha_k) \text{ et } h(\beta_1) = h(\beta_k) , \quad (3.2)$$

on a $H_1(a) = H_1(b)$. Du fait que a et b ne sont pas dans une même classe d'ordre 1, on déduit

$$d(\lambda_{ab}) \geq 1 ,$$

d'où

$$\begin{aligned} \delta_1[a, b] &= d(a, b) \Leftrightarrow B(\alpha_1) \\ &= d(\Pi(a)) + d(\lambda_{ab}) \Leftrightarrow B(\alpha_1) \\ &= d(\lambda_{ab}) \\ &\geq 1 . \end{aligned}$$

Alors

$$\begin{aligned} \delta_k[a, b] &= \delta_1[a, b] + B(\alpha_1) \Leftrightarrow B(\alpha_k) \\ &\geq 1 + B(\alpha_1) \Leftrightarrow B(\alpha_k) \\ &= 1 + h(\alpha_1) \Leftrightarrow [h(\alpha_k) \Leftrightarrow k + 1] \text{ (P3}(k)) \\ &= k . \end{aligned}$$

Si (3.2) n'est pas vrai, soit $i \leq k \Leftrightarrow 1$ le plus grand entier tel que

$$h(\alpha_i) = h(\alpha_k) \Leftrightarrow 1 \text{ ou } h(\beta_i) = h(\beta_k) \Leftrightarrow 1 .$$

Considérons les trois cas suivants.

1) Si $h(\alpha_i) = h(\alpha_k) \Leftrightarrow 1$ et $h(\beta_i) = h(\beta_k) \Leftrightarrow 1$, alors α_i et β_i sont minimales. D'après P5($k \Leftrightarrow 1$)

$$\delta_i[a, b] \geq i .$$

Or si $\delta_i[a, b] = i$, α_i et β_i seront dans une même classe d'ordre k ce qui contredit le fait que $\alpha_k \neq \beta_k$. Donc

$$\delta_i[a, b] \geq i + 1 .$$

D'où

$$\begin{aligned}
\delta_k[a, b] &= \delta_i[a, b] + B(\alpha_i) \Leftrightarrow B(\alpha_k) \\
&\geq i + 1 + B(\alpha_i) \Leftrightarrow B(\alpha_k) \\
&= i + 1 + [h(\alpha_i) \Leftrightarrow i + 1] \Leftrightarrow [h(\alpha_k) \Leftrightarrow k + 1] \quad (\text{P3}(k)) \\
&= k .
\end{aligned}$$

2) Si $h(\alpha_i) = h(\alpha_k) \Leftrightarrow 1$ et $h(\beta_i) = h(\beta_k)$, alors α_i est minimale et β_i non-minimale. On a

$$\begin{aligned}
\delta_i[a, b] &\geq h(\beta_i) \Leftrightarrow B(\alpha_i) \quad (\text{P4}(k \Leftrightarrow 1)) \\
&= h(\beta_i) \Leftrightarrow [h(\alpha_i) \Leftrightarrow i + 1] \quad (\text{P3}(k)) \\
&= i .
\end{aligned}$$

Mais si $\delta_i[a, b] = i$, α_i et β_i seront dans une même classe d'ordre $i + 1$ ce qui est contradictoire. D'où $\delta_i[a, b] \geq i + 1$. Du même raisonnement que le cas 1), on déduit

$$\delta_k[a, b] \geq k .$$

3) Si $h(\alpha_i) = h(\alpha_k)$ et $h(\beta_i) = h(\beta_k) \Leftrightarrow 1$, alors α_i est non-minimale et β_i minimale. On a

$$\begin{aligned}
\delta_i[b, a] &\geq h(\alpha_i) \Leftrightarrow B(\beta_i) \quad (\text{P4}(k \Leftrightarrow 1)) \\
&= h(\alpha_i) \Leftrightarrow [h(\beta_i) \Leftrightarrow i + 1] \quad (\text{P3}(k)) \\
&= i .
\end{aligned}$$

Mais si $\delta_i[a, b] = i$, α_i et β_i seront dans une même classe d'ordre $i + 1$, ce qui est contradictoire. D'où $\delta_i[a, b] \geq i + 1$. Alors

$$\begin{aligned}
\delta_k[a, b] &= \delta_i[b, a] + B(\beta_i) \Leftrightarrow B(\alpha_k) \\
&\geq i + 1 + B(\beta_i) \Leftrightarrow B(\alpha_k) \\
&= i + 1 + [h(\beta_i) \Leftrightarrow i + 1] \Leftrightarrow [h(\alpha_k) \Leftrightarrow k + 1] \quad (\text{P3}(k)) \\
&= k .
\end{aligned}$$

D'où le résultat.

Etape II.5: P3(k) et P4(k) et P5(k) et P1($k \Leftrightarrow 1$) \implies P1(k).

Comme β_k est non-minimale, on a

$$\delta_k[b, a] \leq k \Leftrightarrow 1 . \quad (3.3)$$

Si β_k est de type 2, soit β_{k-1} la classe non-minimale d'ordre $k \Leftrightarrow 1$ telle que

$$\beta_{k-1} = \beta_k .$$

On a

$$\begin{aligned} \delta_{k-1}[b, a] &= \delta_k[b, a] \\ &= \min\{\delta_k[b', a']; b' \in \beta_k, a' \notin \beta_k\} \\ &= \min\{\delta_{k-1}[b', a']; b' \in \beta_{k-1}, a' \notin \beta_{k-1}\}. \end{aligned}$$

Donc

$$\begin{aligned} \delta_k[a, b] &= \delta_{k-1}[a, b] + B(\alpha_{k-1}) \Leftrightarrow B(\alpha_k) \\ &= h(\beta_{k-1}) \Leftrightarrow B(\alpha_k) \quad (\text{P1}(k \Leftrightarrow 1)) \\ &= h(\beta_k) \Leftrightarrow B(\alpha_k) , \end{aligned}$$

ce qu'on voulait démontrer .

Si β_k est de type 1, supposons que $h(\alpha_k) \geq h(\beta_k)$. Quand $h(\alpha_k) > h(\beta_k)$,

on a

$$\begin{aligned} \delta_k[b, a] &= \delta[\beta_k, \alpha_k] \\ &\geq h(\alpha_k) \Leftrightarrow B(\beta_k) \quad (\text{P4}(k)) \\ &\geq h(\alpha_k) \Leftrightarrow [h(\beta_k) \Leftrightarrow k + 1] \quad (\text{P3}(k)) \\ &\geq k , \end{aligned}$$

ce qui contredit l'équation (3.3). Quand $h(\alpha_k) = h(\beta_k)$.

$$\begin{aligned} \delta_k[b, a] &= \delta[\beta_k, \alpha_k] \\ &\geq k \quad (\text{P5}(k)) , \end{aligned}$$

ce qui contredit aussi (3.3). On en déduit

$$h(\alpha_k) < h(\beta_k) .$$

Posons pour tout i compris entre 1 et k ,

$$\alpha_i = \mathcal{E}_i(a) \text{ et } \beta_i = \mathcal{E}_i(b) .$$

Si $h(\beta_1) = h(\beta_k)$, alors

$$\begin{aligned} \delta_1[b, a] &= \delta_k[b, a] + B(\beta_k) \Leftrightarrow B(\beta_1) \\ &\leq k \Leftrightarrow 1 + B(\beta_k) \Leftrightarrow B(\beta_1) \text{ (d'après l'équation (3.3))} \\ &= k \Leftrightarrow 1 + B(\beta_k) \Leftrightarrow h(\beta_k) \\ &= 0 \text{ (P3(k))} . \end{aligned}$$

Par conséquent,

$$\delta_1[b, a] = 0 .$$

On a alors

$$\begin{aligned} \delta_k[a, b] &= \delta_1[b, a] + B(\beta_1) \Leftrightarrow B(\alpha_k) \\ &= h(\beta_k) \Leftrightarrow B(\alpha_k) . \end{aligned}$$

Si $h(\beta_1) \neq h(\beta_k)$, notons $i \leq k \Leftrightarrow 1$ le plus grand entier tel que

$$h(\beta_i) = h(\beta_k) \Leftrightarrow 1 .$$

Nécessairement β_i est une classe minimale d'ordre i . Alors

$$\delta_i[b, a] \geq i$$

soit d'après la définition de classe minimale si $h(\alpha_i) < h(\beta_i)$, soit d'après P5($k \Leftrightarrow 1$) si $h(\alpha_i) = h(\beta_i)$. Alors

$$\begin{aligned} \delta_k[b, a] &= \delta_i[b, a] + B(\beta_i) \Leftrightarrow B(\beta_k) \\ &\geq i + [h(\beta_i) \Leftrightarrow i + 1] \Leftrightarrow [h(\beta_k) \Leftrightarrow k + 1] \text{ (P3(k))} \\ &= k \Leftrightarrow 1 . \end{aligned}$$

En tenant compte de (3.3), on a

$$\delta_k[b, a] = k \Leftrightarrow 1 .$$

D'où

$$\begin{aligned} \delta_k[a, b] &= \delta_k[b, a] + B(\beta_k) \Leftrightarrow B(\alpha_k) \\ &= k \Leftrightarrow 1 + B(\beta_k) \Leftrightarrow B(\alpha_k) \\ &= h(\beta_k) \Leftrightarrow B(\alpha_k) \quad (\text{P3}(k)) . \end{aligned}$$

Pour le fait

$$\delta[\alpha_k, \beta_k] = h(\beta_k) \Leftrightarrow B(\alpha_k) ,$$

c'est un résultat immédiat compte tenu de P4(k).

□

Au travers de la définition des partitions \mathcal{E}_k comme de la proposition 3.1.2 apparaît une structure hiérarchique des classes, minimales ou non. Nous allons maintenant préciser cette structure hiérarchique. Pour cela, introduisons quelques définitions et notations. Pour la suite, α_k désigne toujours une classe minimale d'ordre k , élément de \mathcal{F}_k .

On s'intéresse à l'arbre de hauteur k défini comme suit. La racine de l'arbre, le nœud de hauteur 0, est α_k . Les fils de α_k sont constitués par les classes de \mathcal{E}_{k-1} incluses dans α_k . Récursivement, les nœuds de l'arbre de hauteur $k \Leftrightarrow i$ sont composés de certaines classes de \mathcal{E}_i incluses dans α_k . Soit $\alpha_i \in \mathcal{E}_i$ un nœud de hauteur $k \Leftrightarrow i$. Il n'a pas de fils, si α_i est non-minimale. Et ses fils sont les classes d'ordre $i \Leftrightarrow 1$ incluses dans α_i , si α_i est minimale. Un nœud de l'arbre sera dit minimal s'il correspond à une classe minimale, non-minimal sinon. On remarque que $F_i(\alpha_k)$ est en fait l'ensemble des nœuds minimaux de hauteur $k \Leftrightarrow i$.

Soit $\alpha_i \in F_i(\alpha_k)$. Pour $i + 1 \leq j \leq k$, posons $\alpha_j = \mathcal{E}_j(\alpha_i)$ qui est forcément dans $F_j(\alpha_k)$. Définissons les ensembles $O_j[\alpha_i]$, $U_j[\alpha_i]$ tels que

$$O_j[\alpha_i] = \begin{cases} \alpha_i & \text{si } j = i \\ \alpha_j \Leftrightarrow \alpha_{j-1} & \text{si } i < j \leq k \\ E \Leftrightarrow \alpha_k & \text{si } j = k + 1, \end{cases}$$

et

$$U_j[\alpha_i] = \begin{cases} \alpha_1 & \text{si } j = i = 1 \\ \alpha_j \Leftrightarrow \bigcup_{\beta_{j-1} \in F_{j-1}(\alpha_j)} \beta_{j-1} & \text{si } i \leq j \leq k \text{ et } j \neq 1 \\ E \Leftrightarrow \bigcup_{\beta_k \in \mathcal{F}_k} \beta_k & \text{si } j = k + 1. \end{cases}$$

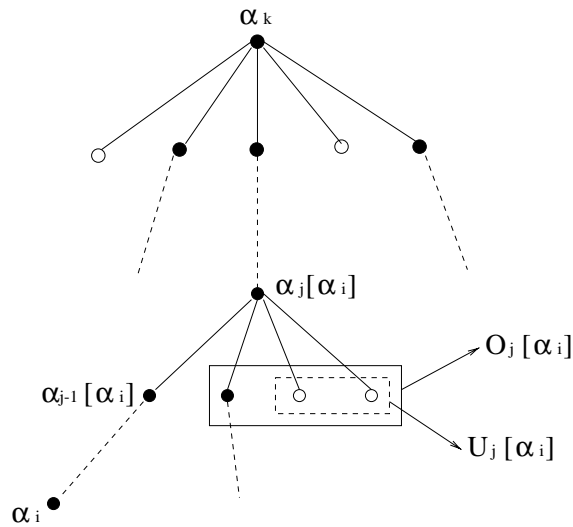


Figure 3.7: Arbre issu de α_k .

Dans la figure 3.7, on donne un exemple d'arbre ayant α_k comme racine. Dans cette figure, les points noirs et blancs représentent respectivement les nœuds minimaux et non-minimaux. Soit α_i un nœud minimal de hauteur $k \Leftrightarrow i$. Alors α_i est un nœud du sous-arbre issu de $\alpha_{j-1}[\alpha_i]$ avec $j \Leftrightarrow 1 \geq i$. Et $O_j[\alpha_i]$ est l'union des nœuds encadrés par le rectangle en trait plein et $U_j[\alpha_i]$ l'union des nœuds encadrés par le rectangle pointillé.

En notant

$$O_j = O_j[\alpha_i] \text{ et } U_j = U_j[\alpha_i],$$

on remarque que

- $\{O_j\}_{l \leq j \leq k+1}$ est une partition de $E \setminus \alpha_{l-1}$ si $l \geq i+1$ et une partition de E si $l = i$,
- et que

$$O_j = U_j \cup \bigcup_{\substack{\beta_{j-1} \in F_{j-1}(\alpha_j) \\ \beta_{j-1} \neq \alpha_{j-1}}} \beta_{j-1} \quad (3.4)$$

- et que l'ensemble

$$\{U_l[\alpha_l]; \alpha_l \in F_l(\alpha_i), 1 \leq l \leq i\}$$

est une partition de α_i .

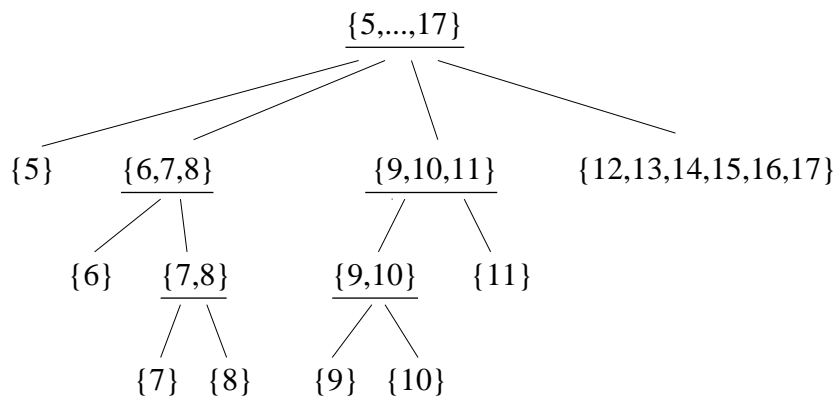


Figure 3.8: Arbre issu de $\alpha_4 = \{5, \dots, 17\}$.

Reprenons l'exemple 3.1.1, et posons par exemple

$$\alpha_4 = \{5, \dots, 17\}.$$

Dans la figure 3.8, c'est l'arbre ayant α_4 comme racine défini ci-dessus où les ensembles soulignés sont les nœuds minimaux de l'arbre. Posons

$$\alpha_1 = \{9\} \text{ et } \alpha_j = \mathcal{E}_j(\alpha_1)$$

pour $1 \leq j \leq 4$. Alors

$$\alpha_2 = \{9, 10\}, \alpha_3 = \{9, 10, 11\}, \alpha_4 = \{5, \dots, 17\}.$$

Notons pour $1 \leq j \leq 4$

$$O_j = O_j[\alpha_1] \text{ et } U_j = U_j[\alpha_1].$$

Alors

$$\begin{aligned} O_1 &= \{9\}, & O_2 &= \{10\}, \\ O_3 &= \{11\}, & O_4 &= \{5, \dots, 8, 12, \dots, 17\}, \end{aligned}$$

et

$$U_1 = \alpha_1, U_2 = \{10\}, U_3 = \{11\}, U_4 = \{5, 12, \dots, 17\}, U_5 = \emptyset.$$

Les difficultés de communication ainsi que les degrés des taux de transition sont liés à la hiérarchie qui vient d'être décrite.

Proposition 3.1.3 *Soit $\alpha_i \in F_i(\alpha_k)$. Soit $a \in \alpha_i$, $b \in O_j[\alpha_i]$ avec $i + 1 \leq j \leq k + 1$. Alors*

$$a) \ d(a, b) \geq \delta_{j-1}[a, b] \geq j \Leftrightarrow 1,$$

$$b) \ d(\lambda_{ab}) \geq j \Leftrightarrow i.$$

Démonstration. Posons

$$\alpha_{j-1} = \mathcal{E}_{j-1}(a) \text{ et } \beta_{j-1} = \mathcal{E}_{j-1}(b).$$

Nécessairement, α_{j-1} est minimale. Montrons

$$\delta_{j-1}[a, b] \geq j \Leftrightarrow 1. \tag{3.5}$$

Si β_{j-1} est minimale, comme

$$\alpha_{j-1}, \beta_{j-1} \subset \alpha_j,$$

on a

$$h(\alpha_{j-1}) = h(\beta_{j-1}).$$

Donc (3.5) est vrai d'après la propriété (5) de la proposition 3.1.2.

Si β_{j-1} est non-minimale, alors

$$h(\alpha_{j-1}) = h(\beta_{j-1}) \Leftrightarrow 1$$

d'où

$$\begin{aligned} \delta_{j-1}[a, b] &\geq h(\beta_{j-1}) \Leftrightarrow B(\alpha_{j-1}) \text{ (d'après (4) de la proposition 3.1.2)} \\ &= h(\beta_{j-1}) \Leftrightarrow [h(\alpha_{j-1}) \Leftrightarrow j + 2] \text{ ((3) de la proposition 3.1.2)} \\ &= j \Leftrightarrow 1. \end{aligned}$$

D'autre part

$$\begin{aligned} d(a, b) &= \delta_{j-1}[a, b] + B(\alpha_{j-1}) \\ &\geq \delta_{j-1}[a, b]. \end{aligned}$$

Ce qui démontre l'inégalité a).

Pour l'inégalité b). On a

$$\begin{aligned} d(\lambda_{ab}) &= \delta_{j-1}[a, b] + B(\alpha_{j-1}) \Leftrightarrow d(\Pi(a)) \\ &\geq j \Leftrightarrow 1 + B(\alpha_{j-1}) \Leftrightarrow d(\Pi(a)) \text{ (d'après l'égalité (3.5)).} \end{aligned}$$

Or

$$\begin{aligned} B(\alpha_{j-1}) &= h(\alpha_{j-1}) \Leftrightarrow j + 2 \text{ ((3) de la proposition 3.1.2)} \\ &= h(\alpha_i) \Leftrightarrow i + 1 \\ &= B(\alpha_i) \text{ ((3) de la proposition 3.1.2)}. \end{aligned}$$

Donc

$$\begin{aligned} d(\lambda_{ab}) &\geq j \Leftrightarrow 1 + B(\alpha_i) \Leftrightarrow d(\Pi(a)) \\ &\geq j \Leftrightarrow 1 + B(\alpha_i) \Leftrightarrow h(\alpha_i) \\ &= j \Leftrightarrow i \text{ ((3) de la proposition 3.1.2)}, \end{aligned}$$

d'où l'inégalité b).

□

3.1.2 Décomposition hiérarchique du générateur

Nous rassemblons dans ce paragraphe les résultats quantitatifs qui relient le générateur Λ et sa mesure réversible Π à la structure hiérarchique de l'espace d'état.

Pour tout sous-ensemble S de E , on note Λ_S le générateur restreint de Λ sur S . Pour tous sous-ensembles S_1, S_2 de E avec $S_1 \cap S_2 = \emptyset$, notons

$$G_{S_1}^{S_2} = (\lambda_{ab})_{a \in S_1, b \in S_2}$$

la matrice composée des taux de transition de S_1 à S_2 . Et posons $D_{S_1}^{S_2}$ la matrice diagonale telle que

$$D_{S_1}^{S_2} \mathbb{1}_{S_1} = G_{S_1}^{S_2} \mathbb{1}_{S_2} .$$

Soit $\alpha_i \in F_i(\alpha_k)$. On note pour $i \leq j \leq k+1$

$$U_j = U_j[\alpha_i] .$$

Si $S_1 \subseteq \alpha_i, S_2 \subseteq O_j$, posons

$$A_{S_1}^{S_2} = \frac{1}{\varepsilon^{j-i}} G_{S_1}^{S_2} .$$

Selon *b*) de la proposition 3.1.3, la matrice $A_{S_1}^{S_2}$ est analytique en $\varepsilon = 0$. Notons $B_{S_1}^{S_2}$ la matrice diagonale telle que

$$B_{S_1}^{S_2} \mathbb{1}_{S_1} = A_{S_1}^{S_2} \mathbb{1}_{S_2} .$$

On pose

$$\alpha_j = \mathcal{E}_j(\alpha_i) .$$

Notons D_{U_j} la matrice diagonale telle que

$$D_{U_j} \mathbb{1}_{U_j} = \begin{cases} \sum_{\beta_{j-1} \in F_{j-1}(\alpha_j)} G_{U_j}^{\beta_{j-1}} \mathbb{1}_{\beta_{j-1}} & \text{si } i \leq j \leq k \\ \sum_{\beta_k \in \mathcal{F}_k} G_{U_{k+1}}^{\beta_k} \mathbb{1}_{\beta_k} & \text{si } j = k+1 . \end{cases}$$

Notons Π_{α_k} le vecteur ligne égal à la mesure de probabilité réversible de Λ_{α_k} . D'après la proposition 2.1.6,

$$\forall a \in E, \Pi_{\alpha_k}(a) = \frac{\Pi(a)}{\sum_{b \in \alpha_k} \Pi(b)} .$$

On en déduit que

$$d(\Pi_{\alpha_k}(a)) = d(\Pi(a)) \Leftrightarrow B(\alpha_k).$$

Définition 3.1.2 Soit $\Lambda = (\lambda_{ab})_{a,b \in E}$ un générateur de Markov irréductible et π -réversible sur l'espace d'états E . On dit que Λ est irréductible au niveau k , si et seulement si pour tout $a, b \in E$, il existe un entier $l \geq 2$ et un ensemble d'états $\{c_1, \dots, c_l\} \subseteq E$ avec $c_1 = a$, $c_l = b$ tels que pour tout $1 \leq i \leq l \Leftrightarrow 1$, on a

$$d(\pi(c_i)) + d \not\lambda_{c_i c_{i+1}} \leq k.$$

Dans la proposition suivante, on montre que Λ_{α_i} est irréductible au niveau $i \Leftrightarrow 1$ pour toute classe α_i d'ordre i et de type 1.

Proposition 3.1.4 Pour tout $1 \leq i \leq K$, soit α_i une classe d'ordre i . Alors pour tous $a, b \in \alpha_i$, il existe un entier $l \geq 2$ et un ensemble d'états $\{c_1, \dots, c_l\} \subseteq \alpha_i$ avec $c_1 = a$, $c_l = b$ tels que pour tout $1 \leq j \leq l \Leftrightarrow 1$,

$$d(\Pi(c_j)) + d \not\lambda_{c_j c_{j+1}} \leq h(\alpha_i).$$

En particulier, si α_i est de type 1, alors Λ_{α_i} est irréductible au niveau $i \Leftrightarrow 1$.

Démonstration. Pour $i = 1$, le résultat découle immédiatement de la définition des classes d'ordre 1.

Supposons que le résultat est vrai pour $m \leq i \Leftrightarrow 1$. Si α_i est de type 2, il existe alors une classe d'ordre $i \Leftrightarrow 1$ identique à α_i , on déduit ainsi le résultat par récurrence.

Si α_i est de type 1, alors pour tous $a, b \in \alpha_i$, il existe un entier $l \geq 2$ et un ensemble d'états $\{c_1, \dots, c_l\} \subseteq \alpha_i$ avec $c_1 = a$, $c_l = b$ tels que pour tout $1 \leq j \leq l \Leftrightarrow 1$, $\lambda_{c_j c_{j+1}} > 0$ et qu'en notant $\alpha_{i-1} = \mathcal{E}_{i-1}(c_j)$, il y a trois cas possibles

- 1) $c_{j+1} \in \alpha_{i-1}$,
- 2) $c_{j+1} \notin \alpha_{i-1}$, $\alpha_{i-1} \in \mathcal{F}_{i-1}$ et $\delta_i[c_j, c_{j+1}] \leq i \Leftrightarrow 1$,

3) $c_{j+1} \in \beta_{i-1} \neq \alpha_{i-1}$ et $\beta_{j-1} \in \mathcal{F}_{i-1}$.

Pour le cas 1), par récurrence,

$$d(\Pi(c_j)) + d \cancel{\chi_{c_j c_{j+1}}} \leq h(\alpha_{i-1}) \leq h(\alpha_i).$$

Pour le cas 2), on en déduit

$$\begin{aligned} & d(\Pi(c_j)) + d \cancel{\chi_{c_j c_{j+1}}} \\ = & \delta_i [c_j, c_{j+1}] + B(\alpha_{i-1}) \\ \leq & i \Leftrightarrow 1 + B(\alpha_{i-1}) \\ = & h(\alpha_{i-1}) + 1 \quad (\text{d'après (3) de la proposition 3.1.2}) \\ = & h(\alpha_i). \end{aligned}$$

Pour le cas 3), on a

$$\begin{aligned} & d(\Pi(c_j)) + d \cancel{\chi_{c_j c_{j+1}}} \\ = & \delta_i [c_{j+1}, c_j] + B(\beta_{i-1}) \\ \leq & i \Leftrightarrow 1 + B(\beta_{i-1}) \\ \leq & h(\beta_{i-1}) + 1 \quad (\text{d'après (3) de la proposition 3.1.2}) \\ = & h(\alpha_i). \end{aligned}$$

On constate de plus que

$$d(\Pi_{\alpha_i}(c_j)) + d \cancel{\chi_{c_j c_{j+1}}} = \delta_i [c_j, c_{j+1}] \leq i \Leftrightarrow 1.$$

D'où le résultat. □

Pour tout sous-ensemble S de α_k , $\Pi_{\alpha_k}^S$ désigne le vecteur restreint de Π_{α_k} sur S . On pose

$$q_{\alpha_k}^S = \lim_{\varepsilon \searrow 0} \Pi_{\alpha_k}^S \mathbb{1}_S.$$

En fait, Π_{α_k} ne charge que les états des classes de $F_1(\alpha_k)$, quand ε tend vers 0.

Proposition 3.1.5 *Soit $\alpha_i \in F_i(\alpha_k)$. Alors*

$$q_{\alpha_k}^{\alpha_i} = \sum_{\alpha_{i-1} \in F_{i-1}(\alpha_i)} q_{\alpha_k}^{\alpha_{i-1}} .$$

Démonstration. Soit $a \in \alpha_k$, $a \notin \bigcup_{\alpha_1 \in F_1(\alpha_k)} \alpha_1$. Soit i l'entier le plus grand tel que a appartienne à un nœud non-minimal d'ordre i noté β_i . Soit

$$\beta_{i+1} = \mathcal{E}_{i+1}(\beta_i) \text{ et } \alpha_i \in F_i(\beta_{i+1}) .$$

Alors

$$\begin{aligned} B(\beta_i) &\geq h(\beta_i) \Leftrightarrow i + 1 \text{ (d'après (3) de la proposition 3.1.2)} \\ &= h(\alpha_i) \Leftrightarrow i + 2 \\ &= B(\alpha_i) + 1 \text{ ((3) de la proposition 3.1.2),} \end{aligned}$$

donc

$$d(\Pi(a)) \geq B(\beta_i) > B(\alpha_i) \geq B(\alpha_k) .$$

D'où

$$\lim_{\varepsilon \searrow 0} \Pi_{\alpha_k}(a) = \lim_{\varepsilon \searrow 0} \frac{\Pi(a)}{\sum_{b \in \alpha_k} \Pi(b)} = 0 .$$

On a alors

$$\begin{aligned} q_{\alpha_k}^{\alpha_i} &= \lim_{\varepsilon \searrow 0} \sum_{\alpha_1 \in F_1(\alpha_i)} \Pi_{\alpha_k}^{\alpha_1} \mathbb{1}_{\alpha_1} \\ &= \lim_{\varepsilon \searrow 0} \sum_{\alpha_{i-1} \in F_{i-1}(\alpha_i)} \sum_{\alpha_1 \in F_1(\alpha_{i-1})} \Pi_{\alpha_k}^{\alpha_1} \mathbb{1}_{\alpha_1} \\ &= \sum_{\alpha_{i-1} \in F_{i-1}(\alpha_i)} q_{\alpha_k}^{\alpha_{i-1}} . \end{aligned}$$

□

Proposition 3.1.6 *Soit $\alpha_k \in \mathcal{F}_k$ et $\alpha_{i+1} \in F_{i+1}(\alpha_k)$ avec $i + 1 \leq k$. Notons*

$$U_{i+1} = U_{i+1}[\alpha_{i+1}] .$$

Alors on a

$$a) \sum_{\alpha_i \in F_i(\alpha_{i+1})} \Pi_{\alpha_k}^{\alpha_i} G_{\alpha_i}^{U_{i+1}} \cancel{N_{U_{i+1}} \Leftrightarrow D_{U_{i+1}}}^{-1} = \Leftrightarrow \Pi_{\alpha_k}^{U_{i+1}},$$

b) pour toute suite de fonctions $\{c_{\alpha_i}\}_{\alpha_i \in F_i(\alpha_{i+1})}$,

$$\begin{aligned} & \Leftrightarrow \sum_{\alpha_i \in F_i(\alpha_{i+1})} c_{\alpha_i} \Pi_{\alpha_k}^{\alpha_i} G_{\alpha_i}^{U_{i+1}} \mathbb{1}_{U_{i+1}} \\ &= \sum_{\alpha_i \in F_i(\alpha_{i+1})} \Pi_{\alpha_k}^{\alpha_i} G_{\alpha_i}^{U_{i+1}} \cancel{N_{U_{i+1}} \Leftrightarrow D_{U_{i+1}}}^{-1} \\ & \times \sum_{\alpha_i \in F_i(\alpha_{i+1})} c_{\alpha_i} G_{U_{i+1}}^{\alpha_i} \mathbb{1}_{\alpha_i}. \end{aligned}$$

Démonstration. De la réversibilité, on déduit

$$\begin{aligned} & \sum_{\alpha_i \in F_i(\alpha_{i+1})} \Pi_{\alpha_k}^{\alpha_i} G_{\alpha_i}^{U_{i+1}} \cancel{N_{U_{i+1}} \Leftrightarrow D_{U_{i+1}}}^{-1} \\ &= \Pi_{\alpha_k}^{U_{i+1}} D_{U_{i+1}} \cancel{N_{U_{i+1}} \Leftrightarrow D_{U_{i+1}}}^{-1} \\ &= \Leftrightarrow \Pi_{\alpha_k}^{U_{i+1}} \cancel{N_{U_{i+1}} \Leftrightarrow D_{U_{i+1}} + \Lambda_{U_{i+1}}} \cancel{N_{U_{i+1}} + D_{U_{i+1}}}^{-1}. \end{aligned}$$

On déduit l'égalité a) du fait que

$$\Pi_{\alpha_k}^{U_{i+1}} \Lambda_{U_{i+1}} = 0.$$

Pour l'égalité b),

$$\begin{aligned} & \sum_{\alpha_i \in F_i(\alpha_{i+1})} \Pi_{\alpha_k}^{\alpha_i} G_{\alpha_i}^{U_{i+1}} \cancel{N_{U_{i+1}} \Leftrightarrow D_{U_{i+1}}}^{-1} \sum_{\alpha_i \in F_i(\alpha_{i+1})} c_{\alpha_i} G_{U_{i+1}}^{\alpha_i} \mathbb{1}_{\alpha_i} \\ &= \Leftrightarrow \sum_{\alpha_i \in F_i(\alpha_{i+1})} c_{\alpha_i} \Pi_{\alpha_k}^{U_{i+1}} G_{U_{i+1}}^{\alpha_i} \mathbb{1}_{\alpha_i} \text{ (selon l'égalité a)} \\ &= \Leftrightarrow \sum_{\alpha_i \in F_i(\alpha_{i+1})} c_{\alpha_i} \Pi_{\alpha_k}^{\alpha_i} G_{\alpha_i}^{U_{i+1}} \mathbb{1}_{U_{i+1}} \text{ (réversibilité)}. \end{aligned}$$

□

3.1.3 Propositions préparatoires

Dans ce paragraphe, nous établirons quelques propositions techniques préparant la démonstration du théorème A.

Proposition 3.1.7 *Soit $\Lambda = (\lambda_{ij})_{i,j \in E}$ un générateur irréductible, Π -réversible et analytique en $\varepsilon = 0$ sur E . Alors les valeurs propres et les vecteurs propres normalisés*

Proposition 3.1.8 *de Λ sont analytiques en $\sqrt{\varepsilon}$ au voisinage de 0.*

Démonstration. Posons

$$D = \text{diag} \left(\sqrt{\Pi(i)} \right)_{i \in E} .$$

Alors l'élément de colonne i et de ligne j de la matrice $D\Lambda D^{-1}$ est $\sqrt{\lambda_{ij}\lambda_{ji}}$. Donc $D\Lambda D^{-1}$ est symétrique et analytique en $\sqrt{\varepsilon}$ au voisinage de 0. D'après le théorème 6.1 du chapitre II de [44], les valeurs propres et vecteurs propres normalisés de $D\Lambda D^{-1}$ sont analytiques en $\sqrt{\varepsilon}$ au voisinage de 0. D'après la proposition 3.1.1, D est analytique en $\sqrt{\varepsilon}$ au voisinage de 0, on en déduit que les valeurs propres et vecteurs propres de Λ sont analytiques en $\sqrt{\varepsilon}$ au voisinage de 0.

□

L'objectif principal du théorème A est de calculer les équivalents des valeurs propres de Λ en $\varepsilon = 0$. Que les ordres de grandeur des ces équivalents soient des puissances entières de ε n'est pas évident a priori.

Proposition 3.1.9 *Soit λ une valeur propre de Λ . Alors $d(\lambda)$ est un entier.*

Démonstration. Soit $n + 1$ le cardinal de l'espace d'états. Notons les valeurs propres de Λ dans l'ordre décroissant.

$$0 = \lambda_0 > \lambda_1 \geq \dots \geq \lambda_n .$$

Supposons que $\{N_l; 1 \leq l \leq m\}$ est une partition de l'ensemble $\{1, \dots, n\}$ et $\{d_l; 1 \leq l \leq m\}$ est un ensemble de réels strictement croissants tels que

$$\forall i \in N_l, d(\lambda_i) = d_l .$$

Posons le polynôme

$$P(x) = \frac{1}{x} |\Lambda \Leftrightarrow xI| .$$

On a alors

$$P(x) = (\Leftrightarrow 1)^n \prod_{i=1}^n (x \Leftrightarrow \lambda_i) .$$

Soit a_i le coefficient de x^i dans le polynôme $P(x)$. On a d'autre part

$$P(x) = \sum_{i=0}^n a_i x^i .$$

En additionnant toutes les autres colonnes de $|\Lambda \Leftrightarrow xI|$ à la première, cette dernière devient $x\mathbb{1}$. Donc $P(x)$ est le déterminant de la matrice obtenue en remplaçant la première colonne de $\Lambda \Leftrightarrow xI$ par $\mathbb{1}$. Ceci implique que les a_i sont analytiques en $\varepsilon = 0$. Par conséquent, $d(a_i)$ est entier. Or

$$a_{n-1} = \Leftrightarrow \sum_{i=1}^n \lambda_i ,$$

d'où

$$d(a_{n-1}) = d_1 ,$$

qui est donc entier.

Supposons que d_k sont entiers pour $k \leq l$. Posons

$$r = \sum_{k=1}^l \#N_k .$$

On a

$$a_{n-r-1} = (\Leftrightarrow 1)^{r+1} \sum_{j_1 < j_2 \cdots < j_{r+1}} \lambda_{j_1} \lambda_{j_2} \cdots \lambda_{j_{r+1}} .$$

Or pour tout état i de N_{l+1} ,

$$\begin{aligned} d(\lambda_1 \lambda_2 \cdots \lambda_r \lambda_i) &= \sum_{k=1}^l \#N_k d_k + d_{l+1} \\ &= \min \left\{ d \lambda_{j_1} \lambda_{j_2} \cdots \lambda_{j_{r+1}} ; j_1 < j_2 \cdots < j_{r+1} \right\} . \end{aligned}$$

On en déduit

$$d(a_{n-r-1}) = \sum_{k=1}^l \#N_k d_k + d_{l+1} ,$$

D'où d_{l+1} est entier.

□

Proposition 3.1.10 Soit Λ un générateur d'ordre n , irréductible, réversible et analytique en $\varepsilon = 0$. Soit G une matrice d'ordre $n \times m$, positive et analytique en $\varepsilon = 0$. Soit D une matrice diagonale d'ordre n et analytique en $\varepsilon = 0$ telle que

$$D \mathbb{I}_n \geq G \mathbb{I}_m .$$

(1) Alors $\Lambda \Leftrightarrow D$ est inversible. De plus, tout coefficient de $(\Lambda \Leftrightarrow D)^{-1}$ est négatif ou nul et supérieur ou égal au coefficient diagonal sur la même colonne. Par conséquent, tout cofacteur du déterminant $|\Lambda \Leftrightarrow D|$ est négatif ou nul.

(2) La matrice $(\Lambda \Leftrightarrow D)^{-1} G$ est analytique en $\varepsilon = 0$.

Démonstration. Montrons la propriété (1).

Notons

$$A = \Lambda \Leftrightarrow D = (a_{ij})_{1 \leq i, j \leq n} .$$

La matrice D étant supposée non nulle, il existe un entier i_0 tel que

$$\sum_{\substack{j=1 \\ j \neq i_0}}^n a_{i_0 j} < \Leftrightarrow a_{i_0 i_0} .$$

Supposons que A n'est pas inversible. Alors il existe un vecteur colonne $X = (x_i)_{1 \leq i \leq n}$ non nul tel que $AX = 0$. On pose

$$m = \max\{|x_i|; 1 \leq i \leq n\},$$

alors $m > 0$. Posons

$$M = \{i; |x_i| = m\} .$$

On a $M \neq \{1, \dots, n\}$, sinon

$$\left| \sum_{\substack{j=1 \\ j \neq i_0}}^n x_j a_{i_0 j} \right| \leq \sum_{\substack{j=1 \\ j \neq i_0}}^n |x_j a_{i_0 j}| \leq |x_{i_0}| \sum_{\substack{j=1 \\ j \neq i_0}}^n a_{i_0 j} < |x_{i_0}| a_{i_0 i_0}$$

ce qui est contradictoire. Comme Λ est irréductible, alors $\exists i \in M$ et $j \in \{1, \dots, n\} \Leftrightarrow M$ tel que $a_{ij} \neq 0$, comme $|x_j| < |x_i|$

$$\left| \sum_{\substack{l=1 \\ l \neq i}}^n x_l a_{il} \right| \leq \sum_{\substack{l=1 \\ l \neq i}}^n |x_l a_{il}| < |x_i| \sum_{\substack{l=1 \\ l \neq i}}^n a_{il} \leq |x_i a_{ii}| ,$$

ce qui est contradictoire. D'où l'inversibilité de A .

Il existe une constante $c > 0$ et une matrice positive B telles que

$$\Lambda \Leftrightarrow D = \Leftrightarrow c (I \Leftrightarrow B) .$$

Il vient que

$$(\Lambda \Leftrightarrow D)^{-1} = \Leftrightarrow c^{-1} \sum_{k=0}^{+\infty} B^k .$$

On en déduit que $\Leftrightarrow (\Lambda \Leftrightarrow D)^{-1}$ est positive.

Montrons

$$\forall 1 \leq i, j \leq n, b_{jj} \leq b_{ij} .$$

Sans perte de généralité, on montre le résultat quand $j = 1$. Il est vrai si tous les b_{i1} sont identiques. Sinon, soit b_{i1} un des éléments les plus petits sur la première colonne de B . Forcément, l'élément sur la première colonne et la i -ème ligne de AB est strictement positif, or $AB = I$, forcément $i = 1$.

Montrons la propriété (2).

Comme on a

$$(\Lambda \Leftrightarrow D) \mathbb{1} = \Leftrightarrow D \mathbb{1} \leq \Leftrightarrow G \mathbb{1} ,$$

il existe alors un vecteur v positif tel que

$$(\Lambda \Leftrightarrow D) \mathbb{1} = \Leftrightarrow v \Leftrightarrow G \mathbb{1} .$$

Par suite

$$\mathbb{1} = \Leftrightarrow (\Lambda \Leftrightarrow D)^{-1} v \Leftrightarrow (\Lambda \Leftrightarrow D)^{-1} G \mathbb{1} .$$

Du fait que $\Leftrightarrow (\Lambda \Leftrightarrow D)^{-1} v$ et $\Leftrightarrow (\Lambda \Leftrightarrow D)^{-1} G$ sont toutes les deux positives ou nulles, on déduit l'analyticité de $(\Lambda \Leftrightarrow D)^{-1} G$.

□

Nous appliquons la proposition 3.1.10 dans la situation suivante.

Corollaire 3.1.1 *Soit $\alpha_i \in F_i(\alpha_k)$, pour $1 \leq i \leq k$. Notons*

$$U_i = U_i[\alpha_i]$$

quand $i \leq k$ et

$$U_{k+1} = U_{k+1}[\alpha_k].$$

a) Alors $\Lambda_{U_i} \Leftrightarrow D_{U_i}$ est inversible pour $\varepsilon > 0$ et tout coefficient de $(\Lambda_{U_i} \Leftrightarrow D_{U_i})^{-1}$ est négatif ou nul et supérieur ou égal au coefficient diagonal de la même colonne.

b) Pour tout sous ensemble S de $\alpha_i \Leftrightarrow U_i$,

$$(\Lambda_{U_i} \Leftrightarrow D_{U_i})^{-1} G_{U_i}^S$$

est analytique en $\varepsilon = 0$.

Démonstration. Notons

$$\mathcal{M} = \{M_j\}_{1 \leq j \leq l}$$

l'ensemble des classes irréductibles de U_i par rapport à Λ_{U_i} . Pour tout j entre 1 et l , soit $\Lambda_{U_i}^j$ le générateur restreint de Λ_{U_i} sur M_j et $D_{U_i}^j$ la matrice diagonale telle que

$$D_{U_i} = \text{diag} \left(D_{U_i}^j \right)_{1 \leq j \leq l}.$$

Alors

$$\Lambda_{U_i} \Leftrightarrow D_{U_i} = \text{diag} \left(\Lambda_{U_i}^j \Leftrightarrow D_{U_i}^j \right)_{1 \leq j \leq l}.$$

Il suffit de vérifier les hypothèses de la proposition 3.1.10 pour $\Lambda_{U_i}^j$ et $D_{U_i}^j$ pour $1 \leq j \leq l$. Notons que le générateur restreint de Λ sur α_i (resp. E , si $i = k$) est irréductible, donc pour chaque classe M_j , il existe des taux de transitions non nuls vers $\alpha_i \setminus U_i$ (resp. $E \setminus U_i$, si $i = k$), ce qui signifie que D_{U_i} est positive. D'où le point a).

Pour le point b), il suffit de constater que

$$(\Lambda_{U_i} \Leftrightarrow D_{U_i})^{-1} G_{U_i}^S = \text{diag} \left(\left(\Lambda_{U_i}^j \Leftrightarrow D_{U_i}^j \right)^{-1} G_{M_j}^S \right)_{1 \leq j \leq l}.$$

En appliquant (2) de la proposition 3.1.10, on en déduit b).

□

La proposition suivante étudiant la limite de certaines matrices quand ε tend vers 0 est importante pour la démonstration du lemme 3.1.3 et celle du théorème A.

Proposition 3.1.11 *Soit $\Lambda = (\lambda_{ij})_{1 \leq i, j \leq n}$ un générateur irréductible, π -réversible et analytique en $\varepsilon = 0$. Supposons que*

$$D = \text{diag}(d_i)_{1 \leq i \leq n} \text{ et } B = \text{diag}(b_i)_{1 \leq i \leq n}$$

sont positives et analytiques en $\varepsilon = 0$. Soit A une matrice positive d'ordre $n \times m$, analytique en $\varepsilon = 0$ telle que

$$B \mathbb{I}_n \geq A \mathbb{I}_m .$$

Supposons que Λ est irréductible au niveau $k \geq 1$ et que pour tout $1 \leq i \leq n$, on a

$$d(\pi(i)b_i) \geq k$$

et il existe i tel que

$$d(\pi(i)d_i) \leq k .$$

On a alors les résultats suivants.

(1) *Il existe un vecteur ligne continu v d'ordre m et une matrice O , d'ordre $m \times n$ et continue telles que*

$$\varepsilon (\Lambda \Leftrightarrow \varepsilon B)^{-1} A = \mathbb{I}_n v + \varepsilon O .$$

(2) *On a*

$$\lim_{\varepsilon \searrow 0} \varepsilon (\Lambda \Leftrightarrow \varepsilon B \Leftrightarrow D)^{-1} A = 0 .$$

(3) *Pour tout vecteur colonne X continu en ε tel que la limite*

$$\lim_{\varepsilon \searrow 0} \varepsilon^{k+1} (\Lambda \Leftrightarrow \varepsilon B)^{-1} X$$

existe, il existe une fonction continue c et un vecteur colonne Y continu telles que

$$\varepsilon^{k+1} (\Lambda \Leftrightarrow \varepsilon B)^{-1} X = c \mathbb{I}_n + \varepsilon Y .$$

Pour démontrer cette proposition, on a besoin des deux lemmes suivants. Introduisons d'abord quelques notations.

Pour tout sous-ensemble S de $\{2, \dots, n\}$, notons

$$\overline{S} = \{2, \dots, n\} \ominus S .$$

Posons L_S le déterminant composé des taux de transition λ_{ij} avec $i, j \in \overline{S}$ ou bien le déterminant de la matrice obtenue en supprimant toute i -ème colonne et ligne avec i appartenant à $\{1\} \cup S$. C'est-à-dire

$$L_S = \left| (\lambda_{ij})_{i, j \in \overline{S}} \right| .$$

Par exemple,

$$L_{\{2, \dots, i-1\}} = \begin{vmatrix} \lambda_{ii} & \cdots & \lambda_{in} \\ \vdots & \ddots & \vdots \\ \lambda_{ni} & \cdots & \lambda_{nn} \end{vmatrix} , \quad L_{\emptyset} = \begin{vmatrix} \lambda_{22} & \cdots & \lambda_{2n} \\ \vdots & \ddots & \vdots \\ \lambda_{n2} & \cdots & \lambda_{nn} \end{vmatrix} .$$

Posons en particulier

$$L_{\{2, \dots, n\}} = 1 .$$

Soit \mathcal{T}_S l'ensemble des applications de \overline{S} dans $\{1, \dots, n\}$ tel que pour tous $T \in \mathcal{T}_S$, $S' \subseteq \overline{S}$ avec $S' \neq \emptyset$, on ait $T(S') \not\subseteq S'$. En d'autres termes, $T \in \mathcal{T}_S$ si et seulement si $T(j) \neq j$ pour tout $j \in \overline{S}$ et que dans le graphe $G(T)$ ayant

$$\overline{S} \cup T(\overline{S})$$

comme ensemble de sommets et

$$\{(j, T(j)) ; j \in \overline{S}\}$$

comme l'ensemble d'arêtes, chaque composante connexe est un arbre, i.e. il n'y pas de boucle. Rappelons que \mathcal{T}_S est l'ensemble des $\{1\} \cup S$ -graphes dans la terminologie de [27].

Lemme 3.1.1 (de la proposition 3.1.11) *Soit $\Lambda = (\lambda_{ij})_{1 \leq i, j \leq n}$ un générateur irréductible et réversible. Alors pour tout sous-ensemble S de $\{2, \dots, n\}$*

différent de $\{2, \dots, n\}$, on a

$$L_S = (\Leftrightarrow 1)^{n-\#S} \sum_{T \in \mathcal{T}_S} \prod_{j \in \bar{S}} \lambda_{jT(j)} .$$

Démonstration. Sans perte de généralité, montrons le lemme dans le cas où

$$S = \emptyset \text{ ou } S = \{2, \dots, i \Leftrightarrow 1\}$$

avec $3 \leq i \leq n$.

Pour alléger les écritures, notons \mathcal{T} l'ensemble des applications de $\{i, \dots, n\}$ dans $\{1, \dots, n\}$ avec $2 \leq i \leq n$ tel que pour tous $T \in \mathcal{T}$, $S \subseteq \{i, \dots, n\}$ avec $S \neq \emptyset$, on ait $T(S) \not\subseteq S$. Posons $\bar{\mathcal{T}}$ l'ensemble des applications de $\{i, \dots, n\}$ dans $\{1, \dots, n\}$ tel que

$$\forall T \in \bar{\mathcal{T}}, \forall j \in \{i, \dots, n\} \quad T(j) \neq j .$$

Evidemment

$$\mathcal{T} \subset \bar{\mathcal{T}} .$$

Si on développe le déterminant $L_{\{2, \dots, i-1\}}$ en remplaçant les λ_{jj} par $\Leftrightarrow \sum_{l \neq j} \lambda_{jl}$, on voit que $L_{\{2, \dots, i-1\}}$ est une combinaison linéaire des éléments de l'ensemble

$$\left\{ \prod_{j=i}^n \lambda_{jT(j)} ; T \in \bar{\mathcal{T}} \right\} ,$$

i.e.

$$L_{\{2, \dots, i-1\}} = \sum_{T \in \bar{\mathcal{T}}} c(T) \prod_{j=i}^n \lambda_{jT(j)}$$

où $c(T)$ est une constante.

1) Montrons que si $T \in \bar{\mathcal{T}} \Leftrightarrow T$, alors $c(T) = 0$.

Comme $G(T)$ n'est pas un arbre, quitte à changer les indices, on peut supposer que $G(T)$ contient la boucle

$$\{(j, j+1); i \leq j \leq l \Leftrightarrow 1\} \cup \{(l, i)\}$$

qui correspond au facteur $\lambda_{li} \prod_{j=i}^{l-1} \lambda_{jj+1}$ dans $c(T) \prod_{j=i}^n \lambda_{jT(j)}$. On remarque que $c(T) \prod_{j=i}^n \lambda_{jT(j)}$ est forcément un terme dans le développement de

$$\left| \begin{array}{cccc} \Leftrightarrow \sum_{j \neq i} \lambda_{ij} & \lambda_{ii+1} & \cdots & \lambda_{il} \\ \lambda_{i+1i} & \Leftrightarrow \sum_{j \neq i+1} \lambda_{i+1j} & \cdots & \lambda_{i+1l} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{li} & \lambda_{li+1} & \cdots & \Leftrightarrow \sum_{j \neq l} \lambda_{lj} \end{array} \right| \left| \begin{array}{ccc} \lambda_{l+l+1} & \cdots & \lambda_{l+1n} \\ \vdots & \ddots & \vdots \\ \lambda_{ni+1} & \cdots & \lambda_{nn} \end{array} \right|.$$

Or tous les termes contenant le facteur $\lambda_{li} \prod_{j=i}^{l-1} \lambda_{jj+1}$ dans le développement de l'expression ci-dessus ont

$$(\Leftrightarrow)^{l-i+1} + (\Leftrightarrow)^{l-i-2} = 0$$

comme coefficient. D'où $c(T) = 0$. On a alors

$$L_{\{2, \dots, i-1\}} = \sum_{T \in \mathcal{T}} c(T) \prod_{j=i}^n \lambda_{jT(j)}$$

2) Montrons

$$\forall T \in \mathcal{T}, c(T) = (\Leftrightarrow)^{n-i+1}.$$

Il suffit de montrer par récurrence $c(T) \prod_{j=i}^n \lambda_{jT(j)}$ est un terme dans le développement de

$$\prod_{j=i}^n \left(\Leftrightarrow \sum_{l \neq j} \lambda_{jl} \right)$$

et par conséquent,

$$c(T) = (\Leftrightarrow)^{n-i+1}.$$

Si $i = n$,

$$L_{\{2, \dots, n-1\}} = \Leftrightarrow \sum_{j \neq n} \lambda_{nj},$$

d'où le résultat est vrai. Supposons qu'il est vrai pour tout $j > i$. Considérons $L_{\{2, \dots, i-1\}}$. Comme

$$T(\{i, \dots, n\}) \not\subseteq \{i, \dots, n\},$$

quitte à changer les indices, il existe alors un entier $l > i$ tel que

$$T(j) \in \begin{cases} \{1, \dots, i \Leftrightarrow 1\} & \text{si } i \leq j \leq l \\ \{i, \dots, n\} & \text{si } l+1 \leq j \leq n. \end{cases}$$

Alors $c(T) \prod_{j=i}^n \lambda_{jT(j)}$ est forcément un terme dans le développement de

$$\left[\prod_{j=i}^l \left(\Leftrightarrow \sum_{k \neq j} \lambda_{jk} \right) \right] L_{\{2, \dots, l\}} .$$

Or par récurrence, dans le développement de $L_{\{2, \dots, l\}}$, le terme

$$(\Leftrightarrow)^{n-l} \prod_{j=l+1}^n \lambda_{jT(j)}$$

est un terme du développement de

$$\prod_{j=l+1}^n \left(\Leftrightarrow \sum_{k \neq j} \lambda_{jk} \right) ,$$

donc $c(T) \prod_{j=i}^n \lambda_{jT(j)}$ est un terme du développement de

$$\prod_{j=i}^n \left(\Leftrightarrow \sum_{l \neq j} \lambda_{jl} \right) .$$

Ce qui termine la démonstration. □

Lemme 3.1.2 (de la proposition 3.1.11) *Soit $\Lambda = (\lambda_{ij})_{1 \leq i, j \leq n}$ un générateur irréductible, π -réversible et analytique en $\varepsilon = 0$. Soit $B = \text{diag}(b_i)_{1 \leq i \leq n}$ une matrice positive et analytique en $\varepsilon = 0$. Supposons que Λ est irréductible au niveau $k \geq 1$ et que*

$$\forall 1 \leq i \leq n, \quad d(\pi(i)b_i) \geq k .$$

Alors pour tous sous-ensemble S de $\{2, \dots, n\}$, on a

$$d \left(L_S \prod_{j \in S} b_j \right) \geq d(L_\emptyset) .$$

Démonstration. Si $S = \emptyset$, il n'y a rien à démontrer. Si $S = \{2, \dots, n\}$, alors

$$L_S \prod_{j \in S} b_j = b_2 b_3 \cdots b_n .$$

Puisque Λ est irréductible au niveau k , il existe au moins un $T \in \mathcal{T}_S$ tel que

$$\prod_{j=2}^n \lambda_{jT(j)} \neq 0 \text{ et } d \left(\prod_{j=2}^n \lambda_{jT(j)} \right) \leq k .$$

Or d'après les conditions données,

$$d(b_j) \geq d(\lambda_{jT(j)}) .$$

On en déduit

$$d \left(L_S \prod_{j \in S} b_j \right) = d(b_2 b_3 \cdots b_n) \geq d \left(\prod_{j=2}^n \lambda_{jT(j)} \right) \geq d(L_\emptyset) .$$

Soit S différent de \emptyset et $\{2, \dots, n\}$, quitte à changer les indices, on peut supposer que $S = \{2, \dots, i \Leftrightarrow 1\}$ avec $2 < i \leq n$. D'après le lemme 3.1.1

$$L_S = (\Leftrightarrow 1)^{n-i+1} \sum_{T \in \mathcal{T}_S} \prod_{j=i}^n \lambda_{jT(j)} ,$$

donc

$$d \left(L_S \prod_{j=2}^{i-1} b_j \right) = \min \left\{ d \left(\prod_{j=2}^{i-1} b_j \prod_{j=i}^n \lambda_{jT(j)} \right) ; T \in \mathcal{T}_S, \prod_{j=i}^n \lambda_{jT(j)} \neq 0 \right\} .$$

De même,

$$d(L_\emptyset) = \min \left\{ d \left(\prod_{j=2}^n \lambda_{jT'(j)} \right) ; T' \in \mathcal{T}_\emptyset, \prod_{j=2}^n \lambda_{jT'(j)} \neq 0 \right\} .$$

Ainsi pour démontrer le lemme, il suffit de montrer que pour tout $T \in \mathcal{T}_S$ avec $\prod_{j=i}^n \lambda_{jT(j)} \neq 0$, il existe $T' \in \mathcal{T}_\emptyset$ avec $\prod_{j=2}^n \lambda_{jT'(j)} \neq 0$, tel que

$$d \left(\prod_{j=2}^{i-1} b_j \prod_{j=i}^n \lambda_{jT(j)} \right) \geq d \left(\prod_{j=2}^n \lambda_{jT'(j)} \right) .$$

Pour chaque $j \in \{1, \dots, i \Leftrightarrow 1\}$, notons

$$P_j = \{l \geq i ; \exists k \geq 1 \text{ t.q. } T^k(l) = j\}$$

et

$$\overline{P_j} = \{j\} \cup P_j .$$

Constatons que

$$\{P_j; 1 \leq j \leq i \Leftrightarrow 1\}$$

est une partition de $\{i, \dots, n\}$ et que

$$\{\overline{P}_j; 1 \leq j \leq i \Leftrightarrow 1\}$$

est une partition de $\{1, \dots, n\}$.

Montrons quitte à changer les indices, pour $2 \leq j \leq i \Leftrightarrow 1$, il existe

$$u_j \in \overline{P}_j \text{ et } v_j \in \bigcup_{l=1}^{j-1} \overline{P}_l$$

tels que

$$d \pi(u_j \lambda_{u_j v_j}) \leq k .$$

Comme Λ est irréductible au niveau k , il est évident pour $j = 2$. Supposons que le résultat est vrai pour $l \leq j \Leftrightarrow 1$. Comme il existe

$$u_j \in \bigcup_{l=j}^{i-1} \overline{P}_l \text{ et } v_j \in \bigcup_{l=1}^{j-1} \overline{P}_l$$

tels que

$$d \pi(u_j \lambda_{u_j v_j}) \leq k .$$

Quitte à changer les indices, on peut supposer $u_j \in \overline{P}_j$.

Posant $T^0(u_j) = u_j$, on définit

$$k_j = \min\{k \in \mathbb{N}; T^k(u_j) = j\} .$$

On a $k_j \geq 0$. Et $k_j = 0$ si et seulement si $u_j = j$.

On constate que

$$\prod_{j=i}^n \lambda_{jT(j)} = f \prod_{\substack{2 \leq j \leq i-1 \\ u_j \neq j}} \prod_{k=1}^{k_j} \lambda_{T^{k-1}(u_j)T^k(u_j)}$$

où

$$f = \prod_{\substack{i \leq j \leq n \\ j \notin U_i}} \lambda_{jT(j)}$$

avec

$$U_i = \{T^k(u_j); 2 \leq j \leq i \Leftrightarrow 1, 0 \leq k \leq k_j\}.$$

Par suite

$$\prod_{j=2}^{i-1} b_j \prod_{j=i}^n \lambda_{jT(j)} = f \left(\prod_{\substack{2 \leq j \leq i-1 \\ u_j=j}} b_j \right) \left(\prod_{\substack{2 \leq j \leq i-1 \\ u_j \neq j}} b_j \prod_{k=1}^{k_j} \lambda_{T^{k-1}(u_j)T^k(u_j)} \right).$$

Puisque pour tous $j, l, l' \in \{1, \dots, n\}$ avec $\lambda_{l'} \neq 0$, on a

$$d(\pi(j)b_j) \geq d(\pi(l)\lambda_{l'}),$$

on en déduit que

$$\begin{aligned} & d \left(\prod_{j=2}^{i-1} \pi(u_j) \prod_{j=2}^{i-1} b_j \prod_{j=i}^n \lambda_{jT(j)} \right) \\ &= d \left(f \prod_{\substack{2 \leq j \leq i-1 \\ u_j=j}} \pi(j)b_j \prod_{\substack{2 \leq j \leq i-1 \\ u_j \neq j}} \pi(u_j)b_j \prod_{k=1}^{k_j} \lambda_{T^{k-1}(u_j)T^k(u_j)} \right) \\ &= d \left(f \prod_{\substack{2 \leq j \leq i-1 \\ u_j=j}} \pi(j)b_j \prod_{\substack{2 \leq j \leq i-1 \\ u_j \neq j}} \pi(j)b_j \prod_{k=1}^{k_j} \lambda_{T^k(u_j)T^{k-1}(u_j)} \right) \text{ (réversibilité)} \\ &\geq d \left(f \prod_{\substack{2 \leq j \leq i-1 \\ u_j=j}} \pi(u_j)\lambda_{u_j v_j} \prod_{\substack{2 \leq j \leq i-1 \\ u_j \neq j}} \pi(u_j)\lambda_{u_j v_j} \prod_{k=1}^{k_j} \lambda_{T^k(u_j)T^{k-1}(u_j)} \right). \end{aligned}$$

D'où

$$d \left(\prod_{j=2}^{i-1} b_j \prod_{j=i}^n \lambda_{jT(j)} \right) \geq d \left(f \prod_{\substack{2 \leq j \leq i-1 \\ u_j=j}} \lambda_{u_j v_j} \prod_{\substack{2 \leq j \leq i-1 \\ u_j \neq j}} \lambda_{u_j v_j} \prod_{k=1}^{k_j} \lambda_{T^k(u_j)T^{k-1}(u_j)} \right).$$

Définissons l'application T' de $\{2, \dots, n\}$ dans $\{1, \dots, n\}$ de la façon suivante:

$$T'(j) = \begin{cases} T^{k-1}(u_l) & \text{si } j = T^k(u_l), 2 \leq l \leq i \Leftrightarrow 1, 1 \leq k \leq k_l \\ v_l & \text{si } j \in \{u_l; 2 \leq l \leq i \Leftrightarrow 1\} \\ T(j) & \text{si } j \in \{2, \dots, n\} \Leftrightarrow U_i . \end{cases}$$

On vérifie facilement que $T' \in \mathcal{T}_\emptyset$ et que

$$\prod_{j=2}^n \lambda_{jT'(j)} = f \prod_{\substack{2 \leq j \leq i-1 \\ u_j=j}} \lambda_{u_j v_j} \prod_{\substack{2 \leq j \leq i-1 \\ u_j \neq j}} \lambda_{u_j v_j} \prod_{k=1}^{k_j} \lambda_{T^k(u_j) T^{k-1}(u_j)} .$$

Ce qui termine la démonstration. □

Démontrons la proposition 3.1.11.

Démonstration. Comme A, B sont positives et analytiques en $\varepsilon = 0$ telles que

$$B \mathbb{1} \geq A \mathbb{1} ,$$

alors il existe une matrice F d'ordre $n \times m$ et analytique en $\varepsilon = 0$ telle que

$$A = BF .$$

Montrons la propriété (1).

Posons

$$C = \varepsilon (\Lambda \Leftrightarrow \varepsilon B)^{-1} B = (c_{ij})_{1 \leq i, j \leq n} .$$

Montrons pour tous $1 \leq i, j \leq n$,

$$\lim_{\varepsilon \searrow 0} \frac{c_{ij}}{c_{jj}} = 1 .$$

Sans perte de généralité, montrons

$$\lim_{\varepsilon \searrow 0} \frac{c_{21}}{c_{11}} = 1 .$$

On a

$$\lim_{\varepsilon \searrow 0} \frac{c_{21}}{c_{11}} = \lim_{\varepsilon \searrow 0} \frac{m_{12}}{m_{11}}$$

où m_{12} et m_{11} sont les cofacteurs (1, 2) et (1, 1) respectivement du déterminant $|\Lambda \Leftrightarrow \varepsilon B|$, c'est-à-dire

$$m_{12} = \Leftrightarrow \begin{vmatrix} \lambda_{21} & \lambda_{23} & \cdots & \lambda_{2n} \\ \lambda_{31} & \lambda_{33} \Leftrightarrow \varepsilon b_3 & \cdots & \lambda_{3n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{n1} & \lambda_{n3} & \cdots & \lambda_{nn} \Leftrightarrow \varepsilon b_n \end{vmatrix}$$

et

$$m_{11} = \begin{vmatrix} \lambda_{22} \Leftrightarrow \varepsilon b_2 & \lambda_{23} & \cdots & \lambda_{2n} \\ \lambda_{32} & \lambda_{33} \Leftrightarrow \varepsilon b_3 & \cdots & \lambda_{3n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{n2} & \lambda_{n3} & \cdots & \lambda_{nn} \Leftrightarrow \varepsilon b_n \end{vmatrix} .$$

Pour tout $S \subseteq \{3, \dots, n\}$, on note L'_S le déterminant obtenu en remplaçant les λ_{i2} dans L_S par λ_{i1} avec $i \in \bar{S}$. C'est-à-dire

$$L'_S = \begin{vmatrix} (\lambda_{ij}) & & & \\ & i \in \bar{S} & & \\ & j \in \{1, 3, \dots, n\} - S & & \end{vmatrix} .$$

Développons m_{11} et m_{12} de la façon suivante

$$m_{11} = \sum_{S \subseteq \{2, \dots, n\}} \prod_{i \in S} (\Leftrightarrow \varepsilon b_i) \cdot L_S ,$$

et

$$m_{12} = \Leftrightarrow \sum_{S \subseteq \{3, \dots, n\}} \prod_{i \in S} (\Leftrightarrow \varepsilon b_i) \cdot L'_S .$$

On constate que pour tout sous-ensemble S de $\{3, \dots, n\}$, L_S et L'_S ne peuvent être différents sur que la première colonne. En additionnant toutes les autres colonnes du déterminant L_S à la première, chaque coefficient sur la première colonne est plus petit que l'opposé de l'élément à la même position dans le déterminant L'_S . D'autre part, d'après la proposition 3.1.10, tout cofacteur de L_S et L'_S est négatif ou nul. En développant L_S et L'_S selon la première colonne, on voit que

$$|L'_S| \leq |L_S| . \tag{3.6}$$

En particulier

$$L'_\emptyset = \Leftrightarrow L_\emptyset .$$

D'après le lemme 3.1.2, on a pour tout $S \subseteq \{2, \dots, n\}$ avec $S \neq \emptyset$,

$$d \left(L_S \prod_{i \in S} b_i \right) \geq d(L_\emptyset) .$$

Ceci implique

$$\lim_{\varepsilon \searrow 0} \frac{m_{11}}{L_\emptyset} = 1 + \sum_{\substack{S \subseteq \{2, \dots, n\} \\ S \neq \emptyset}} \lim_{\varepsilon \searrow 0} \frac{\prod_{i \in S} (\Leftrightarrow \varepsilon b_i) \cdot L_S}{L_\emptyset} = 1 .$$

D'autre part pour tout $S \subseteq \{3, \dots, n\}$ avec $S \neq \emptyset$,

$$0 \leq \lim_{\varepsilon \searrow 0} \left| \frac{\prod_{i \in S} (\Leftrightarrow \varepsilon b_i) \cdot L'_S}{L_\emptyset} \right| \leq \lim_{\varepsilon \searrow 0} \left| \frac{\prod_{i \in S} (\Leftrightarrow \varepsilon b_i) \cdot L_S}{L_\emptyset} \right| = 0 ,$$

et par suite

$$\lim_{\varepsilon \searrow 0} \frac{m_{12}}{L_\emptyset} = \frac{\Leftrightarrow L'_\emptyset}{L_\emptyset} + \sum_{\substack{S \subseteq \{3, \dots, n\} \\ S \neq \emptyset}} \lim_{\varepsilon \searrow 0} \frac{\prod_{i \in S} (\Leftrightarrow \varepsilon b_i) \cdot L'_S}{L_\emptyset} = 1 .$$

D'où

$$\lim_{\varepsilon \searrow 0} \frac{c_{21}}{c_{11}} = \lim_{\varepsilon \searrow 0} \frac{m_{12}}{m_{11}} = \lim_{\varepsilon \searrow 0} \frac{m_{12}/L_\emptyset}{m_{11}/L_\emptyset} = 1 .$$

Posons

$$v = (c_{1i})_{1 \leq i \leq n} F$$

et

$$O = \frac{1}{\varepsilon} (c_{ij} \Leftrightarrow c_{1j})_{1 \leq i, j \leq n} F .$$

Alors

$$\varepsilon (\Lambda \Leftrightarrow \varepsilon B)^{-1} A = \mathbb{1} v + \varepsilon O .$$

D'après la proposition 3.1.10, C est analytique en $\varepsilon = 0$, on en déduit que O l'est aussi. D'où le résultat.

Montrons la propriété (2).

Posons maintenant

$$C = \varepsilon (\Lambda \Leftrightarrow D \Leftrightarrow \varepsilon B)^{-1} B = (c_{ij})_{1 \leq i, j \leq n} .$$

Selon (1) de la proposition 3.1.10, pour tous $1 \leq i, j \leq n$, on a

$$c_{jj} \leq c_{ij} \leq 0 .$$

Donc pour montrer $\lim_{\varepsilon \searrow 0} C = 0$, il suffit de montrer que

$$\lim_{\varepsilon \searrow 0} c_{jj} = 0 ,$$

et par symétrie il suffit de montrer que

$$\lim_{\varepsilon \searrow 0} c_{11} = 0 .$$

Posons maintenant m_{11} le cofacteur $(1, 1)$ du déterminant $|\Lambda \Leftrightarrow D \Leftrightarrow \varepsilon B|$, i.e.

$$m_{11} = \begin{vmatrix} \lambda_{22} \Leftrightarrow d_2 \Leftrightarrow \varepsilon b_2 & \lambda_{23} & \cdots & \lambda_{2n} \\ \lambda_{32} & \lambda_{33} \Leftrightarrow d_3 \Leftrightarrow \varepsilon b_3 & \cdots & \lambda_{3n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{n2} & \lambda_{n3} & \cdots & \lambda_{nn} \Leftrightarrow d_n \Leftrightarrow \varepsilon b_n \end{vmatrix} .$$

Alors

$$c_{11} = \frac{\varepsilon b_1 m_{11}}{|\Lambda \Leftrightarrow D \Leftrightarrow \varepsilon B|} .$$

Définissons le générateur $\tilde{\Lambda} = (\tilde{\lambda}_{ij})$ sur $\{0, \dots, n\}$ tel que Λ est sa restriction sur $\{1, \dots, n\}$ et que le taux de transition de i à 0, noté λ_{i0} est d_i et le taux de transition de 0 à i , noté λ_{0i} , est $\pi(i)d_i$ pour $1 \leq i \leq n$. C'est-à-dire

$$\tilde{\Lambda} = \begin{pmatrix} \Leftrightarrow \sum_{i=1}^n \pi(i)d_i & \pi(1)d_1 \cdots \pi(n)d_n \\ d_1 & \\ \vdots & \Lambda \Leftrightarrow D \\ d_n & \end{pmatrix} ..$$

Evidemment $\tilde{\Lambda}$ est irréductible et réversible. Notons pour tout $S \subseteq \{1, \dots, n\}$,

$$L_S(\tilde{\Lambda}) = \left| \begin{matrix} (\tilde{\lambda}_{ij})_{i \in \{1, \dots, n\} - S} \\ j \in \{1, \dots, n\} - S \end{matrix} \right|.$$

On a

$$b_1 m_{11} = \sum_{\substack{S \subseteq \{1, \dots, n\} \\ 1 \in S}} (\Leftrightarrow \varepsilon)^{\#S-1} \prod_{i \in S} b_i \cdot L_S(\tilde{\Lambda}).$$

Appliquant le lemme 3.1.1 au déterminant $|\tilde{\Lambda}|$, on remarque que tout terme dans la sommation ci-dessus est de même signe. D'autre part, $\tilde{\Lambda}$ est évidemment irréductible au niveau k . D'après le lemme 3.1.2, pour tout $S \subseteq \{1, \dots, n\}$ avec $1 \in S$,

$$d\left(\prod_{i \in S} b_i \cdot L_S(\tilde{\Lambda})\right) \geq d(L_\emptyset(\tilde{\Lambda})),$$

donc

$$d(b_1 m_{11}) \geq d(L_\emptyset(\tilde{\Lambda})).$$

D'autre part, on a

$$\begin{aligned} |\Lambda \Leftrightarrow D \Leftrightarrow \varepsilon B| &= \sum_{S \subseteq \{1, \dots, n\}} (\Leftrightarrow \varepsilon)^{\#S} \prod_{i \in S} b_i \cdot L_S(\tilde{\Lambda}) \\ &= L_\emptyset(\tilde{\Lambda}) + \sum_{\substack{S \subseteq \{1, \dots, n\} \\ S \neq \emptyset}} (\Leftrightarrow \varepsilon)^{\#S} \prod_{i \in S} b_i \cdot L_S(\tilde{\Lambda}). \end{aligned}$$

De plus pour tout $S \subseteq \{1, \dots, n\}$ avec $S \neq \emptyset$, d'après le lemme 3.1.2, on a

$$d\left(\prod_{i \in S} b_i \cdot L_S(\tilde{\Lambda})\right) \geq d(L_\emptyset(\tilde{\Lambda})).$$

On en déduit

$$d(|\Lambda \Leftrightarrow D \Leftrightarrow \varepsilon B|) = d(L_\emptyset(\tilde{\Lambda})).$$

Par conséquent,

$$d(d_1 m_{11}) \geq d(|\Lambda \Leftrightarrow D \Leftrightarrow \varepsilon B|).$$

Il vient que

$$\lim_{\varepsilon \searrow 0} c_{11} = \lim_{\varepsilon \searrow 0} \frac{\varepsilon b_1 m_{11}}{|\Lambda \Leftrightarrow D \Leftrightarrow \varepsilon B|} = 0.$$

D'où

$$\lim_{\varepsilon \searrow 0} \varepsilon (\Lambda \Leftrightarrow D \Leftrightarrow \varepsilon B)^{-1} A = \varepsilon (\Lambda \Leftrightarrow D \Leftrightarrow \varepsilon B)^{-1} BF = 0 .$$

Montrons la propriété (3).

Posons maintenant

$$C = (c_{ij})_{1 \leq i, j \leq n} = (\Lambda \Leftrightarrow \varepsilon B)^{-1} \text{ et } X = (x_i)_{1 \leq i \leq n} .$$

Montrons d'abord que pour $\forall 1 \leq i, l \leq n$

$$\varepsilon^{k+1} \sum_{j=1}^n (c_{ij} \Leftrightarrow c_{lj}) x_j = O(\varepsilon) .$$

D'après (1) de la proposition 3.1.10,

$$c_{jj} \leq c_{ij} \leq 0 ,$$

d'autre part X est continu selon la proposition 3.1.7, alors il suffit de montrer que

$$\varepsilon^{k+1} (c_{ij} \Leftrightarrow c_{jj}) = O(\varepsilon) .$$

Sans perte de généralité, montrons

$$\varepsilon^{k+1} (c_{21} \Leftrightarrow c_{11}) = O(\varepsilon) .$$

Définissons le générateur $\bar{\Lambda}$ sur $\{0, \dots, n\}$ tel que $\Lambda \Leftrightarrow \varepsilon B$ est sa restriction sur $\{1, \dots, n\}$ et que le taux de transition de i à 0, noté λ_{i0} , est εb_i et le taux de transition de 0 à i , noté λ_{i0} , est $\varepsilon b_i \pi(i)$ pour $1 \leq i \leq n$. Evidemment $\bar{\Lambda}$ est irréductible et réversible. Posons \mathcal{T}'_{\emptyset} l'ensemble des applications T de $\{1, \dots, n\}$ dans $\{0, 1, \dots, n\}$ telles qu'on ait $T(S) \neq S$ pour tout $S \subseteq \{1, \dots, n\}$ différent de \emptyset . D'après le lemme 3.1.1, on a

$$|\Lambda \Leftrightarrow \varepsilon B| = (\Leftrightarrow 1)^{n+1} \sum_{T \in \mathcal{T}'_{\emptyset}} \prod_{i=1}^n \lambda_{iT(i)} .$$

D'autre part

$$c_{11} = |\Lambda \Leftrightarrow \varepsilon B|^{-1} \begin{vmatrix} \lambda_{22} \Leftrightarrow \varepsilon b_2 & \lambda_{23} & \cdots & \lambda_{2n} \\ \lambda_{32} & \lambda_{33} \Leftrightarrow \varepsilon b_3 & \cdots & \lambda_{3n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{n2} & \lambda_{n3} & \cdots & \lambda_{nn} \Leftrightarrow \varepsilon b_n \end{vmatrix}$$

$$= |\Lambda \Leftrightarrow \varepsilon B|^{-1} \sum_{S \subseteq \{2, \dots, n\}} \prod_{i \in S} (\Leftrightarrow \varepsilon b_i) \cdot L_S$$

et

$$\begin{aligned} c_{21} &= \Leftrightarrow |\Lambda \Leftrightarrow \varepsilon B|^{-1} \begin{vmatrix} \lambda_{11} \Leftrightarrow \varepsilon b_1 & \lambda_{12} & \cdots & \lambda_{1n} \\ \lambda_{21} & \lambda_{22} \Leftrightarrow \varepsilon b_2 & \cdots & \lambda_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{n1} & \lambda_{n2} & \cdots & \lambda_{nn} \Leftrightarrow \varepsilon b_n \end{vmatrix} \\ &= \Leftrightarrow |\Lambda \Leftrightarrow \varepsilon B|^{-1} \sum_{S \subseteq \{3, \dots, n\}} \prod_{i \in S} (\Leftrightarrow \varepsilon b_i) \cdot L'_S \end{aligned}$$

où L_S, L'_S sont définies comme précédemment. Puisque $L_\emptyset = \Leftrightarrow L'_\emptyset$, on a

$$\begin{aligned} c_{11} \Leftrightarrow c_{21} &= |\Lambda \Leftrightarrow \varepsilon B|^{-1} \times \\ &\left[\sum_{\emptyset \neq S \subseteq \{2, \dots, n\}} \prod_{i \in S} (\Leftrightarrow \varepsilon b_i) \cdot L_S + \sum_{\emptyset \neq S \subseteq \{3, \dots, n\}} \prod_{i \in S} (\Leftrightarrow \varepsilon b_i) \cdot L'_S \right]. \end{aligned}$$

Compte tenu du lemme 3.1.1, dans les sommations ci-dessus, chaque terme a $(\Leftrightarrow 1)^n$ comme signe. On en déduit que

$$\begin{aligned} &d(c_{11} \Leftrightarrow c_{21}) \\ &= \inf_{\emptyset \neq S \subseteq \{2, \dots, n\}} \left\{ d \left(|\Lambda \Leftrightarrow \varepsilon B|^{-1} \prod_{i \in S_1} (\Leftrightarrow \varepsilon b_i) \cdot L_{S_1} \right) \right\} \\ &\quad \bigwedge \inf_{\emptyset \neq S \subseteq \{3, \dots, n\}} \left\{ d \left(|\Lambda \Leftrightarrow \varepsilon B|^{-1} \prod_{i \in S} (\Leftrightarrow \varepsilon b_i) \cdot L'_S \right) \right\}. \end{aligned}$$

Or pour tout $S \subseteq \{3, \dots, n\}$, on a (3.6), donc

$$d(c_{11} \Leftrightarrow c_{21}) = \inf_{\emptyset \neq S \subseteq \{2, \dots, n\}} \left\{ d \left(|\Lambda \Leftrightarrow \varepsilon B|^{-1} \prod_{i \in S} (\Leftrightarrow \varepsilon b_i) \cdot L_S \right) \right\}.$$

En plus, selon le lemme 3.1.1, pour tout $S \subseteq \{2, \dots, n\}$ différent de \emptyset , on a

$$L_S = (\Leftrightarrow 1)^{n-\#S} \sum_{T \in \mathcal{T}_S} \prod_{i \in \overline{S}} \lambda_{iT(i)},$$

ainsi

$$d(c_{11} \Leftrightarrow c_{21}) = \inf_{\emptyset \neq S \subseteq \{2, \dots, n\}} \left\{ d \left(|\Lambda \Leftrightarrow \varepsilon B|^{-1} \prod_{i \in S} \lambda_{i0} \prod_{i \in \overline{S}} \lambda_{iT(i)} \right); T \in \mathcal{T}_S \right\}.$$

Donc pour montrer

$$d \left(\varepsilon^{k+1} (c_{11} \Leftrightarrow c_{21}) \right) = O(\varepsilon) , \quad (3.7)$$

il suffit de montrer pour tous $\emptyset \neq S \subseteq \{2, \dots, n\}$, $T_1 \in \mathcal{T}_S$, il existe un $T_2 \in \mathcal{T}'_\emptyset$ tels que

$$d \left(\prod_{i \in \{1, \dots, n\}} \lambda_{iT_2(i)} \right) \leq d \left(\prod_{i \in S} \lambda_{i0} \prod_{i \in \bar{S}} \lambda_{iT_1(i)} \right) + k .$$

Comme Λ est irréductible au niveau k , il existe $j \in \{1, \dots, n\}$ tel que

$$d(\pi(1)\lambda_{1j}) \leq k .$$

Pour tous $\emptyset \neq S \subseteq \{2, \dots, n\}$, $T_1 \in \mathcal{T}_S$, définissons T_2 une application de $\{1, \dots, n\}$ à $\{0, \dots, n\}$ de la façon suivante, pour tout $i \in \{1, \dots, n\}$

$$T_2(i) = \begin{cases} 0 & \text{si } i \in S \\ T_1(i) & \text{si } i \in \bar{S} \\ j & \text{si } i = 1 . \end{cases}$$

On vérifie que $T_2 \in \mathcal{T}'_\emptyset$, en plus

$$\begin{aligned} d \left(\prod_{i \in \{1, \dots, n\}} \lambda_{iT_2(i)} \right) &= d \left(\prod_{i \in S} \lambda_{i0} \prod_{i \in \bar{S}} \lambda_{iT_1(i)} \right) + d(\lambda_{1j}) \\ &\leq d \left(\prod_{i \in S} \lambda_{i0} \prod_{i \in \bar{S}} \lambda_{iT_1(i)} \right) + k . \end{aligned}$$

On en déduit (3.7).

Posons

$$c = \varepsilon^{k+1} \sum_{i=1}^n c_{1i} x_i$$

et

$$Y = \left(\varepsilon^k \sum_{j=1}^n (c_{ij} \Leftrightarrow c_{1j}) x_j \right)_{1 \leq i \leq n} .$$

Alors Y est continu et

$$\varepsilon^{k+1} (\Lambda \Leftrightarrow \varepsilon B)^{-1} X = c \mathbf{1} + \varepsilon Y .$$

□

3.1.4 Démonstration du théorème A

Le lemme suivant est crucial pour la démonstration du théorème A. Ce lemme montre que le vecteur propre associé à une valeur propre tendant vers 0, disons d'ordre $O(\varepsilon^k)$, est quasiment constant.

Lemme 3.1.3 *Soit $\varepsilon^k \lambda(\varepsilon)$ une valeur propre du générateur Λ avec $k \geq 1$. Soit X un vecteur propre associé. Pour tout sous-ensemble S de l'espace d'états E , on note X_S la restriction de X sur S . Soit α_k une classe minimale de \mathcal{F}_k . Alors il existe un ensemble de fonctions continues*

$$\{c_{\alpha_i} ; \alpha_i \in F_i(\alpha_k), 1 \leq i \leq k\}$$

tel que pour tout $\alpha_i \in F_i(\alpha_k)$, en posant $\alpha_j = \mathcal{E}_j(\alpha_i)$ pour $i \leq j \leq k$,

$$X_{\alpha_i} = \left(\sum_{l=0}^{k-i} \varepsilon^l c_{\alpha_{k-l}} \right) \mathbb{1}_{\alpha_i} + \varepsilon^{k-i+1} Y_{\alpha_i}$$

où Y_{α_i} est un vecteur continu sur α_i et tel que

$$X_{U_{k+1}} = \left(\mathcal{N}_{U_{k+1}} \Leftrightarrow D_{U_{k+1}} \right)^{-1} \sum_{\alpha_k \in \mathcal{F}_k} c_{\alpha_k} G_{U_{k+1}}^{\alpha_k} \mathbb{1}_{\alpha_k} + O(\varepsilon/U_{k+1}).$$

Démonstration. Nous avons

$$\Lambda X = \varepsilon^k \lambda(\varepsilon) X \quad (3.8)$$

Prenant les lignes indicées par α_k , on a

$$\Lambda_{\alpha_k} X_{\alpha_k} \Leftrightarrow \varepsilon B_{\alpha_k}^{O_{k+1}} X_{\alpha_k} + \varepsilon A_{\alpha_k}^{O_{k+1}} X_{O_{k+1}} = \varepsilon^k \lambda(\varepsilon) X_{\alpha_k}.$$

Il vient que

$$X_{\alpha_k} = \Leftrightarrow \varepsilon \left(\mathcal{N}_{\alpha_k} \Leftrightarrow \varepsilon B_{\alpha_k}^{O_{k+1}} \right)^{-1} A_{\alpha_k}^{O_{k+1}} X_{O_{k+1}} + \left(\mathcal{N}_{\alpha_k} \Leftrightarrow \varepsilon B_{\alpha_k}^{O_{k+1}} \right)^{-1} \varepsilon^k \lambda(\varepsilon) X_{\alpha_k}.$$

D'après la proposition 3.1.4, Λ_{α_k} est irréductible au niveau $k \Leftrightarrow 1$. D'autre part, d'après a) de la proposition 3.1.3, pour tous $a \in \alpha_k, b \in O_{k+1}$

$$\begin{aligned} \delta_k[a, b] &= d(\Pi(a)) + d(\lambda_{ab}) \Leftrightarrow B(\alpha_k) \\ &= d(\Pi_{\alpha_k}(a) \lambda_{ab}) \\ &\geq k. \end{aligned}$$

En plus, X_{α_k} est continue selon la proposition 3.1.7 et

$$\varepsilon \Lambda_{\alpha_k} \Leftrightarrow \varepsilon (B_{\alpha_k}^{O_{k+1}})^{-1} A_{\alpha_k}^{O_{k+1}} \quad (3.9)$$

est analytique en $\varepsilon = 0$ selon la proposition 3.1.10, donc la limite du vecteur

$$\varepsilon \Lambda_{\alpha_k} \Leftrightarrow \varepsilon (B_{\alpha_k}^{O_{k+1}})^{-1} A_{\alpha_k}^{O_{k+1}} X_{O_{k+1}} \quad (3.10)$$

existe. Appliquons la propriété (1) de la proposition 3.1.11 à la matrice (3.9) et la propriété (3) au vecteur (3.10) et puis au vecteur

$$\Lambda_{\alpha_k} \Leftrightarrow \varepsilon (B_{\alpha_k}^{O_{k+1}})^{-1} \varepsilon^k \lambda(\varepsilon) X_{\alpha_k} .$$

On conclue qu'il existe une fonction continue qu'on note c_{α_k} et un vecteur continu Y_{α_k} sur α_k tels que

$$Y'_{\alpha_k} = c_{\alpha_k} \mathbb{1}_{\alpha_k} + \varepsilon Y_{\alpha_k} .$$

Supposons qu' il existe un ensemble de fonctions continues

$$\{c_{\alpha_j} ; \alpha_j \in F_j(\alpha_k), i + 1 \leq j \leq k\}$$

tel que pour tout $\alpha_j \in F_j(\alpha_k)$, en posant $\alpha_l = \mathcal{E}_l(\alpha_j)$ pour $j \leq l \leq k$,

$$X_{\alpha_j} = \left(\sum_{l=0}^{k-j} \varepsilon^l c_{\alpha_{k-l}} \right) \mathbb{1}_{\alpha_j} + \varepsilon^{k-j+1} Y_{\alpha_j}$$

où Y_{α_j} est un vecteur continu sur α_j . Soient $\alpha_i \in F_i(\alpha_k)$ et $\alpha_j = \mathcal{E}_j(\alpha_i)$ pour $i \leq j$, comme $\alpha_i \subseteq \alpha_{i+1}$, on a

$$X_{\alpha_i} = \left(\sum_{l=0}^{k-i-1} \varepsilon^l c_{\alpha_{k-l}} \right) \mathbb{1}_{\alpha_i} + \varepsilon^{k-i} Y'_{\alpha_i} \quad (3.11)$$

où Y'_{α_i} est la restriction de $Y_{\alpha_{i+1}}$ sur α_i . Prenant les lignes indicées par α_i de (3.8) et notant

$$O_{i+l} = O_{i+l}[\alpha_i]$$

pour $1 \leq l \leq k \Leftrightarrow i + 1$, on a

$$\Lambda_{\alpha_i} X_{\alpha_i} + \sum_{l=1}^{k-i+1} \varepsilon^l \left(B_{\alpha_i}^{O_{i+l}} X_{\alpha_i} + A_{\alpha_i}^{O_{i+l}} X_{O_{i+l}} \right) = \varepsilon^k \lambda(\varepsilon) X_{\alpha_i} .$$

Compte tenu de (3.11), on a

$$\Lambda_{\alpha_i} X_{\alpha_i} = \left(\sum_{l=0}^{k-i-1} \varepsilon^l c_{\alpha_{k-l}} \right) \Lambda_{\alpha_i} \mathbb{1}_{\alpha_i} + \varepsilon^{k-i} \Lambda_{\alpha_i} Y'_{\alpha_i} = \varepsilon^{k-i} \Lambda_{\alpha_i} Y'_{\alpha_i}.$$

Pour $i+l < k+1$,

$$\begin{aligned} & \Leftrightarrow B_{\alpha_i}^{O_{i+l}} X_{\alpha_i} + A_{\alpha_i}^{O_{i+l}} X_{O_{i+l}} \\ &= \left(\sum_{j=0}^{k-i-l} \varepsilon^j c_{\alpha_{k-j}} \right) \left(\Leftrightarrow B_{\alpha_i}^{O_{i+l}} \mathbb{1}_{\alpha_i} + A_{\alpha_i}^{O_{i+l}} \mathbb{1}_{O_{i+l}} \right) \\ & \Leftrightarrow \left(\sum_{j=k-i-l+1}^{k-i-1} \varepsilon^j c_{\alpha_{k-j}} \right) B_{\alpha_i}^{O_{i+l}} \mathbb{1}_{\alpha_i} \Leftrightarrow \varepsilon^{k-i} B_{\alpha_i}^{O_{i+l}} Y'_{\alpha_i} + \varepsilon^{k-i-l+1} A_{\alpha_i}^{O_{i+l}} Y_{O_{i+l}} \\ &= \Leftrightarrow \sum_{j=k-i-l+1}^{k-i-1} \varepsilon^j c_{\alpha_{k-j}} B_{\alpha_i}^{O_{i+l}} \mathbb{1}_{\alpha_i} \Leftrightarrow \varepsilon^{k-i} B_{\alpha_i}^{O_{i+l}} Y'_{\alpha_i} + \varepsilon^{k-i-l+1} A_{\alpha_i}^{O_{i+l}} Y_{O_{i+l}}, \end{aligned}$$

et pour $i+l = k+1$,

$$\begin{aligned} & \Leftrightarrow B_{\alpha_i}^{O_{k+1}} X_{\alpha_i} + A_{\alpha_i}^{O_{k+1}} X_{O_{k+1}} \\ &= \Leftrightarrow \sum_{j=0}^{k-i-1} \varepsilon^j c_{\alpha_{k-j}} B_{\alpha_i}^{O_{k+1}} \mathbb{1}_{\alpha_i} \Leftrightarrow \varepsilon^{k-i} B_{\alpha_i}^{O_{k+1}} Y'_{\alpha_i} + A_{\alpha_i}^{O_{k+1}} Y_{O_{k+1}} \end{aligned}$$

où $Y_{O_{i+l}}$ est le vecteur restreint de $Y_{\alpha_{i+l}}$ sur O_{i+l} pour $i+l \leq k+1$. On en déduit

$$\begin{aligned} & \varepsilon^{k-i} \left[\cancel{\Lambda_{\alpha_i} \Leftrightarrow \varepsilon B_{\alpha_i}^{O_{i+1}} Y'_{\alpha_i}} \Leftrightarrow \varepsilon \sum_{l=1}^{k-i+1} f_l B_{\alpha_i}^{O_{i+l}} Y_{\alpha_i}^l + \varepsilon \sum_{l=1}^{k-i+1} h_l A_{\alpha_i}^{O_{i+l}} Y_{O_{i+l}} \right] \\ &= \varepsilon^k \lambda(\varepsilon) X_{\alpha_i} \end{aligned}$$

où les f_l, h_l sont des fonctions continues en $\varepsilon = 0$ et les $Y_{\alpha_i}^l$ sont des vecteurs continus définis sur α_i . Par suite,

$$\begin{aligned} Y'_{\alpha_i} &= \sum_{l=1}^{k-i+1} f_l \varepsilon \cancel{\Lambda_{\alpha_i} \Leftrightarrow \varepsilon B_{\alpha_i}^{O_{i+1}}}^{-1} B_{\alpha_i}^{O_{i+l}} Y_{\alpha_i}^l \\ & \Leftrightarrow \sum_{l=1}^{k-i+1} h_l \varepsilon \cancel{\Lambda_{\alpha_i} \Leftrightarrow \varepsilon B_{\alpha_i}^{O_{i+1}}}^{-1} A_{\alpha_i}^{O_{i+l}} Y_{O_{i+l}} \\ & + \lambda(\varepsilon) \varepsilon^i \cancel{\Lambda_{\alpha_i} \Leftrightarrow \varepsilon B_{\alpha_i}^{O_{i+1}}}^{-1} X_{\alpha_i}. \end{aligned}$$

Par un raisonnement analogue au précédent, on peut appliquer la propriété (1) de la proposition 3.1.11 aux matrices

$$\varepsilon \mathcal{N}_{\alpha_i} \Leftrightarrow \varepsilon B_{\alpha_i}^{O_{i+1}})^{-1} B_{\alpha_i}^{O_{i+1}} , \quad \varepsilon \mathcal{N}_{\alpha_i} \Leftrightarrow \varepsilon B_{\alpha_i}^{O_{i+1}})^{-1} A_{\alpha_i}^{O_{i+1}}$$

et la propriété (3) au vecteur

$$\varepsilon^i \mathcal{N}_{\alpha_i} \Leftrightarrow \varepsilon B_{\alpha_i}^{O_{i+1}})^{-1} X_{\alpha_i} ,$$

conclusions qu'il existe une fonction continue qu'on note c_{α_i} et un vecteur continu Y_{α_i} sur α_i tels que

$$X_{\alpha_i} = c_{\alpha_i} \mathbb{1}_{\alpha_i} + \varepsilon Y_{\alpha_i} .$$

D'où

$$X_{\alpha_i} = \left(\sum_{l=0}^{k-i} \varepsilon^l c_{\alpha_{k-l}} \right) \mathbb{1}_{\alpha_i} + \varepsilon^{k-i+1} Y_{\alpha_i} .$$

Il reste à montrer

$$X_{U_{k+1}} = \mathcal{N}_{U_{k+1}} \Leftrightarrow D_{U_{k+1}})^{-1} \sum_{\alpha_k \in \mathcal{F}_k} G_{U_{k+1}}^{\alpha_k} \mathbb{1}_{\alpha_k} c_{\alpha_k} + O(\varepsilon^{U_{k+1}}) .$$

Prenant les lignes indicées par U_{k+1} de (3.8), on a

$$\mathcal{N}_{U_{k+1}} \Leftrightarrow D_{U_{k+1}}) X_{U_{k+1}} + \sum_{\alpha_k \in \mathcal{F}_k} G_{U_{k+1}}^{\alpha_k} X_{\alpha_k} = \varepsilon^k \lambda(\varepsilon) X_{U_{k+1}} .$$

D'où

$$\begin{aligned} X_{U_{k+1}} &= \mathcal{N}_{U_{k+1}} \Leftrightarrow D_{U_{k+1}})^{-1} \sum_{\alpha_k \in \mathcal{F}_k} G_{U_{k+1}}^{\alpha_k} X_{\alpha_k} + O(\varepsilon^{U_{k+1}}) \\ &= \mathcal{N}_{U_{k+1}} \Leftrightarrow D_{U_{k+1}})^{-1} \sum_{\alpha_k \in \mathcal{F}_k} c_{\alpha_k} G_{U_{k+1}}^{\alpha_k} \mathbb{1}_{\alpha_k} + O(\varepsilon^{U_{k+1}}) . \end{aligned}$$

□

Pour $k \geq 0$, notons $S_k(\Lambda)$ le vecteur composé des valeurs propres d'ordre $O(\varepsilon^k)$ de Λ avec multiplicité éventuelle. Le théorème A suivant donne un générateur d'ordre $\#S_k(\Lambda)$ conservant les équivalents des valeurs propres d'ordre $O(\varepsilon^k)$.

Théorème A Soit $1 \leq k \leq K$.

- (1) Pour toutes classes $\alpha_k \neq \beta_k$ de \mathcal{F}_k , définissons le taux de transition de α_k à β_k comme suit,

$$\lambda_{\alpha_k \beta_k} = \frac{1}{\varepsilon^k} \Pi_{\alpha_k} \left[G_{\alpha_k}^{\beta_k} \mathbb{1}_{\beta_k} \Leftrightarrow G_{\alpha_k}^{U_{k+1}} \left(N_{U_{k+1}} \Leftrightarrow D_{U_{k+1}} \right)^{-1} G_{U_{k+1}}^{\beta_k} \mathbb{1}_{\beta_k} \right]$$

Alors $\lim_{\varepsilon \searrow 0} \lambda_{\alpha_k \beta_k}$ existe. Soit $l \geq k+1$ l'entier le plus petit tel que α_k soit contenu dans une classe minimale d'ordre l , notée α_l . Alors si

$$\beta_k \not\subseteq \alpha_l$$

on a en plus

$$\lim_{\varepsilon \searrow 0} \lambda_{\alpha_k \beta_k} = 0 .$$

- (2) Pour tout a de U_{k+1} , notons $p_a^{\beta_k}$ la probabilité que le processus Λ , partant de a , atteigne β_k pour la première fois sans visiter les autres classes de \mathcal{F}_k . Soit $P_{U_{k+1}}^{\beta_k}$ le vecteur colonne composé des $p_a^{\beta_k}$ pour a de U_{k+1} . Alors

$$\lambda_{\alpha_k \beta_k} = \frac{1}{\varepsilon^k} \Pi_{\alpha_k} \left(G_{\alpha_k}^{\beta_k} \mathbb{1}_{\beta_k} + G_{\alpha_k}^{U_{k+1}} P_{U_{k+1}}^{\beta_k} \right) .$$

- (3) Soit Λ_k le générateur sur \mathcal{F}_k composé des taux de transition $\lambda_{\alpha_k \beta_k}$. Alors

$$\lim_{\varepsilon \searrow 0} S_0(\Lambda_k) = \lim_{\varepsilon \searrow 0} \frac{1}{\varepsilon^k} S_k(\Lambda) .$$

Démonstration. Montrons la propriété (1).

Montrons d'abord que $\lim_{\varepsilon \searrow 0} \lambda_{\alpha_k \beta_k}$ existe. D'après b) de la proposition 3.1.3,

$$\Pi_{\alpha_k} G_{\alpha_k}^{U_{k+1}} = O \left(\varepsilon^k \right) .$$

Donc

$$\lim_{\varepsilon \searrow 0} \frac{1}{\varepsilon^k} \Pi_{\alpha_k} G_{\alpha_k}^{\beta_k} \mathbb{1}_{\beta_k} \text{ et } \lim_{\varepsilon \searrow 0} \frac{1}{\varepsilon^k} \Pi_{\alpha_k} G_{\alpha_k}^{U_{k+1}}$$

existent. D'après le corollaire 3.1.1, pour toute classe β_k de \mathcal{F}_k ,

$$\lim_{\varepsilon \searrow 0} \left(N_{U_{k+1}} \Leftrightarrow D_{U_{k+1}} \right)^{-1} G_{U_{k+1}}^{\beta_k} \mathbb{1}_{\beta_k}$$

existe. D'où l'existence de $\lim_{\varepsilon \searrow 0} \lambda_{\alpha_k \beta_k}$.

Soit $l \geq k + 1$ l'entier le plus petit tel que α_k soit contenu dans une classe minimale d'ordre l , notée α_l . Si

$$\beta_k \not\subseteq \alpha_l$$

montrons

$$\lim_{\varepsilon \searrow 0} \lambda_{\alpha_k \beta_k} = 0 .$$

D'après a) de la proposition 3.1.3, on a

$$\forall a \in \alpha_k, b \in \beta_k, \delta_k[a, b] \geq l \geq k + 1 .$$

Par conséquent,

$$\lim_{\varepsilon \searrow 0} \frac{1}{\varepsilon^k} \Pi_{\alpha_k} G_{\alpha_k}^{\beta_k} \mathbf{1}_{\beta_k} = 0 .$$

Il reste à montrer

$$\lim_{\varepsilon \searrow 0} \frac{1}{\varepsilon^k} \Pi_{\alpha_k} G_{\alpha_k}^{U_{k+1}} \not\subseteq_{U_{k+1}} \Leftrightarrow D_{U_{k+1}})^{-1} G_{U_{k+1}}^{\beta_k} \mathbf{1}_{\beta_k} = 0 .$$

Notons

$$\alpha_{k+1} = \mathcal{E}_{k+1}(\alpha_k) \text{ et } \beta_{k+1} = \mathcal{E}_{k+1}(\beta_k) .$$

Posons

$$U_{k+1}^1 = \alpha_{k+1} \Leftrightarrow \bigcup_{\substack{\gamma_k \subseteq \alpha_{k+1} \\ \gamma_k \in \mathcal{F}_k}} \gamma_k \text{ et } U_{k+1}^2 = U_{k+1} \Leftrightarrow U_{k+1}^1 .$$

Alors

$$G_{\alpha_k}^{U_{k+1}} = \begin{pmatrix} G_{\alpha_k}^{U_{k+1}^1} & G_{\alpha_k}^{U_{k+1}^2} \end{pmatrix}$$

et

$$G_{U_{k+1}}^{\beta_k} = \begin{pmatrix} G_{U_{k+1}^1}^{\beta_k} \\ G_{U_{k+1}^2}^{\beta_k} \end{pmatrix} .$$

Soit $D_{U_{k+1}^1}, D_{U_{k+1}^2}$ deux matrices diagonales sur U_{k+1}^1 et U_{k+1}^2 respectivement telles que

$$D_{U_{k+1}} = \begin{pmatrix} D_{U_{k+1}^1} & (0) \\ (0) & D_{U_{k+1}^2} \end{pmatrix} .$$

Alors

$$\begin{aligned}
&= \left(\begin{array}{cc} \mathcal{N}_{U_{k+1}} \Leftrightarrow D_{U_{k+1}} & \varepsilon A_{U_{k+1}}^{U_{k+1}^2} \\ \Lambda_{U_{k+1}}^{U_{k+1}^1} \Leftrightarrow \varepsilon B_{U_{k+1}}^{U_{k+1}^2} \Leftrightarrow D_{U_{k+1}}^{U_{k+1}^1} & \Lambda_{U_{k+1}}^{U_{k+1}^2} \Leftrightarrow D_{U_{k+1}}^{U_{k+1}^1} \Leftrightarrow D_{U_{k+1}}^{U_{k+1}^2} \\ G_{U_{k+1}}^{U_{k+1}^1} & \end{array} \right)^{-1} \\
&= \begin{pmatrix} M_1 & M_2 \\ M_3 & M_4 \end{pmatrix}
\end{aligned}$$

où

$$\begin{aligned}
M_1 &= \left[\left(\Lambda_{U_{k+1}}^{U_{k+1}^1} \Leftrightarrow \varepsilon B_{U_{k+1}}^{U_{k+1}^2} \Leftrightarrow D_{U_{k+1}}^{U_{k+1}^1} \right) \Leftrightarrow \right. \\
&\quad \left. \varepsilon A_{U_{k+1}}^{U_{k+1}^2} \left(\Lambda_{U_{k+1}}^{U_{k+1}^2} \Leftrightarrow D_{U_{k+1}}^{U_{k+1}^1} \Leftrightarrow D_{U_{k+1}}^{U_{k+1}^2} \right)^{-1} G_{U_{k+1}}^{U_{k+1}^1} \right]^{-1}
\end{aligned}$$

et

$$M_2 = M_1 \varepsilon A_{U_{k+1}}^{U_{k+1}^2} \left(\Lambda_{U_{k+1}}^{U_{k+1}^2} \Leftrightarrow D_{U_{k+1}}^{U_{k+1}^1} \Leftrightarrow D_{U_{k+1}}^{U_{k+1}^2} \right)^{-1}.$$

D'après a) de la proposition 3.1.3, on a

$$\lim_{\varepsilon \searrow 0} \frac{1}{\varepsilon^k} \Pi_{\alpha_k} G_{\alpha_k}^{U_{k+1}^2} = 0.$$

Donc

$$\begin{aligned}
&\lim_{\varepsilon \searrow 0} \frac{1}{\varepsilon^k} \Pi_{\alpha_k} G_{\alpha_k}^{U_{k+1}^1} \mathcal{N}_{U_{k+1}} \Leftrightarrow D_{U_{k+1}} \left(G_{U_{k+1}}^{\beta_k} \right) \mathbb{1}_{\beta_k} \\
&= \lim_{\varepsilon \searrow 0} \frac{1}{\varepsilon^k} \Pi_{\alpha_k} G_{\alpha_k}^{U_{k+1}^1} \left(M_1 G_{U_{k+1}}^{\beta_k} + M_2 G_{U_{k+1}}^{\beta_k} \right) \mathbb{1}_{\beta_k}.
\end{aligned}$$

Puisque, d'après toujours a) de la proposition 3.1.3,

$$\lim_{\varepsilon \searrow 0} \frac{1}{\varepsilon^k} \Pi_{\alpha_k} G_{\alpha_k}^{U_{k+1}^1}$$

existe, il suffit de montrer

$$\lim_{\varepsilon \searrow 0} M_1 G_{U_{k+1}}^{\beta_k} = 0 \text{ et } \lim_{\varepsilon \searrow 0} M_2 G_{U_{k+1}}^{\beta_k} = 0.$$

D'après b) de la proposition 3.1.3, $\forall a \in U_{k+1}^1, b \in \beta_k$, on a

$$d(\lambda_{ab}) \geq 1.$$

Par conséquent, si on pose

$$A_{U_{k+1}^1}^{\beta_k} = \varepsilon^{-1} G_{U_{k+1}^1}^{\beta_k} ,$$

alors $A_{U_{k+1}^1}^{\beta_k}$ est analytique en $\varepsilon = 0$. On a alors

$$\begin{aligned} & M_1 G_{U_{k+1}^1}^{\beta_k} \\ &= \left[I \Leftrightarrow \varepsilon \left(\Lambda_{U_{k+1}^1} \Leftrightarrow \varepsilon B_{U_{k+1}^1}^{U_{k+1}^2} \Leftrightarrow D_{U_{k+1}^1} \right)^{-1} A_{U_{k+1}^1}^{U_{k+1}^2} \times \right. \\ & \quad \left. \left(\Lambda_{U_{k+1}^2} \Leftrightarrow D_{U_{k+1}^2}^{U_{k+1}^1} \Leftrightarrow D_{U_{k+1}^2} \right)^{-1} G_{U_{k+1}^2}^{U_{k+1}^1} \right]^{-1} \times \\ & \quad \varepsilon \left(\Lambda_{U_{k+1}^1} \Leftrightarrow \varepsilon B_{U_{k+1}^1}^{U_{k+1}^2} \Leftrightarrow D_{U_{k+1}^1} \right)^{-1} A_{U_{k+1}^1}^{\beta_k} . \end{aligned}$$

Appliquons la propriété (2) de la proposition 3.1.11 aux matrices

$$\varepsilon \left(\Lambda_{U_{k+1}^1} \Leftrightarrow \varepsilon B_{U_{k+1}^1}^{U_{k+1}^2} \Leftrightarrow D_{U_{k+1}^1} \right)^{-1} A_{U_{k+1}^1}^{U_{k+1}^2}$$

et

$$\varepsilon \left(\Lambda_{U_{k+1}^1} \Leftrightarrow \varepsilon B_{U_{k+1}^1}^{U_{k+1}^2} \Leftrightarrow D_{U_{k+1}^1} \right)^{-1} A_{U_{k+1}^1}^{\beta_k} ,$$

nous obtenons

$$\lim_{\varepsilon \searrow 0} M_1 G_{U_{k+1}^1}^{\beta_k} = 0 .$$

Par un raisonnement analogue,

$$\lim_{\varepsilon \searrow 0} M_1 \varepsilon A_{U_{k+1}^1}^{U_{k+1}^2} = 0 .$$

D'autre part,

$$\left(\Lambda_{U_{k+1}^2} \Leftrightarrow D_{U_{k+1}^2}^{U_{k+1}^1} \Leftrightarrow D_{U_{k+1}^2} \right)^{-1} G_{U_{k+1}^2}^{\beta_k}$$

est analytique en $\varepsilon = 0$ selon le corollaire 3.1.1, d'où

$$\lim_{\varepsilon \searrow 0} M_2 G_{U_{k+1}^2}^{\beta_k} = \lim_{\varepsilon \searrow 0} M_1 \varepsilon A_{U_{k+1}^1}^{U_{k+1}^2} \left(\Lambda_{U_{k+1}^2} \Leftrightarrow D_{U_{k+1}^2}^{U_{k+1}^1} \Leftrightarrow D_{U_{k+1}^2} \right)^{-1} G_{U_{k+1}^2}^{\beta_k} = 0 .$$

On a ainsi montré

$$\lim_{\varepsilon \searrow 0} \lambda_{\alpha_k} \beta_k = 0 .$$

Pour la propriété (2), il suffit de montrer

$$P_{U_{k+1}}^{\beta_k} = \Leftrightarrow \cancel{N_{U_{k+1}} \Leftrightarrow D_{U_{k+1}}}^{-1} G_{U_{k+1}}^{\beta_k} \mathbb{1}_{\beta_k} .$$

Pour tout $a \in U_{k+1}$,

$$p_a^{\beta_k} = \sum_{b \in U_{k+1} \setminus \{a\}} \frac{\lambda_{ab}}{\lambda_{a.}} p_b^{\beta_k} + \sum_{b \in \beta_k} \frac{\lambda_{ab}}{\lambda_{a.}}$$

où

$$\lambda_{a.} = \Leftrightarrow \lambda_{aa} = \sum_{b \in U_{k+1} \setminus \{a\}} \lambda_{ab} + \sum_{\substack{b \in \gamma_k \\ \gamma_k \in \mathcal{F}_k}} \lambda_{ab} .$$

D'où

$$\sum_{b \in U_{k+1} \setminus \{a\}} \lambda_{ab} p_b^{\beta_k} \Leftrightarrow \lambda_{a.} p_a^{\beta_k} + \sum_{b \in \beta_k} \lambda_{ab} = 0$$

ce qui équivaut à

$$\cancel{N_{U_{k+1}} \Leftrightarrow D_{U_{k+1}}} P_{U_{k+1}}^{\beta_k} + G_{U_{k+1}}^{\beta_k} \mathbb{1}_{\beta_k} = 0 .$$

D'où (2).

Soient $\varepsilon^k \lambda(\varepsilon)$ une valeur propre de Λ et X un vecteur propre associé. On a alors

$$\Lambda X = \varepsilon^k \lambda(\varepsilon) X . \quad (3.12)$$

Pour montrer la propriété (3). Montrons premièrement que $\lambda(0)$ est une valeur propre du générateur $\lim_{\varepsilon \searrow 0} \Lambda_k$.

Pour toute classe minimale α_i de $F_i(\alpha_k)$, posons $\alpha_j = \mathcal{E}_j(\alpha_i)$ avec $1 \leq i \leq j \leq k$. En tenant compte du lemme 3.1.3,

$$X_{\alpha_j} = \left(\sum_{l=0}^{k-j} \varepsilon^l c_{\alpha_{k-l}} \right) \mathbb{1}_{\alpha_j} + \varepsilon^{k-j+1} Y_{\alpha_j} .$$

Montrons par récurrence, pour $1 \leq i \leq k$,

$$\frac{1}{\varepsilon^{i-1}} \Pi_{\alpha_k}^{\alpha_i} \left(\sum_{j=0}^{k-i-1} T_{ij} + V_i \right) = \lambda(0) c_{\alpha_k} q_{\alpha_k}^{\alpha_i} + O(\varepsilon) \quad (3.13)$$

où la somme par rapport à l'indice j est 0 si $i = k$ et

$$\begin{aligned} T_{ij} &= \sum_{\substack{\beta_{i+j} \in F_{i+j}(\alpha_{i+j+1}) \\ \beta_{i+j} \neq \alpha_{i+j}}} B_{\alpha_i}^{\beta_{i+j}} \mathbb{1}_{\alpha_i} \not/c_{\beta_{i+j}} \Leftrightarrow c_{\alpha_{i+j}} \\ &\Leftrightarrow B_{\alpha_i}^{U_{i+j+1}} \mathbb{1}_{\alpha_i} c_{\alpha_{i+j}} + A_{\alpha_i}^{U_{i+j+1}} Y_{U_{i+j+1}} \end{aligned} \quad (3.14)$$

avec $Y_{U_{i+j+1}}$ étant la restriction de $Y_{\alpha_{i+j+1}}$ sur U_{i+j+1} , et

$$V_i = \Leftrightarrow B_{\alpha_i}^{O_{k+1}} \mathbb{1}_{\alpha_i} c_{\alpha_k} + A_{\alpha_i}^{O_{k+1}} X_{O_{k+1}} . \quad (3.15)$$

Pour $i = 1$, notons

$$U_{j+1} = U_{j+1}[\alpha_1]$$

avec $0 \leq j \leq k$. Considérons les lignes indicées par α_1 de (3.12), nous avons

$$\Lambda_{\alpha_1} X_{\alpha_1} + \sum_{j=1}^k \varepsilon^j \not/\left(B_{\alpha_1}^{O_{j+1}} X_{\alpha_1} + A_{\alpha_1}^{O_{j+1}} X_{O_{j+1}} \right) = \varepsilon^k \lambda(\varepsilon) X_{\alpha_1} .$$

Multiplions l'égalité ci-dessus à gauche par $\Pi_{\alpha_k}^{\alpha_1}$ qui est la restriction à α_1 de Π_{α_k} , la mesure réversible de Λ_{α_k} . Du fait que

$$\Pi_{\alpha_k}^{\alpha_1} \Lambda_{\alpha_1} = 0 ,$$

on déduit

$$\Pi_{\alpha_k}^{\alpha_1} \sum_{j=1}^k \varepsilon^j \not/\left(B_{\alpha_1}^{O_{j+1}} X_{\alpha_1} + A_{\alpha_1}^{O_{j+1}} X_{O_{j+1}} \right) = \varepsilon^k \lambda(\varepsilon) \Pi_{\alpha_k}^{\alpha_1} X_{\alpha_1} .$$

Or d'après le lemme 3.1.3, pour $1 \leq j \leq k \Leftrightarrow 1$, on a

$$\begin{aligned} &\Leftrightarrow B_{\alpha_1}^{O_{j+1}} X_{\alpha_1} + A_{\alpha_1}^{O_{j+1}} X_{O_{j+1}} \\ &= \left(\sum_{l=0}^{k-j-1} \varepsilon^l c_{\alpha_{k-l}} \right) \left(\Leftrightarrow B_{\alpha_1}^{O_{j+1}} \mathbb{1}_{\alpha_1} + A_{\alpha_1}^{O_{j+1}} \mathbb{1}_{O_{j+1}} \right) \\ &\quad + \varepsilon^{k-j} \left(\Leftrightarrow B_{\alpha_1}^{O_{j+1}} \mathbb{1}_{\alpha_1} c_{\alpha_j} + A_{\alpha_1}^{O_{j+1}} Y_{O_{j+1}} \right) + O \left(\varepsilon_{\alpha_1}^{k-j+1} \right) \\ &= \varepsilon^{k-j} \left(\Leftrightarrow B_{\alpha_1}^{O_{j+1}} \mathbb{1}_{\alpha_1} c_{\alpha_j} + A_{\alpha_1}^{O_{j+1}} Y_{O_{j+1}} \right) + O \left(\varepsilon_{\alpha_1}^{k-j+1} \right) \\ &= \varepsilon^{k-j} \sum_{\substack{\beta_j \in F_j(\alpha_{j+1}) \\ \beta_j \neq \alpha_j}} \left(\Leftrightarrow B_{\alpha_1}^{\beta_j} \mathbb{1}_{\alpha_1} c_{\alpha_j} + A_{\alpha_1}^{\beta_j} \mathbb{1}_{\beta_j} c_{\beta_j} \right) \end{aligned}$$

$$\begin{aligned}
 & +\varepsilon^{k-j} \left(\Leftrightarrow B_{\alpha_1}^{U_{j+1}} \mathbb{1}_{\alpha_1} c_{\alpha_j} + A_{\alpha_1}^{U_{j+1}} Y_{U_{j+1}} \right) + O \left(\varepsilon_{\alpha_1}^{k-j+1} \right) \\
 = & \varepsilon^{k-j} \sum_{\substack{\beta_j \in F_j(\alpha_{j+1}) \\ \beta_j \neq \alpha_j}} B_{\alpha_1}^{\beta_j} \mathbb{1}_{\alpha_1} c_{\beta_j} \Leftrightarrow c_{\alpha_j} \\
 & +\varepsilon^{k-j} \left(\Leftrightarrow B_{\alpha_1}^{U_{j+1}} \mathbb{1}_{\alpha_1} c_{\alpha_j} + A_{\alpha_1}^{U_{j+1}} Y_{U_{j+1}} \right) + O \left(\varepsilon_{\alpha_1}^{k-j+1} \right) \\
 = & \varepsilon^{k-j} T_{1j-1} + O \left(\varepsilon_{\alpha_1}^{k-j+1} \right)
 \end{aligned} \tag{3.16}$$

où l'égalité (3.16) provient de (3.4).

Or pour $j = k$, on a

$$\begin{aligned}
 & \Leftrightarrow B_{\alpha_1}^{O_{k+1}} X_{\alpha_1} + A_{\alpha_1}^{O_{k+1}} X_{O_{k+1}} \\
 = & \Leftrightarrow B_{\alpha_1}^{O_{k+1}} \mathbb{1}_{\alpha_1} c_{\alpha_k} + A_{\alpha_1}^{O_{k+1}} X_{O_{k+1}} + O(\varepsilon_{\alpha_1}) \\
 = & V_1 + O(\varepsilon_{\alpha_1}) .
 \end{aligned}$$

D'ailleurs,

$$\Pi_{\alpha_k}^{\alpha_1} X_{\alpha_1} = c_{\alpha_k} \Pi_{\alpha_k}^{\alpha_1} \mathbb{1}_{\alpha_1} + O(\varepsilon_{\alpha_1}) = c_{\alpha_k} q_{\alpha_k}^{\alpha_1} + O(\varepsilon_{\alpha_1}) .$$

On a alors

$$\Pi_{\alpha_k}^{\alpha_1} \left(\sum_{j=0}^k T_{1j} + V_1 \right) = \lambda(0) c_{\alpha_k} q_{\alpha_k}^{\alpha_1} + O(\varepsilon) .$$

Supposons que l'équation (3.13) soit vraie pour $j \leq i$. Quel que soit $\alpha_{i+1} \in F_{i+1}(\alpha_k)$, faisant la somme des (3.13) pour toutes les classes α_i de $F_i(\alpha_{i+1})$, en tenant compte de la proposition 3.1.5, on a

$$\frac{1}{\varepsilon^{i-1}} \sum_{\alpha_i \in F_i(\alpha_{i+1})} \Pi_{\alpha_k}^{\alpha_i} \left(\sum_{j=0}^{k-i-1} T_{ij} + V_i \right) = \lambda(0) c_{\alpha_k} q_{\alpha_k}^{\alpha_{i+1}} + O(\varepsilon) \tag{3.17}$$

où T_{ij} et V_i sont définis par 3.14 et 3.15 respectivement.

Soit $\alpha_i \in F_i(\alpha_{i+1})$, en notant

$$U_{i+1} = U_{i+1}[\alpha_i]$$

et pour $j = 1, \dots, k \Leftrightarrow i \Leftrightarrow 1$

$$O_{i+1+j} = O_{i+1+j}[\alpha_i] ,$$

on remarque qu'à une permutation près,

$$\Pi_{\alpha_k}^{\alpha_{i+1}} = \left(\begin{array}{c} \Pi_{\alpha_k}^{\alpha_i} \\ \alpha_i \in F_i(\alpha_{i+1}) \end{array}, \Pi_{\alpha_k}^{U_{i+1}} \right)$$

et que pour tout sous-ensemble S de O_{i+1+j} , on a

$$A_{\alpha_{i+1}}^S = \left(\begin{array}{c} \varepsilon \\ \left(A_{\alpha_i}^S \right)_{\alpha_i \in F_i(\alpha_{i+1})}^* \\ A_{U_{i+1}}^S \end{array} \right)$$

et

$$B_{\alpha_{i+1}}^S \mathbb{1}_{\alpha_{i+1}} = \left(\begin{array}{c} \varepsilon \\ \left(B_{\alpha_i}^S \mathbb{1}_{\alpha_i} \right)_{\alpha_i \in F_i(\alpha_{i+1})}^* \\ B_{U_{i+1}}^S \mathbb{1}_{U_{i+1}} \end{array} \right)$$

où $\left(A_{\alpha_i}^S \right)_{\alpha_i \in F_i(\alpha_{i+1})}^*$ désigne le vecteur colonne composé des $A_{\alpha_i}^S$ pour $\alpha_i \in F_i(\alpha_{i+1})$ et analoguement pour $\left(B_{\alpha_i}^S \mathbb{1}_{\alpha_i} \right)_{\alpha_i \in F_i(\alpha_{i+1})}^*$. On a donc

$$\Pi_{\alpha_k}^{\alpha_{i+1}} A_{\alpha_{i+1}}^S = \varepsilon \sum_{\alpha_i \in F_i(\alpha_{i+1})} \Pi_{\alpha_k}^{\alpha_i} A_{\alpha_i}^S + \Pi_{\alpha_k}^{U_{i+1}} A_{U_{i+1}}^S$$

et

$$\Pi_{\alpha_k}^{\alpha_{i+1}} B_{\alpha_{i+1}}^S \mathbb{1}_{\alpha_{i+1}} = \varepsilon \sum_{\alpha_i \in F_i(\alpha_{i+1})} \Pi_{\alpha_k}^{\alpha_i} B_{\alpha_i}^S \mathbb{1}_{\alpha_i} + \Pi_{\alpha_k}^{U_{i+1}} B_{U_{i+1}}^S \mathbb{1}_{U_{i+1}}.$$

D'où

$$\frac{1}{\varepsilon^{i-1}} \sum_{\alpha_i \in F_i(\alpha_{i+1})} \Pi_{\alpha_k}^{\alpha_i} A_{\alpha_i}^S = \frac{1}{\varepsilon^i} \Pi_{\alpha_k}^{\alpha_{i+1}} A_{\alpha_{i+1}}^S \Leftrightarrow \frac{1}{\varepsilon^i} \Pi_{\alpha_k}^{U_{i+1}} A_{U_{i+1}}^S$$

et

$$\frac{1}{\varepsilon^{i-1}} \sum_{\alpha_i \in F_i(\alpha_{i+1})} \Pi_{\alpha_k}^{\alpha_i} B_{\alpha_i}^S \mathbb{1}_{\alpha_i} = \frac{1}{\varepsilon^i} \Pi_{\alpha_k}^{\alpha_{i+1}} B_{\alpha_{i+1}}^S \mathbb{1}_{\alpha_{i+1}} \Leftrightarrow \frac{1}{\varepsilon^i} \Pi_{\alpha_k}^{U_{i+1}} B_{U_{i+1}}^S \mathbb{1}_{U_{i+1}}.$$

Considérant les deux équations ci-dessus et les équations (3.14) (3.15) et (3.17),

on a

$$\begin{aligned} & \frac{1}{\varepsilon^{i-1}} \sum_{\alpha_i \in F_i(\alpha_{i+1})} \Pi_{\alpha_k}^{\alpha_i} \left(\sum_{j=0}^{k-i-1} T_{ij} + V_i \right) \\ &= \frac{1}{\varepsilon^{i-1}} \sum_{\alpha_i \in F_i(\alpha_{i+1})} \Pi_{\alpha_k}^{\alpha_i} \left(\sum_{j=1}^{k-i-1} T_{ij} + V_i \right) + \frac{1}{\varepsilon^{i-1}} \sum_{\alpha_i \in F_i(\alpha_{i+1})} \Pi_{\alpha_k}^{\alpha_i} T_{i0} \end{aligned}$$

$$\begin{aligned}
&= \sum_{j=1}^{k-i-1} \frac{1}{\varepsilon^{i-1}} \sum_{\alpha_i \in F_i(\alpha_{i+1})} \Pi_{\alpha_k}^{\alpha_i} (T_{ij} + V_i) + \frac{1}{\varepsilon^{i-1}} \sum_{\alpha_i \in F_i(\alpha_{i+1})} \Pi_{\alpha_k}^{\alpha_i} T_{i0} \\
&= \frac{1}{\varepsilon^i} \Pi_{\alpha_k}^{\alpha_{i+1}} \left(\sum_{j=0}^{k-i-2} T_{i+1j} + V_{i+1} \right) \Leftrightarrow \frac{1}{\varepsilon^i} \Pi_{\alpha_k}^{U_{i+1}} \left(\sum_{j=1}^{k-i-1} M_{ij} + W_i \right) \\
&\quad + \frac{1}{\varepsilon^{i-1}} \sum_{\alpha_i \in F_i(\alpha_{i+1})} \Pi_{\alpha_k}^{\alpha_i} T_{i0}
\end{aligned}$$

avec

$$\begin{aligned}
M_{ij} &= \sum_{\substack{\beta_{i+j} \in F_{i+j}(\alpha_{i+j+1}) \\ \beta_{i+j} \neq \alpha_{i+j}}} B_{U_{i+1}}^{\beta_{i+j}} \mathbb{1}_{U_{i+1}} \not/c_{\beta_{i+j}} \Leftrightarrow c_{\alpha_{i+j}} \\
&\Leftrightarrow B_{U_{i+1}}^{U_{i+j+1}} \mathbb{1}_{U_{i+1}} c_{\alpha_{i+j}} + A_{U_{i+1}}^{U_{i+j+1}} Y_{U_{i+j+1}}
\end{aligned}$$

et

$$W_i = \Leftrightarrow B_{U_{i+1}}^{O_{k+1}} \mathbb{1}_{U_{i+1}} c_{\alpha_k} + A_{U_{i+1}}^{O_{k+1}} X_{O_{k+1}} .$$

Donc pour montrer l'égalité 3.13 pour $i+1$, c'est-à-dire

$$\frac{1}{\varepsilon^i} \Pi_{\alpha_k}^{\alpha_{i+1}} \left(\sum_{j=0}^{k-i-2} T_{i+1j} + V_{i+1} \right) = \lambda(0) c_{\alpha_k} q_{\alpha_k}^{\alpha_i} + O(\varepsilon) ,$$

il suffit de montrer

$$\frac{1}{\varepsilon^{i-1}} \sum_{\alpha_i \in F_i(\alpha_{i+1})} \Pi_{\alpha_k}^{\alpha_i} T_{i0} \Leftrightarrow \frac{1}{\varepsilon^i} \Pi_{\alpha_k}^{U_{i+1}} \left(\sum_{j=1}^{k-i-1} M_{ij} + W_i \right) = O(\varepsilon) ,$$

où

$$\begin{aligned}
T_{i0} &= \sum_{\substack{\beta_i \in F_i(\alpha_{i+1}) \\ \beta_i \neq \alpha_i}} B_{\alpha_i}^{\beta_i} \mathbb{1}_{\alpha_i} (c_{\beta_i} \Leftrightarrow c_{\alpha_i}) \\
&\Leftrightarrow B_{\alpha_i}^{U_{i+1}} \mathbb{1}_{\alpha_i} c_{\alpha_i} + A_{\alpha_i}^{U_{i+1}} Y_{U_{i+1}} . \tag{3.18}
\end{aligned}$$

Soit $\alpha_i \in F_i(\alpha_{i+1})$. Prenons les lignes indicées par U_{i+1} de (3.12), nous obtenons,

$$\begin{aligned}
&\not/X_{U_{i+1}} \Leftrightarrow D_{U_{i+1}} X_{U_{i+1}} + \sum_{\alpha_i \in F_i(\alpha_{i+1})} G_{U_{i+1}}^{\alpha_i} X_{\alpha_i} \\
&+ \sum_{j=1}^{k-i} \varepsilon^j \left(\Leftrightarrow B_{U_{i+1}}^{O_{i+j+1}} X_{U_{i+1}} + A_{U_{i+1}}^{O_{i+j+1}} X_{O_{i+j+1}} \right) \\
&= \varepsilon^k \lambda(\varepsilon) X_{U_{i+1}} .
\end{aligned}$$

D'apr es le lemme 3.1.3, on a

$$\begin{aligned}
\Lambda_{U_{i+1}} X_{U_{i+1}} &= \left(\sum_{l=0}^{k-i-1} \varepsilon^l c_{\alpha_{k-l}} \right) \Lambda_{U_{i+1}} \mathbb{1}_{U_{i+1}} + \varepsilon^{k-i} \Lambda_{U_{i+1}} Y_{U_{i+1}} \\
&= \varepsilon^{k-i} \Lambda_{U_{i+1}} Y_{U_{i+1}}, \\
&\Leftrightarrow D_{U_{i+1}} X_{U_{i+1}} + \sum_{\alpha_i \in F_i(\alpha_{i+1})} G_{U_{i+1}}^{\alpha_i} X_{\alpha_i} \\
&= \left(\sum_{l=0}^{k-i-1} \varepsilon^l c_{\alpha_{k-l}} \right) \left(\Leftrightarrow D_{U_{i+1}} \mathbb{1}_{U_{i+1}} + \sum_{\alpha_i \in F_i(\alpha_{i+1})} G_{U_{i+1}}^{\alpha_i} \mathbb{1}_{\alpha_i} \right) \\
&\quad + \varepsilon^{k-i} \left(\Leftrightarrow D_{U_{i+1}} Y_{U_{i+1}} + \sum_{\alpha_i \in F_i(\alpha_{i+1})} G_{U_{i+1}}^{\alpha_i} \mathbb{1}_{\alpha_i} c_{\alpha_i} \right) + O(\varepsilon_{U_{i+1}}^{k-i+1}) \\
&= \varepsilon^{k-i} \left(\Leftrightarrow D_{U_{i+1}} Y_{U_{i+1}} + \sum_{\alpha_i \in F_i(\alpha_{i+1})} G_{U_{i+1}}^{\alpha_i} \mathbb{1}_{\alpha_i} c_{\alpha_i} \right) + O(\varepsilon_{U_{i+1}}^{k-i+1})
\end{aligned}$$

et pour $1 \leq j \leq k \Leftrightarrow i \Leftrightarrow 1$,

$$\begin{aligned}
&\Leftrightarrow B_{U_{i+1}}^{O_{i+j+1}} X_{U_{i+1}} + A_{U_{i+1}}^{O_{i+j+1}} X_{O_{i+j+1}} \\
&= \left(\sum_{l=0}^{k-i-j-1} \varepsilon^l c_{\alpha_{k-l}} \right) \left(\Leftrightarrow B_{U_{i+1}}^{O_{i+j+1}} \mathbb{1}_{U_{i+1}} + A_{U_{i+1}}^{O_{i+j+1}} \mathbb{1}_{O_{i+j+1}} \right) + \\
&\quad \varepsilon^{k-i-j} \left(\Leftrightarrow B_{U_{i+1}}^{O_{i+j+1}} \mathbb{1}_{U_{i+1}} c_{\alpha_{i+j}} + A_{U_{i+1}}^{O_{i+j+1}} Y_{O_{i+j+1}} \right) + O(\varepsilon_{U_{i+1}}^{k-i-j+1}) \\
&= \varepsilon^{k-i-j} \left(\Leftrightarrow B_{U_{i+1}}^{O_{i+j+1}} \mathbb{1}_{U_{i+1}} c_{\alpha_{i+j}} + A_{U_{i+1}}^{O_{i+j+1}} Y_{O_{i+j+1}} \right) + O(\varepsilon_{U_{i+1}}^{k-i-j+1}) \\
&= \varepsilon^{k-i-j} \sum_{\substack{\beta_{i+j} \in F_{i+j}(\alpha_{i+j+1}) \\ \beta_{i+j} \neq \alpha_{i+j}}} \left(\Leftrightarrow B_{U_{i+1}}^{\beta_{i+j}} \mathbb{1}_{U_{i+1}} c_{\alpha_{i+j}} + A_{U_{i+1}}^{\beta_{i+j}} \mathbb{1}_{\beta_{i+j}} c_{\beta_{i+j}} \right) \\
&\quad + \varepsilon^{k-i-j} \left(\Leftrightarrow B_{U_{i+1}}^{U_{i+j+1}} \mathbb{1}_{U_{i+1}} c_{\alpha_{i+j}} + A_{U_{i+1}}^{U_{i+j+1}} Y_{U_{i+j+1}} \right) + O(\varepsilon_{U_{i+1}}^{k-i-j+1}) \\
&= \varepsilon^{k-i-j} \sum_{\substack{\beta_{i+j} \in F_{i+j}(\alpha_{i+j+1}) \\ \beta_{i+j} \neq \alpha_{i+j}}} B_{U_{i+1}}^{\beta_{i+j}} \mathbb{1}_{U_{i+1}} \cancel{c_{\beta_{i+j}}} \Leftrightarrow c_{\alpha_{i+j}} \\
&\quad \varepsilon^{k-i-j} \left(\Leftrightarrow B_{U_{i+1}}^{U_{i+j+1}} \mathbb{1}_{U_{i+1}} c_{\alpha_{i+j}} + A_{U_{i+1}}^{U_{i+j+1}} Y_{U_{i+j+1}} \right) + O(\varepsilon_{U_{i+1}}^{k-i-j+1}) \\
&= \varepsilon^{k-i-j} M_{ij} + o(\varepsilon_{U_{i+1}}^{k-i-j})
\end{aligned}$$

pour $j = k \Leftrightarrow i$

$$\Leftrightarrow B_{U_{i+1}}^{O_{k+1}} X_{U_{i+1}} + A_{U_{i+1}}^{O_{k+1}} X_{O_{k+1}}$$

$$\begin{aligned}
 &= \cancel{\Leftrightarrow} B_{U_{i+1}}^{O_{k+1}} \mathbb{1}_{U_{i+1}} c_{\alpha_k} + A_{U_{i+1}}^{O_{k+1}} X_{O_{k+1}} + O(\cancel{\varepsilon_{U_{i+1}}}) \\
 &= W_i + O(\cancel{\varepsilon_{U_{i+1}}}) .
 \end{aligned}$$

D'où

$$\begin{aligned}
 &\cancel{N_{U_{i+1}}} \Leftrightarrow D_{U_{i+1}} Y_{U_{i+1}} + \sum_{\alpha_i \in F_i(\alpha_{i+1})} G_{U_{i+1}}^{\alpha_i} \mathbb{1}_{\alpha_i} c_{\alpha_i} + \sum_{j=1}^{k-i-1} M_{ij} + W_i \\
 &= O(\varepsilon_{U_{i+1}}) ,
 \end{aligned}$$

où le terme $O(\varepsilon_{U_{i+1}})$ est déduit de l'égalité a) de la proposition 3.1.3. Du corollaire 3.1.1, on déduit

$$\begin{aligned}
 &Y_{U_{i+1}} \\
 &= \cancel{\Leftrightarrow} \cancel{N_{U_{i+1}}} \Leftrightarrow D_{U_{i+1}}^{-1} \left(\sum_{\alpha_i \in F_i(\alpha_{i+1})} G_{U_{i+1}}^{\alpha_i} \mathbb{1}_{\alpha_i} c_{\alpha_i} + \sum_{j=1}^{k-i-1} M_{ij} + W_i \right) \\
 &\quad + O(\varepsilon_{U_{i+1}}) .
 \end{aligned}$$

Notons

$$O_{i+1} = O_{i+1}[\alpha_i] .$$

Pour tout sous-ensemble S de O_{i+1} , on a

$$\varepsilon B_{\alpha_i}^S \mathbb{1}_{\alpha_i} = \varepsilon A_{\alpha_i}^S \mathbb{1}_S = G_{\alpha_i}^S \mathbb{1}_S .$$

Insérons l'expression ci-dessus de $Y_{U_{i+1}}$ dans T_{i0} donné par (3.18), on a

$$\begin{aligned}
 &\frac{1}{\varepsilon^{i-1}} \sum_{\alpha_i \in F_i(\alpha_{i+1})} \Pi_{\alpha_k}^{\alpha_i} T_{i0} \\
 &= \frac{1}{\varepsilon^i} \sum_{\alpha_i \in F_i(\alpha_{i+1})} \Pi_{\alpha_k}^{\alpha_i} \sum_{\substack{\beta_i \in F_i(\alpha_{i+1}) \\ \beta_i \neq \alpha_i}} G_{\alpha_i}^{\beta_i} \mathbb{1}_{\beta_i} (c_{\beta_i} \Leftrightarrow c_{\alpha_i}) \\
 &\quad \Leftrightarrow \frac{1}{\varepsilon^i} \sum_{\alpha_i \in F_i(\alpha_{i+1})} \Pi_{\alpha_k}^{\alpha_i} G_{\alpha_i}^{U_{i+1}} \mathbb{1}_{U_{i+1}} c_{\alpha_i} \\
 &\quad \Leftrightarrow \frac{1}{\varepsilon^i} \sum_{\alpha_i \in F_i(\alpha_{i+1})} \Pi_{\alpha_k}^{\alpha_i} G_{\alpha_i}^{U_{i+1}} \cancel{N_{U_{i+1}}} + D_{U_{i+1}})^{-1} \sum_{\alpha_i \in F_i(\alpha_{i+1})} G_{U_{i+1}}^{\alpha_i} \mathbb{1}_{\alpha_i} c_{\alpha_i}
 \end{aligned}$$

$$\Leftrightarrow \frac{1}{\varepsilon^i} \sum_{\alpha_i \in F_i(\alpha_{i+1})} \Pi_{\alpha_k}^{\alpha_i} G_{\alpha_i}^{U_{i+1}} \left(N_{U_{i+1}} + D_{U_{i+1}} \right)^{-1} \left(\sum_{j=1}^{k-i-1} M_{ij} + W_i \right) + O(\varepsilon_{\alpha_i}) .$$

Or de la r eversibilit e, on d eduit

$$\begin{aligned} & \sum_{\alpha_i \in F_i(\alpha_{i+1})} \sum_{\substack{\beta_i \in F_i(\alpha_{i+1}) \\ \beta_i \neq \alpha_i}} \Pi_{\alpha_k}^{\alpha_i} G_{\alpha_i}^{\beta_i} \mathbb{1}_{\beta_i} c_{\alpha_i} \\ &= \sum_{\alpha_i \in F_i(\alpha_{i+1})} \sum_{\substack{\beta_i \in F_i(\alpha_{i+1}) \\ \beta_i \neq \alpha_i}} \Pi_{\alpha_k}^{\beta_i} G_{\beta_i}^{\alpha_i} \mathbb{1}_{\alpha_i} c_{\alpha_i} . \end{aligned}$$

D'o u

$$\begin{aligned} & \sum_{\alpha_i \in F_i(\alpha_{i+1})} \Pi_{\alpha_k}^{\alpha_i} \sum_{\substack{\beta_i \in F_i(\alpha_{i+1}) \\ \beta_i \neq \alpha_i}} G_{\alpha_i}^{\beta_i} \mathbb{1}_{\beta_i} (c_{\beta_i} \Leftrightarrow c_{\alpha_i}) \\ &= \sum_{\alpha_i \in F_i(\alpha_{i+1})} \sum_{\substack{\beta_i \in F_i(\alpha_{i+1}) \\ \beta_i \neq \alpha_i}} \Pi_{\alpha_k}^{\alpha_i} G_{\alpha_i}^{\beta_i} \mathbb{1}_{\beta_i} c_{\beta_i} \\ &\Leftrightarrow \sum_{\alpha_i \in F_i(\alpha_{i+1})} \sum_{\substack{\beta_i \in F_i(\alpha_{i+1}) \\ \beta_i \neq \alpha_i}} \Pi_{\alpha_k}^{\alpha_i} G_{\alpha_i}^{\beta_i} \mathbb{1}_{\beta_i} c_{\alpha_i} \\ &= \sum_{\alpha_i \in F_i(\alpha_{i+1})} \sum_{\substack{\beta_i \in F_i(\alpha_{i+1}) \\ \beta_i \neq \alpha_i}} \Pi_{\alpha_k}^{\alpha_i} G_{\alpha_i}^{\beta_i} \mathbb{1}_{\beta_i} c_{\beta_i} \\ &\Leftrightarrow \sum_{\alpha_i \in F_i(\alpha_{i+1})} \sum_{\substack{\beta_i \in F_i(\alpha_{i+1}) \\ \beta_i \neq \alpha_i}} \Pi_{\alpha_k}^{\beta_i} G_{\beta_i}^{\alpha_i} \mathbb{1}_{\alpha_i} c_{\alpha_i} \\ &= 0 . \end{aligned}$$

En tenant compte en plus de la proposition 3.1.6, on obtient

$$\frac{1}{\varepsilon^{i-1}} \sum_{\alpha_i \in F_i(\alpha_{i+1})} \Pi_{\alpha_k}^{\alpha_i} T_{i0} = \frac{1}{\varepsilon^i} \Pi_{\alpha_k}^{U_{i+1}} \left(\sum_{j=1}^{k-i-1} M_{ij} + W_i \right) + O(\varepsilon) .$$

On a ainsi montr e l' egalit e (3.13).

Pour toute classe minimale α_k de \mathcal{F}_k , on a alors

$$\frac{1}{\varepsilon^{k-1}} \Pi_{\alpha_k} \left(\Leftrightarrow B_{\alpha_k}^{O_{k+1}} \mathbb{1}_{\alpha_k} c_{\alpha_k} + A_{\alpha_k}^{O_{k+1}} X_{O_{k+1}} \right) = \lambda(0) c_{\alpha_k} + O(\varepsilon)$$

ou bien

$$\frac{1}{\varepsilon^k} \Pi_{\alpha_k} \left(\Leftrightarrow G_{\alpha_k}^{O_{k+1}} \mathbb{1}_{O_{k+1}} c_{\alpha_k} + G_{\alpha_k}^{O_{k+1}} X_{O_{k+1}} \right) = \lambda(0) c_{\alpha_k} + O(\varepsilon)$$

Or pour toute $\beta_k \in \mathcal{F}_k$, il existe une fonction continue c_{β_k} telle que

$$X_{\beta_k} = \mathbb{1}_{\beta_k} c_{\beta_k} + O(\varepsilon_{\beta_k}) .$$

En tenant compte de a) de la proposition couche,

$$\begin{aligned} & \frac{1}{\varepsilon^k} \Pi_{\alpha_k} \left(\Leftrightarrow G_{\alpha_k}^{O_{k+1}} \mathbb{1}_{O_{k+1}} c_{\alpha_k} + G_{\alpha_k}^{O_{k+1}} X_{O_{k+1}} \right) \\ &= \frac{1}{\varepsilon^k} \Pi_{\alpha_k} \sum_{\substack{\beta_k \in \mathcal{F}_k \\ \beta_k \neq \alpha_k}} \left(\Leftrightarrow G_{\alpha_k}^{\beta_k} \mathbb{1}_{\beta_k} c_{\alpha_k} + G_{\alpha_k}^{\beta_k} \mathbb{1}_{\beta_k} c_{\beta_k} \right) \\ & \frac{1}{\varepsilon^k} \Pi_{\alpha_k} \left(\Leftrightarrow G_{\alpha_k}^{U_{k+1}} \mathbb{1}_{U_{k+1}} c_{\alpha_k} + G_{\alpha_k}^{U_{k+1}} X_{U_{k+1}} \right) + O(\varepsilon) \\ &= \frac{1}{\varepsilon^k} \Pi_{\alpha_k} \sum_{\substack{\beta_k \in \mathcal{F}_k \\ \beta_k \neq \alpha_k}} G_{\alpha_k}^{\beta_k} \mathbb{1}_{\beta_k} (c_{\beta_k} \Leftrightarrow c_{\alpha_k}) \\ & \frac{1}{\varepsilon^k} \Pi_{\alpha_k} \left(\Leftrightarrow G_{\alpha_k}^{U_{k+1}} \mathbb{1}_{U_{k+1}} c_{\alpha_k} + G_{\alpha_k}^{U_{k+1}} X_{U_{k+1}} \right) + O(\varepsilon) . \end{aligned}$$

On a alors

$$\begin{aligned} & \frac{1}{\varepsilon^k} \Pi_{\alpha_k} \left[\sum_{\substack{\beta_k \in \mathcal{F}_k \\ \beta_k \neq \alpha_k}} G_{\alpha_k}^{\beta_k} \mathbb{1}_{\beta_k} (c_{\beta_k} \Leftrightarrow c_{\alpha_k}) \Leftrightarrow G_{\alpha_k}^{U_{k+1}} \mathbb{1}_{U_{k+1}} c_{\alpha_k} + G_{\alpha_k}^{U_{k+1}} X_{U_{k+1}} + O(\varepsilon_{\alpha_k}) \right] \\ &= \lambda(0) c_{\alpha_k} + O(\varepsilon) . \end{aligned} \tag{3.19}$$

D'après le lemme 3.1.3,

$$X_{U_{k+1}} = \Leftrightarrow \cancel{N_{U_{k+1}}} \Leftrightarrow D_{U_{k+1}})^{-1} \sum_{\beta_k \in \mathcal{F}_k} G_{U_{k+1}}^{\beta_k} \mathbb{1}_{\beta_k} c_{\beta_k} + O(\varepsilon_{U_{k+1}}) .$$

Remplaçant $X_{U_{k+1}}$ par l'expression ci-dessus dans (3.19), on obtient

$$\begin{aligned} & \frac{1}{\varepsilon^k} \Pi_{\alpha_k} \left[\sum_{\beta_k \in \mathcal{F}_k, \beta_k \neq \alpha_k} G_{\alpha_k}^{\beta_k} \mathbb{1}_{\beta_k} (c_{\beta_k} \Leftrightarrow c_{\alpha_k}) \right. \\ & \left. \Leftrightarrow G_{\alpha_k}^{U_{k+1}} \mathbb{1}_{U_{k+1}} c_{\alpha_k} \Leftrightarrow A_{\alpha_k}^{U_{k+1}} \cancel{N_{U_{k+1}}} \Leftrightarrow D_{U_{k+1}})^{-1} \sum_{\beta_k \in \mathcal{F}_k} G_{U_{k+1}}^{\beta_k} \mathbb{1}_{\beta_k} c_{\beta_k} \right] \\ &= \sum_{\beta_k \in \mathcal{F}_k, \beta_k \neq \alpha_k} \lambda_{\alpha_k \beta_k} c_{\beta_k} \Leftrightarrow c_{\alpha_k} \frac{1}{\varepsilon^k} \Pi_{\alpha_k} \times \\ & \left[G_{\alpha_k}^{O_{k+1}} \mathbb{1}_{O_{k+1}} + G_{\alpha_k}^{U_{k+1}} \cancel{N_{U_{k+1}}} \Leftrightarrow D_{U_{k+1}})^{-1} G_{U_{k+1}}^{\alpha_k} \mathbb{1}_{\alpha_k} \right] \\ &= \lambda(0) c_{\alpha_k} + O(\varepsilon) . \end{aligned}$$

Comme

$$\mathcal{N}_{U_{k+1}} \Leftrightarrow D_{U_{k+1}} \mathbb{1}_{U_{k+1}} = \Leftrightarrow D_{U_{k+1}} \mathbb{1}_{U_{k+1}} = \Leftrightarrow \sum_{\beta_k \in \mathcal{F}_k} G_{U_{k+1}}^{\beta_k} \mathbb{1}_{\beta_k} ,$$

On a

$$\mathbb{1}_{U_{k+1}} = \Leftrightarrow \mathcal{N}_{U_{k+1}} \Leftrightarrow D_{U_{k+1}})^{-1} \sum_{\beta_k \in \mathcal{F}_k} G_{U_{k+1}}^{\beta_k} \mathbb{1}_{\beta_k} . \quad (3.20)$$

Alors

$$\begin{aligned} & \sum_{\substack{\beta_k \in \mathcal{F}_k \\ \beta_k \neq \alpha_k}} G_{\alpha_k}^{\beta_k} \mathbb{1}_{\beta_k} \Leftrightarrow G_{\alpha_k}^{U_{k+1}} \mathcal{N}_{U_{k+1}} \Leftrightarrow D_{U_{k+1}})^{-1} \sum_{\substack{\beta_k \in \mathcal{F}_k \\ \beta_k \neq \alpha_k}} G_{U_{k+1}}^{\beta_k} \mathbb{1}_{\beta_k} \\ = & \sum_{\substack{\beta_k \in \mathcal{F}_k \\ \beta_k \neq \alpha_k}} G_{\alpha_k}^{\beta_k} \mathbb{1}_{\beta_k} + G_{\alpha_k}^{U_{k+1}} \mathbb{1}_{U_{k+1}} + G_{\alpha_k}^{U_{k+1}} \mathcal{N}_{U_{k+1}} + D_{U_{k+1}})^{-1} G_{U_{k+1}}^{\alpha_k} \mathbb{1}_{\alpha_k} \\ = & G_{\alpha_k}^{O_{k+1}} \mathbb{1}_{O_{k+1}} + G_{\alpha_k}^{U_{k+1}} \mathcal{N}_{U_{k+1}} \Leftrightarrow D_{U_{k+1}})^{-1} G_{U_{k+1}}^{\alpha_k} \mathbb{1}_{\alpha_k} . \end{aligned}$$

D'où

$$\sum_{\beta_k \neq \alpha_k} \lambda_{\alpha_k \beta_k} = \frac{1}{\varepsilon^k} \Pi_{\alpha_k} \left[G_{\alpha_k}^{O_{k+1}} \mathbb{1}_{O_{k+1}} + G_{\alpha_k}^{U_{k+1}} \mathcal{N}_{U_{k+1}} + D_{U_{k+1}})^{-1} G_{U_{k+1}}^{\alpha_k} \mathbb{1}_{\alpha_k} \right] .$$

On en déduit

$$\Lambda_k (c_{\alpha_k})_{\alpha_k \in \mathcal{F}_k}^* = \lambda(0) (c_{\alpha_k})_{\alpha_k \in \mathcal{F}_k}^* + O(\varepsilon_{\mathcal{F}_k}) .$$

Par suite

$$\left(\lim_{\varepsilon \searrow 0} \Lambda_k \right) \lim_{\varepsilon \searrow 0} (c_{\alpha_k})_{\alpha_k \in \mathcal{F}_k}^* = \lambda(0) \lim_{\varepsilon \searrow 0} (c_{\alpha_k})_{\alpha_k \in \mathcal{F}_k}^* .$$

Du fait que la nullité du vecteur $\lim_{\varepsilon \searrow 0} (c_{\alpha_k})_{\alpha_k \in \mathcal{F}_k}^*$ déduit celle de $\lim_{\varepsilon \searrow 0} X$, compte tenu du lemme (3.1.3), on déduit que $\lambda(0)$ est une valeur propre de $\lim_{\varepsilon \searrow 0} \Lambda_k$ ayant $\lim_{\varepsilon \searrow 0} (c_{\alpha_k})_{\alpha_k \in \mathcal{F}_k}^*$ comme vecteur propre associé.

Montrons deuxièmement

$$\#S_k(\Lambda) = \#\mathcal{F}_k .$$

Soit $\Omega_k(\varepsilon)$ l'espace propre associé aux valeurs propres constituant $S_k(\Lambda)$. Soit $\Omega_k(0)$ l'espace engendré par les limites des vecteurs dans $\Omega_k(\varepsilon)$ quand ε

tend vers 0. D'après le lemme 4.10 p34 dans [44], $\Omega_k(\varepsilon)$ et $\Omega_k(0)$ ont la même dimension. Or d'après le lemme 3.1.3, $\Omega_k(0)$ a une dimension plus petite ou égale à $\#\mathcal{F}_k$. Donc il en est de même pour $\Omega_k(\varepsilon)$. C'est-à-dire on a

$$\#S_k(\Lambda) \leq \#\mathcal{F}_k .$$

D'après Kato [44] (cf. II-2.1.4 P.67), $\Omega_1(0)$ est l'espace propre associé à la valeur propre 0 de Λ en $\varepsilon = 0$. Quand $k = 1$, pour tout ensemble de constantes $\{C_{\alpha_1}\}_{\alpha_1 \in \mathcal{F}_1}$, soit X^0 le vecteur sur E tel que pour toute classe α_1 de \mathcal{F}_1

$$X_{\alpha_1}^0 = C_{\alpha_1} \mathbb{1}_{\alpha_1}$$

et

$$X_{U_2}^0 = \Leftrightarrow \lim_{\varepsilon \searrow 0} (\Lambda_{U_2} \Leftrightarrow D_{U_2})^{-1} \sum_{\alpha_1 \in \mathcal{F}_1} C_{\alpha_1} G_{U_2}^{\alpha_1} \mathbb{1}_{\alpha_1}$$

avec $U_2 = E \Leftrightarrow \bigcup_{\alpha_1 \in \mathcal{F}_1} \alpha_1$. D'autre part quand $\varepsilon = 0$, à une permutation près

$$\Lambda(0) = \begin{pmatrix} \Lambda_{U_2}(0) \Leftrightarrow D_{U_2}(0) & G_{U_2}^{\alpha_1}(0)_{\alpha_1 \in \mathcal{F}_1} \\ (0) & \text{diag}(\Lambda_{\alpha_1}(0))_{\alpha_1 \in \mathcal{F}_1} \end{pmatrix} .$$

On en déduit que

$$\Lambda(0)X^0 = 0 .$$

Ce qui signifie que $X^0 \in \Omega_1(0)$ et que la dimension de $\Omega_1(0)$ est plus grande ou égale à $\#\mathcal{F}_1$. Par conséquent,

$$\#S_1(\Lambda) = \#\mathcal{F}_1 .$$

Supposons que pour tout $1 \leq i \leq k \Leftrightarrow 1$, on a

$$\#S_i(\Lambda) = \#\mathcal{F}_i .$$

Considérons le générateur Λ_{k-1} . Soit $\alpha_k \neq \beta_k \in \mathcal{F}_k$, d'après (1), pour tous $\alpha_{k-1} \subseteq \alpha_k, \beta_{k-1} \subseteq \beta_k$ avec $\alpha_{k-1}, \beta_{k-1} \in \mathcal{F}_{k-1}$, on a

$$\lim_{\varepsilon \searrow 0} \lambda_{\alpha_{k-1}\beta_{k-1}} = 0 \text{ et } \lim_{\varepsilon \searrow 0} \lambda_{\beta_{k-1}\alpha_{k-1}} = 0 .$$

Pour toute classe α_k de \mathcal{F}_k , notons $\Lambda_{k-1}^{\alpha_k}$ la restriction de Λ_{k-1} sur l'ensemble des classes de \mathcal{F}_{k-1} contenues dans α_k . On peut alors écrire $\lim_{\varepsilon \searrow 0} \Lambda_{k-1}$ sous la forme suivante.

$$\lim_{\varepsilon \searrow 0} \Lambda_{k-1} = \begin{pmatrix} \text{diag} \left(\lim_{\varepsilon \searrow 0} \Lambda_{k-1}^{\alpha_k} \right)_{\alpha_k \in \mathcal{F}_k} & (0) \\ x & x \end{pmatrix}$$

où x désigne des termes sans intérêt. Donc $\lim_{\varepsilon \searrow 0} \Lambda_{k-1}$ a au moins $\#\mathcal{F}_k$ valeurs propres 0. C'est-à-dire

$$\#S_1(\Lambda_{k-1}) \geq \#\mathcal{F}_k .$$

Par conséquent le nombre des valeurs propres non nulle de $\lim_{\varepsilon \searrow 0} \Lambda_{k-1}$

$$\#S_0(\Lambda_{k-1}) \Leftrightarrow \#S_1(\Lambda_{k-1}) \leq \#S_0(\Lambda_{k-1}) \Leftrightarrow \#\mathcal{F}_k \leq \#\mathcal{F}_{k-1} \Leftrightarrow \#\mathcal{F}_k . \quad (3.21)$$

Puisque toute valeur propre $\varepsilon^{k-1}\lambda(\varepsilon)$ de Λ avec $\lambda(0) \neq 0$ correspond à une valeur propre d'équivalent $\lambda(0)$ de Λ_{k-1} , donc le nombre des valeurs propres $\varepsilon^{k-1}\lambda(\varepsilon)$ avec $\lambda(0) \neq 0$ de Λ est plus petit ou égal au nombre des valeurs propres $\lambda'(0)$ avec $\lambda'(0) \neq 0$ de Λ_{k-1} , i.e.

$$\#S_{k-1}(\Lambda) \Leftrightarrow \#S_k(\Lambda) \leq \#S_0(\Lambda_{k-1}) \Leftrightarrow \#S_1(\Lambda_{k-1}) . \quad (3.22)$$

De (3.21) et (3.22), on déduit

$$\#S_k(\Lambda) \geq \#\mathcal{F}_k .$$

Ce qui termine la démonstration de (3). □

Remarque D'après le théorème A, $\#S_k(\Lambda)$, le nombre de valeurs propres d'ordre $O(\varepsilon^k)$, est égal au nombre de classes minimales d'ordre k , $\#\mathcal{F}_k$. Et donc il y a $\#\mathcal{F}_k \Leftrightarrow \#\mathcal{F}_{k+1}$ valeurs propres d'ordre $\theta(\varepsilon^k)$. Par conséquent, le gap est d'ordre $\theta(\varepsilon^{K-1})$.

Reprenons l'exemple 3.1.1. Du fait que $\#\mathcal{F}_1 = 5$, on déduit qu'il y a 5 valeurs propres d'ordre $O(\varepsilon)$. Déterminons les équivalents des valeurs propres tendant vers 0, quand ε tend vers 0.

Notons

$$\mathcal{F}_1 = \{\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5\}$$

avec

$$\alpha_1 = \{1\}, \alpha_2 = \{8\}, \alpha_3 = \{9\}, \alpha_4 = \{13, 14\}, \alpha_5 = \{17\}.$$

Soit $\Lambda_1 = \lambda_{(\alpha_i \alpha_j)_{1 \leq i, j \leq 5}}$, alors

$$\lim_{\varepsilon \searrow 0} \Lambda_1 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1/4 & \Leftrightarrow & 1/4 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Les valeurs propres de $\lim_{\varepsilon \searrow 0} \Lambda_1$ sont: 0 (de multiplicité 4) et $\Leftrightarrow 0.5$.

Notons

$$\mathcal{F}_2 = \{\beta_1, \beta_2, \beta_3, \beta_4\}$$

avec

$$\beta_1 = \{1, 2\}, \beta_2 = \{7, 8\}, \beta_3 = \{9, 10\}, \beta_4 = \{16, 17\}.$$

Soit $\Lambda_2 = \lambda_{(\beta_i \beta_j)_{1 \leq i, j \leq 4}}$, alors

$$\lim_{\varepsilon \searrow 0} \Lambda_2 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 5/4 & \Leftrightarrow 5/4 \end{pmatrix}.$$

Les valeurs propres de $\lim_{\varepsilon \searrow 0} \Lambda_2$ sont: 0 (de multiplicité 3) et $\Leftrightarrow 1.25$.

Notons

$$\mathcal{F}_3 = \{\gamma_1, \gamma_2, \gamma_3\}$$

avec

$$\gamma_1 = \{1, 2, 3\}, \gamma_2 = \{6, 7, 8\}, \gamma_3 = \{9, 10, 11\}.$$

Soit $\Lambda_3 = \cancel{\lambda_{\gamma_i \gamma_j}}_{1 \leq i, j \leq 3}$, alors

$$\lim_{\varepsilon \searrow 0} \Lambda_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \Leftrightarrow 1 & 1 \\ 0 & 1 & \Leftrightarrow 1 \end{pmatrix}.$$

Les valeurs propres de $\lim_{\varepsilon \searrow 0} \Lambda_3$ sont: 0 (de multiplicité 2) et $\Leftrightarrow 2$.

Notons

$$\mathcal{F}_4 = \{\delta_1, \delta_2\}$$

avec

$$\delta_1 = \{1, 2, 3, 4\}, \delta_2 = \{5, \dots, 17\}.$$

Soit $\Lambda_4 = \cancel{\lambda_{\gamma_i \gamma_j}}_{1 \leq i, j \leq 2}$, alors

$$\lim_{\varepsilon \searrow 0} \Lambda_4 = \begin{pmatrix} \Leftrightarrow 1 & 1 \\ 1/2 & \Leftrightarrow 1/2 \end{pmatrix}.$$

Les valeurs propres de $\lim_{\varepsilon \searrow 0} \Lambda_4$ sont: 0 et $\Leftrightarrow 1.5$.

Finalement les équivalents des 5 valeurs propres d'ordre $O(\varepsilon)$ du générateur initial sont

$$\Leftrightarrow 0.5\varepsilon + o(\varepsilon), \Leftrightarrow 1.25\varepsilon^2 + o(\varepsilon^2), \Leftrightarrow 2\varepsilon^3 + o(\varepsilon^3), \Leftrightarrow 1.5\varepsilon^4 + o(\varepsilon^4), 0.$$

3.1.5 Discussion

De nombreux problèmes, comme le comportement des systèmes complexes ([67, 18]), l'algorithme du recuit simulé à basse température ([35, 57]), le processus de Ising ([64, 65]) et les réseaux neuronaux ([24, 48]), se ramènent à étudier le comportement d'un processus (ou une chaîne) de Markov perturbé qui a ainsi suscité beaucoup d'études depuis les années 60. Les points d'intérêts en sont essentiellement les suivants: le comportement spectral ([44, 19, 27, 57]), le comportement métastable ([27, 11, 59]) et l'évolution aléatoire du processus ([67, 18, 16, 15, 74]). Il nous paraît utile de comparer le théorème A et la technique de base de sa démonstration avec les résultats existants sur le comportement spectral et le même type de technique utilisé par d'autres.

Les résultats dans [19] ont montré l'existence des générateurs $\Lambda_k(\varepsilon)$ conservant les équivalents des valeurs propres d'ordre $\theta(\varepsilon^k)$ de $\Lambda(\varepsilon)$ dans un cadre plus général où $\Lambda(\varepsilon)$ n'est pas forcément irréductible et réversible que le théorème A. Mais pour déduire le nombre de valeurs propres d'ordre $\theta(\varepsilon^k)$ de $\Lambda(\varepsilon)$ et les équivalents de ces valeurs propres, il nécessite la connaissance de $\Lambda_k(\varepsilon)$ dont la définition récursive oblige la détermination de tous les $\Lambda_i(\varepsilon)$ pour $i \leq k$. En fait, pour tout $i \leq k$, le générateur $\Lambda_i(\varepsilon)$ défini par Delebecque a contenu les informations des valeurs propres d'ordre $\theta(\varepsilon^j)$ de $\Lambda(\varepsilon)$ pour tout $j \geq i$. Or ce qui nous faut est seulement les équivalents des valeurs propres d'ordre $\theta(\varepsilon^k)$ de $\Lambda(\varepsilon)$. Dans le cas où $\Lambda(\varepsilon)$ est irréductible et réversible, la complexité de la détermination de $\Lambda_k(0)$ proposée par Delebecque peut être beaucoup simplifiée par le théorème A au travers de l'analyse de la structure hiérarchique de $\Lambda(\varepsilon)$. Dans le théorème A, le nombre des valeurs propres d'ordre $\theta(\varepsilon^k)$ se déduit directement du cardinal de \mathcal{F}_k , la classe minimale d'ordre k par rapport à $\Lambda(\varepsilon)$. Et pour obtenir les équivalents des valeurs propres d'ordre $\theta(\varepsilon^k)$, il suffit de déterminer $\Lambda_k(0)$ qui est défini directement à partir de $\Lambda(\varepsilon)$. Comparons grossièrement la complexité du calcul de $\Lambda_k(0)$ proposé par Delebecque et celle selon le théorème A. On note C_i^j la complexité de tirer les informations des valeurs propres d'ordre $\theta(\varepsilon^j)$ à partir du générateur $\Lambda_i(\varepsilon)$ avec $j \geq i$. Alors la complexité de la détermination de $\Lambda_k(0)$ d'après Delebecque est

$$\sum_{j=1}^K C_0^j + \sum_{j=2}^K C_1^j + \cdots + \sum_{j=k-1}^K C_{k-2}^j + C_{k-1}^k .$$

Et la complexité de la détermination de $\Lambda_k(0)$ selon le théorème A est

$$\text{complexité de détermination de } \mathcal{F}_k + C_0^k .$$

La technique de base qui nous permet de tirer $\Lambda_k(0)$ de $\Lambda(\varepsilon)$ directement est d'analyser la structure hiérarchique de $\Lambda(\varepsilon)$ ce qui n'est pas nouveau. Freidlin-Wentzell [27] a introduit la notion de cycles qui a inspiré par suite beaucoup de travaux dans de nombreux domaines: [5, 37, 54, 57, 8]. Dans [57], Miclo a considéré le recuit simulé à basse température dont le générateur $\Lambda(\varepsilon) =$

$(\lambda_{ab})_{a,b \in E}$ est analytique en $\varepsilon = 0$ mais sous contrainte

$$\forall \lambda_{ab} > 0, \quad d(\lambda_{ab}) \, d(\lambda_{ba}) = 0. \quad (3.23)$$

En utilisant la notion de k -cycle, il a donné aisément le nombre de valeurs propres d'ordre $\theta \varepsilon^k$ de $\Lambda(\varepsilon)$ qui est tout simplement le nombre de k -cycles (voir le paragraphe 2.3.3). En fait, il y a une bijection entre l'ensemble des k -cycles et l'ensemble des classes minimales d'ordre k . Or la notion de k -cycle n'est plus valable pour un générateur sans la contrainte 3.23. La description récurrente des cycles dans le cas du recuit simulé généralisé est due à Hwang et Sheu ([37]). Elle a en suite été reprise par Trouvé ([73])

3.2 Applications

3.2.1 Processus quasi décomposables

La notion de quasi décomposabilité semble avoir été introduite par les économistes Simon et Ando [67] dans le cadre de l'analyse hiérarchique de systèmes complexes. L'idée, fort naturelle, a ensuite été appliquée en sociologie, biologie, et de manière très poussée pour l'analyse de performance des réseaux informatiques (pour un historique complet, voir l'introduction de Courtois [18]). Cette idée est la suivante. Dans un système dynamique complexe coexistent plusieurs échelles de temps. Pour fixer les idées, parlons de secondes, heures et mois. A l'échelle de quelques secondes, le système peut être décomposé en sous systèmes isolés qui n'ont que des interactions internes. Ce n'est qu'à l'échelle de plusieurs heures que des interactions sensibles apparaissent entre des composantes jusque là isolées. Ceci induit un regroupement des anciennes composantes en de nouvelles composantes qui constituent une partition du système global. Certaines de ces composantes interagissent à l'échelle du mois, pour constituer par regroupement une partition plus grossière. On obtient ainsi une hiérarchie de partitions emboîtées dont chacune correspond à une échelle de temps différente.

L'idée est formalisée à l'aide de chaînes de Markov dans le premier chapitre de Courtois [18]. Nous préférons le formalisme des processus à temps continu. Les

résultats ci-dessous s'adaptent de manière immédiate aux matrices de transition des chaînes à temps discret.

Soit $\Lambda' = (\lambda'_{ij})_{i,j \in E}$ un générateur irréductible sur l'espace d'états E . Le générateur Λ' est supposé Π -réversible où $\Pi = (\Pi(i))_{i \in E}$. Considérons $\{\mathcal{F}_k\}_{1 \leq k \leq K}$ un ensemble de partitions de l'espace d'états E tel que les conditions suivantes soient vérifiées.

a) Pour $1 \leq k \leq K \Leftrightarrow 1, \forall \alpha_k \in \mathcal{F}_k, \exists \alpha_{k+1} \in \mathcal{F}_{k+1}$ tel que $\alpha_k \subseteq \alpha_{k+1}$.

Autrement dit, toute classe α_{k+1} de \mathcal{F}_{k+1} est une union de classes de \mathcal{F}_k .

b) Les \mathcal{F}_i sont différentes,

$$i \neq j \iff \mathcal{F}_i \neq \mathcal{F}_j .$$

c) La partition \mathcal{F}_K réduite à un élément,

$$\mathcal{F}_K = \{E\} .$$

Ces partitions \mathcal{F}_k sont donc emboîtées. On suppose que le générateur restreint de Λ' à chaque classe de ces partitions est irréductible. Définissons le nouveau générateur $\Lambda = (\lambda_{ij})_{i,j \in E}$ tel qu'il est analytique en $\varepsilon = 0$ de la façon suivante.

Pour tous i, j de E , posons $k = 0$ si i, j sont dans une même classe d'ordre 1. Sinon, soit k l'entier le plus grand tel que i, j ne sont pas dans une même classe d'ordre k . Alors λ_{ij} est défini par

$$\lambda_{ij} = \varepsilon^k \lambda'_{ij} .$$

On dit que le générateur Λ ainsi défini est *quasi décomposable*.

Puisque le générateur Λ est obtenu en multipliant λ'_{ij} et λ'_{ji} par un même coefficient, on a

$$\Pi(i)\lambda_{ij} = \Pi(j)\lambda_{ji} .$$

Par conséquent, Λ est aussi Π -réversible. Du fait

$$\lambda'_{ij} \neq 0 \iff \lambda_{ij} \neq 0 ,$$

on en déduit que le générateur restreint de Λ sur α_k de \mathcal{F}_k est irréductible et $\Pi_{\alpha_k} \Leftrightarrow$ réversible où Π_{α_k} se définit sur α_k par

$$\forall i \in \alpha_k, \Pi_{\alpha_k}(i) = \frac{\Pi(i)}{\Pi(\alpha_k)} .$$

En plus, on constate que \mathcal{F}_k est exactement l'ensemble des classes minimales d'ordre k associé à Λ défini dans le paragraphe 3.1.1 et que $\mathcal{E}_k = \mathcal{F}_k$. Dans ce cas là, l'expression de la limite du générateur $\Lambda_k = (\lambda_{\alpha_k \beta_k})_{\alpha_k, \beta_k \in \mathcal{F}_k}$ défini dans le théorème A (page 96) est beaucoup simplifiée.

Proposition 3.2.1 *Si α_k, β_k sont dans une même classe d'ordre $k + 1$, alors*

$$\lim_{\varepsilon \searrow 0} \lambda_{\alpha_k \beta_k} = \sum_{\substack{i \in \alpha_k \\ j \in \beta_k}} \Pi_{\alpha_k} \lambda'_{ij} .$$

Si non

$$\lim_{\varepsilon \searrow 0} \lambda_{\alpha_k \beta_k} = 0 .$$

Démonstration. Il suffit de constater que dans l'expression de $\lambda_{\alpha_k \beta_k}$ dans le théorème A, l'ensemble

$$U_{k+1} = E \Leftrightarrow \bigcup_{\alpha_k \in \mathcal{F}_k} \alpha_k = \emptyset ,$$

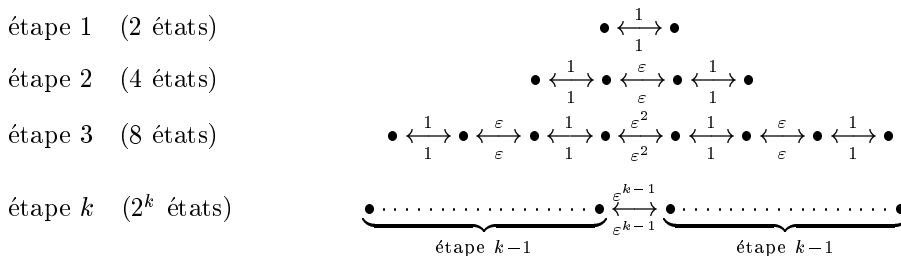
et de plus si α_k, β_k sont dans une même classe d'ordre $k + 1$, on a

$$\forall i \in \alpha_k, \forall j \in \beta_k, \lambda_{ij} = \varepsilon^k \lambda'_{ij} .$$

□

3.2.2 Processus de naissance et de mort à homothétie interne

Comme exemple des processus quasi décomposables, nous décrivons une famille de processus de naissance et de mort sur $\{1, \dots, 2^k\}$, présentant une hiérarchisation des échelles de temps. La famille est construite de manière itérative selon les diagrammes de transition suivants.



Plus précisément, soit $\Lambda(1)$ le générateur symétrique sur $\{1, 2\}$ avec

$$\Lambda(1) = \begin{pmatrix} \Leftrightarrow 1 & 1 \\ 1 & \Leftrightarrow 1 \end{pmatrix} .$$

Pour tout $k \geq 1$, si $\Lambda(k \Leftrightarrow 1)$ est défini sur $\{1, \dots, 2^{k-1}\}$, définissons le générateur $\Lambda(k)$ sur $\{1, \dots, 2^k\}$ par

$$\Lambda(k) = \begin{pmatrix} \Lambda(k \Leftrightarrow 1) & (0) \\ (0) & \Lambda(k \Leftrightarrow 1) \end{pmatrix} + \varepsilon^{k-1} B_{2^k}$$

où B_{2^k} est le générateur dont les taux de transition entre les deux états 2^{k-1} et $2^{k-1} + 1$ valent 1, les autres étant nuls.

Pour tout $k \geq 1$, le générateur $\Lambda(k)$ est irréductible et symétrique. Il est donc réversible par rapport à la mesure uniforme sur $\{1, \dots, 2^k\}$. Pour décrire le comportement des processus correspondants, supposons que l'unité de temps est la seconde et fixons $\varepsilon = 1/3600$. Considérons le processus à 8 états (étape 3) et supposons qu'il parte de 1 à l'instant 0. A l'échelle de la seconde, le processus va osciller entre 1 et 2, et donc se comporter comme le processus de l'étape 1. Ce n'est qu'au bout d'une heure en moyenne qu'il passera de l'état 2 à l'état 3, et se mettra à osciller entre 3 et 4 toutes les secondes environ. A l'échelle de l'heure, le processus se comportera comme celui de l'étape 2. Il faudra attendre environ 3600 heures (150 jours) pour le voir quitter les états $\{1, 2, 3, 4\}$ et visiter l'autre moitié de l'espace d'états.

On vérifie que les générateurs ainsi construits sont quasi décomposables. La proposition suivante décrit le spectre de $\Lambda(k)$ quand ε tend vers 0.

Proposition 3.2.2 *Pour $k \geq 1$, le générateur $\Lambda(k)$ a 2^{k-1} valeurs propres d'équivalent $\Leftrightarrow 2 + O(\varepsilon)$ et 2^{k-i-1} valeurs propres d'équivalent $\Leftrightarrow \varepsilon^i / 2^{i-1} + O(\varepsilon^{i+1})$*

pour $i = 1, \dots, k \Leftrightarrow 1$. Donc

$$\text{gap}(\Lambda(k)) = \frac{\varepsilon^{k-1}}{2^{k-2}} + O(\varepsilon^k) .$$

Démonstration. Pour simplifier, posons $\Lambda = \Lambda(k)$. Constatons que

$$\lim_{\varepsilon \searrow 0} \Lambda = \text{diag}(J)_{2^{k-1} \times 2^{k-1}}$$

avec

$$J = \begin{pmatrix} \Leftrightarrow 1 & 1 \\ 1 & \Leftrightarrow 1 \end{pmatrix} .$$

Comme J a pour valeurs propres 0 et $\Leftrightarrow 2$, $\lim_{\varepsilon \searrow 0} \Lambda$ a alors pour valeurs propres 0 et $\Leftrightarrow 2$ de multiplicité 2^{k-1} toutes les deux. Alors Λ possède 2^{k-1} valeurs propres d'équivalent $\Leftrightarrow 2 + O(\varepsilon)$ et 2^{k-1} d'équivalent $O(\varepsilon)$.

Notons

$$\alpha_i^j = \{2^i j + 1, 2^i j + 2, \dots, 2^i(j+1)\} ,$$

alors pour $i = 1, \dots, k \Leftrightarrow 1$, \mathcal{F}_i l'ensemble de classes minimales d'ordre i et la partition \mathcal{E}_i de E sont les suivantes:

$$\mathcal{F}_i = \left\{ \alpha_i^j ; j = 0, \dots, 2^{k-i} \Leftrightarrow 1 \right\} \text{ et } \mathcal{E}_i = \mathcal{F}_i .$$

D'où il y a 2^{k-i} valeurs propres d'ordre $O(\varepsilon^i)$ et donc 2^{k-i-1} valeurs propres sont égales à ε^i à une constante près. Déterminons ces constantes. Soit $\lambda_{\alpha_i^{j_1} \alpha_i^{j_2}}$ les taux de transition de $\alpha_i^{j_1}$ à $\alpha_i^{j_2}$ qui constituent le générateur Λ_i défini sur \mathcal{F}_i que donne le théorème A. Pour $j_1 < j_2$, on a

$$\lim_{\varepsilon \searrow 0} \lambda_{\alpha_i^{j_1} \alpha_i^{j_2}} = \begin{cases} 1/2^i & \text{si } j_1 \text{ est paire et } j_2 = j_1 + 1 \\ 0 & \text{sinon .} \end{cases}$$

Evidemment Λ_i est symétrique, alors

$$\lim_{\varepsilon \searrow 0} \Lambda_i = \frac{1}{2^i} \text{diag}(J)_{2^{k-i-1} \times 2^{k-i-1}} .$$

D'où $\lim_{\varepsilon \searrow 0} \Lambda_i$ a pour valeurs propres 0 et $\Leftrightarrow 1/2^{i-1}$ de multiplicité 2^{k-i-1} toutes les deux. Donc il y a 2^{k-i-1} valeurs propres d'équivalent

$$\Leftrightarrow \varepsilon^i / 2^{i-1} + o(\varepsilon^i) .$$

On en déduit

$$\text{gap}(\Lambda) = \frac{\varepsilon^{k-1}}{2^{k-2}} + o(\varepsilon^{k-1}) .$$

□

Sur des exemples numériques, on constate que cet équivalent fournit une approximation correcte du gap, même pour des valeurs de ε de l'ordre de plusieurs dixièmes. Ces vérifications numériques ont été faites à l'aide de Mathematica [75]. Les tables ci-dessous donnent des valeurs de $\text{gap}(\Lambda(k))$ et de son estimation par $\varepsilon^{k-1}/2^{k-2}$, pour $k = 3$ et 4 (8 et 16 états respectivement).

ε	0,05	0,10	0,15	0,20	0,25	0,30	0,35	0,40	0,45	0,50
$\text{gap}(\Lambda(3))$	0,0012	0,0047	0,0101	0,0172	0,0255	0,0347	0,0444	0,0544	0,0645	0,0744
$\varepsilon^2/2$	0,0013	0,0050	0,0113	0,0200	0,0312	0,0450	0,0612	0,0800	0,1012	0,1250

ε	0,05	0,10	0,15	0,20	0,25	0,30	0,35	0,40	0,45	0,50
$\text{gap}(\Lambda(4))$	0,0000	0,0002	0,0008	0,0017	0,0031	0,0049	0,0071	0,0096	0,0123	0,0151
$\varepsilon^3/4$	0,0000	0,0003	0,0008	0,0020	0,0039	0,0068	0,0107	0,0160	0,0228	0,0313

3.2.3 Recuit simulé à basse température

L'algorithme de Metropolis [56] a suscité une abondante littérature ces dix dernières années (voir [22] et les références qu'il contient). L'intérêt pour cet algorithme provient de ces nombreuses applications, du recuit simulé à l'échantillonnage de Gibbs ([46, 32, 2, 57, 58, 29, 38]). Le problème théorique le plus important posé par ce type de technique est la détermination précise de la vitesse de convergence. Une présentation générale de l'ensemble des résultats rigoureux connus à ce jour est donnée par Diaconis et Saloff-Coste [22]. La plupart portent sur des minoration du trou spectral des chaînes de Markov simulées ou sur des inégalités de Sobolev logarithmiques. Ces résultats ne sont évidemment que des pis-allers, à défaut de pouvoir connaître l'ensemble du spectre de la chaîne, ce qu'interdit dans le cas général la taille de l'espace des états. Une description complète du spectre a cependant été obtenue par Miclo [57] et dans des cas particuliers par Diaconis et Saloff-Coste [22] et Liu [51]. C'est un résultat de

ce type que nous proposons ici, dans le cas où certaines transitions sont très faibles. Un cas typique d'application est celui de l'algorithme du recuit simulé à basse température. Nous nous placerons donc dans ce cadre et décrivons les équivalents des différentes valeurs propres à basse température en appliquant le théorème A.

Rappelons tout d'abord la description classique de l'algorithme du recuit simulé. Soit $E = \{a, b, \dots\}$ un ensemble fini, et H une fonction de E dans \mathbb{R} que l'on souhaite minimiser. Cette fonction est interprétée comme une énergie et la mesure de Gibbs correspondante, notée Π est définie par

$$\forall a \in E, \quad \Pi(a) = \frac{1}{Z} \exp \left(-\frac{H(a)}{T} \right),$$

où T désigne un paramètre de température, et Z la constante de normalisation (fonction de partition). Quand T tend vers 0, la famille de mesures ainsi définie converge vers la loi uniforme sur l'ensemble des minima globaux de la fonction H (proposition 2.2.1). L'algorithme du recuit simulé est une variante de l'algorithme de Metropolis qui simule une chaîne de Markov pour laquelle Π est une mesure réversible. En toute généralité, l'algorithme de Metropolis (voir 2.2.1) est une méthode de simulation par rejet à partir d'un noyau Markovien Λ' , qui n'est pas nécessairement symétrique (voir [22]). Dans la plupart des applications, le noyau Λ' est celui de la marche aléatoire symétrique sur un graphe non orienté dont E est l'ensemble des sommets. C'est dans ce cas que nous nous placerons. L'ensemble des arêtes de ce graphe est noté A .

$$A \subset \{(a, b); a, b \in E\}.$$

La relation de voisinage sera notée \sim .

$$a \sim b \iff (a, b) \in A.$$

L'algorithme est vu ici comme un processus de Markov à temps continu, et décrit par ses taux de transition, qui sont non nuls seulement sur les arêtes du

graphe.

$$\forall a \sim b \in E, \quad \lambda_{ab} = \begin{cases} 1 & \text{si } H(b) \leq H(a) \\ \exp \left[\frac{H(b) \Leftrightarrow H(a)}{T} \right] & \text{si } H(b) > H(a) . \end{cases}$$

Il est immédiat de vérifier que le générateur ainsi défini est Π -réversible. Afin de se placer dans le cadre où le générateur est analytique, nous supposons que la fonction H ne peut prendre que des valeurs entiers. Une translation de H ne change ni la mesure ni le générateur. Supposons en plus que

$$\min_{a \in E} H(a) = 0 .$$

Posons ρ la plus grande valeur positive telle que pour tous $a, b \in E$,

$$\frac{H(a) \Leftrightarrow H(b)}{\rho} \in \mathbb{N} .$$

Notre but est d'étudier le spectre de ce générateur à basse température. Nous poserons donc

$$\varepsilon = \exp \frac{\rho}{T} .$$

et désignerons par Λ le générateur correspondant qui est défini donc sur E de la façon suivante

$$\forall a, b \in E, \quad \lambda_{ab} = \begin{cases} 0 & \text{si } a \not\sim b \\ 1 & \text{si } a \sim b \text{ et } H(b) \leq H(a) \\ \varepsilon^k & \text{si } a \sim b \text{ et } [H(b) \Leftrightarrow H(a)] / \rho = k . \end{cases}$$

Et la mesure réversible Π de Λ est définie par

$$\forall a \in E, \quad \Pi(a) = \frac{1}{Z} \varepsilon^{H(a)/\rho} .$$

Etudions la structure hiérarchique associée au générateur Λ . En fait dans le cas du recuit simulé la construction des fonctions d'énergie H_k , des partitions de l'espace d'états E , \mathcal{E}_k et des ensembles de classes minimales \mathcal{F}_k est beaucoup simplifiée. On constate que dans ce cas là, le générateur est complètement déterminé par l'ensemble des arêtes A et la fonction d'énergie H . Or les fonctions d'énergie H_1 et H sont identiques à un coefficient près,

$$H_1 = \frac{H}{\rho} .$$

Donc \mathcal{E}_1 et \mathcal{F}_1 sont déterminés par A et H_1 . Itérativement, on montre que les $\mathcal{E}_k, \mathcal{F}_k$ sont déterminés par A et H_k . On dira que deux sous-ensembles S_1, S_2 de E avec $S_1 \cap S_2 = \emptyset$ communiquent, si il existe $a \in S_1, b \in S_2$ tels que $a \sim b$.

Proposition 3.2.3 *Pour le générateur Λ , nous avons les propriétés suivantes.*

(1) *Pour toutes classes α_k, β_k de \mathcal{F}_k avec $\alpha_k \neq \beta_k$,*

$$\delta[\alpha_k, \beta_k] = \infty \text{ et } \delta[\beta_k, \alpha_k] = \infty .$$

Autrement dit, α_k, β_k ne communiquent pas.

(2) *Soit $\beta_k \in \mathcal{E}_k \setminus \mathcal{F}_k$. Si $d(\lambda_{ba}) = 0$ avec $b \in \beta_k, a \notin \beta_k$, alors*

$$\delta_k[b, a] = \min\{\delta_k[b', a']; b' \in \beta_k, a' \notin \beta_k\} .$$

(3) *Définissons la relation \mathcal{R}_k par $\forall a, b \in E, i\mathcal{R}_k j$ si et seulement si*

$$\exists a = c_1 \sim c_2 \sim \dots \sim c_l = b \text{ t.q. } H_k(c_1) = \dots = H_k(c_l) .$$

Alors \mathcal{E}_k est l'ensemble des classes d'équivalence par rapport à \mathcal{R}_k .

(4) *Soit $\alpha_k \in \mathcal{E}_k$. Alors α_k est minimale si et seulement si pour toute classe β_k de \mathcal{E}_k communiquant avec α_k , on a*

$$h(\beta_k) > h(\alpha_k) .$$

Autrement dit, $h(\alpha_k)$ est un minimum local de H_k .

Démonstration. On constate que dans le cas du recuit simulé, pour tous $a, b \in E$, les ordres de grandeur de λ_{ab} et λ_{ba} ont les trois possibilités suivantes: soit

$$d(\lambda_{ab}) = d(\lambda_{ba}) = \infty$$

ce qui signifie que $a \not\sim b$, soit

$$d(\lambda_{ab}) = d(\lambda_{ba}) = 0$$

ce qui signifie que $a \sim b$ et $H_1(a) = H_1(b)$, soit

$$d(\lambda_{ab}) \geq 1, d(\lambda_{ba}) = 0 \text{ ou } d(\lambda_{ab}) = 0, d(\lambda_{ba}) \geq 1$$

ce qui signifie que $a \sim b$ et $H_1(a) \neq H_1(b)$. Donc si $a \sim b$, au moins l'un de $d(\lambda_{ab}), d(\lambda_{ba})$ est 0.

Montrons la propriété (1).

Montrons d'abord

$$\delta[\alpha_k, \beta_k] \geq k \text{ et } \delta[\beta_k, \alpha_k] \geq k .$$

D'après (5) de la proposition 3.1.2, ceci est vrai si $h(\alpha_k) = h(\beta_k)$.

Si $h(\alpha_k) < h(\beta_k)$, on a

$$\begin{aligned} \delta[\alpha_k, \beta_k] &\geq h(\beta_k) \Leftrightarrow B(\alpha_k) \text{ (d'après (4) de la proposition 3.1.2)} \\ &\geq 1 + h(\alpha_k) \Leftrightarrow B(h_k) \\ &\geq k \text{ (3) de la proposition 3.1.2),} \end{aligned}$$

et d'après la définition de classe minimale, on a

$$\delta[\beta_k, \alpha_k] \geq k .$$

Par conséquent, pour tous $a \in \alpha_k, b \in \beta_k$,

$$\delta_k[a, b] \geq k ,$$

par suite

$$\begin{aligned} d(\lambda_{ab}) &= \delta_k[a, b] \Leftrightarrow (H_1(a) \Leftrightarrow B(\alpha_k)) \\ &\geq k \Leftrightarrow [H_1(a) \Leftrightarrow B(\alpha_k)] \\ &\geq k \Leftrightarrow [h(\alpha_k) \Leftrightarrow B(\alpha_k)] \\ &\geq 1 \text{ (3) de la proposition 3.1.2).} \end{aligned}$$

De même,

$$d(\lambda_{ba}) \geq 1 .$$

On en déduit $a \not\sim b$.

Montrons la propriété (2).

Montrons premièrement que pour toute classe non minimale β_k d'ordre k , si $b \in \beta_k, a \notin \beta_k$ sont tels que

$$\delta_k[b, a] = \min\{\delta_k[b', a']; b' \in \beta_k, a' \notin \beta_k\},$$

alors

$$d(\lambda_{ba}) = 0.$$

Quand $k = 1$, comme β_1 est non-minimale,

$$\min\{\delta_1[b', a']; b' \in \beta_1, a' \notin \beta_1\} = 0,$$

donc ce qu'on voulait démontrer est vrai pour $k = 1$.

Supposons que c'est vrai pour $i \leq k \Leftrightarrow 1$. Si β_k est de type 2, posons β_{k-1} la classe d'ordre $k \Leftrightarrow 1$ telle que $\beta_{k-1} = \beta_k$. On a alors

$$\delta_{k-1}[b, a] = \min\{\delta_{k-1}[b', a']; b' \in \beta_{k-1}, a' \notin \beta_{k-1}\}.$$

Par récurrence

$$d(\lambda_{ba}) = 0.$$

Si β_k est de type 1, comme β_k est non-minimale, forcément

$$\min\{\delta_k[b', a']; b' \in \beta_k, a' \notin \beta_k\} \leq k \Leftrightarrow 1.$$

Par suite

$$\delta_k[b, a] \leq k \Leftrightarrow 1. \quad (3.24)$$

Posons

$$\gamma_k = \mathcal{E}_k(a).$$

D'après (1) de la proposition 3.1.2, on a $h(\gamma_k) > h(\beta_k)$. Alors

$$\begin{aligned} & \delta[\beta_k, \gamma_k] \\ & \geq h(\gamma_k) \Leftrightarrow B(\beta_k) \quad ((3) \text{ de la proposition 3.1.2}) \\ & \geq 1 + h(\beta_k) \Leftrightarrow B(\beta_k) \\ & \geq k \quad ((3) \text{ de la proposition 3.1.2}), \end{aligned}$$

ce qui contredit (3.24).

Montrons deuxièmement que si $d(\lambda_{ba}) = 0$ avec $b \in \beta_k, a \notin \beta_k$, alors

$$H_1(b) = h(\beta_k) .$$

Supposons que l'affirmation ci-dessus n'est pas vraie. Soit alors i l'entier le plus grand tel que

$$H_i(b) = H_k(b) \Leftrightarrow 1 .$$

Posons

$$\beta_i = \mathcal{E}_i(b) \text{ et } \gamma_i = \mathcal{E}_i(a) .$$

Forcément β_i est minimale.

Si $h(\gamma_i) > h(\beta_i)$, on a

$$\begin{aligned} d(\lambda_{ba}) &\geq \delta[\beta_i, \gamma_i] \Leftrightarrow H_1(b) + B(\beta_i) \\ &\geq h(\gamma_i) \Leftrightarrow H_1(b) \quad ((4) \text{ de la proposition 3.1.2}) \\ &\geq 1 , \end{aligned}$$

ce qui est contradictoire.

Si $h(\gamma_i) = h(\beta_i)$, d'après (5) de la proposition 3.1.2,

$$\begin{aligned} d(\lambda_{ba}) &\geq \delta[\beta_i, \gamma_i] \Leftrightarrow H_1(b) + B(\beta_i) \\ &\geq i \Leftrightarrow H_1(b) + B(\beta_i) \quad ((5) \text{ de la proposition 3.1.2}) \\ &\geq i \Leftrightarrow h(\beta_i) + B(\beta_i) \\ &\geq 1 \quad ((3) \text{ de la proposition 3.1.2}), \end{aligned}$$

ce qui est contradictoire.

Si $h(\gamma_i) < h(\beta_i)$,

$$\begin{aligned} d(\lambda_{ba}) &\geq \delta[\beta_i, \gamma_i] \Leftrightarrow H_1(b) + B(\beta_i) \\ &\geq i \Leftrightarrow H_1(b) + B(\beta_i) \quad (\text{définition de classe minimale}) \\ &\geq i \Leftrightarrow h(\beta_i) + B(\beta_i) \\ &\geq 1 \quad ((3) \text{ de la proposition 3.1.2}), \end{aligned}$$

ce qui est contradictoire. D'où le deuxième point de la démonstration.

Du fait que pour tous $b \in \beta_k, a \notin \beta_k$ tel que $d(\lambda_{ba}) = 0$,

$$\delta_k[b, a] = h(\beta_k) \Leftrightarrow B(\beta_k)$$

qui est constant, on en déduit la propriété (2).

Montrons la propriété (3).

Evidemment, toute classe de \mathcal{E}_k appartient à une classe d'équivalence par rapport à \mathcal{R}_k . Donc il suffit de montrer que pour toutes classes $\alpha_k \neq \beta_k \in \mathcal{E}_k$ telles que α_k, β_k communiquent, on a $h(\alpha_k) \neq h(\beta_k)$.

Pour $k = 1, \forall a \in \alpha_1, \forall b \in \beta_1$ avec $a \sim b$, on a

$$d(\Pi(a)) = h(\alpha_1) \text{ et } d(\Pi(b)) = h(\beta_1) .$$

Si $h(\alpha_1) = h(\beta_1)$, alors $\Pi(a) = \Pi(b)$, d'où

$$d(\lambda_{ab}) = d(\lambda_{ba}) = 0$$

ce qui signifie que a, b sont dans une même classe d'ordre 1. Ce qui contredit l'hypothèse. Donc on a $h(\alpha_1) \neq h(\beta_1)$.

Supposons que pour toutes classes $\alpha_{k-1} \neq \beta_{k-1} \in \mathcal{E}_{k-1}$ telles que $\alpha_{k-1}, \beta_{k-1}$ communiquent, on a $h(\alpha_{k-1}) \neq h(\beta_{k-1})$. Soit $\alpha_k \neq \beta_k \in \mathcal{E}_k$ telles que α_k, β_k communiquent.

Si α_k, β_k sont toutes les deux de type 2, alors elles sont aussi classes d'ordre $k \Leftrightarrow 1$, par récurrence, on a $h(\alpha_k) \neq h(\beta_k)$.

Si α_k, β_k sont toutes les deux de type 1, supposons que $h(\alpha_k) = h(\beta_k)$. D'après (5) de la proposition 3.1.2, on a

$$\delta[\alpha_k, \beta_k] \geq k .$$

Donc $\forall a \in \alpha_k, \forall b \in \beta_k$,

$$\begin{aligned} d(\lambda_{ab}) &\geq \delta[\alpha_k, \beta_k] \Leftrightarrow H_1(a) + B(\alpha_k) \\ &\geq k \Leftrightarrow h(\alpha_k) + B(\alpha_k) \\ &\geq 1 \text{ ((5) de la proposition 3.1.2).} \end{aligned}$$

De même

$$d(\lambda_{ba}) \geq 1 .$$

On en déduit

$$d(\lambda_{ab}) = d(\lambda_{ba}) = \infty$$

ce qui contredit le fait que α_k, β_k communiquent. Par conséquent, on a $h(\alpha_k) \neq h(\beta_k)$.

Si α_k est de type 1 et β_k de type 2, supposons que $h(\alpha_k) = h(\beta_k)$. D'après (5) de la proposition 3.1.2. Du même raisonnement que précédemment, $\forall a \in \alpha_k, \forall b \in \beta_k$ avec $a \sim b$, on a $d(\lambda_{ab}) \geq 1$. Donc

$$d(\lambda_{ba}) = 0 .$$

Posons

$$\alpha_{k-1} = \mathcal{E}_{k-1}(a) \text{ et } \beta_{k-1} = \mathcal{E}_{k-1}(b) .$$

Comme β_k est de type 2, alors

$$\beta_{k-1} = \beta_k .$$

Puisque $\alpha_{k-1}, \beta_{k-1}$ communiquent, on a $h(\alpha_{k-1}) \neq h(\beta_{k-1})$, forcément

$$h(\alpha_{k-1}) = h(\beta_{k-1}) \Leftrightarrow 1 ,$$

par conséquent α_{k-1} est minimale. Comme $d(\lambda_{ba}) = 0$, d'après (2),

$$\delta_{k-1}[b, a] = \min\{\delta_{k-1}[b', a']; b' \in \beta_{k-1}, a' \notin \beta_{k-1}\} .$$

On a alors

$$\begin{aligned} \delta[\alpha_{k-1}, \beta_{k-1}] &= h(\beta_{k-1}) \Leftrightarrow B(\alpha_{k-1}) \quad ((1) \text{ de la proposition 3.1.2}) \\ &= 1 + h(\alpha_{k-1}) \Leftrightarrow B(\alpha_{k-1}) \\ &= k \Leftrightarrow 1 \quad ((3) \text{ de la proposition 3.1.2}). \end{aligned}$$

On en déduit que $\alpha_{k-1}, \beta_{k-1}$ sont dans une même classe d'ordre k ce qui est contradictoire. Conséquemment, on a $h(\alpha_k) \neq h(\beta_k)$.

Montrons la propriété (4).

D'après (3), $h(\alpha_k) \neq h(\beta_k)$. Supposons $h(\alpha_k) > h(\beta_k)$. Comme α_k est minimale, alors

$$\delta[\alpha_k, \beta_k] \geq k .$$

Soit $a \in \alpha_k, b \in \beta_k$ tels que $a \sim b$, on a

$$d(\lambda_{ab}) \geq \delta[\alpha_k, \beta_k] \Leftrightarrow H_1(a) + B(\alpha_k) \geq 1 ,$$

donc

$$d(\lambda_{ba}) = 0$$

et

$$H_1(a) < H_1(b) .$$

Soient $\alpha_i = \mathcal{E}_i(a)$ et $\beta_i = \mathcal{E}_i(b)$ pour $1 \leq i \leq k \Leftrightarrow 1$. D'après (3),

$$h(\alpha_i) \neq h(\beta_i)$$

donc

$$h(\alpha_i) > h(\beta_i) .$$

D'autre part on a

$$h(\alpha_1) = H_1(a) < H_1(b) = h(\beta_1) .$$

Ce qui est contradictoire. D'où

$$h(\alpha_k) < h(\beta_k) .$$

La réciproque est immédiate et toujours satisfaite.

□

Compte tenu de (3) et (4) de la proposition 3.2.3, la construction itérative des $\mathcal{E}_k, \mathcal{F}_k$ et H_k se simplifie de la manière suivante.

Définissons la fonction d'énergie H_1 par

$$H_1 = \frac{H}{\theta} .$$

Evidemment $H_1 \in \mathbb{N}$. Posons

$$J = \max_{a,b} |H_1(a) \Leftrightarrow H_1(b)| .$$

Alors H_1 prend ses valeurs dans $\{0, 1, \dots, J\}$. Supposons définie la fonction d'énergie H_k , la partition \mathcal{E}_k de E se définit comme l'ensemble des classes d'équivalence par rapport \mathcal{R}_k donnée dans la propriété (3) de la proposition 3.2.3. Il est clair que pour toute classe α_k de \mathcal{E}_k , H_k est constant sur α_k , notons $h(\alpha_k)$ cette constante d'énergie. On dit que $\alpha_k \in \mathcal{E}_k$ est une classe minimale d'ordre k si $h(\alpha_k)$ est un minimum local de H_k . Notons \mathcal{F}_k l'ensemble des classes minimales d'ordre k . Définissons H_{k+1} de la façon suivante. Pour tout a de E

$$H_{k+1}(a) = \begin{cases} H_k(a) + 1 & \text{si } \exists \alpha_k \in \mathcal{F}_k \text{ t.q. } a \in \alpha_k \\ H_k(a) & \text{sinon .} \end{cases}$$

Alors H_{k+1} prend les valeurs $\{k, \dots, J\}$.

En fait, les appellations “fonction d'énergie”, l'ensemble des “classes minimales” qu'on a utilisées pour H_k, \mathcal{F}_k dans le cas général sont inspirées par le cas du recuit simulé.

Puisqu'il n'y a pas de taux de transition entre deux classes minimales α_k, β_k différentes, compte tenu de (1) de la proposition 3.2.3, alors le taux de transition de α_k à β_k dans le générateur Λ_k défini par le théorème A (page 96) se simplifie comme suit:

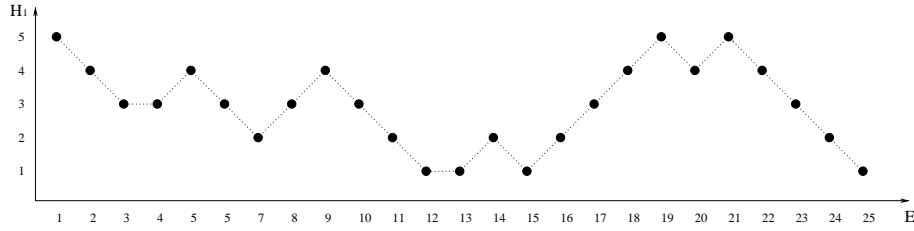
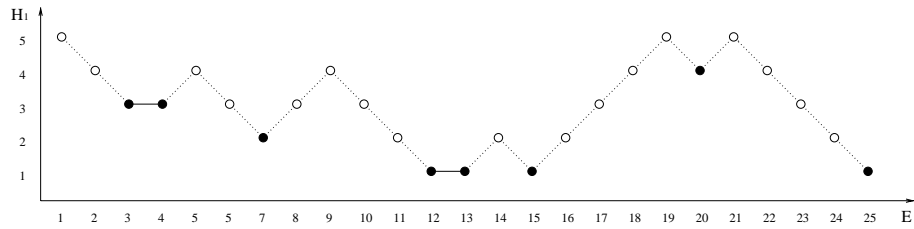
$$\lambda_{\alpha_k \beta_k} = \frac{1}{\varepsilon^k} \Pi_{\alpha_k} G_{\alpha_k}^{U_{k+1}} (\Lambda_{U_{k+1}} \Leftrightarrow D_{U_{k+1}})^{-1} G_{U_{k+1}}^{\beta_k} \mathbb{1}_{\beta_k} .$$

Nous illustrons ci-dessous la structure hiérarchique par un exemple explicite.

Exemple 3.2.1 .

Soit $E = \{1, \dots, 25\}$. Le graphe du générateur $\Lambda = (\lambda_{ij})_{i,j \in E}$ est linéaire. Nous donnons dans la figure 3.9 la fonction d'énergie H_1 qui détermine le générateur. En fait

$$\lambda_{i,j} = \begin{cases} 0 & \text{si } |i \Leftrightarrow j| > 1 \\ 1 & \text{si } |i \Leftrightarrow j| = 1 \text{ et } H(j) \leq H(i) \\ \varepsilon & \text{si } |i \Leftrightarrow j| = 1 \text{ et } H(j) > H(i) . \end{cases}$$

Figure 3.9: Fonction d'énergie H_1 .Figure 3.10: H_1, \mathcal{E}_1 et \mathcal{F}_1 .

Pour $1 \leq k \leq 5$, dans la figure nommée “ H_k, \mathcal{F}_k et \mathcal{E}_k ”, l'énergie H_k des états est donnée sur l'axe vertical, une classe de \mathcal{E}_k est un ensemble de points connectés par des traits pleins, si les points d'une classe sont noirs (resp. blancs), il s'agit alors d'une classe minimale (resp. non-minimale).

À une permutation près, on a

$$\lim_{\varepsilon \searrow 0} \Lambda_1 = \begin{pmatrix} \Leftrightarrow 1/4 & 1/4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \Leftrightarrow 1/4 & 1/4 & 0 & 0 \\ 0 & 0 & 1/2 & \Leftrightarrow 1/2 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & \Leftrightarrow 1 & 1/2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

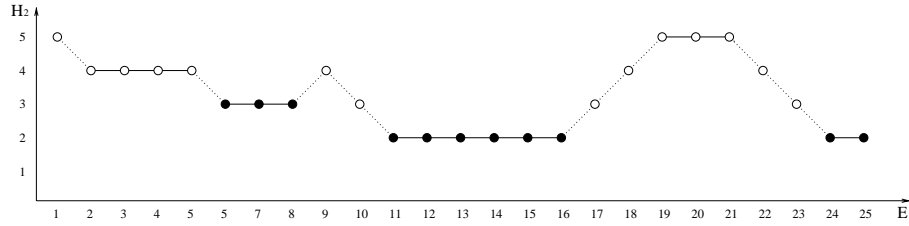


Figure 3.11: H_2, \mathcal{E}_2 et \mathcal{F}_2 .

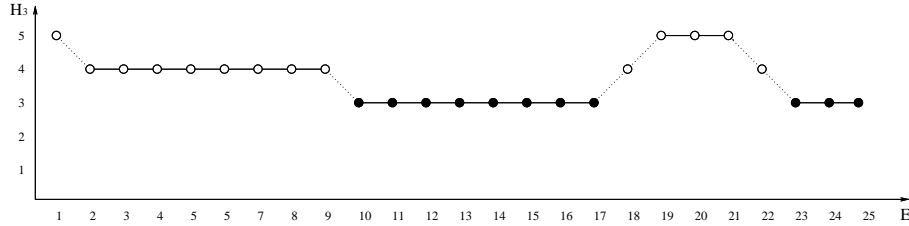


Figure 3.12: H_3, \mathcal{E}_3 et \mathcal{F}_3 .

ses valeurs propres sont: 0 (de multiplicité 3), $\Leftrightarrow \frac{1}{4}$, $\Leftrightarrow \frac{3}{4}$, et $\Leftrightarrow 1$;

$$\lim_{\varepsilon \searrow 0} \Lambda_2 = \begin{pmatrix} \Leftrightarrow 1/2 & 1/2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

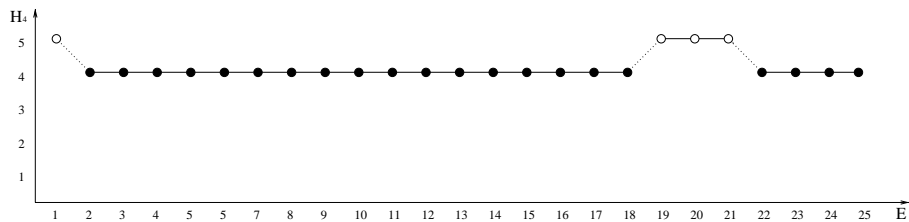
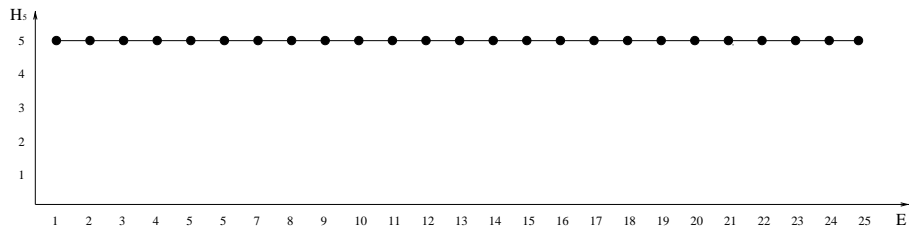
ses valeurs propres sont: 0 (de multiplicité 2) et $\Leftrightarrow \frac{1}{2}$; Et $\lim_{\varepsilon \searrow 0} \Lambda_3$ est nulle;

$$\lim_{\varepsilon \searrow 0} \Lambda_4 = \begin{pmatrix} \Leftrightarrow 1/12 & 1/12 \\ 1/4 & \Leftrightarrow 1/4 \end{pmatrix},$$

ses valeurs propres sont: 0 et $\Leftrightarrow \frac{1}{3}$.

On en déduit que les équivalents des valeurs propres d'ordre $O(\varepsilon)$ du générateur initial Λ sont:

$$0, \Leftrightarrow \frac{1}{4}\varepsilon, \Leftrightarrow \frac{3}{4}\varepsilon, \Leftrightarrow \varepsilon, \Leftrightarrow \frac{1}{2}\varepsilon^3, \Leftrightarrow \frac{1}{3}\varepsilon^4.$$

Figure 3.13: H_4, \mathcal{E}_4 et \mathcal{F}_4 .Figure 3.14: H_5, \mathcal{E}_5 et \mathcal{F}_5 .

3.2.4 Processus d'Ising en dimension un

Adoptons les définitions et notations présentées au paragraphe 2.2.2. On s'intéresse au cas où $d = 1$, l'ensemble de sites $S = \{0, 1, \dots, n \Leftrightarrow 1\}$, l'espace de configurations $\Omega = \{0, 1\}^S$. Pour les trois conditions de bord : périodique, ouvert et unitaire, les générateurs sont notés respectivement par Λ^p, Λ^o et Λ^u . Parmi les processus d'Ising, nous nous intéressons plus particulièrement à deux modèles: l'algorithme de Metropolis et l'échantillonnage de Gibbs. Pour l'algorithme de Metropolis, les taux de transition sont définis par

$$c(x, \eta) = \begin{cases} 1 & \text{si } H(\eta_x) \leq H(\eta) \\ \exp \Leftrightarrow \frac{1}{T} [H(\eta_x) \Leftrightarrow H(\eta)] & \text{sinon} \end{cases} .$$

Pour l'échantillonnage de Gibbs,

$$c(x, \eta) = \left\{ 1 + \exp \Leftrightarrow \frac{1}{T} [H(\eta) \Leftrightarrow H(\eta_x)] \right\}^{-1} .$$

Nous généralisons ces deux définitions de la façon suivante. Posant

$$a(x, \eta) = \exp \Leftrightarrow \frac{1}{T} [H(\eta) \Leftrightarrow H(\eta_x)] ,$$

définissons les taux de transition d'un processus d'Ising par

$$c(x, \eta) = \begin{cases} f(a(x, \eta)) & \text{si } H(\eta_x) < H(\eta) \\ a(x, \eta)^{-1} f(a(x, \eta)^{-1}) & \text{sinon} \end{cases} .$$

où la fonction f est analytique et positive. On retrouve l'algorithme de Metropolis et l'échantillonnage de Gibbs en prenant $f \equiv 1$ et $f(x) = 1/(x+1)$ respectivement.

Pour la famille de processus d'Ising définie ci-dessus, nous avons les résultats suivants sur les équivalents de gaps dans les différentes conditions de bord.

Proposition 3.2.4 *Quand la température T tend vers 0, on a les équivalents de gaps dans les différentes conditions de bord. Pour la condition de bord périodique,*

$$\text{gap}(\Lambda^p) = \frac{4f(0)f(1)n}{4f(1) + f(0)(n \Leftrightarrow 2)} \exp\left(\frac{4}{T}\right) + o\left(\exp\left(\frac{4}{T}\right)\right) .$$

Pour la condition de bord ouverte,

$$\text{gap}(\Lambda^o) = \frac{4f(0)f(1)}{2f(1) + (n \Leftrightarrow 2)f(0)} \exp\left(\frac{2}{T}\right) + o\left(\exp\left(\frac{2}{T}\right)\right) ,$$

Pour la condition de bord unitaire, il existe une constante c strictement positive telle que

$$\text{gap}(\Lambda^u) = c + O\left(\exp\left(\frac{4}{T}\right)\right) .$$

Remarque Comme cas particuliers, les gaps asymptotiques de l'algorithme de Metropolis dans les cas périodique et ouvert sont les suivants.

$$\text{gap}(\Lambda^p) = \left(4 \Leftrightarrow \frac{8}{n+2}\right) \exp\left(\frac{4}{T}\right) + o\left(\exp\left(\frac{4}{T}\right)\right)$$

$$\text{gap}(\Lambda^o) = \frac{4}{n} \exp\left(\frac{2}{T}\right) + o\left(\exp\left(\frac{2}{T}\right)\right) .$$

Pour l'échantillonnage de Gibbs, on a

$$\text{gap}(\Lambda^p) = 2 \exp\left(\frac{4}{T}\right) + o\left(\exp\left(\frac{4}{T}\right)\right)$$

$$\text{gap}(\Lambda^o) = \frac{2}{n \Leftrightarrow 1} \exp\left(\frac{2}{T}\right) + o\left(\exp\left(\frac{2}{T}\right)\right) .$$

Supposons que pour $0 < x \leq 1$, on a $f(x) \geq f(1)$. Dans les cas périodique et ouvert, les horloges sont respectivement

$$\lambda_h^p = \begin{cases} nf\left(\exp\left(\frac{4}{T}\right)\right) & \text{si } n \text{ pair} \\ (n \Leftrightarrow 2)f\left(\exp\left(\frac{4}{T}\right)\right) + 2f(1) & \text{si } n \text{ impair} \end{cases}$$

et

$$\lambda_h^o = (n \Leftrightarrow 2)f\left(\exp\left(\frac{4}{T}\right)\right) + 2f\left(\exp\left(\frac{2}{T}\right)\right) .$$

Pour une température T assez basse et un nombre de sites n assez grand, l'algorithme de Metropolis est meilleur que l'échantillonnage de Gibbs. Cette conclusion figure déjà dans [29] comme conséquence d'un raisonnement différent.

Avant de démontrer le théorème ci-dessus, présentons quelques notations. Pour toute configuration η de Ω , appelons *composante connexe* de η un ensemble de sites voisins (au sens du graphe G qui dépend de la condition de bord) sur lequel η soit constante (un site à 0 entre des sites à 1 est une composante connexe de cardinal 1). Notons Ω_1 l'ensemble des configurations ayant 2 composantes connexes et Ω_2 celui des configurations ayant plus de 2 composantes connexes. Soit $\Lambda_{\Omega_1}, \Lambda_{\Omega_2}$ les générateurs restreints de Λ à respectivement Ω_1 et Ω_2 . Notons η^0, η^1 les configurations dont tous les sites sont à 0 et 1 respectivement qui sont les seules configurations ayant une composante connexe. On pose

$$\alpha_1 = \{\eta^0\}, \beta_1 = \{\eta^1\} .$$

Démonstration. Considérons d'abord la condition de bord périodique. Posons

$$\varepsilon = \exp\left(\frac{4}{T}\right) .$$

Tenons compte du fait que pour toute configuration η et tout site x , $H(\eta) \Leftrightarrow H(\eta_x)$ ne peut valoir que: $\Leftrightarrow 4, 0$ ou 4 . Alors les taux de transition peuvent s'écrire de la façon suivante:

$$c(x, \eta) = \begin{cases} f(\varepsilon) & \text{si } H(\eta) \Leftrightarrow H(\eta_x) = 4 \\ f(1) & \text{si } H(\eta) \Leftrightarrow H(\eta_x) = 0 \\ \varepsilon f(\varepsilon) & \text{si } H(\eta) \Leftrightarrow H(\eta_x) = \Leftrightarrow 4 . \end{cases}$$

Constatons que pour toutes configurations η, ξ de Ω , on a

$$H(\eta) = H(\xi) \iff H_1(\eta) = H_1(\xi)$$

et

$$H(\eta) > H(\xi) \iff H_1(\eta) > H_1(\xi) ,$$

tenant compte du fait que

$$H_1(\eta) = d(\Pi(\eta)) = \frac{H(\eta) \Leftrightarrow H(\eta^0)}{4} .$$

Soit ξ^0 une configuration autre que η^0 et η^1 . On a

$$H(\xi^0) > H(\eta^0) = H(\eta^1)$$

Comme qu'il n'y a pas de taux de transition entre η^0 et η^1 , $\{\eta^0\}$ et $\{\eta^1\}$ sont deux classes minimales d'ordre 1. Si ξ^0 contient un site x à 0 mais entre deux 1 (par exemple $\xi^0 = (1101)$ quand $n = 4$), alors $H(\xi_x^0) < H(\xi^0)$, donc ξ^0 n'est pas dans une classe minimale d'ordre 1. Sinon, ξ^0 contient forcément des sites à 0 et se situant entre 0 et 1 (par exemple $\xi^0 = (1000)$). Notons ξ^1 la configuration obtenue en changeant l'état d'un tel site de 0 à 1. Récursivement, si ξ^i contient des sites à 0 et se situant entre 0 et 1, posons ξ^{i+1} la configuration obtenue en changeant l'état d'un tel site de 0 à 1. Il existe alors un entier l tel que ξ^l contient un site x à 0 mais se situant entre deux 1 (par exemple ξ^l peut être (1110)). Comme $H(\xi^i) = H(\xi^{i+1})$ et les taux de transitions entre ξ^i et ξ^{i+1} sont $f(1)$, donc les ξ^i sont dans une même classe d'ordre 1, pour $i = 0, \dots, l \Leftrightarrow 1$. Or ξ^l n'est pas dans une classe minimale d'ordre 1, donc ξ^0 ne l'est non plus. On a

$$\mathcal{F}_1 = \{\alpha_1, \beta_1\} .$$

Déterminons le générateur Λ_1 défini sur \mathcal{F}_1 que donne le théorème A. Si η possède i composantes connexes (i est nécessairement pair où 1), on a

$$H(\eta) = \begin{cases} \Leftrightarrow n & \text{si } i = 1 \\ 2i \Leftrightarrow n & \text{sinon ,} \end{cases}$$

et

$$H_1(\eta) = \begin{cases} 0 & \text{si } i = 1 \\ i/2 & \text{sinon .} \end{cases}$$

Soit $U_2 = \Omega \setminus (\alpha_1 \cup \beta_1)$. Alors U_2 est l'ensemble des configurations ayant plus de une composante connexe, donc

$$U_2 = \Omega_1 \cup \Omega_2 .$$

Constatons que pour tous $\eta \in \Omega_1, \xi \in \Omega_2$, comme $H(\xi) > H(\eta)$, le taux de transition de η vers ξ tend vers 0,

$$\lambda_{\eta\xi} = O(\varepsilon) .$$

Alors il existe deux matrices M_1, M_2 analytiques en ε telles que Λ_{U_2} peut s'écrire sous la forme,

$$\Lambda_{U_2} = \begin{pmatrix} \Lambda_{\Omega_1} & \varepsilon M_1 \\ M_2 & \Lambda_{\Omega_2} \end{pmatrix} \Leftrightarrow M$$

où la matrice M est diagonale telle que

$$\begin{pmatrix} (0) & \varepsilon M_1 \\ M_2 & (0) \end{pmatrix} \mathbb{1}_{U_2} = M \mathbb{1}_{U_2}$$

En posant

$$D_{U_2} = \begin{pmatrix} D'_{\Omega_1} & (0) \\ (0) & D'_{\Omega_2} \end{pmatrix} ,$$

où les matrices on a

$$\begin{aligned} \lim_{\varepsilon \searrow 0} (\Lambda_{U_2} \Leftrightarrow D_{U_2})^{-1} &= \lim_{\varepsilon \searrow 0} \begin{pmatrix} \Lambda_{\Omega_1} \Leftrightarrow D'_{\Omega_1} & (0) \\ M_2 & \Lambda_{\Omega_2} \Leftrightarrow D'_{\Omega_2} \end{pmatrix}^{-1} \\ &= \lim_{\varepsilon \searrow 0} \begin{pmatrix} \Lambda_{\Omega_1} \Leftrightarrow D'_{\Omega_1} & x \\ x & x \end{pmatrix}^{-1} \end{aligned} \quad (3.25)$$

où x désigne des termes sans intérêt.

Montrons que le cardinal de Ω_1 est $n(n \Leftrightarrow 1)$ et que à une permutation près la limite de la restriction de Λ à Ω_1 s'écrit

$$\lim_{\varepsilon \searrow 0} \Lambda_{\Omega_1} = \begin{pmatrix} \Leftrightarrow 2f(1)I & F & & & (0) \\ F^* & \Leftrightarrow 4f(1)I & F & & \\ & \ddots & \ddots & \ddots & \\ & & F^* & \Leftrightarrow 4f(1)I & F \\ (0) & & & F^* & \Leftrightarrow 2f(1)I \end{pmatrix}_{(n-1) \times (n-1)} \quad (3.26)$$

où

$$F = f(1) \begin{pmatrix} 1 & & & 1 \\ 1 & \ddots & & (0) \\ & \ddots & \ddots & \\ (0) & & 1 & 1 \end{pmatrix}.$$

Soit $\{V_1, \dots, V_{n-1}\}$ une partition de Ω_1 telle que le cardinal de la composante connexe de valeur 1 des configurations de V_i soit i . A toute configuration $\eta \in \Omega_1$, on l'associe le couple d'indices (i, j) si et seulement si

$$\begin{cases} \eta \in V_i \\ \eta(j) = 1 \text{ et } \eta(j \Leftrightarrow 1) = 0 \end{cases}.$$

Alors

$$\Omega_1 = \{\eta_{ij} ; i \in \{1, \dots, n \Leftrightarrow 1\}, j \in \{1, \dots, n\}\}.$$

Donc le cardinal de Ω_1 est $n(n \Leftrightarrow 1)$. A partir de η_{ij} , les seuls taux de transitions non nuls vont vers η_{i+1j-1} , η_{i+1j} , η_{i-1j+1} et η_{i-1j} . Alors à une permutation près, F et F^* sont respectivement les matrices des taux de transition de V_i à V_{i+1} et de V_{i+1} à V_i . D'où Λ_{Ω_1} s'écrit sous la forme (3.26).

On constate que parmi les taux de transition entre α_1 et U_2 , il n'y a de taux de transition non nuls qu'entre α_1 et V_1 . De manière analogue, parmi les taux de transition entre β_1 et U_2 , il n'y a de taux de transition non nuls qu'entre β_1 et V_{n-1} . On a en fait

$$G_{\alpha_1}^{\Omega_1} = \begin{pmatrix} G_{\alpha_1}^{V_1} & (0) \end{pmatrix}$$

avec

$$G_{\alpha_1}^{V_1} = \varepsilon f(\varepsilon) \mathbb{1}_{V_1}^*$$

et

$$G_{\Omega_1}^{\alpha_1} = \begin{pmatrix} G_{V_1}^{\alpha_1} \\ (0) \end{pmatrix}, \quad G_{\Omega_1}^{\beta_1} = \begin{pmatrix} (0) \\ G_{V_{n-1}}^{\beta_1} \end{pmatrix}$$

avec

$$G_{V_1}^{\alpha_1} = f(\varepsilon) \mathbb{1}_{V_1}, \quad G_{V_{n-1}}^{\beta_1} = f(\varepsilon) \mathbb{1}_{V_{n-1}}.$$

D'autre part

$$G_{\alpha_1}^{U_2} = \begin{pmatrix} G_{\alpha_1}^{\Omega_1} & (0) \end{pmatrix}, \quad G_{U_2}^{\alpha_1} = \begin{pmatrix} G_{\Omega_1}^{\alpha_1} \\ (0) \end{pmatrix} \quad \text{et} \quad G_{U_2}^{\beta_1} = \begin{pmatrix} G_{\Omega_1}^{\beta_1} \\ (0) \end{pmatrix}.$$

Donc

$$D'_{\Omega_1} = f(\varepsilon) \begin{pmatrix} I_{V_1} & (0) \\ & (0) \\ (0) & I_{V_{n-1}} \end{pmatrix}$$

En considérant l'équation (3.25), on a

$$\begin{aligned} \lim_{\varepsilon \searrow 0} \lambda_{\alpha_1 \beta_1} &= \Leftrightarrow \lim_{\varepsilon \searrow 0} \frac{1}{\varepsilon} G_{\alpha_1}^{U_2} (\Lambda_{U_2} \Leftrightarrow D_{U_2})^{-1} G_{U_2}^{\beta_1} \\ &= \Leftrightarrow \lim_{\varepsilon \searrow 0} \frac{1}{\varepsilon} G_{\alpha_1}^{\Omega_1} (N_{\Omega_1} \Leftrightarrow D'_{\Omega_1})^{-1} G_{\Omega_1}^{\beta_1} \\ &= \Leftrightarrow f^2(0) \begin{pmatrix} \mathbb{1}_{V_1}^* & (0) \end{pmatrix} Q \begin{pmatrix} (0) \\ \mathbb{1}_{V_{n-1}} \end{pmatrix} \end{aligned}$$

où

$$\begin{aligned} Q &= \lim_{\varepsilon \searrow 0} (N_{\Omega_1} \Leftrightarrow D'_{\Omega_1})^{-1} \\ &= \begin{pmatrix} E_1 & F & & (0) \\ F^* & E_0 & F & \\ & \ddots & \ddots & \ddots \\ & & F^* & E_0 & F \\ (0) & & & F^* & E_1 \end{pmatrix}_{(n-1) \times (n-1)}^{-1} \end{aligned}$$

avec

$$E_1 = \Leftrightarrow (2f(1) + f(0))I \quad \text{et} \quad E_0 = \Leftrightarrow Af(1)I.$$

Posons

$$Q = (Q_{ij})_{1 \leq i, j \leq n-1}$$

où les matrices carrées Q_{ij} sont d'ordre n . Alors

$$\lim_{\varepsilon \searrow 0} \lambda_{\alpha_1 \beta_1} = \Leftrightarrow f^2(0) \mathbb{1}^* Q_{1n-1} \mathbb{1} .$$

Si on pose

$$\begin{cases} E_k = E_0 \Leftrightarrow F E_{k-1}^{-1} F^* & k = 2, \dots, n \Leftrightarrow 2 \\ E_{n-1} = E_1 \Leftrightarrow F E_{n-2}^{-1} F^* \end{cases} ,$$

on vérifie que

$$Q_{1n-1} = (\Leftrightarrow \mathbb{1})^n E_{n-1}^{-1} \prod_{k=1}^{n-2} F E_k^{-1} .$$

D'où

$$\lim_{\varepsilon \searrow 0} \lambda_{\alpha_1 \beta_1} = \Leftrightarrow f^2(0) \mathbb{1}^* E_{n-1}^{-1} \prod_{k=1}^{n-2} F E_{n-k}^{-1} \mathbb{1} .$$

On a

$$F \mathbb{1} = 2f(1) \mathbb{1} , \quad F^* \mathbb{1} = 2f(1) \mathbb{1}$$

et

$$E_1 \mathbb{1} = \Leftrightarrow (2f(1) + f(0)) \mathbb{1} , \quad E_0 \mathbb{1} = \Leftrightarrow 4f(1) \mathbb{1} .$$

Soit $a_1 = \Leftrightarrow (2f(1) + f(0))$. Supposons que $\mathbb{1}$ est un vecteur propre de E_k associé à la valeur propre a_k . Si $k \leq n \Leftrightarrow 3$,

$$E_{k+1} \mathbb{1} = (E_0 \Leftrightarrow F E_k^{-1} F^*) \mathbb{1} = \Leftrightarrow 4f(1) \left(1 + \frac{f(1)}{a_k} \right) \mathbb{1}$$

et

$$E_{n-1} \mathbb{1} = (E_1 \Leftrightarrow F E_{n-2}^{-1} F^*) \mathbb{1} = \Leftrightarrow \left(2f(1) + f(0) + \frac{4f^2(1)}{a_{n-2}} \right) \mathbb{1} .$$

Donc

$$\begin{cases} a_1 = \Leftrightarrow 2f(1) \Leftrightarrow f(0) \\ a_{k+1} = \Leftrightarrow 4f(1) [1 + f(1)a_k^{-1}] & \text{si } 1 \leq k \leq n \Leftrightarrow 3 \\ a_{n-1} = \Leftrightarrow 2f(1) \Leftrightarrow f(0) \Leftrightarrow 4f^2(1)a_{n-2}^{-1} \end{cases} .$$

Par récurrence, on vérifie que

$$a_k^{-1} = \Leftrightarrow \frac{1}{2f(1)} \times \frac{(k \Leftrightarrow 1)f(0) + 2f(1)}{kf(0) + 2f(1)}$$

pour $1 \leq k \leq n \Leftrightarrow 2$ et

$$a_{n-1}^{-1} = \Leftrightarrow \frac{1}{f(0)} \times \frac{(n \Leftrightarrow 2)f(0) + 2f(1)}{4f(1) + f(0)(n \Leftrightarrow 2)} .$$

On a donc

$$\begin{aligned} \lim_{\varepsilon \searrow 0} \lambda_{\alpha_1 \beta_1} &= \Leftrightarrow f^2(0) n a_{n-1}^{-1} (\Leftrightarrow 2f(1))^{n-2} \prod_{k=1}^{n-2} a_k^{-1} \\ &= \frac{2f(0)f(1)n}{4f(1) + f(0)(n \Leftrightarrow 2)} . \end{aligned}$$

Or Λ_1 est symétrique, on en déduit que

$$\text{gap}(\Lambda^p) = \frac{4f(0)f(1)n}{4f(1) + f(0)(n \Leftrightarrow 2)} \varepsilon + o(\varepsilon) .$$

Considérons la condition de bord ouvert.

Posons maintenant

$$\varepsilon = \exp\left(\Leftrightarrow \frac{2}{T}\right) .$$

Les taux de transition s'écrivent

$$c(x, \eta) = \begin{cases} f(\varepsilon^2) & \text{si } H(\eta_x) \Leftrightarrow H(\eta) = \Leftrightarrow 4 \\ f(\varepsilon) & \text{si } H(\eta_x) \Leftrightarrow H(\eta) = \Leftrightarrow 2 \\ f(1) & \text{si } H(\eta_x) \Leftrightarrow H(\eta) = 0 \\ \varepsilon f(\varepsilon) & \text{si } H(\eta_x) \Leftrightarrow H(\eta) = 2 \\ \varepsilon^2 f(\varepsilon^2) & \text{si } H(\eta_x) \Leftrightarrow H(\eta) = 4 \end{cases}$$

Par un raisonnement analogue au cas périodique, on montre que \mathcal{F}_1 est réduit à deux classes.

$$\mathcal{F}_1 = \{\alpha_1, \beta_1\} .$$

Déterminons le générateur Λ_1 défini sur \mathcal{F}_1 . Evidemment, Λ_1 est symétrique, il suffit alors de déterminer le taux de transition $\lambda_{\alpha_1 \beta_1}$ de α_1 à β_1 . Soit $U_2 = \Omega \setminus (\alpha_1 \cup \beta_1)$, alors

$$U_2 = \Omega_1 \cup \Omega_2 .$$

Par un raisonnement analogue au cas périodique, on a

$$\lim_{\varepsilon \searrow 0} \lambda_{\alpha_1 \beta_1} = \Leftrightarrow \lim_{\varepsilon \searrow 0} \frac{1}{\varepsilon} G_{\alpha_1}^{\Omega_1} \cancel{N_{\Omega_1}} (\Leftrightarrow D'_{\Omega_1})^{-1} G_{\Omega_1}^{\beta_1} .$$

Pour $i = 0, 1$, $j = 0, \dots, n \Leftrightarrow 2$, en notant par η^{ij} la configuration dont les sites de 0 à j sont dans l'état i et de $j + 1$ à $n \Leftrightarrow 1$ dans l'état $1 \Leftrightarrow i$, on a

$$\Omega_1 = \{\eta^{ij} ; i = 0, 1 \quad j = 0, \dots, n \Leftrightarrow 2\} \quad .$$

Classons les éléments de Ω_1 selon leurs indices par ordre croissant au sens où

$$ij \leq i'j' \Leftrightarrow i < i' \text{ ou } i = i', j \leq j' \quad .$$

Alors on a

$$G_{\alpha_1}^{\Omega_1} = \varepsilon f(\varepsilon) (0 \cdots 0 \ 11 \ 0 \cdots 0)_{1 \times (2n-2)}$$

où les deux 1 sont le $(n \Leftrightarrow 1)$ -ième et le n -ième élément,

$$\Lambda_{\Omega_1} = \begin{pmatrix} A & (0) \\ (0) & A \end{pmatrix}$$

avec

$$A = f(1) \begin{pmatrix} \Leftrightarrow 1 & 1 & & & (0) \\ 1 & \Leftrightarrow 2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \Leftrightarrow 2 & 1 \\ (0) & & & 1 & \Leftrightarrow 1 \end{pmatrix}_{(n-1) \times (n-1)},$$

et

$$D'_{\Omega_1} = f(\varepsilon) \text{diag} (10 \cdots 0110 \cdots 01)_{(2n-2) \times (2n-2)}$$

où les deux 1 au milieu sont respectivement sur la $(n \Leftrightarrow 1)$ -ième et n -ième ligne.

De plus

$$G_{\Omega_1}^{\beta_1} = f(\varepsilon) (10 \cdots 01)_{1 \times (2n-2)}^* \quad .$$

Posons

$$B = \left(A \Leftrightarrow f(0) \text{diag} (10 \cdots 01)_{(n-1) \times (n-1)} \right)^{-1} = (b_{ij})_{1 \leq i, j \leq n-1},$$

en fait

$$\lim_{\varepsilon \searrow 0} \lambda_{\alpha_1 \beta_1} = \Leftrightarrow 2 f^2(0) b_{n-11} \quad .$$

Des calculs standards nous donnent

$$b_{n-1,1} = (\Leftrightarrow 1)^n c_{n-1}^{-1} f^{n-2}(1) \left(\prod_{k=1}^{n-2} c_k \right)^{-1}$$

avec les c_k vérifiant

$$\begin{cases} c_1 = \Leftrightarrow f(1) \Leftrightarrow f(0) \\ c_{k+1} = \Leftrightarrow 2f(1) \Leftrightarrow f^2(1) c_k^{-1} & 1 \leq k \leq n \Leftrightarrow 3 \\ c_{n-1} = \Leftrightarrow f(1) \Leftrightarrow f(0) \Leftrightarrow f^2(1) c_{n-2}^{-1} \end{cases} .$$

Par récurrence, on a pour $1 \leq k \leq n \Leftrightarrow 2$

$$c_k^{-1} = \Leftrightarrow \frac{1}{f(1)} \times \frac{(k \Leftrightarrow 1)f(0) + f(1)}{kf(0) + f(1)}$$

et

$$c_{n-1}^{-1} = \Leftrightarrow \frac{1}{f(0)} \times \frac{(n \Leftrightarrow 2)f(0) + f(1)}{2f(1) + f(0)(n \Leftrightarrow 2)} .$$

Alors

$$b_{n-1,1} = \Leftrightarrow \frac{f(1)}{f(0) [2f(1) + f(0)(n \Leftrightarrow 2)]} .$$

Doù

$$\lim_{\varepsilon \searrow 0} \lambda_{\alpha_1, \beta_1} = \frac{2f(0)f(1)}{2f(1) + f(0)(n \Leftrightarrow 2)} .$$

On a alors

$$\text{gap}(\Lambda^\circ) = \frac{4f(0)f(1)}{2f(1) + f(0)(n \Leftrightarrow 2)} \varepsilon + O(\varepsilon)$$

ce qui termine la démonstration dans le cas ouvert.

Pour la condition de bord unitaire, les taux de transition s'écrivent de la façon suivante,

$$c(x, \eta) = \begin{cases} f(\varepsilon) & \text{si } H(\eta) \Leftrightarrow H(\eta_x) = 4 \\ f(1) & \text{si } H(\eta) \Leftrightarrow H(\eta_x) = 0 \\ \varepsilon f(\varepsilon) & \text{si } H(\eta) \Leftrightarrow H(\eta_x) = \Leftrightarrow 4 \end{cases}$$

avec

$$\varepsilon = \exp \left(\Leftrightarrow \frac{4}{T} \right) .$$

Dans ce cas, on vérifie que \mathcal{F}_1 est réduit à une seule classe,

$$\mathcal{F}_1 = \{\beta_1\} .$$

C'est-à-dire qu'il n'y a qu'une valeur propre d'ordre $O(\varepsilon)$ qui n'est autre que la valeur propre 0. Donc le $\text{gap}(\Lambda^u)$ est strictement positif quand ε tend vers 0.

□

Chapitre 4

Moyenne harmonique du spectre

Ce chapitre présente une méthode générale d'estimation du gap d'un processus de Markov. Cette méthode utilise la somme des inverses des valeurs propres non nulles d'un générateur obtenu en restreignant le générateur initial à un arbre maximal. Un algorithme de calcul explicite de cette quantité est fourni. Dans la première section (4.1), on démontre le résultat principal. A titre d'application (4.2.1), on donnera une estimation du gap géométrique d'un graphe qui est en fait le gap du générateur de la marche aléatoire symétrique sur le graphe. Ensuite, on reprend la famille des processus d'Ising (4.2.2) qu'on a considéré dans le paragraphe 3.2.4. En utilisant la méthode décrite ci-dessus, on donne une borne inférieure du gap qui est du même ordre de grandeur que l'équivalent du gap donné par la proposition 3.2.4.

4.1 Démonstration du théorème B

On suppose toujours que les processus de Markov sont irréductibles et réversibles. Soit A le générateur d'intérêt défini sur l'espace d'états E de taille $n + 1$.

On suppose que Λ a pour valeurs propres

$$\Leftrightarrow \lambda_0 = 0 > \Leftrightarrow \lambda_1 \geq \Leftrightarrow \lambda_2 \geq \cdots \geq \Leftrightarrow \lambda_n .$$

Alors

$$\text{gap}(\Lambda) = \lambda_1 .$$

Soit S l'opposé de la somme des inverses des valeurs propres non nulles (SIVP) de $\Leftrightarrow \Lambda$,

$$S = \sum_{i=1}^n \frac{1}{\lambda_i} .$$

On a une estimation évidente du gap ,

$$\frac{1}{S} \leq \text{gap}(\Lambda) \leq \frac{n}{S} \quad (4.1)$$

Mais pour un générateur quelconque, la quantité S n'a pas d'expression simple.

La proposition 4.1.1 suivante exprime S en fonction du générateur Λ .

Proposition 4.1.1

$$S = \Leftrightarrow \frac{1}{|\Lambda^1|} \sum_{i=1}^n |{}_i \Lambda^1|$$

où Λ^1 désigne la matrice obtenue en remplaçant la dernière colonne de Λ par le vecteur $\mathbf{1}$, et ${}_i \Lambda^1$ celle que l'on déduit de la précédente en remplaçant la i -ième colonne par le vecteur $(0 \cdots 0 1 0 \cdots 0)^*$ où 1 est le i -ième élément.

Démonstration. Considérons le polynôme caractéristique de Λ ,

$$P(\lambda) = |\Lambda + I\lambda|$$

où I est la matrice identité. En ajoutant toutes les autres colonnes à la dernière, on trouve que la dernière devient $\lambda \mathbf{1}^*$, d'où

$$Q(\lambda) = \frac{P(\lambda)}{\lambda} = |(\Lambda + I\lambda)^1| = \prod_{i=1}^n (\lambda \Leftrightarrow \lambda_i) .$$

Donc

$$\frac{Q'(0)}{Q(0)} = \Leftrightarrow \sum_{i=1}^n \frac{1}{\lambda_i} .$$

Remarquons que

$$Q(0) = |\Lambda^1|$$

et la dérivée de $|(\Lambda + \lambda I)^1|$ par rapport à λ en 0 est la somme des $|_i \Lambda^1|$ qui est la dérivée de $|(\Lambda + \lambda I)^1|$ sur la i -ème colonne en $\lambda = 0$, d'où le résultat. □

Dans le cas particulier où le graphe réduit $G = (E, A)$ de Λ est un arbre, on peut obtenir une expression explicite reliant S aux taux de transition du processus.

Supposons que G est orienté par la racine a_0 . On note B_a le sous-arbre issu de a qui est l'ensemble des sommets de l'arbre supérieurs ou égaux à a .

$$B_a = \{b; a \leq b\}.$$

On note $p(a)$ le père de a et μ_a le taux de transition de a vers $p(a)$. Rappelons que la conductance symétrique de B_a (cf. page 36) est définie comme

$$\phi^*(B_a) = \frac{\pi(a)\mu_a}{\pi(B_a)[1 \leftrightarrow \pi(B_a)]}.$$

Elle s'interprète comme le flux d'échange entre B_a et son complémentaire. Le théorème suivant donne explicitement la SIVP de Λ dans le cas d'un arbre.

Proposition 4.1.2 *Soit $\Lambda = (\lambda_{ij})_{i,j \in E}$ un générateur irréductible et π -réversible tel que le graphe réduit de Λ soit un arbre qui est orienté par rapport à la racine a_0 de E . Alors la SIVP de Λ vaut*

$$S = \sum_{a \in E - \{a_0\}} \frac{1}{\phi^*(B_a)}.$$

Pour un générateur dont le graphe réduit est quelconque, si on annule certains taux d'un processus de manière à préserver l'espace d'états et la mesure réversible, ceci ralentit l'accès à l'équilibre, c'est-à-dire que le gap du nouveau processus est plus faible.

Proposition 4.1.3 *Soit Λ un générateur irréductible π -réversible sur E , et (E, A) son graphe réduit. Soit A_1 un sous-ensemble symétrique de A tel que*

le graphe (E, A_1) soit encore connexe. Considérons les taux de transition λ_{ij}^1 définis par

$$\lambda_{ij}^1 = \begin{cases} \lambda_{ij} & \text{si } (i, j) \in A \\ 0 & \text{sinon.} \end{cases}$$

Soit Λ_1 le générateur correspondant aux taux λ_{ij}^1 . Alors Λ_1 est π -réversible et

$$\text{gap}(\Lambda_1) \leq \text{gap}(\Lambda) \quad .$$

Démonstration. Supprimer des taux de transition de manière symétrique ne modifie pas la condition de bilan détaillé et donc Λ_1 est encore π -réversible. L'inégalité sur les gaps est une conséquence immédiate de l'expression variationnelle du gap (cf. équation 2.1).

□

Cette dernière proposition nous permet d'approcher le gap d'un processus quelconque par celui d'un processus plus simple à traiter, obtenu en annulant des taux de transition jusqu'à ce qu'il n'y ait plus de cycle dans le graphe. Nous nous ramènerons au cas où le graphe réduit du nouveau processus est un arbre maximal du graphe réduit de l'ancien. Nous avons ainsi une minoration du gap d'un générateur quelconque.

Théorème B Soit Λ un générateur irréductible et réversible défini sur E fini. Soit G' un arbre maximal du graphe réduit de Λ orienté par la racine $a_0 \in E$. On a alors

$$\text{gap}(\Lambda) \geq \left[\sum_{\substack{a \in E \\ a \neq a_0}} \frac{1}{\phi^*(B_a)} \right]^{-1}$$

où B_a est le sous-arbre de G' issu du sommet a et $\phi^*(B_a)$ est défini sur le générateur restreint de Λ sur G' .

Pour démontrer la proposition 4.1.2, on donne d'abord les notations et les définitions nécessaires.

Pour la suite, pour toute matrice carrée A , A' désigne la matrice obtenue en supprimant la dernière colonne et la dernière ligne de A , A^1 la matrice obtenue

en remplaçant la dernière colonne de A par le vecteur constant $\mathbb{1}^*$ et ${}_i A$ la matrice obtenue en remplaçant la i -ème colonne de A par le i -ème vecteur de base $(0 \cdots 010 \cdots 0)^*$.

Considérons un générateur Λ sur E dont le graphe réduit est un arbre $G = (E, A)$ orienté par la racine a_0 . Pour tout sommet a de E , nous adoptons les notations et les définitions suivantes.

- Le taux de transition de $p(a)$ à a est noté λ_a , le taux de transition de a à $p(a)$ est noté μ_a ($p(a)$ est le père de a).
- Il n'y a qu'un seul chemin qui joint a et la racine a_0 . On note par $r(a)$ l'ensemble des sommets sur ce chemin, autrement dit, $r(a)$ est l'ensemble des sommets antérieurs ou égaux à a .

$$r(a) = \{c ; c \leq a\} \quad .$$

- L'ensemble des fils de a est noté $\sigma(a)$.

$$\sigma(a) = \{b ; b > a \text{ et } (a, b) \in A\} \quad .$$

- La branche complète issue de a , noté \bar{B}_a , est la réunion de B_a et du père de a ,

$$\bar{B}_a = \{p(a)\} \cup B_a \quad .$$

- La restriction du générateur Λ à B_a est notée Λ_a . La restriction du générateur Λ à \bar{B}_a est notée $\bar{\Lambda}_a$ et $\bar{\Lambda}'_a$ désigne la matrice obtenue en supprimant la dernière colonne et la dernière ligne de $\bar{\Lambda}'_a$.
- La SIVP de Λ_a est notée S_a .
- La mesure de probabilité réversible de Λ_a est notée π_a .
- Le temps moyen de première atteinte de $p(a)$ en partant de a est noté t_a .
- Les temps moyens de parcours de toutes les arêtes de B_a et de \bar{B}_a sont notés T_a et \bar{T}_a respectivement.

$$T_a = \sum_{b \in B_a - \{a\}} t_b \quad ,$$

$$\bar{T}_a = \sum_{b \in B_a} t_b = T_a + t_a \quad .$$

On a

$$T_a = \sum_{b \in \sigma(a)} \bar{T}_b \quad .$$

La mesure réversible de Λ s'exprime facilement à l'aide des taux de transition.

Proposition 4.1.4 *Avec les notations précédentes la mesure réversible du générateur Λ vérifie*

$$\pi(a) = \pi(a_0) \prod_{b \in r(a) \setminus \{a_0\}} \frac{\lambda_b}{\mu_b} \quad .$$

Démonstration. Il suffit de vérifier les conditions de bilan détaillé.

$$\pi(a)\mu_a = \pi(p(a)) \frac{\lambda_a}{\mu_a} \mu_a = \pi(p(a))\lambda_a \quad ,$$

pour tout sommet a différent de a_0 .

□

Le générateur Λ_a étant une troncature du générateur Λ , la proposition 4.1.3 montre que Λ_a est réversible par rapport à la mesure de probabilité π_a suivante. Pour tout sommet a de E et tout sommet b de B_a

$$\pi_a(b) = \frac{\pi(b)}{\pi(B_a)} \quad .$$

Tout arbre est la réunion de plusieurs sous-arbres disjoints et de la racine. Autrement dit, la branche B_a est la réunion de $\{a\}$ et des B_b où b est un fils de a . A une permutation près, tout générateur restreint à une branche est de même forme. Pour la suite on suppose que le sommet a a k fils,

$$\sigma(a) = \{b_1, \dots, b_k\} \quad .$$

On écrit Λ_a , pour tout $a \in E$, en considérant a comme le dernier état et en regroupant les autres états selon les branches B_{b_i} pour $i = 1 \dots k$,

$$\Lambda_a = \begin{pmatrix} \bar{\Lambda}'_{b_1} & (0) & v_1 \\ & \ddots & \vdots \\ (0) & \bar{\Lambda}'_{b_k} & v_k \\ u_1 & \cdots & u_k & \Leftrightarrow \sigma \end{pmatrix} \quad .$$

Et de manière analogue

$$\bar{\Lambda}_a = \begin{pmatrix} \bar{\Lambda}'_{b_1} & (0) & v_1 & 0 \\ & \ddots & \vdots & \vdots \\ (0) & & \bar{\Lambda}'_{b_k} & v_k & 0 \\ u_1 & \cdots & u_k & \Leftrightarrow \sigma \Leftrightarrow \mu_a & \mu_a \\ 0 & \cdots & 0 & \lambda_a & \Leftrightarrow \lambda_a \end{pmatrix},$$

où

$$u_i = (0 \cdots 0 \lambda_{b_i}) \text{ et } v_i = (0 \cdots 0 \mu_{b_i})^*$$

pour $i = 1 \cdots k$,

$$\sigma = \sum_{i=1}^k \lambda_{b_i}$$

et les $\bar{\Lambda}_{b_i}$ sont de même forme que $\bar{\Lambda}_a$.

La proposition suivante donne l'expression du temps de parcours des arêtes.

Proposition 4.1.5 *Pour tout $a \neq a_0 \in E$*

$$t_a = \frac{1}{\mu_a} + \sum_{b \in \sigma(a)} \frac{\lambda_b}{\mu_a} t_b = \frac{\pi(B_a)}{\pi(a)\mu_a} = \frac{1}{\pi_a(a)\mu_a} .$$

Démonstration. Si $\sigma(a) = \emptyset$, la somme est nulle et le résultat est immédiat.

Sinon, on a

$$t_a = \frac{1}{\mu_a + \sum_{b \in \sigma(a)} \lambda_b} + \sum_{b \in \sigma(a)} \frac{\lambda_b}{\mu_a + \sum_{b' \in \sigma(a)} \lambda_{b'}} (t_b + t_a) ,$$

d'où l'expression récursive de t_a . Par récurrence,

$$\begin{aligned} t_a &= \frac{1}{\mu_a} + \sum_{b \in \sigma(a)} \frac{\lambda_b}{\mu_a} t_b \\ &= \frac{1}{\mu_a} + \sum_{b \in \sigma(a)} \frac{\lambda_b}{\mu_a} \frac{\pi(B_b)}{\pi(b)\mu_b} \\ &= \frac{1}{\mu_a} + \frac{1}{\pi(a)\mu_a} \sum_{b \in \sigma(a)} \pi(B_b) \\ &= \frac{\pi(B_a)}{\pi(a)\mu_a} . \end{aligned}$$

□

On en déduit l'expression suivante pour le temps de parcours total des arêtes.

Proposition 4.1.6 *On a*

$$\bar{T}_a = \Leftrightarrow \frac{1}{|\bar{\Lambda}'_a|} \sum_{i=1}^{n_a} |{}_i \bar{\Lambda}'_a|$$

et donc

$$T_a = \Leftrightarrow \frac{1}{|\Lambda'_a|} \sum_{i=1}^{n_a-1} |{}_i \Lambda'_a|$$

où n_a est l'ordre de Λ_a ($n_a = \#B_a$).

Démonstration. On déduit la deuxième égalité de la première en tenant compte de la proposition 4.1.6.

Démontrons la première égalité par récurrence sur la taille de l'espace d'états.

Si $B_a = \{a, b\}$, on vérifie facilement le résultat.

Pour Λ_a quelconque, notons n_i l'ordre de $\bar{\Lambda}'_{b_i}$. Rappelons que

$$\bar{\Lambda}'_a = \begin{pmatrix} \bar{\Lambda}'_{b_1} & (0) & v_1 \\ & \ddots & \vdots \\ (0) & \bar{\Lambda}'_{b_k} & v_k \\ u_1 & \cdots & u_k \Leftrightarrow \sigma \Leftrightarrow \mu_a \end{pmatrix}$$

où $u_i = (0 \cdots 0 \lambda_{b_i})$, $v_i = (0 \cdots 0 \mu_{b_i})^*$ pour $i = 1 \cdots k$, $\sigma = \sum_{i=1}^k \lambda_{b_i}$.

Si $j = n_1 + \cdots + n_{i-1} + l$, $1 \leq l < n_i$, en développant $|{}_j \bar{\Lambda}'_a|$ selon la dernière ligne, on a

$$\begin{aligned} |{}_j \bar{\Lambda}'_a| &= \Leftrightarrow (\sigma + \mu_a) |{}_l \bar{\Lambda}'_{b_i}| \prod_{m \neq i} |\bar{\Lambda}'_{b_m}| \\ &\Leftrightarrow |{}_l \bar{\Lambda}'_{b_i}| \sum_{m \neq i} \lambda_{b_m} \mu_{b_m} |\Lambda'_{b_m}| \prod_{p \notin \{i, m\}} |\bar{\Lambda}'_{b_p}| \\ &\Leftrightarrow \lambda_{b_i} \mu_{b_i} |{}_l \Lambda'_{b_i}| \prod_{m \neq i} |\bar{\Lambda}'_{b_m}| \quad . \end{aligned}$$

Soit $j = n_1 + \cdots + n_i$. Posons $l = n_i$ alors

$$\begin{aligned} |{}_j \bar{\Lambda}'_a| &= \Leftrightarrow (\sigma + \mu_a) |{}_l \bar{\Lambda}'_{b_i}| \prod_{m \neq i} |\bar{\Lambda}'_{b_m}| \\ &\Leftrightarrow |{}_l \bar{\Lambda}'_{b_i}| \sum_{m \neq i} \lambda_{b_m} \mu_{b_m} |\Lambda'_{b_m}| \prod_{p \notin \{i, m\}} |\bar{\Lambda}'_{b_p}| \quad . \end{aligned}$$

En ajoutant toutes les autres colonnes de $|\bar{\Lambda}'_a|$ à la dernière, on a

$$|\bar{\Lambda}'_a| = \Leftrightarrow \mu_a \prod_{i=1}^k |\bar{\Lambda}'_{b_i}| = \Leftrightarrow \mu_a |\Lambda'_a| \quad .$$

Donc

$$\begin{aligned} & \sum_{j=n_1+\dots+n_{i-1}+1}^{n_1+\dots+n_i} \frac{|j\bar{\Lambda}'_a|}{|\bar{\Lambda}'_a|} \\ = & \frac{\sigma + \mu_a}{\mu_a} \sum_{l=1}^{n_i} \frac{|l\bar{\Lambda}'_{b_i}|}{|\bar{\Lambda}'_{b_i}|} + \frac{1}{\mu_a} \sum_{l=1}^{n_i} \frac{|l\bar{\Lambda}'_{b_i}|}{|\bar{\Lambda}'_{b_i}|} \sum_{m \neq i} \lambda_{b_m} \mu_{b_m} \frac{|\Lambda'_{b_m}|}{|\bar{\Lambda}'_{b_m}|} \\ & + \frac{\mu_{b_i} \lambda_{b_i}}{\mu_a} \sum_{l=1}^{n_i-1} \frac{|l\Lambda'_{b_i}|}{|\Lambda'_{b_i}|} \frac{|\Lambda'_{b_i}|}{|\bar{\Lambda}'_{b_i}|} \\ = & \Leftrightarrow \frac{\sigma + \mu_a}{\mu_a} \bar{T}_{b_i} + \frac{1}{\mu_a} \left(\sum_{m \neq i} \lambda_{b_m} \right) \bar{T}_{b_i} + \frac{\lambda_{b_i}}{\mu_a} T_{b_i} \quad \left(\text{car } \frac{|\Lambda'_{b_m}|}{|\bar{\Lambda}'_{b_m}|} = \Leftrightarrow \frac{1}{\mu_{b_m}} \right) \\ = & \Leftrightarrow \bar{T}_{b_i} \Leftrightarrow \frac{\lambda_{b_i}}{\mu_a} \bar{T}_{b_i} \Leftrightarrow T_{b_i} \\ = & \Leftrightarrow \bar{T}_{b_i} \Leftrightarrow \frac{\lambda_{b_i}}{\mu_a} t_{b_i} \quad . \end{aligned}$$

Alors

$$\begin{aligned} \Leftrightarrow \frac{1}{|\bar{\Lambda}'_a|} \sum_{i=1}^{n_a} |i\bar{\Lambda}'_a| &= \Leftrightarrow \frac{|n_a \bar{\Lambda}'_a|}{|\bar{\Lambda}'_a|} \Leftrightarrow \sum_{j=1}^{n_a-1} \frac{|j\bar{\Lambda}'_a|}{|\bar{\Lambda}'_a|} \\ &= \frac{1}{\mu_a} \Leftrightarrow \sum_{j=1}^{n-1} \frac{|j\bar{\Lambda}'_a|}{|\bar{\Lambda}'_a|} \\ &= \frac{1}{\mu_a} \Leftrightarrow \sum_{i=1}^k \sum_{j=n_1+\dots+n_{i-1}+1}^{n_1+\dots+n_i} \frac{|j\bar{\Lambda}'_a|}{|\bar{\Lambda}'_a|} \\ &= \frac{1}{\mu_a} + \sum_{i=1}^k \bar{T}_{b_i} + \frac{1}{\mu_a} \sum_{i=1}^k \lambda_{b_i} t_{b_i} \\ &= \bar{T}_a \quad . \end{aligned}$$

D'où le résultat. □

Proposition 4.1.7

$$\frac{|\Lambda_a^1|}{|\Lambda'_a|} = \frac{1}{\pi_a(a)}$$

Démonstration. On utilise les mêmes notations que dans la proposition 4.1.6. Montrons le résultat par récurrence. Si $B_a = \{a, b\}$, le résultat se vérifie facilement.

Si Λ_a est un générateur quelconque. Comme

$$|\Lambda'_a| = \prod_{i=1}^k |\bar{\Lambda}'_{b_i}| \quad ,$$

En développant $|\Lambda'_a|$ selon la dernière ligne

$$\frac{|\Lambda_a^1|}{|\Lambda'_a|} = 1 \Leftrightarrow \sum_{i=1}^k \lambda_{b_i} \frac{|\Lambda_{b_i}^1|}{|\bar{\Lambda}'_{b_i}|} \quad .$$

Mais

$$\begin{aligned} |\bar{\Lambda}'_{b_i}| &\Leftrightarrow \mu_{b_i} |\Lambda'_{b_i}| \quad , \\ \frac{|\Lambda_{b_i}^1|}{|\Lambda'_{b_i}|} &= \frac{1}{\pi_{b_i}(b_i)} \quad . \end{aligned}$$

D'où

$$\begin{aligned} \frac{|\Lambda_a^1|}{|\Lambda'_a|} &= 1 + \sum_{i=1}^k \frac{\lambda_{b_i}}{\mu_{b_i}} \frac{1}{\pi_{b_i}(b_i)} \\ &= 1 + \sum_{i=1}^k \frac{\lambda_{b_i}}{\mu_{b_i}} \frac{\pi_a(B_{b_i})}{\pi_a(b_i)} \\ &= 1 + \sum_{i=1}^k \frac{\pi_a(B_{b_i})}{\pi_a(a)} \quad (\pi_a(b_i)\mu_{b_i} = \pi_a(a)\lambda_{b_i}) \\ &= \frac{1}{\pi_a(a)} \quad . \end{aligned}$$

□

Pour démontrer la proposition 4.1.2, nous allons démontrer la forme apparemment plus générale suivante

$$S_a = \sum_{b \in B_a \setminus \{a\}} \frac{\pi_a(B_b) [1 \Leftrightarrow \pi_a(B_b)]}{\mu_b \pi_a(b)} \quad .$$

Démonstration. On utilise toujours les mêmes notations. Montrons d'abord

$$S_a = \sum_{b \in \sigma(a)} (1 \Leftrightarrow \pi_a(B_b)) \bar{T}_b + \sum_{b \in \sigma(a)} \pi_a(B_b) S_b \quad .$$

Rappelons que

$$S_a = \Leftrightarrow \frac{1}{|\Lambda_a^1|} \sum_{j=1}^{n_a-1} |{}_j\Lambda_a^1| \quad .$$

Si $j = n_1 + \dots + n_{i-1} + l$ avec $1 \leq l \leq n_i \Leftrightarrow 1$, en développant $|{}_j\Lambda_a^1|$ selon la dernière ligne, on a

$$\begin{aligned} |{}_j\Lambda_a^1| &= |{}_l\bar{\Lambda}'_{b_i}| \prod_{m \neq i} |\bar{\Lambda}'_{b_m}| \Leftrightarrow |{}_l\bar{\Lambda}'_{b_i}| \sum_{m \neq i} \lambda_{b_m} |\Lambda_{b_m}^1| \prod_{p \notin \{m, i\}} |\bar{\Lambda}'_{b_p}| \\ &\Leftrightarrow \lambda_{b_i} |{}_l\Lambda_{b_i}^1| \prod_{m \neq i} |\bar{\Lambda}'_{b_m}| \quad . \end{aligned}$$

Si $k = n_1 + \dots + n_i$, posons $l = n_i$,

$$|{}_j\Lambda_a^1| = |{}_l\bar{\Lambda}'_{b_i}| \prod_{m \neq i} |\bar{\Lambda}'_{b_m}| \Leftrightarrow |{}_l\bar{\Lambda}'_{b_i}| \sum_{m \neq i} \lambda_{b_m} |\Lambda_{b_m}^1| \prod_{p \notin \{m, i\}} |\bar{\Lambda}'_{b_p}| \quad .$$

alors

$$\begin{aligned} &\frac{1}{|\Lambda_a^1|} \sum_{j=n_1+\dots+n_{i-1}+1}^{n_1+\dots+n_i} |{}_j\Lambda_a^1| \\ &= \sum_{l=1}^{n_i} \frac{|{}_l\bar{\Lambda}'_{b_i}|}{|\bar{\Lambda}'_{b_i}|} \Leftrightarrow \left(\sum_{l=1}^{n_i} \frac{|{}_l\bar{\Lambda}'_{b_i}|}{|\bar{\Lambda}'_{b_i}|} \right) \left(\sum_{m \neq i} \lambda_{b_m} \frac{|\Lambda_{b_m}^1|}{|\bar{\Lambda}'_{b_m}|} \right) \Leftrightarrow \lambda_{b_i} \frac{|\Lambda_{b_i}^1|}{|\bar{\Lambda}'_{b_i}|} \sum_{l=1}^{n_i-1} \frac{|{}_l\Lambda_{b_i}^1|}{|\Lambda_{b_i}^1|} \\ &= \Leftrightarrow \bar{T}_{b_i} \left(1 \Leftrightarrow \sum_{m=1}^k \lambda_{b_m} \frac{|\Lambda_{b_m}^1|}{|\bar{\Lambda}'_{b_m}|} \right) \Leftrightarrow \bar{T}_{b_i} \lambda_{b_i} \frac{|\Lambda_{b_i}^1|}{|\bar{\Lambda}'_{b_i}|} + \lambda_{b_i} S_{b_i} \frac{|\Lambda_{b_i}^1|}{|\bar{\Lambda}'_{b_i}|} \\ &= \Leftrightarrow \bar{T}_{b_i} \frac{|\Lambda_a^1|}{|\Lambda_a^1|} \Leftrightarrow \bar{T}_{b_i} \lambda_{b_i} \frac{|\Lambda_{b_i}^1|}{|\bar{\Lambda}'_{b_i}|} + \lambda_{b_i} S_{b_i} \frac{|\Lambda_{b_i}^1|}{|\bar{\Lambda}'_{b_i}|} \\ &= \Leftrightarrow \bar{T}_{b_i} \left(\frac{1}{\pi_a(a)} \Leftrightarrow \frac{\lambda_{b_i}}{\mu_{b_i} \pi_{b_i}(b_i)} \right) \Leftrightarrow S_{b_i} \frac{\lambda_{b_i}}{\mu_{b_i} \pi_{b_i}(b_i)} \\ &= \Leftrightarrow \bar{T}_{b_i} \frac{1 \Leftrightarrow \pi_a(B_{b_i})}{\pi_a(a)} \Leftrightarrow S_{b_i} \frac{\pi_a(B_{b_i})}{\pi_a(a)} \quad . \end{aligned}$$

Alors

$$\Leftrightarrow \frac{1}{|\Lambda_a^1|} \sum_{j=1}^{n_a-1} |{}_j\Lambda_a^1|$$

$$\begin{aligned}
&= \Leftrightarrow \frac{|\Lambda'_a|}{|\Lambda_a^1|} \sum_{j=1}^{n_a-1} \frac{|j\Lambda_a^1|}{|\Lambda'_a|} \\
&= \Leftrightarrow \pi_a(a) \sum_{i=1}^k \sum_{j=n_0+\dots+n_{i-1}+1}^{n_1+\dots+n_i} \frac{|j\Lambda_a^1|}{|\Lambda'_a|} \\
&= \sum_{i=1}^k \{ [1 \Leftrightarrow \pi_a(B_{b_i})] \bar{T}_{b_i} + \pi_a(B_{b_i}) S_{b_i} \} .
\end{aligned}$$

Montrons le résultat par récurrence. Si $B_a = \{a, b\}$, le résultat est vrai.

Pour un sommet a quelconque, si $\sigma(a) \neq \emptyset$,

$$\begin{aligned}
S_a &= \sum_{b \in \sigma(a)} [1 \Leftrightarrow \pi_a(B_b)] \bar{T}_b + \sum_{b \in \sigma(a)} \pi_a(B_b) S_b \\
&= \sum_{b \in \sigma(a)} [1 \Leftrightarrow \pi_a(B_b)] \sum_{c \in B_b} t_c + \sum_{b \in \sigma(a)} \pi_a(B_b) \sum_{c \in B_b \setminus \{b\}} \frac{\pi_b(B_c) [1 \Leftrightarrow \pi_b(B_c)]}{\mu_c \pi_b(c)} \\
&= \sum_{b \in B_a \setminus \{a\}} \frac{\pi_a(B_b) [1 \Leftrightarrow \pi_a(B_b)]}{\mu_b \pi_a(b)} .
\end{aligned}$$

□

Remarque Si on remplace dans le théorème précédent $1/\mu_b$ par

$$t_b \frac{\pi(b)}{\pi(B_b)} ,$$

on obtient une expression de S_a qui relie la moyenne harmonique du spectre d'une branche aux temps de parcours des arêtes.

$$S_a = \sum_{b \in B_a \setminus \{a\}} [1 \Leftrightarrow \pi_a(B_b)] t_b .$$

4.2 Applications

4.2.1 Gap géométrique

La notion de laplacien canonique est classique en théorie des graphes. Considérons un graphe fini sans auto boucle et non orienté $G = (E, A)$ où E est l'ensemble des sommets, A l'ensemble des arêtes. La matrice $\Lambda = (\lambda_{ij})_{i,j \in E}$ est

définie par

$$\lambda_{ij} = \begin{cases} \Leftrightarrow 1 & \text{si } (i, j) \in A \\ \text{degré}(i) & \text{si } i = j \\ 0 & \text{sinon .} \end{cases}$$

On dit que Λ est le *laplacien canonique* sur le graphe G (voir [17]). Si G est connexe, les valeurs propres de Λ vérifient

$$0 = \lambda_0 < \lambda_1 \leq \lambda_2 \cdots \leq \lambda_{\#E-1} .$$

On définit le *gap de G* par

$$\text{gap}(G) = \lambda_1 .$$

En fait, selon le point de vue Markovien, $\Leftrightarrow \Lambda$ est le générateur de la marche aléatoire symétrique sur G et

$$\text{gap}(\Leftrightarrow \Lambda) = \text{gap}(G) .$$

Le résultat suivant est une application immédiate du théorème B en remarquant que la mesure réversible de $\Leftrightarrow \Lambda$ est uniforme.

Proposition 4.2.1 *Soit G un graphe connexe non orienté sur E fini. Si G est un arbre orienté par une racine a_0 , B_a étant la branche de G' issue du sommet a , alors*

$$\left[\sum_{a \in E} \#B_a \left(1 \Leftrightarrow \frac{\#B_a}{\#E} \right) \right]^{-1} \leq \text{gap}(G) \leq (\#E \Leftrightarrow 1) \left[\sum_{a \in E} \#B_a \left(1 \Leftrightarrow \frac{\#B_a}{\#E} \right) \right]^{-1} .$$

Sinon, soit G' un arbre maximal de G orienté, alors

$$\text{gap}(G) \geq \left[\sum_{a \in E} \#B_a \left(1 \Leftrightarrow \frac{\#B_a}{\#E} \right) \right]^{-1} .$$

A titre d'exemple, nous calculons explicitement la SIVP ainsi que la minoration du gap pour le laplacien d'un arbre homogène.

Exemple 4.2.1 .

Soit $G = (E, A)$ l'arbre homogène de degré d de hauteur h . C'est-à-dire les degrés des sommets sauf les feuilles sont tous d et il existe une racine, notée a_0 dont la distance à chaque feuille est h . Alors a_0 a d fils et les autres sommets sauf les feuilles en ont $d \Leftrightarrow 1$. La figure 4.1 représente par exemple l'arbre homogène de degré 3 et de hauteur 2. Pour tout sommet a , la distance entre a et chaque feuille du sous-arbre issu de a , B_a , est la même.

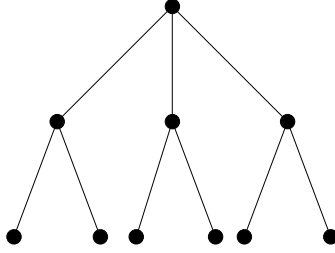


Figure 4.1: Arbre homogène de degré 3 et de hauteur 2

On l'appelle la hauteur de a . Si a est différent de a_0 et de hauteur i , pour $i = 1, \dots, h \Leftrightarrow 1$

$$\#B_a = \sum_{j=0}^i (d \Leftrightarrow 1)^{i-j} = \frac{(d \Leftrightarrow 1)^{i+1} \Leftrightarrow 1}{d \Leftrightarrow 2} .$$

Comme la racine a_0 a d fils,

$$\#E = 1 + d \sum_{i=0}^{h-1} (d \Leftrightarrow 1)^{h-1-i} = \frac{d(d \Leftrightarrow 1)^h \Leftrightarrow 2}{d \Leftrightarrow 2} .$$

Constatons que, pour $i = 0 \dots h \Leftrightarrow 1$, il y a $d(d \Leftrightarrow 1)^{h-1-i}$ sommets de hauteur i . Il vient que la SIVP du laplacien de l'arbre

$$\begin{aligned} S &= \sum_{a \in E} \#B_a \left(1 \Leftrightarrow \frac{\#B_a}{\#E} \right) \\ &= \sum_{i=0}^{h-1} d(d \Leftrightarrow 1)^{h-1-i} \frac{(d \Leftrightarrow 1)^{i+1} \Leftrightarrow 1}{d \Leftrightarrow 2} \left[1 \Leftrightarrow \frac{(d \Leftrightarrow 1)^{i+1} \Leftrightarrow 1}{d(d \Leftrightarrow 1)^h \Leftrightarrow 2} \right] \\ &= \frac{d(d \Leftrightarrow 1)^{h-1}}{(d \Leftrightarrow 2)[d(d \Leftrightarrow 1)^h \Leftrightarrow 2]} \times \\ &\quad \sum_{i=0}^{h-1} \left[d(d \Leftrightarrow 1)^{h+1} \Leftrightarrow d(d \Leftrightarrow 1)^{h-i} \Leftrightarrow (d \Leftrightarrow 1)^{i+2} + \frac{1}{(d \Leftrightarrow 1)^i} \right] \end{aligned}$$

$$\begin{aligned}
 &= \frac{d}{(d \leftrightarrow 2)^2 [d(d \leftrightarrow 1)^h \leftrightarrow 2]} \times \\
 &\quad [hd(d \leftrightarrow 2)(d \leftrightarrow 1)^{2h} \leftrightarrow d(d \leftrightarrow 1)^{2h} + d(d \leftrightarrow 1)^h \\
 &\quad \leftrightarrow (d \leftrightarrow 1)^{2h+1} + (d \leftrightarrow 1)^{h+1} + (d \leftrightarrow 1)^h \leftrightarrow 1] \\
 &= \frac{d\{(d \leftrightarrow 1)^{2h} [hd^2 \leftrightarrow (2h + 2)d + 1] + 2d(d \leftrightarrow 1)^h \leftrightarrow 1\}}{(d \leftrightarrow 2)^2 [d(d \leftrightarrow 1)^h \leftrightarrow 2]} .
 \end{aligned}$$

D'après la proposition 4.2.1,

$$\text{gap}(G) \geq \frac{(d \leftrightarrow 2)^2 [d(d \leftrightarrow 1)^h \leftrightarrow 2]}{d\{(d \leftrightarrow 1)^{2h} [hd^2 \leftrightarrow (2h + 2)d + 1] + 2d(d \leftrightarrow 1)^h \leftrightarrow 1\}} . \quad (4.2)$$

Par exemple pour l'arbre homogène de degré 5 et de hauteur 2, le nombre de sommets est 25. A l'aide de Mathematica, on obtient

$$\text{gap} = 0.17 ,$$

or l'estimation donnée par (4.2) est

$$0.025 .$$

Exemple 4.2.2

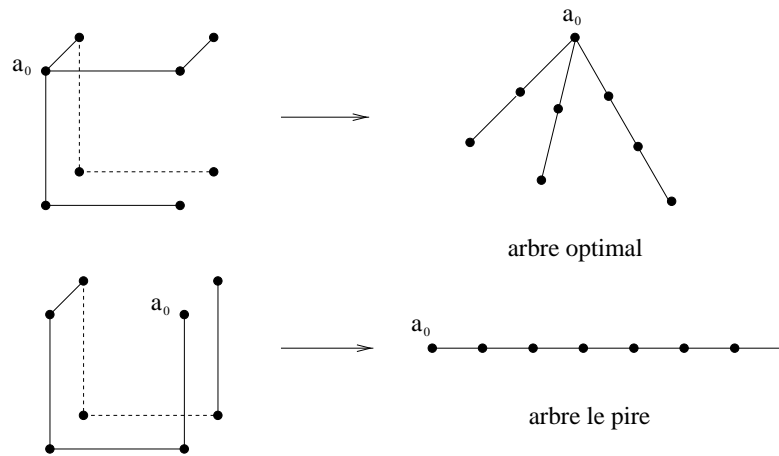


Figure 4.2: Deux arbres maximaux du cube.

Considérons le laplacien sur un cube. Son gap se calcule facilement.

$$\text{gap} = 2 .$$

Estimons le gap selon le théorème B. Dans la figure 4.2, on donne deux arbres maximaux du cube: l'arbre optimal au sens où la SIVP du laplacien sur cet arbre est la plus petite parmi tous les arbres maximaux du cube et l'arbre le pire au sens où sa SIVP est la plus grande. Selon le théorème B, l'estimation du gap donnée par l'inverse du SIVP est

$$0,103$$

pour l'arbre optimal et

$$0,095$$

pour l'arbre le pire.

Remarque On voit que les estimations du gap dans le cas de l'arbre homogène de degré 5 et de hauteur 2 (exemple 4.2.1) et celui du cube (exemple 4.2.2) ne sont pas bonnes. En fait les sources d'erreurs d'estimation sont doubles. La première vient du fait de remplacer le générateur initial par un générateur arborescent. La deuxième se produit quand on estime le gap du générateur arborescent choisi par l'inverse de la SIVP. Pour l'erreur du deuxième type, on ne peut pas la diminuer pour un générateur fixé. Par contre choisir un générateur arborescent tel que sa SIVP soit la plus petite possible diminue l'erreur totale. Malheureusement, on a pas un critère fixe de choix d'un arbre maximal optimal du générateur initial. L'estimation du gap donnée par le théorème B sera bonne si le générateur arborescent extrait du générateur initial conserve les taux de transition rapides et l'arbre associé est assez concentré (ce qui n'est pas toujours possible), c'est le cas pour le processus d'Ising qu'on va traiter au paragraphe suivant. Pour l'arbre homogène de degré 5 et de hauteur 2, l'erreur est évidemment du deuxième type. Par contre dans le cas du cube, les deux types d'erreur se présentent, l'erreur du premier type doit être considérable, parce qu'on est obligé d'enlever des taux de transition rapides du cube.

4.2.2 Processus d'Ising

Le but de cette partie est d'appliquer la technique développée dans les paragraphes précédents à l'estimation du gap des processus d'Ising en dimension 1. La comparaison avec les résultats donnés par la proposition 3.2.4 montre que la minoration donnée par le théorème B fournit dans ce cas l'ordre de grandeur correct du gap.

En posant

$$a(x, \eta) = \exp \frac{H(\eta) \Leftrightarrow H(\eta_x)}{T},$$

considérons la famille des processus d'Ising de taux de transition

$$c(x, \eta) = \begin{cases} f(a(x, \eta)) & \text{si } H(\eta_x) < H(\eta) \\ a(x, \eta)^{-1} f(a(x, \eta)^{-1}) & \text{sinon} \end{cases}$$

où la fonction f est analytique dans un voisinage de 0. Selon les conditions de bord, nous avons les minorations du gap suivantes, quand la température tend vers 0.

Proposition 4.2.2 *Pour la condition de bord périodique, on a*

$$\text{gap}(\Lambda^p) \geq \frac{2f(0)f(1)}{2f(1) + (n \Leftrightarrow 2)f(0)} \exp \left(\frac{4}{T} \right) + o \left(\exp \left(\frac{4}{T} \right) \right).$$

Pour la condition de bord ouverte, on a

$$\text{gap}(\Lambda^o) \geq \frac{2f(0)f(1)}{2f(1) + (n \Leftrightarrow 2)f(0)} \exp \left(\frac{2}{T} \right) + o \left(\exp \left(\frac{2}{T} \right) \right).$$

Pour la condition de bord unitaire,

$$\text{gap}(\Lambda^u) = \theta(1).$$

Ces minorations sont à comparer aux équivalents donnés dans la proposition 3.2.4. Pour les trois conditions de bords, le minorant a le bon ordre de grandeur en $\exp \left(\frac{1}{T} \right)$. Le coefficient diffère presque d'un facteur $2n$ pour la condition de bord périodique, d'un facteur 2 pour la condition de bord ouverte.

Pour démontrer la proposition 4.2.2, on commence par partitionner l'espace d'états selon les niveaux d'énergie. Considérons $\{P_0, \dots, P_m\}$, la partition de Ω

telle que pour tout $0 \leq i \leq m$, l'énergie des configurations de P_i soit constante et que les énergies constantes des P_i soient strictement croissantes. Chaque transition augmente ou diminue le niveau d'énergie d'au plus 1 palier dans les cas périodique et unitaire, au plus 2 paliers dans le cas ouvert. Autrement dit, il n'y a pas de transitions entre deux configurations appartenant à P_i et P_j respectivement avec $|i \Leftrightarrow j| \geq 2$ (resp. $|i \Leftrightarrow j| \geq 3$) dans les cas périodique et unitaire (resp. dans le cas ouvert).

Considérons d'abord les cas des conditions de bord périodique et ouvert. On sait que quand la température tend vers 0, la mesure stationnaire tend vers la mesure uniforme sur P_0 . Intuitivement, la vitesse de parcours de toutes les configurations de P_0 par les chemins les plus rapides du processus mesure en quelque sorte le taux de convergence du processus initial. Notons que P_0 ne contient que deux configurations dont les sites sont soit tous à 0 soit tous à 1. On note ces deux configurations par η^0 et η^1 respectivement. D'ailleurs P_1 contient les configurations ayant deux composantes connexes. On constate aussi que les chemins les plus rapides joignant η^0 et η^1 au sens du temps moyen d'atteinte de η^1 en partant de η^0 sont des chemins dans $P_0 \cup P_1$. On considère ainsi le processus d'Ising restreint sur le sous-ensemble

$$\Omega' = P_0 \cup P_1 .$$

Et on note Λ' la restriction de Λ sur Ω' . Nous montrons d'abord la proposition suivante.

Proposition 4.2.3 *Quand la température tend vers 0, dans le cas périodique,*

$$gap(\Lambda') = gap(\Lambda) + o\left(\exp\left(\frac{4}{T}\right)\right)$$

et dans le cas ouvert

$$gap(\Lambda') = gap(\Lambda) + o\left(\exp\left(\frac{2}{T}\right)\right) .$$

Démonstration. Voici la démonstration du cas périodique. Pour le cas ouvert, la démonstration est analogue.

Posons

$$\varepsilon = \exp\left(\frac{2}{T}\right).$$

Soient π et π' les mesures de probabilité réversibles de Λ et Λ' respectivement.

Pour toute configuration η de P_i avec $i = 0, \dots, m$, on a

$$\pi(\eta) = \frac{1}{2}\varepsilon^{2i} + O\left(\varepsilon^{2(i+1)}\right). \quad (4.3)$$

Alors si $\eta \in \Omega'$, on a

$$\pi'(\eta) = \frac{\pi(\eta)}{\sum_{\xi \in \Omega'} \pi(\xi)} = \frac{\pi(\eta)}{1 + o(\varepsilon^2)} = \pi(\eta) + o(\varepsilon^2).$$

Notons $D_\Lambda, D_{\Lambda'}$ les formes de Dirichlet de Λ et Λ' respectivement. Soit f le vecteur de \mathbb{R}^Ω tel que

$$\pi(f) = 0, \quad \|f\|_\pi^2 = 1$$

et

$$D_\Lambda(f) = \text{gap}(\Lambda).$$

Evidemment, pour tout $\eta \in \Omega$,

$$f(\eta) = O(1).$$

Si $f_{\Omega'}$ désigne la restriction de f sur Ω' , on a

$$\begin{aligned} \pi'(f_{\Omega'}) &= \sum_{\eta \in \Omega'} \pi'(\eta)f(\eta) \\ &= \sum_{\eta \in \Omega'} \pi(\eta)f(\eta) + o(\varepsilon^2) \\ &= \sum_{\eta \in \Omega \setminus \Omega'} \pi(\eta)f(\eta) + o(\varepsilon^2) \\ &= o(\varepsilon^2) \quad (\text{d'après (4.3)}) \end{aligned}$$

et

$$\begin{aligned} \|f_{\Omega'}\|_{\pi'}^2 &= \sum_{\eta \in \Omega'} \pi'(\eta)f^2(\eta) \\ &= \sum_{\eta \in \Omega'} \pi(\eta)f^2(\eta) + o(\varepsilon^2) \\ &= 1 \Leftrightarrow \sum_{\eta \in \Omega \setminus \Omega'} \pi(\eta)f^2(\eta) + o(\varepsilon^2) \\ &= 1 + o(\varepsilon^2) \quad (\text{d'après (4.3)}). \end{aligned}$$

Il existe donc un vecteur g' de $\mathbb{R}^{\Omega'}$ tel que

$$\pi'(g') = 0, \quad \|g'\|_{\pi'}^2 = 1$$

et

$$g' = f_{\Omega'} + o(\varepsilon_{\Omega'}^2).$$

Alors

$$\begin{aligned} D_{\Lambda}(f) &= \frac{1}{2} \sum_{\eta, \xi \in \Omega} [f(\eta) \leftrightarrow f(\xi)]^2 \pi(\eta) \lambda(\eta, \xi) \\ &= \frac{1}{2} \sum_{\eta, \xi \in \Omega'} [f(\eta) \leftrightarrow f(\xi)]^2 \pi(\eta) \lambda(\eta, \xi) \\ &\quad + \frac{1}{2} \sum_{\eta, \xi \in \Omega \setminus \Omega'} [f(\eta) \leftrightarrow f(\xi)]^2 \pi(\eta) \lambda(\eta, \xi) \\ &\quad + \sum_{\substack{\eta \in \Omega' \\ \xi \in \Omega \setminus \Omega'}} [f(\eta) \leftrightarrow f(\xi)]^2 \pi(\eta) \lambda(\eta, \xi). \end{aligned}$$

On remarque qu'il n'y a pas de taux de transition de P_0 vers $\Omega \setminus \Omega'$ et pour toute configuration η de P_1 , ξ de $\Omega \setminus \Omega'$

$$\pi(\eta) = O(\varepsilon^2), \quad \lambda(\eta, \xi) = O(\varepsilon^2) \quad .$$

Donc

$$\begin{aligned} D_{\Lambda}(f) &= D_{\Lambda'}(f_{\Omega'}) + o(\varepsilon^2) \\ &= D_{\Lambda'}(g') + o(\varepsilon^2) \\ &\geq \text{gap}(\Lambda') + o(\varepsilon^2) \quad . \end{aligned}$$

Supposons que f' est le vecteur de $\mathbb{R}^{\Omega'}$ tel que

$$\pi'(f') = 0, \quad \|f'\|_{\pi'}^2 = 1$$

et

$$\text{gap}(\Lambda') = D_{\Lambda'}(f').$$

Soit f'_{Ω} le vecteur prolongé de f' par 0 dans \mathbb{R}^{Ω} . On vérifie que

$$\pi(f'_{\Omega}) = o(\varepsilon^2) \quad \text{et} \quad \|f'_{\Omega}\|_{\pi}^2 = 1 + o(\varepsilon^2) \quad .$$

Il existe donc un vecteur g de \mathbb{R}^Ω tel que

$$\pi(g) = 0 \text{ , } \|g\|_\pi^2 = 1$$

et

$$g = f'_\Omega + o(\varepsilon_\Omega^2) \text{ .}$$

Alors

$$\begin{aligned} D_{\Lambda'}(f') &= \frac{1}{2} \sum_{\eta, \xi \in \Omega'} [f'(\eta) \leftrightarrow f'(\xi)]^2 \pi(\eta) \lambda_{\eta, \xi} \\ &= \frac{1}{2} \sum_{\eta, \xi \in \Omega} [g(\eta) \leftrightarrow g(\xi)]^2 \pi(\eta) \lambda_{\eta, \xi} \\ &\Leftrightarrow \frac{1}{2} \sum_{\eta, \xi \in \Omega \setminus \Omega'} [g(\eta) \leftrightarrow g(\xi)]^2 \pi(\eta) \lambda_{\eta, \xi} \\ &\Leftrightarrow \sum_{\substack{\eta \in \Omega' \\ \xi \in \Omega \setminus \Omega'}} [g(\eta) \leftrightarrow g(\xi)]^2 \pi(\eta) \lambda_{\eta, \xi} + o(\varepsilon^2) \\ &= D_\Lambda(g) + o(\varepsilon^2) \\ &\geq \text{gap}(\Lambda) + o(\varepsilon^2) \text{ .} \end{aligned}$$

d'où le résultat.

□

Démontrons la proposition 4.2.2.

Démonstration. Dans les cas périodique et ouvert, choisissons un arbre maximal noté $G'' = (\Omega', A'')$ du générateur Λ' . Soit Λ'' le générateur restreint du générateur Λ' sur G'' . Il suffit de montrer les minoration dans la proposition 4.2.2 pour $\text{gap}(\Lambda'')$.

Soit ξ_i^j la configuration ayant deux composantes connexes dont celle de valeur 1 débute au site i et termine au site $(i + j \Leftrightarrow 1)$ modulo n .

Dans le cas périodique, définissons l'ensemble des arêtes A'' par

$$\begin{aligned} A'' &= \{ \eta^0, \xi_i^1 \text{ ; } i = 0, \dots, n \Leftrightarrow 1 \} \cup \{ \xi_0^{n-1}, \eta^1 \} \\ &\cup \{ (\xi_i^{j-1}, \xi_i^j) \text{ ; } 0 \leq i \leq n \Leftrightarrow 1 \text{ / } \langle j \leq n \Leftrightarrow 1 \} \text{ .} \end{aligned}$$

On vérifie que le graphe $G'' = (\Omega', A'')$ est un arbre maximal du graphe réduit de Λ' . Choisissons η^0 comme racine de l'arbre. On a alors

$$\phi^* (B_{\xi_0^1}) = 2f(0)\varepsilon^2 + o(\varepsilon^2) \quad ,$$

$$\phi^* \left(\begin{array}{c} B_{\eta^1} \\ / \end{array} \right) = 2f(0)\varepsilon^2 + o(\varepsilon^2)$$

et

$$\phi^* (B_{\xi_i^j}) = \begin{cases} 2f(1)\varepsilon^2 + o(\varepsilon^2) & \text{si } i = 0, j = 2 \cdots n \Leftrightarrow 1 \\ O(1) & \text{si } 1 \leq i \leq n \Leftrightarrow 1, j = 1 \cdots n \Leftrightarrow 1 \quad . \end{cases}$$

On en déduit que

$$\left[\sum_{\substack{\eta \in \Omega' \\ \eta \neq \eta^0}} \frac{1}{\phi^*(B_\eta)} \right]^{-1} = \frac{2f(0)f(1)}{f(0)(n \Leftrightarrow 2) + 2f(1)} \varepsilon^2 + o(\varepsilon^2) \quad .$$

D'où la minoration du gap.

Dans le cas ouvert, définissons A'' par

$$A'' = \left\{ \begin{array}{l} \left\{ \eta^0, \xi_0^1 \right\} \cup \left\{ \left(\xi_0^{j-1}, \xi_0^j \right) ; 2 \leq j \leq n \right\} \\ \cup \left\{ \eta^0, \xi_{n-1}^1 \right\} \cup \left\{ \xi_{i+1}^{n-i-1}, \xi_i^{n-i} ; 1 \leq i \leq n \Leftrightarrow 2 \right\} \quad . \end{array} \right.$$

Pareillement, choisissons η^0 comme racine de l'arbre $G'' = (\Omega', A'')$. On a alors

$$\phi^* (B_{\xi_0^1}) = 2f(0)\varepsilon + o(\varepsilon) \quad , \quad \phi^* \left(\begin{array}{c} B_{\eta^1} \\ / \end{array} \right) = 2f(0)\varepsilon + o(\varepsilon)$$

et

$$\phi^* (B_{\xi_i^j}) = \begin{cases} 2f(1)\varepsilon + o(\varepsilon) & \text{si } i = 0, j = 2 \cdots n \Leftrightarrow 1 \\ O(1) & \text{si } 1 \leq i \leq n \Leftrightarrow 1, j = n \Leftrightarrow i + 1 \quad . \end{cases}$$

On en déduit que

$$\left[\sum_{\substack{\eta \in \Omega' \\ \eta \neq \eta^0}} \frac{1}{\phi^*(B_\eta)} \right]^{-1} = \frac{2f(0)f(1)}{f(0)(n \Leftrightarrow 2) + 2f(1)} \varepsilon + o(\varepsilon) \quad .$$

D'où la minoration du gap.

Dans le cas unitaire, constatons que la restriction de Λ sur P_n est irréductible pour n de 0 à m . Il existe alors un arbre maximal du graphe réduit de Λ qui ne contient qu'une arête liant P_n et P_{n+1} pour $0 \leq n \leq m \Leftrightarrow 1$. Notons $G = (\Omega, A)$ un tel arbre.

On constate que P_0 ne contient qu'une configuration η^1 . Choisissons η^1 comme racine de G . On constate que pour toute configuration η de Ω , si ξ est une configuration de la branche B_η issue de η ,

$$\pi_T(\eta) \leq \pi_T(\xi) \quad .$$

Alors quand la température tend vers 0, pour η différente de η^1 , $\pi_T(B_\eta)$ et $\pi_T(\eta)$ sont de même ordre de grandeur,

$$\frac{\pi_T(\eta)}{\pi_T(B_\eta)} = \theta(1) \quad .$$

De plus

$$1 \Leftrightarrow \pi_T(B_\eta) = 1 + o(1) \quad ,$$

et le taux de transition de η vers son père $p(\eta)$

$$\lambda_{\eta,p(\eta)} = \theta(1) \quad .$$

Alors

$$\phi^*(B_\eta) = \frac{\pi_T(\eta)\lambda_{\eta,p(\eta)}}{\pi_T(B_\eta)[1 \Leftrightarrow \pi_T(B_\eta)]} = \theta(1) \quad .$$

D'où le résultat.

□

Bibliographie

- [1] D. Aldous and M. Brown. Inequalities for rare events in time-reversible Markov chains. *IMS Lecture Notes-Monograph, Stochastic Inequalities*, 22:1–16, 1993.
- [2] D. Bertsimas and J. Tsitsiklis. Simulated annealing. *Statistical Science*, 8:10–15, 1993.
- [3] J. Besag. On the statistical analysis of dirty pictures. *J. R. Statist. Soc.*, 48(3):259–302, 1986.
- [4] J. Besag and P.J. Green. Spatial statistics and Bayesian computation. *J. R. Statist. Soc.*, B 55:25–37, 1993.
- [5] O. Catoni. Applications of sharp large deviations estimates to optimal cooling schedules. *Annales de l'Institut Henri Poincaré*, 27(4):463–518, 1991.
- [6] I. Chavel. *Eigenvalues in Riemannian geometry*. Academic Press, New York, 1984.
- [7] J. Cheeger. A lower bound for the lowest eigenvalue of the Laplacian. *Problems in Analysis: A Symposium in Honor of S. Bocher (R. C. Gunnind, ed.) Princeton Univ. Press, Princeton. N.J.*, pages 195–199, 1970.
- [8] D.Y. Chen, J.F. Feng, and M.P. Qian. Metastability of exponentially perturbed Markov chains. *Science in China*, 39(1):7–28, 1996.

- [9] M. F. Chen. *From Markov chains to non-equilibrium particle systems*. World Scientific, 1992.
- [10] M. F. Chen. Optimal Markovian couplings and applications. *Acta Math. Sin. New Ser.*, 10(3):260–275, 1994.
- [11] M. F. Chen. Estimation of spectral gap for Markov chains. *Acta Math. Sin. New Ser.*, 12(4):337–360, 1996.
- [12] M. F. Chen. Trilogy of couplings and general formulas for lower bound of spectral gap. *Proceedings of the Symposium on probability towards the Year 2000*, pages 337–360, 1996.
- [13] F. R. K. Chung. Diameters and eigenvalues. *Journal of AMS*, 2(2):187–196, 1989.
- [14] E. Çinlar. *Introduction to stochastic processes*. Prentice Hall, New York, 1975.
- [15] M. Coderch, A.S. Willsky, S.S. Sastry, and Castanon D.A. Hierarchical aggregation of linear systems with multiple time scales. *IEEE Trans. Automatic Control*, AC-28(11):1017–1030, 1983.
- [16] M. Coderch, A.S. Willsky, S.S. Sastry, and Castanon D.A. Hierarchical aggregation of singularly perturbed finite state Markov processes. *Stochastics*, 8:259–289, 1983.
- [17] Y. Colin de Verdière. *Spectres de graphes*. Cours de DEA, Université Joseph Fourier -ENS Lyon, 1994.
- [18] P.J. Courtois. *Decomposability, queuing and computer systems applications*. Academic Press, New York, 1977.
- [19] F. Delebecque. A reduction process for perturbed markov chains. *SIAM J. Appl. Math.*, 43(2):325–350, 1983.

- [20] P. Diaconis and L. Saloff-Coste. Comparison techniques for random walk on finite groups. *Ann. Appl. Probab.*, 21:2131–2156, 1993.
- [21] P. Diaconis and L. Saloff-Coste. Comparison theorems for reversible Markov chains. *Ann. Appl. Probab.*, 3:696–730, 1993.
- [22] P. Diaconis and L. Saloff-Coste. What do we know about the Metropolis algorithm. To appear in *J. Comp. Sci. Syst.*, 1996.
- [23] P. Diaconis and D.W. Strook. Geometric bounds for eigenvalues of Markov chains. *Ann. Appl. Probab.*, 1(1):36–61, 1991.
- [24] J. Feng and M. Qian. The role of noise in the neural network. *Proceedings of IJCNN*, II:65–71, 1992.
- [25] J.A. Fill. Eigenvalue bounds on convergence to the stationarity for nonreversible markov chains, with an application to the exclusion process. *Ann. Appl. Probab.*, 1(1):62–87, 1991.
- [26] O. François. Convergence of Monte-Carlo Markov chains by conductances. Technical Report 37, projet MAI, IMAG Grenoble, France, 1997.
- [27] M.I. Freidlin and A.D. Wentzell. *Random perturbations of dynamical systems*. Springer Verlag, New York, 1984.
- [28] A. Frigessi, F. Martinelli, and J. Sander. Computational complexity of MCMC methods. *Prépublication*, 1994.
- [29] A. Frigessi and P. Stefano. Convergence rates of the Gibbs sampler, the Metropolis algorithm and other single-site updating dynamics. *J.R. Statist. Soc.*, B 55:205–219, 1993.
- [30] A. Gelman and D. B. Rubin. Inference from iterative simulation using multiple sequences (with discussion). *Statistical Science*, 7:457–511, 1992.
- [31] D. Geman. Random fields and inverse problems in imaging. *Lect. Notes. Math.*, Ecole d'été de probabilités de Saint-Flour XVIII-1988:113–193, 1990.

- [32] D. Geman and S. Geman. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-6:721–741, 1984.
- [33] R.J. Glauber. Time dependent statistics of the Ising model. *J. Math. Phys.*, 4:294–307, 1963.
- [34] U. Grenander. *Tutorial in pattern theory*. Brown University, Providence, 1984.
- [35] R. Holley, S. Kusuoka, and D. Stroock. Asymptotics of the spectral gap with the applications to the theory of simulated annealing. *J.F.A.*, 93:333–347, 1989.
- [36] R. Horn and C. Johnson. *Matrix analysis*. Cambridge University Press, 1985.
- [37] C.R. Hwang and S.J. Sheu. Singular perturbed markov chains and exact behaviors of simulated annealing processes. *Journal of Theoretical Probability*, 5(2), 1992.
- [38] S. Ingrassia. On the rates of convergence of Metropolis algorithm and Gibbs sampler. *Ann. Appl. Prob.*, 4:347–389, 1994.
- [39] I. Iscoe and D. McDonald. Asymptotics of exit times for Markov jump processes (I). *Ann. App. Prob.*, 22(1):372–397, 1994.
- [40] M. Jerrum and A. Sinclair. The Markov Chain Monte Carlo method: an approach to approximate counting and intergration. *In approximation algorithms for NP-hard problems, D.S. Hochbaum (Ed.), PWS publishing, Boston*, 1997.
- [41] N. Kannan. Markov chains and polynomial time algorithms. *Proceedings of the 35th IEEE symposium on foundations of computer science, Computer Society press*, pages 656–671, 1994.

- [42] S. Karlin. *A first course in stochastic processes*. Academic Press, San Diego, 1966.
- [43] S. Karlin and H.M. Taylor. *A second course in stochastic processes*. Academic Press, San Diego, 1981.
- [44] T. Kato. *Perturbation theory for linear operators*. Springer Verlag, New York, 1978.
- [45] F.P. Kelly. *Reversibility and stochastic networks*. Wiley, London, 1979.
- [46] S. Kirkpatrick, C.D. Gellat, and M.P. Vecchi. Optimisation by simulated annealing. *Science*, 220:671–680, 1983.
- [47] G. Lawler and A. Sokal. Bounds on the l^2 spectrum for Markov chains and Markov processes: a generalization of Cheeger's inequality. *Trans. Amer. Math. Soc.*, 309:557–580, 1988.
- [48] M. Lewenstein and A. Nowak. Fully connected neural networks with self control of noise levels. *Phy. Rev. Letters*, 62(2):225, 1989.
- [49] P. Li and S.T. Yau. Estimates of eigenvalues of a compact Riemannian manifold. *Ann. Math. Soc. Proc. Symp. Pure Math.*, 36:205–240, 1980.
- [50] T. M. Liggett. *Interacting Particle Systems*. Springer, New York, 1985.
- [51] J. Liu. Eigenanalysis for a metropolis sampling scheme with comparisons to rejection sampling and importance sampling. *Statistics and Computing*, A paraitre, 1995.
- [52] S. Maille. Etude de processus de Markov associés au modèle de Ising: attractivité et couplage, existence de densité. *Thèse de l'Université Paul Sabatier Toulouse*, 1994.
- [53] F. Martinelli. On the two-dimensional dynamical Ising model in phase coexistence region. *J. Statist. Phys.*, 76:1179–1246, 1994.

- [54] P. Mathieu. Zero white noise limit through Dirichlet forms. *Probability Theory and Related Fields*, A paraitre, 1995.
- [55] B. M. McCoy and T. T. Wu. *The two-dimensional Ising model*. Harvard Univ. Press, Cambridge, Massachusetts., 1973.
- [56] M. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller. Equation of state calculation by fast computing machines. *J. Chem. Phys.*, 21:1087–1092, 1953.
- [57] L. Miclo. Comportement de spectres d'opérateurs de Schrödinger à basse température. *Bull. Sc. math.*, 119:529–553, 1995.
- [58] L. Miclo. Une étude des algorithmes de recuit simulé sous-admissibles. *Annales de la Faculté des sciences de Toulouse*, 4(4):819–877, 1995.
- [59] E. Neves and R. Schonmann. Critical droplets and metastability for a glauher dynamics at very low tempreture. *Commun. Math. Phys.*, 137:209, 1991.
- [60] C. P. Robert. *Méthodes de Monte-Carlo par chaînes de Markov*. CREST-INSEE, Paris, 1996.
- [61] L. Saloff-Coste. Communication privée. 1995.
- [62] L. Saloff-Coste. Lectures on finite Markov chains. Ecole d'été de probabilités de Saint-Flour, 1996.
- [63] R. Schoen and S. T. Yau. *Differential Geometry*. Science Press, Beijing, China, 1988.
- [64] R. Schonmann. Slow droplet-driven relaxation of stochastic Ising models in the vicinity of the phase coexistence region. *Commun. Math. Phys.*, 161:1–49, 1994.
- [65] R. Schonmann. Theorems and conjectures on the droplet-driven relaxation of stochastic Ising models. Technical report, 1995.

- [66] B. Simon. *The statistical mechanics of lattice gases*. I. Princeton University Press, Princeton, 1993.
- [67] H.A. Simon and A. Ando. Aggregation of variables in dynamic systems. *Econometrica*, 29:111–138, 1961.
- [68] A. Sinclair. Improved bounds for mixing rates of Markov chains and multicommodity flow. *Combinatorics, Probability and Computing*, 1:351–370, 1992.
- [69] A. Sinclair and M. Jerrum. Approximate counting, uniform generation and rapidly mixing Markov chains. *Information and Computing*, 82:93–133, 1989.
- [70] A. Sinclair and M. Jerrum. Polynomial-time approximation algorithms for the Ising model. *SIAM J. Comput.*, 22:1087–1116, 1993.
- [71] A.D. Sokal. *Monte-Carlo methods in statistical mechanics: foundations and algorithms*. Cours de troisième cycle de physique en Suisse Romande, Lausanne.
- [72] A.D. Sokal and L.E. Thomas. Absence of mass gap for a class of stochastic contour models. *J. Statis. Phys.*, 51 5/6:907–947, 1988.
- [73] A. Trouve. *Parallélisation massive du recuit simulé*. PhD thesis, Université Paris 11, Janvier 1993. Thèse de doctorat.
- [74] J.N. Tsitsielis. Chains with rare transitions and simulated annealing. *Maths Oper. Res.*, 14(1):70–90, 1989.
- [75] S. Wolfram. *Mathematica*. Wolfram Research, Inc., 1992.
- [76] B. Ycart. *Systèmes markoviens*. Cours de DEA, Université Joseph Fourier, 1993.
- [77] B. Ycart, W. A. Woyczynski, J. Szulga, S. Reazor, and J. A. Mann. An interacting particle model of adsorption. *Appl. Mathematicae*, 20,3:79–94, 1989.