

Étude quantitative des chaînes de Markov par perturbation de leur noyau.

Pascal Lezaud

► **To cite this version:**

Pascal Lezaud. Étude quantitative des chaînes de Markov par perturbation de leur noyau.. Probabilités [math.PR]. Université Toulouse 3 Paul Sabatier, 1998. Français. tel-01084797

HAL Id: tel-01084797

<https://hal-enac.archives-ouvertes.fr/tel-01084797>

Submitted on 20 Nov 2014

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.

Numéro d'ordre : 2966

THÈSE

présentée en vue
de l'obtention du

DOCTORAT DE L'UNIVERSITÉ PAUL SABATIER

Spécialité : Probabilités

par

Pascal LEZAUD

ETUDE QUANTITATIVE DES CHAÎNES DE MARKOV

PAR PERTURBATION DE LEUR NOYAU

Soutenue le 26 mars 1998 devant le jury composé de

Madame et Messieurs les Professeurs :

Marie	DUFLO	Université de Marne-la-Vallée	Rapporteur
Dominique	BAKRY	Université Paul Sabatier	Examinateur
Jean	DIEBOLT	IMAG	Rapporteur
Michel	LEDOUX	Université Paul Sabatier	Examinateur
Xavier	MILHAUD	Université Paul Sabatier	Examinateur
Laurent	SALOFF-COSTE	Université Paul Sabatier	Directeur de thèse

Laboratoire de Statistique et Probabilités

UMR CNRS 5883, Université Paul Sabatier, Toulouse III

*A Béatrice et Olivier
A la mémoire de Jean-Pierre Lucas*

REMERCIEMENTS

Je tiens à exprimer ma sincère gratitude à Laurent Saloff-Coste pour la confiance qu'il m'a témoignée en acceptant d'encadrer ce travail. J'ai été tout particulièrement sensible à la limpidité de son enseignement, m'incitant à mon tour à une telle démarche de clarification. Bien que très loin d'y arriver, je pense qu'une pensée claire témoigne d'une bonne compréhension du sujet.

Je remercie également Dominique Bakry, Michel Ledoux et Xavier Milhaud pour l'aide qu'ils m'ont apportée et l'honneur qu'ils me font en participant à mon jury de thèse.

Je remercie, tout particulièrement, Marie Duflo et Jean Diebolt pour le temps qu'ils ont consacré à l'examen de ce travail, leurs nombreuses suggestions et l'honneur qu'ils me font en participant à ce jury.

Je suis également très redevable envers le Centre d'Etudes de la Navigation Aérienne, sans l'accord duquel je n'aurais pu effectuer cette thèse. Je tiens tout particulièrement à exprimer ma reconnaissance à Jean-Marc Garot, Bruno Kriner, Alain Peytavin et Michel Sabatier.

Je souhaite aussi témoigner du plaisir que j'ai eu à rencontrer Susan Holmes et Persi Diaconis. Qu'ils soient remerciés pour leur chaleureux accueil et leur intérêt pour ce travail. Ce témoignage s'adresse également à Michel Benaim et Michel Doisy pour les discussions que nous avons eues ensemble.

Merci à Béatrice et Olivier pour leur soutien et leurs encouragements.

Table des matières

Introduction	vii
partie 1. Espace d'états fini	1
Chapitre 1. Perturbation des opérateurs linéaires en dimension finie	3
1. Valeurs propres et résolvante	3
2. Points singuliers de la résolvante et décomposition canonique d'un opérateur	4
3. Opérateurs hermitiens et principe du minmax	6
4. Perturbations analytiques des valeurs propres	7
5. Compléments	14
Chapitre 2. Processus de Markov	17
1. Borne du type Chernoff	17
2. Borne inférieure dans le cas réversible	29
3. Borne du type Berry-Esséen	33
4. Compléments	45
Chapitre 3. Chaîne de Markov	47
1. Borne du type Chernoff dans le cas réversible	47
2. Borne du type Chernoff dans le cas non réversible	53
3. Borne du type Berry-Esséen	56
4. Compléments	59
partie 2. Espace d'états quelconque	61
Chapitre 4. Perturbation des opérateurs linéaires	63
1. Opérateurs fermés	63
2. Perturbation analytique	68
3. Le cas des semi-groupes	73
Chapitre 5. Rappels sur les chaînes de Markov	77
1. Point de vue probabiliste	78
2. Point de vue des opérateurs	85
3. Point de vue du mélange	90
4. Compléments	95
Chapitre 6. Inégalités dans le cas général	97
1. Chaînes de Markov	98
2. Processus markoviens	102
3. Compléments	106
Bibliographie	111

Introduction

Le développement de la modélisation des phénomènes aléatoires par chaînes de Markov pose le problème du contrôle de convergence des algorithmes de simulations. Les méthodes de simulations par chaînes de Markov ergodiques s'appuient sur la loi des grands nombres, qui stipule que pour toute mesure initiale q , la quantité

$$\frac{1}{n} \sum_{i=1}^n f(X_i)$$

converge vers la moyenne

$$\pi f := \int f(x) \pi(dx),$$

où π désigne l'unique probabilité invariante de la chaîne. Il convient alors, de déterminer un nombre suffisant de pas de simulation pour déterminer, de façon suffisamment précise, la moyenne d'une certaine fonction par la moyenne empirique. Au lieu de s'intéresser à la proximité de la distribution d'un échantillon de la chaîne à sa mesure stationnaire, l'objet de cette thèse sera de considérer la probabilité de déviation

$$(0.1) \quad P_q \left[n^{-1} \sum_{i=1}^n f(X_i) - \pi f \geq \gamma \right],$$

où $\gamma > 0$. D. Gillman dans sa thèse [26] fut le premier à déterminer une majoration de cette quantité, ceci dans le cas d'une chaîne de Markov réversible à espace d'états fini. Plus que le résultat obtenu, la méthode employée pour l'obtenir fut le point de départ des travaux présentés dans ce document. En effet, la preuve utilisée par D. Gillman s'appuie sur la théorie des perturbations des opérateurs linéaires. Dès 1957, S. V. Nagaev a employé cette théorie pour obtenir des résultats asymptotiques sur des chaînes de Markov générales vérifiant une condition de Doeblin [42, 43]. Le développement de cette théorie, dont les principaux résultats figurent dans le livre [34] de T. Kato, permet dans un premier temps d'améliorer le résultat obtenu par D. Gillman et de l'étendre aux chaînes non réversibles, puis dans un deuxième temps de considérer le cas d'un espace d'états non dénombrable. Signalons que I.H. Dinwoodie [16] a amélioré le résultat obtenu initialement par D. Gillman tout en l'étendant aux chaînes non réversibles. Pour cela, il a adapté une technique de F. Rellich, l'un des principaux acteurs du développement de la théorie des perturbations.

Les techniques mises en œuvre diffèrent des méthodes propres à la théorie des grandes déviations [13]. Cette théorie fournit une vitesse logarithmique asymptotique de convergence vers zéro de la quantité (0.1), alors que la borne obtenue par D. Gillman est valable pour tout n .

Le chapitre 1 synthétise les résultats de la théorie des perturbations des opérateurs linéaires en dimension finie ; plus particulièrement ceux présentés dans [34]. L'extension de ces résultats aux opérateurs linéaires en dimension infinie est traitée au chapitre 4, avec pour objectif l'étude de la quantité (0.1) pour une chaîne de Markov à espace d'états quelconque. Les opérateurs linéaires fermés ayant une valeur propre

isolée et dont le projecteur propre associé est de rang fini, joueront un rôle important dans le contexte général. Les opérateurs quasi-compacts font partie de cette classe d'opérateurs ; leurs liens avec les chaînes de Markov (plus particulièrement la condition de Doeblin) sont analysés au chapitre 5. Celui-ci rappelle également quelques notions sur les chaînes de Markov, en tentant un parallèle entre le point de vue probabiliste développé dans [41, 17], le point de vue des opérateurs abordé dans [52, 44] et le point de vue du mélange [18].

Le sujet principal des chapitres 2 et 3 se rapporte à la détermination d'une majoration de la vitesse de convergence de la quantité (0.1), lorsque l'espace d'états est fini. Le chapitre 2 étudie les processus markoviens, le cas des chaînes de Markov étant traité au chapitre 3. En adaptant la démonstration d'un théorème de Kolmogorov sur les variables aléatoires indépendantes [36], nous obtenons également une minoration de la vitesse de convergence de (0.1). Enfin l'étude de vitesse de convergence de la somme normalisée

$$(0.2) \quad \frac{1}{\sqrt{n}} \sum_{i=1}^n (f(X_i) - \pi f),$$

vers une loi normale centrée est abordée, via la détermination d'une borne du type Berry-Esséen. En effet, la détermination des intervalles de confiance via le théorème central limite permet d'estimer la précision des résultats d'une simulation [9]. Contrôler si la fonction de répartition de la variable aléatoire (0.2) est suffisamment proche d'une loi normale suppose la détermination d'une borne du type Berry-Esséen et la connaissance de la vitesse de convergence de sa variance vers la variance asymptotique. Dans sa thèse [39], B. Mann obtient une expression explicite, pour tout n , d'une borne du type Berry-Esséen dans le contexte d'une chaîne de Markov à espace d'états dénombrable satisfaisant la condition de mélange suivante :

$$\sup \left\{ \|Pf\|_{\pi} / \|f\|_{\pi}^2, \pi f = 0, f \neq 0 \right\} < 1.$$

La technique employée par B. Mann dérive également de la théorie des perturbations, mais n'utilise pas les outils de [34], contrairement à ce que nous proposons dans ces deux chapitres.

Enfin, dans le chapitre 6, les bornes du type Chernoff obtenues aux chapitres 2 et 3 sont étendues aux processus markoviens et aux chaînes de Markov dont l'espace d'états est quelconque. Pour cela, le générateur infinitésimal et le noyau markovien associés respectivement au processus et à la chaîne de Markov devront avoir respectivement 0 et 1 comme valeur propre simple isolée.

Les chapitres sont divisés en sections et les sections en sous-sections ; par exemple, la première sous-section de la quatrième section du premier chapitre est référencée par §4.1 dans le premier chapitre, et par 1-§4.1 dans les autres chapitres. Les théorèmes, propositions, corollaires, lemmes et équations sont numérotés à l'intérieur de chaque section. Par exemple, l'expression lemme 3-(1.1) réfère au lemme 1.1 du chapitre 3 ; il sera repéré simplement par lemme 1.1 à l'intérieur du chapitre 3.

Première partie

Espace d'états fini

Perturbation des opérateurs linéaires en dimension finie

Ce chapitre synthétise les résultats de la théorie des perturbations des opérateurs linéaires en dimension finie ; résultats qui seront utilisés dans les chapitres ultérieurs. Cette théorie, créée par L. Rayleigh et E. Schrödinger pour des questions de physique théorique, a été principalement développée par F. Rellich. Le résultat principal de F. Rellich peut se formuler de la façon suivante. Soit $T(\chi)$ un opérateur borné autoadjoint dans un espace de Hilbert H (non nécessairement de dimension finie), analytique en le paramètre réel χ . $T(\chi)$ admet donc le développement en série convergente (pour $|\chi|$ assez petit)

$$T(\chi) = T + \chi T^{(1)} + \chi^2 T^{(2)} + \dots$$

Supposons que l'opérateur non perturbé $T = T(0)$ possède une valeur propre λ isolée du reste du spectre de T et de multiplicité m . Alors, $T(\chi)$ possède exactement m valeurs propres $\mu_j(\chi), j = 1, \dots, m$ (non nécessairement distinctes) dans le voisinage de λ , pour $|\chi|$ suffisamment petit. De plus, ces valeurs propres admettent chacune un développement en série convergente

$$\mu_j(\chi) = \lambda + \chi \mu_j^{(1)} + \chi^2 \mu_j^{(2)} + \dots$$

Les résultats présentés ici ne seront pas démontrés. Ils proviennent en grande majorité de [34], où figurent les démonstrations. La méthode utilisée s'appuie sur l'étude de la résolvante comme fonction analytique, et en particulier sur l'expression des projecteurs propres comme intégrale complexe, sur un certain contour, de la résolvante. Ceci est la méthode la plus rapide, pour obtenir des résultats généraux aussi bien que pour en déduire diverses estimations sur les rayons de convergence et les coefficients des développements en série des valeurs propres perturbées.

1. Valeurs propres et résolvante

Nous ne considérerons que des opérateurs linéaires définis et à valeurs sur des espaces vectoriels complexes, normés et de dimension finie. Par conséquent, ces opérateurs ont une représentation matricielle finie et toutes les normes seront équivalentes. Cependant, la norme utilisée pour l'opérateur T sera généralement définie par

$$\|T\| = \sup_{\|u\|=1} \|Tu\|$$

La norme d'un opérateur est donc déterminée en fonction des normes adoptées sur X et Y , celles-ci étant généralement de la forme

$$\|u\|_p = \left(\sum_j |\xi_j|^p \right)^{1/p} \quad p = 1, 2, \dots$$

$$\|u\|_\infty = \max_j |\xi_j|,$$

où les ξ_j sont les coordonnées de u sur une base fixée a priori.

Soit T un endomorphisme sur un espace vectoriel X de dimension finie. Le spectre de T est l'ensemble des $\lambda \in \mathcal{C}$ tels que l'opérateur $T - \lambda I$ ne soit pas inversible ; il sera noté $\Sigma(T)$. Le complémentaire de $\Sigma(T)$ dans \mathcal{C} est appelé ensemble résolvant et noté $P(T)$. Pour $\zeta \in P(T)$, l'opérateur $R(\zeta) = (T - \zeta)^{-1}$ est appelé la résolvante de T . Les seuls points singuliers de $R(\zeta)$ sont les valeurs propres de T et $\lim_{\zeta \rightarrow \infty} R(\zeta) = 0$. Signalons également que la résolvante $R(\zeta)$ de T admet comme seules valeurs propres :

$$(\lambda_h - \zeta)^{-1}, \quad \lambda_h \in \Sigma(T).$$

La résolvante vérifie l'équation suivante, dite équation résolvante

$$(1.1) \quad R(\zeta_1) - R(\zeta_2) = (\zeta_1 - \zeta_2)R(\zeta_1)R(\zeta_2).$$

Elle implique en particulier que $R(\zeta_1)$ et $R(\zeta_2)$ commutent, et la relation :

$$R(\zeta) = [1 - (\zeta - \zeta_0)R(\zeta_0)]^{-1}R(\zeta_0).$$

La résolvante possède donc un développement en série, dite première série de Neumann pour la résolvante, de la forme suivante :

$$(1.2) \quad R(\zeta) = \sum_{n=0}^{\infty} (\zeta - \zeta_0)^n R(\zeta_0)^{n+1},$$

cette série étant absolument convergente si

$$|\zeta - \zeta_0| < \|R(\zeta_0)\|^{-1}.$$

Donc $R(\zeta)$ est analytique en ζ (au sens où chaque coordonnée de $R(\zeta)$ est analytique). Le rayon de convergence de cette série est donné par $1/\text{spr}(R(\zeta_0)) = (\lim_{n \rightarrow \infty} \|R(\zeta_0)^n\|^{1/n})^{-1}$.

Pour ζ grand, $R(\zeta)$ peut s'écrire

$$(1.3) \quad R(\zeta) = -\zeta^{-1}(1 - \zeta^{-1}T)^{-1} = -\sum_{n=0}^{\infty} \zeta^{-n-1}T^n,$$

la série convergeant si et seulement si $|\zeta| > \text{spr}(T)$, d'où l'holomorphie de la résolvante à l'infini.

Le spectre de T n'est jamais vide, il contient au moins une valeur propre, car autrement sa résolvante serait une fonction entière tendant vers zéro pour $\zeta \rightarrow \infty$, et, par le théorème de Liouville, nous aurions $R(\zeta) \equiv 0$, ce qui serait en contradiction avec $I = (T - \zeta)R(\zeta)$. En fait, chaque valeur propre est un pôle de la fonction analytique $R(\zeta)$, et $\text{spr}(T)$ coïncide avec la plus grande (en module) des valeurs propres de T .

2. Points singuliers de la résolvante et décomposition canonique d'un opérateur

Pour chaque $\lambda \in \Sigma(T)$, il existe une décomposition unique de l'espace X en somme directe

$$X = N_\lambda \oplus F_\lambda$$

telle que

- La restriction D_λ de l'opérateur $T - \lambda$ à l'espace N_λ est un opérateur nilpotent (i.e. il existe un entier $k(\lambda)$ tel que $D_\lambda^{k(\lambda)} = 0$)
- Le projecteur P_λ sur N_λ est appelé projecteur propre associé à la valeur propre λ . La dimension de $m(\lambda) = \dim(N_\lambda)$ est appelée multiplicité algébrique de λ . C'est aussi la multiplicité de λ comme racine du polynôme caractéristique $\det(T - \lambda)$. D'autre part, $k(\lambda) \leq m(\lambda)$.
- La restriction de $T - \lambda$ au sous-espace F_λ est inversible ; son inverse S_λ est appelée la résolvante réduite de T associée à la valeur propre λ .

- Le sous-espace $E_\lambda := \text{Ker}(T - \lambda)$ est inclus dans N_λ et sa dimension est appelée multiplicité géométrique de λ . Si λ est semi-simple (i.e. $D_\lambda = 0$), alors $N_\lambda = E_\lambda$.
- Si $\mu \neq \lambda$ est également une valeur propre de T , alors $N_\mu \subset F_\lambda$.

Les points singuliers de la résolvante sont exactement les valeurs propres de T , d'où le développement de la résolvante en série de Laurent autour d'une valeur propre λ ,

$$(2.1) \quad R(\zeta) = -(\zeta - \lambda)^{-1}P_\lambda - \sum_{n=1}^{\infty} (\zeta - \lambda)^{-n-1}D_\lambda^n + \sum_{n=0}^{\infty} (\zeta - \lambda)^n S_\lambda^{n+1}.$$

Notons $\lambda_h, h = 1, \dots, s$ les valeurs propres de T et, pour simplifier, les opérateurs $P_{\lambda_h}, D_{\lambda_h}, S_{\lambda_h}$ respectivement par P_h, D_h, S_h . Nous avons alors les relations suivantes :

$$(2.2) \quad P_h D_k = D_k P_h = \delta_{kh} D_h, \quad P_h S_h = S_h P_h = 0, \quad D_h D_k = 0, \quad h \neq k$$

et

$$(2.3) \quad P_h P_k = \delta_{hk} P_h, \quad \sum_{h=1}^s P_h = I, \quad P_h T = T P_h.$$

Les opérateurs P_h, D_h et S_h peuvent se calculer par une intégrale complexe sur un contour Γ_h positivement orienté, entourant λ_h et excluant toutes les autres valeurs propres :

$$\begin{aligned} P_h &= -\frac{1}{2\pi i} \int_{\Gamma_h} R(\zeta) d\zeta \\ D_h &= -\frac{1}{2\pi i} \int_{\Gamma_h} (\zeta - \lambda_h) R(\zeta) d\zeta \\ S_h &= -\frac{1}{2\pi i} \int_{\Gamma_h} (\zeta - \lambda_h)^{-1} R(\zeta) d\zeta. \end{aligned}$$

Pour tout h , λ_h étant un pôle d'ordre inférieur à m_h , la résolvante est une fonction méromorphe en ζ . Elle est régulière à l'infini, d'où sa décomposition en éléments simples [35][T.1] :

$$(2.4) \quad R(\zeta) = -\sum_{h=1}^s \left[(\zeta - \lambda_h)^{-1} P_h - \sum_{n=1}^{m_h-1} (\zeta - \lambda_h)^{-n-1} D_h^n \right].$$

Ces résultats conduisent à la forme canonique d'un opérateur T . En effet, soient M_h les sous-espaces $P_h X$, $h = 1, \dots, s$. Nous avons

$$X = M_1 \oplus \dots \oplus M_s.$$

M_h est appelé sous-espace propre algébrique pour la valeur propre λ_h de T et $m_h = \dim(M_h)$ la multiplicité algébrique de λ_h . Les relations

$$(2.5) \quad T P_h = P_h T = P_h T P_h = \lambda_h P_h + D_h, \quad h = 1, \dots, s,$$

montrent que la restriction T_{M_h} de l'opérateur T au sous-espace M_h s'exprime comme la somme de l'opérateur scalaire λ_h et de la restriction à M_h de l'opérateur nilpotent D_h . T_{M_h} a donc une et une seule valeur propre λ_h . Par addition des s équations (2.5) et d'après (2.3), nous obtenons

$$(2.6) \quad T = S + D$$

où

$$S = \sum_h \lambda_h P_h$$

$$D = \sum_h D_h$$

D est nilpotent, $D^n = \sum D_h^n = 0$ pour $n \geq \max m_h$ et (2.2) montre que D commute avec S . L'équation (2.6) exprime que tout endomorphisme T peut être décomposé en la somme d'un opérateur diagonalisable S ($\lambda_h \neq \lambda_k$ si $h \neq k$) et d'un opérateur nilpotent D qui commute avec S . Cette décomposition s'appelle la représentation spectrale de T . On démontre qu'elle est unique [34][§I.5.4].

3. Opérateurs hermitiens et principe du minmax

Les valeurs propres d'un opérateur T seront notées $\mu_1, \mu_2, \dots, \mu_N$, avec $\mu_1 = \dots = \mu_{m_1} = \lambda_1, \mu_{m_1+1} = \dots = \mu_{m_1+m_2} = \lambda_2, \dots$, ce qui revient à répéter chaque valeur propre un nombre de fois égal à leur multiplicité algébrique.

Considérons un opérateur T hermitien (c.-à-d. $T^* = T$). Il ne possède que des valeurs propres réelles, que nous classerons dans l'ordre croissant

$$(3.1) \quad \mu_1 \leq \mu_2 \leq \dots \leq \mu_N.$$

Cet opérateur est diagonalisable, plus précisément il existe une matrice unitaire U (c.-à-d. $U^* = U^{-1}$) telle que $T = U \text{diag}(\mu_i) U^*$. Une propriété importante des opérateurs hermitiens est la suivante [34][§I.6.6]

$$\|T^n\|_2 = \|T\|_2^n \quad \text{d'où} \quad \text{spr} T = \max_i |\lambda_i| = \|T\|_2.$$

La norme $\|\cdot\|_2$ est présentement utilisée, car la notion d'opérateur hermitien est liée à celle de produit scalaire et donc à la 2-norme sur X . En fait, ce résultat est vrai pour une classe plus générale d'opérateurs, à savoir les opérateurs normaux, c.-à-d. vérifiant

$$T^*T = TT^*.$$

Il est possible d'en déduire que si T est normal, alors

$$\|R(\zeta)\|_2 = 1 / \min_k |\zeta - \lambda_k|$$

Soit λ une valeur propre d'un opérateur T hermitien, et u un vecteur propre associé, alors $u^* T u = \lambda u^* u$ d'où, comme u est non nul :

$$\lambda = \frac{u^* T u}{u^* u}.$$

Le terme de droite est appelé quotient de Rayleigh.

THÉORÈME 3.1. *Soit $u \neq 0$, alors*

$$(3.2) \quad \mu_1 \leq \frac{u^* T u}{u^* u} \leq \mu_N$$

DÉMONSTRATION. Soient (ϕ_i) une base orthonormée de X formées de vecteurs propres de T et (ξ_i) les coordonnées du vecteur u sur cette base. Alors

$$\frac{u^* T u}{u^* u} = \frac{\sum_i |\xi_i|^2 \mu_i}{\|u\|_2^2},$$

et (3.1) donne le résultat. □

Appliquons le théorème aux vecteurs ϕ_1 et ϕ_n , pour obtenir l'expression de μ_1 et μ_N en termes d'extrema du facteur de Rayleigh

$$\mu_1 = \min_{u \neq 0} \frac{u^* T u}{u^* u} \leq \max_{u \neq 0} \frac{u^* T u}{u^* u} = \mu_N.$$

Une expression identique peut être trouvée pour les autres valeurs propres, en faisant varier u dans certains sous-espaces de X [34][§I.6.10].

THÉORÈME 3.2. *Soit un opérateur hermitien T de valeurs propres ordonnées $\mu_1 \leq \mu_2 \cdots \leq \mu_N$. Alors*

$$\mu_i = \min_{\dim(M)=i} \max_{\substack{u \in M \\ u \neq 0}} \frac{u^* T u}{u^* u}$$

où M est un sous-espace vectoriel de X .

Ce théorème a deux applications importantes.

THÉORÈME 3.3. *Si S et T sont des opérateurs hermitiens tels que $S \leq T$ (c.-à-d. $u^* S u \leq u^* T u$ pour tout u), alors les valeurs propres de S ne sont pas plus grandes que les valeurs propres de T de même rang, soit :*

$$\mu_i[S] \leq \mu_i[T] \quad i = 1, \dots, N.$$

où $\mu_i[T]$ représente la i -ième valeur propre de T dans l'ordre croissant.

THÉORÈME 3.4. *Soient les opérateurs hermitiens A, B et $C = A + B$, alors*

$$\mu_i[A] + \mu_1[B] \leq \mu_i[C] \leq \mu_i[A] + \mu_N[B] \quad (i = 1, \dots, N)$$

En particulier, si B est un opérateur de rang 1, alors il possède une unique valeur propre β (non nulle) et le théorème précédent montre que les valeurs propres de A sont décalées d'une quantité comprise entre zéro et β . Ce résultat ne suppose pas la perturbation B "petite".

4. Perturbations analytiques des valeurs propres

Intéressons-nous maintenant aux valeurs propres et vecteurs propres d'un opérateur linéaire T , défini sur un espace X de dimension finie, modifié par une "petite" perturbation. Pour cela nous considérerons la famille d'opérateurs $T(\chi)$ définie par

$$(4.1) \quad T(\chi) = T + \chi T^{(1)} + \chi^2 T^{(2)} + \dots$$

où χ est un nombre complexe petit (en module). De plus, nous supposerons que $T(\chi)$ est analytique sur un domaine D_0 du plan complexe. Les valeurs propres de $T(\chi)$ sont les solutions de l'équation caractéristique

$$(4.2) \quad \det(T(\chi) - \lambda) = 0,$$

équation algébrique en λ de degré $N = \dim(X)$ dont les coefficients sont des fonctions analytiques en χ .

Selon la théorie des fonctions algébriques [35][T.2 p.121], les racines de (4.2) forment alors des branches de fonctions analytiques. Comme par la suite, nous n'appliquerons cette théorie qu'au cas d'une valeur propre de T simple, nous ne considérerons que ce cas. Ainsi, pour une perturbation suffisamment petite, la valeur propre perturbée sera une fonction analytique de χ . En effet, nous avons le théorème suivant :

THÉORÈME 4.1. *Soit λ une valeur propre simple de T . Alors il existe $\delta_\lambda > 0$ tel que pour tout $|\chi| < \delta_\lambda$, $T(\chi)$ admette une valeur propre simple $\lambda(\chi)$ définie par la série convergente*

$$\lambda(\chi) = \lambda + \chi \lambda^{(1)} + \chi^2 \lambda^{(2)} + \dots$$

La fonction $\chi \rightarrow \lambda(\chi)$ est une fonction analytique dans le domaine $|\chi| < \delta_\lambda$.

DÉMONSTRATION. $\lambda(\chi)$ est une racine du polynôme caractéristique

$$\det(T(\chi) - \zeta I) = (-1)^N \zeta^N + p_{N-1}(\chi) \zeta^{N-1} + \cdots + p_0(\chi) = 0,$$

où les $p_{N-i}(\chi)$ sont des fonctions analytiques en χ . L'application Φ de $\mathcal{C} \times \mathcal{C} \rightarrow \mathcal{C}$ qui à (χ, ζ) associe $\det(T(\chi) - \zeta I)$ est analytique. D'autre part, $\Phi(0, \lambda) = 0$ et λ est une valeur propre simple de T d'où $\Phi'_\zeta(0, \lambda) \neq 0$. Le théorème des fonctions implicites garantit l'existence d'un voisinage de zéro et d'une unique fonction analytique $\lambda(\cdot)$, définie sur ce voisinage, telle que $\lambda(0) = \lambda$ et $\Phi(\chi, \lambda(\chi)) = 0$. Il suffit alors de choisir δ_λ , tel que le disque ouvert de centre 0 et de rayon δ_λ soit inclus dans le voisinage de définition de $\lambda(\cdot)$. \square

4.1. Perturbation de la résolvante et des projecteurs propres. La résolvante

$$(4.3) \quad R(\zeta, \chi) = (T(\chi) - \zeta)^{-1}$$

est définie pour tout ζ différent d'une valeur propre de $T(\chi)$, et est méromorphe en ζ pour tout χ fixé dans D_0 .

THÉORÈME 4.2. $R(\zeta, \chi)$ est une fonction analytique des deux variables ζ et χ dans chaque domaine où ζ est différent d'une valeur propre de $T(\chi)$.

Nous pouvons alors développer la résolvante en double série entière, que nous écrirons sous la forme suivante, dite seconde série de Neumann pour la résolvante :

$$(4.4) \quad \begin{aligned} R(\zeta, \chi) &= R(\zeta)[1 + A(\chi)R(\zeta)]^{-1} \\ &= R(\zeta) \sum_{p=0}^{\infty} [-A(\chi)R(\zeta)]^p \\ &= R(\zeta) + \sum_{p=1}^{\infty} \chi^p R^{(p)}(\zeta), \end{aligned}$$

où $A(\chi) := T(\chi) - T = \sum_{n=1}^{\infty} \chi^n T^{(n)}$.

Cette série est convergente si $\|A(\chi)R(\zeta)\| < 1$. Ceci sera vérifié si en particulier

$$(4.5) \quad \sum_{n=1}^{\infty} |\chi|^n \|T^{(n)}R(\zeta)\| < 1.$$

Soit $r(\zeta)$ la valeur de $|\chi|$ telle que le membre de gauche de (4.5) soit égal à 1. Alors l'inégalité (4.5) est satisfaite dès que $|\chi| < r(\zeta)$.

Soient λ une valeur propre simple de T et Γ un cercle centré sur λ , inclus dans l'ensemble résolvant $P(T)$ et ne contenant aucune autre valeur propre de T . La série (4.5) converge uniformément pour $|\chi| < r(\zeta)$ et $\zeta \in \Gamma$. L'existence de la résolvante $R(\zeta, \chi)$ de $T(\chi)$ implique qu'il n'y a pas de valeur propre de $T(\chi)$ sur Γ . L'opérateur

$$(4.6) \quad P(\chi) = -\frac{1}{2\pi i} \int_{\Gamma} R(\zeta, \chi) d\zeta$$

est un projecteur égal à la somme des projecteurs propres associés à toutes les valeurs propres de $T(\chi)$ situées dans le domaine délimité par Γ . En particulier, $P(0)$ coïncide avec le projecteur propre de T associé à λ . En intégrant terme-à-terme (4.4), nous obtenons

$$(4.7) \quad P(\chi) = P + \sum_{n=1}^{\infty} \chi^n P^{(n)},$$

où les coefficients $P^{(n)}$ s'expriment par

$$(4.8) \quad \begin{aligned} P^{(n)} &= -\frac{1}{2\pi i} \int_{\Gamma} R^{(n)}(\zeta) d\zeta \\ &= -\sum_{p=1}^n (-1)^p \sum_{\substack{\nu_1 + \dots + \nu_p = n \\ k_1 + \dots + k_{p+1} = p \\ \nu_j \geq 1, k_j \geq 0}} S^{(k_1)} T^{(\nu_1)} S^{(k_2)} \dots S^{(k_p)} T^{(\nu_p)} S^{(k_{p+1})}, \end{aligned}$$

avec

$$S^{(0)} = -P, \quad S^{(n)} = S^n, \quad n \geq 1,$$

S désignant la valeur en $\zeta = \lambda$ de la résolvante réduite de T . Par exemple,

$$(4.9) \quad \begin{aligned} P^{(1)} &= -PT^{(1)}S - ST^{(1)}P, \\ P^{(2)} &= -PT^{(2)}S - ST^{(2)}P + PT^{(1)}ST^{(1)}S + ST^{(1)}PT^{(1)}S + \\ &\quad + ST^{(1)}ST^{(1)}P - PT^{(1)}PT^{(1)}S^2 - PT^{(1)}S^2T^{(1)}P - S^2T^{(1)}PT^{(1)}P. \end{aligned}$$

La série (4.7) est uniformément convergente pour tout $\zeta \in \Gamma$ si

$$(4.10) \quad |\chi| < r_0 = \min_{\zeta \in \Gamma} r(\zeta).$$

Le résultat important démontrée dans [34][p.68] est que pour $|\chi|$ suffisamment petit,

$$\dim(\text{Im}P(\chi)) = \dim(\text{Im}P(0)) = 1,$$

ce qui signifie que seule la valeur propre de $T(\chi)$, située à l'intérieur de Γ , est la valeur propre perturbée $\lambda(\chi)$ issue de λ . $P(\chi)$ est donc le projecteur propre de $T(\chi)$ pour la valeur propre $\lambda(\chi)$.

4.2. Développement en série des valeurs propres perturbées. Soient λ , une valeur propre simple de l'opérateur non perturbé $T = T(0)$ et P le projecteur propre associé (l'opérateur nilpotent D est nul par hypothèse). Les relations (4.6) et

$$(T(\chi) - \lambda)R(\zeta, \chi) = 1 + (\zeta - \lambda)R(\zeta, \chi)$$

conduisent à l'expression suivante :

$$(T(\chi) - \lambda)P(\chi) = -\frac{1}{2\pi i} \int_{\Gamma} (\zeta - \lambda)R(\zeta, \chi) d\zeta.$$

En remplaçant la résolvante dans l'intégrale par son développement en série (4.4), et comme $(T - \lambda)P = 0$, nous obtenons le développement en série suivant :

$$(4.11) \quad (T(\chi) - \lambda)P(\chi) = + \sum_{n=1}^{\infty} \chi^n \tilde{T}^{(n)},$$

où

$$\tilde{T}^{(n)} = -\sum_{p=1}^n (-1)^p \sum_{\substack{\nu_1 + \dots + \nu_p = n \\ k_1 + \dots + k_{p+1} = p-1 \\ \nu_j \geq 1, k_j \geq 0}} S^{(k_1)} T^{(\nu_1)} S^{(k_2)} \dots S^{(k_p)} T^{(\nu_p)} S^{(k_{p+1})}.$$

Par exemple,

$$\tilde{T}^{(1)} = PT^{(1)}P,$$

$$\tilde{T}^{(2)} = PT^{(2)}P - PT^{(1)}PT^{(1)}S - PT^{(1)}ST^{(1)}P - ST^{(1)}PT^{(1)}P.$$

De la série (4.11), nous en déduisons le développement en série de $\lambda(\chi)$,

$$(4.12) \quad \lambda(\chi) = \lambda + \sum_{n=1}^{\infty} \chi^n \lambda^{(n)},$$

où

$$(4.13) \quad \begin{aligned} \lambda^{(n)} &= \operatorname{tr}(\tilde{T}^{(n)}) \\ &= \sum_{p=1}^n \frac{(-1)^p}{p} \sum_{\substack{\nu_1 + \dots + \nu_p = n \\ k_1 + \dots + k_p = p-1 \\ \nu_j \geq 1, k_j \geq 0}} \operatorname{tr}(T^{(\nu_1)} S^{(k_1)} \dots T^{(\nu_p)} S^{(k_p)}). \end{aligned}$$

Le passage de la première identité à la seconde n'est pas immédiat, il utilise l'expression suivante de $\lambda(\chi) - \lambda$:

$$\lambda(\chi) - \lambda = -\frac{1}{2i\pi} \operatorname{tr} \int_{\Gamma} \log \left[1 + \left(\sum_{n=1}^{\infty} \chi^n T^{(n)} \right) R(\zeta) \right] d\zeta.$$

Pour la preuve, se reporter à [34][§II.2]. En outre, la relation $\operatorname{tr}(AB) = \operatorname{tr}(BA)$ permet de simplifier les expressions obtenues. Par exemple,

$$\begin{aligned} \lambda^{(1)} &= \operatorname{tr}(T^{(1)} P) \\ \lambda^{(2)} &= \operatorname{tr}[T^{(2)} P - T^{(1)} S T^{(1)} P] \end{aligned}$$

La série (4.13) converge si la condition (4.10) est vérifiée. Pour estimer les coefficients $\lambda^{(n)}$, nous utilisons le fait que, dès que (4.10) est vraie, la valeur propre $\lambda(\chi)$ se situe dans le domaine convexe délimité par la courbe Γ . En posant

$$\rho = \max_{\zeta \in \Gamma} |\zeta - \lambda|,$$

nous voyons que $\lambda(\chi) - \lambda$ est analytique et bornée par ρ . Il s'ensuit, par application des inégalités de Cauchy sur les coefficients du développement en série de Taylor, que

$$(4.14) \quad |\lambda^{(n)}| \leq \rho r_0^{-n} \quad n = 1, 2, \dots$$

REMARQUE 4.3. Si $T(\chi)$ est linéaire, c.-à-d. si $T^{(n)} = 0$ pour $n \geq 2$, alors

$$\lambda^{(n)} = \operatorname{tr}(T^{(1)} P^{(n-1)}).$$

4.3. Méthode de la série majorante. Une autre méthode d'estimation des coefficients et du rayon de convergence du développement en série de $\lambda(\chi) - \lambda$ consiste à introduire une fonction $\Phi(\zeta - \lambda, \chi)$, dont chaque coefficient $c_{k,n}$ du développement en série double

$$\Phi(\zeta - \lambda, \chi) = \sum_{k=-\infty}^{+\infty} \sum_{n=1}^{\infty} c_{k,n} (\zeta - \lambda)^k \chi^n$$

majoré les coefficients respectifs du développement en série de l'opérateur

$$A(\chi)R(\zeta) = \sum_{n=1}^{\infty} \chi^n T^{(n)} R(\zeta),$$

utilisé dans l'étude des valeurs propres perturbées. Les inégalités désirées s'obtiennent à partir du développement (2.1) de $R(\zeta)$. Elles s'expriment de la façon suivante, où nous avons noté λ_h, P_h, D_h, S_h respectivement λ, P, D, S :

$$\begin{aligned} \|T^{(n)} D^k\| &\leq c_{-k-1,n}, & \|T^{(n)} P\| &\leq c_{-1,n}, \\ \|T^{(n)} S^k\| &\leq c_{k-1,n}, & k &> 0. \end{aligned}$$

La construction de cette fonction Φ nous fournit directement la majoration

$$\left\| \sum_{n=1}^{\infty} \chi^n T^{(n)} R(\zeta) \right\| \leq \Phi(|\zeta - \lambda|, |\chi|).$$

Ainsi, la série (4.4) sera convergente si $\Phi(|\zeta - \lambda|, |\chi|) < 1$. Nous appellerons cette fonction Φ , fonction ou série majorante de l'opérateur $A(\chi)R(\zeta)$ [34][p. 89] et nous noterons

$$A(\chi)R(\zeta) \ll \Phi(\zeta - \lambda, \chi).$$

Nous supposons dans la suite que $c_{k,n} = 0$ pour $k < -m$, m étant la multiplicité de la valeur propre λ . Ainsi, $\Phi(z, \chi)$ a un seul pôle en $z = 0$; ceci est toujours possible du fait que $D^m = 0$. Considérons maintenant une valeur propre λ simple et choisissons comme courbe Γ , le cercle $|\zeta - \lambda| = \rho$. Il est alors possible d'obtenir une borne inférieure r du rayon de convergence des développements en série de $P(\chi)$ et de $\lambda(\chi)$ en calculant la plus petite racine positive de l'équation

$$\Phi(\rho, r) = 1.$$

Il conviendra de choisir ρ pour obtenir la plus grande valeur possible de r .

A partir de la série majorante Φ , il est possible de construire une série majorante pour $\lambda(\chi) - \lambda$, permettant ainsi d'obtenir une majoration des coefficients $\lambda^{(n)}$. Une telle série est donnée par [34][p. 90]

$$(4.15) \quad \Psi(\chi) = \frac{1}{2\pi i} \int_{|\zeta - \lambda| = \rho} -\log(1 - \Phi(\zeta - \lambda, \chi)) d\zeta.$$

La complexité apparente de cette formule se simplifie grâce à un théorème d'analyse complexe exprimant (4.15) comme la somme des zéros moins la somme des pôles de la fonction $z \mapsto 1 - \Phi(z, \chi)$ contenus dans le domaine délimité par le cercle $|z| = \rho$. Mais le seul pôle de cette fonction est en $z = 0$, et il ne contribue pas à la valeur de l'expression.

A titre d'exemple, calculons les fonctions Φ et Ψ dans le cas particulier d'une valeur propre simple, en s'inspirant de la démarche employée dans l'exemple 3.3 de [34][p. 90]. Pour cela nous supposons que

$$(4.16) \quad \|T^{(n)}\| \leq ac^{n-1}, \quad n = 1, 2, \dots;$$

ces constantes existent toujours en raison de la condition de convergence de (4.1). D'autre part, pour tout α réel, $S^{k+1} = S(S - \alpha P)^k$, car $SP = PS = 0$. Ainsi, nous pouvons poser

$$\begin{aligned} c_{-1,n} &= ac^{n-1} \|P\| := ac^{n-1} p, \\ c_{k,n} &= ac^{n-1} \|S\| \|S - \alpha P\|^k := ac^{n-1} q s^k, \end{aligned}$$

les autres coefficients étant nuls du fait que $D = 0$. Par un calcul simple, il vient

$$\begin{aligned} \Phi(z, \chi) &= \frac{a\chi}{1 - c\chi} \left(\frac{p}{z} + \frac{q}{1 - zs} \right), \\ \Psi(\chi) &= \frac{ap\chi}{1 - c\chi} + \frac{1}{2s(1 - c\chi)} [(1 - c\chi) - a\chi(q + ps) - \Omega(\chi)] \\ &= ap\chi + \frac{ap\chi^2}{1 - c\chi} \left(c + \frac{2aq}{(1 - c\chi) - a\chi(q + ps) + \Omega(\chi)} \right), \end{aligned}$$

où

$$\Omega(\chi) = \{[(1 - c\chi) - a\chi(q + ps)]^2 - 4a^2\chi^2 pqs\}^{1/2}.$$

Pour obtenir le rayon de convergence, remarquons que $\Psi(\chi)$ a un développement en série convergent si $\Omega(\chi) > 0$. r est donc solution de l'équation $\Omega(r) = 0$ et ρ se calcule en résolvant $\Phi(\rho, r) = 1$. Ceci conduit aux expressions suivantes :

$$(4.17) \quad r = [c + a(\sqrt{ps} + \sqrt{q})^2]^{-1}$$

$$(4.18) \quad \rho = p^{1/2}s^{-1/2}[\sqrt{ps} + \sqrt{q}]^{-1}.$$

Notons que $\rho \leq s^{-1} = \|S - \alpha P\|^{-1} \leq d$, où d est la distance de λ aux autres valeurs propres λ_k de T . L'expression obtenue pour $\Psi(\chi)$ permet en particulier d'obtenir la majoration suivante :

$$\begin{aligned} |\lambda(\chi) - \lambda - \chi\lambda^{(1)}| &\leq \frac{ap\chi^2}{1 - c\chi} \left(c + \frac{2aq}{(1 - c\chi) - a\chi(q + ps) + \Omega(\chi)} \right) \\ &\leq \frac{ap\chi^2}{1 - cr} \left(c + \frac{q^{1/2}}{r(ps)^{1/2}} \right). \end{aligned}$$

Dans le cas particulier d'une perturbation linéaire, les formules figurant au §II.3 de [34] se retrouvent facilement à partir des formules ci-dessus. Pour cela, il suffit dans un premier temps de poser $c = 0$, puis dans un second temps de remplacer p et q par

$$p := \|T^{(1)}P\|, \quad q = \|T^{(1)}S\|,$$

ce qui implique que $a = 1$.

4.4. Estimation des vecteurs propres. Nous aurons besoin dans les chapitres ultérieurs d'estimer un vecteur propre de $T(\chi)$ associé à la valeur propre simple $\lambda(\chi)$. Comme les vecteurs propres ne sont pas uniquement déterminés, il ne sera pas possible d'obtenir une formule générale en fonction de χ . Dans le cas $m = 1$, le choix le plus adéquat du vecteur propre $\varphi(\chi)$ de $T(\chi)$, associé à la valeur propre $\lambda(\chi)$, est donné par

$$(4.19) \quad \varphi(\chi) = (P(\chi)\varphi, \psi)^{-1} P(\chi)\varphi,$$

où φ est le vecteur propre de T pour la valeur propre λ et ψ est le vecteur propre de T^* pour la valeur propre $\bar{\lambda}$ normalisé par $(\varphi, \psi) = 1$. Ainsi,

$$P\varphi = 1, \quad P^*\psi = 1, \quad (\varphi, \psi) = 1.$$

Le choix de la constante de normalisation dans (4.19) est équivalent à chacune des conditions de normalisation suivante :

$$(4.20) \quad (\varphi(\chi), \psi) = 1, \quad (\varphi(\chi) - \varphi, \psi) = 0, \quad P(\varphi(\chi) - \varphi) = 0.$$

La relation $(T(\chi) - \lambda(\chi))\varphi(\chi) = 0$ peut s'écrire

$$(4.21) \quad (T - \lambda)(\varphi(\chi) - \varphi) + (A(\chi) - \lambda(\chi) + \lambda)\varphi(\chi) = 0,$$

où $A(\chi) = T(\chi) - T$. Multiplions (4.21) à gauche par S et utilisons les relations $S(T - \lambda) = I$ et (4.20), nous obtenons

$$(4.22) \quad \varphi(\chi) - \varphi + S[A(\chi) - \lambda(\chi) + \lambda]\varphi(\chi) = 0.$$

Notons en outre que $S\varphi = 0$, d'où en écrivant $\varphi(\chi) = \varphi(\chi) - \varphi + \varphi$ dans le dernier terme de (4.22), nous obtenons

$$\begin{aligned} \varphi(\chi) - \varphi &= -(I + S[A(\chi) - \lambda(\chi) + \lambda])^{-1} SA(\chi)\varphi \\ (4.23) \quad &= -S[I + A(\chi)S - (\lambda(\chi) - \lambda)S]^{-1} A(\chi)\varphi. \end{aligned}$$

Une série majorante pour $\varphi(\chi) - \varphi$ se déduit aisément de (4.23) :

$$(4.24) \quad \varphi(\chi) - \varphi \ll \|S\|(1 - \Phi_2(\chi) - \|S\|\Psi(\chi))^{-1}\Phi_3(\chi),$$

où $\Phi_2(\chi)$ et $\Phi_3(\chi)$ sont respectivement des séries majorantes pour $A(\chi)S$ et $A(\chi)\varphi$, alors que $\Psi(\chi)$ est une série majorante pour $\lambda(\chi) - \lambda$.

Le cas d'une perturbation linéaire est traité en détail dans [34][§II.3]. Il suffit de prendre

$$\Phi_2(\chi) = \chi \|T^{(1)}S\| := \chi q, \quad \Phi_3(\chi) = \chi \|T^{(1)}\varphi\|.$$

Nous obtenons ainsi

$$(4.25) \quad \varphi(\chi) - \varphi \ll \frac{2\chi s \|T^{(1)}\varphi\|}{1 - (ps + q)\chi + \{(1 - (ps + q)\chi)^2 - 4pqs\chi^2\}^{1/2}}$$

où

$$s := \|S\|, \quad p = \|T^{(1)}P\|, \quad q = \|T^{(1)}S\|.$$

Pour $|\chi| < r$, où r est donné par (4.17) (avec $a = 1$ et $c = 0$), nous en déduisons la majoration suivante

$$(4.26) \quad \|\varphi(\chi) - \varphi\| \leq |\chi| \frac{s}{(pqs)^{1/2}} \left((ps)^{1/2} + q^{1/2} \right)^2 \|T^{(1)}\varphi\|.$$

4.5. Le cas d'un opérateur normal. Supposons dans cette partie que X est un espace de Hilbert et que l'opérateur T est normal. Alors la résolvante $R(\zeta)$ est également normale, d'où, pour tout $\zeta \in P(T)$,

$$(4.27) \quad \|R(\zeta)\| = 1/\text{dist}(\zeta, \Sigma(T)).$$

Supposons également que les $T^{(n)}$ satisfassent les inégalités (4.16), alors (4.5) est satisfait si $|\chi| < (a\|R(\zeta)\| + c)^{-1}$. Ainsi nous pouvons prendre

$$r_0 = \min_{\zeta \in \Gamma} (a\|R(\zeta)\| + c)^{-1}.$$

Si nous choisissons pour Γ , le cercle centré en λ de rayon $d/2$, où d est la distance de λ au reste du spectre de T , alors nous obtenons d'après (4.27),

$$(4.28) \quad r_0 = \left(\frac{2a}{d} + c \right)^{-1}.$$

D'autre part, la résolvante réduite S de T pour la valeur propre (simple) λ est également un opérateur normal et $\|S\| = 1/d$. Les séries majorantes obtenues précédemment se simplifient dans le cas considéré. En effet, en tenant compte du fait que $\|P\| = 1$, il suffit de choisir

$$p = \|T^{(1)}\| = a, \quad q = a/d, \quad s = 1/d.$$

La série majorante (4.25) devient

$$(4.29) \quad \begin{aligned} \varphi(\chi) - \varphi &\ll \frac{2\chi \|T^{(1)}\varphi\|/d}{1 - \frac{2a\chi}{d} + \left(1 - \frac{4a\chi}{d}\right)^{1/2}} \\ &= \frac{\|T^{(1)}\varphi\|}{d} \sum_{n=1}^{\infty} (-1)^n \binom{1/2}{n+1} 2^{2n+1} \left(\frac{a}{d}\right)^{n-1} \chi^n. \end{aligned}$$

Nous pouvons obtenir une majoration de la norme des coefficients du développement en série $\varphi(\chi) - \varphi = \sum \chi^n \varphi^{(n)}$, à savoir :

$$\begin{aligned} \|\varphi^{(1)}\| &\leq \|T^{(1)}\varphi\|/d, & \|\varphi^{(2)}\| &\leq 2\|T^{(1)}\varphi\|a/d^2, \\ \|\varphi^{(3)}\| &\leq 5\|T^{(1)}\varphi\|a^2/d^3, & \|\varphi^{(4)}\| &\leq 14\|T^{(1)}\varphi\|a^3/d^4. \end{aligned}$$

L'inégalité

$$\binom{1/2}{n+1} \leq \frac{1}{4(n+1)},$$

permet d'obtenir une formule générale :

$$(4.30) \quad \|\varphi^{(n)}\| \leq \frac{2\|T^{(1)}\varphi\|}{(n+1)d} \left(\frac{4a}{d}\right)^{n-1}.$$

5. Compléments

Lorsque la multiplicité m de la valeur propre λ de T est supérieure à 1, une perturbation analytique de T , aussi petite soit-elle, peut conduire à une séparation de λ en m valeurs propres perturbées distinctes. Cependant, ces valeurs propres perturbées issues de λ constituent des branches d'une ou de plusieurs fonctions analytiques multiformes. En effet, les valeurs propres perturbées sont les racines du polynôme caractéristique de $T(\chi)$; polynôme dont les coefficients sont des fonctions de χ analytiques dans un domaine D_0 . Supposons dans un premier temps que ces coefficients soient des polynômes en χ (ou des fonctions méromorphes). Nous sommes alors conduits à la théorie des fonctions algébriques [35], qui stipule que les zéros d'un polynôme irréductible de degré n constant,

$$P(\chi, z) = a_0(\chi)z^n + a_1(\chi)z^{n-1} + \cdots + a_n(\chi)$$

forment une fonction analytique multiforme. Plus précisément, le polynôme étant irréductible, son discriminant n'est pas identiquement nul ; par conséquent, dans tout compact, il existe un nombre fini de points, dénommés critiques, annulant ce discriminant. En dehors de ces points, $P(\chi, z)$ admet n zéros distincts, fonctions analytiques en χ . Le comportement des zéros autour d'un point critique est plus compliqué. La théorie des fonctions algébriques construit une surface de Riemann compacte (en considérant le point critique ∞) sur laquelle les zéros constituent une fonction analytique. Le nombre de feuillets au-dessus d'un point non critique est constant et égal à n [24]. Au-dessus d'un point critique χ_0 , les zéros se séparent en l ($1 \leq l \leq n$) cycles centrés en $\lambda(\chi_0)$, le nombre d'éléments de chaque cycle étant appelé période du cycle. La surface de Riemann correspondante présente alors l feuillets au-dessus de ce point critique. Le point à l'infini est étudié par le changement de variable z en $1/z$, ou grâce à la formule de Riemann-Hurwitz [24][p. 140]. D'autre part, au voisinage de ce point, chaque zéro appartenant à un cycle de période p admet un développement en série de Puiseux

$$\lambda_h(\chi) = \lambda(\chi_0) + \alpha_1 \omega^h (\chi - \chi_0)^{1/p} + \alpha_2 \omega^{2h} (\chi - \chi_0)^{2/p} + \cdots, \quad h = 1, \dots, p-1,$$

où $\omega = \exp(2\pi i/p)$.

Revenons maintenant au cas d'une perturbation analytique de T , le polynôme caractéristique admettant pour coefficients, des fonctions analytiques dans un domaine D_0 . Les zéros du polynôme caractéristique ne constituent plus une fonction algébrique, mais une fonction dite algébroïdale, dont les propriétés sont localement similaires à celles d'une fonction algébrique [4]. Enfin, le polynôme caractéristique n'étant pas nécessairement irréductible, nous aurons en fait plusieurs fonctions algébroïdales, i.e. les valeurs propres de $T(\chi)$ constituent une ou plusieurs branches d'une ou plusieurs fonctions analytiques multiformes, n'ayant que des singularités algébriques dans D_0 .

Obtenir une estimation d'une valeur propre multiple perturbée comme dans le cas d'une valeur propre simple n'est plus possible. Néanmoins la même démarche permet d'obtenir une estimation de la moyenne pondérée $m^{-1}\text{Tr}(T(\chi)P(\chi))$ des valeurs propres issues d'une même valeur propre λ de T de multiplicité m [34].

Les perturbations étudiées dans [34] sont du type analytique ; elles permettent d'obtenir des développements en série. J. H. Wilkinson, dans [51], traite des perturbations du type $A = B + C$, où B est la matrice non perturbée et C la matrice

perturbatrice ; les estimations sont alors moins précises sauf, lorsque les matrices présentent certaines caractéristiques.

Processus de Markov

Considérons un processus de Markov irréductible, d'espace d'états G fini, dont l'unique probabilité invariante sera notée π . Soit une fonction réelle f définie sur G ; la loi faible des grands nombres stipule que pour toute mesure initiale q , $t^{-1} \int_0^t f(X_s) ds$ converge p.s. vers la moyenne de $\pi f = \sum_x \pi(x) f(x)$. Ce résultat est à la base des méthodes de simulation en vue de l'évaluation de la moyenne de la fonction f . L'objectif de ce chapitre sera de quantifier la vitesse de convergence de ce terme vers la moyenne de f , en établissant une majoration exponentielle de la quantité

$$P_q \left[\frac{1}{t} \int_0^t f(X_s) ds \geq \gamma \right],$$

où P_q désigne la mesure du processus de mesure initiale q , et la déviation γ est en général un nombre assez petit, de l'ordre de $\pi f/10$, ou $\pi f/100$.

Ce travail est motivé par l'éventuelle utilisation de telles bornes pour le contrôle de convergence des simulations. Cette utilisation sera d'autant plus utile que les termes figurant dans ces bornes seront précis et relativement faciles à calculer. La méthode utilisée, pour atteindre ce but, s'appuie sur la théorie des perturbations des opérateurs linéaires. Celle-ci fut employée pour la première fois par Nagaev [42, 43] dans l'étude de certains théorèmes limites pour les chaînes de Markov vérifiant une condition de Doeblin. Nous avons suivi pour notre part, l'approche utilisée par D. Gillman dans [26]; approche basée sur la théorie développée par Kato dans [34]. Signalons que I. H. Dinwoodie [16] a amélioré le résultat obtenu initialement par D. Gillman, mais en adaptant une technique de Rellich. A la différence des travaux précédents, l'utilisation des outils figurant dans [34] permet d'obtenir des bornes suffisamment précises pour faire ressortir un comportement Gaussien, pour les petites valeurs de γ , et un comportement Poissonien, pour les grandes valeurs de γ .

L'utilisation des perturbations s'avère fructueuse pour ce type de problème, puisque par une démarche adaptée de celle de I. H. Dinwoodie [16], B. Mann [39] a obtenue une borne du type Berry-Esséen pour les chaînes de Markov à espace d'états dénombrables. Ce type de borne sera également étudié dans ce chapitre à partir de la démarche développée dans [34]. Le choix délibéré d'aborder l'étude des processus markoviens, préalablement à celle des chaînes de Markov, tient à la plus grande simplicité de la démarche, due à l'absence de périodicité du processus.

1. Borne du type Chernoff

Conformément à la terminologie employée dans [41], nous désignerons par chaîne de Markov une suite de variables aléatoires $X = \{X_n : n \in T\}$, satisfaisant à la propriété de Markov, et telle que T soit dénombrable, en particulier $T = \mathbb{N}$. Le terme de processus markovien sera réservé au cas où T est non dénombrable, par exemple lorsque $T = \mathbb{R}$. Soit $(X_t, t \in \mathbb{R}_+)$ un processus de Markov d'espace d'états G fini de cardinal N . Nous supposons ce processus irréductible et désignerons le semi-groupe associé par $P_t = \exp(-\Lambda t)$. Le processus de Markov, étant irréductible, admet une unique probabilité invariante π ; il sera alors naturel de travailler dans

l'espace de Hilbert $\ell^2(\pi)$ muni du produit scalaire

$$\langle f, g \rangle = \sum_{x \in G} f(x)g(x)\pi(x).$$

L'opérateur associé à P_t agit sur $\ell^2(\pi)$ de la façon suivante

$$P_t f(x) := \sum_{y \in G} P_t(x, y) f(y).$$

Cet opérateur admet 1 comme valeur propre simple et le projecteur propre associé est l'opérateur π défini par

$$\pi f = \sum_x \pi(x) f(x) = E_\pi(f).$$

π est également le projecteur propre de Λ associé à la valeur propre nulle. Nous désignerons par Λ^* l'adjoint de Λ dans $\ell^2(\pi)$.

Soit $f : G \rightarrow \mathbb{R}$ une fonction de moyenne nulle, telle que $\sup_x |f(x)| \leq 1$ et $\|f\|^2 \leq b^2$, où b est une constante positive inférieure à 1. Nous allons démontrer le théorème suivant où la constante N_q désignera la norme dans $\ell^2(\pi)$ de la densité de la mesure initiale q par rapport à la probabilité invariante π , soit $N_q = \|q/\pi\|$:

THÉORÈME 1.1. *Soient (P_t, π) , un noyau markovien irréductible défini sur un espace d'états fini et $-\Lambda$ son générateur infinitésimal. Soit $f : G \rightarrow \mathbb{R}$ telle que $\pi f = 0$, $\sup_x |f(x)| \leq 1$ et $0 < \|f\|^2 \leq b^2$. Alors pour toute distribution initiale q , tout $t > 0$ et tout $0 \leq \gamma \leq 1$,*

$$(1.1) \quad P_q \left[\frac{1}{t} \int_0^t f(X_s) ds \geq \gamma \right] \leq N_q \exp \left[-\frac{\gamma^2 \lambda_1 t}{4b^2(1 + h(4\gamma/b^2))} \right],$$

où λ_1 est la plus petite valeur propre non nulle de $(\Lambda + \Lambda^*)/2$ et

$$h(x) = \frac{1}{2}(\sqrt{1+x} - (1-x/2)).$$

En particulier si $\gamma \leq b^2/2$,

$$P_q \left[\frac{1}{t} \int_0^t f(X_s) ds \geq \gamma \right] \leq N_q \exp \left[-\frac{\gamma^2 \lambda_1 t}{4b^2} \left(1 - \frac{2\gamma}{b^2} \right) \right],$$

et si $\gamma > b^2/2$,

$$P_q \left[\frac{1}{t} \int_0^t f(X_s) ds \geq \gamma \right] \leq N_q \exp \left[-\frac{\gamma \lambda_1 t}{8} \left(1 - \frac{b^2}{2\gamma} \right) \right].$$

D'autre part pour $\gamma \ll b^2$, nous obtenons une borne asymptotique de l'ordre de

$$N_q \exp \left[-\gamma^2 \lambda_1 t / (4b^2) \right].$$

REMARQUE 1.2. Pour $\gamma > 1$, la probabilité de déviation est nulle, ce qui permet de se restreindre au cas $0 \leq \gamma \leq 1$. D'autre part le théorème 1.1 appliqué à la fonction $-f$ conduit aux mêmes conclusions pour l'expression $P_q[t^{-1} \int_0^t f(X_s) ds \leq -\gamma]$. Ainsi $P_q[t^{-1} |\int_0^t f(X_s) ds| \geq \gamma]$ est borné par deux fois le terme de droite de (1.1).

REMARQUE 1.3. Nous pouvons évidemment remplacer le terme b^2 par 1 dans l'inégalité (1.1). Nous obtenons alors une majoration indépendante de la variance de f à savoir

$$P_q \left[\frac{1}{t} \int_0^t f(X_s) ds \geq \gamma \right] \leq N_q \exp \left[-\gamma^2 \lambda_1 t / 11 \right].$$

La démonstration du théorème 1.1 s'appuiera sur plusieurs lemmes dont voici le premier

LEMME 1.4. *Pour tout $r > 0$*

$$(1.2) \quad P_q \left[\frac{1}{t} \int_0^t f(X_s) ds > \gamma \right] \leq e^{-rt\gamma} \langle q/\pi, \exp(-\Lambda(r)t) \cdot \mathbf{1} \rangle,$$

où $\Lambda(r) = \Lambda - rD$, D désignant l'opérateur multiplicatif défini par

$$Dg(x) = f(x)g(x).$$

DÉMONSTRATION. Le processus étant à espace d'états fini, nous pouvons le supposer à trajectoires càdlàg. Par conséquent $f(X_s)$ est continue par morceaux. Ceci permet d'exprimer l'intégrale, figurant dans le terme de gauche de (1.2), comme une limite de sommes de Riemann :

$$\frac{1}{t} \int_0^t f(X_s) ds = \lim_{k \rightarrow \infty} k^{-1} \sum_{i=1}^k f(X_{it/k}).$$

Fixons alors k , et considérons le noyau markovien discret $P_{t/k}$. Pour tout $r > 0$ l'inégalité de Markov donne la majoration suivante :

$$\begin{aligned} P_q \left[k^{-1} \sum_{i=1}^k f(X_{it/k}) > \gamma \right] &\leq e^{-rt\gamma} E_q \left(\exp k^{-1} \sum_{i=1}^k (rt) f(X_{it/k}) \right) \\ &\leq e^{-rt\gamma} \langle q/\pi, P_{t/k}^k(r) \cdot \mathbf{1} \rangle, \end{aligned}$$

où $P_{t/k}(r)g(x) := \sum_y P_{t/k}(x, y) \exp(rt k^{-1} f(y))g(y)$. Cet opérateur peut s'exprimer sous la forme suivante :

$$P_{t/k}(r) = \exp(-\Lambda t/k) \exp(rDt/k),$$

où D est l'opérateur multiplicatif défini dans le lemme 1.4. Le résultat suivant, valable pour toutes matrices A, B ,

$$\lim_{k \rightarrow \infty} \left(\exp(A/k) \exp(B/k) \right)^k = \exp(A + B),$$

nous permet, par passage à la limite et grâce au lemme de Fatou, d'obtenir l'inégalité (1.2) \square

LEMME 1.5. *Pour tout $r > 0$,*

$$(1.3) \quad P_q \left[\frac{1}{t} \int_0^t f(X_s) ds > \gamma \right] \leq N_q \exp \left\{ -t(r\gamma + \lambda_0(r)) \right\},$$

où $\lambda_0(r)$ est la plus petite valeur propre de l'opérateur autoadjoint

$$\tilde{\Lambda}(r) := (\Lambda + \Lambda^*)/2 - rD.$$

DÉMONSTRATION. Notons $P_t(r)$ le semi-groupe (non markovien) engendré par $\Lambda(r)$ et considérons la fonction φ définie par $\varphi(t) = \|P_t(r) \cdot g\|^2$. Nous obtenons, par dérivation, la relation

$$(1.4) \quad \varphi'(t) = -2 \langle \tilde{\Lambda}(r) P_t(r) \cdot g, P_t(r) \cdot g \rangle.$$

L'opérateur $\tilde{\Lambda}(r)$ étant autoadjoint, ses valeurs propres sont réelles. Soit $\lambda_0(r)$ la plus petite ; sa caractérisation extrême

$$\lambda_0(r) = \inf \left\{ \langle \tilde{\Lambda}(r)g, g \rangle / \|g\|^2, g \neq 0 \right\},$$

conduit à l'inégalité $\varphi'(t) \leq -2\lambda_0(r)\|P_t(r)g\|^2$. La fonction $\varphi(t) \exp(2\lambda_0(r)t)$ est donc décroissante et $\varphi(0) = \|g\|^2$, d'où

$$\|P_t(r)\| \leq \exp(-\lambda_0(r)t).$$

L'inégalité de Cauchy-Schwarz appliquée à (1.2), et combinée avec le résultat précédent conduit à l'inégalité (1.3). \square

Nous avons donc ramené notre problème initial à la détermination d'une minoration de la plus petite valeur propre d'une perturbation linéaire du symétrisé additif, $(\Lambda + \Lambda^*)/2$, du générateur du semi-groupe P_t .

REMARQUE 1.6. La propriété réciproque est également vraie, à savoir que si pour toute fonction de moyenne nulle, il existe une constante $\lambda > 0$ telle que $\|P_t f\| \leq e^{-\lambda t} \|f\|$, alors

$$\lambda \leq \inf \left\{ \left\langle \frac{\Lambda + \Lambda^*}{2} f, f \right\rangle / \|f\|^2, \pi \cdot f = 0 \right\}.$$

Il suffit pour cela de considérer la fonction croissante $\varphi(t) := \|P_t f\|^2 \exp(2\lambda t)$.

PREUVE DU THÉORÈME 1.1. D'après les résultats présentés au §1-4.2, $\lambda_0(r)$ correspond à la perturbation de la plus petite valeur propre de $(\Lambda + \Lambda^*)/2$. Par conséquent, une minoration de $\lambda_0(r)$ peut être obtenue à partir des estimations 1-(4.13). En particulier, il existe r_0 tel que pour tout $r < r_0$, $\lambda_0(r)$ admet un développement en série entière convergente

$$\lambda_0(r) = \lambda_0(0) + \sum_{n=1}^{\infty} \lambda_0^{(n)} (-1)^n r^n.$$

Le rayon de convergence r_0 s'exprime ici simplement en fonction du trou spectral de l'opérateur $(\Lambda + \Lambda^*)/2$, défini comme la distance entre $\lambda_0(0) = 0$ et le reste du spectre. Si λ_1 désigne la plus petite valeur propre non nulle de $(\Lambda + \Lambda^*)/2$, le trou spectral sera égal à λ_1 , et $r_0 = \lambda_1/2$.

L'opérateur perturbé étant autoadjoint et la perturbation linéaire, les formules 1-(4.13), exprimant les divers coefficients $\lambda^{(n)}$, se simplifient. Nous obtenons

$$(1.5) \quad \lambda_0^{(n)} = \frac{(-1)^n}{n} \sum_{k_1 + \dots + k_n = n-1} \text{tr} \left(DZ^{(k_1)} DZ^{(k_2)} \dots DZ^{(k_n)} \right),$$

où la somme porte sur tous les n -uplets (k_1, \dots, k_n) d'entiers positifs ou nuls solutions de l'équation $k_1 + \dots + k_n = n - 1$. Rappelons que Z désigne la résolvante réduite de l'opérateur $(\Lambda + \Lambda^*)/2$ pour la valeur propre $\lambda_0 = 0$ et

$$Z^{(0)} := -\pi, \quad Z^{(k)} := Z^k \quad k > 0.$$

Z n'est autre que l'inverse de $(\Lambda + \Lambda^*)/2$ sur le sous-espace des fonctions de moyenne nulle pour la probabilité invariante. Il s'exprime également comme la limite $\lim_{\zeta \rightarrow 0} R(\zeta)$, où $R(\zeta)$ est la résolvante de l'opérateur $(\Lambda + \Lambda^*)/2$. Nous en déduisons l'expression de la restriction de Z aux fonctions de moyenne nulle,

$$Z = \int_0^{\infty} \exp(-t(\Lambda + \Lambda^*)/2) dt.$$

En particulier si le processus de Markov est réversible, donc Λ autoadjoint, alors Z n'est autre que la matrice fondamentale

$$(1.6) \quad Z := \int_0^{\infty} P_t dt.$$

Définissons sur les n -uplets (k_1, \dots, k_n) une relation d'équivalence en posant que (k_1, \dots, k_n) est équivalent à (m_1, \dots, m_n) si et seulement si (m_1, \dots, m_n) est l'image de (k_1, \dots, k_n) par une permutation circulaire; nous noterons par $[k_1, \dots, k_n]$ la classe d'équivalence correspondante. La relation $\text{tr}(AB) = \text{tr}(BA)$ valable pour toutes matrices A, B implique que la valeur de $\text{tr}(A_{k_1} \dots A_{k_n})$ ne dépend que de la classe d'équivalence $[k_1, \dots, k_n]$. D'autre part les n -uplets solutions de l'équation $k_1 + \dots + k_n = n - 1$ contiennent toujours une composante

nulle ; choisissons comme représentant de la classe d'équivalence $[k_1, \dots, k_n]$ le n -uplet dont la première composante est nulle (i.e tel que $k_1 = 0$). Nous pouvons alors exprimer la relation (1.5) sous la forme suivante :

$$(1.7) \quad \lambda_0^{(n)} = (-1)^{n+1} \sum_{[0, k_1, \dots, k_{n-1}]} \text{tr} \left(\pi D Z^{(k_1)} D \dots D Z^{(k_{n-1})} D \right)$$

$$(1.8) \quad = (-1)^{n+1} \sum_{[0, k_1, \dots, k_{n-1}]} \left\langle f, Z^{(k_1)} D \dots D Z^{(k_{n-1})} f \right\rangle,$$

où la somme porte sur toutes les classes d'équivalence des n -uplets (k_1, \dots, k_n) solutions de l'équation $k_1 + \dots + k_n = n - 1$ (d'où la disparition du facteur n^{-1}). En particulier, nous obtenons

$$\begin{aligned} \lambda_0^{(1)} &= 0, \\ \lambda_0^{(2)} &= -\langle f, Zf \rangle, \\ \lambda_0^{(3)} &= \langle f, (ZDZ)f \rangle, \\ \lambda_0^{(4)} &= -\langle f, (ZD)^2 Zf \rangle + \langle f, (Z^2 D \pi D Z)f \rangle. \end{aligned}$$

Dans (1.8), la somme contient des termes nuls ; ce sont ceux correspondant aux classes d'équivalence $[0, k_1, \dots, k_{n-2}, 0]$ qui s'expriment sous la forme $\langle f, M\pi f \rangle = 0$, où M est un certain opérateur. Déterminons le nombre de termes non nuls dans cette somme. Le nombre de solutions de l'équation $k_1 + \dots + k_n = n - 1$ telles que m composantes fixées soient nulles, les autres étant toutes strictement positives, est $\binom{n-2}{n-m-1}$; si $2m > n$, alors le n -uplet aura toujours deux zéros adjacents, et pour $2m \leq n$ le nombre de possibilités de n'avoir aucun zéro adjacent est égal au nombre de solutions de l'équation $r_1 + \dots + r_m = n - m$ avec $r_i > 0$ pour tout $1 \leq i \leq m$, soit $\binom{n-m-1}{m-1}$. Ce nombre donne en particulier le nombre de possibilités de placer les m zéros sur le cercle $\mathbf{Z}/n\mathbf{Z}$ de façon à n'avoir aucun zéro adjacent, sous la restriction qu'un zéro est dans une position fixée a priori. Pour avoir le nombre total de classes d'équivalence, il faut tenir compte des permutations circulaires des différents m -uplets, soit un nombre total de $m^{-1} \binom{n-m-1}{m-1}$. Le nombre de termes non nuls s'exprime donc par la somme suivante

$$\sum_{m=1}^{\lceil n/2 \rceil} \frac{1}{m} \binom{n-m-1}{m-1} \binom{n-2}{n-m-1}.$$

En tenant compte de l'identité $\|Z\| = 1/\lambda_1$, due au fait que Z est un opérateur autoadjoint, nous obtenons par application de l'inégalité de Cauchy-Schwarz la majoration suivante de chaque coefficient

$$(1.9) \quad |\lambda_0^{(n)}| \leq b^2 r_0^{-(n-1)} \beta_n,$$

où

$$\beta_n = 2^{1-n} \sum_{m=1}^{\lceil n/2 \rceil} \frac{1}{m} \binom{n-m-1}{m-1} \binom{n-2}{n-m-1}.$$

Pour n petit, ces coefficients ont pour valeurs :

n	β_n	2^{n-5}
2	1/2 = 0,5000	1/8
3	1/4 = 0,2500	1/4
4	1/4 = 0,2500	1/2
5	1/4 = 0,2500	1
6	9/32 = 0,2812	2
7	21/64 = 0,3281	4
8	51/128 = 0,3984	8
9	127/256 = 0,4961	16
10	323/512 = 0,6308	32
11	835/1024 = 0,8154	64
12	547/512 = 1,0684	128

L'estimation (1.9) est à comparer à celle obtenue par application des inégalités 1-(4.14). Dans l'hypothèse $b^2 = 1$, celle-ci est meilleure pour seulement les 11 premiers termes, mais elle présente le mérite de faire apparaître la variance de f (voir aussi [34] §II.3 inégalité (3.66)).

Comme nous le prouverons ci-dessous, les coefficients β_n satisfont, pour tout $n \geq 3$, à l'inégalité $\beta_n \leq 2^{n-5}$. Ainsi, nous obtenons pour tout $r < r_0/2$ la minoration suivante :

$$\lambda_0(r) \geq -\frac{b^2}{\lambda_1} r^2 - \frac{b^2 \lambda_1}{8} \frac{(r/r_0)^3}{1-2r/r_0} \geq -\frac{b^2}{\lambda_1} r^2 \left(1 - \frac{4r}{\lambda_1}\right)^{-1}.$$

La fonction

$$r \longrightarrow \gamma r - \frac{b^2}{\lambda_1} r^2 \left(1 - \frac{4r}{\lambda_1}\right)^{-1}$$

atteint son maximum pour

$$r = \frac{\gamma \lambda_1}{b^2(1 + 4\gamma/b^2 + \sqrt{1 + 4\gamma/b^2})} < \frac{\lambda_1}{4}.$$

Portant cette valeur dans l'expression (1.3), nous obtenons l'inégalité (1.1). \square

PREUVE DE LA MAJORATION DE β_n . Considérons d'abord le cas n pair ; pour $n = 2p$

$$\beta_{2p} = \frac{1}{2^{2p-1}} \sum_{m=0}^{p-1} \frac{1}{m+1} \frac{(2(p-1))!}{(m!)^2 (2(p-1)-2m)!}.$$

D'après la double inégalité suivante [22][II-9.15],

$$(1.10) \quad \sqrt{2\pi n}^{n+1/2} e^{-n} e^{(12n+1)^{-1}} < n! < \sqrt{2\pi n}^{n+1/2} e^{-n} e^{(12n)^{-1}},$$

nous obtenons pour tout $k \geq 1$

$$(1.11) \quad \frac{2^{2k}}{\sqrt{k\pi}} e^{-1/8} < \binom{2k}{k} < \frac{2^{2k}}{\sqrt{k\pi}}.$$

Par conséquent pour tout $p \geq 2$,

$$\begin{aligned} \beta_{2p} &< \frac{1}{2^{2p-1}} \left\{ \frac{\pi^{1/2} e^{1/8}}{4} (p-1)^{1/2} \sum_{m=0}^{p-2} \frac{(2(p-1))!}{(m!)^2 ((p-1-m)!)^2} + \frac{1}{p} \frac{(2(p-1))!}{((p-1)!)^2} \right\} \\ &< \frac{(p-1)^{1/2}}{2^{2p}} \sum_{m=0}^{p-1} \frac{(2(p-1))!}{(m!)^2 ((p-1-m)!)^2} \\ &= \frac{(p-1)^{1/2}}{2^{2p}} \binom{2(p-1)}{p-1}^2 \\ &< \frac{2^{2p-4}}{\pi(p-1)^{1/2}} \leq 2^{2p-5}, \end{aligned}$$

où nous avons utilisé pour l'avant-dernière ligne, la relation [22][II-12.12]

$$\sum_{j=0}^n \frac{(2n)!}{(j!)^2 (n-j)!^2} = \binom{2n}{n}^2.$$

Une démarche identique donne pour $n = 2p + 1$ la majoration $\beta_{2p+1} < 2^{2p-6}$, valable pour $p \geq 2$. Nous avons ainsi obtenu que pour tout $n \geq 3$, $\beta_n \leq 2^{n-5}$. \square

REMARQUE 1.7. Le théorème 1.1 fournit dans le cas non réversible une majoration de la vitesse de convergence en temps fini, en fonction de la plus petite valeur propre positive du symétrisé additif de Λ . Dans [1][Chap.4 §7], les auteurs expriment leur scepticisme sur la possibilité d'obtenir des informations sur le comportement du processus de Markov par une estimation du trou spectral δ , défini par

$$\delta = \min\{\operatorname{Re}(\lambda), \lambda \neq 0 \text{ valeur propre de } \Lambda\}.$$

En particulier,

$$\rho(t) = \max_{h,g} \frac{E_\pi \left((h(X_0) - \pi.h)(g(X_t) - \pi.g) \right)}{\sqrt{\operatorname{Var}_\pi(h) \operatorname{Var}_\pi(g)}},$$

est asymptotiquement équivalent à $c \exp(-\delta t)$, sans être pour autant borné par $K \exp(-\delta t)$, où K est une constante universelle. Notons que $\lambda_1 \leq \delta$, l'égalité ayant lieu en particulier lorsque Λ est réversible. Ce résultat se déduit de la caractérisation extrémale de λ_1 et de l'identité

$$\operatorname{Re}(\langle \Lambda f, f \rangle) = \left\langle \frac{1}{2} (\Lambda + \Lambda^*) f, f \right\rangle.$$

Il semble donc peu probable d'améliorer, dans le cas non réversible, la majoration obtenue en considérant δ au lieu de λ_1 .

REMARQUE 1.8. Le théorème de Gärtner-Ellis [13, 14] prouve que pour le processus considéré, $t^{-1} \int_0^t f(X_s) ds$ satisfait un principe de grandes déviations de fonction de taux

$$I_f(x) := \sup_r \{rx - \rho(r)\},$$

où

$$(1.12) \quad \rho(r) = \lim_{t \rightarrow \infty} \frac{1}{t} \log E_q \left[\exp \int_0^t r f(X_s) ds \right].$$

Or, $\rho(r)$ est égale au rayon spectral logarithmique

$$\rho(r) := \lim_t \frac{1}{t} \log \|P_t(r)\|_{\infty \rightarrow \infty}$$

de l'opérateur $P_1(r)$ [14][corollaire 4.2.27]. Mais en dimension finie, nous savons, d'après le théorème de Perron-Frobenius, que ce rayon spectral est la plus grande valeur propre de $-\Lambda(r)$. Par exemple, dans le cas réversible $\rho(r) = -\lambda_0(r)$.

1.1. Processus réversible. Nous supposons dorénavant que le processus de Markov considéré est réversible; par conséquent Λ est autoadjoint dans $\ell^2(\pi)$. Soient

$$\lambda_0 = 0 < \lambda_1 \leq \dots \leq \lambda_{N-1}$$

ses valeurs propres et $\{\varphi_i\}$ une base orthonormée constituée des fonctions propres de Λ associées respectivement aux valeurs propres λ_i ; en particulier $\varphi_0 = 1$. Nous avons alors les relations suivantes :

$$\begin{aligned} P_t \varphi_i &= e^{-\lambda_i t} \varphi_i \\ Z \varphi_i &= \lambda_i^{-1} \varphi_i, \quad i > 0 \\ E_\pi f(X_0) f(X_t) &= \langle f, P_t f \rangle = \sum_{i \geq 1} e^{-\lambda_i t} \langle f, \varphi_i \rangle^2. \end{aligned}$$

Suivons maintenant Aldous et Fill dans [1][Chap. 4 prop. 29] pour prouver la proposition suivante :

PROPOSITION 1.9. *Soit $Y_t = \int_0^t f(X_s) ds$ alors $t^{-1} \text{Var}_\pi Y_t \uparrow \sigma^2$ quand $t \uparrow \infty$, où*

$$\sigma^2 := 2 \sum_{i \geq 1} \langle f, \varphi_i \rangle^2 / \lambda_i = 2 \langle f, Z f \rangle.$$

D'autre part

$$\begin{aligned} 2 \|f\|^2 / \lambda_{N-1} &\leq \sigma^2 \leq 2 \|f\|^2 / \lambda_1 \\ A(\lambda_1 t) t \sigma^2 &\leq \text{Var}_\pi Y_t \leq t \sigma^2, \end{aligned}$$

où

$$A(s) := \int_0^s \left(1 - \frac{u}{s}\right) e^{-u} du = 1 - s^{-1} (1 - e^{-s}) \uparrow 1 \text{ quand } s \uparrow \infty$$

DÉMONSTRATION.

$$\begin{aligned} t^{-1} \text{Var}_\pi Y_t &= t^{-1} \int_0^t \int_0^t E_\pi f(X_{t_1}) f(X_{t_2}) dt_1 dt_2 \\ &= 2 \int_0^t \left(1 - \frac{u}{t}\right) E_\pi f(X_0) f(X_u) du \\ (1.13) \quad &= 2 \sum_{i \geq 1} \langle f, \varphi_i \rangle^2 \lambda_i^{-1} A(\lambda_i t). \end{aligned}$$

Le terme de droite de la dernière inégalité croît vers $\sigma^2 := 2 \sum_{i \geq 1} \langle f, \varphi_i \rangle^2 \lambda_i^{-1}$. Nous vérifions aisément que d'une part $\sigma^2 = 2 \langle f, Z f \rangle$ d'après (1.6), et d'autre part que

$$2 \|f\|^2 / \lambda_{N-1} \leq \sigma^2 \leq 2 \|f\|^2 / \lambda_1.$$

Comme $A(\cdot)$ est une fonction croissante la dernière inégalité de la proposition 1.9 s'obtient à partir de (1.13). \square

REMARQUE 1.10. La proposition 1.9 fournit des bornes à la vitesse de convergence de $t^{-1} \text{Var}_\pi Y_t$ vers la variance asymptotique σ^2 . En effet, nous obtenons les inégalités

$$\frac{2 \|f\|^2}{\lambda_{N-1}^2 t} (1 - e^{-\lambda_{N-1} t}) \leq \sigma^2 - t^{-1} \text{Var}_\pi Y_t \leq \frac{2 \|f\|^2}{\lambda_1^2 t} (1 - e^{-\lambda_1 t}).$$

La relation $\lambda_0^{(2)} = -\langle f, Z f \rangle$ implique que $\lambda_0^{(2)} = -\sigma^2/2$.

1.1.1. *Majoration en fonction de σ^2 dans le cas réversible.* Prouvons maintenant dans le cas réversible, une autre version du théorème 1.1 où la borne s'exprimera en fonction de la variance asymptotique σ^2 au lieu de la variance de f . Cette version nous sera utile pour déterminer une borne exponentielle inférieure à la vitesse de convergence.

THÉORÈME 1.11. *Soient (P_t, π) , un noyau markovien en temps continu, irréductible, réversible, défini sur un espace d'états fini et $-\Lambda$ son générateur infinitésimal. Soit $f : G \rightarrow \mathbb{R}$ telle que $\pi.f = 0$, $\sup_x |f(x)| \leq a$ et $\sigma^2 > 0$. Alors pour toute distribution initiale q , tout $t > 0$ et tout $0 < \gamma \leq a$,*

$$(1.14) \quad P_q \left[\frac{1}{t} \int_0^t f(X_s) ds \geq \gamma \right] \leq N_q \exp \left[- \frac{\gamma^2 t}{2\sigma^2(1 + h^*(4\gamma a/\lambda_1\sigma^2))} \right],$$

où

$$h^*(x) = \frac{1}{2} \left(\sqrt{(1-x)^2 + \frac{2\|f\|^2}{\lambda_1\sigma^2}x} - (1-x) \right).$$

DÉMONSTRATION. Nous obtenons, sous les hypothèses du théorème 1.11, les estimations suivantes :

$$(1.15) \quad r_0 = \lambda_1/2a, \quad |\lambda_0^{(n)}| \leq \left(\frac{\|f\|^2}{a^2} \right) \left(\frac{\lambda_1}{2} \right) r_0^{-n} \beta_n,$$

d'où la minoration suivante de $\lambda_0(r)$

$$\lambda_0(r) \geq - \frac{\sigma^2 r^2}{2} \left(1 + \frac{\|f\|^2}{\lambda_1\sigma^2} \frac{r/r_0}{1 - 2r/r_0} \right).$$

Pour conclure, il suffit de poser $r = \gamma/\sigma^2(1 + h^*(4\gamma a/\lambda_1\sigma^2))$ dans l'inégalité (1.3). \square

REMARQUE 1.12. D'après la proposition 1.9, $\sigma^2 \geq 2\|f\|^2/\lambda_{N-1}$, d'où la possibilité de remplacer dans l'inégalité précédente la fonction h^* par la fonction

$$\tilde{h}(x) := \frac{1}{2} \left(\sqrt{(1-x)^2 + \frac{\lambda_{N-1}}{\lambda_1}x} - (1-x) \right).$$

REMARQUE 1.13. Lorsque la fonction f est une fonction propre $f = \varphi_i$, le théorème 1.11 fournit la majoration suivante

$$P_q \left[\frac{1}{t} \int_0^t \varphi_i(X_s) ds \geq \gamma \right] \leq N_q \exp \left[\frac{-\gamma^2 t \lambda_i}{4(1 + h(2\gamma a \lambda_i/\lambda_1))} \right],$$

où $h(x) = \frac{1}{2} \left[\sqrt{(1-x)^2 + \frac{\lambda_i}{\lambda_1}x} - (1-x) \right]$. En particulier, si $\gamma \ll \lambda_1/(2a\lambda_i)$ nous obtenons une borne asymptotique en $N_q e^{-\gamma^2 \lambda_i t/4}$, qui sera d'autant meilleure que la valeur propre λ_i sera grande par rapport à λ_1 .

1.2. Quelques propriétés de la matrice fondamentale Z . Comme nous l'avons vu, la matrice fondamentale Z peut s'interpréter en termes probabilistes. Il est également possible de donner une interprétation probabiliste de certains coefficients de cette matrice, en particulier les résultats suivants figurant dans [1][Chap.2, lemmes 10 et 11]. Nous ne supposerons pas nécessairement la réversibilité du processus de Markov; Z sera alors défini sur les fonctions de moyenne nulle par la formule (1.6).

LEMME 1.14. Soit $T_x = \inf\{t \geq 0 : X_t = x\}$, le premier temps d'atteinte de l'état x , alors

$$(1.16) \quad \pi(x)E_\pi T_x = Z(x, x)$$

$$(1.17) \quad \pi(y)E_x T_y = Z(y, y) - Z(x, y)$$

$$(1.18) \quad \sum_y \pi(y)E_x T_y = \sum_y Z(y, y)$$

$$(1.19) \quad \max_{x,z} \sum_y \pi(y) \left| E_x T_y - E_z T_y \right| = \max_{x,z} \sum_y \left| Z(z, y) - Z(x, y) \right|$$

DÉMONSTRATION. Nous allons donner de ce lemme une preuve différente de celle figurant dans [1]; elle sera basée sur l'approche martingale. En effet, pour tout $g \in \ell^2(\pi)$ le processus

$$M_t^g = g(X_t) - g(X_0) + \int_0^t \Lambda g(X_s) ds$$

est une martingale pour la filtration canonique (\mathcal{F}_t^X) du processus. En particulier si T est un temps d'arrêt p.s. fini, alors le théorème d'arrêt donne, pour toute mesure initiale du processus, la relation suivante :

$$(1.20) \quad E_q g(X_T) = E_q g(X_0) - E_q \int_0^T \Lambda g(X_s) ds.$$

Celle-ci donne pour $g = Zf$, où f est une fonction de moyenne nulle, et $T = T_x$, l'identité

$$(1.21) \quad Zf(x) = -E_\pi \int_0^{T_x} f(X_s) ds.$$

Rappelons d'abord la relation $\sum_y Z(x, y) = 0$, valable pour tout $x \in G$. Les relations (1.16) et (1.17) se déduisent respectivement de l'identité (1.21), appliquée à la fonction $f_x = \delta_x - \pi(x)$, et de l'identité (1.20) avec $T = T_y$, $f_y = \delta_y - \pi(y)$ et $q = \delta_x$.

La relation (1.18) découle de la relation (1.17) après sommation sur les y , tandis que la relation (1.19) s'obtient en notant que la relation (1.17) implique que $Z(z, y) - Z(x, y) = \pi(y)(E_x T_y - E_z T_y)$. \square

Supposons maintenant le processus markovien réversible et introduisons, conformément à la terminologie employée dans [1][Chap 4], les paramètres suivants :

- Le temps moyen maximum d'aller-retour entre deux points

$$\tau^* := \max_{x,y} (E_x T_y + E_y T_x).$$

- Le temps d'atteinte moyen

$$\tau_0 := \sum_y \pi(y) \pi(x) E_x T_y.$$

- Le temps de relaxation

$$\tau_2 := \lambda_1^{-1}.$$

- Le temps minimal d'atteinte de la distribution stationnaire

$$\tau_1^{(2)} := \max_x \min_{U_x} E_x U_x,$$

où le minimum porte sur tous les temps d'arrêt U_x tels que $P_x(X(U_x) \in A) = \pi(A)$. Ce n'est autre que le temps minimal moyen nécessaire pour générer un échantillon de distribution la mesure π .

– Mesure de la variation du temps d'atteinte en fonction du point de départ

$$\tau_1^{(4)} := \max_{x,z} \sum_y \pi(y) \left| E_x T_y - E_z T_y \right|$$

La relation (1.18) montre que $\sum_y \pi(y) E_x T_y$ ne dépend pas de x , donc $\tau_0 = \sum_{i \geq 1} \lambda_i^{-1}$. Nous en déduisons la double inégalité

$$\tau_2 \leq \tau_0 \leq (N-1)\tau_2.$$

Soit la fonction $f = \lambda f_x + \mu f_y$ où λ et μ sont des constantes réelles. L'opérateur Z étant autoadjoint dans $\ell^2(\pi)$ et défini positif, nous obtenons

$$\langle f, Zf \rangle = \lambda^2 \pi(x) Z(x, x) + \mu^2 \pi(y) Z(y, y) + 2\lambda\mu\pi(x) Z(x, y) \geq 0.$$

Comme $Z(x, x) \geq 0$, la condition précédente valable pour tout λ et tout μ implique

$$\left(\lambda \sqrt{Z(x, x)} + \mu \sqrt{\frac{\pi(y)}{\pi(x)} Z(y, y)} \right)^2 + 2\lambda\mu \left(Z(x, y) - \sqrt{\frac{\pi(y)}{\pi(x)} Z(y, y) Z(x, x)} \right) \geq 0.$$

Cette dernière inégalité, appliquée pour $\lambda = -\sqrt{\frac{\pi(y)}{\pi(x)} Z(y, y)}$ et $\mu = \sqrt{Z(x, x)}$, donne la majoration suivante valable pour tout x, y :

$$(1.22) \quad Z^2(x, y) \leq Z(x, x) Z(y, y) \frac{\pi(y)}{\pi(x)}.$$

Ainsi

$$(1.23) \quad \left(\frac{Z(x, y)}{\pi(y)} \right)^2 \leq \left(\frac{Z(x, x)}{\pi(x)} \right) \left(\frac{Z(y, y)}{\pi(y)} \right),$$

d'où

$$-\frac{Z(x, y)}{\pi(y)} \leq \max_x \frac{Z(x, x)}{\pi(x)}.$$

Cette dernière inégalité, combinée avec les relations (1.16) et (1.17), implique

$$E_x T_y \leq 2 \max_x \frac{Z(x, x)}{\pi(x)} = 2 \max_x E_\pi T_x,$$

d'où la double inégalité suivante :

$$\max_x E_\pi T_x \leq \tau^* \leq 4 \max_x E_\pi T_x.$$

Pour comparer τ^* et τ_0 , il suffit de noter que

$$\tau_0 = \frac{1}{2} \sum_{x,y} \pi(x)\pi(y) (E_x T_y + E_y T_x) \leq \frac{1}{2} \tau^*.$$

D'autre part comme $E_x T_y \leq \tau_0/\pi(y)$, il s'ensuit que $\tau^* \leq 2\tau_0/\pi_*$, où $\pi_* := \min_x \pi(x)$. Néanmoins, il existe une meilleure majoration, à savoir que $\tau^* \leq 2\tau_2/\pi_*$ [1][lem. 25]. Nous avons donc le lemme suivant [1][Chap. 4]

LEMME 1.15.

$$\begin{aligned} \tau_2 \leq \tau_0 \leq (N-1)\tau_2, \\ 2\tau_0 \leq \tau^* \leq \frac{2\tau_2}{\pi_*}. \end{aligned}$$

Il est possible d'obtenir une autre majoration de τ_2 , valable même dans le cas non réversible. Pour cela il suffit d'appliquer le théorème 2.10 dans [48], qui pour toute matrice $A = (a_{ij})$ dont la somme des coefficients de chaque ligne est une constante a , fournit l'inégalité suivante valable pour toute valeur propre λ de A différente de a

$$|\lambda| \leq \frac{1}{2} \max_{i,j} \sum_{s=1}^n |a_{is} - a_{js}|.$$

Appliquons ce résultat à la matrice fondamentale Z , qui vérifie les hypothèses du théorème avec $a = 0$. Nous obtenons d'après le lemme 1.14, l'identité suivante :

$$(1.24) \quad 2\tau_2 \leq \max_{x,z} \sum_y \pi(y) |E_x T_y - E_z T_y| = \tau_1^{(4)}.$$

Ce résultat signifie que toutes les valeurs propres de Z se situent dans le cercle centré en zéro et de rayon $\tau_1^{(4)}/2$.

Le temps $\tau_1^{(2)}$ s'exprime d'après [1][Chap. 4] par

$$\tau_1^{(2)} = \max_{x,y} \frac{-Z(x,y)}{\pi(y)}.$$

La relation (1.20), appliquée avec $g = Zf$ (f de moyenne nulle) et $T = U_x$, donne l'identité

$$(1.25) \quad Zf(x) = E_x \int_0^{U_x} f(X_s) ds.$$

En particulier pour $f_y = \delta_y - \pi(y)$ elle implique la relation

$$(1.26) \quad \frac{Z(x,y)}{\pi(y)} = E_x U_x - \frac{1}{\pi(y)} E_x (\text{durée de séjour en } y \text{ avant le temps } U_x),$$

d'où

$$E_x T_y \leq E_\pi T_y + \tau_1^{(2)}.$$

En reprenant les arguments de [1][Chap. 4], nous avons

$$\begin{aligned} \sum_y \pi(y) E_x T_y &= \sum_{x,y} \pi(y) \pi(x) E_x T_y = \sum_y \pi(y) E_\pi T_y \\ \sum_y \pi(y) |E_x T_y - E_\pi T_y| &= 2 \sum_y \pi(y) (E_x T_y - E_\pi T_y)^+ \leq 2\tau_1^{(2)} \end{aligned}$$

d'où par l'inégalité triangulaire

$$\tau_1^{(4)} \leq 4\tau_1^{(2)}$$

Nous avons donc le lemme suivant

LEMME 1.16.

$$(1.27) \quad \tau_2 \leq \tau_1^{(4)}/2 \leq 2\tau_1^{(2)}.$$

Une meilleure relation entre τ_2 et $\tau_1^{(2)}$ s'obtient en appliquant l'identité (1.25) à la fonction φ_1 . En effet, nous obtenons la relation

$$\tau_2 \varphi_1(x) = \min_{U_x} E_x \int_0^{U_x} \varphi_1(X_s) ds.$$

Soit $a = \max_x \varphi_1(x) = \varphi_1(x_0)$. a est positif car sinon la relation (1.21) impliquerait que $\tau_2 < 0$. Nous en déduisons que

$$\tau_2 \varphi_1(x_0) = \min_{U_{x_0}} E_{x_0} \int_0^{U_{x_0}} \varphi_1(X_s) ds \leq a \min_{U_{x_0}} E_{x_0} U_{x_0},$$

d'où

LEMME 1.17.

$$\tau_2 \leq \tau_1^{(2)}.$$

Ce résultat répond par l'affirmative au problème ouvert 48 b) du Chapitre 3 de [1].

REMARQUE 1.18. La relation (1.23) permet de majorer la norme $\|Z\|_{2 \rightarrow \infty}$. En effet

$$\begin{aligned} \|Z\|_{2 \rightarrow \infty} &= \max_x \left(\sum_y \left(\frac{Z(x, y)}{\pi(y)} \right)^2 \pi(y) \right)^{1/2} \\ &\leq \max_x \left(\sum_y \frac{Z(x, x)}{\pi(x)} \frac{Z(y, y)}{\pi(y)} \pi(y) \right)^{1/2} \\ &\leq \max_x \left| \frac{Z(x, x)}{\pi(x)} \right|^{1/2} (\text{Trace}(Z))^{1/2} \\ &\leq \max_x (E_\pi T_x)^{1/2} \tau_0^{1/2} \leq \sqrt{\tau^* \tau_0} \end{aligned}$$

En particulier, le lemme (1.15) conduit à l'inégalité

$$\|Z\|_{2 \rightarrow \infty} \leq \tau^*.$$

2. Borne inférieure dans le cas réversible

Nous considérons dans cette section un processus de Markov réversible et une fonction f de moyenne nulle telle que $\sup_x |f(x)| \leq a$. L'objectif sera de prouver le théorème suivant :

THÉORÈME 2.1. *Pour tout $\epsilon > 0$, il existe des constantes $k(\epsilon)$ (assez petite) et $K(\epsilon)$ (assez grande) telles que : si $\gamma/a \leq k(\epsilon) \left(\frac{\lambda_1 \sigma^2}{2a^2}\right)^3$ et $t \geq K(\epsilon) \left(\frac{\sigma}{\gamma}\right)^2$, alors*

$$P_\pi(t^{-1}Y_t \geq \gamma) \geq \frac{\pi_*}{2} \exp\left\{-t \frac{\gamma^2}{2\sigma^2}(1 + \epsilon)\right\},$$

où $Y_t := \int_0^t f(X_s) ds$ et $\pi_* := \min_x \pi(x)$. En particulier, $k(6) = 4.10^{-3}$.

Ce théorème fournit une borne inférieure à la probabilité $P_\pi(t^{-1}Y_t \geq \gamma)$ pour tout t suffisamment grand et pour tout γ/a suffisamment petit. Cette borne est du même type que celle obtenue par Kolmogorov pour les variables aléatoires indépendantes [49][th. 5.2.2 (iii)]. Nous utiliserons d'ailleurs la preuve figurant dans [49], après adaptation au contexte markovien.

REMARQUE 2.2. L'inégalité $\sigma^2 \geq 2\|f\|^2/\lambda_{N-1}$, démontrée dans la proposition 1.9, permet d'exprimer la borne inférieure en termes de la variance de f et de la plus grande valeur propre de Λ

$$P_\pi(t^{-1}Y_t \geq \gamma) \geq \frac{\pi_*}{2} \exp\left\{-t \frac{7\gamma^2 \lambda_{N-1}}{4\|f\|^2}\right\}.$$

PREUVE DU THÉORÈME 2.1. La preuve se déroulera en deux étapes ; la première consistera à minorer $E_q \exp(rY_t)$ à partir des résultats obtenus dans la section précédente, la deuxième étape consistera à minorer la probabilité de déviation par $E_q \exp(rY_t)$. Dans toute la démonstration nous supposons r fixé ; les conditions qui apparaîtront au fur à mesure de son déroulement porteront alors sur le rapport a/σ . La valeur de r en fonction de γ sera allouée vers la fin de la preuve. A la différence de la preuve donnée dans [49], la valeur de r est bornée par $r_0/2$. Il ne sera donc pas possible de la faire tendre vers l'infini, l'alternative utilisée sera de prendre t assez grand.

Pour minorer $E_q \exp(rY_t)$, introduisons l'opérateur $P(r)$, projecteur propre de $\Lambda(r)$ associé à la valeur propre $\lambda_0(r)$. Le vecteur $u(r) := \langle P(r)1, 1 \rangle^{-1} P(r)1$ est un vecteur propre de $\Lambda(r)$ pour la valeur propre $\lambda_0(r)$, normalisé par la condition $\langle u(r), 1 \rangle = 1$; il suffit de se rappeler les relations $P(r)\Lambda(r) = \Lambda(r)P(r) = \lambda_0(r)P(r)$.

Soit $u_M(r) = \max_x u(r)(x)$, la condition de normalisation implique que $1 \leq u_M(r) \leq \pi_*^{-1}$, où $\pi_* = \min_x \pi(x)$. Nous obtenons ainsi

$$\begin{aligned} E_q \exp(rY_t) &\geq \sum_{x,y} q(x) \exp(-\Lambda(r)t)(x,y) u_r(y) / u_M(r) \\ &= M_q(r) \exp(-\lambda_0(r)t), \end{aligned}$$

où $M_q(r) = \langle q/\pi, u(r) \rangle u_M(r)^{-1}$. D'autre part, les relations (1.15) fournissent, à partir du développement en série de $\lambda_0(r)$, l'inégalité suivante :

$$\lambda_0(r) \leq -\frac{\sigma^2 r^2}{2} (1 - Bv),$$

où $v = (r/r_0)/(1-2r/r_0)$ et $B = a^2/(\lambda_1 \sigma^2)$. Notons que pour tout $r > 0$, $M_q(r) \leq 1$ et $1/2 \leq B$.

Fixons $0 < r < r_0/2$ et soit $0 < \delta < 1$, dont la valeur sera fixée ultérieurement. Supposons a/σ tel que

$$(2.1) \quad (r/r_0)B \leq \delta^2/8,$$

nous obtenons alors la minoration suivante :

$$(2.2) \quad E_q \exp(rY_t) \geq M_q(r) \exp\left[t \frac{\sigma^2 r^2}{2} (1 - \delta^2/4)\right]$$

Introduisons maintenant la fonction de répartition de support $[-a, a]$, $Q(x) = P_q(t^{-1}Y_t \leq x)$. Nous savons, d'après le théorème 1.11, que pour tout $-a \leq x \leq a$,

$$Q(x) \leq N_q \exp\left[-\frac{tx^2}{\sigma^2 \Phi(4Bx/a)}\right] \leq N_q \exp\left[-\frac{tx^2}{\sigma^2 \Phi(4B)}\right],$$

où

$$\Phi(y) = 1 + y + \sqrt{(1-y)^2 + 2By}.$$

Une intégration par parties conduit à l'expression suivante :

$$\begin{aligned} E_q \exp(rY_t) &= rt \int_{-a}^a \exp(rtx) Q(x) dx \\ &= rt \sum_{i=1}^5 \int_{A_i} \exp(rtx) Q(x) dx, \end{aligned}$$

où

$$\begin{aligned} A_1 &= [-a, 0], \\ A_2 &=]0, \sigma^2 r(1-\delta)[, \\ A_3 &= [\sigma^2 r(1-\delta), \sigma^2 r(1+\delta)], \\ A_4 &=]\sigma^2 r(1+\delta), 2\sigma^2 r\Phi(4B)] \\ A_5 &=]2\sigma^2 r\Phi(4B), a], \end{aligned}$$

à condition de supposer a/σ tel que $2\sigma^2 r\Phi(4B) < a$. Or $\sigma^2 r_0 \leq a$, il suffit donc d'avoir $2(r/r_0)\Phi(4B) < 1$. Posons $(4B)/\Phi(4B) = \Psi(4B)$, où

$$\Psi(x) = \frac{x}{1+x+\sqrt{(1-x)^2+x^2/2}}.$$

Cette fonction étant croissante, l'inégalité $B \geq 1/2$ implique que $\Psi(4B) \geq \Psi(2) \geq 0,4$; il suffira donc de supposer a/σ tel que

$$(2.3) \quad (r/r_0)B \leq \delta^2\Psi(2)/8,$$

cette condition remplaçant la condition (2.1).

Sur A_1 :

$$\int_{-a}^0 rt \exp(rtx)Q(x)dx \leq \int_{-\infty}^0 rt \exp(rtx)dx = 1.$$

Sur A_5 , $x \geq 2\sigma^2 r\Phi(4B)$, d'où

$$\int_{A_5} rt \exp(rtx)Q(x)dx \leq N_q \int_0^\infty rt \exp(-rtx)dx \leq N_q.$$

Les deux résultats précédents combinés impliquent la majoration suivante :

$$(2.4) \quad \int_{A_1 \cup A_5} rt \exp(rtx)Q(x)dx \leq 1 + N_q.$$

Sur $A_2 \cup A_4$, nous avons la majoration suivante :

$$\frac{4Bx}{a} \leq 4\frac{r}{r_0}\Phi(4B) \leq \frac{16}{\Psi(2)}\frac{r}{r_0}B.$$

Supposons maintenant que a/σ soit tel que

$$(2.5) \quad (r/r_0)B \leq \delta^2\Psi(2)/16$$

alors $4Bx/a < 1$. Ceci implique que

$$\begin{aligned} \Phi(4Bx/a) &\leq 2\left(1 + B\frac{2Bx/a}{1-4Bx/a}\right) \\ &\leq 2\left(1 + B\frac{2Bx/a}{1-\delta^2}\right) \\ &\leq 2\left(1 + \frac{C(r/r_0)}{1-\delta^2}\right) \end{aligned}$$

avec $C = 8B^2/\Psi(2)$, d'où

$$\exp(rtx)Q(x) \leq N_q \exp\left\{t\left(rx - \frac{x^2}{2\sigma^2}\left(1 - \frac{C(r/r_0)}{1-\delta^2}\right)\right)\right\}.$$

Supposons maintenant que a/σ soit tel que $C(r/r_0) < 1 - \delta^2$. Comme $B \geq 1/2$, il suffit que

$$(2.6) \quad C(r/r_0) < \min(\delta^2/2, 1 - \delta^2).$$

Posons

$$g(x) = rx - \frac{x^2}{2\sigma^2}\left(1 - \frac{C(r/r_0)}{1-\delta^2}\right),$$

pour tout $x \in \mathbb{R}_+$; g atteint son maximum en $x_0 = \sigma^2 r / (1 - C(r/r_0)(1 - \delta^2)^{-1})$. Supposons alors a/σ tel que $x_0 \in A_3$, ou plus précisément a/σ tel que

$$(2.7) \quad \frac{C(r/r_0)}{1-\delta^2} < \frac{1}{2}\left(\frac{\delta}{1+\delta}\right)^2.$$

Cette condition implique la condition (2.6) et $x_0 \in A_3$ (il suffit de remarquer que $(\delta/(1+\delta))^2 \leq \delta/(1+\delta)$). Nous obtenons ainsi pour $x \in A_2$

$$\begin{aligned} g(x) &\leq g(\sigma^2 r(1-\delta)) \\ &= \frac{\sigma^2 r^2}{2}(1-\delta) \left(1 + \delta + \frac{C(r/r_0)(1-\delta)}{1-\delta^2} \right) \\ &< \frac{\sigma^2 r^2}{2}(1-\delta^2/2), \end{aligned}$$

il s'ensuit

$$\int_{A_2} rt \exp(rtx) Q(x) dx \leq N_q \sigma^2 r^2 t \exp\left\{ \frac{t\sigma^2 r^2}{2}(1-\delta^2/2) \right\}.$$

De même, pour tout $x \in A_4$, nous avons

$$g(x) \leq g(\sigma^2 r(1+\delta)) < \frac{\sigma^2 r^2}{2}(1-\delta^2/2),$$

d'où

$$\begin{aligned} \int_{A_4} rt \exp(rtx) Q(x) dx &\leq N_q t (2\sigma^2 r^2 \Phi(4B) - \sigma^2 r(1+\delta)) \exp\left\{ \frac{t\sigma^2 r^2}{2}(1-\delta^2/2) \right\} \\ &\leq N_q t \sigma^2 r^2 (C-1) \exp\left\{ \frac{t\sigma^2 r^2}{2}(1-\delta^2/2) \right\}. \end{aligned}$$

Les deux dernières inégalités combinées conduisent à la majoration suivante :

$$(2.8) \quad \int_{A_2 \cup A_4} rt \exp(rtx) Q(x) dx \leq N_q t C \sigma^2 r^2 \exp\left\{ \frac{t\sigma^2 r^2}{2}(1-\delta^2/2) \right\}.$$

Posons $\gamma = r\sigma^2(1-\delta)$, la condition (2.7) devient alors une condition sur γ en fonction de a, σ et δ , plus précisément

$$(2.9) \quad \frac{\gamma}{a} \left(\frac{a}{\sigma} \right)^6 < \frac{\lambda_1^3 \Psi(2)}{32} \frac{(\delta(1-\delta))^2}{1+\delta}.$$

Notons que le terme de droite, dans l'inégalité précédente, est maximum pour $\delta = (-1 + \sqrt{11/3})/2$, et la condition (2.9) devient avec $\Psi(2) = 2/(3 + \sqrt{3})$

$$(2.10) \quad \frac{\gamma}{a} \leq 4.10^{-3} \left(\frac{\sigma^2 \lambda_1}{2a^2} \right)^3.$$

Ceci posé, nous obtenons alors sur A_3 ,

$$(2.11) \quad rt \int_{A_3} \exp(rtx) Q(x) dx \leq Q(\gamma) \exp(t\sigma^2 r^2(1+\delta)).$$

Choisissons maintenant t suffisamment grand pour que

$$\begin{aligned} \int_{A_1 \cup A_5} rt \exp(rtx) Q(x) dx &\leq (1/4) E_q(e^{rY_t}) \\ \int_{A_2 \cup A_4} rt \exp(rtx) Q(x) dx &\leq (1/4) E_q(e^{rY_t}) \end{aligned}$$

Pour cela il suffit d'après les inégalités (2.2), (2.4) et (2.8) d'avoir

$$(2.12) \quad 1 + N_q \leq \frac{M_q(r)}{4} \exp\left\{ t \frac{\sigma^2 r^2}{2}(1-\delta^2/4) \right\}$$

et

$$(2.13) \quad (tC\sigma^2 r^2) \exp\left\{ -\frac{t\sigma^2 r^2 \delta^2}{8} \right\} \leq M_q(r)/4.$$

Il convient donc de prendre $t\sigma^2r^2 \propto t(\gamma/\sigma)^2$ assez grand pour vérifier (2.12), ensuite il suffit de noter que le terme de gauche dans (2.13) tend vers zéro lorsque $t \rightarrow \infty$. Par conséquent, soit $\epsilon > 0$, il existe une constante assez grande $K(\epsilon)$, telle que $t \geq K(\epsilon)(\sigma/\gamma)^2$, l'inégalité suivante est vérifiée :

$$\begin{aligned} Q(\gamma) &\geq (1/2) \exp\left\{-t\sigma^2r^2(1+\delta)\right\} E_q \exp(tY_t) \\ &\geq (1/2) M_q\left(\frac{\gamma}{\sigma^2(1-\delta)}\right) \exp\left\{-\frac{t\gamma^2}{2\sigma^2} \frac{1+2\delta+\delta^2/4}{(1-\delta)^2}\right\}. \end{aligned}$$

Choisissons δ tel que

$$\frac{1+2\delta+\delta^2/4}{(1-\delta)^2} \leq 1+\epsilon,$$

alors

$$Q(\gamma) \geq (1/2) M_q\left(\frac{2\gamma}{\sigma^2(3-\sqrt{11/3})}\right) \exp\left\{-t\frac{\gamma^2}{2\sigma^2}(1+\epsilon)\right\}.$$

Pour conclure il suffit de noter que $M_\pi(r) \geq \pi_*$.

Notons également que pour $\delta = (-1 + \sqrt{11/3})/2$, nous obtenons

$$Q(\gamma) \geq \frac{\pi_*}{2} \exp\left\{-t\frac{7\gamma^2}{2\sigma^2}\right\}.$$

□

3. Borne du type Berry-Esséen

L'objectif de cette section est d'obtenir une borne du type Berry-Esséen pour les processus de Markov irréductibles et à espace d'états fini. Pour cela, nous allons suivre la preuve de B. Mann [39], qui a obtenu une telle borne pour une chaîne de Markov irréductible et à espace d'états fini ou dénombrable. Il suppose une hypothèse de mélange

$$(3.1) \quad \beta := \sup \frac{\|Pg\|}{\|g\|} < 1,$$

où le supremum est pris sur les fonctions $g \in \ell^2(\pi)$ telles que $E_\pi g = 0$.

La considération du temps continu et de l'utilisation des estimations fournies dans [34] conduira à une preuve simplifiée par rapport à celle de B. Mann. Plus précisément, outre le fait d'étendre son résultat au temps continu, nous améliorons la constante numérique figurant dans la borne obtenue d'un facteur d'environ 200. Le théorème que nous allons prouver est le suivant :

THÉORÈME 3.1. *Soit (P_t, π) un noyau markovien en temps continu irréductible défini sur un espace d'états G fini. Soient $-\Lambda$ son générateur infinitésimal et $f : G \rightarrow \mathbb{R}$ telle que $\pi f = 0$, $\sup_x |f(x)| \leq a$. Alors pour toute distribution initiale q , tout $t > 0$*

$$\left| P_q \left\{ \frac{1}{\sigma\sqrt{t}} \int_0^t f(X_s) ds \leq x \right\} - \Phi(x) \right| \leq N_q \frac{59a\|f\|^2}{\sigma^3\lambda_1^2 t^{1/2}} + \frac{N_q}{2\pi} e^{-\lambda_1 t},$$

où $\Phi(x)$ est la fonction de répartition de la loi normale standard et λ_1 la plus petite valeur propre non nulle de $(\Lambda + \Lambda^*)/2$.

Notons que l'hypothèse de mélange (3.1) devient dans le cas du temps continu

$$\sup \frac{\|P_t g\|}{\|g\|} \leq e^{-\lambda t}, \quad \forall t \geq 0 \text{ et } \lambda > 0.$$

Or, comme nous l'avons déjà remarquée, celle-ci est équivalente à l'existence d'un trou spectral du symétrisé additif du générateur Λ . Dans le cas d'un espace d'états fini et sous l'hypothèse d'irréductibilité du processus, λ peut être pris égal à la

plus petite valeur propre non nulle, λ_1 , du symétrisé additif de Λ . Cette hypothèse de mélange peut également permettre d'obtenir une borne du type Chernoff, sans recours à la théorie des perturbations [38][§4].

DÉMONSTRATION. Tout d'abord pour simplifier les expressions, introduisons les notations suivantes :

$$Y_t := \int_0^t f(X_s) ds, \quad \varphi_q(\omega) := E_q \exp\left(\frac{i\omega}{\sigma t^{1/2}} Y_t\right), \quad \Lambda(\omega) := \Lambda - \frac{i\omega}{\sigma t^{1/2}} D.$$

Notons que nous n'utilisons par le symétrisé additif du générateur du processus. L'idée de la preuve est de majorer

$$|\varphi_q(\omega) - e^{-\omega^2/2}|, \quad \omega \in \mathbb{R},$$

puis d'utiliser l'inégalité (3.13) dans [23][p. 538] qui s'écrit dans notre cas

$$(3.2) \quad \left| P_q \left\{ \frac{Y_t}{\sigma\sqrt{t}} \leq x \right\} - \Phi(x) \right| \leq \frac{1}{\pi} \int_{-W}^W \frac{|\varphi_q(\omega) - e^{-\omega^2/2}|}{|\omega|} d\omega + \frac{24}{\pi\sqrt{2\pi}W}.$$

Définissons la matrice fondamentale Z par l'identité (1.6) lorsque Z agit sur les fonctions de moyenne nulle ; nous en déduisons immédiatement que pour toute fonction f de moyenne nulle $\|Zf\| \leq \lambda_1^{-1}\|f\|$. La variance asymptotique de Y_t s'exprime également par $2\langle f, Zf \rangle$; en effet

$$\begin{aligned} t^{-1} \text{Var}_\pi Y_t &= 2 \int_0^\infty \left(1 - \frac{u}{t}\right) \langle f, P_u f \rangle du \\ &\rightarrow 2\langle f, \int_0^\infty P_u f du \rangle \\ &= 2\langle f, Zf \rangle. \end{aligned}$$

$\Lambda(r) = \Lambda - rD$ étant une perturbation linéaire de Λ , il est alors possible d'estimer la valeur propre $\beta_0(r)$ correspondant à la perturbation de la valeur propre nulle de Λ . $\beta_0(r)$ admet un développement en série dont les coefficients sont réels car s'exprimant par un produit scalaire ne faisant intervenir que f et Z (Z par définition étant à coefficients réels), donc $\beta_0(r)$ est réelle. Ceci peut aussi se déduire du théorème 2.6 dans [48][§ 2.3], puisque $B = -\Lambda(r)$ en tant que matrice ML (i.e telle que $B(x, y) \geq 0$, $x \neq y$) admet une valeur propre réelle majorant la partie réelle de toutes les autres valeurs propres de B . Les mêmes techniques, que dans le cas réversible, conduisent à la minoration suivante valable pour tout $r < 1/(4\|Z\|)$:

$$\beta_0(r) \geq -\frac{\sigma^2}{2} r^2 - \frac{b^2 \|Z\|}{2} \frac{2\|Z\|r}{1 - 4\|Z\|r}.$$

Notons par $\beta_0(\omega)$, la valeur propre de $\Lambda(\omega)$ correspondant à la perturbation de la valeur propre nulle de Λ . La théorie des perturbations, présentée au chapitre 2, s'applique pour une perturbation analytique à coefficient χ complexe. Nous pouvons donc utiliser une démarche analogue à celle du §1, où $\chi = i\omega/(\sigma t^{1/2})$. Celle-ci conduit, pour $|\omega| < \sigma t^{1/2}/(4a\|Z\|)$, à l'inégalité suivante :

$$\begin{aligned} \left| \beta_0(\omega)t - \frac{\omega^2}{2} \right| &\leq \frac{a\|f\|^2 |\omega|^3 \|Z\|^2}{\sigma^3 t^{1/2}} \left(1 - \frac{4a|\omega|\|Z\|}{\sigma t^{1/2}}\right)^{-1} \\ \left| \exp(-\beta_0(\omega)t) - e^{-\omega^2/2} \right| &\leq e^{-\omega^2/2} (\exp H(|\omega|) - 1), \end{aligned}$$

où $H(|\omega|)$ désigne le membre de droite de la première ligne.

Soit $u(\omega)$ le vecteur propre de $\Lambda(\omega)$ associé à la valeur propre $\beta_0(\omega)$ et normalisé par la condition $\langle 1, u(\omega) \rangle_\pi = 1$. En utilisant l'estimation 1-(4.26), nous obtenons

pour tout $|\omega|/\sigma t^{1/2} < 1/(4a\|Z\|)$ la majoration suivante :

$$\|\mathbf{1} - u(\omega)\| \leq \frac{4\|\omega\|\|Z\|}{\sigma t^{1/2}} \|D\mathbf{1}\| = \frac{4\|\omega\|\|f\|\|Z\|}{\sigma t^{1/2}}.$$

Majorons maintenant

$$|\varphi_q(\omega) - e^{-\omega^2/2}| \leq |\varphi_q(\omega) - \exp(-\beta_0(\omega)t)| + |\exp(-\beta_0(\omega)t) - e^{-\omega^2/2}|.$$

Or

$$\begin{aligned} |\varphi_q(\omega) - \exp(-\beta_0(\omega)t)| &\leq N_q \|\exp(-\Lambda(\omega)t) \cdot \mathbf{1} - \exp(-\beta_0(\omega)t) \cdot \mathbf{1}\| \\ &\leq N_q \left(\|e^{-\Lambda(\omega)t}(\mathbf{1} - u(\omega))\| + |e^{-\beta_0(\omega)t}| \|\mathbf{1} - u(\omega)\| \right) \end{aligned}$$

Le vecteur $\mathbf{1} - u(\omega)$ étant de moyenne nulle, un raisonnement identique à celui employé au §1 conduit à l'inégalité

$$\|e^{-\Lambda(\omega)t}(\mathbf{1} - u(\omega))\| \leq \|\mathbf{1} - u(\omega)\| e^{-\lambda_1(\omega)t},$$

où

$$\lambda_1(\omega) = \inf \left\{ \langle \tilde{\Lambda}(\omega)g, g \rangle / \|g\|^2; g \neq 0, \pi \cdot g = 0 \right\}.$$

Mais à la différence du cas traité au §1, le paramètre χ de la perturbation est ici un nombre complexe. Dans le cas général

$$\tilde{\Lambda}(\chi) = (\Lambda + \Lambda^*)/2 - \operatorname{Re}(\chi)D,$$

ce qui donne pour une perturbation purement imaginaire, $\tilde{\Lambda}(\omega) = (\Lambda + \Lambda^*)/2$ d'où $\lambda_1(\omega) = \lambda_1$. Nous obtenons ainsi

$$|\varphi_q(\omega) - \exp(-\beta_0(\omega)t)| \leq N_q \frac{4\|\omega\|\|f\|\|Z\|}{\sigma t^{1/2}} \left(e^{-\lambda_1 t} + e^{H(|\omega|) - \omega^2/2} \right).$$

Supposons dorénavant que

$$|\omega| < \frac{\sigma^3 t^{1/2}}{16a\|f\|^2\|Z\|^2},$$

hypothèse qui implique que

$$H(|\omega|) \leq \frac{2a\|f\|^2|\omega|^3\|Z\|^2}{\sigma^3 t^{1/2}} \leq \omega^2/8.$$

D'autre part pour tout $x \geq 0$, $e^x - 1 \leq xe^x$ d'où les inégalités suivantes :

$$\begin{aligned} |\varphi_q(\omega) - e^{-\omega^2/2}| &\leq N_q \frac{4\|f\|\|\omega\|\|Z\|}{\sigma t^{1/2}} \left(e^{-3\omega^2/8} \left\{ \frac{a\|f\|\omega^2\|Z\|}{2\sigma^2} + 1 \right\} + e^{-\lambda_1 t} \right) \\ (3.3) \quad &\leq N_q \frac{2a\|f\|^2\|Z\|^2}{\sigma^3 t^{1/2}} e^{-3\omega^2/8} (|\omega|^3 + 4|\omega|) + N_q \frac{4\|f\|\|\omega\|\|Z\|}{\sigma t^{1/2}} e^{-\lambda_1 t}, \end{aligned}$$

où nous avons utilisé dans la dernière inégalité le fait que

$$\sigma^2 \leq 2\|f\|^2\|Z\| \leq 2\|f\|^2/\lambda_1.$$

Pour conclure, appliquons l'inégalité (3.2) avec

$$W = \frac{\sigma^3 t^{1/2}}{16a\|f\|^2\|Z\|^2}.$$

□

3.1. Développement asymptotique. Dans le cas de variables aléatoires indépendantes, le théorème 3.1 peut se prolonger en une étude du développement asymptotique de la différence entre la fonction de répartition de la somme normalisée de ces variables et de $\Phi(x)$. Une présentation détaillée de ces résultats pour des variables aléatoires i.i.d. figure dans [27][§41] et pour des variables indépendantes dans [45]. Nagaev dans [42] a étendu ces résultats aux chaînes de Markov définies sur un espace d'états quelconques, mais satisfaisant à la condition de Doeblin (D_0) (cf Chap. 5). L'objectif de cette partie consiste à établir un tel développement asymptotique pour un processus markovien à espace d'états fini, mais toujours à partir des techniques développées dans [34].

Introduisons d'abord quelques définitions et rappels. Les polynômes d'Hermite sont définis par la formule

$$H_k(x) = (-1)^k e^{x^2/2} \frac{d^k}{dx^k} e^{-x^2/2},$$

qui donne pour les premiers indices

$$\begin{cases} H_0(x) &= 1 \\ H_1(x) &= x \\ H_2(x) &= x^2 - 1 \\ H_3(x) &= x^3 - 3x \\ H_4(x) &= x^4 - 6x^2 + 3 \\ H_5(x) &= x^5 - 10x^3 + 15x \end{cases}$$

Les fonctions $\Phi^{(k)}$ définies par

$$\Phi^{(k)}(x) = \frac{(-1)^{k-1}}{\sqrt{2\pi}} H_{k-1}(x) e^{-x^2/2},$$

admettent pour transformée de Fourier-Stieltjes

$$(3.4) \quad \int e^{i\omega x} d\Phi^{(k)}(x) = \frac{(-i\omega)^k}{\sqrt{2\pi}} \int e^{i\omega x - x^2/2} dx = (-i\omega)^k e^{-\omega^2/2}.$$

Par conséquent, si la fonction caractéristique

$$\varphi_q(\omega) := E_q \exp\left(\frac{i\omega}{\sigma t^{1/2}} Y_t\right)$$

peut s'exprimer formellement par la série

$$\varphi_q(\omega) \sim e^{-\omega^2/2} \left(1 + \sum_{j=1}^{\infty} P_j(-i\omega) \left(\frac{1}{\sigma t^{1/2}}\right)^j\right),$$

où les P_j sont des polynômes de degré $3j$, alors

$$(3.5) \quad P_q\left(\frac{1}{\sigma t^{1/2}} Y_t \leq x\right) - \Phi(x) \sim \sum_1^{\infty} P_j(-\Phi) \left(\frac{1}{\sigma t^{1/2}}\right)^j,$$

où les $P_j(-\Phi)$ s'obtiennent en remplaçant ω^r par $\Phi^{(r)}$ dans $P_j(-\omega)$.

Cherchons donc le développement en série de $\varphi_q(\omega)$ jusqu'à l'ordre $k-2$ pour un $k \geq 2$ fixé. Dans la suite, la notation (q, g) désignera pour toute fonction g la moyenne selon la mesure q ,

$$(q, g) := \sum_x q(x) g(x).$$

Soit $u(\omega)$ le vecteur propre de $\Lambda(\omega)$ pour la valeur propre $\beta_0(\omega)$, normalisé par la condition $\langle \mathbf{1}, u(\omega) \rangle = 1$. Ce vecteur propre admet pour développement en série

$$u(\omega) = \mathbf{1} + \sum_{j=1}^{\infty} u_j \left(\frac{-i\omega}{\sigma t^{1/2}} \right)^j,$$

convergente pour $|\omega| < \lambda_1 \sigma t^{1/2} / (4a)$. En particulier

$$u_1 = -Zf, \quad u_2 = ZDZf.$$

L'opérateur Λ étant autoadjoint par hypothèse, l'expression 1-(4.29) s'applique; d'où les inégalités suivantes :

$$(3.6) \quad \begin{aligned} u(\omega) - \mathbf{1} &\ll \frac{\|f\|}{4a} \sum_{j=1}^{\infty} \left(\frac{-i4a\omega}{\lambda_1 \sigma t^{1/2}} \right)^j \\ \|u_n\| &\leq \frac{\|f\|}{4a} \left(\frac{4a}{\lambda_1} \right)^n. \end{aligned}$$

De même, les estimations sur $\beta_0(\omega)$ conduisent aux formules suivantes :

$$\begin{aligned} -\beta_0(\omega)t &= -\frac{\omega^2}{2} + \sum_{j=1}^{\infty} \frac{(-\beta_0^{(j+2)})\omega^2}{\sigma^2} \left(\frac{-i\omega}{\sigma t^{1/2}} \right)^j, \\ e^{-\beta_0(\omega)t} &= e^{-\omega^2/2} \left\{ 1 + \sum_{j=1}^{\infty} Q_j(\omega^2) \left(\frac{-i\omega}{\sigma t^{1/2}} \right)^j \right\}, \end{aligned}$$

où les Q_j s'obtiennent par un développement en série de l'exponentielle selon les puissances de $(-i\omega)/(\sigma t^{1/2})$. Les fonctions $Q_j(x)$ sont des polynômes de degré j dont les coefficients dépendent exclusivement de $-\beta_0^{(3)}/\sigma^2, \dots, -\beta_0^{(j+2)}/\sigma^2$. Par exemple

$$\begin{cases} Q_1(x) &= -\frac{-\beta_0^{(3)}}{\sigma^2} x, \\ Q_2(x) &= -\frac{-\beta_0^{(4)}}{\sigma^2} x + \frac{1}{2} \left(\frac{-\beta_0^{(3)}}{\sigma^2} \right)^2 x^2. \end{cases}$$

Comme seul un développement limité à l'ordre $k-2$ nous intéresse, introduisons les notations suivantes :

$$\begin{aligned} u(\omega) &= \mathbf{1} + \sum_{j=1}^{k-2} u_j \left(\frac{-i\omega}{\sigma t^{1/2}} \right)^j + R_{k-1}(\omega) \\ e^{-\beta_0(\omega)t} &= e^{-\omega^2/2} \left\{ 1 + \sum_{j=1}^{k-2} Q_j(\omega^2) \left(\frac{-i\omega}{\sigma t^{1/2}} \right)^j + S_{k-1}(\omega) \right\} \end{aligned}$$

Ainsi la fonction caractéristique $\varphi_q(\omega)$ s'exprime par les relations

$$\begin{aligned} \varphi_q(\omega) &= e^{-\beta_0(\omega)t} (q, u(\omega)) + (q, P_t(\omega)(\mathbf{1} - u(\omega))) \\ &= e^{-\omega^2/2} \left\{ 1 + \sum_{j=1}^{k-2} \left(\sum_{\substack{j_1+j_2=j \\ 0 \leq j_1, j_2 \leq k-2}} Q_{j_1}(\omega^2)(q, u_{j_2}) \right) \left(\frac{-i\omega}{\sigma t^{1/2}} \right)^j \right\} \\ &\quad + V_{k-1}(\omega) + (q, P_t(\omega)(\mathbf{1} - u(\omega))), \end{aligned}$$

où $Q_0 = 1$, $u_0 = \mathbf{1}$ et

$$\begin{aligned} V_{k-1}(\omega) &= e^{-\omega^2/2} \sum_{j=k-1}^{2(k-2)} \left(\sum_{\substack{j_1+j_2=j \\ 1 \leq j_1, j_2 \leq k-2}} Q_{j_1}(\omega^2)(q, u_{j_2}) \right) \left(\frac{-i\omega}{\sigma t^{1/2}} \right)^j \\ &\quad + e^{-\beta_0(\omega)t} (q, R_{k-1}(\omega)) + e^{-\omega^2/2} S_{k-1}(\omega)(q, u(\omega)). \end{aligned}$$

Nous allons tout d'abord démontrer le lemme suivant :

LEMME 3.2. *Pour tout $k \geq 2$ et pour tout $|\omega| \leq \frac{\lambda_1^2 \sigma^3 t^{1/2}}{16a\|f\|^2}$, l'inégalité suivante est vérifiée*

$$(3.7) \quad \left| \varphi_q(\omega) - e^{-\omega^2/2} \left\{ 1 + \sum_{j=1}^{k-2} P_j(-i\omega) \left(\frac{1}{\sigma t^{1/2}} \right)^j \right\} \right| \leq N_q \left(\frac{4a\|f\|^2|\omega|}{\lambda_1^2 \sigma^3 t^{1/2}} \right) e^{-\lambda_1 t} + 2N_q e^{-3\omega^2/8} \left(\frac{8a\|f\|^2}{\lambda_1^2 \sigma^3 t^{1/2}} \right)^{k-1} T_{k-1}(\omega),$$

où les P_j sont des polynômes de degré $3j$ et

$$\begin{aligned} T_{k-1}(\omega) &= ((k-2)! + 1/4)(2|\omega|)^{k-1} + (|\omega|/2)^{3(k-1)}, \quad k \geq 3, \\ T_1(\omega) &= \frac{1}{2}(|\omega| + |\omega|^3) \end{aligned}$$

DÉMONSTRATION. Dans un premier temps, nous majorerons le reste $V_{k-1}(\omega)$. En particulier, grâce à l'inégalité de Cauchy-Schwarz, nous obtenons

$$\begin{aligned} |V_{k-1}(\omega)| &\leq N_q e^{-\omega^2/2} \sum_{j=k-1}^{2(k-2)} \left(\sum_{\substack{j_1+j_2=j \\ 1 \leq j_1, j_2 \leq k-2}} |Q_{j_1}(\omega^2)| \|u_{j_2}\| \right) \left| \frac{\omega}{\sigma t^{1/2}} \right|^j + \\ &\quad + N_q |e^{-\beta_0(\omega)t}| \|R_{k-1}(\omega)\| + N_q e^{-\omega^2/2} |S_{k-1}(\omega)| \|u(\omega)\|. \end{aligned}$$

La technique de la série majorante sera utile pour majorer $|S_{k-1}(\omega)|$. En effet, posons

$$U := -\beta_0(\omega)t + \frac{\omega^2}{2},$$

et choisissons ω tel que

$$(3.8) \quad |\omega| \leq T := \frac{\lambda_1^2 \sigma^3 t^{1/2}}{16a\|f\|^2}.$$

Nous obtenons alors les relations suivantes :

$$\begin{aligned} U &\ll \left(\frac{-ia\|f\|^2\omega^3}{\lambda_1^2 \sigma^3 t^{1/2}} \right) \left(1 - \frac{i4a\omega}{\lambda_1 \sigma t^{1/2}} \right)^{-1} \\ |U| &\leq \frac{2a\|f\|^2|\omega|^3}{\lambda_1 \sigma^3 t^{1/2}} \leq \frac{\omega^2}{8} \\ U^j &\ll \left(\frac{-ia\|f\|^2\omega^3}{\lambda_1^2 \sigma^3 t^{1/2}} \right)^j \sum_{l=0}^{\infty} \binom{l+j-1}{l} \left(\frac{-i4a\omega}{\lambda_1 \sigma t^{1/2}} \right)^l \end{aligned}$$

Supposons dans un premier temps que $k \geq 3$. La somme des termes du développement en série de U^j , pour $j \leq k-2$, contenant des puissances de $(\omega/(\sigma t^{1/2}))$ d'ordre supérieur à $k-1$ est majorée par

$$(3.9) \quad \left(\frac{\omega^2}{8} \right)^j \left(\frac{8a\|f\|^2|\omega|}{\lambda_1^2 \sigma^3 t^{1/2}} \right)^{k-1} \sum_{l=0}^{\infty} \binom{l+k-2}{j-1} \left(\frac{4a|\omega|}{\sigma \lambda_1 t^{1/2}} \right)^l.$$

Or $1 \leq j \leq k - 2$, d'où

$$\binom{l+k-2}{j-1} = \binom{l+k-2}{l} \frac{l!(k-2)!}{(j-1)!(l+k-1-j)!} \leq \binom{l+k-2}{l} \frac{(k-2)!}{(l+1)(j-1)!}.$$

Ceci conduit à la majoration de (3.9) par

$$(k-2)! \left(\frac{\omega^2}{8}\right)^j \left(\frac{16a\|f\|^2|\omega|}{\lambda_1^2\sigma^3t^{1/2}}\right)^{k-1}.$$

Si $k = 2$, ce terme est nul. Pour en tenir compte, remplaçons $(k-2)!$ par $(k-2)(k-3)!$, avec la convention $(-1)! = 0$. Nous en déduisons la majoration suivante de $|S_{k-1}(\omega)|$:

$$\begin{aligned} |S_{k-1}(\omega)| &\leq \sum_{s=k-1}^{\infty} \frac{1}{s!} \left(\frac{2a|\omega|^3\|f\|^2}{\lambda_1^2\sigma^3t^{1/2}}\right)^s + \\ &\quad + (k-2)(k-3)! \left(\frac{16a\|f\|^2|\omega|}{\lambda_1^2\sigma^3t^{1/2}}\right)^{k-1} \sum_{j=1}^{k-2} \left(\frac{\omega^2}{8}\right)^j \frac{1}{j!} \\ &\leq e^{\omega^2/8} \left(\frac{8a\|f\|^2}{\lambda_1^2\sigma^3t^{1/2}}\right)^{k-1} T_{k-1}^{(1)}(\omega), \end{aligned}$$

où

$$T_{k-1}^{(1)}(\omega) = \frac{1}{4^{k-1}(k-1)!} |\omega|^{3(k-1)} + (k-2)(k-3)!(2|\omega|)^{k-1}.$$

Une majoration des polynômes Q_j s'obtient en considérant la série majorante suivante de $\beta_0(\omega)$:

$$-\beta_0(\omega)t \ll -\frac{\omega^2}{2} + \frac{\|f\|^2(-i\omega)^2}{4\sigma^2\lambda_1} \left(\frac{-i4a\omega}{\lambda_1\sigma t^{1/2}}\right) \left(1 - \frac{-i4a\omega}{\lambda_1\sigma t^{1/2}}\right)^{-1},$$

d'où

$$e^{-\beta_0(\omega)t} \ll e^{-\omega^2/2} \left\{ 1 + \sum_{s=1}^{\infty} \left(\frac{-i4a\omega}{\lambda_1\sigma t^{1/2}}\right)^s \sum_{\substack{j+l=s \\ j \geq 1, l \geq 0}} \binom{j+l-1}{l} \frac{1}{j!} \left(\frac{\|f\|^2(-i\omega)^2}{4\lambda_1\sigma^2}\right)^j \right\}.$$

Ceci implique les relations suivantes :

$$\begin{aligned}
(3.10) \quad |Q_s(\omega^2)| &\leq \left(\frac{4a}{\lambda_1}\right)^s \sum_{\substack{j+l=s \\ j \geq 1, l \geq 0}} \binom{j+l-1}{l} \frac{1}{j!} \left(\frac{\|f\|^2 \omega^2}{4\lambda_1 \sigma^2}\right)^j \\
&= \left(\frac{4a}{\lambda_1}\right)^s \sum_{j=1}^s \binom{s-1}{s-j} \frac{1}{j!} \left(\frac{\|f\|^2 \omega^2}{4\lambda_1 \sigma^2}\right)^j \\
&= \left(\frac{4a}{\lambda_1}\right)^s \left(\frac{\|f\|^2 \omega^2}{4\lambda_1 \sigma^2}\right) \sum_{j=0}^{s-1} \binom{s-1}{j} \frac{1}{(j+1)!} \left(\frac{\|f\|^2 \omega^2}{4\lambda_1 \sigma^2}\right)^j \\
&\leq \left(\frac{4a}{\lambda_1}\right)^s \left(\frac{\|f\|^2 \omega^2}{4\lambda_1 \sigma^2}\right) \sum_{j=0}^{s-1} \binom{s-1}{j} \left(\frac{\|f\|^2 \omega^2}{8\lambda_1 \sigma^2}\right)^j \\
&= \left(\frac{a\|f\|^2 \omega^2}{\lambda_1^2 \sigma^2}\right) \left(\frac{4a}{\lambda_1} + \frac{a\|f\|^2 \omega^2}{2\lambda_1^2 \sigma^2}\right)^{s-1} \\
&\leq \left(\frac{a\|f\|^2 \omega^2}{\lambda_1^2 \sigma^2}\right) \left(\frac{8a\|f\|^2}{\lambda_1^2 \sigma^2} + \frac{a\|f\|^2 \omega^2}{2\lambda_1^2 \sigma^2}\right)^{s-1} \\
&= \left(\frac{a\|f\|^2}{\sigma^2 \lambda_1^2}\right)^s \omega^2 \left(8 + \frac{1}{2}\omega^2\right)^{s-1} \\
&\leq \frac{1}{2} \left(\frac{a\|f\|^2}{\sigma^2 \lambda_1^2}\right)^s \left(16^{s-1} \omega^2 + \omega^{2s}\right).
\end{aligned}$$

En particulier, nous obtenons pour les deux premiers polynômes :

$$\begin{cases} |Q_1(\omega^2)| &\leq \left(\frac{a\|f\|^2}{\sigma^2 \lambda_1^2}\right) \omega^2, \\ |Q_2(\omega^2)| &\leq \frac{1}{2} \left(\frac{a\|f\|^2}{\sigma^2 \lambda_1^2}\right)^2 (16\omega^2 + \omega^4). \end{cases}$$

Rappelons les majorations des $\beta_0^{(n)}$ pour $n \geq 3$,

$$\beta_0^{(n)} \leq \frac{a\|f\|^2}{\lambda_1^2} \left(\frac{4a}{\lambda_1}\right)^{n-3}.$$

Pour sa part, le reste $R_{k-1}(\omega)$ se majore facilement, puisque d'après (3.6) et (3.8)

$$\begin{aligned}
\|R_{k-1}(\omega)\| &\leq \frac{\|f\|}{2a} \sum_{j=k-1}^{\infty} \frac{1}{j+1} \left(\frac{4a|\omega|}{\sigma \lambda_1 t^{1/2}}\right)^j \\
&\leq \frac{1}{2} \left(\frac{8a\|f\|^2 |\omega|}{\lambda_1^2 \sigma^3 t^{1/2}}\right)^{k-1}.
\end{aligned}$$

Pour finir, il reste à majorer le terme

$$A := N_q e^{-\omega^2/2} \sum_{j=k-1}^{2(k-2)} \left(\sum_{\substack{j_1+j_2=j \\ 1 \leq j_1, j_2 \leq k-2}} |Q_{j_1}(\omega^2)| \|u_{j_2}\| \right) \left| \frac{\omega}{\sigma t^{1/2}} \right|^j.$$

Tout d'abord $A = 0$ pour $k = 2$. Supposons donc que $k \geq 3$, les majorations des polynômes $Q_j(\omega^2)$ et des $\|u_j\|$ obtenues ci-dessus permettent d'aboutir à une majoration de A , puisque :

$$A \leq N_q e^{-\omega^2/2} \frac{\omega^2}{4} \sum_{j=k-1}^{2(k-2)} \left(\frac{|\omega|}{\sigma t^{1/2}}\right)^j \sum_{\substack{j_1+j_2=j \\ 1 \leq j_1, j_2 \leq k-2}} \left(\frac{8a\|f\|^2}{\lambda_1^2 \sigma^2}\right)^{j_1} \left(1 + \frac{1}{16}\omega^2\right)^{j_1-1} \left(\frac{4a}{\lambda_1}\right)^{j_2}.$$

Or $4a/\lambda_1 \leq (8a\|f\|^2)/(\lambda_1^2\sigma^2)$, d'où

$$\begin{aligned} A &\leq N_q e^{-\omega^2/2} \frac{\omega^2}{32} \sum_{j=k-1}^{2(k-2)} \left(\frac{8a\|f\|^2|\omega|}{\lambda_1^2\sigma^3 t^{1/2}} \right)^j \sum_{\substack{j_1+j_2=j \\ 1 \leq j_1, j_2 \leq k-2}} \left(1 + \frac{\omega^2}{16}\right)^{j_1-1} \\ &\leq N_q (k-2) e^{-\omega^2/2} \frac{\omega^2}{32} \left(1 + \frac{\omega^2}{16}\right)^{k-3} \sum_{j=k-1}^{2(k-2)} \left(\frac{8a\|f\|^2|\omega|}{\lambda_1^2\sigma^3 t^{1/2}} \right)^j \\ &\leq N_q (k-2) e^{-\omega^2/2} \left(1 + \frac{\omega^2}{16}\right)^{k-1} \left(\frac{8a\|f\|^2|\omega|}{\lambda_1^2\sigma^3 t^{1/2}} \right)^{k-1} \sum_{j=0}^{\infty} 2^{-j} \\ &\leq (k-2) 2^{-1} N_q e^{-\omega^2/2} \left(\frac{8a\|f\|^2}{\lambda_1^2\sigma^3 t^{1/2}} \right)^{k-1} T_{k-1}^{(2)}(\omega), \end{aligned}$$

où

$$T_{k-1}^{(2)}(\omega) = (2|\omega|)^{k-1} + (|\omega|/2)^{3(k-1)}.$$

Tous ces calculs assez fastidieux conduisent à la majoration du reste $|V_{k-1}(\omega)|$, après avoir noté que

$$\left| \exp(-\beta_0(\omega)t + \omega^2/2) \right| \leq \exp(\omega^2/8), \quad \text{et } \|u(\omega)\| \leq \frac{5}{4}.$$

Nous obtenons avec $k \geq 2$,

$$\begin{aligned} |V_{k-1}(\omega)| &\leq N_q e^{-3\omega^2/8} \left(\frac{8a\|f\|^2}{\lambda_1^2\sigma^3 t^{1/2}} \right)^{k-1} \left(\frac{1}{2} |\omega|^{k-1} + \frac{5}{4} T_{k-1}^{(1)}(\omega) + 2^{-1} (k-2) T_{k-1}^{(2)}(\omega) \right) \\ &\leq 2N_q e^{-3\omega^2/8} \left(\frac{8a\|f\|^2}{\lambda_1^2\sigma^3 t^{1/2}} \right)^{k-1} T_{k-1}(\omega). \end{aligned}$$

Une majoration du dernier terme $(q, P_t(\omega)(\mathbf{1}-u(\omega)))$ s'obtient par une démarche analogue à celle utilisée au §3 ;

$$\begin{aligned} (q, P_t(\omega)(\mathbf{1}-u(\omega))) &\leq N_q e^{-\lambda_1 t} \|\mathbf{1}-u(\omega)\| \\ &\leq N_q \left(\frac{4a\|f\|^2|\omega|}{\lambda_1^2\sigma^3 t^{1/2}} \right) e^{-\lambda_1 t}. \end{aligned}$$

Posons

$$(3.11) \quad P_j(-i\omega) = (-i\omega)^j \sum_{\substack{j_1+j_2=j \\ 0 \leq j_1, j_2 \leq k-2}} Q_{j_1}(\omega^2)(q, u_{j_2}),$$

qui sont des polynômes de degré $3j$ en ω . En particulier :

$$\begin{cases} P_0(-i\omega) &= 1, \\ P_1(-i\omega) &= \frac{-\beta_0^{(3)}}{\sigma^2} (-i\omega)^3 - E_q(Zf)(-i\omega). \end{cases}$$

Les résultats précédents permettent d'obtenir l'inégalité (3.7). \square

Comme indiqué en introduction de cette section, nous en déduisons (3.5), où les P_j s'obtiennent à partir de (3.11). Soit $G_k(x)$ la fonction

$$G_k(x) = \Phi(x) + \sum_{j=1}^{k-2} P_j(-\Phi) \left(\frac{1}{\sigma t^{1/2}} \right)^j.$$

Cette fonction satisfait aux hypothèses du lemme 2 dans [23][§XV1.3]. La transformée de Fourier-Stieltjes de G étant la transformée de Fourier de sa dérivée, nous

obtenons par application de (3.2) (légèrement modifiée) avec $W = T^{k-1}$, l'inégalité suivante :

$$\begin{aligned} \left| P_q \left(\frac{Y_t}{\sigma t^{1/2}} \leq x \right) - G_k(x) \right| &\leq \frac{N_q}{2\pi} e^{-\lambda_1 t} + \frac{24m_k}{\pi W} + I \\ &+ \frac{2N_q}{\pi} \left(\frac{8a\|f\|^2}{\lambda_1^2 \sigma^3 t^{1/2}} \right)^{k-1} \int_{-\infty}^{\infty} \frac{T_{k-1}(\omega)}{|\omega|} e^{-3\omega^2/8} d\omega, \end{aligned}$$

où $m_k := \sup_x |G'_k(x)|$ et

$$I = \frac{1}{\pi} \int_{(-W, -T) \cup (T, W)} \frac{|\varphi_q(\omega) - \psi(\omega)|}{|\omega|} d\omega,$$

$\psi(\omega)$ désignant la transformée de Fourier de $G'_k(x)$. Bien entendu I sera prise nulle si $T \leq 1$; nous supposons donc sans perte de généralité que t est suffisamment grand pour avoir $T \geq 1$.

Les moments d'ordre pair d'une loi normale $\mathcal{N}(m, \sigma)$ s'expriment par la relation

$$\mu_s = (s-1)!! \sigma^s,$$

où

$$n!! := n(n-2)(n-4) \cdots .$$

D'autre part les coefficients

$$\bar{\mu}_s := \frac{1}{\sigma\sqrt{2\pi}} \int_0^{\infty} x^s e^{-\frac{x^2}{2\sigma^2}} dx,$$

pour s impair vérifient la relation de récurrence

$$\bar{\mu}_s = (s-1)\sigma^2 \bar{\mu}_{s-2}, \quad s > 1.$$

Or $\bar{\mu}_1 = \sigma/\sqrt{2\pi}$, d'où pour s impair

$$\bar{\mu}_s = \frac{(s-1)!! \sigma^s}{\sqrt{2\pi}}.$$

A titre d'exemple, considérons le cas $k = 3$; l'intégrale est majorée par 7.

LEMME 3.3. *Supposons $T \geq 1$, alors pour tout $k \geq 2$ nous avons l'inégalité suivante :*

$$m_k \leq \frac{N_q}{\sqrt{2\pi}} \left(1 + \frac{1}{2} (k-2)^{2(k-1)} \right).$$

DÉMONSTRATION. Remarquons tout d'abord que $\Phi^{(k)'}(x) = \Phi^{(k+1)}(x)$. Nous en déduisons l'expression suivante de la dérivée de G

$$G'_k(x) = \frac{e^{-x^2/2}}{\sqrt{2\pi}} + \sum_{j=1}^{k-2} \left(\frac{1}{\sigma t^{1/2}} \right)^j \sum_{\substack{j_1+j_2=j \\ 0 \leq j_1, j_2 \leq k-2}} Q_{j_1}^{(j)}(-\Phi)(q, u_{j_2}),$$

où $Q_{j_1}^{(j)}(-\Phi)$ s'obtient à partir de $Q_{j_1}(-\omega)$ en remplaçant les ω^r par les fonctions $\Phi^{(r+j+1)}$. Par exemple

$$\begin{aligned} G_3(x) &= \Phi(x) + \frac{e^{-x^2/2}}{\sqrt{2\pi}\sigma t^{1/2}} \left(E_q(Zf) - \frac{-\beta_0^{(3)}}{\sigma^2} H_2(x) \right) \\ G'_3(x) &= \frac{e^{-x^2/2}}{\sqrt{2\pi}} \left\{ 1 + \frac{-\beta_0^{(3)}}{\sigma^3 t^{1/2}} H_3(x) - E_q(Zf) H_1(x) \right\} \\ m_3(t) &\leq \frac{1}{\sqrt{2\pi}} + \frac{1}{\sqrt{2\pi}\sigma t^{1/2}} \left(\frac{N_q e^{-1/2} \|f\|}{\lambda_1} + \frac{3\|f\|^2}{2\sigma^2 \lambda_1^2} \right) \\ &\leq \frac{N_q}{\sqrt{2\pi}} \left(1 + \frac{4a\|f\|^2}{\lambda_1^2 \sigma^3 t^{1/2}} \right). \end{aligned}$$

Une majoration de m_k s'obtient par une estimation des bornes supérieures des fonctions $\Phi^{(k)}$. Ces fonctions étant continues, la formule d'inversion de la transformée de Fourier s'applique. D'après (3.4), nous obtenons les expressions suivantes :

$$\begin{aligned}\Phi^{(2m+1)}(x) &= \frac{(-1)^m}{2\pi} \int_{-\infty}^{\infty} e^{-u^2/2} u^{2m} \cos(ux) du \\ \Phi^{(2m)}(x) &= \frac{(-1)^m}{2\pi} \int_{-\infty}^{\infty} e^{-u^2/2} u^{2m-1} \sin(ux) du\end{aligned}$$

d'où pour $m \geq 1$,

$$\begin{aligned}|\Phi^{(2m+1)}(x)| &\leq \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u^{2m} \gamma(du) = \frac{(2m-1)!!}{\sqrt{2\pi}} = \frac{(2m)!}{2^m \sqrt{2\pi} m!}, \\ |\Phi^{(2m)}(x)| &\leq \frac{2}{\sqrt{2\pi}} \int_0^{\infty} u^{2m-1} \gamma(du) = \frac{(2m-2)!!}{\pi} = \frac{2^{m-1} (m-1)!}{\pi}.\end{aligned}$$

où $\gamma(dx)$ désigne la mesure gaussienne standard.

La double inégalité (1.10) conduit, pour $m > 0$, aux bornes suivantes

$$\begin{aligned}|\Phi^{(2m+1)}(x)| &\leq \frac{1}{\sqrt{\pi}} 2^{m+1/2} m^m e^{-m}, \\ |\Phi^{(2m)}(x)| &\leq \frac{1}{\sqrt{2\pi}} 2^{m+1/2} m^m e^{-m}.\end{aligned}$$

L'inégalité (3.10) permet de majorer les différents polynômes intervenant dans l'expression de $G'_k(x)$. En effet,

$$|Q_{j_1}^{(2m)}(-\Phi)| \leq \frac{2^{m+1/2}}{\sqrt{2\pi}} \left(\frac{4a}{\lambda_1}\right)^{j_1} \sum_{l=1}^{j_1} \binom{j_1-1}{l-1} \left(\frac{2\|f\|^2}{4\lambda_1\sigma^2}\right)^l \frac{1}{l!} (l+m)^{l+m} e^{-(l+m)}.$$

Or $l \leq j_1 \leq j = 2m$, d'où

$$\begin{aligned}|Q_{j_1}^{(2m)}(-\Phi)| &\leq \frac{2^{m-1/2} (3m)^m}{e^m \sqrt{2\pi}} \left(\frac{4a}{\lambda_1}\right)^{j_1} \sum_{l=1}^{j_1} \binom{j_1-1}{l+1} \left(\frac{3m\|f\|^2}{4e\lambda_1\sigma^2}\right)^l \\ &\leq \frac{2^{m-1/2} (2m)^m}{\sqrt{2\pi}} \left(\frac{4a}{\lambda_1}\right)^{j_1} \sum_{l=1}^{j_1} \binom{j_1-1}{l+1} \left(\frac{3m\|f\|^2}{4e\lambda_1\sigma^2}\right)^l \\ &= \frac{2^{m-1/2} (2m)^m}{\sqrt{2\pi}} \left(\frac{4a}{\lambda_1}\right)^{j_1} \left(\frac{3m\|f\|^2}{4e\lambda_1^2\sigma^2}\right) \left(1 + \frac{3m\|f\|^2}{4e\lambda_1\sigma^2}\right)^{j_1-1} \\ &\leq \frac{2^{m-1/2} (2m)^m}{\sqrt{2\pi}} \left(\frac{3am\|f\|^2}{e\lambda_1^2\sigma^2}\right)^{j_1} \left(1 + \frac{8e}{3m}\right)^{j_1-1} \\ &\leq \frac{2^{m-1/2} (2m)^m}{4e\sqrt{2\pi}} \left(\frac{12am\|f\|^2}{\lambda_1^2\sigma^2}\right)^{j_1}.\end{aligned}$$

Un calcul analogue pour $j = 2m + 1$ permet d'obtenir la majoration suivante pour tout $j \geq 1$:

$$|Q_{j_1}^{(j)}(-\Phi)| \leq \frac{K(j)}{\sqrt{2\pi}} \left(\frac{6aj\|f\|^2}{\lambda_1^2\sigma^2}\right)^{j_1},$$

où

$$4K(j) = 2^{(j-1)/2} j^{j/2}.$$

Ceci conduit à la majoration suivante de m_k ;

$$\begin{aligned}
\sqrt{2\pi}m_k &\leq 1 + \sum_{j=1}^{k-2} \left(\frac{1}{\sigma t^{1/2}} \right)^j \left\{ |Q_j^{(j)}(-\Phi)| + N_q \|u_j\| + N_q \sum_{\substack{j_1+j_2=j \\ 1 \leq j_1, j_2}} |Q_{j_1}^{(j)}(-\Phi)| \|u_{j_2}\| \right\} \\
&\leq 1 + N_q \sum_{j=1}^{k-2} \left(\frac{1}{\sigma t^{1/2}} \right)^j \left\{ K(j) \left(\frac{6aj\|f\|^2}{\lambda_1^2 \sigma^2} \right)^j + \frac{1}{4} \left(\frac{4a}{\lambda_1} \right)^j + \right. \\
&\quad \left. + ((j-1) \wedge 1) \frac{K(j)}{4} \sum_{\substack{j_1+j_2=j \\ 1 \leq j_1, j_2}} \left(\frac{6aj\|f\|^2}{\lambda_1^2 \sigma^2} \right)^{j_1} \left(\frac{4a}{\lambda_1} \right)^{j_2} \right\} \\
&\leq 1 + N_q \sum_{j=1}^{k-2} \left(\frac{8a\|f\|^2}{\lambda_1^2 \sigma^3 t^{1/2}} \right)^j \left\{ K(j)j^j + \frac{1}{4} + ((j-1) \wedge 1) \frac{K(j)}{4} \sum_{l=1}^{j-1} j^l \right\} \\
&\leq 1 + N_q \sum_{j=1}^{k-2} \left(\frac{8a\|f\|^2}{\lambda_1^2 \sigma^3 t^{1/2}} \right)^j \left\{ K(j)j^j + \frac{1}{4} + \frac{K(j)}{4} (j-1)j^{j-1} \right\} \\
&\leq N_q \left(1 + 2 \sum_{j=1}^{k-2} K(j) \left(\frac{8aj\|f\|^2}{\lambda_1^2 \sigma^3 t^{1/2}} \right)^j K(j) \right) \\
&\leq N_q \left(1 + (k-2)^{3(k-2)/2} \left(\frac{8a\|f\|^2}{\lambda_1^2 \sigma^3 t^{1/2}} \right)^{k-3} \sum_{j=0}^{k-3} \left(\frac{16a\|f\|^2}{\lambda_1^2 \sigma^3 t^{1/2}} \right)^j \right).
\end{aligned}$$

Or par hypothèse $T \geq 1$, d'où

$$m_k \leq \frac{N_q}{\sqrt{2\pi}} (1 + (k-2)^{2(k-1)}).$$

□

Il reste maintenant à majorer l'intégrale I . Par symétrie et l'inégalité triangulaire

$$I = 2 \int_T^W \frac{|\psi(\omega)|}{\omega} d\omega + 2 \int_T^W \frac{|\varphi_q(\omega)|}{\omega} d\omega := I_1 + I_2.$$

Par un calcul analogue à celui conduit pour majorer le terme A , nous obtenons :

$$\begin{aligned}
|\psi(\omega)| &\leq N_q e^{-\omega^2/2} \left(1 + \sum_{j=1}^{k-2} \left(\frac{8a\|f\|^2|\omega|}{\lambda_1^2 \sigma^3 t^{1/2}} \right)^j \left\{ \frac{1}{4} + \omega^2 \left(1 + \frac{\omega^2}{16} \right)^{j-1} + \right. \right. \\
&\quad \left. \left. + ((j-1) \wedge 1) \frac{\omega^2}{4} \sum_{l=1}^{j-1} \left(1 + \frac{\omega^2}{16} \right)^{l-1} \right\} \right) \\
&\leq N_q e^{-\omega^2/2} \left(1 + \sum_{j=1}^{k-2} \left(\frac{8a\|f\|^2|\omega|}{\lambda_1^2 \sigma^3 t^{1/2}} \right)^j \left\{ \frac{1}{4} + \omega^2 (k-2) \left(1 + \frac{\omega^2}{16} \right)^{j-1} \right\} \right),
\end{aligned}$$

d'où pour $\omega \geq T > 1$,

$$|\psi(\omega)| \leq N_q e^{-\omega^2/2} \left(1 + (k-1) \sum_{j=1}^{k-2} \left(\frac{8a\|f\|^2|\omega|}{\lambda_1^2 \sigma^3 t^{1/2}} \right)^j \left(1 + \frac{\omega^2}{16} \right)^j \right).$$

Or pour toute série géométrique de raison $q > 0$,

$$\sum_{j=1}^{k-2} q^j \leq (k-2)(q + q^{k-2}),$$

ce qui conduit, pour $k \geq 3$, à la majoration suivante :

$$|\psi(\omega)| \leq N_q e^{-\omega^2/2} \left(1 + 2(k-1)(k-2)(\omega^3/T)^{k-2}\right).$$

Ainsi, I_1 tend vers zéro à vitesse exponentielle lorsque $t \rightarrow \infty$.

Obtenir une borne explicite du terme I_2 est plus difficile, car nous ne pouvons utiliser le développement en série de la valeur propre perturbée, celle-ci n'étant plus convergente. Supposons par exemple que la v.a.r. Y_t admette une densité \mathcal{C}^∞ par rapport à la mesure de Lebesgue, dont les dérivées sont absolument intégrables, alors $|\omega^k \varphi_q(\omega)|$ décroît à l'infini pour tout $k \geq 0$. Par conséquent, $I_2 = O(T^{-(k-1)})$.

Supposons $k = 3$, nous obtenons alors, pour $T \geq 1$, l'inégalité suivante :

$$\left| P_q \left(\frac{Y_t}{\sigma t^{1/2}} \leq x \right) - G_3(x) \right| \leq \frac{N_q}{2\pi} e^{-\lambda_1 t} + \frac{36N_q}{2\pi} \left(\frac{16a\|f\|^2}{\lambda_1^2 \sigma^3 t^{1/2}} \right)^2 (1 + o(1)).$$

Signalons que contrairement à notre objectif, nous ne sommes pas arrivés à obtenir une borne explicite, faute de pouvoir éliminer le terme en $o(1)$.

4. Compléments

Le semi-groupe $P_t(r)$ obtenu par perturbation de P_t agit sur les fonctions réelles g de la façon suivante :

$$P_t(r)g(x) = E_x(\exp(rY_t)g(X_t)),$$

où $Y_t = \int_0^t f(X_s)ds$. Cette identité est appelée formule de Feynman-Kac [14][page 121], et dans le cas du mouvement brownien $P_t(1)$ s'apparente au semi-groupe de Feynman-Kac introduit dans [11]. Nous avons donc obtenu, pour $r < \lambda_1/2$, une majoration du rayon spectral logarithmique de ce semi-groupe. Le générateur de $P_t(r)$ est $-\Lambda(r) = -\Lambda + rf$, ce qui pour $r = 1$ conduit à l'analogue de l'opérateur de Schrödinger dans le cas du mouvement Brownien [11], à savoir l'opérateur $S = -\left(\frac{1}{2}\Delta + f\right)$, où Δ désigne le Laplacien dans \mathbb{R}^d . Soit $u(r)$ un vecteur propre associé à la valeur propre maximale $\lambda_0(r)$; nous pouvons alors introduire l'opérateur $Z(r)$, résolvante réduite de $\Lambda(r)$, qui s'exprime, sur le supplémentaire des fonctions engendrées par $u(r)$, par l'expression [11]

$$Z(r)g(x) = \int_0^\infty P_t(r)g(x)dt.$$

Par une démarche analogue à celle utilisée dans [34][IX-2.6], nous obtenons le développement en série suivant du semi-groupe $P_t(r)$:

$$P_t(r)g = P_t g + r g_t^{(1)} + r^2 g_t^{(2)} + \dots$$

avec

$$g_t^{(1)} = \int_0^t P_{t-s} f P_s g ds, \quad g_t^{(n)} = \int_0^t P_{t-s} f g_s^{(n-1)} ds.$$

Conformément à la définition de $P_t(r)$, un calcul, relativement simple, permet d'obtenir la relation suivante :

$$g_t^{(n)}(x) = E_x \left(\frac{Y_t^n}{n!} g(X_t) \right),$$

D'autre part,

$$\begin{aligned} \sup_x |g_t^{(n)}(x)| &\leq \|g\|_\infty \sup_x E_x \frac{|Y_t|^n}{n!}, \\ \|g_t^{(n)}\|_\pi^2 &\leq \|g\|_\infty^2 \binom{2n}{n} E_\pi \frac{|Y_t|^{2n}}{(2n)!}. \end{aligned}$$

Or d'après l'estimation obtenue dans [38][§4],

$$E_\pi \frac{|Y_t|^{2n}}{(2n)!} \leq 2^{-1} \left(\frac{2}{\lambda_1} \right)^{2n} \sum_{s=1}^n \frac{(\|f\|^2 \lambda_1 t)^s}{s!}.$$

Ceci permet avec l'inégalité (1.11) d'aboutir à la majoration suivante :

$$\|g_t^{(n)}\| \leq \|g\|_\infty e^{\|f\|^2 t \lambda_1 / 2} \left(\frac{4}{\lambda_1} \right)^n,$$

par conséquent le développement en série du semi-groupe P_t converge pour $r < \lambda_1/4$.

Il est possible d'obtenir l'expression de la variance asymptotique d'une manière différente de celle employée dans la preuve de la proposition 1.9. Pour cela utilisons l'identité suivante, valable pour toute fonction g de moyenne nulle :

$$\int_0^s P_u g du = Z \Lambda \int_0^s P_u g du = Z(g - P_s g).$$

Celle-ci implique que

$$\begin{aligned} E_\pi Y_t^2 &= 2 \int_0^t du_1 \langle f, \int_0^{t-u_1} P_{u_2} f du_2 \rangle \\ &= 2t \langle f, Zf \rangle - 2 \langle f, Z \int_0^t P_s f ds \rangle \\ &= 2t \langle f, Zf \rangle + 2 \langle f, Z^2 (P_t - I) f \rangle. \end{aligned}$$

Cette expression de la variance de Y_t en fonction de la matrice fondamentale Z est également valable dans le cas non réversible. Comme $\|P_t f\| \leq e^{-\lambda_1 t} \|f\|$, λ_1 désignant la plus petite valeur propre non nulle du symétrisé additif de Λ , nous en déduisons que la variance asymptotique s'exprime par $2 \langle f, Zf \rangle$. D'autre part, nous obtenons l'inégalité suivante :

$$|t^{-1} \text{Var}_\pi Y_t - \sigma^2| \leq \frac{2\|f\|^2}{\lambda_1^2 t} (e^{-\lambda_1 t} + 1),$$

valable également dans le cas non réversible.

L'objectif de la section 2 était d'obtenir une borne inférieure de la vitesse de convergence de la probabilité de déviation. Ceci est en général possible pour la vitesse de convergence en variation totale vers la mesure stationnaire. La théorie des grandes déviations fournit une borne inférieure de la vitesse asymptotique, mais les méthodes employées ne paraissent pas pouvoir être adaptées à notre problème. Malheureusement, le résultat obtenu n'est valable que pour des γ petits. Signalons que l'objectif de Kolmogorov, en établissant une telle borne pour les v.a.r. indépendantes, était d'établir une loi du logarithme itéré [49]. Une telle loi est valable pour certaines chaînes de Markov [10], d'où la question d'obtenir ce résultat par une démarche identique à celle menée par Kolmogorov, ainsi qu'une vitesse de convergence.

Chaîne de Markov

Les résultats démontrés dans le chapitre précédent sont encore valables pour les chaînes de Markov à espace d'états finis, et s'obtiennent par les mêmes méthodes. Cependant les calculs sont plus compliqués en raison de la périodicité de la chaîne et nécessitent de distinguer le cas réversible du cas non réversible. L'objet de ce chapitre consistera donc en l'étude d'une majoration de la quantité

$$P_q \left[n^{-1} \sum_{i=1}^n f(X_i) - \pi f \geq \gamma \right],$$

où P_q désigne la mesure de la chaîne de mesure initiale q . L'inégalité inférieure obtenue dans le cas d'un processus s'applique également aux chaînes de Markov réversibles. La preuve étant identique, elle ne sera pas reprise dans ce chapitre. Par contre, la vitesse de convergence vers la loi normale est étudiée en détail, à l'exception du développement asymptotique.

1. Borne du type Chernoff dans le cas réversible

Nous considérerons une chaîne de Markov (X_n) d'espace d'états G fini de cardinal N . Nous supposerons également que cette chaîne de Markov est irréductible et admet une probabilité réversible π . Comme dans le cas précédent, nous nous placerons dans $\ell^2(\pi)$ muni du produit scalaire défini par la mesure π ; le noyau de la chaîne de Markov, noté P , agit sur cet espace par la relation

$$Pf(x) := \sum_x P(x, y) f(y).$$

La chaîne étant réversible, le noyau P est autoadjoint et ses valeurs propres sont réelles. Nous les noterons par

$$1 = \beta_0 > \beta_1 \geq \dots \geq \beta_{N-1} \geq -1.$$

Le trou spectral du noyau P sera désigné par $\epsilon(P) := 1 - \beta_1$.

LEMME 1.1. *Pour tout $r > 0$,*

$$\begin{aligned} P_q \left[n^{-1} \sum_{i=1}^n f(X_i) \geq \gamma \right] &\leq e^{-rn\gamma} \langle q/\pi, P(r)^n \cdot \mathbf{1} \rangle \\ &\leq e^r N_q \exp \{ -n(r\gamma - \log \beta_0(P(r))) \} \end{aligned}$$

où $P(r)(x, y) = P(x, y)e^{rf(y)}$, et $\beta_0(r)$ désigne la plus grande valeur propre de $P(r)$.

La preuve s'obtient facilement à partir de l'inégalité de Markov et de l'expression de l'espérance de la somme en fonction de l'opérateur $P(r)$. La dernière ligne découle de l'inégalité de Cauchy-Schwarz, après avoir noté que $P(r)$ est semblable à la matrice autoadjointe $S(r) = \sqrt{E_r} P \sqrt{E_r}$, où $E_r = \text{diag}(\exp(rf(x)))$.

Pour obtenir une majoration de $\beta_0(P(r))$, nous utiliserons également la théorie des perturbations linéaires. En effet $P(r)$ peut s'interpréter comme une perturbation

de P puisque

$$P(r) = P + \sum_{i=1}^{\infty} \frac{r^i}{i!} P D^i,$$

où D a la même signification que dans le cas du temps continu. Donc pour $r < r_0$, $\beta_0(P(r))$ admet un développement en série de Taylor

$$\beta_0(P(r)) = 1 + \sum_{n=1}^{\infty} \beta^{(n)} r^n.$$

D'après l'expression (1-(4.28)), r_0 est supérieur à $\epsilon(P)/3$ et les termes $\beta^{(n)}$ s'expriment par les formules suivantes :

$$\beta^{(n)} = \sum_{p=1}^n \frac{(-1)^p}{p} \sum_{\substack{\nu_1 + \dots + \nu_p = n \\ k_1 + \dots + k_p = p-1 \\ \nu_i \geq 1, k_j \geq 0}} \frac{1}{\nu_1! \dots \nu_p!} \text{tr}[P D^{\nu_1} S^{(k_1)} \dots P D^{\nu_p} S^{(k_p)}]$$

où S désigne la résolvante réduite de l'opérateur markovien P pour la valeur propre 1 et

$$S^{(0)} := -\pi, \quad S^{(k)} := S^k, \quad k > 0.$$

Pour un p fixé, le p -uplet solution de l'équation $k_1 + \dots + k_p = p-1$ admet toujours une composante nulle. Soit $[k_1, \dots, k_p]$ la classe d'équivalence de (k_1, \dots, k_p) pour la relation introduite dans le cas du temps continu, alors en posant $Z = -S$ nous obtenons

$$(1.1) \quad \beta^{(n)} = \sum_{p=1}^n \sum_{\substack{\nu_1 + \dots + \nu_p = n \\ [0, k_1, \dots, k_{p-1}] \\ \nu_i \geq 1, k_j \geq 0}} \frac{1}{\nu_1! \dots \nu_p!} \langle f^{\nu_1}, Z^{(k_1)} P D^{\nu_2} \dots Z^{(k_{p-1})} P f^{\nu_p} \rangle,$$

car pour tout opérateur M

$$\text{tr}(\pi P D^{\nu_1} M D^{\nu_p}) = \text{tr}(\pi D^{\nu_1} M D^{\nu_p}) = \langle f^{\nu_1}, M f^{\nu_p} \rangle.$$

Notons que la matrice Z n'est autre que la matrice fondamentale de la chaîne de Markov

$$Z = \sum_{n=0}^{\infty} (P^n - \pi),$$

avec $P^0 := I$. L'hypothèse de réversibilité de la chaîne de Markov implique que Z est autoadjoint ; par conséquent, sa norme est identique à sa plus grande valeur propre, à savoir $\epsilon(P)^{-1}$. Les deux premiers coefficients du développement en série de $\beta_0(P(r))$ s'expriment facilement. Nous obtenons

$$\begin{aligned} \beta^{(1)} &= \langle f, 1 \rangle = 0, \\ \beta^{(2)} &= (1/2) \langle f, f \rangle + \langle f, Z P f \rangle \\ &= \langle f, Z f \rangle - (1/2) \langle f, f \rangle, \end{aligned}$$

en utilisant pour la dernière inégalité la relation $Z(I - P) = I - \pi$. Nous montrerons ultérieurement que $\beta^{(2)} = \sigma^2/2$ où σ^2 désigne la variance asymptotique, soit

$$\sigma^2 = \lim_{n \rightarrow \infty} n^{-1} E_{\pi} \left(\sum_{i=1}^n f(X_i) \right)^2,$$

résultat qui est donc analogue à celui obtenu dans le cas du temps continu.

Le nombre de termes de la somme de l'identité (1.1) est

$$\sum_{p=1}^n \binom{n-1}{p-1} \binom{2(p-1)}{p-1} \frac{1}{p},$$

par conséquent une majoration de $\beta^{(n)}$ est la suivante :

$$\begin{aligned} |\beta^{(n)}| &\leq \frac{b^2}{2^{n-1}} \sum_{p=1}^n \binom{n-1}{p-1} \binom{2(p-1)}{p-1} \frac{1}{p} \left(\frac{2}{\epsilon(P)} \right)^{p-1}, \\ &\leq \left(\frac{2}{\epsilon(P)} \right)^{n-1} b^2 C(n) \end{aligned}$$

où

$$C(n) = 2^{1-n} \sum_{p=1}^n \binom{n-1}{p-1} \binom{2(p-1)}{p-1} \frac{1}{p}.$$

Nous avons utilisé d'une part l'inégalité de Cauchy-Schwarz et l'inégalité $1/n! \leq 2^{1-n}$ et d'autre part le fait que $\epsilon(P) \leq 2$.

La majoration (2-(1.11)) donne pour tout $n \geq 3$

$$\begin{aligned} C(n) &< 2^{1-n} \left\{ 1 + \pi^{-1/2} \sum_{p=1}^{n-1} \binom{n-1}{p} \frac{4^p}{p+1} \right\} \\ &= 2^{1-n} \left\{ 1 + \pi^{-1/2} \left(\frac{5^n - 1}{4n} - 1 \right) \right\} \\ &\leq \frac{25}{8n} \left(\frac{5}{2} \right)^{n-2} \end{aligned}$$

Pour $n \geq 7$, nous avons $C(n) \leq (1/2)(5/2)^{n-2}$; un calcul direct montre que cette majoration reste valable pour $n = 3, 4, 5, 6$. En particulier $C(3) = 5/4$, $C(4) = 15/8$, $C(5) = 51/16$, $C(6) = 47/8$. Cette majoration de $C(n)$ implique que pour tout $n \geq 3$, $\beta^{(n)} \leq (b^2/5)(5/\epsilon(P))^{n-1}$. Par conséquent

$$\beta_0(r) \leq 1 + \frac{b^2}{\epsilon(P)} r^2 \left(1 - \frac{5r}{\epsilon(P)} \right)^{-1} \leq \exp \left\{ \frac{b^2}{\epsilon(P)} r^2 \left(1 - \frac{5r}{\epsilon(P)} \right)^{-1} \right\}.$$

La fonction

$$r \longrightarrow \gamma r - \frac{b^2}{\epsilon(P)} r^2 \left(1 - \frac{5r}{\epsilon(P)} \right)^{-1}$$

atteint son maximum en

$$r = \frac{\gamma \epsilon(P)}{b^2(1 + 5\gamma/b^2 + \sqrt{1 + 5\gamma/b^2})} < \frac{\epsilon(P)}{5}.$$

Pour cette valeur de r , la dernière inégalité du lemme 1.1 conduit au résultat suivant :

THÉORÈME 1.2. *Soit (P, π) un noyau markovien irréductible, réversible, défini sur un espace d'états G fini. Soit $f : G \rightarrow \mathbb{R}$ telle que $\pi f = 0$, $\sup_x |f(x)| \leq 1$ et $0 < \|f\|^2 \leq b^2$. Alors pour toute distribution initiale q , tout $n > 0$ et tout $0 < \gamma \leq 1$,*

$$(1.2) \quad P_q \left[n^{-1} \sum_{i=1}^n f(X_i) \geq \gamma \right] \leq e^{\epsilon(P)/5} N_q \exp \left[-\frac{n\gamma^2 \epsilon(P)}{4b^2(1 + h(5\gamma/b^2))} \right],$$

où $\epsilon(P) = 1 - \beta_1$ est le trou spectral de P et

$$h(x) = \frac{1}{2} (\sqrt{1+x} - (1-x/2)).$$

En particulier si $\gamma \leq 2b^2/5$,

$$P_q \left[n^{-1} \sum_{i=1}^n f(X_i) \geq \gamma \right] \leq e^{\epsilon(P)/5} N_q \exp \left[-\frac{n\gamma^2 \epsilon(P)}{4b^2} \left(1 - \frac{5\gamma}{2b^2} \right) \right],$$

si $\gamma > 2b^2/5$,

$$P_q \left[n^{-1} \sum_{i=1}^n f(X_i) \geq \gamma \right] \leq e^{\epsilon(P)/5} N_q \exp \left[-\frac{n\gamma\epsilon(P)}{10} \left(1 - \frac{2b^2}{5\gamma} \right) \right].$$

D'autre part si $\gamma \ll b^2$ et $\epsilon(P) \sim 0$ nous obtenons une borne asymptotique de l'ordre de

$$N_q \exp \left[-n\gamma^2\epsilon(P)/(4b^2) \right].$$

REMARQUE 1.3. Comme dans le cas du temps continu, la constante b^2 peut être majorée par 1, d'où la majoration suivante indépendante de la variance de f

$$P_q \left[n^{-1} \sum_{i=1}^n f(X_i) \geq \gamma \right] \leq e^{\epsilon(P)/5} N_q \exp \left[-\frac{n\gamma^2\epsilon(P)}{12} \right].$$

REMARQUE 1.4. Le théorème 1.2 appliqué à la fonction $-f$ donne la même majoration pour la déviation inférieure. Ceci implique l'inégalité suivante :

$$(1.3) \quad P_q \left[\left| n^{-1} \sum_{i=1}^n f(X_i) \right| \geq \gamma \right] \leq 3N_q \exp \left[-\frac{n\gamma^2\epsilon(P)}{4b^2(1+h(5\gamma/b^2))} \right].$$

car $\epsilon(P) \leq 2$.

PREUVE DE L'IDENTITÉ $\beta^{(2)} = \sigma^2/2$: Pour cela, considérons une base orthonormée $\{\varphi_i\}$ constituée des fonctions propres du noyau markovien P , associées respectivement aux valeurs propres β_i ; en particulier $\varphi_0 = 1$. Les vecteurs de cette base satisfont les relations suivantes :

$$\begin{aligned} P\varphi_i &= \beta_i\varphi_i \\ Z\varphi_i &= (1 - \beta_i)^{-1}\varphi_i, \quad i > 0 \\ E_\pi f(X_0)f(X_l) &= \langle f, P^l f \rangle = \sum_{i \geq 1} \beta_i^l \langle f, \varphi_i \rangle^2. \end{aligned}$$

Prouvons la proposition suivante [1][Chap. 4 prop. 29] :

PROPOSITION 1.5. Soit $S_n = \sum_{i=1}^n f(X_i)$ alors

$$\lim_{n \rightarrow \infty} n^{-1} \text{Var}_\pi S_n = \sigma^2 \leq 2\|f\|^2 \epsilon(P)^{-1},$$

où

$$\sigma^2 = \sum_{i \geq 1} \langle f, \varphi_i \rangle^2 \frac{1 + \beta_i}{1 - \beta_i} = 2\langle f, Zf \rangle - \|f\|^2.$$

D'autre part pour $n > 1$

$$n\sigma^2 \left(1 - \frac{2}{n\epsilon(P)} \right) \leq \text{Var}_\pi S_n \leq n\sigma^2 + \|f\|^2 \leq \frac{2n\|f\|^2}{\epsilon(P)}.$$

DÉMONSTRATION. Nous avons

$$\begin{aligned} n^{-1} \text{Var}_\pi S_n &= \|f\|^2 + 2 \sum_{l=1}^{n-1} \left(1 - \frac{l}{n} \right) E_\pi [f(X_0)f(X_l)] \\ &= \|f\|^2 + 2 \sum_{i \geq 1} \langle f, \varphi_i \rangle^2 \frac{\beta_i}{1 - \beta_i} - 2 \sum_{i \geq 1} \langle f, \varphi_i \rangle^2 \frac{\beta_i(1 - \beta_i^n)}{n(1 - \beta_i)^2} \end{aligned}$$

Le dernier terme de droite tend vers zéro quand n tend vers l'infini, donc $n^{-1}\text{Var}_\pi S_n$ converge vers σ^2 , où

$$\begin{aligned}\sigma^2 &= \|f\|^2 + 2 \sum_{i \geq 1} \langle f, \varphi_i \rangle^2 \frac{\beta_i}{1 - \beta_i} = \sum_{i \geq 1} \langle f, \varphi_i \rangle^2 \frac{1 + \beta_i}{1 - \beta_i} \\ &= 2\langle f, Zf \rangle - \|f\|^2 \leq 2\|f\|^2 \epsilon(P)^{-1}\end{aligned}$$

Notons que si $\beta_{N-1} = -1$ et $f = \varphi_{N-1}$ alors $\sigma^2 = 0$. La majoration de $\text{Var}_\pi S_n - n\sigma^2$ découle de l'inégalité

$$\inf_{-1 \leq \beta_i < 1} \frac{\beta_i(1 - \beta_i^n)}{(1 - \beta_i)^2} \geq -\frac{1}{2}.$$

Nous obtenons ainsi pour $n > 1$,

$$\begin{aligned}\text{Var}_\pi S_n &\leq 2n\langle f, Zf \rangle - (n-1)\|f\|^2 \\ &\leq 2n\|f\|^2 \epsilon(P)^{-1}\end{aligned}$$

Pour la minoration, nous suivons la preuve figurant dans [1]. Soit

$$C := \sup_{-1 \leq \beta_i < \beta_1} \frac{2\beta_i(1 - \beta_i^n)}{(1 - \beta_i)(1 + \beta_i)} = \frac{2\beta_1(1 - \beta_1^n)}{(1 - \beta_1)(1 + \beta_1)} \leq \frac{2}{1 - \beta_1},$$

alors

$$n^{-1}\text{Var}_\pi S_n = 2 \sum_{i \geq 1} \frac{\langle f, \varphi_i \rangle^2}{1 - \beta_i} B(\beta_i, n)$$

où

$$\begin{aligned}B(\beta_i, n) &= \frac{1 + \beta_i}{2} - \frac{\beta_i(1 - \beta_i^n)}{n(1 - \beta_i)} \\ &\geq \frac{1 + \beta_i}{2} \left(1 - \frac{C}{n}\right) \\ &\geq \frac{1 + \beta_i}{2} \left(1 - \frac{2}{n\epsilon(P)}\right),\end{aligned}$$

donc

$$\text{Var}_\pi S_n \geq n\sigma^2 \left(1 - \frac{2}{n\epsilon(P)}\right).$$

Notons également que

$$\sigma^2 \geq \|f\|^2 \frac{1 + \beta_{N-1}}{1 - \beta_{N-1}} \geq \frac{\|f\|^2}{2} (1 + \beta_{N-1}).$$

□

1.1. Algorithme de simulation. Supposons que nous voulions estimer la moyenne $\pi.f := \sum_x \pi(x)f(x)$ d'une fonction f définie sur un ensemble fini G . Une des méthodes consistera à construire une chaîne de Markov réversible de probabilité stationnaire π (algorithme de Métropolis par exemple), puis de la simuler. Le problème qui se pose alors est d'estimer le nombre de pas suffisant pour obtenir une certaine précision. L'idée est de trouver un temps d'arrêt U tel qu'à cet instant, X_U ait pour loi la distribution stationnaire, puis de simuler la chaîne à partir de l'état initial X_U . Dans le pire cas, $\tau_1^{(2)}$ (cf Remarque 1 section 2-(1.2)) est le temps moyen minimum pour obtenir un échantillon de distribution π .

Plus pratiquement, l'algorithme proposé dans [1][Chap. 4 §4.2] est le suivant : pour un nombre réel $t_1 > 0$ et un entier $n \geq 1$,

- générons une variable aléatoire $M(t_1)$ de loi de Poisson de paramètre t_1 ,

- simulons la chaîne X_n pendant $M(t_1) + n$ pas,
- calculons

$$S(t_1, n) = \sum_{i=M(t_1)+1}^{M(t_1)+n} f(X_i).$$

Introduisons pour un processus de Markov X_t , le paramètre de séparation suivant :

$$s(t) = \min \{s : P_t(x, y) \geq (1 - s)\pi(y); \text{ pour tout } x, y\}.$$

Ce paramètre fournit, dans un certain sens, une borne inférieure à la vitesse de convergence de $P(x, y)/\pi(y) \rightarrow 1$. Comme $s(\cdot)$ vérifie la relation $s(t+u) \leq s(t)s(u)$, nous pouvons définir le temps $\tau_1^{(1)}$ suivant [1]

$$\tau_1^{(1)} := \min\{t : s(t) \leq e^{-1}\}.$$

Nous utiliserons également la notation

$$\tau_2 = \epsilon(P)^{-1} = (1 - \beta_1)^{-1}.$$

La variable aléatoire $X_{M(t_1)}$ a pour loi

$$P_x(X_{M(t_1)} = y) = \sum_{k=0}^{\infty} e^{-t_1} \frac{t_1^k}{k!} P^k(x, y) = \exp(-(I - P)t_1)(x, y),$$

ce qui permet d'écrire l'identité suivante :

$$P_q(X_{M(t_1)} = y) = (1 - s(t_1))\pi(y) + s(t_1)\rho(y),$$

où ρ est une certaine probabilité. Il s'ensuit alors que

$$\begin{aligned} P_q(|S(t_1, n) - \pi \cdot f| \geq \|f\|^2 \gamma) &= \sum_x q(x) E_x(P_{X_{M(t_1)}}(|S(t_1, n) - \pi \cdot f| \geq \|f\|^2 \gamma)) \\ &\leq s(t_1) + (1 - s(t_1)) P_\pi(|S(t_1, n) - \pi \cdot f| \geq \|f\|^2 \gamma) \\ &\leq s(t_1) + P_\pi(|S(t_1, n) - \pi \cdot f| \geq \|f\|^2 \gamma) \end{aligned}$$

La définition de $\tau_1^{(1)}$ et l'inégalité (1.3) indiquent qu'il suffira de choisir

$$t_1 = \tau_1^{(1)} \log(2/\delta), \quad n = \left\lceil \frac{12\tau_2}{\gamma^2} \log(6/\delta) \right\rceil,$$

pour que le terme de gauche de l'inégalité précédente soit inférieur à $\delta > 0$. Cet algorithme consiste donc, avant de commencer le calcul de la moyenne empirique, à simuler la chaîne de Markov durant un temps aléatoire de loi de Poisson de paramètre t_1 . A la fin de cette phase "dite de chauffage", les échantillons de la chaîne ont une distribution proche de la mesure invariante π . Plus précisément, l'objectif cherché est de diminuer la constante N_q qui peut être très grande, mais égale à 1 lorsque $q = \pi$. Une autre méthode consisterait à évaluer le temps

$$\tau = \min\{n : \|(\pi_n/\pi) - 1\|_2 \leq 1/e\},$$

où π_n est la distribution de X_n pour la mesure initiale q [38]. Dans ce cas $N_n = \|\pi_n/\pi\| \leq (1 + 1/e^2)^{1/2}$, et

$$P_q\left[\left|n^{-1} \sum_{i=\tau+1}^{\tau+n} f(X_i)\right| \geq \gamma\right] \leq 2e^{-n\gamma^2\epsilon(P)/12}.$$

Bien entendu, cette méthode suppose la connaissance du noyau P de la chaîne de Markov et de la probabilité π , éventuellement à une constante multiplicative près. Cette connaissance permet alors d'estimer $\epsilon(P)$ et $\tau_1^{(1)}$; se reporter par exemple à [15] pour le cas des algorithmes de Metropolis.

REMARQUE 1.6. En combinant les lemmes 11, 12 et 23 dans [1] et l'inégalité (2-(1.27)) nous obtenons la double inégalité suivante :

$$\frac{e-1}{2e}\tau_2 \leq \tau_1^{(1)} \leq 4\tau_2\left(1 + \frac{1}{2}\log\frac{1}{\pi_*}\right).$$

1.2. Borne inférieure. Par une démonstration analogue à celle utilisée pour un processus markovien, nous obtenons le résultat suivant :

THÉORÈME 1.7. *Soit (P, π) un noyau markovien irréductible et réversible. Soit $f : G \rightarrow \mathbb{R}$ telle que $\|f\|_\infty \leq a$, $\pi \cdot f = 0$ et $\sigma^2 > 0$. Pour tout $\lambda > 0$, il existe des constantes $k(\lambda)$ (assez petite) et $K(\lambda)$ (assez grande) telles que : si $\gamma/a \leq k(\lambda)\left(\frac{\epsilon(P)\sigma^2}{2a^2}\right)^3$ et $n \geq K(\lambda)\left(\frac{\sigma}{\gamma}\right)^2$, alors*

$$P_\pi(n^{-1}S_n \geq \gamma) \geq \frac{\pi_*}{2} \exp\left\{-n\frac{\gamma^2}{2\sigma^2}(1+\lambda)\right\},$$

où $\pi_* := \min_x \pi(x)$. En particulier, $k(6) = 1,810^{-6}$.

2. Borne du type Chernoff dans le cas non réversible

Considérons à présent une chaîne de Markov (X_n) d'espace d'états fini, mais non nécessairement réversible. Le noyau P n'étant plus autoadjoint, nous utiliserons son symétrisé multiplicatif

$$K := P^*P,$$

où P^* désigne l'adjoint de P dans $\ell^2(\pi)$. Nous supposerons que le noyau K est également irréductible ; ainsi ses valeurs propres réelles seront de la forme :

$$1 = \beta_0 > \beta_1 \geq \dots \geq \beta_{N-1} \geq 0.$$

Nous voulons démontrer le théorème suivant qui est l'analogue au théorème 2-(1.2)

THÉORÈME 2.1. *Soit (P, π) un noyau markovien irréductible défini sur un espace d'états fini tel que son symétrisé multiplicatif soit irréductible. Soit $f : G \rightarrow \mathbb{R}$ telle que $\pi f = 0$, $\sup_x |f(x)| \leq 1$ et $0 < \|f\|^2 \leq b^2$. Alors pour toute distribution initiale q , tout $n > 0$ et tout $0 < \gamma \leq 1$,*

$$(2.1) \quad P_q\left[n^{-1}\sum_{i=1}^n f(X_i) \geq \gamma\right] \leq N_q \exp\left[-\frac{n\gamma^2\epsilon(K)}{8b^2(1+h(5\gamma/b^2))}\right],$$

où $\epsilon(K) = 1 - \beta_1$ est le trou spectral de K et

$$h(x) = \frac{1}{2}(\sqrt{1+x} - (1-x/2)).$$

En particulier si $\gamma \leq 2b^2/5$,

$$P_q\left[n^{-1}\sum_{i=1}^n f(X_i) \geq \gamma\right] \leq N_q \exp\left[-\frac{n\gamma^2\epsilon(K)}{8b^2}\left(1 - \frac{5\gamma}{2b^2}\right)\right],$$

et pour $\gamma > 2b^2/5$,

$$P_q\left[n^{-1}\sum_{i=1}^n f(X_i) \geq \gamma\right] \leq N_q \exp\left[-\frac{n\gamma\epsilon(K)}{20}\left(1 - \frac{2b^2}{5\gamma}\right)\right].$$

D'autre part si $\gamma \ll b^2$ nous obtenons une borne asymptotique de l'ordre de

$$N_q \exp[-n\gamma^2\epsilon(K)/(8b^2)].$$

PREUVE DU THÉORÈME 2.1. Le point de départ de la preuve est le lemme suivant

LEMME 2.2. Pour tout $r > 0$

$$(2.2) \quad P_q [n^{-1} S_n \geq \gamma] \leq e^{-rn\gamma} N_q \beta_0(r)^{n/2},$$

où $\beta_0(r)$ désigne la plus grande valeur propre de l'opérateur $K(r) = P^*(r)P(r)$

DÉMONSTRATION. De façon identique au cas réversible, nous avons

$$\begin{aligned} P_q [n^{-1} S_n \geq \gamma] &\leq e^{-rn\gamma} N_q \|P(r)^n\|_{2 \rightarrow 2} \\ &= e^{-rn\gamma} N_q \|P((r)^*)^n P(r)^n\|_{2 \rightarrow 2}^{1/2} \\ &= e^{-rn\gamma} N_q \sqrt{\beta_{0,n}(r)}, \end{aligned}$$

où $\beta_{0,n}$ désigne la plus grande valeur propre de $P((r)^*)^n P(r)^n$. Nous pouvons conclure grâce au théorème de Marcus [40][9.H.2a] qui stipule que $\beta_{0,n}(r) \leq \beta_0(r)^n$. \square

L'opérateur $K(r)$ peut s'interpréter comme une perturbation linéaire de K car

$$(2.3) \quad K(r) = K + \sum_{i=1}^{\infty} r^i \left(\sum_{j=0}^i \frac{1}{j!(i-j)!} D^j K D^{i-j} \right).$$

Nous en déduisons l'expression suivante des coefficients $\beta^{(n)}$ du développement en série de Taylor de $\beta_0(r)$

$$\beta^{(n)} = \sum_{p=1}^n \frac{(-1)^p}{p} \sum_{\substack{\nu_1 + \dots + \nu_p = n \\ k_1 + \dots + k_p = p-1 \\ \nu_i \geq 1, k_j \geq 0}} \text{tr} \left[T^{\nu_1} S^{(k_1)} \dots T^{\nu_p} S^{(k_p)} \right],$$

où

$$T^{(k)} = \sum_{j=0}^k \frac{1}{j!(k-j)!} D^j K D^{k-j}.$$

Pour un p fixé, le p -uplet solution de l'équation $k_1 + \dots + k_p = p - 1$ admet toujours une composante nulle, d'où l'expression suivante avec $Z = -S$:

$$(2.4) \quad \beta^{(n)} = \sum_{p=1}^n \sum_{\substack{\nu_1 + \dots + \nu_p = n \\ [0, k_1, \dots, k_{p-1}] \\ \nu_i \geq 1, k_j \geq 0}} \left(\sum_{j_1=0}^{\nu_1} \dots \sum_{j_p=0}^{\nu_p} \frac{1}{j_1! \dots j_p! (\nu_1 - j_1)! \dots (\nu_p - j_p)!} \right)$$

$$(2.5) \quad \left\langle f^{j_1}, K D^{(\nu_1 - j_1)} Z^{(k_1)} D^{j_2} K D^{(\nu_2 - j_2)} \dots Z^{(k_{p-1})} D^{j_p} K f^{(\nu_p - j_p)} \right\rangle,$$

En particulier,

$$\begin{aligned} \beta^{(1)} &= 0 \\ \beta^{(2)} &= \|f\|^2 + 2\langle f, ZKf \rangle + \langle f, Zf \rangle + \langle f, KZKf \rangle + \langle f, Kf \rangle \\ &= \|f\|^2 + \langle f, Zf \rangle + 3\langle f, ZKf \rangle \\ &= 4\langle f, Zf \rangle - 2\|f\|^2, \end{aligned}$$

la dernière inégalité étant obtenue par l'identité $Z(I - K) = I - \pi$.

MAJORATION DES $\beta^{(n)}$: Notons d'abord que si $j_1 = 0$, le terme figurant sous le signe somme devient

$$\langle f^{\nu_1}, Z^{(k_1)} D^{j_2} K D^{(\nu_2 - j_2)} \dots Z^{(k_{p-1})} D^{j_p} K f^{(\nu_p - j_p)} \rangle,$$

où $\nu_1 \geq 1$. Il en est de même pour j_p ; par conséquent l'inégalité de Cauchy-Schwarz conduit aux inégalités suivantes :

$$\begin{aligned} |\beta^{(n)}| &\leq \|f\|^2 \sum_{p=1}^n \frac{\epsilon(K)^{1-p}}{p} \sum_{\substack{\nu_1+\dots+\nu_p=n \\ k_1+\dots+k_p=p-1 \\ \nu_i \geq 1, k_j \geq 0}} \sum_{j_1=0}^{\nu_1} \dots \sum_{j_p=0}^{\nu_p} \frac{1}{j_1! \dots j_p! (\nu_1 - j_1)! \dots (\nu_p - j_p)!} \\ &= \|f\|^2 \sum_{p=1}^n \frac{1}{\epsilon(K)^{p-1}} \frac{1}{p} \sum_{\substack{\nu_1+\dots+\nu_p=n \\ k_1+\dots+k_p=p-1 \\ \nu_i \geq 1, k_j \geq 0}} \frac{2^{\nu_1+\dots+\nu_p}}{\nu_1! \dots \nu_p!} \\ &= 2^n \|f\|^2 \sum_{p=1}^n \frac{1}{\epsilon(K)^{p-1}} \frac{1}{p} \sum_{\substack{\nu_1+\dots+\nu_p=n \\ k_1+\dots+k_p=p-1 \\ \nu_i \geq 1, k_j \geq 0}} \frac{1}{\nu_1! \dots \nu_p!}. \end{aligned}$$

Il apparaît un terme identique à celui obtenu dans le cas réversible, au facteur 2^n près. Ceci conduit à la majoration suivante :

$$\begin{aligned} \beta^{(n)} &\leq 2b^2 \left(\frac{4}{\epsilon(K)} \right)^{n-1} C(n) \\ &\leq \frac{2b^2}{5} \left(\frac{10}{\epsilon(K)} \right)^{n-1}, \end{aligned}$$

d'où

$$\beta_0(r) \leq 1 + \frac{4b^2}{\epsilon(K)} r^2 \left(1 - \frac{10r}{\epsilon(K)} \right)^{-1}.$$

La fonction

$$r \longrightarrow \gamma r - \frac{2b^2}{\epsilon(K)} r^2 \left(1 - \frac{10r}{\epsilon(K)} \right)^{-1}$$

atteint son maximum en

$$r = \frac{\gamma \epsilon(K)}{2b^2 (1 + 5\gamma/b^2 + \sqrt{1 + 5\gamma/b^2})},$$

d'où la majoration désirée en portant cette valeur dans (2.2). \square

REMARQUE 2.3. Que peut-on dire de la vitesse de convergence lorsque le noyau K n'est pas irréductible ? L'introduction du symétrisé multiplicatif n'est pas nécessaire dans le cas d'un processus markovien ; par conséquent remplacer la chaîne par une version continue permettrait de s'affranchir du problème. Pour cela posons

$$Z_t = X_{N(t)},$$

où $N(t)$ désigne une v.a.r. indépendante de la chaîne (X_n) et de loi de Poisson de paramètre 1. Le processus $(Z_t, t \geq 0)$ est markovien de générateur infinitésimal $-(I - P)$. Le théorème 2-(1.1) fournit une majoration de la vitesse de convergence en fonction du trou spectral du symétrisé additif.

Les valeurs propres de P étant rangées dans l'ordre décroissant suivant :

$$1 > \operatorname{Re}(\beta_1) \geq \dots \geq \operatorname{Re}(\beta_{N-1}),$$

nous avons les inégalités [40][Ch.9 Sect.F]

$$\operatorname{Re}(\beta_1) \leq \beta_1 \left(\frac{P + P^*}{2} \right) \leq \beta_1 (P^* P)^{1/2}.$$

Néanmoins, je ne connais aucun résultat sur le lien direct entre les vitesses de convergence de ces deux processus.

REMARQUE 2.4. Soit $\rho(r)$ la valeur propre positive, maximale en module de la matrice $P(r)$, dont l'existence découle du théorème de Perron-Frobenius. D'après [13][théorème 3.1.2], le logarithme de cette valeur propre vérifie la relation suivante :

$$\log \rho(r) = \lim_{n \rightarrow \infty} n^{-1} \log E_q \exp(rS_n),$$

identité valable pour toute distribution initiale q .

L'application du théorème de Gärtner-Ellis aux chaînes de Markov à espace d'états fini [13][§3.1] conduit au principe des grandes déviations pour $n^{-1}S_n$, avec pour fonction de taux

$$I(x) = \sup_{r \geq 0} \{rx - \log \rho(r)\}.$$

3. Borne du type Berry–Esséen

Comme dans le cas du temps continu, il est possible d'obtenir une borne du type Berry–Esséen pour les chaînes de Markov en temps discret. Les étapes de la preuve seront identiques à celles utilisées dans le cas du temps continu. Les techniques de majorations s'inspireront de la preuve de B. Mann [39], excepté pour les estimations de la valeur propre perturbée $\beta_0(\omega)$ et de $\|1 - u(\omega)\|$, où $u(\omega)$ est un vecteur propre de $P(\omega)$ associé à la valeur propre $\beta_0(\omega)$.

3.1. Cas réversible : Supposons dans un premier temps la chaîne de Markov réversible, nous allons prouver le théorème suivant :

THÉORÈME 3.1. Soit (P, π) un noyau markovien irréductible, réversible, défini sur un espace d'états G fini. Soit $f : G \rightarrow \mathbb{R}$ telle que $\pi f = 0$, $\sup_x |f(x)| \leq a$ et $\sigma > 0$. Alors pour toute distribution initiale q , tout $n > 0$

$$\left| P_q \left\{ \frac{1}{\sigma\sqrt{n}} \sum_{i=1}^n f(X_i) \leq x \right\} - \Phi(x) \right| \leq N_q \frac{159a\|f\|^2}{\sigma^3\epsilon(P)^2} n^{-1/2},$$

où $\Phi(x)$ est la fonction de répartition de la loi normale et $\epsilon(P)$ le trou spectral de l'opérateur P .

DÉMONSTRATION. Pour simplifier les notations, posons $S_n = \sum_{i=1}^n f(X_i)$ et désignons $P(\frac{i\omega}{\sigma n^{1/2}})$, $\beta_0(\frac{i\omega}{\sigma n^{1/2}})$ et $E_q \exp(\frac{i\omega}{\sigma n^{1/2}} S_n)$ respectivement par $P(\omega)$, $\beta_0(\omega)$ et $\varphi_q(\omega)$. La preuve consistera à majorer $|E_q \exp(\frac{i\omega}{\sigma n^{1/2}} S_n) - e^{-\omega^2/2}|$, puis à utiliser l'inégalité (1-(3.2)), qui s'exprime ici sous la forme

$$(3.1) \quad \left| P_q \left\{ \frac{S_n}{\sigma\sqrt{n}} \leq x \right\} - \Phi(x) \right| \leq \frac{1}{\pi} \int_{-W}^W \frac{|\varphi_q(\omega) - e^{-\omega^2/2}|}{|\omega|} d\omega + \frac{24}{\pi\sqrt{2\pi}W}.$$

Les estimations obtenues dans la section précédente conduisent à la majoration suivante :

$$\left| \beta_0(\omega) - \left(1 - \frac{\omega^2}{2n}\right) \right| \leq \frac{5a\|f\|^2|\omega|^3}{\sigma^3 n^{3/2} \epsilon(P)^2} \left(1 - \frac{5a|\omega|}{\sigma n^{1/2} \epsilon(P)}\right)^{-1}.$$

D'après le théorème de la valeur moyenne, nous avons

$$\left| \beta_0^n(\omega) - \left(1 - \frac{\omega^2}{2n}\right)^n \right| \leq n\theta^{n-1} \left| \beta_0(\omega) - \left(1 - \frac{\omega^2}{2n}\right) \right|,$$

où θ est compris entre $\beta_0(\omega)$ et $1 - \omega^2/2n$. Choisissons ω de telle façon que

$$\frac{5a\|f\|^2|\omega|^3}{\sigma^3 n^{3/2} \epsilon(P)^2} \left(1 - \frac{5a|\omega|}{\sigma n^{1/2} \epsilon(P)}\right)^{-1} \leq \frac{\omega^2}{4n}.$$

D'après l'inégalité $\epsilon(P)\sigma^2 \leq 2\|f\|^2$, il suffira que

$$(3.2) \quad |\omega| \leq \frac{\epsilon(P)^2\sigma^3\sqrt{n}}{30a\|f\|^2} \leq \frac{\epsilon(P)\sigma\sqrt{n}}{15a}.$$

$\beta_0(\omega)$ est alors majoré par $1 - \omega^2/4n$, d'où l'inégalité

$$(3.3) \quad \left| \beta_0^n(\omega) - \left(1 - \frac{\omega^2}{2n}\right)^n \right| \leq \left(1 - \frac{\omega^2}{4n}\right)^{n-1} \frac{15a\|f\|^2|\omega|^3}{2\sigma^3n^{1/2}\epsilon(P)^2}.$$

Pour obtenir une majoration de $|\beta_0^n(\omega) - e^{-\omega^2/2}|$, nous avons besoin du lemme suivant dont la démonstration figure dans [39][lemme 3.4]

LEMME 3.2. *Pour tout $0 < y < n/2$,*

$$\begin{aligned} \left| e^{-y} - \left(1 - \frac{y}{n}\right)^n \right| &\leq \frac{y^2}{n} e^{-y}, \\ \left(1 - \frac{y}{n}\right)^{n-1} &\leq e^{-y/2}, \quad n \geq 1. \end{aligned}$$

D'après la condition (3.2), $\omega^2 \leq 2n\epsilon(P)/15^2 \leq n$. Par conséquent, la première inégalité du lemme 3.2 conduit, pour $y = \omega^2/2$, à la majoration :

$$\left| \left(1 - \frac{\omega^2}{2n}\right)^n - e^{-\omega^2/2} \right| \leq \frac{\omega^4}{4n} e^{-\omega^2/2} \leq \frac{|\omega|^3}{4\sqrt{n}} e^{-\omega^2/2},$$

tandis que la deuxième inégalité du même lemme appliquée à (3.3) donne

$$\left| \beta_0^n(\omega) - \left(1 - \frac{\omega^2}{2n}\right)^n \right| \leq \frac{15a\|f\|^2|\omega|^3}{2\sigma^3n^{1/2}\epsilon(P)^2} e^{-\omega^2/8}.$$

Nous en déduisons alors les inégalités

$$(3.4) \quad \begin{aligned} \left| \beta_0^n(\omega) - e^{-\omega^2/2} \right| &\leq e^{-\omega^2/8} \frac{|\omega|^3}{\sqrt{n}} \left(\frac{1}{4} + \frac{15a\|f\|^2}{2\sigma^3\epsilon(P)^2} \right) \\ &\leq e^{-\omega^2/8} \frac{9a\|f\|^2|\omega|^3}{\epsilon(P)^2\sigma^3\sqrt{n}}, \end{aligned}$$

où nous avons utilisé pour la dernière ligne, les inégalités $\sigma^2\epsilon(P) \leq 2\|f\|^2$, $\epsilon(P) \leq 2$ et $\|f\| \leq a$.

Soit $u(\omega)$ le vecteur propre de $P(\omega)$ associé à la valeur propre $\beta_0(\omega)$ et normalisé par la condition $\langle 1, u(\omega) \rangle_\pi = 1$. L'inégalité 1-(4.24) permet d'obtenir une majoration de $\|1 - u(\omega)\|$. En effet, prenons

$$\Phi_2(|\chi|) = \|Z\|(e^{a|\chi|} - 1), \quad \Phi_3(|\chi|) = e^{a|\chi|} - 1, \quad \Psi(|\chi|) = \beta_0(|\chi|) - 1$$

et utilisons les inégalités $a|\omega|/(\epsilon(P)\sigma n^{1/2}) \leq 1/15$ et $e^x - 1 \leq 4x/3$ pour $x \leq 1/2$. Nous obtenons après quelques calculs :

$$\|1 - u(\omega)\| \leq \frac{3a|\omega|}{2\epsilon(P)\sigma\sqrt{n}}.$$

D'autre part, notons que la matrice $P(\omega)$ n'est plus semblable à une matrice symétrique ; le paramètre de perturbation étant complexe. Cependant,

$$P(\omega) = E_{i\omega/2} (E_{-i\omega/2} P E_{i\omega/2}) E_{i\omega/2} := E_{i\omega/2} S(\omega) E_{i\omega/2},$$

où $S(\omega)$ est autoadjointe dans $\ell^2(\pi)$. Nous en déduisons que $\|P(\omega)\| \leq |\beta_{0,S}(\omega)|$; $\beta_{0,S}(\omega)$ désignant le rayon spectral de $S(\omega)$. Or, en raison de la continuité des valeurs propres perturbées, pour ω assez petit, $\beta_{0,S}(\omega)$ n'est autre que la perturbation

de la valeur propre 1 de P . Une estimation de celle-ci peut donc s'obtenir à partir des techniques du chapitre 1, appliquées à l'opérateur

$$S(\omega) = P + \sum_{l=1}^{\infty} \left(\frac{i\omega}{2\sigma n^{1/2}} \right)^l \sum_{j=0}^l \frac{(-1)^j}{j!(l-j)!} D^j P D^{l-j}.$$

En particulier, nous en déduisons l'inégalité suivante :

$$\left| \beta_{0,S}(\omega) - \left(1 - \frac{\omega^2}{2n}\right) \right| \leq \frac{5a\|f\|^2|\omega|^3}{\sigma^3 n^{3/2} \epsilon(P)^2} \left(1 - \frac{5a|\omega|}{\sigma n^{1/2} \epsilon(P)}\right)^{-1},$$

identique à celle obtenue, ci-dessus, pour $\beta_0(\omega)$.

Nous pouvons ainsi écrire

$$\begin{aligned} |\varphi_q(\omega) - \beta_0^n(\omega)| &\leq |\langle q/\pi, P^n(\omega)1 - \beta_0^n(\omega)1 \rangle| \\ &\leq N_q \|1 - u(\omega)\| (|\beta_0^n(\omega)| + |\beta_{0,S}(\omega)|) \\ &\leq 3N_q e^{-\omega^2/4} \frac{a|\omega|}{\epsilon(P)\sigma\sqrt{n}}, \end{aligned}$$

où nous avons utilisé dans la dernière ligne, l'inégalité $|\beta_0^n(\omega)| \leq (1 - \omega^2/4n)^n \leq e^{-\omega^2/4}$.

Les résultats obtenus impliquent les inégalités suivantes :

$$\begin{aligned} |\varphi_q(\omega) - e^{-\omega^2/2}| &\leq |\varphi_q(\omega) - \beta_0^n(\omega)| + |\beta_0^n(\omega) - e^{-\omega^2/2}| \\ &\leq 3N_q \frac{a\|f\|^2}{\epsilon(P)^2 \sigma^3 \sqrt{n}} e^{-\omega^2/8} \left(3|\omega|^3 + \frac{\epsilon(P)\sigma^2}{\|f\|^2} |\omega|\right) \\ (3.5) \quad &\leq 3N_q \frac{a\|f\|^2}{\epsilon(P)^2 \sigma^3 \sqrt{n}} e^{-\omega^2/8} (3|\omega|^3 + 2|\omega|). \end{aligned}$$

Il est maintenant possible de conclure la preuve du théorème 3.1, par combinaison des inégalités (3.13) dans [23][p. 538] et (3.5) avec

$$W = \frac{\epsilon(P)^2 \sigma^3 \sqrt{n}}{30a\|f\|^2}.$$

□

3.2. Cas non réversible : Dans le cas non réversible, une démarche identique fournit l'inégalité suivante :

$$|\beta_0^n(\omega) - e^{-\omega^2/2}| \leq e^{-\omega^2/8} \frac{9a\|f\|^2|\omega|^3\|Z\|^2}{\sigma^3 n^{1/2}},$$

à condition de supposer que

$$(3.6) \quad |\omega| \leq \frac{\sigma^3 n^{1/2}}{30a\|f\|^2\|Z\|^2}.$$

Nous avons d'autre part les relations

$$|\varphi_q(\omega) - \beta_0^n(\omega)| \leq N_q \|1 - u(\omega)\| \beta_0^n(\omega) + N_q \|P^n(\omega)(1 - u(\omega))\|.$$

Or pour tout x, y $|K(\omega)(x, y)| \leq K(x, y)$; comme d'autre part $1 - u(\omega)$ est centré pour la probabilité π , nous en déduisons que

$$\|P^n(\omega)(1 - u(\omega))\| \leq \beta_1(K)^{n/2} \|1 - u(\omega)\|,$$

d'où

$$|\varphi_q(\omega) - e^{-\omega^2/2}| \leq N_q e^{-\omega^2/8} \frac{3a|\omega|\|Z\|^2\|f\|^2}{\sigma^3 n^{1/2}} (1 + 3|\omega|^2) + N_q \frac{3a|\omega|\|Z\|^2\|f\|^2}{\sigma^3 n^{1/2}} \beta_1(K)^{n/2}.$$

Le trou spectral du noyau K , que nous supposons irréductible, vérifie la relation

$$\epsilon(K) = \inf \left\{ \langle (I - K)g, g \rangle, \pi.g = 0, \|g\| = 1 \right\}.$$

Ceci implique que $\|Pg\| \leq \sqrt{\beta_1(K)} := \mu$, d'où $\|Z\| \leq (1 - \mu)^{-1}$. Or $\epsilon(K) = 1 - \mu^2 \leq 2(1 - \mu)$, par conséquent

$$(3.7) \quad \|Z\| \leq \frac{2}{\epsilon(K)}$$

Appliquons alors l'inégalité (3.1) avec

$$W = \frac{\sigma^3 \epsilon(K)^2 n^{1/2}}{120a \|f\|^2},$$

pour obtenir le résultat suivant :

THÉORÈME 3.3. *Soit (P, π) un noyau markovien irréductible défini sur un espace d'états fini, tel que son symétrisé multiplicatif K soit irréductible. Soit $f : G \rightarrow \mathbb{R}$ telle que $\pi f = 0$, $\sup_x |f(x)| \leq a$ et $\sigma > 0$. Alors pour toute distribution initiale q , tout $n > 0$*

$$\left| P_q \left\{ \frac{1}{\sigma\sqrt{n}} \sum_{i=1}^n f(X_i) \leq x \right\} - \Phi(x) \right| \leq N_q \frac{615a \|f\|^2}{\sigma^3 \epsilon(K)^2} n^{-1/2} + \frac{N_q}{5\pi} (1 - \epsilon(K))^{n/2},$$

où $\Phi(x)$ est la fonction de répartition de la loi normale et $\epsilon(K)$ le trou spectral de l'opérateur K .

REMARQUE 3.4. La borne obtenue par B. Mann s'applique au cas d'une chaîne de Markov dénombrable satisfaisant à la condition

$$\beta := \sup \left\{ \frac{\|Pg\|}{\|g\|}; \pi.g = 0 \right\} < 1.$$

Or $\|Pg\| = \langle Kg, g \rangle^{1/2} \leq \sqrt{\beta_1(K)} \|g\|$, par conséquent $\mu < \beta$. Puisque $\epsilon(K) \geq 0$, nous en déduisons que $\epsilon(K) > (1 - \beta)$.

4. Compléments

Les estimations exponentielles pour la distribution d'une somme de v.a.r indépendantes ont été largement étudiées; se reporter par exemple aux §4,5 du chap. III dans [45] pour un aperçu des principaux résultats. L'inégalité de Bennett [5] permet d'obtenir le comportement Gaussien et Poissonien; elle stipule que pour une suite X_1, \dots, X_n de v.a.r. indépendantes centrées et bornées par $a > 0$, alors pour tout $0 < \gamma < a$

$$P(S_n \geq n\gamma) \leq \exp \left\{ -\frac{n\bar{\sigma}^2}{a^2} h\left(\frac{a\gamma}{\bar{\sigma}^2}\right) \right\},$$

où $h(x) = (1 + x) \log(1 + x) - x$ et $\bar{\sigma}^2 = n^{-1} \sum_{i=1}^n EX_i^2$. Or pour tout $x > 0$, $\log(1 + x) \geq x - x^2/2$. D'où $h(x) \geq x^2(1 - x)/2$, et la majoration suivante pour $\gamma \leq \bar{\sigma}^2/a$:

$$P(S_n \geq n\gamma) \leq \exp \left\{ -\frac{n\gamma^2}{2\bar{\sigma}^2} \left(1 - \frac{a\gamma}{\bar{\sigma}^2} \right) \right\}.$$

D'autre part, pour $x > 1$, $h(x) \geq x \log x - x$, d'où pour $\gamma \geq \bar{\sigma}^2/a$,

$$P(S_n \geq n\gamma) \leq \exp \left\{ -\frac{n\gamma}{a} \left(\log \left(\frac{a\gamma}{\bar{\sigma}^2} \right) - 1 \right) \right\}.$$

QUESTION : Peut-on obtenir une inégalité du type de Bennett pour les chaînes de Markov? Soit encore, avec les hypothèses du théorème 1.2, avons nous l'inégalité

$$P_q(S_n \geq n\gamma) \leq CN_q \exp \left\{ -\frac{n\sigma^2}{\|f\|_\infty^2} h\left(\frac{\|f\|_\infty \gamma}{\sigma^2}\right) \right\} ?$$

Heuristiquement, soient $x \in G$ fixé et $T(x) = \inf\{n > 0 : X_n = x\}$, le premier temps de retour en x . Supposons la chaîne de Markov apériodique, telle que $X_0 = x$, et considérons

$$S_n = \sum_{i=0}^{T_n(x)-1} f(X_i) = \sum_{k=1}^n \sum_{i=T_{k-1}(x)}^{T_k(x)-1} f(X_i) := \sum_{k=1}^n Y_k(x),$$

où les $T_n(x)$ désignent les temps de retour successifs en x et $T_0(x) = 0$, par convention. Les $Y_k(x)$ sont des variables aléatoires i.i.d., dont les deux premiers moments s'expriment par :

$$E_x Y_k(x) = \pi(x)^{-1} \pi \cdot f, \quad \text{Var}_x Y_k(x) = \pi(x)^{-1} \sigma^2,$$

(ceci peut s'obtenir en particulier par le théorème 3.5 et la proposition 4.6 (iii) dans [46]). Notons que les identités précédentes sont encore valables pour toute distribution initiale de la chaîne, à condition de supposer $k \geq 2$. La chaîne de Markov étant récurrente, $T(x)$ est p.s. finie. L'inégalité de Bennett s'applique et nous obtenons, pour f de moyenne nulle, la majoration suivante :

$$P_x(S_n \geq n\gamma) \leq P_x(T(x) \geq \pi(x)^{-1/2}) + \exp\left\{-\frac{n\sigma^2}{\|f\|_\infty^2} h\left(\frac{\|f\|_\infty \gamma}{\sigma^2}\right)\right\}.$$

De la même façon, $\sqrt{\pi(x)}S_n/(\sigma\sqrt{n})$ converge en loi vers une gaussienne standard, et si $E|S_n|^3 < \infty$, l'inégalité de Berry–Esséen pour les v.a.r. i.i.d. donne [9] :

$$\begin{aligned} \sup_t \left| P_q \left[\frac{\sqrt{\pi(x)}S_n}{\sigma\sqrt{n}} \leq t \right] - \Phi(t) \right| &\leq \frac{CE_q|S_n|^3}{\text{Var}_q[S_n]^{3/2}\sqrt{n}} \\ &\leq \frac{C\|f\|_\infty^3 \pi(x)^{3/2} E[T(x)^3]}{\sigma^3 \sqrt{n}} \end{aligned}$$

où $C \leq 0,7915$. Ce résultat, nous permet d'envisager une éventuelle amélioration de la constante numérique présente dans la borne obtenue au théorème 3.1.

Signalons les travaux de E. Bolthausen sur la borne du type Berry–Esséen pour les chaînes de Markov à espace d'états dénombrable [6]. Il obtient le résultat suivant. Soient x un état quelconque, $T_x = \inf\{n \geq 0 : X_n = x\}$ et τ le temps de récurrence entre deux passages successifs par x ; si

$$\begin{aligned} E_x(\tau^3) < \infty \quad E_x\left(\sum_{i=1}^{\tau} |f|(X_i)\right)^3 < \infty, \\ E_q(T_x) < \infty \quad E_q\left(\sum_{i=1}^{T_x} |f|(X_i)\right) < \infty, \end{aligned}$$

alors,

$$\sup_n \left| P_q \left(\frac{\sqrt{\alpha}}{\sigma n^{1/2}} \sum_{i=1}^n (f(X_i) - \pi f) \leq t \right) - \Phi(t) \right| = O(n^{-1/2}),$$

où $\alpha := E_x \tau$, et $\sigma^2 := E_x \left(\sum_{i=1}^{\tau} (f(X_i) - \pi f) \right)^2 > 0$.

Deuxième partie

Espace d'états quelconque

Perturbation des opérateurs linéaires

Pour étendre aux processus markoviens à espace d'états quelconque, les bornes obtenues précédemment, la solution la plus rapide sera d'étudier sous quelles conditions étendre au cas d'un espace quelconque, les techniques de perturbation des opérateurs linéaires en dimension finie. Lors de l'estimation des coefficients du développement en série d'une valeur propre perturbée simple, le rôle primordial joué par la distance entre la valeur propre non perturbée et son complémentaire dans le spectre, nous conduit à une première condition ; la valeur propre non perturbée devra être isolée. Ceci n'est pas vrai en général pour un opérateur linéaire défini sur un espace quelconque. Une deuxième condition provient des estimations elles-mêmes, puisqu'elles font appel à la trace d'un produit d'opérateurs dans lequel figure toujours le projecteur propre associé à la valeur propre non perturbée (cf I-(4.13)). Il faudra donc pouvoir définir la trace de tels opérateurs ; un moyen étant que le projecteur propre soit de rang fini et il en sera alors de même pour les opérateurs intervenant dans les diverses estimations. Signalons que la dernière hypothèse est naturelle pour une chaîne de Markov ergodique, lorsque la valeur propre simple considérée est $\lambda = 1$. La première correspond à la notion de trou spectral.

L'objectif de ce chapitre consistera à préciser ces notions et à prouver que les hypothèses intuitives, énoncées ci-dessus, suffisent pour étendre les résultats du chapitre 1 aux opérateurs linéaires définis sur un espace de Banach général. Nous introduirons également la notion d'opérateurs fermés, car nous aborderons la perturbation du générateur infinitésimal d'un semi-groupe. Ceci nous sera utile, lorsque nous étudierons le cas d'un processus markovien.

Les résultats présentés dans ce chapitre proviennent presque exclusivement du livre de Kato [34].

1. Opérateurs fermés

1.1. Définitions. Soient X et Y des espaces de Banach et T un opérateur linéaire de X dans Y . Comme T n'est pas nécessairement défini sur tout l'espace X , on désignera par $D(T)$ le sous-espace vectoriel sur lequel il est défini ; il sera appelé domaine de définition, ou simplement domaine de T . Son image sera notée $R(T)$. Si S et T sont deux opérateurs linéaires de X dans Y tels que $D(S) \subset D(T)$ et $Su = Tu$ pour tout $u \in D(S)$, alors T sera appelé une extension de S et S une restriction de T . Cette relation sera notée $T \supset S$ ou $S \subset T$.

Pour tout sous-ensemble M de X , l'image de $M \cap D(T)$ sera notée TM ; pour tout sous-ensemble M' de Y , l'image inverse $T^{-1}M'$ sera l'ensemble des vecteurs $u \in D(T)$ tels que $Tu \in M'$. Si l'opérateur T est bijectif, ce qui est équivalent à $Tu = 0$ implique que $u = 0$, l'opérateur inverse T^{-1} est défini par

$$\begin{aligned} D(T^{-1}) &= R(T), & R(T^{-1}) &= D(T) \\ T^{-1}(Tu) &= u, \quad u \in D(T), & T(T^{-1}v) &= v, \quad v \in R(T). \end{aligned}$$

Nous noterons dans la suite par $\mathcal{B}(X, Y)$ l'espace de Banach des opérateurs linéaires bornés définis sur tout X (donc $D(T) = X$) à valeurs dans Y . Ce sont des opérateurs continus; la continuité en $u_0 \in D(T)$ étant définie pour un opérateur linéaire T par la condition suivante : si pour toute suite $u_n \in D(T)$ telle que $\|u_n - u_0\| \rightarrow 0$ alors $\|Tu_n - Tu_0\| \rightarrow 0$.

Un opérateur $P \in \mathcal{B}(X) := \mathcal{B}(X, X)$ idempotent ($P^2 = P$) est appelé projecteur. Nous avons alors la décomposition

$$X = M \oplus N,$$

où $M = PX$ et $N = (I - P)X$. Les sous-espaces vectoriels M et N sont fermés dans X . De façon réciproque, une décomposition de X en somme directe de deux sous-espaces vectoriels fermés M et N permet de définir le projecteur P sur M parallèlement à N . P est un opérateur borné (voir ci-dessous).

Parmi les opérateurs non bornés, il existe une classe importante pour les applications, à savoir celle des opérateurs fermés. Soit T un opérateur de X dans Y , une suite $u_n \in D$ sera dite T -convergente vers $u \in X$ si $\{u_n\}$ et $\{Tu_n\}$ sont des suites de Cauchy et que $u_n \xrightarrow{T} u$; nous noterons cette condition par $u_n \xrightarrow{T} u$. T sera dit fermé si $u_n \xrightarrow{T} u$ implique que $u \in D(T)$ et $Tu = \lim Tu_n$; soit encore si pour toute suite $u_n \in D(T)$, telle que $u_n \rightarrow u$ et $Tu_n \rightarrow v$, alors u appartient à $D(T)$ et $Tu = v$. La différence entre un opérateur fermé et un opérateur borné est la suivante : si T est borné alors l'existence d'une suite $u_n \in D(T)$ convergente implique l'existence de $\lim Tu_n$, tandis que si T est seulement fermé la convergence de la suite $u_n \in D(T)$ n'implique pas en général la convergence de la suite Tu_n . Cependant si T est fermé et si $u_n \in D(T)$ et $v_n \in D(T)$ sont deux suites qui convergent vers la même limite, alors les suites Tu_n et Tv_n ne peuvent converger vers des limites différentes. Désignons par $\mathcal{C}(X, Y)$ (resp. $\mathcal{C}(X)$) le sous-espace des opérateurs fermés de X dans Y (resp. de X dans X). Notons qu'un opérateur borné est fermé si et seulement si $D(T)$ est fermé dans X car la continuité de T implique que $Tu = v$. En particulier tout $T \in \mathcal{B}(X, Y)$ est fermé.

La propriété pour un opérateur T d'être fermé peut se traduire par une propriété sur le graphe de T . Pour cela considérons l'espace produit $X \times Y$ muni de la norme $\|(u, v)\| = (\|u\|^2 + \|v\|^2)^{1/2}$; cet espace est un espace de Banach. Le graphe $G(T)$ d'un opérateur T de X dans Y est le sous-ensemble des (u, Tu) pour tout $u \in D(T)$. $G(T)$ est un sous-espace vectoriel de $X \times Y$. Nous avons alors la propriété que T est un opérateur fermé si et seulement si $G(T)$ est un sous-espace fermé de $X \times Y$; en effet T est fermé si et seulement si (u_n, Tu_n) est une suite de Cauchy dans $X \times Y$. Le théorème suivant [34][th. III-5.20] énonce une condition suffisante pour qu'un opérateur fermé soit borné.

THÉORÈME 1.1. *Un opérateur fermé T de X dans Y de domaine X est borné. Autrement dit, $T \in \mathcal{C}(X, Y)$ et $D(T) = X$ impliquent que $T \in \mathcal{B}(X, Y)$.*

Comme application de ce théorème citons la preuve qu'un projecteur P sur M parallèlement à N est borné, en prouvant que P est fermé [34][p.167].

1.1.1. *Commutativité et décomposition.* Deux opérateurs $S, T \in \mathcal{B}(X)$ commutent si $ST = TS$. En raison des domaines respectifs de chaque opérateur, cette définition ne peut pas être prolongée facilement aux opérateurs non bornés. Une extension partielle est la suivante : un opérateur T dans X commute avec un opérateur $A \in \mathcal{B}(X)$ si

$$AT \subset TA.$$

Cela signifie donc que dès que $u \in D(T)$, Au appartient à $D(T)$ et $TAu = ATu$. De la même façon la notion de sous-espace M invariant par T (i.e $TM \subset M$) est difficile à étendre car il se peut que $M \cap D(T) = \emptyset$. Cependant la notion de

décomposition de T par deux sous-espaces M et N supplémentaires l'un de l'autre peut être étendue. T sera dit décomposable selon $X = M \oplus N$ si

$$PD(T) \subset D(T), \quad TM \subset M, \quad TN \subset N,$$

où P est le projecteur sur M parallèlement à N . Cette définition est équivalente à la condition que T commute avec P . Lorsque T se décompose comme précédemment, nous pouvons définir les restrictions T_M et T_N de l'opérateur T dans M et N . T_M est alors un opérateur linéaire dans l'espace de Banach M avec $D(T_M) = D(T) \cap M$, tel que $T_M u = Tu \in M$. Si T est fermé, il en sera de même de T_M et de T_N , car $G(T_M)$ est l'intersection de $G(T)$ avec l'ensemble fermé $M \times M$.

1.1.2. *Résolvante et spectre.* Supposons que $T \in \mathcal{C}(X)$, il en sera alors de même pour $T - \zeta$ pour tout $\zeta \in \mathcal{C}$. Si $T - \zeta$ est inversible avec

$$R(\zeta) = R(\zeta, T) = (T - \zeta)^{-1} \in \mathcal{B}(X),$$

nous dirons que ζ appartient à l'ensemble résolvant de T (noté $P(T)$). La fonction $R(\zeta)$ a valeurs dans $\mathcal{B}(X)$ est la résolvante de T . Ainsi pour tout $\zeta \in P(T)$, $R(\zeta)$ a pour domaine l'espace X et pour image $D(T)$. La résolvante vérifie l'équation suivante valable pour tout ζ_1, ζ_2 dans $P(T)$:

$$(1.1) \quad R(\zeta_1) - R(\zeta_2) = (\zeta_1 - \zeta_2)R(\zeta_1)R(\zeta_2).$$

La résolvante d'un opérateur linéaire possède les propriétés suivantes [34][th. 6.5 et 6.7] :

THÉORÈME 1.2. *Supposons que $P(T)$ soit non vide. Pour que T commute avec $A \in \mathcal{B}(X)$, il est nécessaire que*

$$R(\zeta)A = AR(\zeta),$$

pour tout $\zeta \in P(T)$, et il est suffisant que l'égalité ci-dessus soit vérifiée pour au moins un $\zeta \in P(T)$.

L'équation (1.1) conduit à la même série de Neumann que celle obtenue dans le cas de la dimension finie

$$R(\zeta) = \sum_{n=0}^{\infty} (\zeta - \zeta_0)^n R(\zeta_0)^{n+1},$$

[34][p.173-174]. Ceci conduit au théorème suivant :

THÉORÈME 1.3. *$P(T)$ est un ensemble ouvert du plan complexe, pouvant être non connexe. $R(\zeta)$ est holomorphe sur chaque composante connexe de $P(T)$, et ne peut pas être prolongé analytiquement à travers les frontières des composantes connexes.*

Le complémentaire de $P(T)$ dans \mathcal{C} sera appelé le spectre de T et noté $\Sigma(T)$. $\Sigma(T)$ peut être vide ou égal à \mathcal{C} . Le théorème suivant [34][th. 6.13] caractérise la résolvante en l'infini :

THÉORÈME 1.4. *Soit $T \in \mathcal{C}(X)$, tel que $P(T)$ contienne l'extérieur d'un cercle. Alors seules deux alternatives sont possibles :*

- $T \in \mathcal{B}(X)$; $R(\zeta)$ est holomorphe au point $\zeta = \infty$ et $R(\infty) = 0$.
- $R(\zeta)$ a une singularité essentielle au point $\zeta = \infty$.

Il sera donc naturel d'inclure le point $\zeta = \infty$ dans l'ensemble résolvant de T ou dans son spectre, selon que $T \in \mathcal{B}(X)$ ou non. Nous désignerons alors respectivement par $\tilde{P}(T)$ et $\tilde{\Sigma}(T)$ l'ensemble résolvant étendu et le spectre étendu de T . Comme un opérateur $T \in \mathcal{B}(X)$ a son spectre inclus dans le cercle centré en l'origine et de rayon $\|T\|$, le théorème 1.4 s'applique et par conséquent le point

$\zeta = \infty \in \tilde{P}(T)$ si et seulement si $T \in \mathcal{B}(X)$. Donc un opérateur non borné aura toujours le point à l'infini dans son spectre étendu ; si ce point est isolé dans le spectre étendu, cela sera alors une singularité essentielle de $R(\zeta)$. Si $\dim X < \infty$, les points du spectre sont des pôles de la résolvante, qui est donc une fonction méromorphe sur \mathcal{C} . Dans le cas de la dimension infinie, les points du spectre d'un opérateur, même borné, peuvent être essentiels ; par exemple l'opérateur de translation sur ℓ^p défini par $Tx_1 = 0, Tx_n = x_{n-1} (n \geq 2)$ est borné et de spectre le disque unité. Cependant dans le cas des opérateurs compacts, les points du spectre non nul sont des pôles de la résolvante ; il en est de même pour un opérateur T dont la résolvante est compacte pour au moins un ζ , sinon qu'alors 0 n'appartient pas au spectre de T .

REMARQUE 1.5. La notion de spectre d'un opérateur peut s'étendre à un opérateur linéaire quelconque. Si T admet une extension fermée \tilde{T} alors $P(T) = P(\tilde{T})$, et $\Sigma(T) = \Sigma(\tilde{T})$, si T n'admet pas d'extension fermée alors $P(T)$ est vide et $\Sigma(T) = \mathcal{C}$.

1.2. Séparation du spectre. Supposons que le spectre d'un opérateur fermé T contienne une partie bornée Σ' séparée du reste du spectre, noté Σ'' , de telle façon qu'il soit possible de tracer une courbe rectifiable Γ dans $P(T)$ (ou plus généralement un nombre fini de courbes si $P(T)$ n'est pas connexe) qui entoure un ouvert contenant Σ' tout en laissant Σ'' à l'extérieur.

THÉORÈME 1.6. *Soit $\Sigma(T)$ séparable en deux parties Σ' et Σ'' de la façon décrite ci-dessus. Alors T admet une décomposition selon $X = M' \oplus M''$ de telle façon que les spectres des restriction $T_{M'}$ et $T_{M''}$ coïncident respectivement avec Σ' et Σ'' et que $T_{M'} \in \mathcal{B}(M')$. Ainsi $\tilde{\Sigma}(T_{M'}) = \Sigma'$, tandis que $\tilde{\Sigma}(T_{M''})$ peut contenir $\zeta = \infty$.*

DÉMONSTRATION. [34][p. 178] L'opérateur borné

$$P = -\frac{1}{2i\pi} \int_{\Gamma} R(\zeta) d\zeta$$

est un projecteur. En effet, comme la valeur de cette intégrale ne dépend que de la classe d'homotopie de la courbe Γ , choisissons une autre courbe Γ' dans $P(T)$ englobant Γ . Nous obtenons ainsi

$$\begin{aligned} P^2 &= \left(\frac{1}{2i\pi} \right)^2 \int_{\Gamma'} \int_{\Gamma} R(\zeta) R(\zeta') d\zeta d\zeta' \\ &= \left(\frac{1}{2i\pi} \right)^2 \int_{\Gamma'} \int_{\Gamma} (\zeta' - \zeta)^{-1} (R(\zeta') - R(\zeta)) d\zeta d\zeta' \end{aligned}$$

d'après (1.1). D'autre part

$$\left(\frac{1}{2i\pi} \right)^2 \int_{\Gamma} (\zeta' - \zeta)^{-1} d\zeta = 0, \quad \left(\frac{1}{2i\pi} \right)^2 \int_{\Gamma'} (\zeta' - \zeta)^{-1} d\zeta' = 1,$$

d'où $P^2 = P$. Ainsi P est un projecteur sur $M' = PX$, parallèlement à $M'' = (I - P)X$.

Comme pour tout $\zeta \in P(T)$, $PR(\zeta) = R(\zeta)P$, le théorème 1.2 implique que T commute avec P ; par conséquent T se décompose selon $X = M' \oplus M''$. Notons par $R_{M'}(\zeta)$ et $R_{M''}(\zeta)$ les restrictions de la résolvante au sous espaces respectifs M' et M'' (i.e $R_{M'}(\zeta)$ est définie sur l'espace de Banach M' et non pas sur X). Ce sont les inverses respectifs de $T_{M'} - \zeta$ et $T_{M''} - \zeta$; ceci montre que $P(T)$ est inclus dans $P(T_{M'})$ et $P(T_{M''})$. Mais pour tout $\zeta \in P(T) \setminus \Gamma$, nous avons

$$(1.2) \quad R(\zeta)P = -\frac{1}{2i\pi} \int_{\Gamma} R(\zeta)R(\zeta')d\zeta' = -\frac{1}{2i\pi} \int_{\Gamma} (R(\zeta) - R(\zeta'))(\zeta - \zeta')^{-1}d\zeta'.$$

Si ζ est à l'extérieur de Γ , nous obtenons

$$R(\zeta)P = \frac{1}{2i\pi} \int_{\Gamma} R(\zeta')(\zeta - \zeta')^{-1} d\zeta'.$$

Comme le membre de droite de l'expression précédente est une fonction holomorphe à l'extérieur de Γ , il s'ensuit que $R(\zeta)P$ et donc $R_{M'}(\zeta)$ peuvent être prolongés analytiquement à l'extérieur de Γ . Le théorème 1.3 implique que ce prolongement analytique de $R_{M'}(\zeta)$ est la résolvante de $T_{M'}$; ainsi $P(T_{M'})$ contient l'extérieur de Γ et donc $\Sigma(T_{M'}) \subset \Sigma'$. De façon similaire, si ζ est à l'intérieur de Γ , l'équation (1.2) donne

$$R(\zeta)P = R(\zeta) + \frac{1}{2i\pi} \int_{\Gamma} R(\zeta')(\zeta - \zeta')^{-1} d\zeta';$$

ce qui montre que $R(\zeta)(I - P)$ peut être prolongé analytiquement dans l'intérieur de Γ , impliquant que $\Sigma(T_{M''}) \subset \Sigma''$. D'autre part, un point $\zeta \in \Sigma(T)$ ne peut appartenir à $P(T_{M'}) \cap P(T_{M''})$, car $R_{M'}(\zeta)P + R_{M''}(\zeta)(I - P)$ étant l'inverse de $T - \zeta$, cela impliquerait que $\zeta \in P(T)$. Nous en déduisons que $\Sigma(M') = \Sigma'$ et $\Sigma(M'') = \Sigma''$. Pour conclure, une approximation de l'intégrale par une somme finie et l'utilisation du fait que T est fermé conduisent à l'expression suivante :

$$TP = -\frac{1}{2i\pi} \int_{\Gamma} TR(\zeta)d\zeta = -\frac{1}{2i\pi} \int_{\Gamma} \zeta R(\zeta)d\zeta.$$

Cette expression implique que $T_{M'} \in \mathcal{B}(M')$, ce qui complète la preuve du théorème. Notons également que la résolvante $R(\zeta)$ peut s'écrire sous la forme

$$\begin{aligned} R(\zeta) &= R'(\zeta) + R''(\zeta) \\ R'(\zeta) &= R(\zeta)P, \quad R''(\zeta) = R(\zeta)(I - P) \end{aligned}$$

où $R'(\zeta)$, holomorphe dans $\mathcal{C} \setminus \Sigma'$, coïncide avec $R_{M'}(\zeta)$ sur M' et s'annule sur M'' (résultats similaires pour $R''(\zeta)$). \square

Le théorème 1.6 s'étend facilement au cas où $\Sigma(T)$ se décompose en plusieurs parties bornées $\Sigma_1, \dots, \Sigma_s$ séparées entre elles et du reste du spectre.

1.2.1. *Cas d'un point isolé.* Supposons maintenant que le spectre $\Sigma(T)$ de l'opérateur fermé T contienne un point isolé λ . Les résultats précédents s'appliquent avec $\Sigma'(T) = \{\lambda\}$. L'opérateur $T_{M'}$ a donc un spectre constitué d'un seul point; ainsi $T_{M'} - \lambda$ est quasi-nilpotent (i.e son rayon spectral est nul). Le développement en série de la résolvante de $T_{M'}$ convergente pour tout $\zeta \neq \lambda$,

$$R_{M'}(\zeta) = (\zeta - \lambda)^{-1} \left[I - \frac{T_{M'} - \lambda}{\zeta - \lambda} \right]^n = -\sum_{n=0}^{\infty} (\zeta - \lambda)^{-n-1} (T_{M'} - \lambda)^n,$$

est équivalent à l'expression suivante :

$$(1.3) \quad R'(\zeta) = R(\zeta)P = -(\zeta - \lambda)^{-1}P - \sum_{n=1}^{\infty} (\zeta - \lambda)^{-n-1} D^n$$

où

$$(1.4) \quad D = (T - \lambda)P = -\frac{1}{2i\pi} \int_{\Gamma} (\zeta - \lambda)R(\zeta)d\zeta \in \mathcal{B}(X),$$

est un opérateur quasi-nilpotent, vérifiant $D = DP = PD$. D'autre part, $R_{M''}(\zeta)$, étant holomorphe en $\zeta = \lambda$, admet le développement en série suivant :

$$R_{M''}(\zeta) = \sum_{n=0}^{\infty} (\zeta - \lambda)^n R_{M''}(\lambda)^{n+1},$$

qui est équivalent à

$$(1.5) \quad R''(\zeta) = R(\zeta)(I - P) = \sum_{n=0}^{\infty} (\zeta - \lambda)^n S^{n+1}$$

où

$$(1.6) \quad S = R_{M''}(\lambda)(I - P) = \lim_{\zeta \rightarrow \lambda} R(\zeta)(I - P).$$

S est appelée la résolvante réduite de T pour le point λ . L'opérateur S possède des propriétés analogues à l'opérateur du même type introduit dans le cas de la dimension finie; en particulier

$$S = \frac{1}{2i\pi} \int_{\Gamma} (\zeta - \lambda)^{-1} R(\zeta) d\zeta \in \mathcal{B}(X)$$

$$ST \subset TS \in \mathcal{B}(X), \quad (T - \lambda)S = I - P, \quad SP = PS = 0.$$

En combinant (1.3) et (1.5), nous obtenons

$$R(\zeta) = -(\zeta - \lambda)^{-1}P - \sum_{n=1}^{\infty} (\zeta - \lambda)^{-n-1}D^n + \sum_{n=0}^{\infty} (\zeta - \lambda)^n S^{n+1},$$

expression similaire au cas de la dimension finie, sinon que la partie principale peut contenir une infinité de termes. Néanmoins si M' est de dimension finie, alors $D_{M'} = T_{M'} - \lambda$ est nilpotent et par conséquent il en est de même de D ($D = D_{M'}$ sur M' et 0 sur M''). Dans ce cas λ est une valeur propre de $T_{M'}$ qui est un opérateur de rang fini, et donc λ est une valeur propre de T . La dimension de M' sera également appelée multiplicité algébrique de λ et P sera le projecteur propre associé à λ . Notons que si $\dim M' = \infty$, le point isolé λ ne sera pas nécessairement une valeur propre de T . Ces résultats peuvent être étendus au cas de plusieurs points isolés.

Les résultats énoncés ci-dessus suggèrent d'étendre la théorie des perturbations analytiques, obtenue dans le cas de la dimension finie, aux opérateurs fermés qui ont un point isolé dans le spectre et dont le projecteur propre associé est de rang fini. Ce sera en particulier le cas pour les opérateurs compacts (qui sont par définition dans $\mathcal{B}(X)$), les opérateurs quasi-compacts (cf définition 5-(2.1)) et les opérateurs fermés dont la résolvante est un opérateur compact pour au moins un ζ_0 ; en effet, le théorème de l'image spectrale implique que le spectre de T est l'image du spectre discret de sa résolvante par l'application $\zeta \rightarrow (\zeta - \zeta_0)^{-1}$, notons que le point zéro du spectre de la résolvante est envoyé en l'infini, donc le spectre de T est sans point d'accumulation.

2. Perturbation analytique

Considérons une famille d'opérateurs $T(\chi) \in \mathcal{B}(X, Y)$ dépendant de la variable complexe χ . $T(\chi)$ est holomorphe s'il est différentiable en norme pour tout χ dans un domaine du plan complexe. On a équivalence entre cette propriété d'holomorphicité en norme et l'holomorphicité faible définie par : toutes les fonctions $(T(\chi)u, g)$ sont holomorphes pour tout $u \in X$ et tout $g \in Y^*$, où Y^* désigne le dual topologique de Y . Il est possible d'étendre cette notion aux opérateurs fermés; pour faire la distinction nous dirons que $T(\chi)$ est holomorphiquement borné si $T(\chi) \in \mathcal{B}(X, Y)$ et $T(\chi)$ est holomorphe dans le sens précédent.

Une famille d'opérateurs $T(\chi) \in \mathcal{C}(X, Y)$, définis dans un voisinage de $\chi = 0$, est dite holomorphe en $\chi = 0$ s'il existe un troisième espace de Banach Z et

deux familles d'opérateurs, $U(\chi) \in \mathcal{B}(Z, X)$ et $V(\chi) \in \mathcal{B}(Z, Y)$ holomorphiquement bornés en $\chi = 0$, telles que $U(\chi)$ envoie injectivement Z sur $D(T(\chi))$ et telles que

$$(2.1) \quad T(\chi)U(\chi) = V(\chi).$$

$T(\chi)$ sera dit holomorphe dans un domaine $D \subset \mathcal{C}$ s'il est holomorphe en chaque $\chi \in D$. Cette définition n'est pas ambiguë, car nous déduisons des hypothèses que $T(\chi)U(\chi) \in \mathcal{C}(Z, Y)$ et $D(T(\chi)U(\chi)) = Z$, d'où d'après le théorème 1.1 $T(\chi)U(\chi) \in \mathcal{B}(Z, Y)$. D'autre part, $T(\chi) \in \mathcal{B}(X, Y)$ est holomorphiquement borné si et seulement s'il est holomorphe dans le sens précédent ; en effet si $T(\chi)$ est holomorphiquement borné, il suffit de prendre $Z = G(T)$ muni de la norme $\|(u, Tu)\| = (\|u\|^2 + \|Tu\|^2)^{1/2}$, $U(\chi)(u, Tu) = u$ et $V(\chi)(u, Tu) = Tu$. Réciproquement, comme $U(\chi)$ envoie injectivement Z sur $D(T(\chi))$, alors $U(\chi)^{-1}$ existe et a pour domaine $D(T(\chi))$ et pour image Z . D'après le théorème 1.1, $U(\chi)^{-1} \in \mathcal{B}(X, Z)$ et $U(\chi)^{-1}$ est holomorphiquement borné [34][VII-1.1 p. 365] ; il en est alors de même de $T(\chi) = V(\chi)U(\chi)^{-1}$. Nous avons le théorème suivant [34][th. VII-1.3] :

THÉORÈME 2.1. *Soit $T(\chi) \in \mathcal{C}(X)$ défini dans un voisinage de $\chi = 0$ et soit $\zeta \in P(T(0))$. Alors $T(\chi)$ est holomorphe en $\chi = 0$ si et seulement si pour $|\chi|$ suffisamment petit, $\zeta \in P(T(\chi))$ et la résolvante $R(\zeta, \chi) = (T(\chi) - \zeta)^{-1}$ est holomorphiquement bornée. Dans ce cas $R(\zeta, \chi)$ est holomorphiquement bornée pour les deux variables sur l'ensemble des ζ, χ tels que $\zeta \in P(T(0))$ et $|\chi|$ suffisamment petit (dépendant de ζ).*

Sans faire la preuve, signalons que si $R(\zeta, \chi)$ est holomorphiquement borné en χ , alors pour montrer que $T(\chi)$ est holomorphe en χ , il suffit de choisir $Z = X$, $R(\zeta, \chi) = U(\chi)$ et $V(\chi) = 1 + \zeta U(\chi)$ pour vérifier (2.1).

2.1. Séparation du spectre pour une famille holomorphe.

THÉORÈME 2.2. *Soit $T(\chi) \in \mathcal{C}(X)$ une famille d'opérateurs fermés holomorphes dans un voisinage de $\chi = 0$. Supposons que le spectre $\Sigma(T(0))$ se décompose en deux parties $\Sigma'(0)$ et $\Sigma''(0)$ séparées par une courbe fermée Γ , alors $\Gamma \subset P(T(\chi))$ pour tout $|\chi|$ suffisamment petit et $\Sigma(T(\chi))$ se décompose également en deux parties $\Sigma'(\chi)$ et $\Sigma''(\chi)$ séparées par Γ .*

Nous obtenons alors une décomposition $X = M'(\chi) \oplus M''(\chi)$, où le projecteur sur $M'(\chi)$ parallèlement à $M''(\chi)$ est donné par

$$(2.2) \quad P(\chi) = -\frac{1}{2\pi i} \int_{\Gamma} R(\zeta, \chi) d\zeta.$$

Comme $R(\zeta, \chi)$ est holomorphiquement borné en χ et ζ , il en découle que $P(\chi)$ est aussi holomorphiquement borné près de $\chi = 0$.

De cette façon l'étude du spectre de $T(\chi)$ est réduite à l'étude du spectre des restrictions de $T(\chi)$ sur $M'(\chi)$ et $M''(\chi)$. Mais ces espaces dépendent de χ ; cependant cet inconvénient peut être évité de la façon suivante. Puisque le projecteur $P(\chi)$ sur $M'(\chi)$ est holomorphe en χ , il existe une transformation $U(\chi) \in \mathcal{B}(X)$ holomorphiquement bornée, inversible et d'inverse $U(\chi)^{-1} \in \mathcal{B}(X)$ également holomorphiquement borné, qui transforme $P(0)$ en $P(\chi)$ [34][II-§4.2]

$$P(\chi) = U(\chi)P(0)U(\chi)^{-1}.$$

Puisque $T(\chi)$ commute avec $P(\chi)$, alors l'opérateur

$$\tilde{T}(\chi) = U(\chi)^{-1}T(\chi)U(\chi)$$

commute avec $P(0)$. Ainsi $M'(0) = P(0)X$ et $M''(0) = (I - P(0))X$ décompose $\tilde{T}(\chi)$ pour tout χ . D'autre part, chercher le spectre de la restriction de $T(\chi)$ à $M'(\chi)$ est équivalent à rechercher le spectre de la restriction $\tilde{T}(\chi)$ à $M'(0)$ (car

$\check{T}(\chi)P(0) = U(\chi)^{-1}T(\chi)P(\chi)U(\chi)$). En particulier si $\dim M'(0) < \infty$, alors il en sera de même de $M'(\chi)$ car le rang de $P(\chi)$ sera fini ; $\Sigma'(0)$ se compose alors d'un nombre fini de valeurs propres. Supposons par exemple que le spectre de la restriction de $\check{T}(\chi)$ à $M'(0)$ soit constitué des valeurs propres $\lambda_h(\chi)$, $h = 1, \dots, s$; si $\check{P}_h(\chi)$ et $\check{D}_h(\chi)$ désignent respectivement le projecteur propre et l'opérateur nilpotent associés à la valeur propre $\lambda_h(\chi)$, nous obtenons les relations suivantes :

$$\check{T}(\chi)\check{P}_h(\chi) = \lambda_h(\chi)\check{P}_h(\chi) + \check{D}_h(\chi), \quad h = 1, \dots, s,$$

où

$$\begin{aligned} \check{P}_h(\chi)P(0) &= P(0)\check{P}_h(\chi) = \check{P}_h(\chi) \\ \check{D}_h(\chi)D(0) &= D(0)\check{D}_h(\chi) = \check{D}_h(\chi). \end{aligned}$$

D'autre part, la restriction de $T(\chi)$ au sous-espace $M'(\chi)$ vérifie les relations suivantes :

$$T(\chi)P_h(\chi) = \lambda_h(\chi)P_h(\chi) + D_h(\chi), \quad h = 1, \dots, s,$$

avec

$$P_h(\chi) = U(\chi)\check{P}_h(\chi)U(\chi)^{-1}, \quad D_h(\chi) = U(\chi)\check{D}_h(\chi)U(\chi)^{-1}.$$

Comme les résultats établis dans le cas d'un espace X de dimension finie s'appliquent à la restriction de $\check{T}(\chi)$ au sous-espace $M'(0)$, la recherche des valeurs propres (en nombre fini) de la restriction de $T(\chi)$ au sous-espace $M'(\chi)$ se réduit à un problème en dimension finie.

2.2. Développement en série des valeurs propres perturbées. Considérons une famille $T(\chi)$ d'opérateurs fermés holomorphes dans un voisinage de $\chi = 0$, telle que $T(0)$ admette une valeur propre λ isolée dont le projecteur propre est de rang fini. Pour obtenir des formules explicites permettant d'estimer les coefficients du développement en série, selon les puissances de χ , de la valeur propre perturbée $\lambda(\chi)$ il convient de considérer le développement en série

$$\check{T}(\chi)P(0) = \check{T}_{M'(0)}(0) + \sum_{n \geq 1} \chi^n \check{T}^{(n)},$$

où $\check{T}^{(n)} \in \mathcal{B}(X)$. La relation $\check{T}(\chi)P(0) = U(\chi)^{-1}T(\chi)P(\chi)U(\chi)$ permet d'obtenir le développement en série suivant

$$(2.3) \quad T_r(\chi) := T(\chi)P(\chi) = T_r + \chi T_r^{(1)} + \chi^2 T_r^{(2)} + \dots,$$

avec $T_r = U(\chi)\check{T}(0)P(0)U(\chi)^{-1}$, et $T_r^{(n)} = U(\chi)\check{T}^{(n)}U(\chi)^{-1}$.

Il n'est pas possible de considérer une telle série directement pour l'opérateur $T(\chi)$, celui-ci n'étant pas nécessairement holomorphiquement borné. Pour estimer $\lambda(\chi)$ l'opérateur $U(\chi)$ n'a pas besoin d'être introduit explicitement, il suffit d'utiliser les estimations obtenues dans le cas de la dimension finie directement pour l'opérateur $T_r(\chi)$. Dans ce cas il convient de remplacer dans les diverses expressions, les opérateurs $T^{(n)}$ par les opérateurs $T_r^{(n)}$, ceux-ci étant également holomorphiquement bornés. D'autre part l'opérateur trace est bien défini (par ex. $\text{tr}T(\chi)P(\chi)$) car les opérateurs considérés sont de rang fini [34][§III.4.3].

Par exemple si λ est de multiplicité algébrique m , la formule II-(2.5) dans [34] devient pour $|\chi|$ suffisamment petit

$$\hat{\lambda}(\chi) := \frac{1}{m} \text{tr}(T_r(\chi)) = \lambda + \sum_{n=1}^{\infty} \chi^n \hat{\lambda}^{(n)},$$

où

$$\hat{\lambda}^{(n)} = \frac{1}{m} \sum_{p=1}^n \frac{(-1)^p}{p} \sum_{\substack{\nu_1 + \dots + \nu_p = n \\ k_1 + \dots + k_p = p-1 \\ \nu_i \geq 1, k_j \geq 0}} \text{tr} \left[T_r^{(\nu_1)} S^{(k_1)} \dots T_r^{(\nu_p)} S^{(k_p)} \right]$$

avec $S^{(0)} = -P(0)$, et $S^{(k)} = S^k$ pour $k > 0$, S étant donné par la formule (1.6).

Les résultats obtenus, dans le cas de la dimension finie, peuvent être appliqués également pour une autre catégorie de famille d'opérateurs $T(\chi)$; les opérateurs fermés holomorphes de type (A) [34][VII §2.1]. Une famille $T(\chi) \in \mathcal{C}(X, Y)$ définie pour χ dans un domaine complexe D_0 sera dite holomorphe de type (A) si

- (1) chaque domaine $D(T(\chi)) = D$ est indépendant de χ ,
- (2) $T(\chi)u$ est holomorphe pour $\chi \in D_0$ et pour tout $u \in D$.

Dans ce cas $T(\chi)u$ admet un développement en série de Taylor pour chaque $\chi \in D_0$, en particulier si $\chi = 0 \in D_0$ celui-ci s'écrira

$$(2.4) \quad T(\chi)u = Tu + \chi T^{(1)}u + \chi^2 T^{(2)}u + \dots, \quad u \in D.$$

Cette série converge pour $|\chi| < r$, r étant indépendant de u , d'autre part $T = T(0)$ et les $T^{(n)}$ sont des opérateurs linéaires de X dans Y de domaine D .

Une famille $T(\chi)$ holomorphe de type (A) est holomorphe dans le sens général introduit auparavant [34][p. 375]. Plusieurs critères, permettant d'affirmer qu'une famille d'opérateurs $T(\chi)$ est holomorphe de type (A), figurent dans [34][VII-§2]. Un de ces critères est le suivant :

THÉORÈME 2.3. *Soit T un opérateur de X dans Y admettant une extension fermée. Soient $T^{(n)}$, $n = 1, 2, \dots$ des opérateurs de X dans Y dont les domaines contiennent $D := D(T)$, tels que*

$$\|T^{(n)}u\| \leq c^{n-1} (a\|u\| + b\|Tu\|), \quad u \in D, n = 1, 2, \dots,$$

pour des constantes positives a, b, c . Alors la série (2.4) définit un opérateur $T(\chi)$ de domaine D pour $|\chi| < 1/c$. Si $|\chi| < (b+c)^{-1}$, $T(\chi)$ admet une extension fermée $\tilde{T}(\chi)$ qui est holomorphe de type (A).

Notons que si $T^{(n)} = 0$, pour tout $n \geq 2$, c pourra être pris égal 0 si

$$\|T^{(1)}u\| \leq a\|u\| + b\|Tu\|.$$

Supposons maintenant que le spectre $\Sigma(T)$ puisse être scindé en deux parties par une courbe fermée Γ . Alors d'après le théorème 2.2, le spectre $\Sigma(T(\chi))$ est également constitué de deux parties séparables par Γ pour $|\chi|$ suffisamment petit. Dans le cas d'une famille d'opérateurs holomorphes de type (A), il est possible d'estimer l'ordre de grandeur de $|\chi|$. En effet posons $A(\chi) = T(\chi) - T$, nous obtenons alors

$$\|A(\chi)u\| \leq |\chi|(1 - c|\chi|)^{-1} (a\|u\| + b\|Tu\|), \quad u \in D$$

$$\|R(\zeta, T(\chi))\| \leq \|R(\zeta, T)\| (1 - a(\chi)\|R(\zeta, T)\| - b(\chi)\|TR(\zeta, T)\|)^{-1}, \quad \zeta \in P(T)$$

avec $a(\chi) = a|\chi|(1 - c|\chi|)^{-1}$ et $b(\chi) = b|\chi|(1 - c|\chi|)^{-1}$. Nous avons utilisé le fait que $R(\zeta, T)$ est un opérateur borné, ce qui permet dans la deuxième inégalité de sortir le vecteur u des diverses normes.

Par conséquent, la séparation du spectre est conservée dès que

$$(2.5) \quad |\chi| < r_0 := \min_{\zeta \in \Gamma} (a\|R(\zeta, T)\| + b\|TR(\zeta, T)\| + c)^{-1},$$

et le projecteur $P(\chi)$ donné par la formule (2.2) est alors holomorphe en χ . En particulier, si la partie du spectre de T , à l'intérieur de Γ , est constituée exclusivement d'un nombre fini de valeurs propres, l'existence d'un développement en série

pour $T(\chi)$ rend possible l'application directe des résultats obtenus dans le cadre de la dimension finie. Il n'est donc pas nécessaire de remplacer les $T^{(n)}$ par $T_r^{(n)}$. Ceci vient du fait que malgré le caractère non borné des $T^{(n)}$, la résolvante de l'opérateur perturbé $T(\chi)$ admet un développement en série, comme dans le cas de la dimension finie (seconde série de Neumann). En outre, les formules d'estimations des coefficients du développement en série, de la valeur propre perturbée, ne font apparaître que des termes de la forme $T^{(n)}P, T^{(n)}D = T^{(n)}PD$, qui sont tous des opérateurs bornés. Notons néanmoins, que l'utilisation de la trace ne peut se faire que sur des opérateurs de rang finis, en particulier celle-ci ne peut porter que sur des termes dans lesquels figure le projecteur propre P .

2.3. Opérateurs non bornés dans un espace de Hilbert. Les résultats précédents se simplifient dans le cas $X = H$ où H est un espace de Hilbert et $T = T(0)$ est un opérateur normal. Rappelons d'abord, quelques définitions et propriétés des opérateurs linéaires définis sur un espace de Hilbert. Pour tout opérateur linéaire T défini sur un espace de Hilbert H , le domaine numérique (numerical range) $\Theta(T)$ est défini comme l'ensemble de tous les nombres complexes (Tu, u) , où u varie dans $D(T)$ avec $\|u\| = 1$. En général, $\Theta(T)$ n'est ni ouvert ni fermé, même lorsque T est fermé. Cependant Hausdorff a prouvé que $\Theta(T)$ est un ensemble convexe. Soit Γ la fermeture de $\Theta(T)$ et Δ le complémentaire de Γ dans \mathcal{C} . Δ est alors un ensemble ouvert connexe, sauf lorsque Γ est une bande bornée par deux lignes parallèles. Dans ce cas, Δ est formé de deux composantes connexes Δ_1, Δ_2 qui sont des demi-plans. Le théorème suivant, prouvé dans [34][V-§3.2 Th 3.2], caractérise Δ pour un opérateur fermé :

THÉORÈME 2.4. *Soient $T \in \mathcal{C}(H)$ et $\Gamma, \Delta, \Delta_1, \Delta_2$ définis comme précédemment. Pour tout $\zeta \in \Delta$, $R(T - \zeta)$ est fermé, $\dim N(T - \zeta) = 0$ (où $N(T)$ désigne le noyau de T) et $\text{codim} R(T - \zeta)$ est constante, sauf pour le cas spécial mentionné ci-dessus, où $\text{codim} R(T - \zeta)$ est constante dans chaque composante connexe Δ_1, Δ_2 . Si $\text{codim} R(T - \zeta) = 0$ pour $\zeta \in \Delta$ ($\zeta \in \Delta_1$, ou $\zeta \in \Delta_2$) alors Δ (Δ_1 , ou Δ_2) est un sous-ensemble de $P(T)$ et $\|R(\zeta, T)\| \leq 1/\text{dist}(\zeta, \Gamma)$.*

COROLLAIRE 2.5. *Si $T \in \mathcal{B}(H)$, alors $\Sigma(T)$ est un sous-ensemble de la fermeture de $\Theta(T)$.*

L'opérateur adjoint d'un opérateur linéaire T est un opérateur S tel que

$$(g, Tu) = (Sg, u), \quad u \in D(T), g \in D(S).$$

Il peut cependant y avoir plusieurs opérateurs adjoints pour un même opérateur T . Néanmoins si T a un domaine dense dans H il existe un unique opérateur maximal T^* adjoint de T . Ceci signifie que tout opérateur S adjoint de T est une restriction de T^* . Dans ce cas T^* sera appelé adjoint de T . On démontre que T^* est toujours un opérateur fermé, même si T n'est ni fermé ni à extension fermée, mais il se peut que $D(T^*) = \{0\}$ [34][III-§5.5]. Soit T un opérateur fermé dans H et T^* son adjoint, alors $P(T^*)$ et $\Sigma(T^*)$ sont les images respectives de l'ensemble résolvant et du spectre de T par l'application $z \rightarrow \bar{z}$ et

$$R(\zeta, T^*) = R(\bar{\zeta}, T)^*, \quad \bar{\zeta} \in P(T).$$

Un opérateur T sera appelé autoadjoint si $T^* = T$. Dans ce cas son spectre $\Sigma(T)$ est inclus dans \mathbb{R} ainsi que $\Theta(T)$. Sa résolvante est un opérateur normal (borné), car $R(\zeta)^* = R(\bar{\zeta})$ et $R(\zeta)R(\zeta') = R(\zeta')R(\zeta)$ pour $\zeta \neq \zeta'$. Un opérateur

autoadjoint et sa résolvante vérifient les relations suivantes :

$$(2.6) \quad \begin{aligned} \|(T - \zeta)u\|^2 &= \|(T - \operatorname{Re}\zeta)u\|^2 + (\operatorname{Im}\zeta)^2\|u\|^2 \\ \|R(\zeta)\| &= 1/\operatorname{dist}(\zeta, \Sigma(T)) \leq |\operatorname{Im}\zeta|^{-1} \\ \|TR(\zeta)\| &= \sup_{\lambda \in \Sigma(T)} |\lambda| |\lambda - \zeta|^{-1} \end{aligned}$$

Supposons maintenant que $\Sigma(T)$ contienne un point isolé λ , alors le projecteur propre P , la résolvante réduite S et l'opérateur quasi-nilpotent D sont également autoadjoints. En particulier ceci implique que $D = 0$ et par conséquent que λ est une valeur propre de T . De plus, nous avons l'identité suivante :

$$\|S\| = 1/d, \quad \text{où } d = \operatorname{dist}(\lambda, \Sigma(T) \setminus \{\lambda\}).$$

Comme en dimension finie, il est possible de construire un opérateur autoadjoint à partir d'un opérateur T fermé de domaine dense dans H . En effet selon le théorème de Von Neumann [34][V-§3.7], l'opérateur T^*T est autoadjoint dans H et $D(T^*T)$ est un noyau (traduction du terme anglais core) de T . D est appelé noyau de T si l'ensemble des éléments $\{u, Tu\}$ pour $u \in D$ est dense dans le graphe $G(T)$. Si D est un noyau de T alors D est dense dans $D(T)$ et la fermeture de la restriction de T à D est égale à T .

Un opérateur T , non nécessairement borné, est dit normal dans H si T est fermé, de domaine dense dans H et $T^*T = TT^*$. Si T est normal, alors pour tout $u \in D := D(T^*T) = D(TT^*)$ nous avons $\|Tu\| = \|T^*u\|$. D'autre part, D étant un noyau de T et T^* donc dense dans $D(T)$ et $D(T^*)$, nous en déduisons que $D(T) = D(T^*) := D_1(\supset D)$, d'où $\|Tu\| = \|T^*u\|$ pour tout $u \in D_1$. Comme dans le cas d'un opérateur autoadjoint (qui n'est qu'un cas particulier d'un opérateur normal), la résolvante est un opérateur normal et vérifie les mêmes relations. Si le spectre de T contient un point isolé λ , alors le projecteur propre associé, la résolvante réduite S et l'opérateur nilpotent D sont normaux. Ceci implique que $D = 0$ donc que λ est une valeur propre de T .

Si dans les hypothèses du théorème 2.3, l'opérateur T est supposé normal, alors la condition (2.5) devient

$$r_0 := \min_{\zeta \in \Gamma} \left(a \sup_{\lambda' \in \Sigma(T)} |\lambda' - \zeta|^{-1} + b \sup_{\lambda' \in \Sigma(T)} |\lambda'| |\lambda' - \zeta|^{-1} + c \right)^{-1}.$$

Considérons le cas d'une valeur propre isolée λ de T , il est approprié de prendre pour courbe Γ un cercle centré en λ et de rayon $d/2$ où d désigne la distance de λ au reste du spectre de T . Alors $|\lambda' - \zeta|^{-1} \leq 2/d$ et $|\lambda'| |\lambda' - \zeta|^{-1} \leq 1 + |\zeta| |\lambda' - \zeta|^{-1} \leq 1 + (|\lambda| + 2^{-1}d)2d^{-1} \leq 2 + 2|\lambda|d^{-1}$, pour $\lambda' \in \Sigma(T)$ et $\zeta \in \Gamma$. Il est alors possible de minorer r_0 par

$$r_0 \geq [2(a + b|\lambda|)d^{-1} + 2b + c]^{-1}.$$

3. Le cas des semi-groupes

Les noyaux de transition d'un processus de Markov en temps continu, défini sur un espace mesuré $(\mathcal{X}, \mathcal{F}, \pi)$ est un semi-groupe de contraction agissant sur l'espace de Banach X des fonctions mesurables bornées. Le générateur infinitésimal du semi-groupe sera noté $-\Lambda$ donc $P_t = \exp(-t\Lambda)$. Ceci nous amène à l'étude des semi-groupes bornés et à des perturbations des générateurs de ces semi-groupes. Un semi-groupe $U(t)$ est défini par la limite suivante :

$$U(t) := e^{-t\Lambda} := \lim_{n \rightarrow \infty} \left(I + \frac{t}{n} \Lambda \right)^{-n}.$$

Cette définition est valable même lorsque l'opérateur Λ est non borné puisque $(I + \frac{t}{n}\Lambda)^{-1}$ est la résolvante de $-\Lambda$ à un facteur près, qui est donc un opérateur borné

pouvant être réitéré. Dans [34], il est prouvé que la limite précédente existe dès que les conditions suivantes sont vérifiées :

- (1) $\Lambda \in \mathcal{C}(X)$ et $D(\Lambda)$ est dense dans X .
- (2) Il existe $\beta \in \mathbb{R}$ tel que l'intervalle (β, ∞) appartienne à l'ensemble résolvant de $-\Lambda$ et la résolvante vérifie l'inégalité

$$\|(\Lambda + \zeta)^{-k}\| \leq M(\zeta - \beta)^{-k}, \quad \zeta > \beta, \quad k = 1, 2, 3, \dots,$$

où M est une constante indépendante de ζ et de k .

L'ensemble des opérateurs Λ satisfaisant aux conditions (i) et (ii) ci-dessus sera noté $\mathcal{G}(M, \beta)$. Nous obtenons alors la majoration suivante :

$$\|U(t)\| := \|e^{-t\Lambda}\| \leq Me^{\beta t},$$

et le nombre $\beta = \lim_{t \rightarrow \infty} t^{-1} \log \|\exp -t\Lambda\|$ est appelé le type du semi-groupe $\{e^{-t\Lambda}\}$. Lorsque $\|U(t)\| \leq 1$, le semi-groupe $\{U(t)\}$ est dit de contraction. Par exemple le générateur du semi-groupe associé au noyau de transition d'un processus markovien appartient à $\mathcal{G}(1, 0)$.

Si $\beta > 0$, la norme du semi-groupe n'est pas uniformément bornée ; un tel semi-groupe est dit quasi-borné. Ils sont également fortement continus en 0, i.e pour tout $u \in X, U(t)u \rightarrow 0$ quand $t \searrow 0$, et

$$\frac{d}{dt}U(t)u = -U(t)\Lambda u = -\Lambda U(t)u, \quad t \geq 0, \quad u \in D(T).$$

D'autre part, le demi-plan complexe $\operatorname{Re}\zeta > \beta$ appartient à l'ensemble résolvant de $-\Lambda$ et la résolvante correspondante s'exprime par la formule

$$(\Lambda + \zeta)^{-1} = \int_0^\infty e^{-\zeta t} e^{-t\Lambda} dt, \quad \operatorname{Re}\zeta > \beta.$$

3.1. Perturbation analytique d'un semi-groupe quasi-borné. Le théorème suivant affirme qu'un semi-groupe quasi-borné, perturbé par un opérateur borné, reste un semi-groupe quasi-borné ; sa démonstration figure dans [34][Th 2.1 IX-§2].

THÉORÈME 3.1. *Soient $\Lambda \in \mathcal{G}(M, \beta)$ et $A \in \mathcal{B}(X)$. Alors $\Lambda + A \in \mathcal{G}(M, \beta + M\|A\|)$ et $e^{-t(\Lambda + \chi A)}$ est une fonction entière de la variable complexe χ .*

Il est possible, dans le cas d'un semi-groupe de contraction, d'obtenir un résultat plus général. Nous dirons qu'un opérateur A est relativement borné par rapport à l'opérateur T (ou plus simplement T -borné) si $D(T) \subset D(A)$ et

$$\|Au\| \leq a\|u\| + b\|Tu\|, \quad u \in D(T),$$

où a et b sont des constantes positives. La plus grande borne inférieure b_0 de l'ensemble des constantes b admissibles dans l'inégalité précédente est appelé la T -borne de A . Plus b sera choisie proche de b_0 plus a sera grande ; il est donc en général impossible de prendre $b = b_0$. Nous avons [34][Th 2.7 IX-§2]

THÉORÈME 3.2. *Soient $\Lambda \in \mathcal{G}(1, \beta)$ et $A \in \mathcal{G}(1, \beta')$. Supposons que A soit relativement borné par rapport à Λ avec une Λ -borne $< 1/2$. Alors $\Lambda + A \in \mathcal{G}(1, \beta + \beta')$.*

REMARQUE 3.3. Si X est un espace de Hilbert H , la borne $1/2$ pour la condition du théorème 3.2 peut être remplacée par 1.

Soit un semi-groupe markovien engendré par $-\Lambda$ qui agit sur l'espace de Hilbert $H := L^2(\pi)$. Ce semi-groupe est de contraction et 0 est une valeur propre de Λ . Supposons que 0 soit un point isolé du spectre de Λ et désignons par E_π le projecteur propre associé. L'espace H admet alors la décomposition $H = M' \oplus M''$, où $M' =$

$E_\pi H$ et $M'' = (I - E_\pi)H$. Soit Λ'' la restriction de l'opérateur Λ sur le sous-espace M'' ; c'est un opérateur linéaire fermé sur l'espace de Hilbert M'' et son domaine est dense dans M'' . Plus précisément Λ'' génère un semi-groupe de contraction. Soit ω le type du semi-groupe engendré par Λ'' , nous avons $\omega \leq 0$. Lorsque $\omega < 0$ nous obtenons pour tout $f \in D(\Lambda'') \cap M''$ l'inégalité $\|P_t f\| \leq \exp(-|\omega|t)$. Dans ce cas le spectre de $P_t - E_\pi$ est, pour tout $t > 0$, strictement inclus dans le cercle unité.

Si Λ est autoadjoint, alors son spectre est inclus dans la droite réelle positive, il existe donc un $\beta > 0$ tel que $\Sigma(\Lambda'') \subset [\beta, \infty)$, par conséquent $\Lambda'' \in \mathcal{G}(1, -\beta)$. Dans ce cas pour tout $f \in D(\Lambda'') \cap M''$ nous avons

$$\|P_t f\| \leq e^{-\beta t}.$$

Le nombre positif β s'exprime par la relation

$$\beta = \inf\{(\Lambda'' f, f)/\|f\|^2; f \in D(\Lambda'') \cap M'', f \neq 0\},$$

i.e que le domaine numérique $\Theta(\Lambda'')$ (qui est inclus dans l'axe réel positif) est borné inférieurement par β . Ceci implique bien que $\Sigma(\Lambda'')$ est borné inférieurement par β . En effet si un opérateur T autoadjoint vérifie

$$(Tu, u) \geq \beta(u, u), \quad u \in D(T),$$

alors l'ensemble ouvert Δ complémentaire de la fermeture de $\Theta(T)$ est connexe et contient tous les nombres réels $\zeta < \beta$, ceux-ci sont donc inclus dans $P(T)$ (cf théorème 2.4). Ainsi $\Sigma(T)$ est borné inférieurement par $\beta_\Sigma \geq \beta$. D'autre part si $\Sigma(T)$ est borné inférieurement par β_Σ , alors l'opérateur $T' = T - \beta_\Sigma$ a un spectre inclus dans l'axe réel positif, donc pour tout $\alpha > 0$ nous avons $\|(T' + \alpha)^{-1}\| = \alpha^{-1}$ d'après la relation (2.6). Pour tout $u \in D(T)$, nous obtenons $\|u\| \leq \alpha^{-1}\|(T' + \alpha)u\|$ d'où

$$\begin{aligned} \|u\|^2 &\leq \alpha^{-2}\|T'u\|^2 + 2\alpha^{-1}(T'u, u) + \|u\|^2 \\ 0 &\leq \alpha^{-1}\|T'u\|^2 + 2(T'u, u). \end{aligned}$$

Pour conclure, il suffit de faire tendre α vers l'infini; ce qui conduit à $(Tu, u) \leq \beta_\Sigma(u, u)$. Ainsi T est borné inférieurement et $\beta \geq \beta_\Sigma$. Nous en déduisons que $\beta_\Sigma = \beta$. Notons également que Λ'' est inversible, $\Lambda''^{-1} \in \mathcal{B}(H)$ et $\|\Lambda''^{-1}\| \leq \beta$.

Si Λ n'est pas autoadjoint, un raisonnement identique à celui employé dans la section 1 montre que pour tout $f \in D(\Lambda'') \cap M''$, nous avons également

$$\|P_t f\| \leq e^{-\beta t},$$

où

$$\beta = \inf\{((\Lambda'' + \Lambda''^*)/2f, f)/\|f\|^2; f \in D(\Lambda'') \cap M'', f \neq 0\}.$$

L'opérateur fermé Λ ayant un domaine dense dans X , il admet un unique opérateur adjoint. Un espace de Hilbert étant réflexif, Λ^* possède également un domaine dense [34][th. III-5.29]. D'après les propriétés de la résolvante de l'adjoint d'un opérateur, énoncées dans la section 2.3, nous en déduisons que si $\Lambda \in \mathcal{G}(1, \beta)$ alors Λ^* appartient également à $\mathcal{G}(1, \beta)$. D'autre part pour un espace de Hilbert il est possible d'identifier H et H^* donc $\Lambda + \Lambda^*$ est autoadjoint de domaine $D(\Lambda) \cap D(\Lambda^*)$ dense dans H .

Supposons maintenant que $\dim M' < \infty$ et considérons la perturbation $\Lambda(\chi) := \Lambda - \chi A$ où $A \in \mathcal{G}(1, \beta)$ est relativement borné par rapport à Λ ,

$$\|Au\| \leq a\|u\| + b\|Tu\|, \quad u \in D(A).$$

Selon le critère (2.3), la famille d'opérateurs $\Lambda(\chi)$ est holomorphe de type (A) pour tout $|\chi| < b^{-1}$, et d'après (2.5) la séparation du spectre de Λ est conservée si,

$$|\chi| < r_0 = \min_{\zeta \in \Gamma} (a\|R(\zeta, \Lambda)\| + b\|TR(\zeta, \Lambda)\|)^{-1}.$$

En général le théorème 3.2 ne s'applique pas pour $\chi < 0$. En effet, pour avoir $\chi A \in \mathcal{G}(1, \beta)$, il n'est pas en général possible de choisir $\chi < 0$. L'hypothèse supplémentaire $\chi A \in \mathcal{G}(1, \beta)$ pour tout $\chi \in D_0$, où D_0 est un domaine de \mathcal{C} inclus dans le disque $|\chi| < b^{-1}$, permet de conclure que pour tout $\chi \in D_0$, $-\Lambda(\chi)$ engendre un semi-groupe. Lorsque A est borné, il suffit d'appliquer directement le théorème 3.1.

Il est alors possible d'utiliser les estimations de la moyenne pondérée des valeurs propres perturbées issues de 0, telles qu'obtenues en dimension finie.

Rappels sur les chaînes de Markov

Soit $(\mathcal{X}, \mathcal{F})$ un espace mesurable quelconque, la fonction $P(., .)$ définie sur $\mathcal{X} \times \mathcal{F}$ sera appelée fonction de transition si

- (1) Pour chaque x fixé, $P(x, .)$ détermine une probabilité.
- (2) Pour chaque $A \in \mathcal{F}$, $P(., A)$ détermine une fonction mesurable.

Cette fonction de transition définit une chaîne de Markov homogène. Il est alors possible de construire pour chaque instant n et pour tout état x_0 , la loi de probabilité de l'évolution de la chaîne à partir de l'instant n (cf [44], théorème de Ionescu Tulcea); cette probabilité que nous désignerons par P_x est définie sur $(\Omega, \mathcal{A}) = (\mathcal{X}, \mathcal{F})^{\mathbb{N}}$, où Ω est l'ensemble des suites (x_0, x_1, \dots) , $x_i \in \mathcal{X}$, et \mathcal{A} la tribu produit. La valeur de P_x sur tout pavé $\prod_m F_m$ est égale à

$$P_x \left[\prod_m F_m \right] = 1_{F_0}(x) \int_{F_1} P(x_1, dx_2) \cdots \int_{F_T} P(x_{T-1}, x_T),$$

pourvu que T soit tel que $F_m = \mathcal{X}$, si $m > T$. Si $p(.)$ est une probabilité sur $(\mathcal{X}, \mathcal{F})$, alors la formule

$$P(A) = \int_X P_x(A) p(dx)$$

définit une probabilité sur (Ω, \mathcal{A}) .

Une fonction de transition $P(., .)$ sur $(\mathcal{X}, \mathcal{F})$, définit un endomorphisme positif

$$Pf(x) = \int P(x, dy) f(y),$$

sur l'espace de Banach $B(X, \mathcal{F})$ des fonctions mesurables bornées muni de la norme $\|f\| = \sup_x |f(x)|$. Comme $P1 = 1$, on a $\|P\| = 1$. D'autre part la formule :

$$\mu P(F) = \int \mu(dx) P(x, F),$$

où F varie dans \mathcal{F} , définit un endomorphisme positif (opérant à gauche) sur l'espace de Banach $M(\mathcal{X}, \mathcal{F})$ des mesures bornées sur $(\mathcal{X}, \mathcal{F})$. De plus dans la dualité des espaces $B(\mathcal{X}, \mathcal{F})$ et $M(\mathcal{X}, \mathcal{F})$ établie par la forme bilinéaire $\mu(f) = \int \mu(dx) f(x)$, les opérateurs P sont transposés l'un de l'autre : $\mu(Pf) = \mu P(f)$, quels que soient $f \in B(\mathcal{X}, \mathcal{F})$ et $\mu \in M(\mathcal{X}, \mathcal{F})$. Notons que $M(\mathcal{X}, \mathcal{F})$ est muni de la norme duale, ou norme de la variation totale

$$\|\mu\| := \sup_{f: |f| \leq 1} \left| \int \mu(dx) f(x) \right| = \sup_{A \in \mathcal{F}} \mu(A) - \inf_{A \in \mathcal{F}} \mu(A).$$

Une fonction de transition définit donc, par l'intermédiaire de l'opérateur P , un système dynamique sur l'espace $M(\mathcal{X}, \mathcal{F})$ ou $B(\mathcal{X}, \mathcal{F})$; un système dynamique étant défini par un triplet (T, X, d) , où (X, d) est un espace métrique et $T : X \rightarrow X$, un opérateur continu. Une propriété importante d'un tel système dynamique est la structure des trajectoires $T^k x : k \in \mathbb{N}$ où x est la condition initiale. Dans le cas

d'une chaîne de Markov, posons $P^{(n)}(x, A) = P(X_n \in A | X_0 = x)$, alors nous obtenons

$$P^{(n+1)}(x, A) = \int_X P^{(n)}(y, A)P(x, dy),$$

$$\mu P^n(F) = \int \mu(dx)P^{(n)}(x, F).$$

L'étude du comportement asymptotique de $P^{(n)}(x, A)$ et de l'existence d'une probabilité invariante π pour la chaîne de Markov est l'objet de la théorie ergodique. Un premier mode de comportement, en rapport avec la loi des grands nombres, porte sur l'existence d'une limite de la suite $\mu_n = n^{-1}\mu(P + P^2 + \dots + P^n)$ pour tout μ . Cette convergence forte de la moyenne de Cesaro des P^k , pour la norme définie sur $M(\mathcal{X}, \mathcal{F})$ est l'objet du théorème ergodique en moyenne. Notons que si $\pi = \lim_{n \rightarrow \infty} \mu_n$ alors $\pi P = \pi$, π est donc une probabilité invariante. Un autre résultat plus fort, connu sous le nom de théorème ergodique uniforme [52] ou théorème ergodique fort [44], concerne la convergence en norme des itérés de P . Il implique comme précédemment l'existence d'une probabilité invariante π , mais cette fois indépendante de la mesure initiale μ . Deux approches permettent d'obtenir des conditions nécessaires et/ou suffisantes sur la fonction de transition, pour que le théorème ergodique uniforme soit vérifié : une approche probabiliste traitée dans [41] et dans [17] et une approche opérateur développée dans [52] et [44].

L'emploi des termes "ergodique" et "ergodicité uniforme" dans le contexte des chaînes de Markov peut porter à confusion, selon que nous étudions le comportement asymptotique de $P^{(n)}(x, A)$, celui de la moyenne de Cesaro des itérés de P ou une transformation laissant invariante une mesure fixée a priori. Par exemple dans [41], une chaîne de Markov sera dite uniformément ergodique si $\sup_x \|P^{(n)}(x, \cdot) - \pi\| \rightarrow 0$ quand $n \rightarrow \infty$. Cette propriété est un cas particulier du théorème ergodique uniforme, puisqu'il implique l'apériodicité et l'irréductibilité de la chaîne. De même, les propriétés mélangeantes d'une chaîne de Markov font référence à l'ergodicité d'une chaîne de Markov stationnaire (la distribution initiale étant alors la probabilité invariante π) lorsque la tribu des événements invariants (ou asymptotiques) par l'opérateur de translation est triviale. Dans la suite, nous utiliserons le terme d'ergodicité uniforme conformément à la définition donnée dans [41].

1. Point de vue probabiliste

Une chaîne de Markov définie sur un espace d'états dénombrable vérifiera le théorème ergodique uniforme et admettra une unique probabilité invariante si elle est irréductible, récurrente et positive. Dans ce cas, la condition d'irréductibilité se définit à partir de la relation d'équivalence définie entre deux états quelconques par, $x \leftrightarrow y$ si et seulement s'il existe $n(x, y) \in \mathbb{N}$ et $m(y, x) \in \mathbb{N}$ tels que $P^{n(x, y)}(x, y) > 0$ et $P^{m(y, x)}(y, x) > 0$. Une chaîne de Markov sera dite irréductible sur \mathcal{X} dénombrable, si \mathcal{X} contient une seule classe d'équivalence. Malheureusement, cette relation ne peut pas se définir pour une chaîne de Markov définie sur un espace quelconque, car $P(x, \{y\})$ est en général nul. Le concept similaire, développé dans [41] est celui de φ -irréductibilité. Pour cela introduisons les notations suivantes :

$$\tau_A := \min\{n \geq 1 : X_n \in A\}, \quad A \in \mathcal{F},$$

$$L(x, A) := P_x(\tau_A < \infty), \quad x \in \mathcal{X}, A \in \mathcal{F}.$$

DÉFINITION 1.1. Une chaîne de Markov X sera dite φ -irréductible s'il existe une mesure σ -finie φ sur \mathcal{F} telle que $L(x, A) > 0$ pour tout $x \in \mathcal{X}$ dès que $\varphi(A) > 0$.

L'hypothèse de φ -irréductibilité signifie que les ensembles non négligeables pour la mesure φ sont toujours accessibles, indépendamment du point initial. La proposition 4.2.2 dans [41] prouve que si X est φ -irréductible, il existe une mesure ψ maximale telle que X soit ψ -irréductible; i.e. que toute mesure φ' telle que X soit φ' -irréductible est absolument continue par rapport à ψ ($\varphi' \prec \psi$). D'autre part si $\psi(A) = 0$, alors $\psi\{y : L(y, A) > 0\} = 0$. Par conséquent il est toujours possible d'étendre la mesure φ de telle façon que les ensembles négligeables (pour la mesure ψ) ne soient pas accessibles, avec probabilité un. Nous désignerons par $\mathcal{F}^+ := \{B \in \mathcal{F} : \psi(B) > 0\}$.

Comme dans le cas dénombrable, la récurrence se définit en fonction du nombre de passages η_A de la chaîne dans un ensemble A

$$\eta_A := \sum_{n=1}^{\infty} \mathbf{1}(X_n \in A).$$

- DÉFINITION 1.2. (1) Un ensemble A est dit uniformément transient s'il existe $M < \infty$ tel que $E_x[\eta_A] \leq M$ pour tout $x \in A$. Si $A \in \mathcal{F}$ peut être recouvert par un nombre fini d'ensembles uniformément transients alors A sera appelé ensemble transient.
- (2) Un ensemble A sera dit récurrent si $E_x[\eta_A] = \infty$, pour tout $x \in A$.
- (3) Une chaîne de Markov X sera dite récurrente si elle est ψ -irréductible et $E_x[\eta_A] = \infty$ pour tout $x \in X$ et tout $A \in \mathcal{F}^+$
- (4) Une chaîne de Markov X sera dite transiente si elle est ψ -irréductible et X est un ensemble transient

Selon le théorème 8.2.5 dans [41], si X est ψ -irréductible, alors elle est soit récurrente, soit transiente. Nous obtenons donc un résultat analogue aux chaînes de Markov irréductibles définies sur un espace dénombrable.

Ces notions, comme dans le cas dénombrable, impliquent l'existence d'une mesure invariante pour la chaîne, condition nécessaire pour que les théorèmes ergodiques soient vérifiés. En effet d'après le théorème 10.4.4 dans [41] la récurrence de la chaîne implique l'existence et l'unicité (à une constante multiplicative près) d'une mesure invariante. Si la mesure invariante est finie, nous dirons que la chaîne est récurrente positive.

Historiquement, l'une des conditions, garantissant l'ergodicité uniforme pour une chaîne de Markov, est la condition de Doeblin. Cette condition introduite, en 1937, par Doeblin est traitée en détail dans [17]. Cette condition impose une certaine uniformité en x , sur la petitesse de $P(x, A)$ pour des petits ensembles A . La condition de Doeblin est la suivante :

HYPOTHÈSE (D) : Il existe une mesure φ finie sur $(\mathcal{X}, \mathcal{F})$ telle que $\varphi(\mathcal{X}) > 0$, un entier $\nu \geq 1$ et un réel positif ϵ , tels que

$$P^{(\nu)}(x, A) \leq 1 - \epsilon, \quad \text{si} \quad \varphi(A) \leq \epsilon.$$

Cette condition peut aussi se formuler de la façon suivante [41][§16.2] :

Il existe une probabilité φ , un entier m , $\epsilon < 1, \delta > 0$ tels que pour tout $x \in \mathcal{X}$

$$\varphi(A) > \epsilon \implies P^{(m)}(x, A) \geq \delta.$$

La condition de Doeblin n'implique pas la φ -irréductibilité; en effet d'après la proposition 4.2.1 dans [41] la φ -irréductibilité est équivalente à : pour tout $x \in \mathcal{X}$, dès que $\varphi(A) > 0$ il existe un entier $n(x, A) > 0$ tel que $P^{(n)}(x, A) > 0$; alors que l'hypothèse (D) ne porte que sur les ensembles vérifiant $\varphi(A) > \epsilon$.

La condition de Doeblin sera traitée en détail ci-dessous, en raison de son équivalence avec la propriété de quasi-compacité de l'opérateur markovien P , propriété

utilisée dans l'approche opérateur du théorème ergodique uniforme. Un opérateur linéaire P sera dit quasi-compact, s'il existe un itéré P^n de P dont la distance à l'ensemble des opérateurs compacts est strictement inférieure à 1 (cf définition 2.1). Cette condition implique également certaines propriétés du spectre de P , qui permettront d'étendre à cet opérateur la théorie des perturbations développée en dimension finie. Nous montrerons qu'une chaîne de Markov apériodique, ψ -irréductible et satisfaisant l'hypothèse (D) est uniformément ergodique. En réalité, il y a équivalence entre ces deux propriétés [41][th. 16.2.3].

La récurrence peut également se définir d'une façon différente, qui conduit à la notion de chaîne Harris récurrente. Un ensemble A sera dit Harris récurrent si $Q(x, A) := P_x(\eta_A = \infty) = 1$ pour tout $x \in A$. Une chaîne de Markov sera dite Harris récurrente si elle est ψ -irréductible et si chaque ensemble de \mathcal{F}^+ est Harris récurrent. La proposition 9.1.1 dans [41] prouve que si $L(x, A) = 1$ pour tout $x \in A$, alors $Q(x, A) = L(x, A)$ pour tout $x \in \mathcal{X}$ et donc A est Harris récurrent. Ceci signifie que la définition d'une chaîne Harris récurrente en fonction de Q est équivalente à une définition en fonction de L . Un ensemble Harris récurrent est récurrent, donc la condition de Harris récurrence est plus forte que celle de récurrence. Réciproquement, d'après le théorème 9.1.5 dans [41], si une chaîne X est récurrente alors l'espace d'états \mathcal{X} se décompose en $\mathcal{X} = H \cup N$, où N est transient et ψ -nul (i.e. $\psi(N) = 0$) et H est absorbant (i.e. $P(x, H) = 1, x \in H$), non vide et tel que tout sous-ensemble de $H \cap \mathcal{F}^+$ est Harris récurrent. Par conséquent une chaîne récurrente diffère d'une chaîne Harris récurrente seulement par un ensemble ψ -nul. Pour un espace dénombrable, N est vide et donc une chaîne récurrente est Harris récurrente. Les conditions de récurrence et de Harris récurrence impliquent une propriété d'ergodicité en convergence forte ; implication énoncée par le théorème 13.3.4 dans [41]

THÉORÈME 1.3. (1) *Si X est Harris récurrente et positive de période $d \geq 1$ alors pour toute distribution initiale λ*

$$\|d^{-1} \sum_{i=0}^{d-1} \lambda P^{nd+i} - \pi\| \rightarrow 0, \quad n \rightarrow \infty,$$

(2) *Si X est positive récurrente de période $d \geq 1$ alors il existe un ensemble N π -nul tel que pour toute mesure initiale λ vérifiant $\lambda(N) = 0$*

$$\|d^{-1} \sum_{i=0}^{d-1} \lambda P^{nd+i} - \pi\| \rightarrow 0, \quad n \rightarrow \infty.$$

1.1. Condition de Doeblin. Nous allons étudier le comportement asymptotique de $P^{(n)}(x, A)$, lorsque la condition de Doeblin est satisfaite, et plus particulièrement, le partitionnement de l'espace \mathcal{X} en sous-ensembles absorbants de mesure non nulle par rapport à φ .

Notons que si (D) est satisfaite pour un triplet (φ, ν, ϵ) , alors elle est satisfaite pour tout $n > \nu$, car si $\varphi(A) \leq \epsilon$,

$$P^{\nu+m}(x, A) = \int P^{(\nu)}(y, A) P^{(m)}(x, dy) \leq (1 - \epsilon) P^{(m)}(x, \mathcal{X}) = 1 - \epsilon.$$

DÉFINITION 1.4. Un ensemble E sera dit consécutif à x_0 si pour tout $n \geq 1$, $P^{(n)}(x_0, E) = 1$. La condition (D) implique que, si un ensemble E est consécutif à un point de \mathcal{X} , alors $\varphi(E) > \epsilon$. Un ensemble E sera dit invariant s'il est consécutif à chacun de ses points.

Un ensemble consécutif à un point x_0 est donc un ensemble accessible de ce point dès la première transition, avec probabilité un. Cette définition est plus stricte

que la définition d'un ensemble E accessible à partir de $\{x_0\}$ donnée dans [41][p. 91]. Un ensemble invariant est récurrent, il est soit vide soit de mesure $\varphi(E) \geq \epsilon$. Si E est consécutif à x_0 et si F_n désigne l'ensemble des $y \in E$ tels que $P^{(n)}(y, E) < 1$, alors $E \setminus F_n$ est consécutif à x_0 , car sinon $P^{(m)}(x_0, F_n) > 0$ pour un certain m et alors

$$\begin{aligned} 1 = P^{(m+n)}(x_0, E) &= \int_{F_n} P^{(n)}(y, E)P^{(m)}(x_0, dy) + \int_{F_n^c} P^{(n)}(y, E)P^{(m)}(x_0, dy) \\ &< P^{(m)}(x_0, F_n) + P^{(m)}(x_0, E \setminus F_n) \\ &= P^{(m)}(x_0, E) = 1 \end{aligned}$$

Intuitivement, ceci signifie que, si partant de x_0 , la trajectoire de la chaîne arrive dans F_n , alors au bout de $n + 1$ étapes, la probabilité de sortir de E est non nulle, en contradiction avec la définition de E . Donc si E est consécutif à un point, il contient l'ensemble non vide invariant $\Lambda = \bigcap_1^\infty (E \setminus F_n)$. Notons que l'hypothèse (D) n'intervient que pour assurer que les ensembles consécutifs, à un point donné, sont de φ -mesure non nulle.

DÉFINITION 1.5. Un ensemble invariant sera dit minimal, s'il ne contient aucun sous-ensemble invariant non vide de φ -mesure non nulle.

Chaque ensemble invariant E contient un sous-ensemble minimal. En effet, celui-ci s'obtient en prenant la limite inférieure de la suite, décroissante par inclusion, des sous-ensembles invariants inclus dans E , celle-ci étant de φ -mesure non nulle, car minorée par ϵ .

Si E_1 et E_2 sont invariants, alors $E_1 E_2$ est invariant ; en particulier si E_1 est minimal, alors $E_1 E_2$ est minimal et donc $\varphi(E_1) = \varphi(E_1 E_2)$ à moins que $E_1 E_2 = \emptyset$. Ainsi deux ensembles minimaux sont soit disjoints, soit diffèrent d'un ensemble de φ -mesure nulle. Soit E_1, E_2, \dots , une énumération d'ensembles minimaux non vides, alors $\varphi(E_i E_j) = 0$ ($i \neq j$), $\varphi(E_i) \geq \epsilon$, et si E est minimal, il diffère d'un certain E_j d'au plus un ensemble de φ -mesure nulle. Il y a donc au plus $\varphi(\mathcal{X})/\epsilon$ ensembles minimaux dans \mathcal{X} .

Posons $F = \mathcal{X} \setminus \bigcup_j E_j$, alors pour tout $x \in \mathcal{X}$, $\limsup P^{(n)}(x, \bigcup_j E_j) > 0$, car sinon F serait consécutif à x , en contradiction avec sa construction. Ainsi toute trajectoire du système finira par entrer dans un certain E_i et y restera. Nous avons le théorème suivant ([17] théorème 5-6) :

THÉORÈME 1.6. *Pour tout x ,*

$$\lim_{n \rightarrow \infty} P^{(n)}(x, \bigcup_j E_j) = 1,$$

et plus précisément, il existe $\rho < 1$ tel que

$$1 - P^{(n)}(x, \bigcup_j E_j) \leq \text{const. } \rho^n, \quad n = 1, 2, \dots,$$

uniformément en x . Ainsi chaque trajectoire, issue de x , restera (avec probabilité 1) en dehors des E_j , seulement un nombre fini de transitions.

L'ensemble F est appelé transient dans [17], car $\lim_{n \rightarrow \infty} P^{(n)}(x, F) = 0$, pour tout $x \in \mathcal{X}$. L'étude de la dynamique de notre système peut donc se restreindre à un E_i . Pour chaque a , la restriction de la fonction de transition $P(x, E)$ au sous-ensemble E_a est également une fonction de transition. D'autre part, pour tout E tel que $\varphi(E) > 0$ et tout $x \in E_a$ nous avons

$$\limsup P^{(n)}(x, E) > 0, \quad \text{pour tout } x \in E_a,$$

car sinon $E_a \setminus E$ serait invariant dans E_a qui est minimal par construction. Par conséquent, tout sous-ensemble de E_a est atteignable à partir de tout point de E_a ; E_a est donc absorbant et si $\mathcal{X} = E_a$ la chaîne est φ -irréductible.

Nous avons également les résultats suivants [17] : chaque ensemble minimal E_a se décompose en d_a classes périodiques $C_{a,j}$ caractérisées par

$$P^{(nd_a+m)}(x, E) = P^{(nd_a+m)}(x, EC_{a,\beta}), \quad x \in C_{a,\alpha}, \quad \beta \equiv \alpha + m \pmod{d_a},$$

et

$$\lim_{n \rightarrow \infty} P^{(nd_a+m)}(x, E) = \pi_{a,\beta}(E), \quad x \in C_{a,\alpha}, \quad \beta \equiv \alpha + m \pmod{d_a},$$

où la convergence est uniforme et à vitesse exponentielle, et les $\pi_{a,\beta}$ sont des probabilités sur les ensembles $E \subset E_a$, vérifiant

$$\pi_{a,\beta}(C_{a,\beta}) = 1, \quad \pi_{a,\beta}(E) > 0 \quad \text{si } \varphi(EC_{a,\beta}) > 0.$$

Signalons qu'en général les $C_{a,j}$ ne sont pas disjoints, mais que pour tout n , $P^{(n)}(x, C_{a,j}C_{a,k}) = 0$, $j \neq k$, à l'exception peut-être d'un ensemble de points de φ -mesure nulle, et $\varphi(C_{a,j}C_{a,k}) = 0$. Mais il est possible par une procédure de *réduction*, décrite dans [17][p 204], de se ramener à un noyau "équivalent" pour lequel les classes périodiques sont disjointes.

Si l'état initial est x , la probabilité que le système entre finalement dans l'ensemble invariant E_a est

$$\rho(x, E_a) = \lim_{n \rightarrow \infty} P^{(n)}(x, E_a),$$

où $P^{(n)}(x, E_a)$ est croissant en n . Si $x \in E_a$, alors $\rho(x, E_a) = 1$ et plus généralement, en appliquant le théorème 1.6, $\sum_a \rho(x, E_a) = 1$. D'autre part, conditionnellement à l'état initial x , la probabilité que le système atteigne finalement $C_{a,j}$ est

$$\rho(x, C_{a,j}) = \lim_{n \rightarrow \infty} P^{(nd_a)}(x, C_{a,j}),$$

où $\rho(x, C_{a,j}) = 1$ si $x \in C_{a,j}$. Alors

$$\rho(x, E_a) = \sum_{\alpha=1}^{d_a} \rho(x, C_{a,\alpha}), \quad x \in E_a,$$

d'où pour tout $x \in \mathcal{X}$

$$\lim_{n \rightarrow \infty} P^{(nd_a+m)}(x, E) = \sum_{a,\alpha} \rho(x, C_{a,\alpha}) \pi_{a,\alpha+m}(E),$$

$\alpha + m$ étant calculée mod d_a . La convergence est uniforme et exponentielle. Finalement

$$(1.1) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{m=1}^n P^{(m)}(x, E) = \sum_a \rho(x, E_a) \pi_a(E),$$

où

$$(1.2) \quad \pi_a(E) = \sum_{\alpha=1}^{d_a} \frac{\pi_{a,\alpha}(E)}{d_a} \equiv q(x, E).$$

Ainsi π_a est une probabilité telle que

$$\pi_a(E_a) = 1, \quad \pi_a(E) > 0 \quad \text{si } \varphi(EE_a) > 0.$$

Bien entendu, pour chaque $x \in \mathcal{X}$, la probabilité $q(x, \cdot)$ est invariante pour la chaîne de Markov, et si $x \in E_a$, alors $q(x, E) = \pi_a(E)$ est indépendant de x .

La question que l'on peut se poser concerne la dépendance de la décomposition de \mathcal{X} et des mesures $\pi_{a,\alpha}$ en fonction du triplet (ν, φ, ϵ) . En fait le nombre d'ensembles ergodiques E_a et les $\pi_{a,\alpha}$ sont uniquement déterminés. A partir des π_a ,

il est possible de décomposer \mathcal{X} en définissant \bar{E}_a comme l'ensemble des x tels que $\lim_{n \rightarrow \infty} n^{-1} \sum_{m=1}^n P^{(m)}(x, E) = \pi_a(E)$, pour tout $E \in \mathcal{F}$. De même pour les classes de périodicité, $\bar{C}_{a,\alpha}$ sera l'ensemble des x tels que $\lim_{n \rightarrow \infty} P^{(nd_a)}(x, E) = \pi_{a,\alpha}(E)$. Posons $\bar{E}_a = \bigcup_{\alpha} \bar{C}_{a,\alpha}$, alors

$$\mathcal{X} = \bigcup_a \bar{E}_a \cup \bar{F},$$

où $\bar{F} = \mathcal{X} \setminus \bigcup_a \bar{E}_a$. Si E_1, \dots, F est une autre décomposition de \mathcal{X} alors pour tout i , $E_i \subset \bar{E}_i$ (après permutation éventuelle sur les indices), et $C_{a,\alpha} \subset \bar{C}_{a,\alpha}$.

Supposons la condition (D) vérifiée et définissons la mesure

$$\bar{\varphi}(E) = \sum_a \pi_a(E).$$

Si

$$\begin{aligned} E_a &\subset \bar{E}_a \subset E_a + F \\ C_{a,\alpha} &\subset \bar{C}_{a,\alpha} \subset C_{a,\alpha} + F \end{aligned}$$

avec $\bar{\varphi}(F) = 0$, alors la classe des mesures φ intervenant dans la condition (D) se caractérise par la relation

$$\varphi(E) = \int_E f(y) \bar{\varphi}(dy) + \psi(E),$$

où ψ est une mesure confinée sur l'ensemble transient F et f est une fonction positive et non nulle sur un sous-ensemble de $\bar{\varphi}$ -mesure non nulle de chaque $\bar{C}_{a,\alpha}$. Le théorème suivant résume les résultats obtenus :

THÉORÈME 1.7. *Sous la condition (D), $q(x, E)$ définit, pour chaque x , une probabilité invariante pour le processus de Markov; si $x \in E_a$, $q(\cdot, E)$ est indépendant de x , $q(x, E) = \pi_a(E)$. Réciproquement, toute probabilité invariante est une combinaison linéaire convexe (coefficients positifs de somme 1) des π_a .*

Par conséquent, l'enveloppe convexe des π_a définit l'ensemble des probabilités invariantes d'un processus markovien satisfaisant à la condition de Doeblin.

La condition (D) regroupe plusieurs cas particuliers :

- (1) La limite $q(x, E)$ est indépendante de x si et seulement s'il n'y a qu'un seul ensemble ergodique.
- (2) La limite $q(x, E)$ existe comme limite ordinaire (plutôt qu'en moyenne de Cesaro) si et seulement s'il n'y a aucune classe périodique dans aucun ensemble ergodique ($d_a = 1$ pour tout a).
- (3) La limite $q(x, E)$ est strictement positive pour tout x dès que $\varphi(E) > 0$, si et seulement si l'ensemble transient est de φ -mesure nulle, et s'il n'y a qu'un seul ensemble ergodique. Dans ce cas la chaîne est φ -irréductible et récurrente.

Appelons condition (D_0) , les conditions suivantes

- La condition (D) est satisfaite
- Il n'y a qu'un seul ensemble ergodique et cet ensemble ne contient aucune classe cyclique.

Une chaîne de Markov vérifiant la condition (D_0) est φ -irréductible, apériodique et satisfait à la condition de Doeblin. Dans ce cas, il existe une seule probabilité invariante π , limite ordinaire des $P^{(n)}(x, \cdot)$. De plus il existe des constantes γ et $\rho < 1$ telle que

$$\sup_{x, E} |P^{(n)}(x, E) - \pi(E)| \leq \gamma \rho^n, \quad n = 1, 2, \dots$$

La chaîne est donc uniformément ergodique avec une vitesse de convergence géométrique.

LEMME 1.8. *La condition (D_0) est équivalente à la condition suivante : il existe n tel que*

$$(1.3) \quad \sup_{x,y,E} |P^{(n)}(x,E) - P^{(n)}(y,E)| = \delta < 1.$$

La condition (D_0) implique l'existence d'une probabilité invariante π et des constantes γ, ρ , avec $0 < \rho < 1$, telle que

$$|P^{(n)}(x,E) - \pi(E)| \leq \gamma \rho^n, \quad n = 1, 2, \dots$$

L'introduction des fonctions mesurables bornées $\rho_a(\cdot) = \rho(\cdot, E_a)$ permet d'exprimer les expressions (1.1) et (1.2) sous la forme suivante :

$$\lim_{n \rightarrow \infty} \sup_{x,E} \left| \frac{1}{n} \sum_{m=1}^n P^{(m)}(x,E) - \sum_a \rho_a(x) \pi_a(E) \right| = 0,$$

avec $\rho_a = \sum_{\alpha=1}^{d_a} \rho_{a,\alpha}$ et $\pi_a = d_a^{-1} \sum_{\alpha=1}^{d_a} \pi_{a,\alpha}$. D'autre part, si d est le ppcm des d_a , alors

$$\lim_{n \rightarrow \infty} P^{(nd+m)}(x,E) = \sum_a \sum_{\alpha=1}^{d_a} \rho_{a,\alpha}(x) \pi_{a,\alpha+m}(E),$$

la convergence étant uniforme en x et en E .

1.2. Loi des grands nombres et théorème central limite. La loi forte des grands nombres pour les processus (strictement) stationnaires s'applique aux chaînes de Markov dont la mesure initiale est une probabilité invariante π . Si f est une fonction définie sur \mathcal{X} , \mathcal{F} -mesurable telle que $E_\pi |f| < \infty$, alors la limite $\lim_n n^{-1} \sum_{m=1}^n f(X_m)$ existe avec probabilité 1. Si la condition (D) est vérifiée par le processus de Markov, les probabilités stationnaires générales sont de la forme $\sum_a q_a \pi_a$, avec $0 \leq q_a$ et $\sum_a q_a = 1$, d'où le théorème suivant [17][th. 6.2]

THÉORÈME 1.9. *Soit f une fonction définie sur \mathcal{X} , mesurable par rapport à \mathcal{F} telle que*

$$\int_{E_a} |f(x)| d\pi_a(x) < \infty, \quad a = 1, 2, \dots$$

Alors, sous l'hypothèse (D) et pour toute distribution initiale,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{m=1}^n f(X_m)$$

existe avec probabilité 1, et est donnée par $\int_{E_a} f(x) d\pi_a(x)$, si $X_1 \in E_a$ avec probabilité 1.

Le théorème central limite s'applique également aux chaînes de Markov satisfaisant à la condition de Doeblin [17][th. 7.5] :

THÉORÈME 1.10. *Supposons la condition (D_0) vérifiée et considérons une fonction f définie sur \mathcal{X} à valeurs réelles, mesurable pour \mathcal{F} telle que*

$$E_\pi |f(X_1)|^{2+\delta} < \infty,$$

pour un certain $\delta > 0$. Alors

$$\lim_{n \rightarrow \infty} E_\pi \left\{ \left[\frac{1}{\sqrt{n}} \sum_{m=1}^n (f(X_m) - E_\pi f(X_m)) \right]^2 \right\} = \sigma_1^2$$

existe ; si $\sigma_1^2 > 0$, et si $q(\cdot)$ est une mesure initiale quelconque (de X_1) alors,

$$\lim_{n \rightarrow \infty} P_q \left\{ \frac{1}{\sqrt{n}} \sum_{m=1}^n (f(X_m) - E_\pi f(X_m)) \leq \epsilon \right\} = \frac{1}{\sigma_1 \sqrt{2\pi}} \int_{-\infty}^{\epsilon} e^{-t^2/2\sigma_1^2} dt,$$

uniformément en ϵ .

Le théorème précédent s'applique également lorsque la condition (D) est vérifiée avec un seul ensemble ergodique mais plusieurs classes cycliques. Si plusieurs ensembles ergodiques existent dans \mathcal{X} , la distribution limite sera une moyenne pondérée des distributions normales.

2. Point de vue des opérateurs

Dans la partie précédente, la décomposition ergodique de l'espace \mathcal{X} , sous la condition de Doeblin, fut obtenue par l'utilisation de méthodes probabilistes. Une approche par la théorie des opérateurs linéaires définis sur des espaces de Banach est également possible. Cette approche est particulièrement bien présentée dans l'article [52] de Yosida et Kakutani, lequel développe les liens entre la notion d'opérateur quasi-compact et la théorie ergodique. Cette approche est également traitée par Neveu dans [44][§V.3].

2.1. Théorème ergodique uniforme. Le théorème ergodique uniforme est un résultat sur la convergence en norme des itérés de l'opérateur P . Il décrit également la classe des événements asymptotiques de la chaîne de Markov associée. Dans [52], ce théorème est appelé théorème ergodique uniforme ; la convergence en norme étant appelée convergence uniforme. Dans ce même article est démontré un théorème de convergence forte ("mean ergodic theorem") des itérés de P sous des hypothèses plus faibles. Précisons que ces théorèmes sont valables pour des opérateurs linéaires plus généraux que les opérateurs markoviens, en particulier ils ne supposent pas la positivité de l'opérateur.

Supposons maintenant que la fonction de transition vérifie la condition (D) et montrons que l'opérateur $S = \lim_{n \rightarrow \infty} P^{nd}$ est compact, où d est le ppcm des d_a . Notons que S est un endomorphisme positif, de norme 1 et que pour tout entier m ,

$$SP^{md} = P^{md}S = S.$$

La condition (D) étant vérifiée, nous obtenons l'expression

$$Sf(x) = \sum_a \sum_{\alpha=1}^{d_a} \rho_{a,\alpha}(x) c_{a,\alpha}(f),$$

où $c_{a,\alpha}(f) = \int f(x) \pi_{a,\alpha}(dx)$.

Cette convergence uniforme implique que P^{nd} converge en norme vers S . En effet, notons qu'il est possible d'associer à l'opérateur S la probabilité de transition

$$S(x, dy) = \sum_a \sum_{\alpha=1}^{d_a} \rho_{a,\alpha}(x) \pi_{a,\alpha}(dy),$$

ainsi

$$\begin{aligned} \sup_x |P^{nd}f(x) - Sf(x)| &= \sup_x \left| \int f(y) [P^{(nd)}(x, dy) - S(x, dy)] \right| \\ &\leq \|f\| \sup_x \|P^{nd}(x, \cdot) - S(x, \cdot)\| \end{aligned}$$

où la norme dans le dernier terme de droite est la norme en variation totale. Pour conclure, il suffit d'utiliser la convergence uniforme démontrée pour la condition (D). D'autre part l'opérateur S , engendré par un nombre fini de fonctions $\rho_{a,\alpha}$, est

de rang fini, il est donc compact. Nous avons ainsi montré que si la fonction de transition $P(., .)$ vérifie la condition (D), alors il existe un entier n_0 et un opérateur compact Q tel que $\|P^{n_0} - Q\| < 1$.

DÉFINITION 2.1. Un endomorphisme P , défini sur un espace de Banach B , est dit quasi-compact s'il existe une suite $\{Q_n, n \geq 1\}$ d'endomorphismes compacts Q_n sur B telle que les itérés P^n de P vérifient $\lim_{n \rightarrow \infty} \|P^n - Q_n\| = 0$. Cette définition est équivalente à : il existe un entier $n_0 \geq 1$ et un endomorphisme compact Q tels que $\|P^{n_0} - Q\| < 1$.

Désignons par \mathcal{K} l'ensemble des opérateurs compacts de B dans B . Ce sous-espace vectoriel est fermé, nous avons alors l'équivalence entre P est quasi-compact, il existe n_0 tel que $d(P^{n_0}, \mathcal{K}) < 1$ et $\lim_{n \rightarrow \infty} d(P^n, \mathcal{K}) = 0$. Comme \mathcal{K} est fermé, si la suite des itérés de P (ou une suite extraite) converge en norme vers un opérateur S celui-ci sera forcément compact. Bien entendu si la suite des itérés de P converge vers un opérateur compact, alors P est quasi-compact. Dans [44], il est montré qu'un opérateur quasi-compact admet une décomposition du type de celle trouvée dans le cas de la condition de Doeblin. Plus précisément

THÉORÈME 2.2 (théorème ergodique uniforme). *Soit une probabilité de transition sur l'espace mesurable $(\mathcal{X}, \mathcal{F})$ telle que l'opérateur P associé sur $B(\mathcal{X}, \mathcal{F})$ soit quasi-compact. Il existe alors des entiers $d_a \geq 1$, $\sum d_a$ fonctions mesurables positives bornées et $\sum d_a$ probabilités sur $(\mathcal{X}, \mathcal{F})$ notées respectivement $\rho_{a,\alpha}, \pi_{a,\alpha}$ (où α est un entier modulo d_a) telles que :*

$$\begin{aligned} \rho_{a,\alpha} &= P u_{a,\alpha+1}, \quad \pi_{a,\alpha} = \pi_{a,\alpha-1} P, \\ \int_{\mathcal{X}} \pi_{a',\alpha'}(dx) \rho_{a,\alpha}(x) &= \begin{cases} 1 & \text{si } a' = a \text{ et } \alpha' = \alpha(d_a) \\ 0 & \text{sinon} \end{cases} \\ \sum_a \sum_{\alpha} \rho_{a,\alpha} &= 1, \end{aligned}$$

et telles que si d désigne le ppcm des d_a :

$$\lim_{n \rightarrow \infty} \left\| P^{nd+t} - \sum_a \sum_{\alpha=1}^{d_a} \rho_{a,\alpha-t} \otimes \pi_{a,\alpha} \right\| = 0.$$

Il résulte de cette convergence que :

$$\lim_{n \rightarrow \infty} \left\| \frac{1}{n} \sum_1^n P^s - \sum_a \sum_{\alpha=1}^{d_a} \rho_a \otimes \pi_a \right\| = 0,$$

si on a posé

$$\rho_a = \sum_{\alpha=1}^{d_a} \rho_{a,\alpha}, \quad \text{et} \quad \pi_a = \frac{1}{d_a} \sum_{\alpha=1}^{d_a} \pi_{a,\alpha}.$$

La décomposition de l'espace \mathcal{X} et la forme des événements asymptotiques sont fournies par le corollaire suivant [44][§V.3] :

COROLLAIRE 2.3. *Les ensembles $C_{a,\alpha} := \{\rho_{a,\alpha} = 1\}$ sont disjoints 2 à 2 et sont respectivement des supports de $\pi_{a,\alpha}$. De plus $P(x, C_{a,\alpha+1}) = 1$ si et seulement si $x \in C_{a,\alpha}$; par conséquent*

$$P_x(X_n \in C_{a,\alpha+n}, (n \in \mathbb{N})) = 1 \quad \text{si } x \in C_{a,\alpha}.$$

Si les $A_{a,\alpha}$ désignent les événements asymptotiques 2 à 2 disjoints

$$A_{a,\alpha} = \liminf_n \{X_n \in C_{a,\alpha+n}\},$$

alors

$$P_x(A_{a,\alpha}) = \rho_{a,\alpha}(x), \quad x \in \mathcal{X},$$

et d'ailleurs $A_{a,\alpha} = \lim_{n \rightarrow \infty} p.s. \{X_n \in C_{a,\alpha+n}\}$, pour tout P_x . De plus, tout événement asymptotique est équivalent à une somme d'événements $A_{a,\alpha}$; en particulier $\sum_{a,\alpha} A_{a,\alpha} = \mathcal{X}$, P_x p.s.

Nous en déduisons que si la chaîne de Markov satisfait à la condition (D_0) , qui implique que $a = \alpha = 1$, alors elle sera ergodique dans le sens où la tribu des événements asymptotiques est triviale. La réciproque est fautive en général.

En construisant le triplet φ, ν, ϵ à partir de la décomposition de \mathcal{X} en ensembles ergodiques, comme dans la section précédente, nous obtenons que la condition (D) est vérifiée, d'où la proposition suivante :

PROPOSITION 2.4. *L'opérateur P associé à la fonction de transition $P(x, dy)$, défini sur $B(\mathcal{X}, \mathcal{F})$ par*

$$Pf(x) = \int P(x, dy)f(y)$$

est quasi-compact si et seulement si la fonction de transition vérifie la condition de Doeblin (D).

Le caractère de quasi-compactité pour un opérateur dont le spectre est contenu dans le disque unité, implique certaines propriétés sur ses valeurs propres de module l'unité.

LEMME 2.5. *Soit T un opérateur borné, dont le spectre est contenu dans le disque unité et tel que $\|T^n - Q\| < 1$, pour un certain $n > 0$ et un certain opérateur compact Q . Alors tout $\lambda \in \Sigma(T)$ tel que $|\lambda|^n > \|T^n - Q\|$ est isolé et le projecteur propre correspondant P_λ est de rang fini.*

La preuve figure dans [19][p. 709]. Ainsi pour un opérateur quasi-compact induit par une probabilité de transition, les valeurs propres de module unité sont isolées et les projecteurs propres correspondants sont de rang fini. L'ensemble des points du spectre de module unité est fini, car compact (dans $\Sigma(T)$ compact) et discret.

Si T^n/n converge faiblement vers zéro (i.e. pour tout $x \in \mathcal{X}$ et pour tout $f \in \mathcal{X}^*$, $(n^{-1}T^n x, f)$ converge vers zéro) alors ces points sont des pôles d'ordre 1 de la résolvante de T [19][§ VIII.8.1, lemme 1 et théorème 3]. En particulier ceci sera vérifié si T est un opérateur associé à une probabilité de transition. Notons également que lorsque la condition (D_0) est vérifiée, alors 1 est la seule valeur propre de module unité, et le projecteur propre associé est de rang 1. Celui-ci n'est autre que l'opérateur P_1 défini par $P_1 f(x) = \int f(y)d\pi(y)$, où π désigne l'unique probabilité invariante de la chaîne.

2.2. Conditions en norme L^p et théorème central limite. Les résultats présentés proviennent de [47][§ VII.4]. Nous supposons que la chaîne de Markov admet une probabilité invariante π . Dans la suite les espaces $L^p(\pi)$, seront notés L^p . Il est également possible de définir l'opérateur adjoint de P , agissant dans L^∞ , de la façon suivante : soit $h \in L^\infty$, comme π est invariante, la mesure

$$\nu(A; h) = \int h(x)P(x, A)\pi(dx), \quad A \in \mathcal{F},$$

est absolument continue par rapport à π . $\nu(A; h)$ n'est autre que $\mu P(A)$ où μ est une mesure de densité h par rapport à la mesure π . La dérivée de Radon-Nikodym de $\nu(\cdot; h)$ par rapport à $\pi(\cdot)$ existe. Posons

$$P^* h = \frac{\nu(dx; h)}{\pi(dx)}.$$

Alors P^* est un opérateur positif borné de L^∞ dans L^∞ , tel que $P^*1 = 1$. En outre, π est une probabilité invariante pour P^* et pour tout $f \in L^1$

$$\int f(x)P^*h(x)\pi(dx) = \int f(x)\nu(dx; h) = \int h(y)Pf(y)\pi(dy).$$

Il est possible d'étendre, par une procédure de troncature, P^* à l'espace L^1 de telle façon que π soit invariante pour P^* . Notons que cette construction est compatible avec l'expression classique obtenue dans le cas d'une chaîne de Markov à espace d'états fini; à savoir $P^*(x, y) = P(y, x)\pi(y)/\pi(x)$.

DÉFINITION 2.6. Une chaîne de Markov est dite satisfaisante une condition en norme L^p , $1 \leq p \leq \infty$, (ou condition L^p) si l'opérateur P associé vérifie

$$\lim_{n \rightarrow \infty} \sup_{f \perp 1} \frac{\|P^n f\|_p}{\|f\|_p} \rightarrow 0,$$

où $f \perp 1$ signifie que $\int f(x)\pi(dx) = 0$.

Nous avons alors le théorème suivant ([47] § VII.4 théorème 1) :

THÉORÈME 2.7. *Considérons une fonction de transition P de mesure de probabilité invariante π . Les conditions en norme L^p , $1 < p < \infty$, pour l'opérateur P associé sont toutes équivalentes les unes aux autres. De plus, la condition L^∞ pour l'opérateur P et la condition L^1 pour l'opérateur P^* sont mutuellement équivalentes et sont équivalentes à la condition (D_0) . Cependant elles sont plus fortes que la condition L^2 .*

REMARQUE 2.8. Comme pour $1 < p < \infty$, toutes les conditions L^p sont équivalentes, nous ne considérerons que la condition L^2 . Si la chaîne est réversible, i.e $P = P^*$, alors la condition L^∞ pour l'opérateur P est équivalente à la condition L^∞ pour l'opérateur adjoint.

EXEMPLE 1. Rosenblatt [47] donne un exemple de chaîne de Markov vérifiant la condition (D_0) , mais dont le processus inversé dans le temps ne vérifie pas la condition (D_0) . Par conséquent l'opérateur P (resp. l'opérateur P^*) vérifie la condition L^∞ (resp. la condition L^1), mais pas la condition L^1 (resp. la condition L^∞). Il donne également l'exemple suivant d'une chaîne de Markov réversible satisfaisant à la condition L^2 , mais dont l'opérateur $P = P^*$ ne vérifie ni la condition L^∞ ni la condition L^1 . Considérons une suite (X_n) de variables aléatoires gaussiennes telles que $E(X_n) = 0$ et $E(X_{n+m}X_n) = |\eta|^m$, où $0 < |\eta| < 1$. Cette suite forme une chaîne de Markov sur \mathbb{R} et l'espérance conditionnelle $E(X_{n+1}|X_1 = x_1)$ est une variable aléatoire gaussienne de moyenne ηx_1 et de variance $1 - \eta^{2n}$. La fonction de transition de la chaîne de Markov admet donc pour densité

$$p(x, y) = \frac{1}{\sqrt{2\pi(1-\eta^2)}} \exp\left(-\frac{(y-\eta x)^2}{2(1-\eta^2)}\right),$$

d'où

$$P^{(n)}(x, A) = \frac{1}{\sqrt{2\pi(1-\eta^{2n})}} \int_A \exp\left(-\frac{(y-\eta^n x)^2}{2(1-\eta^{2n})}\right) dy,$$

et

$$\lim_{n \rightarrow \infty} P^{(n)}(x, A) = \frac{1}{\sqrt{2\pi}} \int_A e^{-y^2/2} dy.$$

Cette chaîne de Markov admet donc pour probabilité invariante, la mesure gaussienne, que nous désignerons par γ . Cependant, la convergence n'est pas uniforme car $P^{(n)}(x, A)$ peut être rendu arbitrairement petit pour x suffisamment grand (A étant un intervalle fini). Cette chaîne de Markov est irréductible pour la mesure de Lebesgue, récurrente positive mais ne vérifie pas la condition (D_0) de Doeblin. Par

contre, elle satisfait à une condition L^2 , car en introduisant la base orthonormée des polynômes d'Hermite $h_i(x)$ sur $L^2(\gamma)$, nous obtenons l'expression suivante

$$p(x, y) = \sum_{i=0}^{\infty} \eta^i h_i(x) h_i(y) \frac{e^{-y^2/2}}{\sqrt{2\pi}}.$$

Par conséquent, pour toute fonction $f \in L^2(\gamma)$ centrée nous avons

$$\|P^n f\|_2^2 = \sum_{i=1}^{\infty} \eta^{2ni} \langle h_i, f \rangle^2 \leq \eta^{2n} \|f\|_2^2.$$

Plus précisément, $\sup_{f \perp 1} \|P^n f\|_2 / \|f\|_2 = |\eta|^n$ tend vers zéro quand n tend vers l'infini. Notons également que l'opérateur P est un opérateur de Hilbert-Schmidt dans $L^2(\gamma)$ car

$$\int p(x, y)^2 \gamma(dx) \gamma(dy) \leq \frac{1}{2\pi(1-\eta^2)} < \infty,$$

par conséquent il est compact dans $L^2(\gamma)$. Cependant, il ne peut être compact dans $B(\mathbb{R}, \mathcal{B})$ où \mathcal{B} désigne la tribu des Boréliens dans \mathbb{R} . En effet, il serait alors quasi-compact et vérifierait alors la condition de Doeblin, en contradiction avec l'observation précédente.

EXEMPLE 2. B. Mann [39] donne un exemple de chaîne de Markov finie qui ne vérifie pas la condition L^2 . Soit (X_n) une chaîne de Markov irréductible réversible et apériodique définie sur un espace d'états fini G . Soient P son noyau de transition et π son unique probabilité invariante; considérons alors la chaîne de Markov (Y_n) définie sur l'espace d'états

$$S = \left\{ (x, y) \in G \times G : P(x, y) > 0 \right\},$$

par $Y_n = (X_n, X_{n+1})$. Cette chaîne est également irréductible et apériodique, mais non nécessairement réversible. Son noyau de transition s'exprime par

$$\tilde{P}((x, y), (z, \omega)) = \begin{cases} P(z, \omega) & \text{si } y = z, \\ 0 & \text{sinon} \end{cases}$$

et la probabilité invariante est

$$\mu(x, y) = \pi(x) P(x, y).$$

Soit une fonction $g : G \rightarrow \mathbb{R}$ de moyenne nulle pour la mesure π , construisons la fonction $f : S \rightarrow \mathbb{R}$ par $f(x, y) = g(x)$. Cette fonction f est de moyenne nulle pour la mesure μ et $\|f\|_{2, \mu} = \|\tilde{P}f\|_{2, \mu}$. La chaîne de Markov (Y_n) ne satisfait donc pas à une condition L^2 , la condition de Doeblin (D_0) n'est donc pas satisfaite. Signalons que le symétrisé multiplicatif n'est pas irréductible; les sous-ensembles ergodiques sont déterminés par la première coordonnée du couple (x, y) . Le théorème 2.1 ne s'applique pas dans ce cas, néanmoins le symétrisé de \tilde{P}^2 est irréductible. Cette remarque est générale, si 1 est une valeur propre multiple du symétrisé multiplicatif d'un opérateur markovien P alors celui-ci ne vérifie pas de condition L^2 .

Nous avons vu dans la section précédente que la condition (D_0) implique le théorème central limite. En fait la condition L^2 , qui est plus faible, suffit à obtenir un théorème central limite. Considérons une chaîne de Markov stationnaire, de probabilité invariante π , satisfaisant à la condition L^2 et une suite de fonctions réelles \mathcal{F} -mesurables

$$f_{k,n}, \quad k = 1, \dots, k_n, \quad n = 1, 2, \dots$$

telles que $E_\pi f_{k,n} = 0$ et $k_n \rightarrow \infty$, quand $n \rightarrow \infty$. Posons $S_n = \sum_{k=1}^{k_n} f_{k,n}(X_k)$, nous avons le théorème suivant [47] :

THÉORÈME 2.9. Soit $(X_k; k = 0 \pm 1, \dots)$ une chaîne de Markov stationnaire de probabilité invariante π . Si les fonctions $f_{k,n}$ sont uniformément de carré intégrable en k, n avec

$$E_\pi |S_n|^2 \cong k_n \sigma^2, \quad \sigma^2 > 0,$$

quand $n \rightarrow \infty$, et si le processus satisfait à la condition L^2 , alors

$$\lim_{n \rightarrow \infty} P\{k_n^{-1/2} \sigma^{-1} S_n \leq x\} = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du.$$

Dans le cas $f_{k,n} = f_n$, et $k_n = n$, B. Mann [39] donne une borne explicite du type Berry-Esséen pour une chaîne de Markov définie sur un espace dénombrable et vérifiant la condition

$$\sup_{f \perp 1} \frac{\|Pf\|_2}{\|f\|_2} = \beta < 1.$$

Cette condition implique une condition L^2 , car si nous posons

$$\varphi(n) := \sup_{f \perp 1} \frac{\|P^n f\|_2}{\|f\|_2},$$

alors $\varphi(n+m) \leq \varphi(n)\varphi(m)$. Par conséquent si $\varphi(1) \leq \beta$ alors $\varphi(n) \leq \beta^n$. Notons que la réciproque est également vraie ([47][p. 216]), à savoir qu'une condition L^2 est satisfaite si et seulement s'il existe deux constantes K et $0 < \eta < 1$ telles que pour tout n

$$\sup_{f \perp 1} \frac{\|P^n f\|_2}{\|f\|_2} \leq K\eta^n.$$

B. Mann [39] montre que la chaîne de Markov construite dans l'exemple 2 de la section précédente ne vérifie pas le théorème central limite. En effet, soit la fonction $f(x, y) = g(x) - g(y)$ où g est une fonction à valeurs réelles définie sur G et de moyenne nulle. f est de moyenne nulle pour la probabilité invariante μ et $S_n = g(X_0) - g(X_n)$.

3. Point de vue du mélange

Les résultats présentés dans cette partie proviennent en partie de [18]. Ce point de vue, qui dépasse le cadre des chaînes de Markov, traite des propriétés des suites de variables aléatoires dépendantes et stationnaires.

3.1. Mesures de dépendance. Soient $(\mathcal{X}, \mathcal{F}, P)$ un espace probabilisé et deux sous tribus \mathcal{A}, \mathcal{B} de \mathcal{F} . Pour mesurer la dépendance de ces deux tribus, on définit plusieurs coefficients :

$$(3.1) \quad \alpha(\mathcal{A}, \mathcal{B}) = \sup |P(AB) - P(A)P(B)|, \quad A \in \mathcal{A}, B \in \mathcal{B}$$

$$(3.2) \quad \phi(\mathcal{A}, \mathcal{B}) = \sup |P(B|A) - P(B)|, \quad A \in \mathcal{A}, B \in \mathcal{B}, P(A) > 0$$

$$(3.3) \quad \phi_{\text{rev}}(\mathcal{A}, \mathcal{B}) = \phi(\mathcal{B}, \mathcal{A})$$

$$(3.4) \quad \psi(\mathcal{A}, \mathcal{B}) = \sup \frac{|P(AB) - P(A)P(B)|}{P(A)P(B)}, \quad A \in \mathcal{A}, B \in \mathcal{B}$$

$$(3.5) \quad \rho(\mathcal{A}, \mathcal{B}) = \sup |\text{Corr}(X, Y)|, \quad X \in L^2(\mathcal{A}), Y \in L^2(\mathcal{B})$$

$$(3.6) \quad \beta(\mathcal{A}, \mathcal{B}) = \sup \frac{1}{2} \sum_{i=1}^I \sum_{j=1}^J |P(A_i B_j) - P(A_i)P(B_j)|$$

où dans la dernière expression, le supremum est pris sur toutes les partitions (A_i) , (B_j) de \mathcal{X} avec $A_i \in \mathcal{A}$, et $B_j \in \mathcal{B}$. En particulier, notons par $P_{\mathcal{A}}$ (resp. $P_{\mathcal{B}}$) la probabilité restreinte à la tribu \mathcal{A} (resp. \mathcal{B}), alors le coefficient $\beta(\mathcal{A}, \mathcal{B})$ peut s'exprimer par

$$\beta(\mathcal{A}, \mathcal{B}) = \|P_{\mathcal{A}} \otimes P_{\mathcal{B}} - P_{\mathcal{A} \otimes \mathcal{B}}\|_{\text{Var}}.$$

Ces divers coefficients sont reliés entre eux par les relations suivantes :

$$\begin{aligned} 2\alpha(\mathcal{A}, \mathcal{B}) &\leq \beta(\mathcal{A}, \mathcal{B}) \leq \phi(\mathcal{A}, \mathcal{B}) \leq \frac{1}{2}\psi(\mathcal{A}, \mathcal{B}), \\ 4\alpha(\mathcal{A}, \mathcal{B}) &\leq \rho(\mathcal{A}, \mathcal{B}) \leq \psi(\mathcal{A}, \mathcal{B}), \\ \rho(\mathcal{A}, \mathcal{B}) &\leq 2\phi^{1/2}(\mathcal{A}, \mathcal{B}) \cdot \phi_{\text{rev}}^{1/2}(\mathcal{A}, \mathcal{B}), \\ \rho(\mathcal{A}, \mathcal{B}) &= \sup \left\{ \|E(f|\mathcal{B}) - Ef\|_2 / \|f\|_2, f \in L^2(\mathcal{A}), f \text{ réelle} \right\}, \\ \alpha(\mathcal{A}, \mathcal{B}) &\leq 1/4, \beta(\mathcal{A}, \mathcal{B}) \leq 1, \phi(\mathcal{A}, \mathcal{B}) \leq 1, \rho(\mathcal{A}, \mathcal{B}) \leq 1. \end{aligned}$$

D'autre part si X, Y désignent deux v.a.r. mesurables respectivement par rapport à \mathcal{A} et \mathcal{B} , la covariance entre X et Y vérifie les inégalités suivantes :

$$\begin{aligned} |\text{Cov}(X, Y)| &\leq 8\alpha^{1/r}(\mathcal{A}, \mathcal{B}) \|X\|_p \|Y\|_q, \quad p, q, r \geq 1, \quad \frac{1}{p} + \frac{1}{q} + \frac{1}{r} = 1, \\ |\text{Cov}(X, Y)| &\leq 2\phi^{1/p}(\mathcal{A}, \mathcal{B}) \|X\|_p \|Y\|_q, \quad p, q \geq 1, \quad \frac{1}{p} + \frac{1}{q} = 1, \\ |\text{Cov}(X, Y)| &\leq \psi(\mathcal{A}, \mathcal{B}) \|X\|_1 \|Y\|_1, \\ |\text{Cov}(X, Y)| &\leq \rho(\mathcal{A}, \mathcal{B}) \|X\|_2 \|Y\|_2. \end{aligned}$$

3.2. Conditions de mélange. Soit $(X_k, k \in \mathbb{Z})$ une suite de v.a.r., introduisons pour tout $-\infty \leq n \leq m \leq \infty$ la tribu \mathcal{F}_n^m engendrée par les $X_k, n \leq k \leq m$. On définit alors pour tout $n \leq 1$ les quantités suivantes :

$$\begin{aligned} \alpha(n) &:= \sup_{m \in \mathbb{Z}} \alpha(\mathcal{F}_{-\infty}^m, \mathcal{F}_{m+n}^\infty), \\ \phi(n) &:= \sup_{m \in \mathbb{Z}} \phi(\mathcal{F}_{-\infty}^m, \mathcal{F}_{m+n}^\infty), \\ \psi(n) &:= \sup_{m \in \mathbb{Z}} \psi(\mathcal{F}_{-\infty}^m, \mathcal{F}_{m+n}^\infty), \\ \rho(n) &:= \sup_{m \in \mathbb{Z}} \rho(\mathcal{F}_{-\infty}^m, \mathcal{F}_{m+n}^\infty), \\ \beta(n) &:= \sup_{m \in \mathbb{Z}} \beta(\mathcal{F}_{-\infty}^m, \mathcal{F}_{m+n}^\infty) \end{aligned}$$

La suite (X_k) sera dite

- fortement mélangeante si $\lim_{n \rightarrow \infty} \alpha(n) = 0$,
- ϕ -mélangeante si $\lim_{n \rightarrow \infty} \phi(n) = 0$,
- ψ -mélangeante si $\lim_{n \rightarrow \infty} \psi(n) = 0$,
- ρ -mélangeante si $\lim_{n \rightarrow \infty} \rho(n) = 0$,
- absolument régulière si $\lim_{n \rightarrow \infty} \beta(n) = 0$.

Il est possible d'adapter les définitions précédentes aux suites (X_k) définies seulement pour les $k \in \mathbb{N}$, en modifiant la définition de $\alpha(n)$ de la façon suivante : $\alpha(n) = \sup_{m \geq 1} \alpha(\mathcal{F}_1^m, \mathcal{F}_{m+n}^\infty)$; de même pour les autres coefficients.

REMARQUE 3.1. Si la suite $(X_k, k \in \mathbb{Z})$ est stationnaire, chaque coefficient peut être défini plus simplement; par exemple de la façon suivante pour α ,

$$\alpha(n) := \alpha(\mathcal{F}_{-\infty}^0, \mathcal{F}_n^\infty).$$

Une suite stationnaire (X_1, X_2, \dots) peut être prolongée en une suite $(X_k, k \in \mathbb{Z})$ également stationnaire. Les diverses propriétés de mélange se conservent avec les mêmes coefficients.

La condition de ϕ -mélange n'est pas nécessairement préservée lors d'un changement de direction du "temps" dans la suite $(X_k, k \in \mathbb{Z})$. En effet, celle-ci est caractérisée par $\phi_{\text{rev}}(n) := \sup_{m \in \mathbb{Z}} \phi(\mathcal{F}_{m+n}^\infty, \mathcal{F}_{-\infty}^m)$. Il existe des exemples de chaînes de Markov à espace d'états fini, stationnaires telles que $\phi(n) \rightarrow 0$ quand $n \rightarrow \infty$ et $\phi_{\text{rev}}(n) = 1$, pour tout $n \geq 1$.

Nous avons les implications suivantes entre les diverses conditions de mélange

$$\psi - \text{mélange} \implies \phi - \text{mélange} \implies \begin{cases} \beta - \text{mélange} & \implies \alpha - \text{mélange}, \\ \rho - \text{mélange} & \implies \alpha - \text{mélange}. \end{cases}$$

Signalons qu'il y a équivalence pour les processus stationnaires Gaussien entre les propriétés de ρ -mélange et de α -mélange. En effet, un théorème de Kolmogorov-Rozanov démontre la double inégalité suivante :

$$\alpha(n) \leq \rho(n) \leq 2\pi\alpha(n).$$

D'autres conditions, dites de mélange, sont définies dans le cadre de la théorie ergodique. Soit $X = (X_n)_{n \in \mathbb{Z}}$ une suite stationnaire, considérons alors l'espace de probabilité $(\mathbb{R}^{\mathbb{Z}}, \mathcal{B}^{\mathbb{Z}}, P)$ où \mathcal{B} désigne la tribu des boréliens de \mathbb{R} et P la distribution de X . Soit θ l'opérateur de translation défini par $(\theta X)_n = X_{n+1}$.

- (X_n) sera dite mélangeante (au sens de la théorie ergodique) si pour tout $A, B \in \mathcal{F}_{-\infty}^{\infty} := \mathcal{B}^{\mathbb{Z}}$

$$\lim_{n \rightarrow \infty} P(A \cap \theta^n B) = P(A)P(B).$$

- (X_n) sera dite régulière si la tribu $\bigcap_{n=1}^{\infty} \mathcal{F}_{-\infty}^{-n}$ est triviale (i.e. ne contient que des événements de probabilité 0 ou 1). D'après le théorème 17.1.1 dans [30], une condition nécessaire et suffisante pour qu'une suite stationnaire soit régulière, est que pour tout $B \in \mathcal{F}_{\infty} := \mathcal{F}_{-\infty}^{\infty}$,

$$\lim_{n \rightarrow -\infty} \sup_{A \in \mathcal{F}_{-\infty}^n} |P(AB) - P(A)P(B)| = 0.$$

- (X_n) sera dite ergodique si la tribu des événements invariants par θ est triviale.

Nous avons alors les implications

$$\alpha - \text{mélange} \implies \text{régulière} \implies \text{ergodique mélangeante} \implies \text{ergodicité}.$$

3.3. Mélange et théorème de la limite centrale. Soit $X = (X_n, n \in \mathbb{Z})$ une suite stationnaire centrée, posons $S_n = X_1 + \dots + X_n$, et $\sigma_n^2 = E|S_n|^2$. Si X est ϕ -mélangeante et si $\lim_{n \rightarrow \infty} \sigma_n^2 = \infty$, alors $\sigma_n^2 = nL(n)$, où $L(n)$ est une fonction variant lentement, i.e. que pour tout $x \geq 0$, $\lim_{t \rightarrow \infty} L(tx)/L(t) = 1$ [30][th. 18.2.3].

Dans le cas d'une suite stationnaire ϕ -mélangeante les théorèmes 18.5.1 et 18.5.2 dans [30] fournissent une condition suffisante pour que X satisfasse un théorème de la limite centrale

THÉORÈME 3.2. *Soit X une suite stationnaire ϕ -mélangeante.*

- Si $E|X_j|^{2+\delta} < \infty$ pour un certain $\delta > 0$ et $\sigma_n^2 \rightarrow \infty$, quand $n \rightarrow \infty$, alors S_n/σ_n converge vers une variable aléatoire de loi normale.
- Si $\sum_n \phi(n)^{1/2} < \infty$, alors σ^2/n converge vers

$$\sigma^2 = E(X_0^2) + 2 \sum_{k=1}^{\infty} E(X_0 X_k) < \infty.$$

Si $\sigma^2 \neq 0$, alors $S_n/(\sigma n^{1/2})$ converge vers une variable aléatoire de loi normale.

Pour une suite stationnaire α -mélangeante, nous avons le théorème suivant [30][th. 18.5.3, th. 18.5.4] :

THÉORÈME 3.3. *Soit X une suite stationnaire α -mélangeante.*

- Si $E|X_j|^{2+\delta} < \infty$ pour un certain $\delta > 0$ et

$$\sum_{n=0}^{\infty} \alpha(n)^{\delta/(\delta+2)} < \infty,$$

pour un certain $\delta > 0$, alors σ_n^2/n converge vers $\sigma^2 < \infty$ (défini au théorème précédent). Si de plus $\sigma^2 \neq 0$ alors $S_n/(\sigma n^{1/2})$ converge vers une variable aléatoire de loi normale.

- Si les X_j sont bornées, i.e. qu'il existe une constante c telle $P(|X_j| < c) = 1$ alors $\sigma^2 < \infty$, et si

$$\sum_{n=1}^{\infty} \alpha(n) < \infty,$$

alors lorsque $\sigma^2 \neq 0$, $S_n/(\sigma n^{1/2})$ converge vers une variable aléatoire de loi normale.

Le cas d'une suite ρ -mélangeante fait l'objet du théorème suivant [31, 32] :

THÉORÈME 3.4. *Une suite stationnaire X vérifie les propriétés suivantes :*

- (1) Si $\lim_{n \rightarrow \infty} \rho(n) = 0$, alors $\sup_n \sigma_n^2 < \infty$, ou $\sigma_n^2 = nL(n)$, $L(n)$ étant une fonction variant lentement. Si $\lim_{n \rightarrow \infty} \sigma_n = \infty$ et $E|X_j|^{2+\delta} < \infty$ pour un certain $\delta > 0$, alors S_n/σ_n converge vers une variable aléatoire de loi normale.
- (2) Si $\sum_{n=1}^{\infty} \rho(2^n)^{1/2} < \infty$, alors $\sigma_n^2 = 2\pi n f(0)(1 + o(1))$, où f est la densité spectrale de X et $f(0) \neq 0$. De plus S_n/σ_n converge vers une variable aléatoire de loi normale.

Nous verrons ci-dessous qu'une chaîne de Markov vérifiant la condition (D_0) est ϕ -mélangeante et que $\phi(n)$ est à décroissance géométrique. Nous retrouvons ainsi que le théorème de la limite centrale s'applique pour une telle chaîne.

3.4. Conditions de mélange pour les chaînes de Markov stationnaires.

Nous avons les résultats suivants

THÉORÈME 3.5. *Soit $(X_k, k \in \mathbb{Z})$ une chaîne de Markov stationnaire, alors nous avons pour tout $n \geq 1$,*

- (1) $\alpha(n) := \alpha(\sigma(X_0), \sigma(X_n))$,
- (2) $\phi(n) := \phi(\sigma(X_0), \sigma(X_n))$,
- (3) $\psi(n) := \psi(\sigma(X_0), \sigma(X_n))$,
- (4) $\rho(n) := \rho(\sigma(X_0), \sigma(X_n))$,
- (5) $\beta(n) := \beta(\sigma(X_0), \sigma(X_n))$.

En outre, si $\rho(n) \rightarrow 0$, $\phi(n) \rightarrow 0$, $\psi(n) \rightarrow 0$, alors respectivement $\rho(n)$, $\phi(n)$, $\psi(n)$ tendent vers 0 exponentiellement vite.

Il est possible de construire une chaîne de Markov stationnaire d'espace d'états dénombrable dont la décroissance de $\alpha(n)$ n'est pas géométrique. Un tel processus est α -mélangeant, mais pas ρ -mélangeant (car $4\alpha(n) \leq \rho(n)$).

Rosenblatt [47] compare certaines conditions de mélange pour une chaîne de Markov aux conditions sur les normes de l'opérateur markovien associé. En particulier il démontre les lemmes suivants :

LEMME 3.6. *Une chaîne de Markov d'opérateur associé P et de probabilité invariante π est fortement mélangeante si et seulement si*

$$\lim_{n \rightarrow \infty} \sup_{f \perp 1} \frac{\|P^n f\|_1}{\|f\|_\infty} = 0.$$

Comme pour tout $1 \leq p \leq \infty$, nous avons les inégalités $\|f\|_1 \leq \|f\|_p \leq \|f\|_\infty$, nous obtenons le lemme suivant :

LEMME 3.7. *Soit une chaîne de Markov d'opérateur associé P et de probabilité invariante π . Si*

$$\lim_{n \rightarrow \infty} \sup_{f \perp 1} \frac{\|P^n f\|_p}{\|f\|_p} = 0,$$

pour au moins un p , $1 \leq p \leq \infty$, alors la chaîne est fortement mélangeante.

LEMME 3.8. *Une chaîne de Markov stationnaire $(X_k, k \in \mathbf{Z})$ est ρ -mélangeante si et seulement si elle satisfait à une L^2 -condition.*

Nous retrouvons bien le fait que la propriété de ρ -mélange implique la propriété d'être fortement mélangeant. La réciproque n'est pas vraie en général; d'après le théorème 2.7 il faut pour cela que la chaîne de Markov soit fortement mélangeante et ne vérifie aucune L^p condition, $1 < p \leq \infty$. Elle peut vérifier la condition en norme L^1 mais une telle chaîne ne pourra vérifier la condition (D_0) . Nous en déduisons donc que la condition (D_0) implique la propriété de ρ -mélange; en fait nous avons une équivalence plus forte [47][pp 212-213] :

LEMME 3.9. *Une chaîne de Markov stationnaire satisfait à la condition (D_0) si et seulement si elle est ϕ -mélangeante.*

Le processus gaussien de l'exemple 1 est une chaîne de Markov ρ -mélangeante mais non ϕ -mélangeante.

Le cas d'une chaîne de Markov Harris récurrente est traité par le théorème suivant :

THÉORÈME 3.10. *Toute chaîne de Markov sur \mathbf{R} stationnaire aperiodique et Harris récurrente est absolument régulière. Plus généralement, pour toute chaîne de Harris récurrente réelle et stationnaire, $\lim_{n \rightarrow \infty} \beta(n) = 1 - 1/p$, où p est la période.*

3.5. Inégalités exponentielles. Les conditions de α -mélange et de ϕ -mélange permettent, sous certaines conditions, d'obtenir une inégalité exponentielle pour la probabilité de déviation de la somme de variables aléatoires dépendantes. Nous n'explicitons ici que certains des résultats présentés dans [18][§1.4.2]. En particulier le résultat suivant obtenu par Collomb [12] :

PROPOSITION 3.11. *Soit (X_n) une suite de v.a.r. centrées et ϕ -mélangeante telle que $|X_n| \leq 1$ et $\sum_{n=1}^{\infty} \phi(n) < \infty$. Posons*

$$b = 8 \left(1 + 4 \sum_{n=1}^{\infty} \phi(n) \right), \quad a = 2 \exp(3\sqrt{e}), \quad \sigma^2 = \sup_n E|X_n|^2.$$

Alors si $n\sigma^2 \geq 1$ et $0 \leq x \leq b\sigma^2/(8k_n)$, où $k_n = \inf\{k; \phi(k)/k \leq n^{-1}\}$,

$$P\left(\left|n^{-1} \sum_{i=1}^n X_i\right| \geq x\right) \leq a \exp\left\{-\frac{nx^2}{b\sigma^2}\right\}.$$

Appliquons ce résultat au cas d'une chaîne de Markov vérifiant la condition de Doeblin (D_0) . Elle est alors ϕ -mélangeante et $\phi(n) \leq 2\eta^n$ où $0 < \eta < 1$. Soit f une fonction réelle, bornée par 1 et de moyenne nulle pour la probabilité invariante π de la chaîne. La suite de v.a.r. $Y_n := f(X_n)$ satisfait aux hypothèses de la proposition précédente. Nous avons

$$8 \leq b \leq \frac{8(1+7\eta)}{1-\eta}, \quad \sigma^2 = \|f\|_{\pi}^2.$$

Nous obtenons ainsi l'inégalité suivante, pour tout $0 \leq x \leq \|f\|^2/(8bk_n)$ et $n\|f\|^2 \geq 1$:

$$P\left(\left|n^{-1} \sum_{i=1}^n X_i\right| \geq x\right) \leq a \exp\left\{-\frac{nx^2(1-\eta)}{8\|f\|^2(1+7\eta)}\right\}.$$

Nous verrons au chapitre 6 que nous pouvons dans ce cas, obtenir par la technique des perturbations une meilleure borne (voir aussi [18][th. 5 §1.4.2]).

Le cas d'une suite fortement mélangeante est traité en particulier dans [8]. Posons $I = P(|\sum_{i=1}^n X_i| \geq x\sqrt{n})$. Pour tout $0 < a < 1$, il existe un $b > 0$ tel que pour tout n suffisamment grand :

- si $\alpha(n) \leq v e^n$, pour $0 \leq v < 1$, alors $I \leq 2 \exp\{-bn^{1/2-a}x\}$,
- si $\alpha(n) \leq v^n$, pour $0 \leq v < 1$, alors $I \leq 2 \exp\{-bn^{-a/2}x\}$,
- si $\alpha(n) \leq n^{-v}$, pour $v > 0$, alors $I \leq 2 \exp\{-bx(\log n)^{1-a}/\sqrt{n}\}$.

4. Compléments

Les définitions de l'irréductibilité d'une chaîne de Markov, de la notion de chaîne Harris récurrente et les résultats correspondant proviennent de [41]. L'étude de la condition de Doeblin est développée dans [17], où la partition de l'espace d'états en sous-classes ergodiques et en un ensemble transient est démontrée.

Les liens entre la notion d'opérateur quasi-compact introduite en 1937 par N. Kryloff & N. Bogoliouboff, l'ergodicité uniforme et la condition de Doeblin sont développés dans l'article [52]. Notons cependant que la définition de la condition de Doeblin utilisée dans cet article est légèrement différente de celle donnée dans [17]. La démonstration de l'ergodicité uniforme pour un noyau quasi-compact figure également dans [44].

Le caractère discret du spectre d'un opérateur quasi-compact borné est prouvé dans [19]. En particulier, si T est un noyau markovien quasi-compact, alors 1 est une valeur propre isolée. Cette propriété des opérateurs quasi-compacts s'obtient également en considérant la notion de spectre essentiel d'un opérateur $T \in \mathcal{B}(X)$, algèbre des opérateurs linéaires bornés définis sur un espace de Banach X , [34, 20]. Le spectre essentiel d'un opérateur $T \in \mathcal{B}(X)$ est défini comme le spectre de l'image de T dans l'algèbre de Calkin, quotient de $\mathcal{B}(X)$ par l'idéal des opérateurs compacts muni de la norme quotient. Alors, un point du spectre de T n'appartenant pas au spectre essentiel est une valeur propre isolée de multiplicité algébrique finie [20]. Dans le cas d'un opérateur quasi-compact, il existe n_0 tel que T^{n_0} admette un rayon spectral essentiel strictement inférieur à 1. Une démarche réciproque conduirait à considérer des opérateurs bornés T dont le rayon spectral essentiel de T^m , pour un certain $m \geq 1$, est strictement inférieur à 1. Dans ce cas, $I - T$ est un opérateur de Fredholm d'indice nul [20][Th I-4.4] et si 1 est une valeur propre de T alors 1 sera une valeur propre isolée de multiplicité algébrique finie. Montrons que T est alors quasi-compact. Pour cela, il faut faire appel aux notions de mesure de non-compacité introduites dans [20], en particulier :

$$\tilde{\beta}(T) = \inf\{\delta > 0 : T(B) \text{ admet un recouvrement fini par des boules de rayon } \delta\},$$

où B est la boule fermée (ou ouverte) unité, et

$$c(T) = \inf\{\epsilon > 0 : \text{il existe un s.e.v. } M \text{ tel que} \\ \text{codim}M < \infty \text{ et tel que } \|Tx\| \leq \epsilon\|x\| \text{ pour tout } x \in M\}.$$

Notons qu'il est possible de ne considérer que des sous-espaces M fermés. Il est prouvé dans [20][§ I-2,4] que $\frac{1}{2}\tilde{\beta}(T) \leq c(T) \leq 2\tilde{\beta}(T)$, que $c(T)$ est une semi-norme et que le rayon spectral essentiel est donné par

$$r_e(T) := \lim_{n \rightarrow \infty} [c(T^n)]^{1/n} = \lim_{n \rightarrow \infty} [\tilde{\beta}(T^n)]^{1/n}.$$

Ainsi, si le rayon spectral essentiel de T est strictement inférieur à 1, il existe alors un m tel que $c(T^m) = \alpha < (1/2)$, d'où un sous-espace vectoriel fermé M de codimension finie tel que $\|T^m x\| < (c(T) + \epsilon)\|x\|$ pour tout $x \in M$ et $\epsilon > 0$. M étant de codimension finie, il existe un sous-espace vectoriel L de dimension

finie tel que $X = L \oplus M$. Soit P le projecteur continu sur L , qui est un opérateur compact, car de rang fini, alors pour tout $x \in X$, nous avons $\|T^m(I - P)x\| < (1/2 + \epsilon)\|(I - P)x\| < (1 + \epsilon)\|x\|$. Comme $T^m P$ est compact, nous avons prouvé que T est quasi-compact.

Les définitions et les principales propriétés des suites stationnaires mélangeantes sont issues de [18]. Voir également [7] pour une bonne introduction. Le coefficient ρ fut introduit par Hirschfeld (1935) et Gebelein (1941), le coefficient α par Rosenblatt en 1956, le coefficient β par Kolmogorov avant 1959, le coefficient ψ par Ibragimov en 1962 et le coefficient ϕ par Bum, Hanson et Koopmans en 1963. Les liens entre les propriétés mélangeantes et les conditions en norme L^p du noyau markovien d'une chaîne de Markov sont démontrés dans [47], qui contient également de nombreux autres résultats. Les suites stationnaires mélangeantes sont étudiées en détail dans [30], dont la représentation spectrale et le théorème de la limite centrale.

Toutes les définitions des propriétés mélangeantes s'étendent au cas du temps continu. Cependant, les notions d'hypercontractivité et d'ultracontractivité semblent plus intéressantes pour les processus markoviens. Dans [3], un processus $(X_t, t \geq 0)$ défini sur un espace de probabilité (Ω, P) , est dit hypercontractif s'il existe une constante $\lambda > 0$ telle que pour tous $p \geq 1, q \geq 1$ et $t \geq 0$,

$$q - 1 \leq (p - 1)e^{\lambda t} \implies \forall f, \quad \|E(f(X_t)|X_0)\|_q \leq \|f(X_t)\|_p.$$

Dans ce cas, $|E[f(X_t)g(X_0)]| \leq \|g(X_0)\|_{q'} \|f(X_t)\|_p$, avec $1/q' + 1/q = 1$ et $(p - 1)(q' - 1) \geq e^{-\lambda t}$. En particulier, pour un semi-groupe markovien de probabilité invariante π , l'hypercontractivité s'écrit simplement par

$$q - 1 \leq (p - 1)e^{\lambda t} \implies \forall f, \quad \|P_t f\|_{L^q(\pi)} \leq \|f\|_{L^p(\pi)}.$$

Pour plus de détails, se reporter à [3, 28], où l'équivalence entre l'hypercontractivité et les inégalités de Sobolev logarithmiques est décrite, et à [2] pour le lien entre certaines inégalités de Sobolev logarithmiques et l'existence d'un trou spectral.

Dans [28], un semi-groupe P_t de mesure réversible π est dit ultracontractif, si pour tout $t > 0$ et tout $1 \leq p \leq \infty$, P_t est contractif de $L^p(\pi)$ dans $L^p(\pi)$ et si pour tout $t > 0$, $\|P_t\|_{2 \rightarrow \infty} = C_t < \infty$. Le semi-groupe P_t étant autoadjoint, nous en déduisons en passant dans les espaces duaux que $\|P_t\|_{1 \rightarrow 2} \leq C_t$. Ainsi

$$\|P_t\|_{1 \rightarrow \infty} \leq \|P_{t/2}\|_{2 \rightarrow \infty} \|P_{t/2}\|_{1 \rightarrow 2} \leq C_{t/2}^2.$$

Supposons que le semi-groupe admette un noyau continu $p_t(x, y)$ tel que pour tout $f \in L^1(\pi)$

$$P_t f(x) = \int p_t(x, y) f(y) \pi(dy),$$

alors ce noyau est uniformément borné et

$$|p_t(x, y)| \leq C_{t/2}^2, \quad t > 0, \quad \forall x, y.$$

En particulier, ceci implique que l'opérateur P_t est de Hilbert-Schmidt, donc compact dans $L^2(\pi)$, [28, 2].

Inégalités dans le cas général

L'objectif de cette partie est d'étendre, dans un premier temps, aux chaînes de Markov définies sur un espace d'états quelconque, les bornes exponentielles du type Chernoff obtenues dans le cas fini. Nous ne considérerons que des chaînes de Markov ψ -irréductibles, positives récurrentes, i.e. possédant une unique probabilité invariante π . Une première méthode pour obtenir une telle borne consiste à supposer une propriété de ρ -mélange sur le processus. Cette propriété étant équivalente à une L^2 -condition (cf lemme 5-(3.8)), elle se ramène à

$$\beta = \sup \left\{ \frac{\|Pg\|_2}{\|g\|_2}; \pi \cdot g = 0, g \neq 0 \right\} < 1.$$

Notons que cette condition implique que le rayon spectral de P_\perp , restriction de P au sous-espace fermé des fonctions de moyenne nulle est strictement inférieur à 1. Comme P_\perp est borné, son spectre est strictement inclus dans le disque unité. La distance (ou trou spectral) entre la valeur propre 1 de P et le spectre de P_\perp est alors inférieure à $1 - \beta$.

Nagaev [42] semble avoir été le premier à utiliser la méthode des perturbations pour obtenir des résultats asymptotiques dans le cadre d'une chaîne de Markov satisfaisant à la condition de Doeblin (D_0). En particulier il obtient un théorème central limite sous la condition que $f \in L^2(\pi)$ [42][th. 2.2]. Ce résultat améliore celui du théorème 5-(1.10). Nagaev obtient également, sous certaines hypothèses, par la méthode des perturbations un développement asymptotique au premier ordre de la différence $P(n^{-1/2}\sigma^{-1}S_n \leq x) - \Phi(x)$, $\Phi(x)$ étant la fonction de répartition de la loi normale [43][th. 2].

Lorsqu'une fonction mesurable f vérifie les conditions suivantes :

$$\int f(x)\pi(dx) = 0, \quad \sup_x |f(x)| \leq 1, \quad \int f^2(x)\pi(dx) \leq b^2,$$

il est alors possible d'obtenir la majoration suivante [38] :

$$P_\pi \left(n^{-1} \sum_{i=1}^n f(X_i) \geq \gamma \right) \leq \exp \left\{ -\frac{n(1-\beta)\gamma^2}{8b^2} \left(1 - \frac{\gamma}{2b^2} \right) \right\}, \quad \gamma \leq 2b^2.$$

Cette borne s'étend également aux processus markoviens, en supposant qu'il existe $\lambda > 0$ tel que

$$\|P_t g\|_2 \leq e^{-\lambda t} \|g\|_2, \quad \forall t \geq 0, \pi g = 0.$$

La deuxième méthode consistera à étendre la méthode des perturbations utilisée dans le cas d'un espace fini, au cas général. Pour appliquer les résultats décrits dans la partie précédente, il suffira de supposer d'une part que 1 est une valeur propre isolée du spectre du noyau markovien et d'autre part que le projecteur propre associé est de rang fini (resp. 0 est une valeur propre isolée du générateur infinitésimal, dont le projecteur propre associé est de rang fini). Or nous avons vu (cf lemme 5-(2.5)) qu'il suffit pour cela que le noyau markovien de la chaîne soit quasi-compact, condition elle-même équivalente à ce que la chaîne satisfasse à la

condition de Doeblin. Rappelons que la condition (D_0) est équivalente à la propriété de ϕ -mélange (cf lemme 5-(3.9)) condition plus forte que celle de ρ -mélange.

1. Chaînes de Markov

Soit un noyau markovien P défini sur un espace \mathcal{X} , dont la valeur propre 1 est isolée. Notons par $\epsilon(P)$ la distance de 1 au reste du spectre de P . Supposons également qu'il existe une unique probabilité invariante π pour le noyau markovien et que ce noyau agit sur $L^2(\pi)$. P est donc un opérateur borné défini sur l'espace de Hilbert $L^2(\pi)$ et l'existence d'une unique probabilité invariante implique que 1 est une valeur propre de multiplicité 1 (relation entre le spectre de P et de son adjoint) donc que le projecteur propre noté E_π est de rang fini 1.

L'espace $H = L^2(\pi)$ admet alors la décomposition $H = M' \oplus M''$, où $M' = E_\pi H$, et les restrictions $P_{M'}$ et $P_{M''}$ de P à M' et M'' admettent pour spectre respectif $\{1\}$ et $\Sigma(P) \setminus \{1\}$. Le spectre de $P_{M'}$ est constitué d'un seul point, donc $P_{M'} - I$ est quasi-nilpotent (i.e. son rayon spectral est nul). La résolvante $R(\zeta)$ de P peut s'écrire comme la somme

$$\begin{aligned} R(\zeta) &= R'(\zeta) + R''(\zeta) \\ R'(\zeta) &= R(\zeta)E_\pi, \quad R''(\zeta) = R(\zeta)(1 - E_\pi), \end{aligned}$$

où $R''(\zeta)$ s'appelle la résolvante réduite de P pour la valeur propre 1. En particulier $S \equiv R''(1)$ est l'inverse de la restriction de $P - 1$ à M'' , i.e.

$$(P - 1)S = S(P - 1) = 1 - E_\pi, \quad SE_\pi = E_\pi S = 0.$$

Soit f une fonction réelle mesurable définie sur \mathcal{X}, \mathcal{F} telle que

$$\pi f = 0, \quad \|f\|_\infty \leq 1, \quad \|f\|_2^2 \leq b^2 \leq 1.$$

Soit q la mesure initiale du processus; nous la supposons absolument continue par rapport à π et nous noterons φ sa densité. Nous voulons donc borner la vitesse de convergence de la probabilité

$$P_q \left[n^{-1} \sum_{i=1}^n f(X_i) - \pi f \geq \gamma \right].$$

De façon identique au cas d'un espace d'états fini, nous obtenons grâce à l'inégalité de Chebyshev la majoration suivante pour $r > 0$:

$$\begin{aligned} P_q \left[n^{-1} \sum_{i=1}^n f(X_i) \geq \gamma \right] &\leq e^{-rn\gamma} E_q \exp \left(r \sum_{i=1}^n f(X_i) \right) \\ &= e^{-rn\gamma} \int \varphi(x_0) \pi(dx_0) \int P(x_0, dx_1) e^{rf(x_1)} \dots \\ &\dots \int P(x_{n-1}, dx_n) e^{rf(x_n)} \\ &= e^{-rn\gamma} \int \varphi(x) P_f^n(r) 1(x) \pi(dx), \end{aligned}$$

où $P_f(r)g(x) := \int e^{rf(y)} g(y) P(x, dy)$, et $1(x) = 1$ pour tout $x \in X$. Supposons que $\varphi \in L^2(\pi)$, nous pouvons donc appliquer l'inégalité de Cauchy-Schwarz, ce qui conduit à

$$P_q \left[n^{-1} \sum_{i=1}^n f(X_i) \geq \gamma \right] \leq \|\varphi\|_2 e^{-rn\gamma} \|P_f^n(r) 1\|_2.$$

Ceci ne s'applique pas lorsque la mesure q est concentrée en un point x , i.e. lorsque le processus démarre du point x . Dans ce cas nous avons la majoration

suivante

$$P_x \left[n^{-1} \sum_{i=1}^n f(X_i) \geq \gamma \right] \leq e^{-rn\gamma} P_f^n(r) 1(x).$$

Notons que la distribution marginale de la chaîne après une étape est $P(x, \cdot)$. L'hypothèse supplémentaire sur les fonctions de transition

$$P(x, \cdot) \prec \pi, \quad \text{pour tout } x \in \mathcal{X},$$

résout le problème. Cela sera en particulier le cas si $P(x, dy)$ admet une densité $p(x, y)$ par rapport à la mesure π , $\mathcal{F} \otimes \mathcal{F}$ mesurables et vérifiant la condition de Hilbert-Schmidt

$$\int p(x, y)^2 \pi(dx) \pi(dy) < \infty.$$

L'opérateur $P_f(r)$ agit sur $L^2(\pi)$, il est borné par $\|P_f(r)\|_{2 \rightarrow 2} \leq \exp r \|f\|_\infty$. Désignons par D l'opérateur multiplicatif défini par $Dg(x) = f(x)g(x)$. Cet opérateur est borné dans $L^2(\pi)$ de norme $\|f\|_\infty$. Nous pouvons donc exprimer $P_f(r)$ sous la forme suivante :

$$P_f(r) = P + \sum_{i=1}^{\infty} \frac{r^i}{i!} P D^i,$$

où $\|P D^n\|_{2 \rightarrow 2} \leq n!(1/2)^{n-1}$. Nous obtenons ainsi une famille d'opérateurs holomorphes de type (A). Il est possible d'utiliser directement les expressions obtenues en dimension finie, pour estimer les coefficients du développement en série de la valeur propre perturbée $\beta_0(r)$, issue de 1. Nous pouvons prendre

$$r_0 = \left(\frac{2}{\epsilon(P)} + \frac{1}{2} \right)^{-1} \geq \frac{\epsilon(P)}{3}.$$

QUESTION : Peut-on assouplir la condition sur f ? Par exemple pour que $P_f(r)g \in L^2(\pi)$, dès que $g \in L^2(\pi)$, il suffit que $\sup_x \int P(x, dy) \exp(rf(y)) < \infty$, pour r petit. En particulier, l'emploi de la formule de Stirling conduit aux conditions suffisantes pour $r < 1/(2e\rho)$:

$$\sup_x \left| \int P(x, dy) f^n(y) \right| = M_n < \infty$$

$$\limsup_n \frac{1}{n} M_n^{1/n} = \rho$$

1.1. Chaînes réversibles. Si la mesure π est réversible alors l'opérateur P est autoadjoint dans $L^2(\pi)$. L'opérateur $\tilde{P}_f(r)$ défini par

$$\tilde{P}_f(r)g(x) = \int e^{\frac{r}{2}f(x)} e^{\frac{r}{2}f(y)} g(y) P(x, dy) = e^{\frac{r}{2}f(x)} P_f(r/2)g(x),$$

est également autoadjoint dans $L^2(\pi)$ et vérifie la relation

$$P_f^n(r) 1(x) = e^{-\frac{r}{2}f(x)} \tilde{P}_f^n(r) e^{\frac{r}{2}f(x)}.$$

Introduisons l'opérateur multiplicatif E_r défini par $E_r g(x) = g(x) \exp(rf(x))$; nous pouvons alors écrire l'identité précédente sous la forme

$$P_f^n(r) = E_{-r/2} \tilde{P}_f^n(r) E_{r/2}.$$

Nous en déduisons que le rayon spectral de $P_f(r)$ est identique au rayon spectral de $\tilde{P}_f(r)$. Nous obtenons la majoration suivante

$$P_q \left[n^{-1} \sum_{i=1}^n f(X_i) \geq \gamma \right] \leq \|\varphi\|_2 e^{r\|f\|_\infty} e^{-rn\gamma} \|\tilde{P}_f^n(r)\|_2$$

$$= \|\varphi\|_2 e^{r\|f\|_\infty} e^{-rn\gamma} \beta_0^n(r)$$

où $\beta_0(r)$ est le rayon spectral de $\tilde{P}_f(r)$. Or $\tilde{P}_f(r)$ étant autoadjoint possède un spectre réel; donc pour $r > 0$ suffisamment petit, le rayon spectral $\beta_0(r)$ sera identique à la valeur propre perturbée issue de 1 (cf th. 4-(2.2)). Or pour tout r , $\beta_0(r)$ est aussi une valeur propre de $P_f(r)$ telle que $\beta_0(0) = 1$, c'est donc la valeur propre 1 perturbée. Nous sommes donc ramenés à majorer $\beta_0(r)$, il suffit pour cela d'utiliser les estimations obtenues dans le cas fini. Nous obtenons alors le théorème suivant

THÉORÈME 1.1. *Soit $P(x, dy)$ un noyau markovien défini sur un espace probabilisé $(\mathcal{X}, \mathcal{F}, \pi)$ où π est l'unique probabilité invariante de P . Soient $f : \mathcal{X} \rightarrow \mathbb{R}$ une fonction bornée telle que $\pi f = 0$, $\|f\|_\infty \leq 1$ et $0 < \|f\|^2 \leq b^2$ et q la distribution initiale du processus markovien ayant pour densité relativement à π , une fonction $\varphi \in L^2(\pi)$. Supposons que l'opérateur linéaire P , défini par le noyau de transition du processus, vérifie les hypothèses suivantes :*

- (1) P se prolonge à $L^2(\pi)$ et est autoadjoint,
- (2) 1 est un point isolé du spectre de P ; $\Sigma(P) = \{1\} + \Sigma'(P)$.

alors pour tout $n > 0$ et tout $0 < \gamma \leq 1$,

$$(1.1) \quad P_q \left[n^{-1} \sum_{i=1}^n f(X_i) \geq \gamma \right] \leq e^{\epsilon(P)/5} \|\varphi\|_2 \exp \left[-\frac{n\gamma^2 \epsilon(P)}{4b^2(1 + h(5\gamma/b^2))} \right],$$

où $\epsilon(P) = \text{dist}(1, \Sigma'(P)) > 0$ et

$$h(x) = \frac{1}{2}(\sqrt{1+x} - (1-x/2)).$$

REMARQUE 1.2. Nous avons supposé que π était une probabilité invariante, ceci pour que les fonctions constantes appartiennent à $L^2(\pi)$.

Le théorème n'est pas valable si la mesure initiale q est concentrée en un point x , à moins de supposer que pour tout $x \in \mathcal{X}$ la mesure $P(x, \cdot) \prec \pi$.

EXEMPLE 3. Le processus gaussien de l'exemple (5-(1)) vérifie une condition L^2 , par conséquent la chaîne de Markov correspondante satisfait aux hypothèses du théorème 1.1. La chaîne étant réversible pour la mesure gaussienne standard γ , le trou spectral est égal à $|\eta|$. D'autre part pour tout $x \in \mathbb{R}$, $P(x, \cdot) \prec \gamma$; cette mesure admet pour densité par rapport à la mesure γ la fonction

$$\varphi_x(y) = e^{x^2/2} \frac{1}{\sqrt{1-\eta^2}} \exp \left(-\frac{\eta^2}{2(1-\eta^2)} \left(y - \frac{x}{\eta} \right)^2 \right),$$

dont la norme L^2 est

$$\|\varphi_x\|_2 = (1-\eta^4)^{-1/4} e^{x^2(1+\eta^2)/2}.$$

Nous obtenons ainsi la majoration suivante :

$$P_x \left[n^{-1} \sum_{i=1}^n f(X_i) \geq \gamma \right] \leq \frac{e^{|\eta|/5}}{(1-\eta^4)^{1/4}} e^{x^2(1+\eta^2)/2} \exp \left[-\frac{n\gamma^2 |\eta|}{4b^2(1 + h(5\gamma/b^2))} \right].$$

EXEMPLE 4. Supposons que la chaîne de Markov vérifie la condition de Doeblin (D_0) c'est-à-dire qu'il existe un entier k et un $\delta < 1$ tels que

$$\sup_{x,y} \|P^{(k)}(x, \cdot) - P^{(k)}(y, \cdot)\|_{VT} = \delta.$$

Cette chaîne admet une unique probabilité invariante π et

$$\|P^{(n)}(x, \cdot) - \pi\|_{VT} \leq \kappa \rho^n,$$

avec $\kappa = \delta^{-1}$ et $\rho = \delta^{1/k}$. L'opérateur linéaire P présente un trou spectral $\epsilon(P) \geq 1 - \rho$, donc le théorème 1.1 s'applique dans le cas réversible .

Pour les mêmes raisons, la borne du type Berry–Esséen s'étend au cas des chaînes de Markov réversibles vérifiant les hypothèses du théorème 1.1. Le résultat obtenu par B. Mann peut donc être élargi au cas d'un espace non dénombrable.

REMARQUE 1.3. Dans le cas d'une matrice positive irréductible, le théorème de Perron-Frobenius garantit l'existence d'un vecteur propre strictement positif associé à la valeur propre maximale. En est-il de même pour un opérateur positif général ? Le théorème III.10.1 dans [29][Thm. III.10.1] apporte une réponse positive sous certaines conditions. En particulier, considérons un opérateur P à noyau admettant une densité $p(x, y)$ par rapport à une certaine probabilité $\mu(dx)$. S'il existe un entier n_0 , et $0 < a < b < \infty$ tels que $a \leq p^{(n_0)}(x, y) \leq b$, alors P admet une valeur propre positive, maximale en module, et une fonction propre associée $u(x)$ uniformément positive ; i.e. il existe $c > 0$ telle que pour tout x , $u(x) \geq c$. Dans ce cas, la preuve de la borne inférieure, établie au chapitre 2, s'applique.

1.2. Chaîne non réversible. Dans cette section, nous allons étendre le résultat précédent au cas des chaînes de Markov non nécessairement réversibles. Par analogie avec la dimension finie nous aurons besoin d'étendre le théorème de Marcus au cas des opérateurs linéaires bornés.

LEMME 1.4. *Soit $T \in \mathcal{B}(H)$, alors pour tout $n > 0$*

$$\|(T^*)^n T^n\| \leq \|(T^* T)\|^n.$$

DÉMONSTRATION. Soient U et V deux éléments de $\mathcal{B}(H)$, comme $(UV)^n = U(VU)^{n-1}V$, le rayon spectral $\rho(UV) = \rho(VU)$. Soit $T \in \mathcal{B}(H)$, prenons $U = T^*$ et $V = T$ alors $U^n V^n$ est autoadjoint donc

$$\begin{aligned} \|U^n V^n\| &= \rho(U^n V^n) \\ &= \rho((VU)U^{n-1}V^{n-1}) \\ &\leq \|(VU)U^{n-1}V^{n-1}\| \\ &\leq \|VU\| \|U^{n-1}V^{n-1}\| \\ &\leq \|TT^*\|^n. \end{aligned}$$

□

Ce lemme nous permet d'étendre directement le résultat obtenu en dimension finie au cas d'un espace quelconque sous certaines hypothèses ; d'où le théorème suivant :

THÉORÈME 1.5. *Soit $P(x, dy)$ un noyau markovien défini sur un espace probabilisé $(\mathcal{X}, \mathcal{F}, \pi)$ où π est l'unique probabilité invariante de P . Soient $f : \mathcal{X} \rightarrow \mathbb{R}$ une fonction bornée telle que $\pi f = 0$, $\|f\|_\infty \leq 1$ et $0 < \|f\|^2 \leq b^2$ et q la distribution initiale du processus markovien ayant pour densité relativement à π , une fonction $\varphi \in L^2(\pi)$. Supposons que l'opérateur linéaire P , défini par le noyau de transition du processus, vérifie les hypothèses suivantes :*

- (1) P se prolonge à $L^2(\pi)$
- (2) 1 est un point isolé du spectre de $K := P^* P$;
- (3) La dimension du projecteur propre de K associé à la valeur propre 1 est finie et égale à 1,

alors pour tout $n > 0$ et tout $0 < \gamma \leq 1$,

$$(1.2) \quad P_q \left[n^{-1} \sum_{i=1}^n f(X_i) \geq \gamma \right] \leq \|\varphi\|_2 \exp \left[-\frac{n\gamma^2 \epsilon(K)}{8b^2(1 + h(5\gamma/b^2))} \right],$$

où $\epsilon(K)$ désigne le trou spectral de K et

$$h(x) = \frac{1}{2}(\sqrt{1+x} - (1-x/2)).$$

2. Processus markoviens

2.1. Rappels. Etant donné un espace probabilisé $(\mathcal{X}, \mathcal{F}, \pi)$, nous considérons un semi-groupe d'opérateurs P_t agissant sur l'espace $B(\mathcal{X}, \mathcal{F})$ des fonctions mesurables bornées à l'aide d'un noyau de transition $p_t(x, dy)$:

$$P_t f(x) = \int_{\mathcal{X}} f(y) p_t(x, dy).$$

Nous supposons toujours que les mesures $p_t(x, dy)$ sont positives et bornées et que les opérateurs P_t se prolongent en opérateurs bornés sur $L^2(\pi)$. Nous supposons de plus les propriétés suivantes :

- (1) (Propriété de semi-groupe) : $P_t P_s = P_{t+s}$; et $P_0 = I$.
- (2) (Continuité forte en 0 dans $L^2(\pi)$) : Pour tout $f \in L^2(\pi)$, $P_t f \rightarrow f$ dans $L^2(\pi)$ lorsque $t \rightarrow 0$.
- (3) $P_t 1 = 1$.

Les semi-groupes markoviens sont associés aux processus de Markov (X_t) d'espace d'états l'espace \mathcal{X} par la formule

$$E_x f(X_t) = P_t f(x).$$

Nous noterons D_2 le domaine dans $L^2(\pi)$ du générateur infinitésimal $-\Lambda$ du semi-groupe P_t : D_2 est l'espace des fonctions f de $L^2(\pi)$ pour lesquelles la limite dans $L^2(\pi)$

$$\Lambda f = \lim_{t \rightarrow 0} \frac{1}{t} (f - P_t f),$$

existe. L'espace D_2 est dense dans $L^2(\pi)$ et est stable par action du semi-groupe P_t . Réciproquement, l'opérateur $-\Lambda$ et son domaine D_2 déterminent entièrement le semi-groupe borné P_t (cf section 4-(3)) satisfaisant à l'équation

$$\partial_t P_t f = -P_t \Lambda f = -\Lambda P_t f.$$

Pour tout $f \in D_2(\Lambda)$, le processus

$$M_t = f(X_t) - f(X_0) + \int_0^t \Lambda f(X_s) ds$$

est une martingale pour la filtration engendrée par le processus X_t . Elle admet alors une version càdlàg, par conséquent le processus $f(X_t)$ en tant que semi-martingale admet une version càdlàg. Il est alors possible de considérer l'intégrale de Riemann $\int_0^t f(X_s) ds$.

2.2. Inégalités pour les processus markoviens. La preuve du théorème 2-(1.1) dans la section 2-(1) est basée sur la propriété suivante valable pour toutes matrices A, B

$$\lim_{k \rightarrow \infty} \left(\exp(A/k) \exp(B/k) \right)^k = \exp(A + B).$$

Pour adapter la démonstration de ce théorème dans le cas général, il sera nécessaire d'étendre cette propriété aux générateurs infinitésimaux des semi-groupes markoviens. Bien évidemment, pour cela il faudra que la somme des deux générateurs soit également un générateur ; cette extension est connue sous le nom de formule produit de Trotter.

THÉORÈME 2.1 (Trotter). *Soient $\Lambda \in \mathcal{G}(1, \beta)$, $A \in \mathcal{G}(1, \beta')$ et r un réel positif tels que $D(\Lambda + rA) = D(\Lambda) \cap D(A)$ soit dense dans X . Alors (la fermeture de) $-(\Lambda + rA)$ engendre un semi-groupe fortement continu en zéro si et seulement si $R(\lambda + \Lambda + rA)$ est (dense dans) X pour un $\lambda > \beta + r\beta'$. Si $-(\Lambda + rA)$ (ou sa fermeture) engendre un semi-groupe fortement continu en zéro, celui-ci est donné par*

$$S_{r,t} = \lim_{h \rightarrow 0} \left(\exp(-\Lambda h) \exp(-rAh) \right)^{\lceil t/h \rceil}.$$

La preuve de ce théorème figure dans l'article original de Trotter [50], ou dans [21]. En particulier, d'après le théorème 4-(3.1) si $A \in \mathcal{B}(X)$ alors pour tout $r \geq 0$, $-(\Lambda - rA)$ engendre un semi-groupe qui s'exprime par

$$e^{-t(\Lambda - rA)} = \lim_{n \rightarrow \infty} \left(e^{-(t/n)\Lambda} e^{(t/n)rA} \right)^n.$$

Considérons un processus markovien irréductible ayant pour unique probabilité invariante π . Supposons que le semi-groupe markovien associé agisse sur $L^2(\pi)$, et que son générateur infinitésimal $-\Lambda$ soit tel que le symétrisé additif $(\Lambda + \Lambda^*)/2$ admette 0 comme valeur propre isolée simple. Le projecteur propre associé sera par conséquent de rang 1. Conformément à ce qui a été vu à la fin de la section 4-(3.1), il existe un réel $\lambda_1 > 0$ tel que pour tout $f \in D_2(\Lambda)$ de moyenne nulle pour la mesure π , $\|P_t f\| \leq \exp(-\lambda_1 t)$. λ_1 sera appelé trou spectral du symétrisé additif du générateur.

La formule de Trotter appliquée à $\Lambda \in \mathcal{G}(1, 0)$ et à l'opérateur multiplicatif, par la fonction f , permet de se ramener à la détermination d'une majoration de la norme du semi-groupe $P_t(r)$ engendré par $-\Lambda(r) = -(\lambda - rD)$. Posons $\tilde{\Lambda}(r) = (\Lambda + \Lambda^*)/2 - rD$, alors $\|P_t(r)\| \leq \exp(-\lambda_0(r)t)$, où

$$\lambda_0(r) = \inf \left\{ \langle \tilde{\Lambda}(r)g, g \rangle / \|g\|^2, g \neq 0 \right\}.$$

Mais comme $\tilde{\Lambda}(r)$ est autoadjoint, $\lambda_0(r)$ sera égal à la valeur propre perturbée issue de 0. Nous obtenons donc le théorème suivant

THÉORÈME 2.2. *Soit un processus markovien irréductible et ergodique d'espace d'états \mathcal{X} et de semi-groupe (P_t) . Supposons que son générateur infinitésimal $-\Lambda$ vérifie les conditions suivantes :*

- (1) Λ se prolonge à $L^2(\pi)$, où π désigne l'unique probabilité invariante du processus,
- (2) 0 est un point isolé du spectre de Λ .

Soient $f \in D_2(\Lambda) : \mathcal{X} \rightarrow \mathbb{R}$ telle que $\pi f = 0$, $\sup_x |f(x)| \leq 1$ et $0 < \|f\|^2 \leq b^2$ et q la distribution initiale du processus markovien ayant pour densité relativement à π , une fonction $\varphi \in L^2(\pi)$. Alors pour tout $t > 0$ et tout $0 < \gamma \leq 1$,

$$(2.1) \quad P_q \left[\frac{1}{t} \int_0^t f(X_s) ds \geq \gamma \right] \leq \|\varphi\|_2 \exp \left[-\frac{\gamma^2 \lambda_1 t}{4b^2(1 + h(4\gamma/b^2))} \right],$$

où λ_1 est le trou spectral de $(\Lambda + \Lambda^*)/2$ et

$$h(x) = \frac{1}{2}(\sqrt{1+x} - (1-x/2)).$$

REMARQUE 2.3. Supposons Λ réversible; lorsque la distribution initiale est concentrée en un point x donné, le résultat précédent ne s'applique pas directement. Supposons de plus le semi-groupe P_t ultracontractif et admettant un noyau continu

$p_t(x, y)$ par rapport à la mesure π . Alors, le semi-groupe $P_t(r)$ est également ultra-contractif, car pour $g \in L^2(\pi)$ et $e_h(t) := \exp \int_0^t h(X_s) ds$

$$\begin{aligned} P_t(r)g(x) &\leq E_x[e_{2rf}(t)]^{1/2} E_x[|g(X_t)|]^{1/2} \\ &\leq \sup_x E_x[e_{2rf}(t)]^{1/2} \left[\sup_{x,y} |p_t(x, y)| \right]^{1/2} \|g\|_2 \\ &\leq e^{r\|f\|_\infty t} C_{t/2} \|g\|_2, \end{aligned}$$

la dernière inégalité provenant de la majoration uniforme du noyau de P_t . Ainsi, $\|P_t(r)\|_{2 \rightarrow \infty} \leq C'_t$ et

$$\begin{aligned} \sup_x P_x \left[\frac{1}{t} \int_0^t f(X_s) ds \geq \gamma \right] &\leq e^{-r\gamma} \|P_t(r)\mathbf{1}\|_\infty \\ &\leq e^{-r\gamma} \|P_t(r)\|_{2 \rightarrow \infty} \\ &\leq e^{-r\gamma} \|P_{t_0}(r)\|_{2 \rightarrow \infty} \|P_{t-t_0}(r)\|_{2 \rightarrow 2} \\ &\leq C'_{t_0} \exp \left[-\frac{\gamma^2 \lambda_1(t-t_0)}{4b^2(1+h(2\gamma/b^2))} \right], \end{aligned}$$

où $0 < t_0 < t$.

Par une preuve identique à celle du théorème 2-(3.1), nous obtenons, sous les hypothèses du théorème précédent, une borne de Berry-Esséen identique à celle obtenue dans le cas fini.

EXEMPLE 5 (Processus d'Ornstein-Uhlenbeck sur \mathbb{R}^n). Soit $\gamma_n(dx)$ la mesure Gaussienne dans \mathbb{R}^n ,

$$\gamma_n(dx) = (2\pi)^{-n/2} e^{-|x|^2/2} dx.$$

Considérons l'opérateur différentiel L défini sur l'ensemble $\mathcal{C}_c^\infty(\mathbb{R}^n)$ des fonctions indéfiniment dérivables sur \mathbb{R}^n à support compact, par

$$L := \Delta - x \cdot \nabla,$$

où

$$\Delta := \sum_{i=1}^n (\partial^2 / \partial x_i^2), \quad x \cdot \nabla := \sum_{i=1}^n x_i (\partial / \partial x_i).$$

L'opérateur L est symétrique dans $L^2(\gamma_n)$ d'après la formule de Green, d'où l'existence d'une extension fermée. Le semi-groupe associé à ce générateur est appelé semi-groupe d'Ornstein-Uhlenbeck. Il admet une densité absolument continue par rapport à la mesure de Lebesgue et s'exprime par

$$P_t f(x) = \int_{\mathbb{R}^n} f(e^{-t}x + (1 - e^{-2t})^{1/2}y) \gamma_n(dy).$$

Ce semi-groupe s'étend à $L^2(\gamma_n)$; considérons alors l'extension \dot{L} de L au domaine $D_2(\dot{L})$ des fonctions $f \in L^2(\gamma_n)$ telle que la limite suivante existe

$$\lim_{t \rightarrow 0} \frac{P_t f - f}{t}.$$

Notons que $\mathcal{C}_c^\infty(\mathbb{R}^n)$ est dense dans $D_2(\dot{L})$. D'autre part le semi-groupe est autoadjoint dans $L^2(\gamma_n)$. Une dérivation en zéro de l'identité

$$\langle P_t f, g \rangle = \langle f, P_t g \rangle, \quad f, g \in D_2(\dot{L}),$$

prouve que \dot{L} est symétrique dans $D_2(\dot{L})$. \dot{L} admet ainsi une extension fermée que nous désignerons par \tilde{L} .

Soient $h_k, k \in \mathbb{N}$ les polynômes d'Hermite

$$h_k = (-1)^k e^{x^2/2} \frac{d^k}{dx^k} e^{-x^2/2}.$$

Les polynômes $h_k/\sqrt{k!}$ forment une famille complète orthogonale dans $L^2(\gamma_1)$. Les polynômes

$$H_k(x) = \prod_{i=1}^n h_{k_i}(x_i)/\sqrt{k_i!},$$

où $k = (k_1, \dots, k_n)$, forment une base orthonormée dans $L^2(\gamma_n)$ de fonctions propres de \tilde{L} pour les valeurs propres respectives $k_1.k_2 \dots k_n$. Ainsi \tilde{L} est autoadjoint et admet un trou spectral égal à 1. Le théorème 2.2 s'applique donc avec $\lambda_1 = 1$.

Signalons que le semi-groupe d'Ornstein-Uhlenbeck est hypercontractif (théorème de E. Nelson) : pour $1 < p < q < \infty$ et $t > 0$ vérifiant $e^t \geq [(q-1)/(p-1)]^{1/2}$, alors pour toutes fonctions $f \in L^p(\gamma_n)$,

$$\|P_t f\|_q \leq \|f\|_p,$$

où $\|\cdot\|_p$ désigne la norme dans $L^p(\gamma_n)$.

EXEMPLE 6 (Processus de sauts). Un processus de sauts d'espace d'états \mathcal{X} est défini à partir d'un noyau de transition markovien $\mu(x, dy)$ de probabilité invariante π et du générateur infinitésimal

$$\Lambda f(x) = \int (f(x) - f(y))\mu(x, dy), \quad f \in L^2(\pi).$$

Cet opérateur est borné dans $L^2(\pi)$ (de norme ≤ 2). Le processus de sauts associé peut se construire de la façon suivante : soit $Y = (Y_n)$ une chaîne de Markov à valeurs dans \mathcal{X} de noyau de transition μ et soit $V = (V(t))$ un processus de Poisson de paramètre 1 indépendant de la chaîne Y . Posons $X_t = Y_{V(t)}$, ce processus est alors un processus de Markov de semi-groupe

$$U(t) = \sum_{k=0}^{\infty} e^{-t} \frac{t^k}{k!} P^k,$$

où P est l'opérateur associé au noyau de transition μ , i.e $Pf(x) = \int f(y)\mu(x, dy)$ [21][§4.2]. Dans l'article [37], les auteurs donnent des bornes du spectre de l'opérateur Λ sur les fonctions de moyenne nulle pour la mesure π . La borne inférieure fait intervenir la constante de Cheeger définie par

$$k := \inf_{\substack{A \in \mathcal{F} \\ 0 < \pi(A) < 1}} k(A),$$

où

$$(2.2) \quad k(A) := \frac{\int \pi(dx) I_A(x) \mu(x, A^c)}{\pi(A) \pi(A^c)}$$

$$(2.3) \quad = \frac{-\langle I_A, \Lambda I_{A^c} \rangle_{\pi}}{\pi(A) \pi(A^c)} = \frac{\langle I_A, \Lambda I_A \rangle_{\pi}}{\pi(A) \pi(A^c)}.$$

Supposons dans un premier temps la mesure π réversible, le spectre de Λ est donc réel. Le théorème 2.1 et la proposition 2.2 dans [37] conduisent aux estimations du trou spectral

$$\frac{k^2}{8} \leq \lambda_1 \leq k.$$

L'expression (2.3) implique que la constante de Cheeger du symétrisé additif de Λ est la même que celle de Λ . Par conséquent le théorème 2.2 s'applique lorsque

la constante de Cheeger du processus est non nulle. Notons également que si π n'est pas réversible, le théorème 2.3 dans [37] dit que le spectre de Λ restreint aux fonctions de moyenne nulle est inclus dans l'ensemble

$$\{\lambda : |\lambda| \leq 2 \text{ et } \operatorname{Re} \lambda \geq k^2/8\}.$$

EXEMPLE 7 (Mouvement brownien sur une variété riemannienne compacte).

Soit M une variété riemannienne compacte, connexe et de dimension n . Désignons par g_{ij} le tenseur métrique covariant, par g le déterminant de la matrice (g_{ij}) et par $\mu(dx)$ la mesure riemannienne définie par

$$\mu(dx) = \sqrt{g} dx^1 \cdots dx^n.$$

Considérons le semi-groupe P_t de générateur infinitésimal l'opérateur de Laplace-Beltrami

$$\Delta f := \frac{1}{\sqrt{g}} \frac{\partial}{\partial x^i} \left(\sqrt{g} g^{ik} \frac{\partial f}{\partial x^k} \right), \quad f \in C^\infty(M),$$

écrit avec la convention de sommation d'Einstein.

La formule de Green [25][prop. 4.11] montre que le générateur est autoadjoint dans $L^2(\mu)$ et admet un spectre discret

$$\lambda_0 = 0 < \lambda_1 \leq \lambda_2 \leq \cdots \lambda_n \leq \cdots,$$

tel que chaque espace propre E_{λ_i} est de rang fini [25][th. 4.43]. Ainsi le théorème 2.2 s'applique. Par exemple, dans le cas de la sphère S^N dans \mathbb{R}^{N+1} les valeurs propres du Laplacien sont $k(k+N-1)$, avec $k \in \mathbb{N}$, et donc le trou spectral a pour valeur N [25][cor. 4.49].

3. Compléments

Au chapitre 2, nous avons étudié le développement d'Edgeworth de la probabilité de déviation. Bien entendu, le résultat obtenu s'étend au cas d'une chaîne de Markov à espace d'états quelconque ; en particulier, le cas d'une chaîne de Markov satisfaisant à la condition de Doeblin (D_0) est traité dans [42]. Un autre type de développement asymptotique peut être obtenu dans le cas des chaînes de Markov ; il s'agit de l'approximation du point de selle. Exposons les principes de cette méthode présentée dans [33].

Soit P un noyau markovien satisfaisant à la condition suivante : il existe une probabilité $\nu(dx)$ sur l'espace d'états $(\mathcal{X}, \mathcal{F})$ et des constantes $0 < a < b$, telles que

$$a\nu(A) \leq P(x; A) \leq b\nu(A), \quad x \in \mathcal{X}, \quad A \in \mathcal{F}.$$

Soit $p(x, dy) = \frac{dP(x, \cdot)}{d\nu}(y)$, alors nous avons $a \leq p(x, y) \leq b$. Soit f une fonction mesurable de \mathcal{X} dans \mathbb{R} , posons

$$P_\theta(x, dy) = \exp\{\theta f(y)\} P(x, dy).$$

D'après la remarque 1.3, l'opérateur P_θ , associé au noyau introduit ci-dessus, admet une valeur propre maximale $\lambda(\theta)$ et une fonction propre $u(\cdot; \theta)$ uniformément bornée et positive. Il existe également une mesure propre $L(dx; \theta)$, i.e. $\int P_\theta(x; A) L(dx; \theta) = \lambda(\theta) L(A; \theta)$, telle que sa densité par rapport à la mesure ν soit uniformément positive. Normalisons $u(\cdot; \theta)$ et $L(\cdot; \theta)$ de telle façon que

$$\pi_\theta(A) := \int_A u(x; \theta) L(dx; \theta)$$

soit une probabilité sur \mathcal{X} . Suivant [33][chap. 9], nous pouvons construire un nouveau noyau markovien dépendant de θ de la façon suivante :

$$Q_\theta(x, dy) = \lambda(\theta)^{-1} \frac{u(y; \theta)}{u(x; \theta)} P_\theta(x, dy).$$

Ce noyau admet pour probabilité invariante la probabilité π_θ définie ci-dessus ; en outre $\pi_\theta f := \int f(x)\pi_\theta(dx) = \kappa'(\theta)$, où $\kappa(\theta) := \log \lambda(\theta)$. Soient $S_n = \sum_{i=1}^n f(X_i)$ et q la mesure initiale de la chaîne de Markov (X_i) de noyau P , nous obtenons par changement de mesure, l'expression suivante :

$$P_q(n^{-1}S_n > x) = \exp\{-n(\theta x - \kappa(\theta))\} E_{\theta,q} \left\{ \frac{u(X_0; \theta)}{u(X_n; \theta)} \exp[-\theta(S_n - nx)] \mathbf{1}_{\{S_n \geq nx\}} \right\},$$

où $E_{\theta,q}$ désigne l'espérance sous la mesure engendrée par le noyau Q_θ et la mesure initiale q .

Etant donnée la probabilité

$$Q_{n,\theta}(C) = \int \mathbf{1}_C(x_0, \dots, x_n) q(dx_0) \prod_{i=1}^n Q_\theta(x_{i-1}, dx_i),$$

introduisons une nouvelle probabilité $P_{n,\theta}$ de densité par rapport à $Q_{n,\theta}$,

$$\frac{dP_{n,\theta}}{dQ_{n,\theta}} = \psi_n(\theta)^{-1} \frac{u(x_0; \theta)}{u(x_n; \theta)},$$

où $\psi_n(\theta)$ est une constante de normalisation. Notons que

$$\frac{dP_{n,\theta}}{dP_q} = (\lambda^n(\theta)\psi_n(\theta))^{-1} \exp\{\theta S_n\}.$$

Désignant par $E_{n,\theta}$ l'espérance sous cette nouvelle probabilité, nous avons

$$P_q(n^{-1}S_n > x) = \exp\{-n(\theta x - \kappa(\theta))\} \psi_n(\theta) E_{n,\theta} \left[\exp\{-\theta(S_n - nx)\} \mathbf{1}_{\{S_n \geq nx\}} \right].$$

Supposons que $P_{n,\theta}$ admette une densité $f_{n,\theta}(y)$ par rapport à la mesure de Lebesgue sur \mathbb{R} et que $\gamma(t) := E_{n,\theta} \exp\{itS_n\} \in L^\xi(\mathbb{R})$ pour un $\xi \geq 1$. Alors la formule d'inversion s'applique [33][théorème 1.2.1], d'où en posant

$$\varphi_\theta\left(\frac{t}{\delta}\right) := E_{n,\theta} \exp\left\{\frac{it}{\delta}(S_n - \mu_0(\theta) - n\mu_1(\theta))\right\},$$

nous obtenons

$$\begin{aligned} f_{n,\theta}(y) &= \frac{1}{2\pi\delta} \int_{\mathbb{R}} \exp\left\{\frac{it}{\delta}(\mu_0(\theta) + n\mu_1(\theta))\right\} \exp\left(-\frac{ity}{\delta}\right) \varphi_\theta\left(\frac{t}{\delta}\right) dt, \\ P_q(n^{-1}S_n > x) &= \exp\{-n(\theta x - \kappa(\theta))\} \psi_n(\theta) \\ &\quad \times \frac{1}{2\pi\alpha} \int_{\mathbb{R}} \frac{1}{1 + it/\alpha} \varphi_\theta\left(\frac{t}{\delta}\right) \exp\left\{\frac{it}{\delta}(\mu_0(\theta) + n(\mu_1(\theta) - x))\right\} dt, \end{aligned}$$

où $\alpha := \delta\theta$. Choisissons $\hat{\theta}$ tel que $\kappa'(\hat{\theta}) = x$; alors $\hat{\theta}x - \kappa(\hat{\theta}) = \sup_{\theta > 0}(\theta x - \kappa(\theta)) := k^*(x)$. D'autre part, si

$$T_\theta(t)(x, dy) := Q_\theta(x, dy) \exp\left\{\frac{it}{\delta}(f(y) - \mu_1(\theta))\right\},$$

désigne la perturbation de l'opérateur associé au noyau Q_θ et si $U(\theta)$ désigne l'opérateur multiplicatif par la fonction $u(\cdot; \theta)$, alors

$$\psi_n(\theta)\varphi_\theta\left(\frac{t}{\delta}\right) = e^{-\frac{it\mu_0}{\delta}} q.U(\theta)T_\theta(t)^n U(\theta)^{-1} \mathbf{1},$$

avec les notations usuelles. En choisissant $\mu_1(\hat{\theta}) = \int f(x)\pi_\theta(dx) = \kappa'(\hat{\theta}) = x$, nous obtenons l'expression suivante :

$$P_q(n^{-1}S_n > x) = e^{-n\kappa^*(x)} \frac{1}{2\pi\alpha} \int_{\mathbb{R}} \frac{1}{1 + it/\alpha} q.U(\hat{\theta})T_{\hat{\theta}}(t)^n U(\hat{\theta})^{-1} \mathbf{1} dt.$$

Soient $\lambda(t)$ la valeur propre maximale de $T_\theta(t)$, $T_1(t)$ le projecteur propre associée à $\lambda(t)$ et $T_2(t) = I - T_1(t)$, nous avons

$$\lambda(t) = \frac{\lambda(\hat{\theta} + it/\lambda)}{\lambda(\hat{\theta})} e^{\frac{it}{\delta} \mu_1(\hat{\theta})},$$

$$q.U(\hat{\theta})T_{\hat{\theta}}(t)^n U(\hat{\theta})^{-1} \mathbf{1} = \lambda(t)^n q.U(\hat{\theta})T_1(t)U(\hat{\theta})^{-1} \mathbf{1} + q.U(\hat{\theta})T_{\hat{\theta}}(t)^n T_2(t)U(\hat{\theta})^{-1} \mathbf{1}.$$

Prenons $\delta = \sqrt{n\kappa''(\hat{\theta})}$, nous avons alors, pour $|t|$ suffisamment petit, les développements suivants [33][chap. 9] :

$$q.U(\hat{\theta})T_1(t)U(\hat{\theta})^{-1} \mathbf{1} = \gamma_0(\hat{\theta}) + \frac{it\gamma_1(\hat{\theta})}{\sqrt{n\kappa''(\hat{\theta})}} - \frac{t^2\gamma_2(\hat{\theta})}{2n\kappa''(\hat{\theta})} + \omega_{c_1}(\hat{\theta}) \frac{|t|^3}{n^{3/2}},$$

$$\left| \frac{d^k}{dt^k} q.U(\hat{\theta})T_{\hat{\theta}}(t)^n T_2(t)U(\hat{\theta})^{-1} \mathbf{1} \right| \leq \omega_{c_2}(\hat{\theta}) \rho^n |t|, \quad \rho < 1,$$

$$n \log \lambda(t) = -\frac{t^2}{2} - \frac{it^3}{6\sqrt{n}} \zeta_3(\hat{\theta}) + \frac{t^4}{24n} \zeta_4(\hat{\theta}) + \omega_{c_3}(\hat{\theta}) \frac{|t|^5}{n^{3/2}},$$

où $|\omega| \leq 1$ et $\zeta_k(\hat{\theta}) = \kappa^{(k)}(\hat{\theta})/(\kappa^{(2)}(\hat{\theta}))^{k/2}$. Nous obtenons ainsi, pour $|t|$ petit

$$\begin{aligned} q.U(\hat{\theta})T_{\hat{\theta}}(t)^n U(\hat{\theta})^{-1} \mathbf{1} &= \exp\left(-\frac{t^2}{2}\right) \left\{ \gamma_0 + \frac{1}{\sqrt{n}} \left[it\gamma_1 - \frac{it^3}{6} \gamma_0 \zeta_3 \right] \right. \\ &\quad \left. + \frac{1}{n} \left[-\frac{t^2}{2} \gamma_2 + t^4 \left(\frac{\gamma_0 \zeta_4}{24} + \frac{\gamma_1 \zeta_3}{6} \right) - \frac{t^6}{72} \gamma_0 \zeta_3^2 \right] \right\} + \text{Erreur.} \end{aligned}$$

Pour conclure, introduisons les fonctions d'Esscher [33][chap. 2]

$$B_k(\alpha) = \frac{1}{2\pi} \int_{\mathbf{R}} \exp\left(\frac{t^2}{2}\right) \frac{(it)^k}{1 + it/\alpha} dt, \quad \alpha > 0, \quad k = 0, 1, \dots$$

Nous obtenons, par une technique identique à celle utilisée pour obtenir un développement d'Edgeworth, l'approximation du point de selle

$$\begin{aligned} P_q(n^{-1}S_n > x) &= e^{-n\kappa^*(x)} \frac{1}{\alpha} \left\{ B_0(\alpha)\gamma_0 + \frac{1}{\sqrt{n}} \left[\gamma_1 B_1(\alpha) + \frac{\gamma_0 \zeta_3}{6} B_3(\alpha) \right] \right. \\ &\quad \left. + \frac{1}{n} \left[\frac{\gamma_2}{2} B_2(\alpha) + \left(\frac{\gamma_0 \zeta_4}{24} + \frac{\gamma_1 \zeta_3}{6} \right) B_4(\alpha) + \frac{\gamma_0 \zeta_3^2}{72} B_6(\alpha) \right] \right\} + \text{Erreur.} \end{aligned}$$

Les hypothèses justifiant ces résultats figurent dans [33][chap. 9].

Le coefficient γ_0 est positif, puisque

$$\gamma_0 = \int u(x; \hat{\theta}) q(dx) \geq \min u(x; \hat{\theta}) > 0.$$

Posons $R^2 = 2(n\kappa^*(x) - \log \gamma_0)$, alors d'après [33][p. 82, 83],

$$\begin{aligned} \frac{1}{\alpha} B_0(\alpha) &= \exp\left(\frac{\alpha^2}{2}\right) [1 - \Phi(\alpha)] \\ &= \exp\left(\frac{R^2}{2}\right) [1 - \Phi(R)] + \frac{1}{\sqrt{2\pi}} \left[\frac{1}{\alpha} - \frac{1}{R} \right] \\ &\quad - \frac{\alpha - R}{\alpha^3} B_3(\alpha) - \frac{(\alpha - R)^2}{2\alpha^4} (\alpha^{-1} B_6(\alpha) - B_3(\alpha)) + O(n^{-3/2}) \end{aligned}$$

$$P_q(n^{-1}S_n > x) = \{1 - \Phi(R)\} + \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{R^2}{2}\right) \left\{ \frac{1}{\alpha} - \frac{1}{R} \right\} + O\left(\frac{\exp(-R^2/2)}{n^{3/2}}\right),$$

où Φ désigne, comme d'habitude, la fonction de répartition de la loi gaussienne. Soit

$$R_L := R + \frac{1}{R} \log \left(\frac{\alpha}{R} \right),$$

nous obtenons, d'après [33][th. 5.1.1], l'approximation suivante, dénommée également formule de Lugannani–Rice :

$$P_q(n^{-1}S_n > x) \approx 1 - \Phi(R_L).$$

La combinaison des estimations des valeurs propres perturbées et des techniques d'approximation du point de selle pourrait permettre d'améliorer les bornes du type Chernoff obtenues auparavant. Cette perspective débouchera certainement sur des calculs fastidieux et techniques, mais les enjeux en valent la peine.

Bibliographie

- [1] D. Aldous and J. Fill. *Reversible Markov Chains and Random Walks on Graphs*. En préparation.
- [2] D. Bakry. L'hypercontractivité et son utilisation en théorie des semigroupes. In *Ecole d'Été de Probabilités de St-Flour*, volume 1581 of *Lecture Notes in Math.*, pages 1–114. Springer-Verlag, 1994.
- [3] D. Bakry and M. Emery. Diffusions hypercontractives. In *Séminaire de probabilités XIX*, volume 1123 of *Lecture Notes in Math.*, pages 177–206. Springer-Verlag, 1985.
- [4] H. Baumgärtel. *Analytic Perturbation Theory for Matrices and Operators*. Birkhäuser, 1985.
- [5] G. Bennett. Probability inequalities for sums of independent random variables. *J. Amer. Statist. Assoc.*, **57** :33–45, 1962.
- [6] E. Bolthausen. The Berry–Esseen Theorem for Functionals of Discrete Markov Chains. *Z. Wahrscheinlichkeitstheorie verw.*, **54** :59–73, 1980.
- [7] R. C. Bradley. Basic properties of strong mixing conditions. In P. Huber and M. Rosenblatt, editors, *Dependence in Probability and Statistics, a Survey of Recent Results*. Birkhäuser, 1986.
- [8] M. Carbon. Inégalité de Bernstein pour les processus fortement mélangés non nécessairement stationnaires. *C. R. Acad. Paris, Série A(297)* :303–306, 1983.
- [9] D. Chauveau and Diebolt J. MCMC Convergence diagnostic via the Central Limit Theorem. *preprint*, 1997.
- [10] K.L. Chung. *Markov chains with stationary transition probabilities*. Springer-Verlag, 1960.
- [11] K.L. Chung and Z. Zhao. *From Brownian Motion to Schrödinger's Equation*. Springer-Verlag, 1995.
- [12] G. Collomb. Propriétés de convergence presque complète du prédicteur à noyau. *Z. Wahrsch. Verw. Gebiete*, (66) :441–460, 1984.
- [13] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications*. Jones and Bartlett Publishers, 1993.
- [14] J. D. Deuschel and D. W. Stroock. *Large Deviations*. Academic Press, Boston, 1989.
- [15] P. Diaconis and L. Saloff-Coste. What do we know about the Metropolis algorithm ? Technical report, 1995.
- [16] I.H. Dinwoodie. A probability inequality for the occupation measure of a reversible Markov chain. *Ann. Appl. Probab.*, **5** :37–43, 1995.
- [17] J. L. Doob. *Stochastic Processes*. Wiley, New York, 1953.
- [18] P. Doukhan. *Mixing : Properties and Examples*. Springer-Verlag, 1994.
- [19] N. Dunford and J. T. Schwartz. *Linear Operators. Part I. General Theory*. Wiley, New York, 1958.
- [20] D. E. Edmunds and W. D. Evans. *Spectral Theory and Differential Operators*. Oxford Mathematical Monograph, 1990.
- [21] Stewart.P. Ethier and Thomas. G. Kurtz. *Markov processes ; Characterization and convergence*. Wiley, N.Y., 1986.
- [22] W. Feller. *An Introduction to Probability Theory and Its Applications*, volume 1. Wiley & Sons, 3ème édition, 1971.
- [23] W. Feller. *An Introduction to Probability Theory and Its Applications*, volume 2. Wiley & Sons, 2ème édition, 1971.
- [24] O. Forster. *Lectures on Riemann Surfaces*, volume 81 of *Graduate texts in Mathematics*. Springer-Verlag, 1993.

- [25] S. Gallot, D. Hulin, and J. Lafontaine. *Riemannian Geometry*. Universitext. Springer-Verlag, 1990.
- [26] David Gillman. *Hidden Markov Chains : Rates of Convergence and the Complexity of Inference*. PhD thesis, Massachusetts Institute of Technology, July 1993.
- [27] B. V. Gnedenko and A. N. Kolmogorov. *Limit Distributions for Sums of Independent Random Variables*. Addison-Wesley, 1954.
- [28] L. Gross. Logarithmic Sobolev Inequalities and Contractivity Properties of Semigroups. In *Dirichlet forms, Varenna (Italy)*, volume 1563 of *Lecture Notes in Math.*, pages 54–88. Springer-Verlag, 1992.
- [29] T. E. Harris. *The Theory of Branching Processes*. Springer, Berlin, 1963.
- [30] I. A. Ibragimov and Yu. V. Linnik. *Independent and stationary sequences of random variables*. Wolters-Noordhoff Publishing Groningen, 1971.
- [31] I.A. Ibragimov. Some limit theorems for stationary processes. *Theory Probab. Appl.*, 7 :349–382, 1962.
- [32] I.A. Ibragimov. A note on the central limit theorem for dependent random variables. *Theory Probab. Appl.*, 20(1) :135–141, 1975.
- [33] J. L. Jensen. *Saddlepoint Approximations*, volume 16 of *Oxford Statistical Science*.
- [34] T. Kato. *Perturbation theory for linear operators*. Springer, 1966.
- [35] K. Knopp. *Theory of functions*. Dover, 1947.
- [36] A.N. Kolmogorov. Über das Gesetz des iterierten Logarithmus. *Math. Ann.*, **99** :309–319, 1929.
- [37] G. F. Lawler and A. D. 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.
- [38] P. Lezaud. Chernoff-type Bound for Finite Markov Chains. *Ann. Appl. Probab.*, 8(3) :849–867, 1998.
- [39] Brad Mann. *Berry-Esseen Central Limit Theorem For Markov chains*. PhD thesis, Harvard University, 1996.
- [40] A. W. Marshall and I. Olkin. *Inequalities : Theory of Majorization and Its Applications*. Academic Press, New York, 1979.
- [41] S.P. Meyn and R.L. Tweedie. *Markov Chains and Stochastic Stability*. Springer Verlag, 1993.
- [42] S. V. Nagaev. Some limit theorems for stationary Markov chains. *Theory Probab. Appl.*, 2 :378–406, 1957.
- [43] S. V. Nagaev. More exact statements of limits theorems for homogeneous Markov chains. *Theory Probab. Appl.*, 6 :62–81, 1961.
- [44] J. Neveu. *Bases mathématiques du calcul des probabilités*. Masson, 2ème édition, 1980.
- [45] V. V. Petrov. *Sums of Independent Random Variables*. Springer-Verlag, 1975.
- [46] J. W. Pitman. Occupation measures for Markov chains. *Adv. in Appl. Probab.*, **9** :69–86, 1977.
- [47] M. Rosenblatt. *Markov Processes. Structure And Asymptotic Behavior*. Springer-Verlag, 1971.
- [48] E. Seneta. *Non-negative Matrices and Markov Chains*. Springer-Verlag, New York, 1980.
- [49] W. F. Stout. *Almost Sure Convergence*. Academic Press, New York, 1974.
- [50] H. F. Trotter. On the product of semi-groups of operators. *Proc. Amer. Math. Soc.*, 10 :545–551, 1959.
- [51] J.H. Wilkinson. *The Algebraic Eigenvalue Problem*. Clarendon Press, 1965.
- [52] K. Yosida and S. Kakutani. Operator-theoretical treatment of Markoff’s process and mean ergodic theorem. *Annals of Mathematics*, 42(1), 1941.