



THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par : *l'Université Toulouse 3 Paul Sabatier (UT3 Paul Sabatier)*

Présentée et soutenue le *19/07/2018* par :

GUILLAUME COPROS

Stationnarité forte sur des graphes discrets ou quantiques.

JURY

MARC ARNAUDON
REDA CHHAIBI
LAURE COUTIN
ANTOINE LEJAY
LAURENT MICLO
PIERRE PATIE

Professeur des Universités
Maitre de Conférences
Professeure des Universités
Directeur de Recherche
Directeur de Recherche
Associate Professor

Membre du Jury
Membre du Jury
Membre du Jury
Rapporteur
Directeur de Thèse
Rapporteur

École doctorale et spécialité :

MITT : Domaine Mathématiques : Mathématiques appliquées

Unité de Recherche :

Institut de Mathématiques de Toulouse (UMR 5219)

Directeur de Thèse :

Laurent Miclo

Rapporteurs :

Antoine Lejay et Pierre Patie

Remerciements

Dirigé par :

Laurent « Markovitch » Miclo

Rapporteurs :

Pierre Patie & Antoine Lejay

Jury :

Marc Arnaudon & Reda Chhaibi & Laure Coutin

Avec le soutien administratif de :

Marie-Line Domenjole

Agnès Requis & Martine Labruyère

L'incroyable Sylvie Crabos

Toute l'équipe administrative et informatique de l'IMT

Les co-bureau :

La fille toujours dans mon dos pendant trois ans, presque quatre	<i>Laure</i>
L'écolo... en Kangoo	<i>Hugo</i>
La basketteuse toujours blessée, jamais fâchée	<i>Laure (encore)</i>
Le stalker, presque footballeur, presque bilingue (aveyronnais-français)	<i>Hugo (bis)</i>
Le grand frère de thèse	<i>Pierre M.</i>
La startupeuse	<i>Anna</i>
Les gens de passage	<i>Cyril</i> <i>Andressa</i> <i>Kiki</i> ¹ <i>Maël</i> <i>Pascal</i>
Les animaux de compagnie	<i>Feu le poisson rouge d'Hugo</i> (de passage lui aussi) <i>Feu le ver d'Hugo (pareil)</i>

Le groupe de travail sur les permutations d'un ensemble à 2^5 éléments :

110 avec 7-belote *Valentin*

1. Albertus, Mickaël (1992-). Mathématicien, nageur et amateur de pâte à tartiner français.

« Mais Val! »	<i>Maylis</i>
« Ouais, on a gagné les annonces! »	<i>Fabien</i>
Vainqueurs des deux coupes	<i>Kévin F.</i>
	<i>Pierre M. (mais pas le même)</i>
« On peut parler à atout-merde? »	<i>Aurore</i>
Chevalier-Dieu-Philosophe	<i>Anton</i>
Les gens de passage	<i>Théo</i>
	<i>Maxime</i>
	<i>et tous les autres</i>
Le roi de carreau	<i>César</i>

Les « anciens » de Master :

Mon co-bonobo catalan	<i>Loulou</i>
L'entremetteuse	<i>Anne</i>
La néo-parisienne	<i>Mélanie G.</i>
L'ingénieur revenu sur le droit chemin (même si c'est pas assez payé)	<i>Marc</i>
Le coloc éphémère	<i>Clément</i>
Mon partenaire de chansons Disney ...et de randonnée casse-gueule!	<i>Benoît D.</i>
Le rasta parti trop loin	<i>Rémi B.</i>
Expert d'Ethier-Kurtz et relecteur occasionnel	<i>Sylvain</i>
Nounours	<i>Stéphane</i>
Avec également	<i>Maiï</i>
	<i>Ibrahim</i>
	<i>Hue</i>
	<i>Lauren</i>
Mais aussi de vrais acteurs comme ²	<i>Anton (oui, encore)</i>

Les doctorants (enfin... les autres quoi) :

Organisatrice de randonnées	<i>Mélanie B.</i>
Gestionnaire de cadeaux	<i>Malika</i>
Un relecteur inattendu	<i>Pan Zhao</i>
Celui qui m'a ouvert les yeux sur la pub	<i>Nil</i>
La trinité	<i>Claire-s</i>
Les végés	<i>Benoît</i>
	<i>Antoine B.</i>
Le végété ++	<i>Anton (toujours le même)</i>
Le bordelais quand même sympa	<i>William</i>
Le suédois d'en face également roumanophile	<i>Zak</i>
Et bien d'autres encore	<i>Fanny, Magali, Maria, Laura, Damien, Clément, Silvère, Yuriy, ...</i>

2. Voir [7], si vous ne l'avez pas vu

Les profs :

Car on ne peut décidément pas rire de tout, je fais une petite parenthèse dans ce générique pour exprimer ma peine due à la disparition précoce de deux personnes dont les enseignements m'ont particulièrement marqués, ainsi que ma gratitude à leur égard. Je pense bien sûr à Katy Paroux, dont je garderai comme un symbole qu'elle fut la première à m'annoncer les résultats de l'agreg, et à Naoufel Ben Abdallah, qui a su me faire apprécier le calcul différentiel à une époque où ce n'était pas gagné d'avance.

Les responsables de la prépa agreg	<i>Stéphane Lamy</i> <i>Marc Perret</i>
Combo M1-M2-Jury de thèse	<i>Laure Coutin</i>
L'auteur de la loi uniforme sur un disque (l'inoubliable)	<i>Sébastien Gadat</i>
Et tous les supers enseignants d'agreg et de master	<i>Clément Pellegrini</i> <i>Patrick Cattiaux</i> <i>Joseph Najnudel</i> <i>Joan Millès</i> <i>Thomas Dedieu</i> <i>Michel Boileau</i>
	...
Les collègues de l'INSA	<i>Olivier</i> <i>Stéphane</i> <i>Simona</i> <i>Lorick</i> <i>Benjamin</i> <i>Géraldine</i> <i>Sandrine</i>
	...

Ceux qu'on se tape dessus (avec un sabre!)

Celui qui me supporte depuis 20 ans presque sans s'en plaindre	<i>Fabrice (le grand maître)</i>
Le futur papa, compagnon de route pendant... longtemps!	<i>Rémi</i>
La future mariée-retraîtée	<i>Charlotte</i>
L'autre grand maître	<i>Damien</i>
et tous les autres...	<i>Brieuc, Romain, Marco, Alban, Lionel, Sylvie, Tristan, Fanny, Guillaume, ...</i>

Les inclassables :

Les anciens colocs	<i>Romain</i> <i>Antoine</i> <i>Manni</i>
Le basque accueilli et accueillant	<i>Maxime</i>
L'ex-hôtesse dont la vie est un roman (en plusieurs tomes)	<i>Julie</i>
Celle qu'elle est toujours là depuis 15 ans (et un centre de vacance à Moliets)	<i>Béné</i>
Les sœurs aux Jérém(y/ie)	<i>Sandrine</i> <i>Valou</i>
La joyeuse petite nordiste	<i>Lise</i>

Un cervidé affectif *Manu*
Le père, la mère, les fils et le cousin *Michou, Nini, Théo, Hugo et Anthony*

La famille (presque par ordre d'apparition) :

Les indispensables *Mes parents*
Les grands-parents *Grand-Père, Mamette, Papie et Mamie*
PP *Parrain Pierre*
SMS *Super Marraine Sophie*
Tontons et (grand-)taties *Sabine, Thierry, Sandrine, François,
Michèle, Michèle*
Sœur *Madeline*
Les cousins *Emilie, Adrien, Nathan, Noémie, Lison,
Baptiste, Lola*
Le « beau-frère » *Cédric*
Les animaux de compagnies Le chat *Lumo* (même s'il est gros)

Echipa din România (familia a doua):

Și indispensabilii Mali, Ghiță și Tată Ionel
Domnul și Doamna Pis Ama și Andrei
Flăcău, flăchea și flăcăuș Răzvan, Alexandra și Tudor
"Ciun tit"și parinții Andra, Lau și Victor
Tinerii bunici Cornel și Tatiana

Ioana:

♡ *Ioana*

L'inconvénient de faire le malin avec ses remerciements, c'est qu'on ne peut pas dire à chaque personne pourquoi on la remercie. Alors... J'espère que vous le savez ! (Un indice si vous ne le savez pas : ça peut être lié à votre rôle dans cette histoire) Donc à tou-te-s celles et ceux et que j'ai cité-e-s, et à celles et ceux que j'ai sûrement oublié-e-s je dirai simplement : Merci ! *Mulțumesc!*

(mais aussi : *Obrigado! Спасибо! Danke schöne! Gràcies! Arigato! Cám ơn bạn! Xièxiè! Tack så mycket! Eskerrik asko!*)

Table des matières

Notations	vii
Présentation générale	ix
0 Introduction	1
0.1 Outils généraux	1
0.1.1 Processus stochastiques	1
0.1.2 Caractérisation des processus de Markov homogènes	4
0.1.3 Convergence de mesures	7
0.2 Temps forts de stationnarité et dualité	10
0.2.1 Définition d'un temps fort de stationnarité	10
0.2.2 Dualité	11
0.2.3 Entrelacement	12
0.3 Historique	14
0.3.1 Développements théoriques	14
0.3.2 Exemples d'applications	16
0.4 Présentation des travaux	18
I Stationnarité pour une chaîne de Markov	25
1 Dualité et entrelacement	27
1.1 Processus minimaux	27
1.2 Dualité avec explosion	38
2 Application à la recherche de temps forts de stationnarité finis	47
2.1 Processus de vie et de mort	47
2.2 Marche aléatoire sur un graphe	54
2.2.1 Définition du cadre et notations	54
2.2.2 Générateur dual	57
2.2.3 Comportement du processus dual	60
2.2.4 Preuve du théorème 0.4.1	67
3 Dualité pour une marche aléatoire non-réversible sur un graphe	71
II Dualité pour une diffusion sur un graphe quantique	81
4 Diffusion sur un graphe	83
4.1 Cadre théorique	83
4.2 Réversibilité et mesure invariante	86
4.3 Espace d'état dual et formule de Stokes	89

4.4	Générateur dual	93
5	Étude d'un cas particulier	99
5.1	Diffusion avec une dérive singulière	99
5.2	Graphe à deux branches sans dérive	107
5.3	Adaptation au graphe en étoile	112
6	Approche par des processus de Markov finis	115
6.1	Approximation d'une diffusion	115
6.2	Approximation du dual d'une diffusion	126
6.2.1	Redéfinition du générateur dual	126
6.2.2	Mise en place de la convergence	130

Notations

\mathbb{N} , \mathbb{Z} , \mathbb{Q} , \mathbb{R} : ensemble des entiers positifs, des entiers relatifs, des rationnels, des réels

\mathbb{k}_+ , \mathbb{k}_- , \mathbb{k}^* : ensemble des éléments positifs, négatifs, non-nuls de $\mathbb{k} = \mathbb{N}$, \mathbb{Z} , \mathbb{Q} , \mathbb{R}

$\llbracket m, n \rrbracket$: intervalle d'entiers

$\mathcal{P}(E)$: ensemble des parties de E

$\mathcal{P}(E)$: ensemble des mesures de probabilités sur (un espace mesurable) E

$\mathcal{M}(E)$: ensemble des mesures signées sur E

$\|\cdot\|_{VT}$, d_{VT} : norme de la variation totale sur $\mathcal{M}(E)$, et distance induite sur $\mathcal{P}(E)$

$\rho(\cdot, \cdot)$: distance de Prokhorov sur $\mathcal{P}(E)$

$\text{sep}(\cdot, \cdot)$: séparation sur $\mathcal{P}(E)$

$\delta_{i,j}$: symbole de Kronecker (1 si $i = j$, 0 sinon)

δ_x : mesure de Dirac en x

$\mathbf{1}$: fonction constante égale à un

$\mathbf{1}_A$: fonction indicatrice de l'ensemble A

$(\Omega, \mathcal{A}, \mathbb{P})$: l'espace de probabilités sous-jacent

\mathbb{E} : l'espérance sous \mathbb{P}

\mathcal{F}^X : filtration engendrée par un processus X

$\mathcal{L}(X)$: loi d'une variable aléatoire X

$\xrightarrow{\mathbb{P}}$, $\xrightarrow{\mathcal{L}}$, $\xrightarrow{\text{p.s.}}$: convergence en probabilité, en loi, presque sûre

$\mathcal{B}(E)$: tribu borélienne d'un espace topologique E

$\mathcal{B}(E)$: ensemble des fonctions bornées de E dans \mathbb{R}

$\mathcal{C}(E)$, $\mathcal{C}^n(E)$, $\mathcal{C}^\infty(E)$: ensemble des fonctions réelles continues, n fois continûment différentiables, indéfiniment différentiables sur E

$\mathcal{C}_b(E)$, $\mathcal{C}_b^n(E)$, $\mathcal{C}_b^\infty(E)$: ensemble des fonctions de $\mathcal{C}(E)$, $\mathcal{C}^n(E)$, $\mathcal{C}^\infty(E)$, bornées ainsi que leurs dérivées partielles

$D_E(I)$: ensemble des fonction càdlàg de I dans E

\bar{A} , $\overset{\circ}{A}$, ∂A : fermeture, intérieur, frontière d'un ensemble A pour la topologie ambiante

A^c : complémentaire de A

Présentation générale

Considérons un paquet de n cartes, que l'on souhaite mélanger en suivant la procédure suivante : prendre la carte du dessus, l'insérer dans le paquet à une position choisie aléatoirement et uniformément sur toutes les positions possibles (y compris sur le dessus du paquet, auquel cas la configuration reste inchangée), et répéter ces deux étapes autant de fois que nécessaire. Cette méthode porte le nom de *top in at random* et, vous l'aurez compris, n'est pas des plus efficaces lorsque n est grand ; mais peu importe, là n'est pas son intérêt. Pour ce mélange comme pour les autres, une question qui intéresse le joueur de cartes honnête est « quand arrêter de mélanger, de manière à ce que le mélange soit suffisamment uniforme ? ».

Imaginons que l'on colle (temporairement) une étiquette sur la dernière carte du paquet avant de commencer à mélanger, afin de pouvoir suivre son parcours. Il va arriver un instant, que l'on appelle T_1 , où une carte sera insérée en dernière position, ce qui aura pour conséquence de faire remonter d'un rang notre carte-repère, qui se trouvera donc en position $n - 1$. Puis, au temps T_2 , une deuxième carte sera insérée dessous (en dernière ou avant dernière position), notre carte-repère va monter en position $n - 2$, et ainsi de suite jusqu'au temps T_{n-1} , où la carte initialement en-dessous aura atteint la surface. À chacun de ces temps T_i , on ne peut pas dire grand chose sur l'identité des cartes en-dessous de la carte étiquetée (ce n'est même pas un tirage uniforme de i cartes parmi les $n - 1$ restantes). En revanche, ce que l'on sait c'est qu'elles sont ordonnées de manière totalement arbitraire, ou autrement dit, que chaque arrangement des i dernières cartes est équiprobable. Au temps T_{n-1} , les $n - 1$ cartes non étiquetées sont donc uniformément mélangées. Au temps $T = T_{n-1} + 1$, le mélange est terminé. Le lecteur avisé (et honnête) aura anticipé, et retiré l'étiquette entre les temps T_{n-1} et T .

Le processus de mélange peut être vu comme une marche aléatoire sur le groupe \mathfrak{S}_n des permutations d'un ensemble à n éléments, et au temps T la loi du processus est uniforme sur les permutations. Ce problème est étroitement lié avec celui du collectionneur de coupons, et notamment chaque $T_i - T_{i-1}$ suit une loi géométrique de paramètre $\frac{i}{n}$. Le temps T étant la somme de ces variables géométriques, son espérance est de l'ordre de $n \ln(n)$.

Cet exemple, ainsi que d'autres exemples de marches aléatoires sur des groupes, a été introduit par Aldous et Diaconis [1] en 1986, et le temps T , auquel la position du processus suit une loi uniforme, a d'abord été baptisé « temps fort uniforme ». En 1990, ce concept a été étendu par Diaconis et Fill aux chaînes de Markov (en temps discret [23] ou continu [32]) dont la loi stationnaire (ou « mesure invariante ») n'est pas forcément uniforme, et la terminologie de « temps fort de stationnarité » a été adoptée. Cette généralisation a été accompagnée du développement de la notion de *dual de stationnarité forte* et en particulier de *dual Λ -lié*. C'est cette seconde notion qui va particulièrement nous intéresser, car elle permet d'obtenir des informations pertinentes sur l'évolution de la loi du processus étudié, par l'observation directe d'un second processus. En particulier, on va demander à ce processus dual d'avoir un état absorbant, et le temps d'absorption constituera un temps fort de stationnarité pour le processus primal.

En pratique, un choix opportun consiste à prendre le processus dual comme un sous-ensemble de l'espace d'état³, évoluant avec le temps, et à le coupler avec le processus primal de manière à imposer à celui-ci d'être toujours à l'intérieur de son dual. Encore mieux, on peut les coupler de manière à ce qu'à chaque instant la loi du primal soit sa loi stationnaire restreinte au sous-ensemble dual. Si l'on choisit bien la dynamique de ce dernier, il va s'agrandir jusqu'à remplir l'espace entier, ce qui signifie que le processus dual sera distribué selon sa loi stationnaire. La problématique consiste donc, étant donné *a priori* un processus primal, à choisir une bonne dynamique pour un processus dual que l'on va construire de manière à respecter les conditions que l'on vient d'énoncer.

Pour ce faire, on va pouvoir s'appuyer sur des outils classiques dans la théorie des processus de Markov, que sont les matrices de transition dans le cadre des chaînes de Markov, ou les semi-groupes et les générateurs dans un cadre plus général. Sous de bonnes conditions, chacun de ces objets caractérise la loi complète d'un processus, à condition de se donner en plus une loi de probabilité pour l'état initial. Dans des cadres relativement simples, il a été montré par Aldous, Diaconis et Fill, que le problème de recherche d'un dual pouvait se ramener à la recherche d'un de ces objets le caractérisant, et satisfaisant une relation algébrique (que l'on appellera « entrelacement ») avec l'objet correspondant pour le processus primal :

$$A^* \Lambda = \Lambda A$$

où Λ est le noyau de transition permettant de passer du dual au primal, et A et A^* désignent, au choix, les matrices de transition, les semi-groupes ou les générateurs des processus primal et dual respectivement.

L'objectif dans cette thèse sera de déterminer des générateurs satisfaisant cette relation d'entrelacement dans différents contextes, de vérifier qu'ils caractérisent bien des processus duaux, et dans certains cas, d'en tirer des conditions d'existence d'un temps fort de stationnarité fini.

Dans le chapitre d'introduction, après avoir rappelé les définitions et propriétés de base liées aux processus stochastiques, ainsi que la caractérisation des processus de Markov par leurs matrices de transition, semi-groupes et générateurs, on s'attachera à présenter une synthèse de la théorie des temps forts de stationnarité. On définira notamment les objets présentés ici de manière rigoureuse, en étendant à chaque fois les définitions et résultats existants à un cadre aussi général que possible. On évoquera également quelques utilisations de cette théorie dans la littérature, notamment pour l'étude du phénomène de cut-off ou pour le perfect sampling.

Dans la première partie, on s'intéresse à une marche aléatoire à temps continu sur un graphe discret. Un tel processus peut être défini assez simplement par son générateur, et l'on se pose la question « quelle(s) condition(s) doit satisfaire le générateur pour que ce processus admettent un temps fort de stationnarité fini ? ». Si le graphe est supposé fini, cette question a déjà été traitée par Fill [32], et on le supposera donc infini, avec quelques hypothèses supplémentaires. On supposera notamment dans la section 2.2 que le nombre de directions infinies est fini, ce qui exclut les marches aléatoires sur \mathbb{Z}^n ou même \mathbb{N}^n , dès que n est supérieur ou égal à 2. Dans le chapitre 1, on vérifiera que la condition d'entrelacement des générateurs est bien suffisante dans ce contexte pour l'existence d'un dual Λ -lié, et on l'appliquera à une marche aléatoire, sur \mathbb{Z} dans la section 2.1 et sur un graphe plus général dans la section 2.2, pour trouver un dual particulier. Ce dual nous

3. Espace « d'état » ou « d'états » ? L'usage hésite. Dans cette thèse nous avons fait le choix de privilégier la première orthographe, par opposition à « espace des états », et par analogie avec l'anglais « state space ».

permettra de montrer qu'il existe un temps fort de stationnarité fini si et seulement si les frontières de l'espace sont des frontières entrantes (*i.e.* que l'on n'atteint pas en temps fini, mais dont on revient rapidement). Le troisième chapitre propose une vue d'ensemble des différents duaux que l'on peut construire pour une même marche aléatoire.

Dans la deuxième partie, on considère une diffusion sur un graphe continu et compact. On explicite notamment sa mesure invariante dans la section 4.2, et on donne un générateur satisfaisant la relation d'entrelacement. Le problème à ce stade est que ce type de générateur n'a, à notre connaissance, pas été étudié jusqu'à présent, et que l'existence d'un processus associé n'est pas garantie. Une utilisation du théorème de Hille-Yosida devrait permettre de régler ce problème, mais nous ne sommes pas encore en mesure de le démontrer, et espérons avancer dans ce sens prochainement. Dans le chapitre 6, nous proposons une autre alternative, consistant à approximer la diffusion par une suite de marches aléatoires sur des sous-graphes discrets, et faisant donc le lien entre les parties I et II. Si la convergence du processus primal est établie, celle du dual est, là aussi, incomplète.

Chapitre 0

Introduction

La convergence des processus de Markov vers leur loi stationnaire a été beaucoup étudiée par des outils analytiques. Ainsi, des théorèmes classiques permettent de trouver un temps t_0 déterministe auquel la loi d'un processus X est arbitrairement proche de sa loi stationnaire, au sens par exemple de la distance en variation totale.

Une approche plus probabiliste consiste à chercher un temps T aléatoire pour lequel X_T est distribué *exactement* selon sa loi stationnaire, et ce indépendamment de T . Un tel temps aléatoire est appelé *temps fort de stationnarité* (en anglais : *strong stationary time*), et sera défini plus précisément en (6). Cet outil a été introduit pour la première fois par Aldous et Diaconis [1] en 1986.

Dans cette introduction, nous commencerons par fixer la terminologie et rappellerons les notions générales utilisées dans la théorie des temps forts de stationnarité. Après avoir introduit précisément cette théorie et présenté les principaux résultats existants, nous ferons une synthèse historique de ses développements successifs et de ses applications. Enfin, nous présenterons de manière succincte les résultats exposés dans cette thèse.

0.1 Outils généraux

Dans toute cette thèse, $(\Omega, \mathcal{A}, \mathbb{P})$ désignera un espace de probabilités, que l'on supposera « suffisamment grand » pour contenir tout l'aléa nécessaire, notamment lorsqu'il s'agira de coupler des variables aléatoires. Cet espace sera généralement sous-entendu et en particulier, en l'absence de précision, « presque sûr(ement) » signifiera « \mathbb{P} -presque sûr(ement) ».

0.1.1 Processus stochastiques

Définitions générales

On rappelle ici quelques définitions concernant la théorie générale des processus stochastiques, formalisée par Doob [27] en 1953. La plupart de ces définitions se trouvent dans les livres de Ethier et Kurtz [29] et de Dellacherie et Meyer [19]. Toutes ces notions sont également présentées de manière moins formelle (sans utilisation de la théorie de la mesure) dans les livres de Karlin et Taylor ([46] et [47]).

Soient I un ensemble d'indices totalement ordonné (en général $I = \mathbb{N}$ ou \mathbb{R}_+) et $(\mathcal{S}, \mathcal{S})$ un espace mesurable. Un *processus stochastique* X , indexé par I et d'espace d'état \mathcal{S} est une famille de variables aléatoires $(X_i)_{i \in I}$ à valeurs dans \mathcal{S} . Pour tout $i_1 \leq i_2 \in I$, on

notera $X_{[i_1, i_2]}$ pour $\{X_i, i_1 \leq i \leq i_2\}$. Pour tout $\omega \in \Omega$, la fonction $i \in I \mapsto X_i(\omega)$ est une *trajectoire* de X .

Une *filtration* $\mathcal{F} = (\mathcal{F}_i)_{i \in I}$ est une suite croissante (pour l'inclusion) de sous-tribus de \mathcal{A} , et on notera :

$$\mathcal{F}_\infty := \bigvee_{i \in I} \mathcal{F}_i = \sigma \left(\bigcup_{i \in I} \mathcal{F}_i \right)$$

Le processus X est dit *adapté* à \mathcal{F} si pour tout $i \in I$, X_i est \mathcal{F}_i -mesurable. On appelle *filtration engendrée par X* et on notera $\mathcal{F}^X = (\mathcal{F}_i^X)_{i \in I}$ la plus petite filtration à laquelle X est adapté, c'est-à-dire :

$$\forall i \in I \quad \mathcal{F}_i^X := \sigma(\{X_j, j \leq i\})$$

D'un point de vue heuristique, une tribu représente une certaine « quantité d'information disponible » et une filtration représente l'évolution au cours du temps de cette quantité d'information. La croissance de la famille de tribus composant la filtration traduit donc l'absence de perte d'information. Un processus est adapté à une filtration si celle-ci contient à chaque instant toute les informations sur le passé du processus. La filtration engendrée est celle qui n'en contient pas plus.

Une variable aléatoire τ à valeurs dans I est un *\mathcal{F} -temps d'arrêt* si à chaque instant la connaissance du passé suffit pour déterminer si le temps τ est déjà atteint, ou en termes mathématiques si pour tout $i \in I$, l'événement $\{\tau \leq i\}$ est \mathcal{F}_i -mesurable. On dira que τ est un temps d'arrêt pour X si c'est un \mathcal{F}^X -temps d'arrêt. On notera \mathcal{F}_τ la tribu constituée des événements $A \in \mathcal{S}$ tels que $A \cap \{\tau \leq i\} \in \mathcal{F}_i$, pour tout $i \in I$. L'intuition au sujet de la tribu \mathcal{F}_τ est la même que celle pour \mathcal{F}_i , $i \in I$: c'est l'information disponible au temps τ .

Une variable aléatoire T à valeurs dans I est appelé *\mathcal{F} -temps d'arrêt randomisé* s'il existe une filtration $\mathcal{F}^* = (\mathcal{F}_i^*)_{i \in I}$ telle que $\mathcal{F}_i \subset \mathcal{F}_i^*$ pour tout $i \in I$ et que T est un \mathcal{F}^* -temps d'arrêt. De manière équivalente, T est un \mathcal{F} -temps d'arrêt randomisé si il existe une sous-tribu \mathcal{G} de \mathcal{S} , indépendante de \mathcal{F}_∞ , telle que T est un temps d'arrêt relativement à la filtration $(\sigma(\mathcal{G} \cup \mathcal{F}_i))_{i \in I}$ (cf Fill [31], section 2.2). On peut interpréter cette définition de la manière suivante : la connaissance du passé ne suffit peut-être pas pour déterminer si le temps T a été atteint, mais le futur n'apporte pas plus d'information à ce sujet.

Un processus M à valeurs dans \mathbb{R} et adapté à la filtration \mathcal{F} est une *\mathcal{F} -martingale* si pour tout $i \leq j \in I$, M_j est intégrable ($\mathbb{E}(|M_j|) < +\infty$) et :

$$\mathbb{E}(M_j | \mathcal{F}_i) = M_i$$

C'est une *\mathcal{F} -sous-martingale* (resp. *sur-martingale*) si le symbole $=$ est remplacé par \geq (resp. \leq), une *\mathcal{F} -martingale locale* s'il existe une suite croissante de \mathcal{F} -temps d'arrêts $(\tau_n)_{n \in \mathbb{N}}$ telle que $\tau_n \rightarrow \infty$ et $(M_{t \wedge \tau})_{t \geq 0}$ est une martingale pour tout $n \in \mathbb{N}$, et enfin c'est une *\mathcal{F} -semi-martingale* si c'est la somme d'une martingale locale et d'un processus à variation bornée. Dans ces définitions, on omettra la filtration \mathcal{F} si celle-ci est égale à \mathcal{F}^M .

Un processus stochastique X adapté à la filtration \mathcal{F} est un *processus de Markov relativement à \mathcal{F}* si pour tout $i \leq j \in I$ et tout $A \in \mathcal{A}$ on a :

$$\mathbb{P}(X_j \in A | \mathcal{F}_i) = \mathbb{P}(X_j \in A | X_i)$$

c'est-à-dire si le futur ne dépend du passé qu'à travers le présent : le processus n'a en quelque sorte pas de mémoire de son évolution passée. Si $\mathcal{F} = \mathcal{F}^X$, on dira simplement que c'est un *processus de Markov*. Si I est un sous-ensemble de \mathbb{R} (ce qui sera généralement le cas), X est dit *homogène* si la probabilité précédente ne dépend de i et j qu'à travers de $j - i$. Enfin, si \mathcal{S} est dénombrable on parle de *chaîne de Markov*.

Régularité des trajectoires

Toutes ces définitions ne sont en général pas très pratiques sans quelques restrictions sur les objets considérés. À partir de maintenant, on suppose que \mathcal{S} est un espace polonais (*i.e.* métrique séparable et complet), muni de la tribu borélienne $\mathcal{B}(\mathcal{S})$, et $I = \mathbb{N}$ ou \mathbb{R}_+ .

On rappelle qu'une fonction de I dans \mathcal{S} est dite *càdlàg* si elle est continue à droite en tout point $x \in I$ et admet une limite à gauche en tout point $x \in I \setminus \{0\}$. Bien sûr, si $I = \mathbb{N}$, cette notion est triviale. On note $D_{\mathcal{S}}(I)$ l'ensemble des fonction càdlàg de I dans \mathcal{S} . On munit $D_{\mathcal{S}}(I)$ de la distance de Skorokhod : sa définition étant technique et ne servant pas dans la suite, nous renvoyons au livre de Billingsley [8] pour plus de précisions. Rappelons tout de même quelques points intéressants sur cet espace métrique.

D'abord, puisqu'on a supposé \mathcal{S} polonais, alors $D_{\mathcal{S}}(I)$ l'est aussi. Plus généralement, si on n'avait supposé \mathcal{S} que séparable, alors $D_{\mathcal{S}}(I)$ serait séparable, et de même pour la complétude.

Tout élément de $D_{\mathcal{S}}(I)$ admet au maximum une infinité dénombrable de points de discontinuité.

Sur le sous-espace $C_{\mathcal{S}}(I)$ des fonctions continues de I dans \mathcal{S} , la topologie induite par la distance de Skorokhod coïncide avec celle de la convergence uniforme. Cela n'est en revanche pas vrai sur l'espace $D_{\mathcal{S}}(I)$, car deux fonctions légèrement décalées en temps seront proches au sens de la distance de Skorokhod ce qui ne sera pas le cas pour la distance de la norme uniforme si ces fonctions ne sont pas continues (autrement dit, la distance de Skorokhod « tolère » de petits écarts en temps et en espace, au contraire de celle de la convergence uniforme qui ne tolère qu'une déformation en espace).

Un processus stochastiques est à *trajectoires càdlàg* si pour presque tout $\omega \in \Omega$, la fonction $t \mapsto X_t(\omega)$ est càdlàg. Un tel processus peut être vu comme une variable aléatoire à valeurs dans $D_{\mathcal{S}}(I)$ muni de sa tribu borélienne. On parlera alors de la *loi* d'un processus pour désigner l'élément de $\mathcal{P}(D_{\mathcal{S}}(I))$ associé à cette variable aléatoire (la notation $\mathcal{P}(\cdot)$ sera utilisé dans toute cette thèse pour désigner l'ensemble des mesures de probabilité sur un espace mesurable). La loi d'un processus X est entièrement caractérisée par ses *lois fini-dimensionnelles*, c'est à dire l'ensemble des probabilités de la forme :

$$\mathbb{P}(X_{t_1} \in A_1, \dots, X_{t_n} \in A_n) \tag{1}$$

avec $n \in \mathbb{N}$, $A_1, \dots, A_n \in \mathcal{B}(\mathcal{S})$ et $0 \leq t_1 < \dots < t_n < \infty$.

Enfin, la filtration engendrée par un processus X à trajectoires càdlàg est *continue à droite* si $I = \mathbb{R}_+$:

$$\forall t \geq 0, \quad \bigcap_{s>t} \mathcal{F}_s^X = \mathcal{F}_t^X$$

On rappelle les propriétés suivantes des temps d'arrêts lorsque la filtration est continue à droite, même si quelques-unes sont vraies pour toute filtration :

Proposition 0.1.1. *Si \mathcal{F} est continue à droite, (ou si $I = \mathbb{N}$), alors pour tout $i \in I$ et toute famille dénombrable $(\tau_n)_{n \in \mathbb{N}}$ de \mathcal{F} -temps d'arrêts :*

$$i, \tau_1 + i, \inf_n \tau_n, \sup_n \tau_n, \liminf_n \tau_n \text{ et } \limsup_n \tau_n$$

sont des \mathcal{F} -temps d'arrêts.

À partir de maintenant, tous les processus considérés seront supposés à trajectoire càdlàg, et nous ne le mentionnerons pas à chaque fois.

Semi-groupe et générateur

Soit $X = (X_i)_{i \in I}$ un processus de Markov relativement à une filtration $\mathcal{F} = (\mathcal{F}_i)_{i \in I}$, homogène (à trajectoires càdlàg) sur \mathcal{S} . On définit successivement la *fonction de transition* de X :

$$\forall i \in I, x \in \mathcal{S}, A \in \mathcal{S}, \quad P(i, x, A) = \mathbb{P}(X_i \in A \mid X_0 = x)$$

son *semi-groupe* $P(\cdot)$ agissant sur l'ensemble des fonctions réelles bornées de \mathcal{S} :

$$\forall f \in \mathcal{B}(\mathcal{S}), i \in I, x \in \mathcal{S}, \quad P(i)f(x) := \int_{\mathcal{S}} f(y)P(i, x, dy)$$

et si $I = \mathbb{R}_+$, son *générateur* :

$$\begin{aligned} \mathcal{L} : \mathcal{D}(\mathcal{L}) &\rightarrow \mathcal{B}(\mathcal{S}) \\ f &\mapsto \lim_{t \rightarrow 0} \frac{P(t)f - f}{t} \end{aligned}$$

où $\mathcal{D}(\mathcal{L})$ est le sous-ensemble de $\mathcal{B}(\mathcal{S})$ sur lequel cette limite existe (au sens de la convergence uniforme).

Ces trois outils ont une importance particulière pour caractériser la loi du processus, et nous reviendrons dessus plus en détails dans la section suivante.

Pour tout \mathcal{F} -temps d'arrêt τ , on dit que X satisfait la *propriété de Markov forte en τ* si :

$$\forall A \in \mathcal{S}, i \in I, \quad \mathbb{P}(X_{\tau+i} \in A \mid \mathcal{F}_\tau) = P(i, X_\tau, A)$$

Si le processus X satisfait la propriété de Markov forte en tout \mathcal{F} -temps d'arrêt, il est dit *fortement Markovien*.

Une mesure de probabilité μ sur \mathcal{S} est dite *invariante* pour X si pour tout $f \in \mathcal{B}(\mathcal{S})$ et tout $i \in I$:

$$\mu(P(i)f) = \mu(f) \tag{2}$$

et elle est dite *réversible* si pour tout $f, g \in \mathcal{B}(\mathcal{S})$ et tout $i \in I$:

$$\mu(fP(i)g) = \mu(gP(i)f) \tag{3}$$

En terme de générateur, les équations (2) et (3) se traduisent respectivement par :

$$\begin{aligned} \mu(\mathcal{L}f) &= 0 \\ \mu(f\mathcal{L}g) &= \mu(g\mathcal{L}f) \end{aligned}$$

pour tout $f, g \in \mathcal{D}(\mathcal{L})$.

0.1.2 Caractérisation des processus de Markov homogènes

Soit $X = (X_i)_{i \in \mathbb{N}}$ un processus de Markov homogène sur \mathcal{S} , $P(\cdot, \cdot, \cdot)$ sa fonction de transition et $P(\cdot)$ son semi-groupe. Si $I = \mathbb{N}$ et \mathcal{S} est dénombrable, la matrice $P = (P_{x,y})_{(x,y) \in \mathcal{S}^2}$ définie par :

$$\forall x, y \in \mathcal{S}, \quad P_{x,y} = P(1, x, \{y\})$$

est appelée *matrice de transition* de la chaîne de Markov X et caractérise celle-ci de manière unique. Dans toute cette thèse, les chaînes de Markov à temps discret ne seront évoquées que dans l'introduction pour énoncer la théorie générale des temps forts de stationnarité, ou pour construire une chaîne de Markov à temps continu. La littérature sur ce sujet étant plus qu'abondante ([10],[6],[69],...) nous ne nous y attarderons pas plus. Nous supposons donc dans cette section que $I = \mathbb{R}$, et on note \mathcal{L} le générateur de X .

Caractérisation générale et problème de martingale

On rappelle que l'on ne considère que des processus à trajectoires càdlàg, vus comme des variables aléatoires à valeurs dans $D_S(I)$. La loi d'un processus est donc entièrement caractérisée par ses lois fini-dimensionnelles, mais cela n'est en général pas beaucoup plus pratique que la loi elle-même (comment décrire toutes les probabilités impliquées?). Une première observation est que, grâce à la propriété de Markov et à l'homogénéité en temps, ces lois fini-dimensionnelles sont entièrement caractérisées par la loi μ_0 de X_0 et sa fonction de transition. On l'explique ici avec $n = 2$ dans (1), sa généralisation découlant d'une simple récurrence :

$$\mathbb{P}(X_{t_1} \in A_1, X_{t_2} \in A_2) = \int_{\mathcal{S}} \left(\int_{A_1} P(t_2 - t_1, y, A_2) P(t_1, x, dy) \right) \mu_0(dx)$$

De manière équivalente, on peut remplacer la fonction de transition par le semi-groupe, puisque pour tout $t \geq 0$, $x \in \mathcal{S}$, $A \in \mathcal{B}(\mathcal{S})$:

$$P(t, x, A) = P(t)f(x), \quad \text{avec } f = \mathbf{1}_A$$

Sous certaines conditions (théorème de Hille-Yosida, voir le livre de Ethier-Kurtz [29] ou les articles originaux de Hille [43] et Yosida [93]), le générateur et son domaine caractérisent le semi-groupe, et donc le processus. L'expression du générateur est en général plus simple que celle du semi-groupe, mais son domaine complet n'est pas toujours accessible.

Une autre caractérisation qui va alors être utilisée est liée à l'observation du fait que pour tout $f \in \mathcal{D}(\mathcal{L})$, le processus $M^f = (M_t^f)_{t \geq 0}$, défini par :

$$M_t^f = f(X_t) - f(X_0) - \int_0^t \mathcal{L}f(X_s)ds$$

est une \mathcal{F}^X -martingale. On va donc caractériser X à partir d'un opérateur linéaire $A : \mathcal{D}(A) \subset \mathcal{B}(\mathcal{S}) \rightarrow \mathcal{B}(\mathcal{S})$, par l'intermédiaire du *problème de martingale* de Stroock et Varadhan ([82] et [83]) : X est dit *solution du problème de martingale associé à A* si pour tout $f \in \mathcal{D}(A)$ il existe une \mathcal{F}^X -martingale M telle que :

$$f(X_t) = f(X_0) - \int_0^t Af(X_s)ds + M_t$$

et *solution du problème de martingale associé à (A, μ)* , avec $\mu \in \mathcal{P}(\mathcal{S})$, si de plus $\mathcal{L}(X_0) = \mu$. Le problème de martingale associé à (A, μ) est dit *bien posé* si sa solution est unique (en loi), et il est dit bien posé pour A si pour tout $\mu \in \mathcal{P}(\mathcal{S})$ il est bien posé pour (A, μ) . Dans ce cas, on dira encore que A est le générateur du processus X .

Les différentes approches pour caractériser X par un générateur et son domaine sont développées plus en détails dans [29].

Chaînes de Markov

Dans le cas des chaînes de Markov à temps continu, les générateurs et les semi-groupes peuvent être vus comme des matrices, et la caractérisation de la loi du processus en est simplifiée. Ce qui suit rappelle brièvement cette caractérisation. Pour plus de détails, on renvoie à Norris [69], Anderson [3] ou encore Liggett [54].

Soit \mathcal{S} un espace topologique dénombrable (pas nécessairement muni de la topologie discrète). On ajoute à \mathcal{S} un point isolé Δ , appelé *point cimetièr*e et on prolonge n'importe quelle fonction réelle f de \mathcal{S} à $\mathcal{S} \cup \{\Delta\}$ en posant $f(\Delta) = 0$. Une *Q-matrice stable et conservative* sur \mathcal{S} est une matrice $(\mathcal{L}_{x,y})_{(x,y) \in \mathcal{S}^2}$ vérifiant :

$$\begin{aligned} 0 \leq \mathcal{L}_{x,y} < +\infty, \quad \forall x \neq y \in \mathcal{S} \\ \sum_{y \neq x} \mathcal{L}_{x,y} = -\mathcal{L}_{x,x} < +\infty, \quad \forall x \in \mathcal{S} \end{aligned}$$

Pour toute Q-matrice \mathcal{L} , on notera $\mathcal{L}_x := -\mathcal{L}_{x,x}$, et on prolonge \mathcal{L} à $\mathcal{S} \cup \{\Delta\}$ en posant $\mathcal{L}_{\Delta,x} = \mathcal{L}_{x,\Delta} = \mathcal{L}_{\Delta,\Delta} = 0$. Ajoutée à une mesure de probabilité μ_0 sur \mathcal{S} , une Q-matrice \mathcal{L} définit un processus X à valeur dans $\mathcal{S} \cup \{\Delta\}$, où Δ est un point cimetièr e n'appartenant pas à \mathcal{S} , par l'intermédiaire de sa chaîne de sauts et de ses temps de sauts :

- $(Y_n)_{n \in \mathbb{N}}$ est une chaîne de Markov à temps discret, de loi initiale μ_0 et de matrice de transition $\left((1 - \delta_{x,y}) \frac{\mathcal{L}_{x,y}}{\mathcal{L}_x} \right)_{(x,y) \in \mathcal{S}^2}$, où $\delta_{x,y}$ est le symbole de Kronecker, qui vaut 1 lorsque $x = y$ et 0 sinon.
- $T_0 = 0$, et pour tout $n \in \mathbb{N}^*$, connaissant Y_0, \dots, Y_n les variable $T_1 - T_0, \dots, T_n - T_{n-1}$ sont indépendantes, de loi exponentielle de paramètres respectifs $\mathcal{L}_{Y_0}, \dots, \mathcal{L}_{Y_{n-1}}$.
- Pour tout $i \in \mathbb{N}$ et tout $t \in [T_i, T_{i+1}[$, $X_t = Y_i$.
- On définit $T := \lim_{n \rightarrow \infty} T_n$. Si $T < \infty$, on pose $X_t = \Delta$ pour tout $t \geq T$.

Ce processus est markovien, homogène, et ses trajectoires sont càdlàg. De plus, cette construction détermine de manière unique la loi de X . Dans la suite, on dira simplement que \mathcal{L} est un générateur sur \mathcal{S} , et que X est un *processus minimal* de générateur \mathcal{L} et de loi initiale μ_0 , ou seulement de générateur \mathcal{L} lorsque la loi initiale n'a pas d'importance. Le temps T est appelé *temps d'explosion* de X , et l'on verra dans la section 1.2 comment dans certains cas on peut définir un processus non-minimal, c'est-à-dire qui ne soit pas « tué » au temps d'explosion. Dans le cas où $T = \infty$ presque sûrement pour toute loi initiale, le générateur \mathcal{L} (ou le processus X) est dit non explosif. Il est dit récurrent positif (respectivement irréductible, réversible) si la chaîne $(Y_n)_{n \in \mathbb{N}}$ est récurrente positive (resp. irréductible, réversible).

Comme mentionné dans le cas général, la loi d'un tel processus est aussi caractérisée par ses lois fini-dimensionnelles, c'est-à-dire l'ensemble des

$$\mathbb{P}(X_{t_1} = x_1, \dots, X_{t_n} = x_n)$$

avec $n \in \mathbb{N}$, $0 \leq t_1 < \dots < t_n < \infty$ et $x_1, \dots, x_n \in \mathcal{S}$. La fonction de transition de X sera notée :

$$P_{x,y}(t) = \mathbb{P}(X_t = y \mid X_0 = x)$$

et constitue (associée à une loi initiale μ_0) une autre manière de caractériser la loi de X , par l'intermédiaire de ses lois fini-dimensionnelles. Une fonction de transition définit un semi-groupe $(P(t))_{t \geq 0}$ de matrices sous-stochastiques, qui peuvent être élargies sur $\mathcal{S} \cup \{\Delta\}$ de manière à être rendues stochastiques.

Une semi-groupe $P(\cdot)$ est une *solution des équations de Kolmogorov* associées à \mathcal{L} si pour tout $t \geq 0$ on a les égalités matricielles :

$$P'(t) = \mathcal{L}P(t) \quad (\text{équation backward}) \quad (4)$$

$$P'(t) = P(t)\mathcal{L} \quad (\text{équation forward}) \quad (5)$$

et c'est la solution minimale de l'équation backward (respectivement forward) si pour toute autre solution $\tilde{P}(\cdot)$ on a $P_{x,y}(t) \leq \tilde{P}_{x,y}(t)$, pour tout $x, y \in \mathcal{S}$ et tout $t \geq 0$. On rappelle que la solution minimale existe (théorème 2.8.3 et 2.8.6 de [69]), et est nécessairement unique par définition.

Chacune des trois conditions suivantes est impliquée par les deux autres :

- X est un processus minimal associé à \mathcal{L}
- X est un processus minimal de semi-groupe $P(\cdot)$
- $P(\cdot)$ est la solution minimale de l'équation backward (respectivement forward) associée à \mathcal{L}

Bien que le point de vue adopté dans le cas des chaîne de Markov soit celui que l'on vient de présenter, on rappelle également un résultat utile faisant le lien avec le point de vue fonctionnel. On identifie une fonction $f \in \mathcal{B}(\mathcal{S})$ avec un vecteur colonne indexé par \mathcal{S} , et dans ce cas $P(\cdot)$ et \mathcal{L} agissent sur $\mathcal{B}(\mathcal{S})$ par multiplication matricielle.

Si X est un processus minimal de générateur \mathcal{L} et de temps de saut τ_n , $n \in \mathbb{N}$, alors pour toute fonction $f \in \mathcal{B}(\mathcal{S})$ et tout $n \in \mathbb{N}$, le processus :

$$f(X_{t \wedge \tau_n}) - f(X_0) - \int_0^{t \wedge \tau_n} \mathcal{L}[f](X_s) ds$$

est une \mathcal{F}^X -martingale.

0.1.3 Convergence de mesures

Sous de bonnes conditions, un processus de Markov admettant une unique mesure invariante μ_∞ converge vers celle-ci, dans le sens (intuitif) où pour tout $x \in \mathcal{S}$ et tout $A \in \mathcal{B}(\mathcal{S})$:

$$P(t, x, A) \rightarrow \mu_\infty(A), \quad \text{quand } t \rightarrow \infty.$$

Pour cette raison, et pas seulement, nous avons besoin d'outils pour quantifier la convergence de mesures en général, et de mesures de probabilités en particulier. Pour cela, il existe de nombreux moyens, parmi lesquels nous nous intéresserons à la distance de Prokhorov, qui est la distance naturelle pour la convergence en loi de variables aléatoires, à la distance en variation totale, dont le nom est relativement explicite et à la séparation, moins répandue mais plus utile dans le cadre des temps forts de stationnarité.

On pourrait également citer, entre autres, la divergence de Kullback-Leibler, utilisée pour quantifier la convergence d'un processus de Markov vers sa loi stationnaire grâce à l'inégalité Sobolev logarithmique, et la distance du χ^2 , utilisée dans le même but avec l'inégalité de Poincaré. Ces derniers ne seront pas exposés ici, et nous renvoyons à [5] ou [50] pour plus de détails à leur sujet.

Distance de Prokhorov

La distance de Prokhorov (ou de Lévy-Prokhorov) a été introduite en 1956 par Iouri Prokhorov [73] pour généraliser la distance de Lévy ([53]). Si pour tout $\varepsilon > 0$ on note A^ε le ε -élargi d'un ensemble $A \in \mathcal{S}$:

$$A^\varepsilon = \{x \in \mathcal{S} : d(x, A) < \varepsilon\}$$

alors la distance de Prokhorov est définie par :

$$\forall \mu, \nu \in \mathcal{P}(\mathbb{G}), \quad \rho(\mu, \nu) := \inf\{\varepsilon > 0 : \mu(A) \leq \nu(A^\varepsilon) + \varepsilon, \forall A \in \mathcal{B}(\mathbb{G})\}$$

et engendre la topologie de la convergence faible, dans le sens où pour toute famille $(\mu_n)_{n \in \mathbb{N}} \subset \mathcal{P}(\mathcal{S})$ et $\mu \in \mathcal{P}(\mathcal{S})$, on a $\rho(\mu_n, \mu) \rightarrow 0$ si et seulement si :

$$\forall f \in \mathcal{B}(\mathcal{S}) \cap \mathcal{C}(\mathcal{S}), \quad \mu_n(f) \rightarrow \mu(f)$$

où de manière équivalente, si et seulement si toute suite de variables aléatoires associée à la suite $(\mu_n)_{n \in \mathbb{N}}$ converge en loi vers une variable aléatoire de loi μ (voir par exemple Billingsley [8] ou Ethier et Kurtz [29] pour plus de détails).

Cette distance ne sera pas vraiment utilisée dans cette thèse dans le contexte décrit en début de section (convergence d'un processus vers sa loi stationnaire), mais plutôt pour montrer la convergence en loi dans $D_{\mathcal{S}}(I)$ d'une suite de processus (chapitre 6).

Distance en variation totale

Si μ est une *mesure signée* sur \mathcal{S} , c'est-à-dire satisfaisant toutes les propriétés d'une mesure sauf éventuellement la positivité, on définit sa *partie positive* et *partie négative* par :

$$\begin{aligned} \mu^+ &:= \sup_{A \in \mathcal{S}} \mu(A) \geq 0 \\ \mu^- &:= \inf_{A \in \mathcal{S}} \mu(A) \leq 0 \end{aligned}$$

La *norme en variation totale* de μ est alors donnée par :

$$\|\mu\|_{VT} := \mu^+ - \mu^-$$

L'espace $\mathcal{M}(\mathcal{S})$ des mesures signées sur \mathcal{S} , muni de la norme en variation totale est un espace de Banach (voir Cohn [16] pour plus de détails).

La *distance en variation totale* sur $\mathcal{P}(\mathcal{S})$ est la distance induite par la norme de la variation totale sur $\mathcal{M}(\mathcal{S})$. Elle peut être caractérisée de plusieurs manières :

$$\begin{aligned} \forall \mu, \nu \in \mathcal{P}(\mathcal{S}), \quad d_{VT}(\mu, \nu) &= \sup_{A \in \mathcal{S}} |\mu(A) - \nu(A)| \\ &= \sup_{A \in \mathcal{S}} (\mu(A) - \nu(A)) \\ &= \sup_{A \in \mathcal{S}} (\nu(A) - \mu(A)) \\ &= \frac{1}{2} \int_{\mathcal{S}} |\mu(dx) - \nu(dx)| \end{aligned}$$

Sa compréhension à partir de la première caractérisation devrait être relativement intuitive.

Séparation

Soient μ et ν deux mesures de probabilité sur \mathcal{S} . Si μ est absolument continue par rapport à ν (*i.e.* pour tout $A \in \mathcal{S}$, $(\nu(A) = 0 \Rightarrow \mu(A) = 0)$), noté $\mu \ll \nu$, on définit la *séparation* entre μ et ν par :

$$\text{sep}(\mu, \nu) := \sup_{A \in \mathcal{S} : \nu(A) > 0} \left(1 - \frac{\mu(A)}{\nu(A)} \right)$$

Sinon, par le théorème de Lebesgue-Radon-Nikodym (voir Rudin [78]), il existe deux mesures positives μ_1 et μ_2 sur \mathcal{S} (appelées respectivement *partie absolument continue* et *partie singulière* de μ par rapport à ν), satisfaisant $\mu = \mu_1 + \mu_2$, $\mu_1 \ll \nu$ et $\mu_2 \perp \nu$ (i.e. $\exists A \in \mathcal{S}$ tel que $\mu_2(A) = 0$ et $\nu(A^c) = 0$), et dans ce cas on pose :

$$\text{sep}(\mu, \nu) := \text{sep}(\mu_1, \nu)$$

Remarque 0.1.2. Notons que cette définition admet une formulation équivalente en terme de densité, donnée par Miclo [65] :

$$\text{sep}(\mu, \nu) = \text{ess sup}_{\nu} \left(1 - \frac{d\mu_1}{d\nu} \right)$$

où $\frac{d\mu}{d\nu}$ désigne la dérivée de Radon-Nikodym de μ par rapport à ν , et ess sup_{ν} le supremum essentiel par rapport à la mesure ν :

$$\forall g \in \mathbb{L}^{\infty}(\nu), \quad \text{ess sup}_{\nu} g = \inf\{a \in \mathcal{S} : \nu(g(x) > a) = 0\}$$

Pour voir que ces formulations sont équivalentes, prenons le cas où $\mu \ll \nu$, et posons $\alpha = \text{ess inf}_{\nu} \frac{d\mu}{d\nu}$ et $\mathcal{N} = \{x \in \mathcal{S} : \frac{d\mu}{d\nu}(x) < \alpha\} \in \mathcal{S}$. On a donc $\nu(\mathcal{N}) = 0$ et pour tout $A \in \mathcal{S}$ tel que $\nu(A) > 0$:

$$\frac{\mu(A)}{\nu(A)} = \frac{1}{\nu(A)} \int_A \frac{d\mu}{d\nu}(x) \nu(dx) \geq \alpha$$

Inversement, pour tout $\varepsilon > 0$, il existe $A_{\varepsilon} \in \mathcal{S}$ tel que $\nu(A_{\varepsilon}) > 0$ et $\frac{d\mu}{d\nu} \leq \alpha + \varepsilon$ sur A_{ε} , donc $\frac{\mu(A_{\varepsilon})}{\nu(A_{\varepsilon})} \leq \alpha + \varepsilon$.

Dans le cas simple où \mathcal{S} est un espace dénombrable (ou si μ est absolument continue par rapport à une mesure de comptage finie sur \mathcal{S}), la séparation peut aussi s'écrire :

$$\text{sep}(\mu, \nu) = \sup_{x \in \mathcal{S} : \nu(\{x\}) > 0} \left(1 - \frac{\mu(\{x\})}{\nu(\{x\})} \right)$$

et c'est d'ailleurs ainsi qu'elle a été formulée initialement par Aldous et Diaconis [1].

Cet outil permet donc en quelque sorte de quantifier l'écart entre deux mesures, mais n'est pas symétrique et n'est donc pas une distance (on dit que c'est une *quasi-distance*, voir Steen et Seebach [80]). En revanche, c'est un majorant de la distance en variation totale comme le montre un rapide calcul :

$$\begin{aligned} \|\mu - \nu\|_{VT} &= \sup_{A \in \mathcal{S}} (\nu(A) - \mu(A)) \\ &= \sup_{A \in \mathcal{S} : \nu(A) > 0} \nu(A) \left(1 - \frac{\mu(A)}{\nu(A)} \right) \\ &\leq \sup_{A \in \mathcal{S} : \nu(A) > 0} \left(1 - \frac{\mu_1(A)}{\nu(A)} \right) = \text{sep}(\mu, \nu) \end{aligned}$$

et on a donc $0 \leq \text{sep}(\mu, \nu) \leq 1$.

La séparation est beaucoup plus sensible aux variations sur de petits intervalles que la distance en variation totale : par exemple, si $\mu = \mathcal{U}([0, 1 - \varepsilon])$ et $\nu = \mathcal{U}([\varepsilon, 1])$, $\varepsilon \in]0, 1[$, on a $\|\mu - \nu\|_{VT} = 2\varepsilon$, tandis que $\text{sep}(\mu, \nu) = \text{sep}(\nu, \mu) = 1$.

0.2 Temps forts de stationnarité et dualité

On rentre maintenant plus spécifiquement dans la théorie de la stationnarité forte, en rappelant les principales définitions et propriétés. Cette théorie a été formalisée par Aldous et Diaconis [2] pour les chaînes de Markov à temps discret, puis généralisée par Fill [31] en temps continu et par Miclo [65] et Lyzinski [58] pour les diffusions réelles ; nous l'énonçons ici dans le cadre général des processus de Markov ergodiques.

0.2.1 Définition d'un temps fort de stationnarité

On considère un ensemble mesurable $(\mathcal{S}, \mathcal{S})$, et un processus de Markov $(X_i)_{i \in I}$ sur \mathcal{S} , à temps discret ($I = \mathbb{N}$) ou continu ($I = \mathbb{R}_+$). On note μ_i sa loi au temps $i \in I$ et μ_∞ sa loi stationnaire, dont on suppose l'existence et l'unicité. On suppose également de manière assez classique que pour tout $i > 0$, μ_i est absolument continue par rapport à μ_∞ .

Un \mathcal{F}^X -temps d'arrêt randomisé T est appelé *temps fort* (*strong time*) si la loi de X_T ne dépend pas de T (lorsque ce dernier est fini). C'est un *temps fort de stationnarité* (*strong stationary time*) si de plus X est distribué selon sa loi stationnaire au temps T :

$$\mathcal{L}(X_T | T, T < \infty) = \mathcal{L}(X_T | T < \infty) = \mu_\infty \quad (6)$$

Si T est un temps fort et X_T est distribué selon $\nu \in \mathcal{P}(\mathcal{S})$, alors T est aussi appelé *future independent ν -time* (voir Thorisson [88]). Comme pour les temps d'arrêts, on dit que T est un temps fort de stationnarité fini si $T < +\infty$ presque sûrement. Si T est un temps fort de stationnarité et que X est fortement markovien en T (ce qui est souvent le cas en pratique), alors de manière assez immédiate, $T + i$ est aussi un temps fort de stationnarité pour tout $i \in I$.

Outre les temps forts de stationnarité, on peut mentionner d'autres types de temps forts (sur lesquels nous ne nous attarderons pas ici), le plus simple étant le temps d'atteinte d'un point : la loi du processus à ce moment est alors une dirac en ce point. Un autre exemple plus intéressant et plus proche de celui qui nous concerne est celui des temps forts de quasi-stationnarité pour un processus ayant un état absorbant, pour lesquels, comme son nom l'indique, la loi du processus arrêté est sa loi quasi-stationnaire. Ce dernier exemple de temps fort a notamment été utilisé par Diaconis et Miclo [24] pour calculer de manière probabiliste le temps d'atteinte de $N \in \mathbb{N}$ pour un processus de vie et de mort sur \mathbb{N} partant de 0.

Comme observé par Diaconis et Fill [23], les temps forts permettent de majorer la distance en variation totale entre un processus au temps $i \in I$ et sa loi stationnaire, puisque pour tout temps fort T on a :

$$\| \mu_i - \mu_\infty \|_{VT} \leq \| \mathcal{L}(X_T) - \mu_\infty \|_{VT} + (1 - \| \mathcal{L}(X_T) - \mu_\infty \|_{VT}) \mathbb{P}(T > i) \quad (7)$$

et cette majoration gagne encore en intérêt si T est un temps fort de stationnarité, puisqu'on a dans ce cas une simplification de l'inégalité précédente :

$$\| \mu_i - \mu_\infty \|_{VT} \leq \mathbb{P}(T > i) \quad (8)$$

Cette quantification de la convergence peut être raffinée grâce à la notion de séparation définie plus haut. On notera par $\mathfrak{s}(t)$ la séparation entre la loi de X au temps $i \in I$ et sa loi stationnaire, la dépendance en X étant implicite :

$$\mathfrak{s}(i) := \text{sep}(\mu_i, \mu_\infty)$$

et on a la proposition suivante :

Proposition 0.2.1. (a) Si T est un temps fort de stationnarité et X satisfait la propriété de Markov forte en T , alors on a :

$$\mathfrak{s}(t) \leq \mathbb{P}(T > t) \tag{9}$$

(b) Pour tout $\varepsilon > 0$, il existe un temps fort de stationnarité T tel que :

$$\mathbb{P}(T > t) \leq \mathfrak{s}(t) + \varepsilon$$

(c) Si \mathcal{S} est fini, alors il existe un temps fort de stationnarité T tel que :

$$\mathbb{P}(T > t) = \mathfrak{s}(t)$$

T est alors appelé temps à la stationnarité

Démonstration. Voir les calculs de [2] et [32], l'adaptation au cas général étant immédiate. \square

Notons que l'inégalité (8) peut être vue comme une implication immédiate de (9).

Il existe un parallèle intéressant avec les temps de couplage : un temps d'arrêt T est un *temps de couplage* pour X s'il existe un processus Y , de même loi d'évolution que X , tel que $Y_0 \sim \mu_\infty$ et $T = \inf\{i \in I : X_i = Y_i\}$ presque sûrement. Alors dans les points (a) et (c) de la proposition précédente, on peut remplacer la séparation par la distance en variation totale et « temps fort de stationnarité » par « temps de couplage » (voir Griffeath [42]).

0.2.2 Dualité

En pratique, pour construire un temps fort de stationnarité, une notion qui s'avère souvent utile (et sur laquelle nous nous focaliserons dans cette thèse) est celle de dual de stationnarité forte (*strong stationary dual*), introduite par Diaconis et Fill [23] pour des chaînes de Markov à temps discret et étendue au cas du temps continu par Fill [32], section 2.1, et par Miclo [65] pour des diffusions.

On considère maintenant un deuxième espace mesurable $(\mathcal{S}^*, \mathcal{S}^*)$. Un processus $X^* = (X_i^*)_{i \in I}$ sur \mathcal{S}^* , ayant un état absorbant que l'on notera $\infty \in \mathcal{S}^*$, est un *dual de stationnarité forte* pour X si les deux conditions suivantes sont satisfaites pour tout $i \in I$:

$$\mathcal{L}(X_i^* | \mathcal{F}_\infty^X) = \mathcal{L}(X_i^* | \mathcal{F}_i^X) \tag{10}$$

$$\mathcal{L}(X_i | \mathcal{F}_i^{X^*}) = \mu_\infty \quad \text{p.s. sur l'ensemble } \{X_i^* = \infty\} \tag{11}$$

La relation (10) exprime le fait que X^* est indépendant du futur de X connaissant son passé, et implique notamment que le temps d'absorption de X^* est un temps d'arrêt randomisé pour X (voir aussi remarque ci-après). Sous cette hypothèse, la condition (11) signifie alors que le temps d'absorption est un temps fort de stationnarité. Réciproquement, à partir d'un temps fort de stationnarité, on peut construire un dual de stationnarité forte de manière canonique (voir Fill [32], théorème 1). Lorsque le temps d'absorption du dual est un temps à la stationnarité, on dit que le dual est *fin* (anglais : *sharp*). Il ne sera pas question d'autre type de dualité dans cette thèse, et donc « processus dual » ou simplement « dual » signifiera toujours « dual de stationnarité forte ».

Remarque 0.2.2. La condition (10) est exprimée par Fill [32] (2.1) sous la forme suivante : il existe une sous-tribu \mathcal{G} de \mathcal{A} , indépendante de \mathcal{F}_∞^X , telle que :

$$\forall i \in I, \quad X_i^* \text{ est } \sigma(\mathcal{G}, \mathcal{F}_i^X)\text{-mesurable} \quad (10')$$

Cette formulation met en évidence le fait que la trajectoire $X_{[0,i]}^*$, $i \in I$, peut être générée à partir de $X_{[0,i]}$ et d'un aléa supplémentaire indépendant de X .

Elle est en fait équivalente à (10). En effet l'implication (10') \Rightarrow (10) est relativement immédiate, tandis que l'implication réciproque découle de l'argument suivant : on pose

$$\mathcal{G}_i = \sigma(\{f(X_i^*) - \mathbb{E}(f(X_i^*) | \mathcal{F}_i^X), f \in \mathcal{B}(\mathcal{S}^*)\})$$

et sous l'hypothèse (10) on a alors que la tribu $\mathcal{G} := \bigcup_{i \in I} \mathcal{G}_i$ est indépendante de \mathcal{F}_∞^X , et pour tout $f \in \mathcal{B}(\mathcal{S}^*)$:

$$\begin{aligned} \mathbb{E}(f(X_i^*) | \mathcal{F}_i^X) &\text{ est } \mathcal{F}_i^X\text{-mesurable} \\ f(X_i^*) - \mathbb{E}(f(X_i^*) | \mathcal{F}_i^X) &\text{ est } \mathcal{G}\text{-mesurable} \end{aligned}$$

Donc $f(X_i^*)$ est $\sigma(\mathcal{G}, \mathcal{F}_i^X)$ -mesurable pour tout $f \in \mathcal{B}(\mathcal{S})$ ce qui implique que X_i^* aussi.

0.2.3 Entrelacement

Soit Λ un noyau de transition markovien de \mathcal{S}^* vers \mathcal{S} , c'est-à-dire :

$$\begin{aligned} \forall x^* \in \mathcal{S}^*, \quad \Lambda(x^*, \cdot) &\in \mathcal{P}(\mathcal{S}) \\ \forall A \in \mathcal{S}, \quad \mathcal{S}^* \ni x^* \mapsto \Lambda(x^*, A) &\text{ est } \mathcal{S}^*\text{-mesurable} \end{aligned}$$

et soit X^* un processus de Markov à valeurs dans \mathcal{S}^* . On dit que X et X^* sont Λ -liés ou entrelacés par Λ si le couple (X, X^*) est markovien et :

$$\forall i \in I \quad \mathcal{L}(X_i | \mathcal{F}_i^{X^*}) = \Lambda(X_i^*, \cdot) \text{ p.s.} \quad (12)$$

Quitte à rajouter un point isolé à \mathcal{S}^* , inaccessible pour X^* , on peut toujours supposer que ce dernier possède un état absorbant ∞ satisfaisant :

$$\Lambda(\infty, \cdot) = \mu_\infty \quad (13)$$

Nous ferons donc cette supposition, et en conséquence, (12) implique (11).

Nous désignerons donc par **H** l'hypothèse « il existe un couplage (X, X^*) satisfaisant (10) et (12) ». Cette formulation est légèrement abusive, et signifie en réalité qu'il existe un couple (Y, Y^*) , dont les marginales ont même loi que X et X^* respectivement, et satisfaisant ces conditions. Dans ce cas, X^* (ou plus exactement sa copie Y^*) est un dual de stationnarité forte Λ -lié pour X , que l'on abrégera en *dual Λ -lié*. Un tel couplage fournit donc un temps fort de stationnarité par l'intermédiaire du temps d'absorption de X^* (fini ou non), mais possède également des propriétés supplémentaires intéressantes pour l'étude de la loi de X .

En prenant $i = 0$ dans (12), on observe que sous l'hypothèse **H** on a :

$$\forall A \in \mathcal{S}, \quad \mu_0(A) = \mathbb{P}(X_0 \in A) = \int_{\mathcal{S}^*} \Lambda(x^*, A) \mathbb{P}(X_0^* \in dx^*)$$

Si on note μ_0^* la loi de X_0^* , on trouve donc que :

$$\mu_0 = \mu_0^* \Lambda \quad (14)$$

est une condition nécessaire à l'entrelacement de X et X^* , que l'on appellera *entrelacement des lois initiales*. Insistons dès à présent sur le fait que cette égalité ne concerne que des lois et ne garantit donc pas l'égalité (12) pour $i = 0$, qui est une égalité presque sûre. Cependant, sous cette condition, il sera toujours possible de coupler X et X^* de manière à ce qu'ils vérifient (12) au temps $i = 0$.

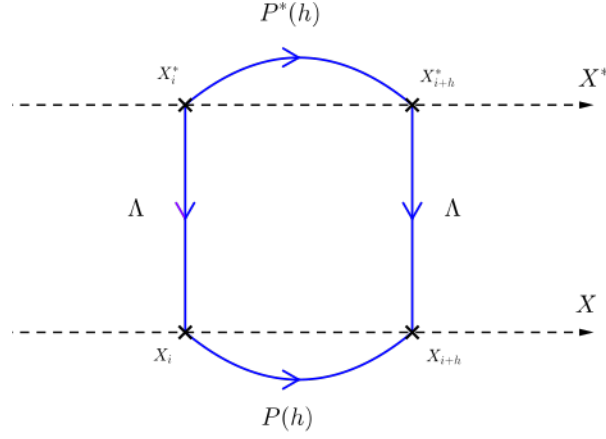


FIGURE 1 – Transitions de la chaîne (X, X^*)

A partir de maintenant, nous supposons donc que μ_0^* et μ_0 vérifient (14). Pour obtenir l'entrelacement complet des processus c'est un bon point de départ, mais il est ensuite nécessaire de prendre en compte les semi-groupes $P(\cdot)$ et $P^*(\cdot)$ de X et X^* respectivement. Comme pour les lois initiales, on dit que P et P^* sont *entrelacés* (ou *algébriquement duaux* selon la terminologie de Diaconis et Fill) s'ils vérifient :

$$P^*(h)\Lambda = \Lambda P(h), \quad \forall h \in I \quad (15)$$

La nécessité de cette condition n'est pas difficile à vérifier par le calcul si l'on raisonne en termes de fonctions de transitions associées. La figure 1 devrait suffire à s'en convaincre : on passe de X_i^* à X_{i+h}^* par la fonction de transition P^* , puis à X_{i+h} par le noyau de transition Λ , ou de X_i^* à X_i par Λ et de X_i à X_{i+h} par P . La relation (15) peut alors s'interpréter comme une condition de commutativité du diagramme. Pour finir avec cette égalité, notons que si son sens général est celui d'une égalité d'opérateurs sur $\mathcal{B}(\mathcal{S})$, elle peut aussi s'interpréter comme une égalité matricielle lorsque \mathcal{S} est dénombrable.

En dehors du cadre de la stationnarité forte, on peut mentionner le papier de Pitman et Rogers [76] qui utilise l'entrelacement de semi-groupes pour donner une condition de markoviannité de $f(X)$ lorsque f est une fonction de \mathcal{S} dans un autre ensemble \mathcal{S}' . D'autres exemples ont également été étudiés notamment par Dynkin [28] ou Yor et al. [92], [13]. Enfin, une autre notion d'entrelacement, proche de celle-ci, est obtenue en remplaçant Λ par un opérateur différentiel, et est étudiée notamment par Chafaï et Joulin [14] ou Cloez et Delplancke [15].

En temps discret, si la condition (15) est remplie pour $h = 1$, alors une simple récurrence suffit pour vérifier qu'elle l'est pour tout $h \in \mathbb{N}$. Elle est alors suffisante pour affirmer qu'il existe un couplage (X, X^*) tel que X et X^* soient entrelacés par Λ . En fait, on a même un résultat plus fort que ça : Diaconis et Fill [23] ont explicité une méthode pour construire, X^* à partir de X déjà construit, de sorte que X et X^* soient entrelacés par Λ .

En revanche, en temps continu on ne peut pas se contenter d'une relation sur les semi-groupes. En effet, à quelques exceptions près, ceux-ci ne sont pas connus explicitement et on préfère généralement travailler avec les générateurs. Heuristiquement, les générateurs \mathcal{L} et \mathcal{L}^* de X et X^* correspondent aux dérivées au temps 0 des semi-groupes associées, et en appliquant ceci à (15), il apparaît alors naturel d'étudier la condition d'entrelacement pour les générateurs :

$$\mathcal{L}^*\Lambda = \Lambda\mathcal{L}$$

Contrairement à l'entrelacement des semi-groupes, cette égalité peut dans certains cas se comprendre de deux manières *a priori* pas tout à fait équivalentes. Naturellement, on

peut la considérer comme une égalité d'opérateurs sur $\mathcal{D}(\mathcal{L})$. Pour avoir un sens, cette égalité sous-entend alors que pour tout $f \in \mathcal{D}(\mathcal{L})$, on a $\Lambda f \in \mathcal{D}(\mathcal{L}^*)$. D'autre part, si les espaces d'état sont dénombrables, on peut aussi la voir comme une égalité matricielle, ce que nous ferons. On reformule alors l'entrelacement en fonction de son interprétation fonctionnelle ou matricielle :

$$\forall f \in \mathcal{D}(\mathcal{L}), \quad \begin{cases} \Lambda f \in \mathcal{D}(\mathcal{L}^*) \\ \mathcal{L}^*[\Lambda f] = \Lambda[\mathcal{L}f] \end{cases} \quad (16)$$

$$\forall x^* \in \mathcal{S}^*, y \in \mathcal{S}, \quad \sum_{y^* \in \mathcal{S}^*} \mathcal{L}_{x^*, y^*}^* \Lambda(y^*, y) = \sum_{x \in \mathcal{S}} \Lambda(x^*, x) \mathcal{L}_{x, y} \quad (16m)$$

Notons que (16) implique (16m).

On a essayé de synthétiser dans la proposition suivante les relations entre \mathbf{H} , (15) et (16)/(16m) dans différents contextes.

Proposition 0.2.3. *Si $I = \mathbb{R}_+$ et $\mu_0 = \mu_0^* \Lambda$, alors :*

(a) *Comme observé précédemment, $\mathbf{H} \Rightarrow (15)$.*

(b) *Miclo [65] a montré en utilisant une approximation par des chaînes de Markov à temps discret (proposition 24) que :*

$$(15) \Rightarrow \mathbf{H}$$

Même si la proposition concerne le cadre spécifique des diffusions réelles, la démonstration est valable en toute généralité.

(c) *Lyzinski [58] a montré que si X et X^* sont deux diffusions sur un segment, alors :*

$$(15) \Leftrightarrow (16)$$

(d) *Une partie des calculs pour montrer l'équivalence précédente reste en fait valable dans le cas général, et on a :*

$$(15) \Leftarrow (16)$$

(e) *Si X et X^* sont deux chaînes de Markov à temps continu, non-explosives, et pour tout $x^* \neq \infty \in \mathcal{S}^*$, l'ensemble $\{x \in \mathcal{S} : \Lambda(x^*, x) > 0\}$ est fini, alors Fill [32] a montré que :*

$$\mathbf{H} \Leftrightarrow (15) \Leftrightarrow (16m)$$

(f) *Nous verrons (théorème 1.2.7) que ces trois conditions restent équivalentes dans un contexte un peu plus général.*

0.3 Historique

0.3.1 Développements théoriques

Les temps forts de stationnarité apparaissent pour la première fois en 1986 dans l'article de Aldous et Diaconis [1]. Ils sont à l'origine introduits pour étudier des exemples du phénomène de cut-off (voir section suivante) pour différentes méthodes de mélange de cartes, modélisées par des marches aléatoires sur des groupes. Dans ce contexte, par le concept même de mélange d'un paquet de carte, la loi stationnaire à atteindre est la loi uniforme. Les temps forts de stationnarité sont donc initialement nommés temps forts uniformes (*strong uniform times*), et plusieurs exemples en sont présentés dans l'optique

de répondre à la question « au bout de combien de temps le paquet est-il suffisamment mélangé ? ».

Une étude plus approfondie des temps forts uniformes pour des marches aléatoires sur des groupes et publiée un an plus tard, toujours par Aldous et Diaconis [2], et développe notamment le lien avec la séparation, le parallèle avec la méthode de couplage et la notion d'optimalité pour un temps fort uniforme, (qui donnera naissance aux temps à la stationnarité). Ils y démontrent également l'inégalité (7) (sous le nom de *early stopping theorem*).

Dans le même contexte, de nombreux exemples sont ensuite étudiés par Diaconis [20] et Matthews [60], [61], [62] entre 1987 et 1989.

Les notions de dual de stationnarité forte et de dual Λ -lié font leur apparition en 1990 dans le papier de Diaconis et Fill [23], qui établissent par la même occasion la terminologie « temps fort de stationnarité ». Une méthode systématique de construction utilisant ces notions de dualité y est proposée, pour des chaînes de Markov à temps discret et espace d'état fini.

Une version de ce travail sur un espace d'état dénombrable est présentée dans le même temps par ces mêmes auteurs [22]. Dans ce cadre, l'état absorbant du processus dual est inaccessible en temps fini et ne peut donc servir à produire un temps fort de stationnarité (fini). L'intérêt réside alors dans l'utilisation de l'entrelacement par un noyau bien choisi (voir (17)), combinée avec le *early stopping theorem* de Aldous et Diaconis, qui permet de quantifier efficacement la convergence d'un processus.

Cette théorie a été étendue aux chaînes de Markov à temps continu par Fill [31], [32]. Dans le premier article, en 1991, il transcrit la partie concernant les temps forts de stationnarité, et montre que la queue de distribution de ceux-ci majore la séparation, ainsi que l'existence d'un temps à la stationnarité lorsque l'espace d'état est fini. Dans le deuxième, en 1992, il définit les notions de dualité et montre l'équivalence entre l'entrelacement de deux processus et celui de leurs semi-groupes ou de leurs générateurs respectifs. Le fait de considérer des processus à temps continu permet de rendre accessible un état absorbant qui serait inaccessible en temps discret, et donc de construire des temps forts de stationnarité finis sur des espaces infinis (ce que l'on exploitera dans cette thèse).

Plus récemment, la théorie de la stationnarité forte a été appliquée à des diffusions sur un intervalle réel par Lyzinski dans sa thèse [58] en 2012, et par Miclo [65] en 2017. La principale différence entre les deux réside dans le fait que Lyzinski considère une diffusion avec une certaine loi initiale, et un dual qui est aussi un processus de diffusion sur \mathbb{R} , tandis que Miclo considère une loi initiale quelconque et construit un dual dont l'espace d'état est l'ensemble des intervalles réels. Il en déduit également une condition pour que la diffusion primale admette un temps fort de stationnarité fini.

La démarche de Lyzinski consistant à trouver un dual ayant même espace d'état que le processus primal est à relier avec celle déjà évoquée par Diaconis et Fill [23] dans l'article mentionné plus haut. En effet, ceux-ci ont montré qu'une chaîne de Markov sur un espace fini totalement ordonné admet un dual sur ce même espace si et seulement si sa matrice de transition renversée en temps est stochastiquement monotone (voir par exemple Van Doorn [89] pour plus d'informations sur cette notion) et le rapport de sa loi initiale et de sa loi stationnaire est décroissant en tant que fonction de l'espace d'état. Ce même résultat apparaît également dans le cas dénombrable ([22]), et a ensuite été étendu en temps continu ([32]).

En 2012, Lorek et Szekli [55] élargissent encore le cadre de ce résultat, en considérant cette fois un espace fini partiellement ordonné, et un processus en temps discret. Ils prouvent alors qu'en remplaçant la monotonie et la monotonie stochastique dans les conditions précédentes par la Möbius-monotonie (voir par exemple Rota [77]), il existe un unique dual sur le même espace, et montre comment le construire. Ils appliquent ce

résultat à une marche aléatoire sur les sommets d'un cube et donnent également plusieurs exemples illustratifs dans un second papier [56].

Enfin, en 2016, Mao et Zhao [59] donnent une condition équivalente à l'existence de ce genre de dual pour une chaîne de Markov sur \mathbb{N}^d , toujours en terme de Möbius-monotonie, et le construisent également.

0.3.2 Exemples d'applications

Nous présentons ici trois exemples d'applications particuliers, que sont le phénomène de cut-off, la loi des temps de passage d'un état à un autre, et l'échantillonnage parfait. De manière plus générale, et comme mentionné dans la partie précédente, l'utilisation d'un dual Λ -lié pour un bon noyau Λ permet, *via* le early stopping theorem, de majorer la séparation (et donc aussi la distance en variation totale) entre la loi instantanée et la loi stationnaire. Cette technique a été l'une des premières applications par les auteurs de cette théorie (voir *early stopping bounds* dans l'article de Diaconis et Fill [23]), et a également été utilisée, outre les exemples cités ci-après, par Fill et Kahn [36] pour comparer la vitesse de mélange de plusieurs processus (utile notamment pour choisir le plus rapide dans la méthode MCMC).

Phénomène de cut-off

Lors de l'étude de la convergence d'un processus de Markov vers sa loi stationnaire, on s'aperçoit parfois que celle-ci se fait de manière brutale, dans le sens où la distance en variation totale entre la loi instantanée du processus et sa loi stationnaire stagne un certain temps autour de 1 avant de décroître exponentiellement à partir d'un certain moment. Ce phénomène a été observé en 1981 par Diaconis et Shahshahani [26] pour un modèle de mélange de paquet de cartes par inversion de la position de deux d'entre elles à chaque étape du mélange. En 1986, il a été étudié par Aldous et Diaconis [1] pour d'autres modèles de mélange et baptisé *cut-off phenomenon*. Ce fut donc la première application des temps forts de stationnarité, puisque ceux-ci ont été introduits à cette occasion.

Une définition plus précise du cut-off (ainsi que plusieurs exemples et contre-exemples) est proposée par Diaconis [21] en 1996, et porte non pas sur un processus, mais sur une suite de processus. Nous donnons ici une version simplifiée de celle présentée par Diaconis et Saloff-Coste [25] en 2006, plus générale car n'utilisant pas nécessairement la distance en variation totale pour quantifier la convergence.

Soient \mathcal{S} un espace dénombrable, $D : \mathcal{P}(\mathcal{S}) \times \mathcal{P}(\mathcal{S}) \rightarrow [0, 1]$ une application satisfaisant la propriété de séparation ($D(\mu, \nu) = 0$ si et seulement si $\mu = \nu$) et telle que $\sup D = 1$ (par exemple, la séparation ou la distance en variation totale) et une famille $(\mathcal{S}_n, \nu_n, (\mu_{k,n})_{k \in \mathbb{N}})_{n \in \mathbb{N}}$ telle que pour tout $k, n \in \mathbb{N}$, $\mathcal{S}_n \subset \mathcal{S}$ est fini et $\nu_n(\mathcal{S}_n) = \mu_{k,n}(\mathcal{S}_n) = 1$. On dit que cette famille présente un *D-cut-off* s'il existe une suite $(t_n)_{n \in \mathbb{N}} \subset \mathbb{R}_+$ telle que pour tout $\varepsilon \in]0, 1[$:

- $\lim_{n \rightarrow \infty} D(\mu_{k_n, n}, \nu_n) = 0$ si $k_n > (1 + \varepsilon)t_n$ pour n assez grand
- $\lim_{n \rightarrow \infty} D(\mu_{k_n, n}, \nu_n) = 1$ si $k_n < (1 - \varepsilon)t_n$ pour n assez grand

Dans la famille considérée, la suite $(\mu_{k,n})_{k \in \mathbb{N}}$, $n \in \mathbb{N}$ correspond en général à la suite des loi instantanées d'un processus X_n , de loi stationnaire ν_n et d'espace d'état \mathcal{S}_n , et on dira alors simplement que la suite $(X_n)_{n \in \mathbb{N}}$ présente un *D-cut-off*. Nous n'avons à ce jour pas connaissance d'une théorie équivalente en temps continu.

Outre la première utilisation qui en a été faite par Aldous et Diaconis, Matthews [61] a également utilisé les temps forts de stationnarité pour étudier la convergence d'une suite de permutations générée par une suite de transpositions, et montrer l'existence d'un cut-off (pour la distance en variation totale). Diaconis et Saloff-Coste [25] ont eux utilisé un

dual de stationnarité forte pour donner une condition nécessaire et suffisante à l'existence d'un sep-cut-off pour une famille de processus de vie et mort (sep désignant, on le rappelle, la séparation). En 2017, Nestoridis [68] a également étudié le sep-cut-off, pour une marche aléatoire sur les chambres d'un arrangement d'hyperplans, en construisant un temps à la stationnarité pour cette chaîne de Markov.

Temps de passage

Étant donné un espace dénombrable \mathcal{S} , deux états $i, j \in \mathcal{S}$ et une chaîne de Markov X sur \mathcal{S} , partant de i , le temps de passage $T_{i,j}$ de i à j est définie comme la loi du premier temps d'atteinte de l'élément j . Si la loi des temps de passage reste, dans un cadre général, un problème ouvert, elle est en revanche parfaitement connue dans certains cas particuliers, et peut notamment être étudiée à l'aide des outils liés à la stationnarité forte. À notre connaissance, les études utilisant ces outils ont jusqu'à présent porté sur des cas où l'espace d'état est un intervalle de \mathbb{Z} .

Si l'on considère un certain $N \in \mathbb{N}$, et si X est un processus irréductible de vie et mort (à temps continu) sur \mathbb{N} , (ou sur un intervalle $[[0, m]]$ contenant N), de Q-matrice \mathcal{L} , alors un résultat remontant au moins à 1959 (Karlin et McGregor [45]) énonce que $T_{0,N}$ est égal en loi à une somme de variable aléatoires exponentielles indépendantes, dont les paramètres sont les valeurs propres non nulles de la matrice $(\mathcal{L}_{x,y} \mathbf{1}_{x < N})_{x,y \in \mathbb{N}}$ (c'est-à-dire, la Q-matrice du processus X absorbé en N). Dans le cas discret, si les valeurs propres de la matrice de transition sont positives, alors $T_{0,N}$ est égal en loi à une somme de variables aléatoires géométriques indépendantes, dont les paramètres sont les valeurs propres différentes de 1. En 2009, Diaconis et Miclo [24] ont donné une preuve probabiliste de ce résultat en temps continu, faisant intervenir la construction de duaux de stationnarité forte, de même que Fill [35] pour les deux résultats.

Dans un second papier [34] la même année, Fill a montré dans le même esprit un résultat plus général dû à Brown et Shao [12], où le processus de vie et mort (à temps discret ou continu) est remplacé par un processus pouvant « monter » d'une unité, ou « descendre » sans restriction. Le résultat dans ce cadre donne une expression de la transformée de Laplace de $T_{0,N}$ (dans le cas continu) ou de sa fonction génératrice des probabilités (dans le cas discret).

Enfin, en 2012, Gong, Mao et Zhang [41] ont étendu le premier résultat en temps continu en considérant n'importe quel temps de passage $T_{i,n}$, $i, n \in \mathbb{N}$, pour un processus de vie et mort sur \mathbb{N} pour lequel l'état 0 n'est pas forcément conservatif (autrement dit, il existe un état absorbant -1 , et $\mathcal{L}_{0,-1} \geq 0$). Cette fois également, la caractérisation de la loi du temps de passage est donnée par sa transformée de Laplace, et est obtenue par l'intermédiaire de duaux Λ -liés.

Échantillonnage parfait

Une application classique des processus (et en particulier des chaînes) de Markov réside dans les algorithmes stochastiques, et les temps forts de stationnarité y trouvent eux-mêmes une utilité naturelle. Parmi les algorithmes stochastiques, la méthode de Monte-Carlo par chaînes de Markov (MCMC, pour *Markov chain Monte-Carlo*) consiste à utiliser une chaîne de Markov bien choisie pour échantillonner sa loi stationnaire, en la laissant converger « suffisamment longtemps ». C'est ce qu'on appelle une *méthode d'échantillonnage*, en ce qu'elle permet de simuler une variable aléatoire selon une certaine loi. Bien que cette méthode soit généralement suffisante en pratique, on peut cependant lui opposer un argument théorique qui est que la convergence d'une chaîne de Markov ne garantit pas (en général) l'atteinte de sa loi stationnaire en un temps (déterministe) fini ; hors, on peut

difficilement se permettre de la laisser converger à l’infini, et la loi obtenue en temps fini est donc différente, bien qu’arbitrairement proche, de la loi ciblée.

On voit alors apparaître l’idée d’utiliser des temps aléatoires au lieu de temps déterministes. En 1996, Propp et Wilson [74] exploitèrent le principe des temps de couplage pour construire un algorithme en ce sens. Lorsque l’algorithme s’arrête, on obtient donc une réalisation exacte de la loi à simuler, et ce genre de méthode est appelé échantillonnage parfait (ou *perfect sampling*). L’inconvénient évident de ce genre de méthode est qu’on n’a pas de garantie sur le temps de calcul.

En 1998, Fill [33] propose un algorithme dans le même esprit, s’arrêtant cette fois lorsqu’un temps fort de stationnarité minimal (ou temps à la stationnarité) est atteint. Par rapport à l’algorithme de Propp-Wilson, celui de Fill présente l’avantage d’être « interruptible », en ce qu’une interruption précoce de celui-ci ne biaise pas le résultat (on obtient dans ce cas une réalisation d’une loi proche de la loi stationnaire). Cette algorithme a connu d’autres développements par la suite, notamment par Møller et Schladitz [66], Thönnies [87] ou Fill, Machida, Murdoch et Rosenthal [38].

On pourra également mentionner une utilisation récente de ce principe pour la création d’algorithme de cryptographie, par Lorek, Zagorski et Kulis en 2017 [57].

0.4 Présentation des travaux

Cette thèse est divisée en deux grandes parties. La première comprend les chapitres 1, 2 et 3 et traite de chaînes de Markov à temps continu. Elle peut être considérée comme un prolongement des travaux de Fill dans ce même cadre ([31],[32]). Les chapitres 1 et 2 ont fait l’objet d’une publication [17]. Dans la deuxième partie, on s’intéresse à une diffusion sur un graphe quantique ainsi qu’à un lien, dans le chapitre 6, avec les chaînes de Markov de la première partie, via une approximation discrète. Le chapitre 4 est le fruit d’un travail mené conjointement avec Laurent Miclo. Cette deuxième partie peut être reliée aux travaux de Miclo ou Lyzinsky sur des diffusions réelles ([58],[65]) ou ceux de Miclo sur des graphes ([64]) ou des variétés riemanniennes ([18]).

Partie I : Stationnarité pour une chaîne de Markov

Une question qui se pose naturellement est : existe-t-il toujours un temps fort de stationnarité fini ? Diaconis et Fill [23] ont prouvé que la réponse est oui lorsque l’espace d’état est fini, mais malheureusement cela n’est plus vrai sur un espace infini, même dénombrable. Bien sûr, l’existence ou non d’un temps fort de stationnarité fini dépend de la distribution initiale : lorsque celle-ci est la même que la loi stationnaire, alors 0 est un temps fort de stationnarité ! On peut alors se demander à quelle condition cette existence est possible, pour un processus donné mais pour toute loi initiale. Miclo [65] a ainsi exhibé ces conditions d’existence dans le cadre des diffusions sur \mathbb{R} . Le but ici est de les donner pour des marches aléatoires à temps continu sur une certaine classe de graphes discrets.

Cette partie s’articule donc autour d’un objectif principal : étant donné un générateur \mathcal{L} sur un espace dénombrable \mathcal{S} , admettant une unique mesure invariante μ , donner une condition nécessaire et suffisante sur ses coefficients pour que tout processus associé admette un temps fort de stationnarité fini. La méthode adoptée consiste à choisir un second espace dénombrable \mathcal{S}^* et un noyau de transition Λ de \mathcal{S}^* vers \mathcal{S} , puis à construire, pour un processus X donné (de générateur \mathcal{L} et de loi initiale μ_0), un dual Λ -lié X^* , et à étudier le temps d’absorption de ce dernier.

Un exemple de choix naturel et pratique est celui qui consiste à prendre \mathcal{S}^* comme une sous-partie de $\mathcal{P}(\mathcal{S})$, et Λ le noyau de transition défini par :

$$\forall Q \in \mathcal{P}(\mathcal{S}), \quad \Lambda(Q, \cdot) = \frac{\mu(Q \cap \cdot)}{\mu(Q)} \quad (17)$$

On veut alors définir un processus X^* sur \mathcal{S}^* , ayant \mathcal{S} comme état absorbant et Λ -lié avec X . Cependant, dans les exemples étudiés, le dual ne peut s'agrandir que d'un nombre fini de points à chaque saut. L'objectif étant d'atteindre l'état absorbant \mathcal{S} en temps fini, cela passe donc par une infinité de sauts en temps fini, autrement dit une explosion. On ne peut donc pas se placer dans le cadre des *General settings* de Fill (correspondant aux hypothèses de la proposition 0.2.3 (e)). De plus, on aimerait aussi pouvoir autoriser X^* à prendre comme valeurs des parties infinies de \mathcal{S} autre que \mathcal{S} lui-même, ce qui n'est pas non plus possible dans ce contexte, à cause de la définition de Λ .

Chapitre 1

Dans ce chapitre, on met en place les outils indispensables pour s'affranchir de certaines contraintes imposées par les *General settings*. En effet, on a besoin de savoir à quelle condition un processus X^* sur \mathcal{S}^* est un dual Λ -lié de X , et il s'agit donc au minimum de généraliser l'implication (16m) \Rightarrow H.

On supposera naturellement que le générateur \mathcal{L} est non-explosif, et l'on rajoute une condition de Λ -intégrabilité de sa diagonale :

$$\forall x^* \in \mathcal{S}^*, \quad \Lambda(x^*, x)\mathcal{L}_x < +\infty$$

Cette dernière condition paraît raisonnable : d'une part, elle affaiblit un peu la contrainte par rapport à l'hypothèse de finitude des ensembles $\{x \in \mathcal{S} : \Lambda(x^*, x) > 0\}$, $x^* \in \mathcal{S}^*$, d'autre part elle équivaut à la condition que le produit (matriciel) $\Lambda\mathcal{L}f$ est bien défini pour toute fonction $f \in \mathcal{B}(\mathcal{S})$. En revanche, on ne fait pas encore d'hypothèse particulière sur le choix de \mathcal{S}^* et Λ .

Sous ces hypothèses, on considère un générateur \mathcal{L}^* et une mesure de probabilités $\mu_0^* \in \mathcal{P}(\mathcal{S}^*)$, entrelacés avec \mathcal{L} et μ_0 respectivement. Dans la section 1.1, on montre qu'il existe un processus minimal X^* associé à \mathcal{L}^* et μ_0^* , qui est un dual Λ -lié pour X jusqu'à son premier temps d'explosion. Ce résultat constitue déjà une première généralisation, puisque si X^* est non-explosif on a bien l'implication (16) \Rightarrow H dans un cadre légèrement plus général que celui dans lequel elle était valable jusqu'à présent.

Dans la section 1.2 on commence par introduire la notion de *processus stratifié*, qui apparaît comme un moyen naturel de prolonger un processus explosif admettant une limite en son temps d'explosion. Concrètement, si Y est une chaîne de Markov explosive et τ son temps d'explosion, on définit $Y_\tau = \lim_{t \rightarrow \tau} Y_t$, à condition que cette limite existe, et on définit $Y_{\tau+}$ comme un processus minimal de même générateur que Y . On applique ensuite la même méthode au second temps d'explosion et ainsi de suite. Du point de vue du problème de martingale, cette construction est naturelle puisque le processus ainsi obtenu est bien solution du problème de martingale associé au générateur de Y , et par la propriété de *quasi-continuité à gauche*, garantissant que toute solution du problème de martingale est continue en tout temps d'arrêt prévisible (voir Ethier-Kurtz [29], chapitre 4, théorème 3.12), c'est aussi le seul. Signalons toutefois que le processus obtenu n'est pas nécessairement défini jusqu'à n'importe quel $t \in \mathbb{R}_+$. En effet, de la même manière qu'un processus minimal peut voir une accumulation de sauts en un temps fini, celui-ci peut avoir une accumulation d'explosions. On pourrait alors définir un temps de double explosion et

recommencer de la même manière indéfiniment, mais ce degré de généralisation ne nous a pas semblé pertinent ici¹.

On prolonge donc le processus minimal X^* déjà construit en un processus stratifié. Afin de pouvoir étendre la propriété d'entrelacement de X et X^* au temps d'explosion τ , il faut que la fonction $x^* \rightarrow \Lambda(x^*, x)$ soit continue en X_τ^* , pour tout $x \in \mathcal{S}$. Par simplicité, on supposera donc que ces fonctions sont continues sur \mathcal{S}^* . Sous cette hypothèse, on prolonge la propriété d'entrelacement au processus stratifié, et on montre que l'on a l'équivalence souhaité : $\mathbf{H} \Leftrightarrow (15) \Leftrightarrow (16m)$. Ce prolongement se rapproche dans l'idée, bien que dans un contexte différent, de celui donné pour l'entrelacement de semi-groupes par le théorème 2.1 de Patie, Savov et Zhao [70].

Chapitre 2

Dans ce chapitre, on répond à la question de l'existence de temps forts de stationnarité fini, pour des processus réversibles sur un certain type d'ensembles, en utilisant les résultats du chapitre 1. On considère un graphe discret G , simple, connexe et non orienté. On appelle *branches infinies* de G les parties connexes formées d'un nombre infini de points de degré 2, maximales au sens de l'inclusion. Une définition plus précise sera proposée en section 2.2, ainsi qu'un rappel du vocabulaire des graphes qui sera utilisé. Pour une introduction plus détaillée à la théorie des graphes (en français), on renvoie à Bretto et Faisant [11].

On suppose que G est constitué d'un nombre fini de branches infinies, plus éventuellement un nombre fini de points de degrés finis. S'il n'y a que des branches infinies, alors il y en a une seule et $G = \mathbb{Z}$. Ce cas particulier est traité dans la section 2.1. Sinon, chacune des N branches infinies (notées Q_i , $i \in I := \llbracket 1, N \rrbracket$) est isomorphe en tant que graphe à \mathbb{N} , et on note $\varphi_i : \mathbb{N} \rightarrow G$ les isomorphismes correspondant. Un tel graphe est représenté dans la figure 2.1.

On se donne ensuite un générateur \mathcal{L} sur G , tel que pour tout $x \neq y \in G$, $\mathcal{L}_{x,y} \neq 0$ si et seulement si x et y sont adjacents. La structure de graphe de G n'apparaissant qu'à travers le générateur, on peut aussi faire le choix de se donner *a priori* un générateur sur un ensemble dénombrable, et de munir cet ensemble de la bonne structure de graphe. C'est la démarche qui est choisie dans la section 2.2. On suppose que \mathcal{L} est récurrent positif et non explosif. Cela se traduit par les conditions :

$$\sum_{n \in \mathbb{N}} \mu^i(n) < \infty, \quad \forall i \in \llbracket 1, N \rrbracket \quad (18)$$

$$\sum_{n \in \mathbb{N}} \frac{1}{\mu^i(n) \mathcal{L}_{\varphi_i(n), \varphi_i(n+1)}} \sum_{m=0}^n \mu^i(m) = \infty \quad \forall i \in \llbracket 1, N \rrbracket \quad (19)$$

où :

$$\mu^i(n) := \prod_{k=0}^{n-1} \frac{\mathcal{L}_{\varphi_i(k), \varphi_i(k+1)}}{\mathcal{L}_{\varphi_i(k+1), \varphi_i(k)}} \quad (20)$$

Il existe alors une unique mesure stationnaire μ , et tout processus X de générateur \mathcal{L} converge en loi vers μ . On suppose de plus que \mathcal{L} est réversible, c'est-à-dire que sa mesure invariante μ satisfait :

$$\forall p, q \in G, \quad \mu(p) \mathcal{L}_{p,q} = \mu(q) \mathcal{L}_{q,p} \quad (21)$$

1. Quel niveau d'explosion un processus peut-il atteindre, tout en gardant un espace d'état dénombrable ? Si l'on appelle temps d'explosion de niveau 1 un temps d'explosion classique, et temps d'explosion de niveau n un temps d'accumulation d'explosions de niveau $n - 1$, peut-il arriver un moment où l'on aura franchi tous les niveaux ?

Dans ce contexte on va construire de manière classique un dual à valeurs dans $\mathcal{P}(G)$ l'ensemble des parties de G , entrelacé avec X par le noyau Λ défini en (17). Dans l'idée des *evolving sets* de Morris et Peres [67] (voir aussi le lien qui est fait entre ces *evolving sets* et les duals de stationnarité en temps discret, dans le livre de Levin, Peres et Wilmer [52]) et de ce qui a été fait par Diaconis et Fill [23] sur des ensembles finis ou par Miclo [65] ou Fill et Lyzinski [37] pour des diffusions réelles, ce dual peut s'agrandir en ajoutant des points voisins. Le principal problème est qu'à chaque instant il doit aussi pouvoir perdre des points de son bord. Afin d'éviter qu'il ne se sépare en plusieurs morceaux, lorsque ceci arrive on le force à choisir une de ses composantes connexes avec une probabilité proportionnelle à sa masse relativement à μ . On obtient ainsi un dual connexe.

En comparant le dual avec des processus de vie et mort, on obtient des conditions d'explosion de celui-ci, le long de chaque branche infinie. Autrement dit, on arrive à déterminer si le dual va pouvoir « remplir » chaque branche infinie en un temps fini. Dans ce cas, on montre qu'il est absorbé en temps fini, ce qui fournit un temps fort de stationnarité fini pour X . En utilisant la séparation de mesures, on montre ensuite pour certaines lois initiales sur X que si le temps fort de stationnarité fourni par le dual n'est pas fini, alors aucun autre ne le sera. On obtient alors le résultat principal :

Théorème 0.4.1. *Le processus X admet un temps fort de stationnarité fini pour toute loi initiale μ_0 , si et seulement si, pour tout $i \in I$:*

$$\sum_{j=1}^{\infty} \mu^i(j+1) \sum_{k=1}^j \frac{1}{\mu^i(k) \mathcal{L}_{\varphi_i(k), \varphi_i(k+1)}} < \infty \quad (22)$$

Miclo [64] a donné une autre condition équivalente à ceci : si on considère la restriction $\mathcal{L}_{\mathcal{F}}$ de l'opérateur défini par \mathcal{L} sur l'ensemble :

$$\mathcal{F} := \{f \in \mathbb{L}^2(\mu) : \mu[f] = 0\},$$

alors (22) est satisfaite pour tout $i \in I$ si et seulement si $\mathcal{L}_{\mathcal{F}}^{-1}$ est à trace finie.

Ce résultat peut aussi s'interpréter en terme de classification des frontières, telle qu'introduite par Feller [30] (voir Anderson [3] pour les chaînes de Markov) : la condition (22) signifie que chaque extrémité d'une branche infinie est une frontière entrante pour le processus X .

Remarque 0.4.2. *Contrairement à ce qui est fait pour les diffusions réelles par exemple, on ne montre pas que le dual que l'on a construit est fin pour certaines lois initiales. En fait, nous pensons qu'il n'existe pas de dual à la fois Λ -lié et fin dès que le nombre de branches infinies est supérieur ou égal à trois (sauf bien sûr dans le cas trivial où la loi initiale est la loi stationnaire).*

Chapitre 3

La recherche et le calcul d'un dual approprié au processus du chapitre 2 a donné lieu à plusieurs observations sur les propriétés que celui-ci peut ou doit satisfaire. Le choix qui a été retenu est principalement lié à des questions de simplicité de calculs, mais n'est pas canonique et peut même sembler peu naturelle. La meilleure illustration de ceci est le fait d'imposer au dual de rester connexe : si les calculs en sont grandement facilités par la suite, c'est au prix de « l'évaporation » pure et simple d'une partie non-négligeable du dual lors de certaines transitions ! Si au lieu de critères de finitude des temps forts de stationnarité, on cherche des informations précises sur la vitesse de convergence du processus par l'intermédiaire du temps d'absorption de son dual, ce choix peut alors sembler totalement farfelu.

L'idée du chapitre 3 est alors de regrouper un certain nombre de propriétés (dont certaines sont à l'état de conjectures) des processus duaux Λ -liés sur des espaces dénombrables assez généraux. On s'affranchit par exemple de l'hypothèse de réversibilité, et on donne une généralisation du générateur dual du chapitre 4. On garde en revanche le même noyau de transition Λ .

On montre par exemple qu'à chaque transition, un dual ne peut s'agrandir que dans la limite des points qui lui sont frontaliers (on ne peut pas passer en un saut d'un état Q à un état $Q' \ni p$ si p n'est pas voisin de Q). À l'inverse, il y a une grande liberté sur les points qu'il peut perdre à chaque transition, mais il est nécessaire que n'importe quel point du bord puisse être enlevé au cours de certaines des transitions possibles.

Ce chapitre est donc plus court que les autres, et d'une importance moindre, mais peut servir de point de départ à d'éventuels développements futurs, notamment pour étendre ou améliorer les résultats du chapitre 2.

Partie II : Dualité pour une diffusion sur un graphe quantique

Cette partie a pour objectif de construire le dual d'une diffusion sur un graphe continu, par analogie avec le chapitre 2. Une fois la dualité obtenue, l'existence d'un temps fort de stationnarité ne pose pas de difficulté et nous avons donc laissé cette problématique de côté.

Chapitre 4

Ce chapitre suit en quelque sorte une procédure standard : on définit un processus par son générateur, on caractérise sa mesure invariante, on définit un second générateur et on vérifie que les deux sont bien entrelacés.

On commence dans la section 4.1 par définir un graphe quantique \mathbb{G} à partir d'un graphe discret fini. Concrètement, cela revient à remplacer les arêtes, abstraites, par des segments dont les extrémités sont des sommets du graphe. En identifiant ces segments à des segments réels par homéomorphisme, on définit naturellement une structure différentielle et métrique. On suit ensuite la définition de Freidlin et Wentzell [40] d'une diffusion sur ce graphe : on définit d'abord son comportement à l'intérieur de chaque arête, par l'intermédiaire d'opérateurs différentiels d'ordre 2 classiques, puis sur chacun des sommets, en posant des conditions sur ceux-ci pour les fonctions du domaine d'un générateur global \mathcal{L} .

Le premier résultat arrive dans la section 4.2, où l'on donne une forme explicite de la mesure invariante μ de \mathcal{L} . Cette mesure est définie par une fonction de densité sur chaque arête, dépendant de la mesure invariante d'un processus de saut sur le graphe discret sous-jacent. On rappelle également un résultat de Miclo [64] donnant une condition équivalente pour sa réversibilité.

Dans la section 4.3, on définit un espace d'intégration \mathcal{G} , ayant vocation à être l'espace d'état d'un processus dual, et on le munit d'une distance le rendant isométrique à une sous-partie de $\mathcal{P}(\mathbb{G})$ munie de la distance de Prokhorov, l'isométrie étant donnée par l'application $A \mapsto \Lambda(A, \cdot)$. On établit ensuite une formule de Stokes pour l'intégrale de \mathcal{L} sur un domaine appartenant à \mathcal{G} , en la ramenant à une intégrale sur la frontière de ce domaine.

Enfin, dans la section 4.4, on définit un générateur dual \mathfrak{L} sur un sous-ensemble \mathcal{D} de \mathcal{G} . Ce générateur est défini par son action sur des fonctionnelles élémentaires de la forme :

$$\mathcal{D} \ni D \mapsto \int_D f d\mu$$

où f est une fonction du domaine de \mathcal{L} . La raison pour laquelle on restreint la définition au sous-ensemble \mathcal{D} est que celle-ci pose problème sur certains éléments de \mathcal{G} , à savoir ceux dont la frontière rencontre un ou plusieurs sommets du graphe. Pour ces derniers, on définit, plus ou moins arbitrairement, la valeur du générateur comme étant égale à sa limite croissante. On espère que cette définition amènera à l'existence (et l'unicité) d'un processus associé au générateur *via* le problème de martingale. On montre que ce générateur est bien entrelacé avec celui de la diffusion, au sens de la relation (16), et que le processus associé (s'il existe) est absorbé en temps fini. Sous condition d'existence, un processus associé au générateur dual peut donc être couplé avec le processus de diffusion de manière à être un dual Λ -lié. Malheureusement, la question d'existence de ce processus n'est à ce jour pas complètement résolue. Les deux chapitres suivants sont orientés vers cette question.

Chapitre 5

Dans ce chapitre, on montre dans le cas d'un mouvement brownien sur un graphe un peu plus simple que le problème de martingale associé au générateur dual est bien posé, à condition d'élargir son domaine. Le graphe considéré ici est composé de n branches de même longueur et ayant un sommet commun O . L'idée est alors de ramener l'étude de ce graphe à celle du segment $[-1, 1]$ avec une singularité en 0, dont la partie $[-1, 0]$ correspond à $n - 1$ branches de paramétrisation identique, la moitié $[0, 1]$ à la dernière branche et 0 naturellement à leur sommet commun O .

L'étude du dual sur ce segment $[-1, 1]$ passe par l'utilisation d'un résultat d'existence et d'unicité de la solution d'une équation différentielle stochastique avec temps local, énoncé par Takanobu [84], [85]. Dans une première section on montre donc comment formuler cette EDS de manière équivalente par un problème de martingale, dans un cadre de généralité intermédiaire entre celui utilisé dans la suite, et celui dans lequel le résultat original est énoncé.

On définit ensuite un générateur sur le segment $[-1, 1]$, correspondant à un processus évoluant parmi les sous-intervalles de ce segment, à la manière du dual défini par Miclo [65]. Contrairement à la diffusion duale sur \mathbb{R} , le dual défini ici ne s'identifie pas à une classique diffusion sur un sous-ensemble de \mathbb{R}^2 . En effet, la présence de conditions particulières en O pose problème lorsque l'une des extrémités du dual arrive en 0. L'équation différentielle stochastique étudiée par Takanobu ne permettant *a priori* d'envisager une singularité qu'en une seule des variables, on montre dans la section 5.2 comment contourner ce problème, en considérant des intervalles de temps où une seule des deux variables passe en 0.

Enfin, dans la section 5.3, on retranscrit ce résultat sur le graphe à n branches, et on en déduit l'existence d'un dual pour le mouvement brownien, caractérisé comme l'unique solution du problème de martingale considéré.

Chapitre 6

Ce chapitre a pour double vocation de proposer une démonstration, pas totalement aboutie, de l'existence d'une solution au problème de martingale associé au générateur dual dans le cadre général, et de faire le lien avec la première partie. Pour cela, on tentera d'approximer le modèle continu à l'aide du discret. Notons que cette approximation est faite par des processus homogènes et à temps continu, contrairement à ce qui est fait en général (par exemple par Pitman et Rogers [71] pour l'approximation d'un processus de Bessel, ou par Karlin et Taylor [46] pour une des constructions classiques du mouvement brownien).

Dans la section 6.1, on définit une suite de discrétisation du graphe. Cette discrétisation présente une difficulté technique par rapport à un unique segment, liée à la non-uniformité de la longueur des arêtes. On contournera cette difficulté par l'utilisation d'un théorème de Dirichlet d'approximation par des nombres rationnels. Sur cette suite d'approximations, on définira ensuite des chaînes de Markov à temps continu, et en utilisant des outils classiques de convergence de mesure, ainsi que la convergence uniforme des générateurs, on montrera que la suite de processus converge en loi vers la diffusion sur le graphe quantique. On montrera également que la suite des mesures invariantes associées converge faiblement vers la mesure invariante de la diffusion.

Dans la section 6.2, on commence par redéfinir le générateur dual \mathfrak{L} d'une manière à la fois plus pratique à approximer et plus similaire à la définition d'un générateur de diffusion multi-dimensionnelle classique, et en même temps beaucoup moins agréable à manipuler. Sa nouvelle expression implique notamment de définir une notion de dérivées directionnelles en chaque point de la frontière d'un ensemble $D \in \mathcal{D}$, pour toute fonction réelle sur \mathcal{D} (on remarque notamment que le nombre de points sur la frontière est variable en fonction de l'ensemble D considéré et que ces dérivées ne peuvent donc pas être définies comme des dérivées partielles classiques). On utilise ensuite un exemple de dual discret donné dans le chapitre 3, et on l'applique à la suite des chaînes de Markov définie précédemment. On montre que les générateurs convergent en un certain sens vers le générateur dual étudié ; malheureusement, cette convergence n'est pas (complètement) uniforme, ce qui ne permet pas de déduire facilement la convergence en loi de la suite des processus duaux. La fin du chapitre est consacrée à des pistes de réflexion pour remédier à ce problème.

Première partie

Stationnarité pour une chaîne de Markov

Chapitre 1

Dualité et entrelacement

Cette section présente des outils généraux faisant le lien entre le couplage par dualité de deux processus et l'entrelacement de leurs générateurs et semi-groupes respectifs. Elle reprend en grande partie les résultats et les démonstrations de Fill [32] (sections 2.2 et 2.3) dans le cadre défini par ses *General Settings* et les adapte au cas infini, d'abord non-explosif puis explosif.

Dans tout ce qui suit, on travaillera sur les objets suivants :

- \mathcal{S} et \mathcal{S}^* sont deux espaces topologiques dénombrables (non nécessairement munis de la topologie discrète). Les éléments de ces ensembles seront généralement désignés par les lettres x, y, z pour \mathcal{S} et x^*, y^*, z^* pour \mathcal{S}^* .
- Λ est un noyau de transition de \mathcal{S}^* vers \mathcal{S} .
- \mathcal{L} et \mathcal{L}^* sont des générateurs respectivement sur \mathcal{S} et \mathcal{S}^* et on suppose \mathcal{L} non-explosif.
- $\mathbf{P}(\cdot)$ et $\mathbf{P}^*(\cdot)$ sont les solutions minimales des équations de Kolmogorov de \mathcal{L} et \mathcal{L}^* respectivement.
- μ_0 et μ_0^* sont deux mesures de probabilité sur \mathcal{S} et \mathcal{S}^* respectivement.
- $(X_t)_{t \geq 0}$ est un processus de Markov, de générateur \mathcal{L} et de loi initiale μ_0 .

Au lieu de l'hypothèse de finitude des ensembles $\{x : \Lambda(x^*, x) > 0\}$ qui est faite dans le *General Settings* de Fill [32] (section 2.2B), on supposera que la diagonale de \mathcal{L} est Λ -intégrable, c'est-à-dire :

$$\forall x^* \in \mathcal{S}^*, \quad \sum_{x \in \mathcal{S}} \Lambda(x^*, x) \mathcal{L}_x < +\infty$$

En particulier, pour tout $f \in \mathcal{B}(\mathcal{S})$, le produit matriciel $\Lambda \mathcal{L} f$ est bien défini, associatif et fini.

1.1 Processus minimaux

Cette section est une reformulation détaillée et légèrement adaptée de [32]. Son but est de donner des conditions suffisantes pour pouvoir construire un couplage de deux processus minimaux X et X^* de générateurs respectifs \mathcal{L} et \mathcal{L}^* , de manière à ce que X^* et X soient Λ -liés. Plus précisément, on va montrer le :

Théorème 1.1.1. *Si (μ_0, \mathcal{L}) et (μ_0^*, \mathcal{L}^*) satisfont la double condition :*

$$\begin{cases} \mu_0^* \Lambda = \mu_0 \\ \mathcal{L}^* \Lambda = \Lambda \mathcal{L} \end{cases} \quad (1.1)$$

alors il existe un processus minimal $(X_t^*)_{t \geq 0}$ sur \mathcal{S}^* de générateur \mathcal{L}^* et de loi initiale μ_0^* , satisfaisant (10) et (12) jusqu'à son temps d'explosion T , c'est-à-dire, $\forall t \geq 0$:

$$\mathcal{L}(X_t | \mathcal{F}_t^{X^*}, T > t) = \Lambda(X_t^*, \cdot) \quad (1.2)$$

$$\mathcal{L}(X_t^* | X, T > t) = \mathcal{L}(X_t^* | \mathcal{F}_t^X, T > t) \quad (1.3)$$

Pour tout $f \in \mathcal{B}(\mathcal{S})$ et tout $x^* \in \mathcal{S}^*$, on a :

$$\sum_{y^* \in \mathcal{S}^*} \sum_{y \in \mathcal{S}} |\mathcal{L}_{x^*, y^*}^* \Lambda(y^*, y) f(y)| \leq \sup_{y \in \mathcal{S}} f(y) \sum_{y^* \in \mathcal{S}^*} |\mathcal{L}_{x^*, y^*}^*| = 2 \sup_{y \in \mathcal{S}} f(y) \mathcal{L}_{x^*}^*$$

donc le produit matriciel $\mathcal{L}^* \Lambda f$ est toujours commutatif et à valeur finies. Sous la condition (1.1), on pose $\Gamma := \mathcal{L}^* \Lambda = \Lambda \mathcal{L}$ et on définit la matrice $\bar{\mathcal{L}}$, qui sera celle d'un générateur de couplage, par :

$$\bar{\mathcal{L}}_{\bar{x}, \bar{y}} := \begin{cases} - \left(\mathcal{L}_x + \mathcal{L}_{x^*}^* + \frac{\Gamma(x^*, x)}{\Lambda(x^*, x)} \right) & , \text{ si } \bar{x} = \bar{y} \\ \mathcal{L}_{x, y} & , \text{ si } y \neq x, y^* = x^* \\ \frac{\mathcal{L}_{x^*, y^*}^* \Lambda(y^*, x)}{\Lambda(x^*, x)} & , \text{ si } y = x, y^* \neq x^* \\ \frac{\mathcal{L}_{x, y} \mathcal{L}_{x^*, y^*}^* \Lambda(y^*, y)}{\Gamma(x^*, y)} & , \text{ si } \Lambda(x^*, y) = 0, \Gamma(x^*, y) > 0 \\ 0 & , \text{ sinon} \end{cases} \quad (1.4)$$

avec :

$$\bar{x} = (x, x^*), \bar{y} = (y, y^*) \in \bar{\mathcal{S}} := \{(z, z^*) \in \mathcal{S} \times \mathcal{S}^*; \Lambda(z^*, z) > 0\}$$

Quelques remarques sur cette définition :

Remarque 1.1.2. (a) $\Lambda(x^*, y) = 0$ implique à la fois $x \neq y$ et $x^* \neq y^*$, donc ces cinq cas sont mutuellement exclusifs (et évidemment exhaustifs).

(b) Si $\mathcal{L}_{x, y} > 0$ et $\Lambda(x^*, y) = 0$, alors d'après la remarque précédente :

$$\begin{aligned} \Gamma(x^*, y) &= \sum_z \Lambda(x^*, z) \mathcal{L}_{z, y} \\ &= \sum_{z \neq y} \Lambda(x^*, z) \mathcal{L}_{z, y} \\ &\geq \Lambda(x^*, x) \mathcal{L}_{x, y} > 0 \end{aligned}$$

(c) En utilisant le même argument pour \mathcal{L}_{x^*, y^*}^* , et le fait que le quatrième coefficient dans (1.4) vaut 0 dès lors que $\mathcal{L}_{x, y}$ ou \mathcal{L}_{x^*, y^*}^* est nul, on peut remplacer le quatrième cas par « $\Lambda(x^*, y) = 0, \mathcal{L}_{x, y} \mathcal{L}_{x^*, y^*}^* > 0$ » sans changer le générateur (dans ce cas on a toujours $\Gamma(x^*, y) > 0$, et dans les cas qui ont été rajoutés au cinquième, le coefficient $\bar{\mathcal{L}}_{\bar{x}, \bar{y}}$ vaut bien 0).

Lemme 1.1.3. $\bar{\mathcal{L}}$ définit un générateur sur $\bar{\mathcal{S}}$.

Démonstration. Pour tout $\bar{x} = (x, x^*) \in \bar{\mathcal{S}}$, on a :

$$\bar{\mathcal{L}}_{\bar{x}, \bar{y}} \geq 0, \quad \forall \bar{y} \in \bar{\mathcal{S}}, \bar{y} \neq \bar{x}$$

et :

$$\begin{aligned}
\sum_{\substack{\bar{y} \in \bar{\mathcal{S}} \\ y^* \neq x, y^* = x^*}} \bar{\mathcal{L}}_{\bar{x}, \bar{y}} &= \sum_{\substack{y \neq x \\ \Lambda(x^*, y) > 0}} \mathcal{L}_{x, y} \\
&= \sum_{y \neq x} \mathcal{L}_{x, y} - \sum_{\substack{y \neq x \\ \Lambda(x^*, y) = 0}} \mathcal{L}_{x, y} \\
&= \mathcal{L}_x - \sum_{y | \Lambda(x^*, y) = 0} \mathcal{L}_{x, y} \\
\sum_{\substack{\bar{y} \in \bar{\mathcal{S}} \\ y^* \neq x^*, y = x}} \bar{\mathcal{L}}_{\bar{x}, \bar{y}} &= \sum_{\substack{y^* \neq x^* \\ \Lambda(y^*, x) > 0}} \frac{\mathcal{L}_{x^*, y^*} \Lambda(y^*, x)}{\Lambda(x^*, x)} \\
&= \sum_{y^*} \frac{\mathcal{L}_{x^*, y^*} \Lambda(y^*, x)}{\Lambda(x^*, x)} - \frac{\mathcal{L}_{x^*, x^*} \Lambda(x^*, x)}{\Lambda(x^*, x)} \\
&= \frac{\Gamma(x^*, x)}{\Lambda(x^*, x)} + \mathcal{L}_{x^*}
\end{aligned}$$

(Dans la seconde inégalité ci-dessus, en supprimant la condition $\Lambda(y^*, x) > 0$, on a juste rajouté des termes nuls à la somme)

$$\begin{aligned}
\sum_{\substack{\bar{y} | \Lambda(x^*, y) = 0 \\ \Gamma(x^*, y) > 0}} \bar{\mathcal{L}}_{\bar{x}, \bar{y}} &= \sum_{\substack{y | \Lambda(x^*, y) = 0 \\ \Gamma(x^*, y) > 0}} \mathcal{L}_{x, y} \sum_{y^* | \Lambda(y^*, y) > 0} \frac{\mathcal{L}_{x^*, y^*} \Lambda(y^*, y)}{\Gamma(x^*, y)} \\
&= \sum_{y | \Lambda(x^*, y) = 0} \mathcal{L}_{x, y}
\end{aligned}$$

(d'après la remarque (b))

Donc :

$$\begin{aligned}
\sum_{\bar{y} \neq \bar{x}} \bar{\mathcal{L}}_{\bar{x}, \bar{y}} &= \sum_{\substack{\bar{y} \in \bar{\mathcal{S}} \\ y^* \neq x, y^* = x^*}} \bar{\mathcal{L}}_{\bar{x}, \bar{y}} + \sum_{\substack{\bar{y} \in \bar{\mathcal{S}} \\ y^* \neq x^*, y = x}} \bar{\mathcal{L}}_{\bar{x}, \bar{y}} + \sum_{\substack{\bar{y} | \Lambda(x^*, y) = 0 \\ \Gamma(x^*, y) > 0}} \bar{\mathcal{L}}_{\bar{x}, \bar{y}} \\
&= \mathcal{L}_x - \sum_{y | \Lambda(x^*, y) = 0} \mathcal{L}_{x, y} + \frac{\Gamma(x^*, x)}{\Lambda(x^*, x)} + \mathcal{L}_{x^*} + \sum_{y | \Lambda(x^*, y) = 0} \mathcal{L}_{x, y} \\
&= -\bar{\mathcal{L}}_{\bar{x}, \bar{x}}
\end{aligned}$$

Ainsi, $\bar{\mathcal{L}}$ est bien le générateur d'un processus markovien sur $\bar{\mathcal{S}}$. □

Démonstration du théorème 1.1.1. La démonstration est fortement inspirée de Fill [32], Proposition 2 et 4 et Section 2.3.

Supposons que la condition (1.1) soit réalisée.

Dans un premier temps, on va construire X^* couplé avec X , de manière à ce que le couple (X, X^*) ait $\bar{\mathcal{L}}$ comme générateur. On vérifiera ensuite que (X, X^*) satisfait bien (1.2) et (1.3).

Couplage des deux processus : Ayant observé $X_0 = x_0 \in \mathcal{S}$, on pose pour tout $x_0^* \in \mathcal{S}^*$:

$$X_0^* = x_0^* \in \mathcal{S}^* \quad \text{avec probabilité} \quad \frac{\mu_0^*(x_0^*)\Lambda(x_0^*, x_0)}{\mu_0(x_0)} \quad (1.5)$$

La loi de (X_0, X_0^*) est donnée par :

$$\begin{aligned} \bar{\mu}_0(x_0, x_0^*) &= \mathbb{P}(X_0 = x_0)\mathbb{P}(X_0^* = x_0^* | X_0 = x_0) \\ &= \mu_0(x_0) \frac{\mu_0^*(x_0^*)\Lambda(x_0^*, x_0)}{\mu_0(x_0)} \\ &= \mu_0^*(x_0^*)\Lambda(x_0^*, x_0) \end{aligned} \quad (1.6)$$

qui est bien une mesure de probabilité sur $\bar{\mathcal{S}}$ car :

$$\begin{aligned} \sum_{x_0, x_0^*} \bar{\mu}_0(x_0, x_0^*) &= \sum_{x_0^*} \mu_0^*(x_0^*) \sum_{x_0} \Lambda(x_0^*, x_0) \\ &= \sum_{x_0^*} \mu_0^*(x_0^*) \\ &= 1 \end{aligned}$$

De plus, $X_0^* \sim \mu_0^*$.

On pose $\tau_0 = 0$, donc pour l'instant X^* est construit jusqu'à ce temps (on rappelle que X est déjà construit). On va maintenant construire X^* étape par étape jusqu'aux temps τ_k , $k \in \mathbb{N}$, où τ_k est la k -ème transition de la chaîne bivariée (X, X^*) (en particulier, X^* peut ne pas sauter au temps τ_k).

Soit $k \geq 1$, et supposons par récurrence que $\tau_0, \dots, \tau_{k-1}$ et $X_{[0, \tau_{k-1}]}$ sont déjà construits. On appelle σ_k le temps de la première transition de X après τ_{k-1} (qui peut être égal à σ_{k-1} si X n'a pas sauté au temps τ_{k-1}). Soit $\bar{x} = (x, x^*) := (X, X^*)_{\tau_{k-1}}$, et ε_{k-1} une variable exponentielle de paramètre $\bar{\mathcal{L}}_{\bar{x}} - \mathcal{L}_x$, indépendante de $(\varepsilon_i)_{i \leq k-2}$ et X .

1. Si $\tau_{k-1} + \varepsilon_{k-1} > \sigma_k$, on pose $\tau_k = \sigma_k$, et

$$X_{\tau_k}^* = y^* \quad \text{avec probabilité} \quad \frac{\bar{\mathcal{L}}_{\bar{x}, \bar{y}}}{\mathcal{L}_{x, y}} \quad (1.7)$$

où $y = X_{\tau_k}$ et $\bar{y} = (y, y^*)$.

2. Sinon, on pose $\tau_k = \tau_{k-1} + \varepsilon_{k-1}$, et

$$X_{\tau_k}^* = y^* \neq x^* \quad \text{avec probabilité} \quad \frac{\bar{\mathcal{L}}_{(x, x^*), (x, y^*)}}{\bar{\mathcal{L}}_{\bar{x}} - \mathcal{L}_x} \quad (1.8)$$

Ces probabilités sont bien définies, puisque $\mathcal{L}_{x, y}$ et $\bar{\mathcal{L}}_{\bar{x}} - \mathcal{L}_x$ sont évidemment non nuls si $X_{\tau_k} = y$ ou si $\varepsilon_{k-1} < \infty$ respectivement. De plus, dans le premier cas, on obtient comme annoncé $X_{\tau_k}^* = x^*$ p.s. si $\Lambda(x^*, y) \neq 0$, et

$$X_{\tau_k}^* = y^* \neq x^* \quad \text{avec probabilité} \quad \frac{\mathcal{L}_{x^*, y^*}^* \Lambda(y^*, y)}{\Gamma(x^*, y)}$$

si $\Lambda(x^*, y) = 0$, la somme de ces probabilités valant bien 1.

On a donc $\tau_k - \tau_{k-1} = \min(\varepsilon_{k-1}, \sigma_k - \tau_{k-1})$. Or, $\varepsilon_{k-1} \sim \exp(\bar{\mathcal{L}}_{\bar{x}} - \mathcal{L}_x)$ et $\sigma_k - \tau_{k-1} \sim \exp(\mathcal{L}_x)$ sont indépendantes, donc $\tau_k - \tau_{k-1} \sim \exp(\bar{\mathcal{L}}_{\bar{x}})$. De plus, considérant les événements $A := \{(X, X^*)_{\tau_{k-1}} = \bar{x}\}$ et $B := \{(X, X^*)_{\tau_k} = \bar{y}\}$, on a les probabilités de transition suivantes :

- Si $y = x$, $y^* \neq x^*$:

$$\begin{aligned} \mathbb{P}(B | A) &= \mathbb{P}(X_{\tau_k} = x | A) \mathbb{P}(X_{\tau_k}^* = y^* | X_{\tau_k} = x, A) \\ &= \mathbb{P}(\varepsilon_{k-1} < \sigma_k - \tau_{k-1}) \mathbb{P}(X_{\tau_k}^* = y^* | \varepsilon_{k-1} < \sigma_k - \tau_{k-1}, A) \\ &= \left(\frac{\bar{\mathcal{L}}_{\bar{x}} - \mathcal{L}_x}{\bar{\mathcal{L}}_{\bar{x}}} \right) \left(\frac{\bar{\mathcal{L}}_{(x, x^*), (x, y^*)}}{\bar{\mathcal{L}}_{\bar{x}} - \mathcal{L}_x} \right) \\ &= \frac{\bar{\mathcal{L}}_{(x, x^*), (x, y^*)}}{\bar{\mathcal{L}}_{\bar{x}}} \end{aligned}$$

- Si $y \neq x$:

$$\begin{aligned} \mathbb{P}(B | A) &= \mathbb{P}(X_{\tau_k} = y | A) \mathbb{P}(X_{\tau_k}^* = y^* | X_{\tau_k} = y, A) \\ &= \frac{\mathcal{L}_{x, y}}{\mathcal{L}_x} \mathbb{P}(\varepsilon_{k-1} > \sigma_k - \tau_{k-1} | A) \mathbb{P}(X_{\tau_k}^* = x^* | X_{\tau_k} = x, A) \\ &= \frac{\mathcal{L}_{x, y}}{\mathcal{L}_x} \left(\frac{\mathcal{L}_x}{\bar{\mathcal{L}}_{\bar{x}}} \right) \left(\frac{\bar{\mathcal{L}}_{\bar{x}, \bar{y}}}{\mathcal{L}_{x, y}} \right) \\ &= \frac{\bar{\mathcal{L}}_{\bar{x}, \bar{y}}}{\bar{\mathcal{L}}_{\bar{x}}} \end{aligned}$$

Par construction, le processus (X, X^*) est un processus de saut markovien, défini jusqu'au premier temps d'explosion, de générateur $\bar{\mathcal{L}}$, et satisfaisant (1.3). De plus, le temps de chaque saut est strictement positif donc les trajectoires de (X, X^*) sont continues à droite.

Vérification de l'entrelacement (relation (1.2)) : Cette vérification s'effectue en trois temps. Dans un premier temps on va montrer une relation sur les générateurs :

$$\forall x^* \in \mathcal{S}^*, f \in \mathcal{B}(\bar{\mathcal{S}}), \quad \sum_{x \in \mathcal{S}} \sum_{\bar{y} \in \bar{\mathcal{S}}} \Lambda(x^*, x) \bar{\mathcal{L}}_{\bar{x}, \bar{y}} f(\bar{y}) = \sum_{\bar{y} \in \bar{\mathcal{S}}} \mathcal{L}_{x^*, y^*}^* \Lambda(y^*, y) f(\bar{y}) \quad (1.9)$$

Intuitivement, le terme de gauche correspond à toutes les transitions d'un $x^* \in \mathcal{S}^*$ donné vers un $x \in \mathcal{S}$, et ensuite de $(x, x^*) \in \bar{\mathcal{S}}$ vers $(y, y^*) \in \bar{\mathcal{S}}$, tandis que le terme de droite correspond à une transition de x^* vers y^* , puis vers y , sans se préoccuper de x .

Dans un second temps, on s'en servira pour montrer une inégalité au niveau des semi-groupes correspondants :

$$\forall x^* \in \mathcal{S}^*, \bar{y} \in \bar{\mathcal{S}} \quad \sum_{x \in \mathcal{S}} \Lambda(x^*, x) \bar{\mathbf{P}}_{\bar{x}, \bar{y}}(t) \geq \mathbf{P}_{x^*, y^*}^*(t) \Lambda(y^*, y) \quad (1.10)$$

et enfin, on en déduira par récurrence sur k :

$$\mathbb{P}(X_{t_0}^* = x_0^*, \dots, X_{t_k}^* = x_k^*, X_{t_k} = x_k) = \mu_0^*(x_0^*) \prod_{i=1}^k \mathbf{P}_{x_{i-1}^*, x_i^*}^*(t_i - t_{i-1}) \Lambda(x_k^*, x_k) \quad (1.11)$$

$\forall k \geq 0, \forall 0 = t_0 \leq \dots \leq t_k, \forall x_0^*, \dots, x_k^* \in \mathcal{S}^*, \forall y_k \in \mathcal{S}$, ce qui implique à la fois (1.2) et que $\mathbf{P}^*(\cdot)$ est le semi-groupe de X^* .

Preuve de (1.9) : On commence par montrer l'égalité matricielle suivante :

$$\sum_{x \in \mathcal{S}} \Lambda(x^*, x) \bar{\mathcal{L}}_{\bar{x}, \bar{y}} = \mathcal{L}_{x^*, y^*}^* \Lambda(y^*, y)$$

Soit $x^* \in \mathcal{S}^*$, $\bar{y} \in \bar{\mathcal{S}}$, avec $x^* \neq y^*$. Si $\Lambda(y, x^*) \neq 0$, alors d'après les coefficients du générateur $\bar{\mathcal{L}}$, $\bar{\mathcal{L}}_{\bar{x}, \bar{y}} \neq 0$ implique $y = x$. D'où :

$$\begin{aligned} \sum_{x \in \mathcal{S}} \Lambda(x^*, x) \bar{\mathcal{L}}_{\bar{x}, \bar{y}} &= \begin{cases} \Lambda(x^*, y) \bar{\mathcal{L}}_{(y, x^*), \bar{y}} & \text{si } (y, x^*) \in \bar{\mathcal{S}} \\ \sum_{x \neq y} \Lambda(x^*, x) \bar{\mathcal{L}}_{\bar{x}, \bar{y}} & \text{sinon} \end{cases} \\ &= \begin{cases} \Lambda(x^*, y) \frac{\mathcal{L}_{x^*, y^*}^* \Lambda(y^*, y)}{\Lambda(x^*, y)} & \text{si } (y, x^*) \in \bar{\mathcal{S}} \\ \sum_{x \neq y} \Lambda(x^*, x) \frac{\mathcal{L}_{x, y} \mathcal{L}_{x^*, y^*}^* \Lambda(y^*, y)}{\Gamma(x^*, y)} & \text{sinon} \end{cases} \\ &= \mathcal{L}_{x^*, y^*}^* \Lambda(y^*, y) \end{aligned}$$

Et si $x^* \in \mathcal{S}^*$, $\bar{y} \in \bar{\mathcal{S}}$, avec $x^* = y^*$, alors :

$$\begin{aligned} \sum_{x: \bar{x} \in \bar{\mathcal{S}}} \Lambda(x^*, x) \bar{\mathcal{L}}_{\bar{x}, \bar{y}} &= -\Lambda(x^*, y) \bar{\mathcal{L}}_{(y, x^*)} + \sum_{x \neq y} \Lambda(x^*, x) \bar{\mathcal{L}}_{\bar{x}, \bar{y}} \\ &= -\Lambda(x^*, y) \left(\mathcal{L}_y + \mathcal{L}_{x^*}^* + \frac{\Gamma(x^*, y)}{\Lambda(x^*, y)} \right) + \sum_{x \neq y} \Lambda(x^*, x) \mathcal{L}_{x, y} \\ &= -\Lambda(x^*, y) (\mathcal{L}_y + \mathcal{L}_{x^*}^*) - \Gamma(x^*, y) + \Gamma(x^*, y) + \Lambda(x^*, y) \mathcal{L}_y \\ &= -\mathcal{L}_{x^*}^* \Lambda(x^*, y) \\ &= \mathcal{L}_{x^*, y^*}^* \Lambda(y^*, y) \end{aligned}$$

Maintenant, pour $f \in \mathcal{B}(\bar{\mathcal{S}})$, on a :

$$\begin{aligned} \sum_{x: \bar{x} \in \bar{\mathcal{S}}} \sum_{\bar{y} \in \bar{\mathcal{S}}} \Lambda(x^*, x) \bar{\mathcal{L}}_{\bar{x}, \bar{y}} f(\bar{y}) &= \sum_{x: \bar{x} \in \bar{\mathcal{S}}} \left[\Lambda(x^*, x) \bar{\mathcal{L}}_{\bar{x}, \bar{x}} f(\bar{x}) + \sum_{\substack{y: (y, x^*) \in \bar{\mathcal{S}} \\ y \neq x}} \Lambda(x^*, x) \mathcal{L}_{x, y} f((y, x^*)) \right. \\ &\quad \left. + \sum_{\substack{y^*: (x, y^*) \in \bar{\mathcal{S}} \\ y^* \neq x^*}} \mathcal{L}_{x^*, y^*}^* \Lambda(y^*, x) f((x, y^*)) + \sum_{\bar{y}: (y, x^*) \notin \bar{\mathcal{S}}} \frac{\Lambda(x^*, x) \mathcal{L}_{x, y} \mathcal{L}_{x^*, y^*}^* \Lambda(y^*, y) f(\bar{y})}{\Gamma(x^*, y)} \right] \\ &= \sum_{x: \bar{x} \in \bar{\mathcal{S}}} \Lambda(x^*, x) \bar{\mathcal{L}}_{\bar{x}, \bar{x}} f(\bar{x}) + \sum_{x: \bar{x} \in \bar{\mathcal{S}}} \sum_{\substack{y: (y, x^*) \in \bar{\mathcal{S}} \\ y \neq x}} \Lambda(x^*, x) \mathcal{L}_{x, y} f((y, x^*)) \\ &\quad + \sum_{x: \bar{x} \in \bar{\mathcal{S}}} \sum_{\substack{y^*: (x, y^*) \in \bar{\mathcal{S}} \\ y^* \neq x^*}} \mathcal{L}_{x^*, y^*}^* \Lambda(y^*, x) f((x, y^*)) + \sum_{x: \bar{x} \in \bar{\mathcal{S}}} \sum_{\bar{y}: (y, x^*) \notin \bar{\mathcal{S}}} \frac{\Lambda(x^*, x) \mathcal{L}_{x, y} \mathcal{L}_{x^*, y^*}^* \Lambda(y^*, y) f(\bar{y})}{\Gamma(x^*, y)} \end{aligned}$$

dès que les quatre sommes dans la dernière égalité convergent. Pour la première, rappelons que :

$$\bar{\mathcal{L}}_{\bar{x}, \bar{x}} = \mathcal{L}_{x, x} + \mathcal{L}_{x^*, x^*}^* - \frac{\Gamma(x^*, x)}{\Lambda(x^*, x)}$$

donc on voit que la première somme converge, puisque la diagonale de \mathcal{L} est μ -intégrable. La deuxième converge pour la même raison, la troisième parce que f est bornée et :

$$\sum_{y^* \in \mathcal{S}^*} \sum_{x \in \mathcal{S}} |\mathcal{L}_{x^*, y^*}^*| \Lambda(y^*, x) = 2\mathcal{L}_{x^*}^* < +\infty,$$

et la dernière parce que f est bornée et :

$$\begin{aligned} \sum_{y^* \in \mathcal{S}^*} \sum_{y: (y, x^*) \notin \bar{\mathcal{S}}} \sum_{x \in \mathcal{S}} \frac{\Lambda(x^*, x) |\mathcal{L}_{x,y} \mathcal{L}_{x^*, y^*}^*| \Lambda(y^*, y)}{\Gamma(x^*, y)} &= \sum_{y^* \in \mathcal{S}^*} \sum_{y: (y, x^*) \notin \bar{\mathcal{S}}} |\mathcal{L}_{x^*, y^*}^*| \Lambda(y^*, y) \\ &\leq 2\mathcal{L}_{x^*}^* \end{aligned}$$

Revenant à nos calculs précédents, on déduit :

$$\begin{aligned} \sum_{x: \bar{x} \in \bar{\mathcal{S}}} \sum_{\bar{y} \in \bar{\mathcal{S}}} \Lambda(x^*, x) \bar{\mathcal{L}}_{\bar{x}, \bar{y}} f(\bar{y}) &= \sum_{x: \bar{x} \in \bar{\mathcal{S}}} \Lambda(x^*, x) \bar{\mathcal{L}}_{\bar{x}, \bar{x}} f(\bar{x}) + \sum_{x: \bar{x} \in \bar{\mathcal{S}}} \sum_{\substack{y: (y, x^*) \in \bar{\mathcal{S}} \\ y \neq x}} \Lambda(x^*, x) \mathcal{L}_{x,y} f((y, x^*)) \\ &+ \sum_{x: \bar{x} \in \bar{\mathcal{S}}} \sum_{\substack{y^*: (x, y^*) \in \bar{\mathcal{S}} \\ y^* \neq x^*}} \mathcal{L}_{x^*, y^*}^* \Lambda(y^*, x) f((x, y^*)) + \sum_{x: \bar{x} \in \bar{\mathcal{S}}} \sum_{\bar{y}: (y, x^*) \notin \bar{\mathcal{S}}} \frac{\Lambda(x^*, x) \mathcal{L}_{x,y} \mathcal{L}_{x^*, y^*}^* \Lambda(y^*, y) f(\bar{y})}{\Gamma(x^*, y)} \\ &= \sum_{x: \bar{x} \in \bar{\mathcal{S}}} \Lambda(x^*, x) \bar{\mathcal{L}}_{\bar{x}, \bar{x}} f(\bar{x}) + \sum_{y: (y, x^*) \in \bar{\mathcal{S}}} \sum_{x \neq y \in \mathcal{S}} \Lambda(x^*, x) \mathcal{L}_{x,y} f((y, x^*)) \\ &+ \sum_{y^* \in \mathcal{S}^*} \sum_{x: (x, y^*) \in \bar{\mathcal{S}}} \mathcal{L}_{x^*, y^*}^* \Lambda(y^*, x) f((x, y^*)) + \sum_{\bar{y}: (y, x^*) \notin \bar{\mathcal{S}}} \sum_{x \neq y \in \mathcal{S}} \frac{\Lambda(x^*, x) \mathcal{L}_{x,y} \mathcal{L}_{x^*, y^*}^* \Lambda(y^*, y) f(\bar{y})}{\Gamma(x^*, y)} \\ &= \sum_{\bar{y} \in \bar{\mathcal{S}}} \sum_{x: \bar{x} \in \bar{\mathcal{S}}} \Lambda(x^*, x) \bar{\mathcal{L}}_{\bar{x}, \bar{y}} f(\bar{y}) \\ &= \sum_{\bar{y} \in \bar{\mathcal{S}}} \mathcal{L}_{x^*, y^*}^* \Lambda(y^*, y) f(\bar{y}) \end{aligned}$$

Preuve de (1.10) : Soit $\bar{\mathbf{P}}(\cdot)$ la solution minimale des équations de Kolmogorov associées à $\bar{\mathcal{L}}$. Pour tout $t > 0$, on définit la matrice $(Q_{x^*, \bar{y}}(t))_{(x^*, \bar{y}) \in \mathcal{S}^* \times \bar{\mathcal{S}}}$ de la manière suivante :

$$Q_{x^*, \bar{y}}(t) := \sum_{x: \bar{x} \in \bar{\mathcal{S}}} \Lambda(x^*, x) \bar{\mathbf{P}}_{\bar{x}, \bar{y}}(t)$$

Soit $(A_n)_n$ une suite croissante de parties finies de \mathcal{S} , telle que $\lim A_n = \mathcal{S}$. Si on définit aussi pour tout $n \in \mathbb{N}$, $(x^*, \bar{y}) \in \mathcal{S}^* \times \bar{\mathcal{S}}$ et $t > 0$:

$$Q_{x^*, \bar{y}}^{(n)}(t) := \sum_{x \in A_n} \Lambda(x^*, x) \bar{\mathbf{P}}_{\bar{x}, \bar{y}}(t)$$

alors on a :

$$\sup_{t \geq 0} |Q_{x^*, \bar{y}}^{(n)}(t) - Q_{x^*, \bar{y}}(t)| \leq \left| \sum_{x \in \mathcal{S} \setminus A_n} \Lambda(x^*, x) \right| \xrightarrow{n \rightarrow \infty} 0$$

donc Q est limite uniforme (en t) de $(Q^{(n)})_n$, et on obtient en utilisant (1.9) :

$$\begin{aligned} Q'_{x^*, \bar{y}}(t) &= \lim_{n \rightarrow \infty} Q^{(n)'}_{x^*, \bar{y}}(t) \\ &= \sum_{x: \bar{x} \in \bar{\mathcal{S}}} \Lambda(x^*, x) \bar{\mathbf{P}}'_{\bar{x}, \bar{y}}(t) \\ &= \sum_{x: \bar{x} \in \bar{\mathcal{S}}} \Lambda(x^*, x) \sum_{\bar{z}} \bar{\mathcal{L}}_{\bar{x}, \bar{z}} \bar{\mathbf{P}}_{\bar{z}, \bar{y}}(t) \\ &= \sum_{\bar{z}} \mathcal{L}_{x^*, z^*}^* \Lambda(z^*, z) \bar{\mathbf{P}}_{\bar{z}, \bar{y}}(t) \\ &= \sum_{z^* \in \mathcal{S}^*} \mathcal{L}_{x^*, z^*}^* Q_{z^*, \bar{y}}(t) \end{aligned}$$

On a alors, $\forall (x^*, \bar{y}) \in \mathcal{S}^* \times \bar{\mathcal{S}}$:

$$\begin{aligned} \frac{d}{dt} [\exp(\mathcal{L}_{x^*}^* t) Q_{x^*, \bar{y}}(t)] &= [Q'_{x^*, \bar{y}}(t) + \mathcal{L}_{x^*}^* Q_{x^*, \bar{y}}(t)] \exp(\mathcal{L}_{x^*}^* t) \\ &= \left[\sum_{z^*} \mathcal{L}_{x^*, z^*}^* Q_{z^*, \bar{y}}(t) + \mathcal{L}_{x^*}^* Q_{x^*, \bar{y}}(t) \right] \exp(\mathcal{L}_{x^*}^* t) \\ &= \sum_{z^* \neq x^*} \mathcal{L}_{x^*, z^*}^* Q_{z^*, \bar{y}}(t) \exp(\mathcal{L}_{x^*}^* t) \end{aligned}$$

Cette somme est finie donc la fonction est continue en t . En appliquant le théorème fondamental de l'analyse et le fait que $Q_{x^*, \bar{y}}(0) = \delta_{x^*, y^*} \Lambda(y^*, y)$, on obtient :

$$\begin{aligned} \exp(\mathcal{L}_{x^*}^* t) Q_{x^*, \bar{y}}(t) &= \delta_{x^*, y^*} \Lambda(x^*, y) + \int_0^t \sum_{z^* \neq x^*} \mathcal{L}_{x^*, z^*}^* Q_{z^*, \bar{y}}(s) \exp(\mathcal{L}_{x^*}^* s) ds \\ Q_{x^*, \bar{y}}(t) &= \delta_{x^*, y^*} \exp(-\mathcal{L}_{x^*}^* t) \Lambda(x^*, y) \\ &\quad + \int_0^t \sum_{z^* \neq x^*} \mathcal{L}_{x^*, z^*}^* Q_{z^*, \bar{y}}(s) \exp(\mathcal{L}_{x^*}^* (s - t)) ds \end{aligned} \quad (1.12)$$

On a exprimé Q comme un point fixe de la fonction f définie par :

$$f: \mathcal{M}_{\mathcal{S}^*, \bar{\mathcal{S}}}(\mathcal{C}(\mathbb{R}^+, \mathbb{R}^+)) \rightarrow \mathcal{M}_{\mathcal{S}^*, \bar{\mathcal{S}}}(\mathcal{C}(\mathbb{R}^+, \mathbb{R}^+))$$

$$f(R)_{x^*, \bar{y}}(t) = \delta_{x^*, y^*} \exp(-\mathcal{L}_{x^*}^* t) \Lambda(x^*, y) + \int_0^t \sum_{z^* \neq x^*} \mathcal{L}_{x^*, z^*}^* R_{z^*, \bar{y}}(s) \exp(\mathcal{L}_{x^*}^* (s - t)) ds \quad (1.13)$$

de sorte que $Q = f(Q)$. On appelle :

- Y^* un processus minimal de générateur \mathcal{L}^*
- J la chaîne des temps de sauts du processus Y^*
- $P_{x^*, y^*}^{*(N)}(t)$ la probabilité pour Y^* de passer de x^* à y^* en temps t , en $N - 1$ sauts maximum :

$$P_{x^*, y^*}^{*(N)}(t) := \mathbb{P}_{x^*}(Y_t^* = y^*, t < J_N)$$

- $Q_{x^*, \bar{y}}^{*(N)}(t) := P_{x^*, y^*}^{*(N)}(t) \Lambda(y^*, y)$

on a alors :

$$\begin{aligned} P_{x^*, y^*}^{*(1)}(t) &= \mathbb{P}_{x^*}(t < J_1) \delta_{x^*, y^*} \\ &= \exp(-\mathcal{L}_{x^*}^* t) \delta_{x^*, y^*} \end{aligned}$$

et donc :

$$Q_{x^*, \bar{y}}^{*(1)}(t) = \delta_{x^*, y^*} \exp(-\mathcal{L}_{x^*}^* t) \Lambda(x^*, y) \leq Q_{x^*, \bar{y}}(t)$$

On va montrer par récurrence que $Q^{*(N)} = f(Q^{*(N-1)})$, où f est la fonction définie en (1.13), pour en déduire l'inégalité

$$Q^* \leq Q \quad (i.e. \quad Q_{x^*, \bar{y}}^*(t) \leq Q_{x^*, \bar{y}}(t), \quad \forall x^*, \bar{y}, t)$$

avec :

$$Q_{x^*, \bar{y}}^*(t) := \mathbf{P}_{x^*, y^*}^*(t) \Lambda(y^*, y)$$

Pour ça, on conditionne par $Y_0^* = x^*$ et on fait une décomposition d'événement en union de deux événements disjoints :

$$\{Y_t^* = y^*, t < J_N\} = \{Y_t^* = y^*, t < J_1\} \cup \{Y_t^* = y^*, J_1 \leq t < J_N\}$$

le premier ayant comme probabilité $P_{x^*, y^*}^{*(1)}(t)$. Pour le second on conditionne par rapport à J_1 (qui suit une loi exponentielle) :

$$\begin{aligned} \mathbb{P}_{x^*}(Y_t^* = y^*, J_1 \leq t < J_N) &= \mathbb{E}_{x^*}(\mathbf{1}_{t < J_N} \mathbf{1}_{Y_t^* = y^*} \mathbf{1}_{J_1 < t}) \\ &= \mathbb{E}_{x^*} \left[\mathbb{E}_{x^*}(\mathbf{1}_{t < J_N} \mathbf{1}_{Y_t^* = y^*} \mid J_1) \mathbf{1}_{J_1 < t} \right] \\ &= \mathbb{E}_{x^*} [\mathbb{P}_{x^*}(t < J_N, Y_t^* = y^* \mid J_1) \mathbf{1}_{J_1 < t}] \end{aligned}$$

puis par X_{J_1} , qui est indépendant de J_1 et vaut $z^* \neq x^*$ avec probabilité $\frac{\mathcal{L}_{x^*, z^*}^*}{\mathcal{L}_{x^*}^*}$, et on utilise l'homogénéité de Y^* :

$$\begin{aligned} &\mathbb{P}_{x^*}(Y_t^* = y^*, J_1 < t < J_N) \\ &= \mathbb{E}_{x^*} [\mathbb{P}_{x^*}(t < J_N, Y_t^* = y^* \mid J_1) \mathbf{1}_{J_1 < t}] \\ &= \int_0^t \mathcal{L}_{x^*}^* e^{-s\mathcal{L}_{x^*}^*} \mathbb{P}_{x^*}(t < J_N, Y_t^* = y^* \mid J_1 = s) ds \\ &= \int_0^t \sum_{z^* \in S^*} \mathcal{L}_{x^*, z^*}^* e^{-s\mathcal{L}_{x^*}^*} \mathbb{P}_{x^*}(t < J_N, Y_t^* = y^* \mid J_1 = s, X_{J_1} = z^*) \mathbb{P}_{x^*}(Y_{J_1}^* = z^*) ds \\ &= \int_0^t \sum_{z^* \neq x^*} \mathcal{L}_{x^*, z^*}^* e^{-s\mathcal{L}_{x^*}^*} \mathbb{P}_{z^*}(Y_{t-s}^* = y^*, t-s \leq J_{N-1}) ds \\ &= \int_0^t \sum_{z^* \neq x^*} \mathcal{L}_{x^*, z^*}^* e^{-s\mathcal{L}_{x^*}^*} P_{z^*, y^*}^{*(N-1)}(t-s) ds \\ &= \int_0^t \sum_{z^* \neq x^*} \mathcal{L}_{x^*, z^*}^* e^{-(t-s)\mathcal{L}_{x^*}^*} P_{z^*, x^*}^{*(N-1)}(s) ds \end{aligned}$$

En regroupant les termes, on obtient :

$$P_{x^*, y^*}^{*(N)}(t) = \exp(-\mathcal{L}_{x^*}^* t) \delta_{x^*, y^*} + \int_0^t \sum_{z^* \neq x^*} \mathcal{L}_{x^*, z^*}^* e^{-(t-s)\mathcal{L}_{x^*}^*} P_{z^*, x^*}^{*(N-1)}(s) ds$$

donc $Q_{x^*, \bar{y}}^{*(N)}(t) = f(Q_{x^*, \bar{y}}^{*(N-1)}(t))$. Puisque f est croissante, on obtient par récurrence $Q_{x^*, \bar{y}}^{*(N)}(t) \leq Q_{x^*, \bar{y}}^{*(N-1)}(t)$. Par convergence monotone, $(Q_{x^*, \bar{y}}^{*(N)}(t))_N$ tend vers $Q_{x^*, \bar{y}}^*(t) \leq Q_{x^*, \bar{y}}(t)$, ce qui conclut la preuve de (1.10).

Lorsque \mathcal{L}^* est non-explosif, Q^* est une matrice stochastique, l'inégalité ci-dessus implique donc l'égalité $Q^* = Q$ (puisque Q est sous-stochastique) et on peut donc montrer (1.11) par une simple récurrence sur k . Malheureusement, cet argument n'est pas valable dans le cas explosif. Le raisonnement qui suit montre le cas non-explosif et s'en sert pour montrer le cas explosif.

Preuve de (1.11) : On rappelle que l'on veut montrer par récurrence sur k :

$$\mathbb{P}(X_{t_0}^* = x_0^*, \dots, X_{t_k}^* = x_k^*, X_{t_k} = x_k) = \mu_0^*(x_0^*) \prod_{i=1}^k \mathbf{P}_{x_{i-1}^*, x_i^*}^*(t_i - t_{i-1}) \Lambda(x_k^*, x_k)$$

T est le temps d'explosion du processus X^* et donc également de (X, X^*) puisqu'on a supposé que X était non-explosif. On rappelle (1.6) :

$$\bar{\mu}_0(x_0, x_0^*) = \mu_0^*(x_0^*)\Lambda(x_0^*, x_0)$$

ce qui nous donne immédiatement l'initialisation de la récurrence pour $k = 0$. Supposons donc maintenant que (1.11) est vrai pour un $k \geq 0$. Soient $x_0^*, \dots, x_{k+1}^* \in \mathcal{S}^*$ et $x_{k+1} \in \mathcal{S}$. On a alors

$$\begin{aligned} & \mathbb{P}(X_{t_0}^* = x_0^*, \dots, X_{t_{k+1}}^* = x_{k+1}^*, X_{t_{k+1}} = x_{k+1}) \\ &= \sum_{x_k \in \mathcal{S}} \mathbb{P}(X_{t_0}^* = x_0^*, \dots, X_{t_k}^* = x_k^*, X_{t_k} = x_k) \\ & \quad \times \mathbb{P}(X_{t_{k+1}}^* = x_{k+1}^*, X_{t_{k+1}} = x_{k+1} \mid X_{t_0}^* = x_0^*, \dots, X_{t_k}^* = x_k^*, X_{t_k} = x_k) \\ &= \mu_0^*(x_0^*) \prod_{i=1}^k \mathbf{P}_{x_{i-1}^*, x_i^*}^*(t_i - t_{i-1}) \sum_{x_k \in \mathcal{S}} \Lambda(x_k^*, x_k) \bar{\mathbf{P}}_{\bar{x}_k, \bar{x}_{k+1}}(t_{k+1} - t_k) \\ &\geq \mu_0^*(x_0^*) \prod_{i=1}^k \mathbf{P}_{x_{i-1}^*, x_i^*}^*(t_i - t_{i-1}) \mathbf{P}_{x_k^*, x_{k+1}^*}^*(t_{k+1} - t_k) \Lambda(x_{k+1}^*, x_{k+1}) \\ &= \mu_0^*(x_0^*) \prod_{i=1}^{k+1} \mathbf{P}_{x_{i-1}^*, x_i^*}^*(t_i - t_{i-1}) \Lambda(x_{k+1}^*, x_{k+1}) \end{aligned}$$

où l'on a appliqué successivement la formule des probabilités totales pour la première égalité, la markovianité de (X, X^*) et l'hypothèse de récurrence pour la seconde, et (1.10) pour l'inégalité.

On a vu que si \mathcal{L}^* est non-explosif, l'inégalité est en fait une égalité et la récurrence est terminée. En particulier, cela montre que $\mathbf{P}^*(\cdot)$ est le semi-groupe de la marginale X^* et donc \mathcal{L}^* est son générateur. Le processus X^* a été construit trajectoriellement à partir de X , indépendamment du caractère explosif ou non de son générateur \mathcal{L}^* . Ses intensités de sauts sont donc données par les coefficients \mathcal{L}_{x^*, y^*}^* , $x^*, y^* \in \mathcal{S}^*$ et on en déduit que \mathcal{L}^* est aussi le générateur du processus minimal X^* lorsque celui-ci est explosif.

Supposons donc maintenant que \mathcal{L}^* est explosif. La quantité de départ et celle d'arrivée dans le calcul précédent sont toutes deux positives pour tout choix des x_i^*, x_{k+1} . Or, dans les deux cas, leur somme en $x_0^*, \dots, x_{k+1}^*, x_{k+1}$ vaut $1 - \mathbb{P}(T \leq t_{k+1})$ (puisque $\mathbf{P}^*(\cdot)$ est le semi-groupe minimal associé à \mathcal{L}^*), ce qui montre l'égalité et termine la récurrence. \square

Corollaire 1.1.4. *Si la condition (1.1) est satisfaite, alors le processus (X, X^*) vérifie aussi :*

$$\mathcal{L}(X_\tau \mid \mathcal{F}_\tau^{X^*}, \tau < T) = \Lambda(X_\tau^*, \cdot) \quad (1.14)$$

pour tout $(\mathcal{F}_t^{X^*})$ -temps d'arrêt τ .

Ici, l'inégalité $\tau < T$ est à voir dans $\bar{\mathbb{R}}$. En particulier, T et τ sont potentiellement infinis, mais l'événement $\{\tau < T\}$ est inclus dans $\{\tau < \infty\}$.

Démonstration. Soit τ un $(\mathcal{F}_t^{X^*})$ -temps d'arrêt, $f : \mathcal{S} \rightarrow \mathbb{R}$ mesurable, continue et bornée et G une variable aléatoire réelle, $\mathcal{F}_\tau^{X^*}$ -mesurable et bornée. Par définition de la loi conditionnelle, on veut montrer :

$$\mathbb{E}(G \mathbf{1}_{\tau < T} f(X_\tau)) = \mathbb{E}(G \mathbf{1}_{\tau < T} \Lambda[f](X_\tau^*)) \quad (1.15)$$

Pour tout $n \in \mathbb{N}$, on pose $\tau_n := \frac{1}{n} \lceil n\tau \rceil$, où $\lceil \cdot \rceil$ désigne la partie entière par excès. On a alors :

$$\begin{aligned} \{\tau_n \leq t\} &= \{\lceil n\tau \rceil \leq nt\} \\ &= \bigcup_{\substack{m \in \mathbb{N} \\ m \leq nt}} \{m = \lceil n\tau \rceil\} \\ &= \bigcup_{\substack{m \in \mathbb{N} \\ m \leq nt}} \{m - 1 < n\tau \leq m\} \\ &= \bigcup_{\substack{m \in \mathbb{N} \\ \frac{m}{n} \leq t}} \left\{ \frac{m-1}{n} < \tau \leq \frac{m}{n} \right\} \in \mathcal{F}_t^{X^*} \end{aligned}$$

et donc τ_n est un $(\mathcal{F}_t^{X^*})$ -temps d'arrêt. τ_n est à valeur dans $\frac{1}{n}\mathbb{N} \cup \{\infty\}$, et $\tau_n = k/n$ si et seulement si $\frac{k-1}{n} < \tau \leq \frac{k}{n}$ donc :

$$G\mathbf{1}_{\tau_n = k/n} = G\mathbf{1}_{\tau \leq k/n} - G\mathbf{1}_{\tau \leq (k-1)/n}$$

qui est $\mathcal{F}_{k/n}^{X^*}$ -mesurable.

On rappelle que l'on a (1.2), soit :

$$\forall t > 0, \quad \mathcal{L}(X_t \mid \mathcal{F}_t^{X^*}, t < T) = \Lambda(X_t^*, \cdot)$$

c'est-à-dire, par définition, pour toute fonction $h : \mathcal{S} \rightarrow \mathbb{R}$ mesurable, continue et bornée, pour toute variable aléatoire réelle $H \mathcal{F}_t^{X^*}$ -mesurable bornée, et tout $t > 0$:

$$\mathbb{E}(H\mathbf{1}_{t < T} h(X_t)) = \mathbb{E}(H\mathbf{1}_{t < T} \Lambda[h](X_t^*))$$

D'où :

$$\begin{aligned} \mathbb{E}(\mathbf{1}_{\tau_n < T} Gf(X_{\tau_n})) &= \mathbb{E}\left(\sum_{k=0}^{\infty} \mathbf{1}_{k/n < T} \mathbf{1}_{\tau_n = k/n} Gf(X_{k/n})\right) \\ &= \sum_{k=0}^{\infty} \mathbb{E}(\mathbf{1}_{k/n < T} \mathbf{1}_{\tau_n = k/n} Gf(X_{k/n})) \\ &= \sum_{k=0}^{\infty} \mathbb{E}(\mathbf{1}_{k/n < T} \mathbf{1}_{\tau_n = k/n} G\Lambda[f](X_{k/n}^*)) \\ &= \mathbb{E}\left(\sum_{k=0}^{\infty} \mathbf{1}_{k/n < T} \mathbf{1}_{\tau_n = k/n} G\Lambda[f](X_{k/n}^*)\right) \\ &= \mathbb{E}(\mathbf{1}_{\tau_n < T} G\Lambda[f](X_{\tau_n}^*)) \end{aligned}$$

Or, presque sûrement, $\lim_{n \rightarrow \infty} f(X_{\tau_n}) = f(X_\tau)$ par continuité à droite des trajectoires de (X, X^*) et continuité de f , $\lim_{n \rightarrow \infty} \mathbf{1}_{\tau_n < T} = \mathbf{1}_{\tau < T}$ et $\lim_{n \rightarrow \infty} \Lambda[f](X_{\tau_n}^*) = \Lambda[f](X_\tau^*)$ par continuité de $\Lambda[f]$, donc on en conclut par convergence dominée :

$$\begin{aligned} \mathbb{E}(G\mathbf{1}_{\tau < T} f(X_\tau)) &= \lim_{n \rightarrow \infty} \mathbb{E}(\mathbf{1}_{\tau_n < T} Gf(X_{\tau_n})) \\ &= \mathbb{E}(G\mathbf{1}_{\tau < T} \Lambda[f](X_\tau^*)) \end{aligned}$$

□

1.2 Dualité avec explosion

Jusqu'à maintenant, nous avons considéré uniquement des processus minimaux, c'est-à-dire que nous ne nous sommes intéressé qu'à ce qui se passe avant le premier temps d'explosion. Cela n'est généralement pas suffisant pour que le processus Λ -lié construit dans la section précédente soit un dual de stationnarité forte pour X .

En effet, prenons l'exemple où X est une chaîne de vie et de mort sur \mathbb{Z} , récurrente positive. On cherche à construire, suivant la méthodologie de [23], un dual X^* dont les états sont les intervalles de \mathbb{Z} , et lié à X par un noyau Λ qui n'est autre que la mesure stationnaire restreinte aux intervalles. Dans ce cas, l'état $[-\infty, +\infty]$ ne peut être atteint par X^* qu'après un nombre infini de sauts. Le temps d'atteinte de cet état étant aussi le temps de stationnarité recherché pour la chaîne X , on aimerait qu'il soit fini et on a donc besoin de s'intéresser à ce qui se passe après un temps d'explosion. Cet exemple sera traité plus en détail dans la section suivante.

Nous allons donc maintenant spécifier un peu plus le cadre de travail. On se place sur un espace topologique dénombrable E , que l'on supposera séparé.

Définition 1.2.1. *On dit qu'un générateur $\hat{\mathcal{L}}$ sur E est stratifié s'il existe $N \in \mathbb{N}$ et un N -uplet $(E_i)_{i \in \llbracket 1, N \rrbracket}$ de parties de E , appelée stratification de $\hat{\mathcal{L}}$, tels que :*

- Les E_i sont deux à deux disjointes.
- $\forall i \in \llbracket 1, N \rrbracket$, $\forall x \in E_i$, si $(Y_t^x)_{0 \leq t \leq \tau}$ est un processus minimal issu de x et associé à $\hat{\mathcal{L}}$, avec τ son temps d'explosion, alors :

$$\mathbb{P}(Y_t^x \in \bigsqcup_{j=1}^i E_j \mid t < \tau) = 1$$

- Si de plus $\tau < +\infty$, alors presque sûrement il existe $j \leq i$ et ε tels que $Y_t^x \in E_j$, $\forall t \in]\tau - \varepsilon, \tau[$, et $\lim_{t \rightarrow \tau} Y_t^x$ existe et appartient à $\bigsqcup_{k=j+1}^N E_k$. Cette limite sera notée Y_τ^x .

Cette définition peut se comprendre ainsi : un processus de saut minimal associé à un générateur stratifié peut sauter d'une strate à une autre uniquement si la strate d'arrivée a un rang inférieur ou égal à la strate de départ. A l'inverse, au temps d'explosion, le processus doit avoir une limite, située dans une strate de rang supérieur (strictement) à celui de la strate où il se trouvait juste avant. Il est alors naturel de définir le processus au temps d'explosion comme étant égal à cette limite, et de le faire ensuite repartir de cette limite. C'est ce qui est fait dans la construction qui suit.

Remarque 1.2.2. *Pour satisfaire la deuxième condition, on doit nécessairement avoir $\hat{\mathcal{L}}_{y,z} = 0$ dès que $y \in E_j$, $z \in E_k$, $k > j$. La Q -matrice $(\hat{\mathcal{L}}_{x,y})_{(x,y) \in E^2}$ est donc triangulaire inférieure par bloc, chaque bloc correspondant à une strate différente.*

Soient $\hat{\mathcal{L}}$ un générateur stratifié sur E , $(E_i)_{i \in \llbracket 1, N \rrbracket}$ une stratification de $\hat{\mathcal{L}}$, avec $N \in \mathbb{N}$, et μ une mesure de probabilités sur E . On construit un processus non-minimal Z de générateur $\hat{\mathcal{L}}$ et de loi initial μ de la manière suivante :

- On pose $T_0 = 0$ et $Z_{T_0} = x$ avec probabilité $\mu(x)$.
- Par récurrence sur $i \geq 0$, on suppose Z construit jusqu'au temps T_i . Si $T_i = \infty$, la construction est terminée. Sinon, on simule un processus minimal $Y^{(i)}$ associé à $\hat{\mathcal{L}}$ et issu de Z_{T_i} , indépendamment de $Z_{[0, T_i]}$ sachant Z_{T_i} . Soit T'_{i+1} son premier temps d'explosion, on pose :

$$\begin{aligned} Z_{T_i+t} &= Y_t^{(i)}, \quad \forall t < T'_{i+1} \\ T_{i+1} &= T_i + T'_{i+1} \\ Z_{T_{i+1}} &= Y_{T'_{i+1}}^{(i)} \quad \text{si } T'_{i+1} < \infty \end{aligned}$$

Définition 1.2.3. On appelle processus de saut (resp. semi-groupe) stratifié associé à $(\nu, \hat{\mathcal{L}})$ un processus Z (resp. son semi-groupe) obtenu par la construction précédente. Ce processus est défini jusqu'à un temps dit de double explosion, possiblement fini :

$$T := \lim_{i \rightarrow +\infty} T_i$$

Comme pour les processus minimaux, on le prolonge après le temps de double explosion en posant $X_t = \Delta$ pour tout $t \geq T$, où Δ est un point cimetièrre n'appartenant pas à E .

Par construction, un tel processus satisfait la propriété de Markov forte, et d'après Anderson [3], Proposition 1.2.7, sa fonction de transition satisfait l'équation backward pour $\bar{\mathcal{L}}$ (mais pas forcément l'équation forward).

Remarque 1.2.4. On pourrait également définir le processus par passage à la limite après le temps de double explosion, jusqu'à un temps de triple explosion, et ainsi de suite jusqu'à des temps de multi-explosion, sous réserve que la limite du processus existe à chaque fois. On verra que pour l'étude que l'on fait ici, cela n'est pas nécessaire.

A chaque temps d'explosion, le processus passe d'un espace E_i à un espace E_j , avec $j > i$, et une telle transition n'est possible qu'après une explosion.

La proposition suivante justifie la construction que l'on vient de faire :

Proposition 1.2.5. Soit $\hat{\mathcal{L}}$ un générateur stratifié sur E , Z un processus de saut stratifié associé à $\hat{\mathcal{L}}$ et T_i , $i \in \mathbb{N}$ ses temps d'explosion. Pour toute fonction $f \in \mathcal{B}(E)$, on définit le processus :

$$M_i^f = f(Z_t) - f(Z_0) - \int_0^t \hat{\mathcal{L}}f(Z_s) ds$$

Si f est continue et $\mathcal{L}f$ est bornée, alors pour tout $i \in \mathbb{N}$, $(M_{t \wedge T_i}^f)_{t \geq 0}$ est une martingale relativement à la filtration $(\mathcal{F}_t^Z)_{t > 0}$. En particulier, si le temps de double-explosion est infini presque sûrement, alors M^f est une martingale.

Démonstration. On prouve que les processus $(M_{t \wedge T_i}^f)_{t \geq 0}$ sont des martingales par récurrence sur i , avec T_i le i -ème temps d'explosion de Z , pour $i \in \mathbb{N}$. Pour $i = 0$, c'est trivial ($T_0 = 0$). Supposons donc que ce soit vrai pour $i \geq 0$. On pose :

$$Y_t = \begin{cases} Z_{T_i+t} & \text{si } T_i + t < T_{i+1} \\ \Delta & \text{sinon} \end{cases}$$

Alors Y est un processus minimal sur E associé au générateur $\hat{\mathcal{L}}$. Si on appelle τ_n son n -ième temps de saut, $n \in \mathbb{N}$, on a alors, pour tout $0 < s < t$:

$$\begin{aligned} \mathbb{E}(M_{T_i+t \wedge \tau_n}^f - M_{T_i+s \wedge \tau_n}^f \mid \mathcal{F}_{T_i+s \wedge \tau_n}^Z) &= \mathbb{E}(M_{T_i+t \wedge \tau_n}^f - M_{T_i+s \wedge \tau_n}^f \mid Z_{T_i+s \wedge \tau_n}) \\ &= \mathbb{E}(f(Y_{t \wedge \tau_n}) - f(Y_{s \wedge \tau_n}) - \int_{s \wedge \tau_n}^{t \wedge \tau_n} \hat{\mathcal{L}}f(Y_r) dr \mid Y_{s \wedge \tau_n}) \\ &= 0 \end{aligned}$$

car Y est solution du problème de martingale pour $\hat{\mathcal{L}}$.

En faisant tendre n vers l'infini, on obtient par le théorème de convergence dominée que le processus $(M_{(T_i+t) \wedge T_{i+1}}^f)_{t \geq 0}$ est une martingale. On a alors :

$$\begin{aligned}
\mathbb{E}(M_{t \wedge T_{i+1}}^f - M_{s \wedge T_{i+1}}^f \mid \mathcal{F}_s) &= \mathbb{E}(M_{t \wedge T_{i+1}} - M_{s \wedge T_{i+1}}^f \mid \mathcal{F}_s) \mathbf{1}_{s < T_i} \\
&= \mathbb{E}(\mathbb{E}(M_{t \wedge T_{i+1}}^f - M_{s \wedge T_{i+1}}^f \mid \mathcal{F}_{T_i}) \mid \mathcal{F}_s) \mathbf{1}_{s < T_i} \\
&= \mathbb{E}(\mathbb{E}(M_{t \wedge T_{i+1}}^f \mid \mathcal{F}_{T_i}) - M_{s \wedge T_{i+1}}^f \mid \mathcal{F}_s) \mathbf{1}_{s < T_i} \\
&= \mathbb{E}(M_{T_i}^f \mathbf{1}_{t > T_i} + M_t^f \mathbf{1}_{t < T_i} - M_{s \wedge T_{i+1}}^f \mid \mathcal{F}_s) \mathbf{1}_{s < T_i} \\
&= \mathbb{E}(M_{t \wedge T_i}^f - M_{s \wedge T_i}^f \mid \mathcal{F}_s) \mathbf{1}_{s < T_i} \\
&= \mathbb{E}(M_{t \wedge T_i}^f - M_{s \wedge T_i}^f \mid \mathcal{F}_s) \mathbf{1}_{s < T_i} \\
&= 0
\end{aligned}$$

Par convergence dominée encore une fois, avec $i \rightarrow \infty$, on obtient que M^f est une martingale si $T_i \rightarrow +\infty$. \square

On suppose dorénavant que :

- \mathcal{S} est un espace muni de la topologie discrète.
- Il existe une stratification $(\mathcal{S}_i^*)_{i \in \llbracket 1, N \rrbracket}$, $N \in \mathbb{N}$ de \mathcal{L}^* sur \mathcal{S}^* . On note $P^*(\cdot)$ le semi-groupe stratifié associé à \mathcal{L}^* .
- Le temps de double-explosion d'un processus stratifié associé à \mathcal{L}^* est infini presque sûrement. La matrice $(P_{x^*, y^*}^*(t))_{x^*, y^* \in \mathcal{S}^*}$ est alors stochastique pour tout $t > 0$.
- Pour tout $x \in \mathcal{S}$, la fonction $\mathcal{S}^* \ni x^* \mapsto \Lambda(x^*, x)$ est continue.

Lemme 1.2.6. *Sous ces hypothèses, $(\mathcal{S} \times \mathcal{S}_i^*)_{i \in \llbracket 1, N \rrbracket}$ est une stratification de $\bar{\mathcal{L}}$ sur $\mathcal{S} \times \mathcal{S}^*$ muni de la topologie produit.*

Démonstration. Ceci découle trivialement du fait que \mathcal{L} est non-explosif et que les marginales d'un processus minimal (X, X^*) de générateur $\bar{\mathcal{L}}$ ont pour générateurs \mathcal{L} et \mathcal{L}^* respectivement. \square

Le théorème suivant sera un outil indispensable pour construire des duals de stationnarité forte sur des espaces dénombrables.

Théorème 1.2.7. *Sous les hypothèses précédentes, les trois conditions suivantes sont équivalentes :*

- (i) *Il existe un processus stratifié $(X_t^*)_{t > 0}$ sur \mathcal{S}^* de générateur \mathcal{L}^* et de lois initiale μ_0^* , tel que $\forall t \geq 0$:*

$$\mathcal{L}(X_t \mid \mathcal{F}_t^{X^*}) = \Lambda(X_t^*, \cdot) \quad (1.16)$$

$$\mathcal{L}(X_t^* \mid X) = \mathcal{L}(X_t^* \mid \mathcal{F}_t^X) \quad (1.17)$$

$$(ii) \begin{cases} \mu_0^* \Lambda = \mu_0 \\ P^*(t) \Lambda = \Lambda P(t) \quad \forall t \geq 0 \end{cases}$$

$$(iii) \begin{cases} \mu_0^* \Lambda = \mu_0 \\ \mathcal{L}^* \Lambda = \Lambda \mathcal{L} \end{cases}$$

Démonstration. Nous allons montrer (i) \Rightarrow (ii) \Rightarrow (iii) \Rightarrow (i).

(i)⇒(ii) La preuve n'est qu'un calcul de probabilité conditionnelle par deux moyens différents. En effet, soit $t \geq 0$, $x^* \in \mathcal{S}^*$, $y \in \mathcal{S}$. D'un côté, d'après (1.16), X_t est indépendant de X_0^* conditionnellement à X_t^* et donc on a :

$$\begin{aligned} \mathbb{P}(X_t = y \mid X_0^* = x^*) &= \sum_{y^* \in \mathcal{S}^*} \mathbb{P}(X_t^* = y^* \mid X_0^* = x^*) \mathbb{P}(X_t = y \mid X_0^* = x^*, X_t^* = y^*) \\ &= \sum_{y^* \in \mathcal{S}^*} P_{x^*, y^*}^*(t) \mathbb{P}(X_t = y \mid X_t^* = y^*) \\ &= \sum_{y^* \in \mathcal{S}^*} P_{x^*, y^*}^*(t) \Lambda(y^*, y) \end{aligned}$$

et d'autre part, X_0^* est indépendant de X_t conditionnellement à X_0 d'après (1.17), d'où :

$$\begin{aligned} \mathbb{P}(X_t = y \mid X_0^* = x^*) &= \sum_{x \in \mathcal{S}} \mathbb{P}(X_0 = x \mid X_0^* = x^*) \mathbb{P}(X_t = y \mid X_0^* = x^*, X_0 = x) \\ &= \sum_{x \in \mathcal{S}} \Lambda(x^*, x) \mathbb{P}(X_t = y \mid X_0 = x) \\ &= \sum_{x \in \mathcal{S}} \Lambda(x^*, x) \mathbf{P}_{x,y}(t) \end{aligned}$$

ce qui montre que $P^* \Lambda = \Lambda \mathbf{P}$. De plus, on a :

$$\begin{aligned} \mu_0(y) &= \mathbb{P}(X_0 = y) \\ &= \sum_{y^* \in \mathcal{S}^*} \mathbb{P}(X_0^* = y^*) \mathbb{P}(X_0 = y \mid X_0^* = y^*) \\ &= \sum_{y^* \in \mathcal{S}^*} \mu_0^*(y^*) \Lambda(y^*, y) \end{aligned}$$

donc $\mu_0 = \mu_0^* \Lambda$.

(ii)⇒(iii) La preuve est quasiment identique à celle de Fill ([32]), Proposition 2. Cependant, certaines étapes méritent une justification supplémentaire, puisque les hypothèses de départ sont légèrement différentes, c'est pourquoi cette preuve est reproduite ici.

On pose $Q(t) = \Lambda \mathbf{P}(t)$. Comme $\mathbf{P}'(t) = \mathbf{P}(t) \mathcal{L}$, on a, $\forall \bar{x} \in \bar{\mathcal{S}}, \forall y \in \mathcal{S}$:

$$\begin{aligned} |\Lambda(x^*, x) \mathbf{P}'_{x,y}(t)| &= \left| \sum_{z \in \mathcal{S}} \Lambda(x^*, x) \mathbf{P}_{x,z}(t) \mathcal{L}_{z,y} \right| \\ &\leq \sum_{z \in \mathcal{S}} \Lambda(x^*, x) |\mathcal{L}_{z,y}| \end{aligned}$$

et cette dernière somme est finie d'après l'hypothèse de μ -intégrabilité de la diagonale de \mathcal{L} . On utilise cette fois l'équation backward de Kolmogorov pour obtenir :

$$\begin{aligned} Q'_{x^*, y}(t) &= \sum_{x \in \mathcal{S}} \Lambda(x^*, x) \mathbf{P}'_{x,y}(t) \\ &= \sum_{x \in \mathcal{S}} \Lambda(x^*, x) \sum_{z \in \mathcal{S}} \mathcal{L}_{x,z} \mathbf{P}_{z,y}(t) \end{aligned}$$

En particulier, $Q'(0) = \Lambda \mathcal{L}$

D'après (ii), on a aussi $Q(t) \equiv Q^*(t) := P^*(t)\Lambda$, et P^* est solution de l'équation backward intégrale pour \mathcal{L}^* (cf Anderson [3], chapitre 2, proposition 1.1) :

$$P_{x^*,y^*}^*(t) = \delta_{x^*,y^*} e^{-\mathcal{L}_{x^*}^* t} + \int_0^t \sum_{z^* \neq x^*} \mathcal{L}_{x^*,z^*}^* P_{z^*,y^*}^*(s) e^{-(t-s)\mathcal{L}_{x^*}^*} ds$$

Donc en utilisant le théorème de Fubini-Tonelli :

$$Q_{x^*,y}^*(t) = \Lambda(x^*, y) e^{-\mathcal{L}_{x^*}^* t} + \int_0^t \sum_{z^* \neq x^*} \mathcal{L}_{x^*,z^*}^* Q_{z^*,y}^*(s) e^{-(t-s)\mathcal{L}_{x^*}^*} ds$$

Pour tout x^*, y , la fonction $t \rightarrow Q_{x^*,y}^*(t)$ est continue (car $Q^*(t) = Q(t) = \Lambda P(t)$ est dérivable en t donc continue), et le nombre de termes dans la somme est finie par hypothèse donc l'intégrande est continue. Par le théorème fondamental de l'analyse, on obtient donc d'une part :

$$\frac{d}{dt} \left(e^{\mathcal{L}_{x^*}^* t} Q_{x^*,y}^*(t) \right) = \left(\sum_{z^* \neq x^*} \mathcal{L}_{x^*,z^*}^* Q_{z^*,y}^*(t) \right) e^{t\mathcal{L}_{x^*}^*}$$

et d'autre part, par dérivation d'un produit de fonction :

$$\frac{d}{dt} \left(e^{\mathcal{L}_{x^*}^* t} Q_{x^*,y}^*(t) \right) = \left(\mathcal{L}_{x^*}^* Q_{x^*,y}^*(t) + Q_{x^*,y}^{\prime}(t) \right) e^{\mathcal{L}_{x^*}^* t}$$

d'où :

$$Q_{x^*,y}^{\prime}(t) = \sum_{z^*} \sum_{y^* \in \mathcal{S}^*} \mathcal{L}_{x^*,z^*}^* P_{z^*,y^*}^*(t) \Lambda(y^*, y)$$

D'où $Q^{\prime}(0) = \mathcal{L}^* \Lambda$, or $Q^* \equiv Q$ donc $\mathcal{L}^* \Lambda = \Lambda \mathcal{L}$.

(iii) \Rightarrow (i) Soit X^* un processus de saut stratifié tel que (X, X^*) ait pour générateur $\bar{\mathcal{L}}$ et pour loi initiale $\bar{\mu}$ (on rappelle que $\bar{\mu}(x_0, x_0^*) = \mu_0^*(x_0^*) \Lambda(x_0^*, x_0)$). Par construction, un tel processus vérifie (1.17). Montrons qu'il vérifie (1.16).

On pose $T_0 = 0$ et $\forall i \geq 0$ on note T_{i+1}^k le k -ième temps de saut de (X, X^*) suivant T_i , $k \in \mathbb{N}$ et :

$$T_{i+1} := \lim_{k \rightarrow \infty} T_{i+1}^k$$

le $(i+1)$ -ième temps d'explosion. D'après les hypothèses,

$$\lim_{i \rightarrow +\infty} T_i = +\infty$$

On prend comme convention $T_j^k = +\infty$, $\forall k, \forall j > i$ si $T_i = +\infty$.

On commence par prouver que le processus (X, X^*) est continu presque sûrement à chaque temps d'explosion T_i . En effet, X^* est continu en T_i par définition d'un processus stratifié, et X est discontinu seulement quand il saute. Appelons σ_n les temps de sauts de X , supposons par récurrence n que (X, X^*) est continu à chaque temps d'explosion jusqu'à σ_n , et appelons T_i le dernier temps d'explosion avant σ_n . Soit $\tilde{\mathcal{L}}$ un générateur sur \mathcal{S}^* défini par $\tilde{\mathcal{L}}_{x^*,y^*} = \bar{\mathcal{L}}_{(X_{\sigma_n}, x^*), (X_{\sigma_n}, y^*)}$ et Y un processus stratifié partant de $X_{\sigma_n}^*$ et de générateur $\tilde{\mathcal{L}}$. En regardant la construction de X^* que l'on a fait dans la preuve du théorème 1.1.1, on s'aperçoit que $(X_{\sigma_n+t}^*)_t$ est égal en loi à Y jusqu'au temps σ_{n+1} . En notant τ_k les temps de saut de Y avant sa première explosion, on obtient :

$$\begin{aligned}
\mathbb{P}(T_{i+1} = \sigma_{n+1} \mid \mathcal{F}_{\sigma_n}^{\bar{X}}) &= \mathbb{P}(\sigma_n + \tau_k \xrightarrow{k \rightarrow \infty} \sigma_{n+1} \mid \mathcal{F}_{\sigma_n}^{\bar{X}}) \\
&= \mathbb{P}(\sigma_n + \tau_1 + \sum_{k=1}^{\infty} (\tau_{k+1} - \tau_k) = \sigma_{n+1} \mid \mathcal{F}_{\sigma_n}^{\bar{X}}) \\
&= 0
\end{aligned}$$

car τ_1 est une variable aléatoire continue, indépendante de $\tau_{k+1} - \tau_k$, σ_n et σ_{n+1} conditionnellement à $\mathcal{F}_{\sigma_n}^{\bar{X}}$. En procédant de la même manière (par un second niveau de récurrence), en commençant de T_{i+j} , $j \geq 0$, si ce dernier est plus petit que σ_{n+1} , on termine la récurrence. En conclusion, on obtient $\sigma_n \neq T_m$, $\forall m, n \geq 0$.

On va maintenant montrer par récurrence sur i que pour tout $0 = t_1 < \dots < t_n < \infty$ et tout $(x_1^*, \dots, x_n^*, x_n) \in \{\mathcal{S}^*\}^n \times \mathcal{S}$, $n \in \mathbb{N}$, on a :

$$\begin{aligned}
&\mathbb{P}(X_{t_1}^* = x_1^*, \dots, X_{T_i}^* = x_n^*, X_{T_i} = x_n \mid t_{n-1} < T_i \leq t_n) \\
&= \mathbb{P}(X_{t_1}^* = x_1^*, \dots, X_{T_i}^* = x_n^* \mid t_{n-1} < T_i \leq t_n) \Lambda(x_n^*, x_n)
\end{aligned} \tag{1.18}$$

et :

$$\begin{aligned}
&\mathbb{P}(X_{t_1}^* = x_1^*, \dots, X_{t_n}^* = x_n^*, X_{t_n} = x_n \mid T_{i+1} > t_n) \\
&= \mathbb{P}(X_{t_1}^* = x_1^*, \dots, X_{t_n}^* = x_n^* \mid T_{i+1} > t_n) \Lambda(x_n^*, x_n)
\end{aligned} \tag{1.19}$$

Pour $i = 0$, on considère simplement un processus minimal associé à $\bar{\mathcal{L}}$ et la récurrence est initialisée par la loi initiale de (X, X^*) (pour (1.18)) et le théorème 1.1.1 (pour (1.19)). Supposons donc la récurrence vraie pour $i - 1 \geq 0$. De (1.19) on déduit :

$$\mathcal{L}(X_t \mid \mathcal{F}_t^{X^*}, T_i > t) = \Lambda(X_{T_i}^*, \cdot), \quad \forall t > 0$$

et de la même manière que pour le corollaire 1.1.4 :

$$\mathcal{L}(X_S \mid \mathcal{F}_S^{X^*}) = \Lambda(X_S^*, \cdot)$$

pour tout temps d'arrêt $S < T_i$ presque sûrement.

On pose :

$$\mathcal{A}_i := \bigvee_{k \geq 0} \mathcal{F}_{T_i^k}^{X^*}$$

la tribu engendrée par la réunion des $\mathcal{F}_{T_i^k}^{X^*}$.

Soit G une variable aléatoire réelle \mathcal{A}_i -mesurable bornée et $f : \mathcal{S} \rightarrow \mathbb{R}$ continue bornée. D'après le théorème de convergence des martingales fermées, on a $\mathbb{E}(G \mid \mathcal{F}_{T_i^k}^{X^*}) \rightarrow G$ p.s. et donc en utilisant le théorème de convergence dominée :

$$\begin{aligned}
\mathbb{E}(Gf(X_{T_i})) &= \lim_{k \rightarrow \infty} \mathbb{E}(\mathbb{E}(G \mid \mathcal{F}_{T_i^k}^{X^*}) f(X_{T_i^k})) \\
&= \lim_{k \rightarrow \infty} \mathbb{E}(\mathbb{E}(G \mid \mathcal{F}_{T_i^k}^{X^*}) \Lambda[f](X_{T_i^k}^*)) \\
&= \mathbb{E}(G \Lambda[f](X_{T_i}^*))
\end{aligned}$$

Maintenant, si on se donne $0 = t_1 < \dots < t_n < \infty$ et $(x_1^*, \dots, x_n^*) \in \{\mathcal{S}^*\}^n$, $n \in \mathbb{N}$ alors il existe une suite décroissante $(V_m)_{m \in \mathbb{N}}$ de voisinages de x_n^* telle que $\bigcap_m V_m = \{x_n^*\}$ et on a :

$$\begin{aligned}
& \{X_{t_1}^* = x_1^*, \dots, X_{T_i}^* = x_n^*, T_i \in]t_{n-1}, t_n]\} \\
&= \{X_{t_1}^* = x_1^*, \dots, \lim_{k \rightarrow \infty} X_{T_i^k}^* = x_n^*, \lim_{k \rightarrow \infty} T_i^k \in]t_{n-1}, t_n]\} \\
&= \bigcap_{m \geq 0} \bigcup_{K \geq 0} \bigcap_{k \geq K} \{X_{t_1}^* = x_1^*, \dots, X_{T_i^k}^* \in V_m, T_i^k \in]t_{n-1}, t_n]\} \\
&= \bigcap_{m \leq 0} \bigcup_{K \geq 0} \bigcap_{k \geq K} \{X_{t_1}^* = x_1^*, \dots, X_{T_i^k}^* \in V_m, T_i^k \in]t_{n-1}, t_n]\}
\end{aligned}$$

c'est-à-dire :

$$\{X_{t_1}^* = x_1^*, \dots, X_{t_n}^* = x_n^*, T_i = t_n\} \in \mathcal{A}_i$$

et on en déduit (1.18). Autrement dit, la relation d'entrelacement est conservée « au temps d'explosion T_i ».

Par construction, le processus $((X, X^*)_{T_i+t})_{t \geq 0}$ a la loi d'un processus minimal (Y, Y^*) de générateur $\bar{\mathcal{L}}$ jusqu'au temps d'explosion T_{i+1} . Ce processus est indépendant de $((X, X^*)_{T_i \wedge t})_{t \geq 0}$ conditionnellement à $(X, X^*)_{T_i}$ et grâce à ce qui précède, on sait aussi que la loi initiale de ce processus minimal vérifie :

$$\mathbb{P}(Y_0 = x_0 \mid Y_0^* = x_0^*) = \Lambda(x_0^*, x_0)$$

pour tout $(x_0, x_0^*) \in \bar{\mathcal{S}}$. Du théorème 1.1.1 on déduit que pour tout $0 = t_1 < \dots < t_n < \infty$ et tout $(x_1^*, \dots, x_n^*, x_n) \in \{\mathcal{S}^*\}^n \times \mathcal{S}$, $n \in \mathbb{N}$:

$$\begin{aligned}
& \mathbb{P}(X_{t_1}^* = x_1^*, \dots, X_{t_n}^* = x_n^*, X_{t_n} = x_n \mid T_{i+1} > t_n, T_i \leq t_1) \\
&= \mathbb{P}(Y_{t_1-T_i}^* = x_1^*, \dots, Y_{t_n-T_i}^* = x_n^*, Y_{t_n-T_i} = x_n \mid T^Y > t_n - T_i, T_i \leq t_1) \\
&= \mathbb{P}(Y_{t_1-T_i}^* = x_1^*, \dots, Y_{t_n-T_i}^* = x_n^* \mid T^Y > t_n - T_i, T_i \leq t_1) \Lambda(x_n^*, x_n) \\
&= \mathbb{P}(X_{t_1}^* = x_1^*, \dots, X_{t_n}^* = x_n^* \mid T_{i+1} > t_n) \Lambda(x_n^*, x_n)
\end{aligned}$$

où T^Y est le premier temps d'explosion du processus (Y, Y^*) . On montre ainsi que l'entrelacement- est conservée après le temps d'explosion T_i . En effet, si $k \in \llbracket 1, n-1 \rrbracket$ et $\bar{x} \in \bar{\mathcal{S}}$, alors en utilisant :

- i. le caractère markovien de (X, X^*)
- ii. l'indépendance de X_{T_i} et $X_{[0, T_i]}^*$ conditionnellement à $X_{T_i}^*$
- iii. la loi d'un processus minimal $(Y, Y^*) := (X, X^*)_{t+}$ de générateur $\bar{\mathcal{L}}$, donnée par le théorème 1.1.1

on obtient :

$$\begin{aligned}
& \mathbb{P}(X_{t_1}^* = x_1^*, \dots, X_{t_n}^* = x_n^*, X_{t_n} = x_n \mid T_i \in]t_k, t_{k+1}], T_{i+1} > t_n) \\
= & \sum_{\bar{x} \in \bar{\mathcal{S}}} \mathbb{P}(X_{t_1}^* = x_1^*, \dots, X_{t_n}^* = x_n^*, X_{t_n} = x_n, \bar{X}_{T_i} = \bar{x} \mid T_i \in]t_k, t_{k+1}], T_{i+1} > t_n) \\
\stackrel{i.}{=} & \sum_{\bar{x} \in \bar{\mathcal{S}}} \mathbb{P}(X_{t_1}^* = x_1^*, \dots, X_{t_k}^* = x_k^*, \bar{X}_{T_i} = \bar{x} \mid T_i \in]t_k, t_{k+1}], T_{i+1} > t_n) \\
& \times \mathbb{P}(X_{t_{k+1}}^* = x_{k+1}^*, \dots, X_{t_n}^* = x_n^*, X_{t_n} = x_n \mid \bar{X}_{T_i} = \bar{x}, T_{i+1} > t_n, T_i \in]t_k, t_{k+1}]) \\
\stackrel{ii.}{=} & \sum_{\bar{x} \in \bar{\mathcal{S}}} \mathbb{P}(X_{t_1}^* = x_1^*, \dots, X_{t_k}^* = x_k^*, X_{T_i}^* = x^* \mid T_i \in]t_k, t_{k+1}], T_{i+1} > t_n) \\
& \times \mathbb{P}(X_{T_i} = x \mid X_{T_i}^* = x^*, T_{i+1} > t_n, T_i \in]t_k, t_{k+1}]) \\
& \times \frac{\mathbb{P}(X_{T_i} = x, X_{t_{k+1}}^* = x_{k+1}^*, \dots, X_{t_n} = x_n \mid X_{T_i}^* = x^*, T_{i+1} > t_n, T_i \in]t_k, t_{k+1}])}{\mathbb{P}(X_{T_i} = x \mid X_{T_i}^* = x^*, T_{i+1} > t_n, T_i \in]t_k, t_{k+1}])} \\
= & \sum_{\bar{x} \in \bar{\mathcal{S}}} \mathbb{P}(X_{t_1}^* = x_1^*, \dots, X_{t_k}^* = x_k^*, X_{T_i}^* = x^* \mid T_i \in]t_k, t_{k+1}], T_{i+1} > t_n) \\
& \times \mathbb{P}(X_{T_i} = x, X_{t_{k+1}}^* = x_{k+1}^*, \dots, X_{t_n} = x_n \mid X_{T_i}^* = x^*, T_{i+1} > t_n, T_i \in]t_k, t_{k+1}]) \\
= & \sum_{x^* \in \mathcal{S}^*} \sum_{x \in \mathcal{S}} \mathbb{P}(X_{t_1}^* = x_1^*, \dots, X_{t_k}^* = x_k^*, X_{T_i}^* = x^* \mid T_i \in]t_k, t_{k+1}], T_{i+1} > t_n) \\
& \times \mathbb{P}(X_{T_i} = x, X_{t_{k+1}}^* = x_{k+1}^*, \dots, X_{t_n} = x_n \mid X_{T_i}^* = x^*, T_{i+1} > t_n, T_i \in]t_k, t_{k+1}]) \\
= & \sum_{x^* \in \mathcal{S}^*} \mathbb{P}(X_{t_1}^* = x_1^*, \dots, X_{t_k}^* = x_k^*, X_{T_i}^* = x^* \mid T_i \in]t_k, t_{k+1}], T_{i+1} > t_n) \\
& \times \mathbb{P}(X_{t_{k+1}}^* = x_{k+1}^*, \dots, X_{t_n} = x_n \mid X_{T_i}^* = x^*, T_{i+1} > t_n, T_i \in]t_k, t_{k+1}]) \\
= & \sum_{x^* \in \mathcal{S}^*} \mathbb{P}(X_{t_1}^* = x_1^*, \dots, X_{t_k}^* = x_k^*, X_{T_i}^* = x^* \mid T_i \in]t_k, t_{k+1}], T_{i+1} > t_n) \\
& \times \int_{t_k}^{t_{k+1}} \mathbb{P}(X_{t_{k+1}}^* = x_{k+1}^*, \dots, X_{t_n} = x_n \mid X_{T_i}^* = x^*, T_{i+1} > t_n, T_i = t) dt \\
\stackrel{iii.}{=} & \sum_{x^* \in \mathcal{S}^*} \mathbb{P}(X_{t_1}^* = x_1^*, \dots, X_{t_k}^* = x_k^*, X_{T_i}^* = x^* \mid T_i \in]t_k, t_{k+1}], T_{i+1} > t_n) \\
& \times \mathbb{P}(X_{t_{k+1}}^* = x_{k+1}^*, \dots, X_{t_n}^* = x_n^* \mid X_{T_i}^* = x^*, T_{i+1} > t_n, T_i \in]t_k, t_{k+1}]) \Lambda(x_n^*, x_n)
\end{aligned}$$

Puis en gardant le terme $\Lambda(x_n^*, x_n)$ de côté, on refait les mêmes étapes dans l'autre sens :

$$\begin{aligned}
&= \Lambda(x_n^*, x_n) \sum_{\bar{x} \in \mathcal{S}} \mathbb{P}(X_{t_1}^* = x_1^*, \dots, X_{t_k}^* = x_k^*, X_{T_i}^* = x^* \mid T_i \in]t_k, t_{k+1}], T_{i+1} > t_n) \\
&\quad \times \mathbb{P}(X_{T_i} = x, X_{t_{k+1}}^* = x_1^*, \dots, X_{t_n}^* = x_n^* \mid X_{T_i}^* = x^*, T_{i+1} > t_n, T_i \in]t_k, t_{k+1}]) \\
&= \Lambda(x_n^*, x_n) \sum_{\bar{x} \in \mathcal{S}} \mathbb{P}(X_{t_1}^* = x_1^*, \dots, X_{t_k}^* = x_k^*, X_{T_i}^* = x^* \mid T_i \in]t_k, t_{k+1}], T_{i+1} > t_n) \\
&\quad \times \mathbb{P}(X_{T_i} = x \mid X_{T_i}^* = x^*, T_{i+1} > t_n, T_i \in]t_k, t_{k+1}]) \\
&\quad \times \frac{\mathbb{P}(X_{T_i} = x, X_{t_{k+1}}^* = x_1^*, \dots, X_{t_n}^* = x_n^* \mid X_{T_i}^* = x^*, T_{i+1} > t_n, T_i \in]t_k, t_{k+1}])}{\mathbb{P}(X_{T_i} = x \mid X_{T_i}^* = x^*, T_{i+1} > t_n, T_i \in]t_k, t_{k+1}])} \\
&= \Lambda(x_n^*, x_n) \sum_{\bar{x} \in \mathcal{S}} \mathbb{P}(X_{t_1}^* = x_1^*, \dots, X_{t_k}^* = x_k^*, \bar{X}_{T_i} = \bar{x} \mid T_i \in]t_k, t_{k+1}], T_{i+1} > t_n) \\
&\quad \times \mathbb{P}(X_{t_{k+1}}^* = x_1^*, \dots, X_{t_n}^* = x_n^* \mid \bar{X}_{T_i} = \bar{x}, T_{i+1} > t_n, T_i \in]t_k, t_{k+1}]) \\
&= \Lambda(x_n^*, x_n) \sum_{\bar{x} \in \mathcal{S}} \mathbb{P}(X_{t_1}^* = x_1^*, \dots, X_{t_n}^* = x_n^*, \bar{X}_{T_i} = \bar{x} \mid T_i \in]t_k, t_{k+1}], T_{i+1} > t_n) \\
&= \mathbb{P}(X_{t_1}^* = x_1^*, \dots, X_{t_n}^* = x_n^* \mid T_i \in]t_k, t_{k+1}], T_{i+1} > t_n) \Lambda(x_n^*, x_n)
\end{aligned}$$

et donc il n'y a plus qu'à sommer sur les intervalles possibles pour T_i :

$$\begin{aligned}
&\mathbb{P}(X_{t_1}^* = x_1^*, \dots, X_{t_n}^* = x_n^*, X_{t_n} = x_n \mid T_{i+1} > t_n) \\
&= \mathbb{P}(X_{t_1}^* = x_1^*, \dots, X_{t_n}^* = x_n^*, X_{t_n} = x_n \mid T_{i+1} > t_n, T_i > t_n) \mathbb{P}(T_i > t_n \mid T_{i+1} > t_n) \\
&\quad + \sum_{k=1}^{n-1} \mathbb{P}(X_{t_1}^* = x_1^*, \dots, X_{t_n}^* = x_n^*, X_{t_n} = x_n, t_k < T_i \leq t_{k+1} \mid T_{i+1} > t_n) \\
&= \mathbb{P}(X_{t_1}^* = x_1^*, \dots, X_{t_n}^* = x_n^* \mid T_i > t_n) \Lambda(x_n^*, x_n) \mathbb{P}(T_i > t_n \mid T_{i+1} > t_n) \\
&\quad + \sum_{k=1}^{n-1} \mathbb{P}(X_{t_1}^* = x_1^*, \dots, X_{t_n}^* = x_n^*, t_k < T_i \leq t_{k+1} \mid T_{i+1} > t_n) \Lambda(x_n^*, x_n) \\
&= \mathbb{P}(X_{t_1}^* = x_1^*, \dots, X_{t_n}^* = x_n^* \mid T_{i+1} > t_n) \Lambda(x_n^*, x_n)
\end{aligned}$$

ce qui achève la démonstration de la récurrence. On en déduit :

$$\mathcal{L}(X_t \mid \mathcal{F}_t^{X^*}, T_i > t) = \Lambda(X_t^*, \cdot)$$

pour tout $i \in \mathbb{N}$, c'est-à-dire :

$$\mathbb{E}(\mathbf{1}_{T_i > t} G f(X_t)) = \mathbb{E}(\mathbf{1}_{T_i > t} G \Lambda[f](X_t^*))$$

pour toute fonction f continue bornée et pour tout variable aléatoire G bornée et $\mathcal{F}_t^{X^*}$ -mesurable. Or, $\lim_{t \rightarrow \infty} \mathbf{1}_{T_i > t} = 1$ presque sûrement et on conclut par convergence dominée que :

$$\mathcal{L}(X_t \mid \mathcal{F}_t^{X^*}) = \Lambda(X_t^*, \cdot)$$

□

Corollaire 1.2.8. *La condition (1.16) peut-être remplacée par :*

$$\mathcal{L}(X_T \mid \mathcal{F}_T^{X^*}, T < \infty) = \Lambda(X_T^*, \cdot) \tag{1.20}$$

pour tout $\mathcal{F}_t^{X^*}$ -temps d'arrêt T .

Démonstration. La condition (1.20) implique trivialement (1.16). L'autre sens est identique à la preuve du corollaire 1.1.4. □

Chapitre 2

Application à la recherche de temps forts de stationnarité finis

2.1 Processus de vie et de mort

Soit $(X_t)_{t \geq 0}$ un processus de vie et de mort sur \mathbb{Z} , de taux de décès $(a_n)_{n \in \mathbb{Z}}$ et de naissance $(b_n)_{n \in \mathbb{Z}}$, tous strictement positifs. On suppose que les coefficients sont tels que X est non-explosif et récurrent positif, c'est-à-dire (cf Anderson [3] Chapitre 8) :

$$\sum_{i=0}^{+\infty} \prod_{j=1}^i \frac{b_j}{a_{j+1}} + \sum_{i=0}^{\infty} \prod_{j=1}^i \frac{a_{-j}}{b_{-j-1}} < \infty \quad (2.1)$$

$$\left\{ \begin{array}{l} \sum_{i=0}^{+\infty} \frac{1}{b_i \mu(i)} = \infty \\ \sum_{i=-\infty}^0 \frac{1}{b_i \mu(i)} = \infty \end{array} \right. \quad (2.2)$$

où μ est la mesure stationnaire, définie de manière unique par les relations :

$$\mu(n)b_n = \mu(n+1)a_{n+1} \quad (2.3)$$

$$\sum_{n \in \mathbb{Z}} \mu(n) = 1 \quad (2.4)$$

soit, pour tout $n \in \mathbb{N}^*$:

$$\begin{aligned} \mu(n) &= \mu(0) \prod_{i=0}^{n-1} \frac{b_i}{a_{i+1}} \\ \mu(-n) &= \mu(0) \prod_{i=0}^{n-1} \frac{a_{-i}}{b_{-i-1}} \end{aligned}$$

$$\text{avec } \mu(0) = \frac{1}{1 + \sum_{n \in \mathbb{N}^*} \left(\prod_{i=0}^{n-1} \frac{b_i}{a_{i+1}} + \prod_{i=0}^{n-1} \frac{a_{-i}}{b_{-i-1}} \right)}.$$

On suppose que les coefficients du générateurs sont tels que :

$$\sum_{n \in \mathbb{Z}} \mu(n)b_n < \infty$$

Le générateur $(\mathcal{L}_{i,j})_{(i,j) \in \mathbb{Z}^2}$ du processus X est défini par :

$$\forall i, j \in \mathbb{Z}, \quad \mathcal{L}_{i,j} = \begin{cases} b_i & \text{si } j = i + 1 \\ a_i & \text{si } j = i - 1 \\ -a_i - b_i & \text{si } j = i \\ 0 & \text{sinon} \end{cases} \quad (2.5)$$

On note μ_t la distribution de X_t , $\forall t \geq 0$.

Partant de ce processus, on cherche à construire un dual de stationnarité forte X^* , à valeurs dans :

$$\mathbf{E} := \{(p, q) \in (\mathbb{Z} \cup \{-\infty\}) \times (\mathbb{Z} \cup \{+\infty\}); p \leq q\}$$

l'ensemble des intervalles de \mathbb{Z} , lié à X par le noyau Λ défini par :

$$\Lambda((p, q), n) = \frac{\mu(n)}{\mu(\llbracket p, q \rrbracket)}, \quad \forall (p, q) \in \mathbf{E} \quad \forall n \in \llbracket p, q \rrbracket$$

(on remarquera au passage que \mathcal{L} est bien Λ -intégrable comme requis dans la section 2). Ici et dans toute la suite, pour $(p, q) \in \mathbf{E}$, $n \in \llbracket p, q \rrbracket$ signifie $n \in [p, q] \cap \mathbb{Z}$ (on exclut notamment $n = \pm\infty$).

On munit \mathbf{E} de la topologie engendrée par les ouverts suivants :

$$\begin{aligned} & \{(p, q)\} \\ & \{(p, q + i), i \in \mathbb{N}\} \cup \{(p, +\infty)\} \\ & \{(p - i, q), i \in \mathbb{N}\} \cup \{(-\infty, q)\} \\ & \{(-\infty, q + i), i \in \mathbb{N}\} \cup \{\mathbb{Z}\} \\ & \{(p - i, +\infty), i \in \mathbb{N}\} \cup \{\mathbb{Z}\} \\ & \{(p - i, q + j), i, j \in \mathbb{N}\} \cup \{\mathbb{Z}\} \end{aligned}$$

avec $-\infty < p \leq q < +\infty$. On pose :

$$\mathbf{E}_1 := \{(p, q) \in \mathbb{Z}^2; p \leq q\}$$

$$\mathbf{E}_2 := \{(-\infty, q) \in \mathbf{E}\} \cup \{(p, +\infty) \in \mathbf{E}\}$$

$$\mathbf{E}_3 := \{\mathbb{Z}\}$$

Soit \mathcal{L}^* un générateur sur \mathbf{E} défini de la manière suivante :

$$\mathcal{L}^*((p, q), (p_i, q_i)) = \begin{cases} \frac{\mu(\llbracket p-1, q \rrbracket)}{\mu(\llbracket p, q \rrbracket)} b_{p-1} =: a'_{p,q} & (p_i, q_i) = (p-1, q) \\ \frac{\mu(\llbracket p, q+1 \rrbracket)}{\mu(\llbracket p, q \rrbracket)} a_{q+1} =: b_{p,q} & (p_i, q_i) = (p, q+1) \\ \frac{\mu(\llbracket p+1, q \rrbracket)}{\mu(\llbracket p, q \rrbracket)} a_p =: b'_{p,q} & (p_i, q_i) = (p+1, q) \\ \frac{\mu(\llbracket p, q-1 \rrbracket)}{\mu(\llbracket p, q \rrbracket)} b_q =: a_{b,q} & (p_i, q_i) = (p, q-1) \\ 0 & \text{pour tout autre } (p_i, q_i) \neq (p, q) \end{cases} \quad (2.6)$$

pour tout $(p, q) \in \mathbf{E}$, avec comme convention $a_\infty = b_\infty = a_{-\infty} = b_{-\infty} = 0$. Alors, on a la :

Proposition 2.1.1. *Il existe une mesure μ_0^* sur \mathbf{E} et un processus minimal X^* de distribution initiale μ_0^* et de générateur \mathcal{L}^* tel que X^* soit Λ -lié avec X .*

Démonstration. D'après le théorème 1.1.1, il suffit d'avoir $\mu_0^* \Lambda = \mu_0$ et $\mathcal{L}^* \Lambda = \Lambda \mathcal{L}$. Pour le premier c'est immédiat si on prend la mesure qui ne charge que des singletons : $\mu_0^*((m, n)) = \delta_{m,n} \mu_0(n)$. Pour le deuxième : $\forall (p, q) \in \mathbf{E}, \forall n \in \mathbb{Z}$

$$\begin{aligned}
(\Lambda \mathcal{L})_{(p,q),n} &= \sum_{m \in \mathbb{Z}} \Lambda((p, q), m) \mathcal{L}(m, n) \\
&= \sum_{m \in \llbracket p, q \rrbracket} \Lambda((p, q), m) \mathcal{L}(m, n) \\
&= \frac{\mu(n-1)}{\mu(\llbracket p, q \rrbracket)} \delta_{n-1}(\llbracket p, q \rrbracket) b_{n-1} \\
+ \frac{\mu(n+1)}{\mu(\llbracket p, q \rrbracket)} \delta_{n+1}(\llbracket p, q \rrbracket) a_{n+1} &- \frac{\mu(n)}{\mu(\llbracket p, q \rrbracket)} \delta_n(\llbracket p, q \rrbracket) (b_n + a_n) \\
&= \frac{\mu(n)[b_n(\delta_{n,p-1} - \delta_{n,q}) + a_n(\delta_{n,q+1} - \delta_{n,p})]}{\mu(\llbracket p, q \rrbracket)}
\end{aligned}$$

et :

$$\begin{aligned}
(\mathcal{L}^* \Lambda)_{(p,q),n} &= \sum_{(p',q') \in \mathbf{E}} \mathcal{L}^*((p, q), (p', q')) \Lambda((p', q'), n) \\
&= \frac{\mu(n)}{\mu(\llbracket p, q \rrbracket)} (b_{p-1} \delta_n(\llbracket p-1, q \rrbracket) - \delta_n(\llbracket p, q \rrbracket) b_{p-1} \\
&\quad + a_{q+1} \delta_n(\llbracket p, q+1 \rrbracket) - \delta_n(\llbracket p, q \rrbracket) a_{q+1} \\
&\quad + a_p \delta_n(\llbracket p+1, q \rrbracket) - \delta_n(\llbracket p, q \rrbracket) a_p \\
&\quad + b_q \delta_n(\llbracket p, q-1 \rrbracket) - \delta_n(\llbracket p, q \rrbracket) b_q) \\
&= \frac{\mu(n)[b_n(\delta_{n,p-1} - \delta_{n,q}) + a_n(\delta_{n,q+1} - \delta_{n,p})]}{\mu(\llbracket p, q \rrbracket)} \\
&= (\Lambda \mathcal{L})_{(p,q),n}
\end{aligned}$$

□

On a donc un processus minimal $(X_t^*) = (P_t, Q_t) \in \mathbf{E}$. Les bornes P et Q évoluent comme des processus de vie et de mort individuellement non-markoviens sur \mathbb{Z} , jusqu'à un possible temps d'explosion T . Ici, le seul état absorbant pour X^* est \mathbb{Z} , mais il se trouve qu'il ne peut pas être atteint en un nombre fini de sauts depuis un autre état. Il nous faut donc considérer un processus de saut stratifié associé à \mathcal{L}^* , ce qui est rendu possible par la :

Proposition 2.1.2. \mathcal{L}^* est stratifié sur \mathbf{E} , et $(\mathbf{E}_1, \mathbf{E}_2, \mathbf{E}_3)$ est une stratification de \mathcal{L}^* .

Démonstration. Pour $(X_t^*) = (P_t, Q_t)$ un processus minimal associé à \mathcal{L}^* et de loi initiale μ_0^* , on appelle T son premier temps d'explosion, et on pose :

$$T_1 := \begin{cases} T & \text{si } T \text{ est un temps d'explosion pour } P \\ +\infty & \text{sinon} \end{cases}$$

$$T_2 := \begin{cases} T & \text{si } T \text{ est un temps d'explosion pour } Q \\ +\infty & \text{sinon} \end{cases}$$

de sorte que $T = T_1 \wedge T_2$.

On va montrer que $\lim_{t \rightarrow T} Q_t = +\infty$ p.s. si $T = T_2$, et on en déduit par symétrie que $\lim_{t \rightarrow T} P_t = -\infty$ si $T = T_1$. Pour cela, on considère le processus $X^* = (P, Q)$ couplé avec un processus de vie et de mort (homogène) Z de même loi initiale que Q , correspondant à l'évolution de la borne droite de X^* lorsque la borne gauche est en $-\infty$. Un tel couple est donné par le :

Lemme 2.1.3. *Si $T = T_2$, il existe un couplage (X^*, Z) tel que $Q_t \geq Z_t, \forall 0 \leq t \leq T$.*

Démonstration. On va construire $\tilde{X}_t^* := (-\infty, Z_t)$ un processus de Markov de générateur \mathcal{L}^* . On simule $X_0^* = (P_0, Q_0)$ selon sa loi initiale μ_0^* , et on pose $\tilde{X}_0^* = (-\infty, Q_0)$. Le processus Z est un processus de vie et de mort, de taux de décès et de naissance respectivement :

$$\tilde{b}_n = \frac{\mu(\llbracket -\infty, n+1 \rrbracket)}{\mu(\llbracket -\infty, n \rrbracket)} a_{n+1} \quad \text{et} \quad \tilde{a}_n = \frac{\mu(\llbracket -\infty, n-1 \rrbracket)}{\mu(\llbracket -\infty, n \rrbracket)} b_n$$

et on a, $\forall (p, q) \in \mathbf{E}$:

$$\begin{aligned} b_{p,q} &= \frac{\mu(\llbracket p, q+1 \rrbracket)}{\mu(\llbracket p, q \rrbracket)} a_{q+1} \\ &= \left(1 + \frac{\mu(q+1)}{\mu(\llbracket p, q \rrbracket)}\right) a_p \\ &\geq \left(1 + \frac{\mu(q+1)}{\mu(\llbracket -\infty, q \rrbracket)}\right) a_p \\ &= \frac{\mu(\llbracket -\infty, q+1 \rrbracket)}{\mu(\llbracket -\infty, q \rrbracket)} a_p \\ &= \tilde{b}_q \\ a_{p,q} &= \frac{\mu(\llbracket p, q-1 \rrbracket)}{\mu(\llbracket p, q \rrbracket)} b_q \\ &\leq \frac{\mu(\llbracket -\infty, q-1 \rrbracket)}{\mu(\llbracket -\infty, q \rrbracket)} b_q \\ &= \tilde{a}_q \end{aligned}$$

c'est-à-dire que les naissances se font en moyenne plus rapidement pour le processus $(Q_t)_t$, et les décès plus lentement. Pour effectuer le couplage, on procède alors en laissant les deux processus évoluer indépendamment tant que $Q_t > Z_t$ puis, lorsqu'ils se rejoignent, en « forçant » Z_t à descendre si Q_t descend, et Q_t à monter si Z_t monte.

Pour cela, on définit par récurrence la suite (T_i) des temps de « jonction » de Q_t et Z_t et (T'_i) celle des temps de saut suivant la jonction (on autorise une jonction instantanée, c'est-à-dire qu'on peut avoir $T_{i+1} = T'_i$), et on décrit le comportement du couple (X^*, Z) entre ces temps. Soit $T_0 = 0$. $\forall i \geq 0$, on définit T'_i comme étant une variable aléatoire telle que $T'_i - T_i$ suit une loi exponentielle de paramètre $\lambda_i := \tilde{a}_{Q_{T_i}} + b_{P_{T_i}, Q_{T_i}} + a'_{P_{T_i}, Q_{T_i}} + b'_{P_{T_i}, Q_{T_i}}$. Soit T_{i+1} le temps d'arrêt :

$$T_{i+1} := \inf\{t \geq T'_i; Q_t = Z_t\}$$

On impose $Q_t = Z_t = Q_{T_i}$ et $P_t = P_{T_i}$, $\forall T_i < t < T'_i$, puis :

$$((P_{T'_i}, Q_{T'_i}), Z_{T'_i}) = \begin{cases} ((P_{T_i}, Q_{T_i} - 1), Z_{T_i} - 1) & \text{avec probabilité } \frac{a_{P_{T_i}, Q_{T_i}}}{\lambda_i} \\ ((P_{T_i}, Q_{T_i} + 1), Z_{T_i} + 1) & \text{avec probabilité } \frac{\tilde{b}_{Q_{T_i}}}{\lambda_i} \\ ((P_{T_i}, Q_{T_i}), Z_{T_i} - 1) & \text{avec probabilité } \frac{\tilde{a}_{Q_{T_i}} - a_{P_{T_i}, Q_{T_i}}}{\lambda_i} \\ ((P_{T_i}, Q_{T_i} + 1), Z_{T_i}) & \text{avec probabilité } \frac{b_{P_{T_i}, Q_{T_i}} - \tilde{b}_{Q_{T_i}}}{\lambda_i} \\ ((P_{T_i} - 1, Q_{T_i}), Z_{T_i}) & \text{avec probabilité } \frac{a'_{P_{T_i}, Q_{T_i}}}{\lambda_i} \\ ((P_{T_i} + 1, Q_{T_i}), Z_{T_i}) & \text{avec probabilité } \frac{b'_{P_{T_i}, Q_{T_i}}}{\lambda_i} \end{cases}$$

Ensuite, X_t^* et Z_t évoluent indépendamment entre les temps T'_i et T_{i+1} . On a clairement $Q_t \geq Z_t$, $\forall 0 \leq t \leq T$, il ne reste qu'à vérifier que les marginales (X_t^*) et (Z_t) du processus ainsi construit correspondent bien aux processus souhaités. Lorsque $Z_t < Q_t$ c'est le cas par définition. Sinon, soit $\bar{\mathcal{L}}^*$ le générateur du processus couplé (X^*, Z) . On a, pour toute fonction f sur $\mathbf{E} \times \mathbb{Z}$, pour tout $(p, q) \in \mathbf{E}$:

$$\begin{aligned} \bar{\mathcal{L}}^* f((p, q), q) &= a_{p,q}(f((p, q - 1), q - 1) - f((p, q), q)) \\ &\quad + \tilde{b}_q(f((p, q + 1), q + 1) - f((p, q), q)) \\ &\quad + (\tilde{a}_q - a_{p,q})(f((p, q), q - 1) - f((p, q), q)) \\ &\quad + (b_{p,q} - \tilde{b}_q)(f((p, q + 1), q) - f((p, q), q)) \\ &\quad + a'_{p,q}(f((p - 1, q), q) - f((p, q), q)) \\ &\quad + b_{p,q}(f((p + 1, q), q) - f((p, q), q)) \end{aligned}$$

Ainsi, si f ne dépend pas de la première variable, on a :

$$\bar{\mathcal{L}}^* f((p, q), q) = \tilde{b}_q(f(q + 1) - f(q)) + \tilde{a}_q(f(q - 1) - f(q))$$

et si f ne dépend pas de la deuxième variable :

$$\begin{aligned} \bar{\mathcal{L}}^* f((p, q), q) &= a_{p,q}(f((p, q - 1)) - f((p, q))) + (b_{p,q}(f((p, q + 1)) - f((p, q))) \\ &\quad + a'_{p,q}(f((p - 1, q)) - f((p, q))) + b_{p,q}(f((p + 1, q)) - f((p, q))) \end{aligned}$$

ce qui correspond bien aux générateurs respectifs de Z et X^* . \square

Il ne reste donc plus qu'à montrer :

$$Z_t \xrightarrow[t \rightarrow +\infty]{} +\infty \quad \text{p.s.} \quad (2.7)$$

$\tilde{X}_t^* = (-\infty, Z_t)$ est égal en loi à un dual de \tilde{X}_t . Ce dernier est récurrent positif, et pour tout $t \geq 0$, $\tilde{X}_t \leq Z_t$, donc $Z_t > 0$ infiniment souvent (si on considère la chaîne de saut sous-jacente). Ceci implique notamment que $Z_t \xrightarrow[t \rightarrow +\infty]{} -\infty$ est impossible p.s..

Si on considère le même processus \tilde{Z}_t que l'on force à rester sur \mathbb{N} en imposant $a_0 = 0$, alors la condition pour $\tilde{Z}_t \xrightarrow[t \rightarrow +\infty]{} +\infty$ est :

$$\sum_{i=0}^{\infty} \prod_{j=1}^i \frac{\tilde{a}_j}{\tilde{b}_j} < \infty$$

Or :

$$\begin{aligned} \sum_{i=0}^{\infty} \prod_{j=1}^i \frac{\tilde{a}_j}{\tilde{b}_j} &= \sum_{i=0}^{\infty} \prod_{j=1}^i \frac{\mu(\llbracket -\infty, j-1 \rrbracket) b_j}{\mu(\llbracket -\infty, j+1 \rrbracket) a_{j+1}} \\ &\leq \sum_{i=0}^{\infty} \prod_{j=1}^i \frac{b_j}{a_{j+1}} \\ &< \infty \end{aligned}$$

Par conséquent, $\tilde{Z}_t \xrightarrow{t \rightarrow +\infty} \infty$ presque sûrement, et donc Z_t aussi. En effet, tant que $Z_t > 0$, le processus Z se comporte comme \tilde{Z} . Si l'on pose $\tau_0 = 0$, $\sigma_0 = \inf\{t \geq \tau_0, Z_t = 0\}$, et pour tout $i > 0$:

$$\tau_i = \inf\{t > \sigma_{i-1}, Z_t = 1\} \quad \text{et} \quad \sigma_i = \inf\{t > \tau_i, Z_t = 0\}$$

(avec comme convention $\inf \emptyset = +\infty$), on a $\mathbb{P}(\tau_i < \infty \mid \sigma_{i-1} < \infty) = 1$ car $Z_t > 0$ infiniment souvent et

$$\mathbb{P}(\sigma_i < \infty \mid \tau_i < \infty) = 1 - \mathbb{P}(\tilde{Z}_t \neq 0, \forall t > 0 \mid \tilde{Z}_0 = 1) < 1$$

Il existe donc presque sûrement un temps τ_i pour lequel $Z_t > 0$ pour tout $t \geq \tau_i$. Comme :

$$\mathcal{L}(Z_{\tau_i+} \mid \tau_i < \infty, \sigma_i = \infty) = \mathcal{L}(\tilde{Z} \mid \tilde{Z}_0 = 1, \tilde{Z}_t > 0, \forall t > 0)$$

il vient donc $Z_t \xrightarrow{t \rightarrow +\infty} \infty$ presque sûrement.

Si au contraire $T < T_2$ alors Q n'explode pas au temps T et admet donc un dernier temps de saut T_k avant T et $\lim_{t \rightarrow T} Q_t = Q_{T_k} \in \mathbb{Z}$ presque sûrement. En faisant le même raisonnement avec P , on conclut que $\lim_{t \rightarrow T} X_t^*$ existe presque sûrement.

On vérifie alors facilement que les trois conditions pour que $\mathbf{E}_1, \mathbf{E}_2, \mathbf{E}_3$ forment une stratification de \mathcal{L}^* sont vérifiées. \square

Par le Théorème 1.2.7, on est maintenant autorisé à considérer un processus de saut stratifié X^* associé au générateur \mathcal{L}^* et à une loi initiale μ_0^* vérifiant $\mu_0^* \Lambda = \mu_0$, tel que X^* soit un dual de stationnarité forte. L'intérêt de ce processus est que l'existence d'un temps fort de stationnarité pour toute distribution initiale μ_0 de X est équivalente à l'absorption en temps fini de X^* en \mathbb{Z} , pour au moins une distribution initiale μ_0^* telle que $\mu_0^* \Lambda = \mu_0$. L'implication dans un sens est immédiate d'après la définition du noyau Λ et le corollaire 1.2.8, dans l'autre elle découle de la proposition suivante :

Proposition 2.1.4. *Si $\mu_0 = \mu_{\llbracket -\infty, 0 \rrbracket}$ et $X_0^* = (-\infty, 0)$ p.s., alors le dual X^* est optimal.*

La preuve est identique à celle de Miclo [65], lemme 26.

Démonstration. On utilise la fonction de séparation \mathfrak{s} (on rappelle que $\mathfrak{s}(t) := \sup_m \left(1 - \frac{\mu_t(m)}{\mu(m)}\right)$, voir [32] (1.3), avec $\mu_t = \mathcal{L}(X_t)$), et le noyau Λ . Si on appelle $T := \inf\{t > 0; Q_t = \infty\}$ le temps d'absorption de X^* , alors :

$$\begin{aligned}
\mathfrak{s}(t) &= \sup_{n \in \mathbb{Z}} \left(1 - \frac{\mu_t(n)}{\mu(n)} \right) \\
&= \sup_{n \in \mathbb{Z}} \mathbb{E} \left(1 - \frac{\Lambda(\cdot - \infty, Q_t], n)}{\mu(n)} \right) \\
&= 1 - \inf_{n \in \mathbb{Z}} \mathbb{E} \left(\frac{\delta_n(\cdot - \infty, Q_t]}{\mu(\cdot - \infty, Q_t]} \right) \\
&= 1 - \lim_{n \rightarrow \infty} \mathbb{E} \left(\frac{\delta_n(\cdot - \infty, Q_t]}{\mu(\cdot - \infty, Q_t]} \right) \\
&= 1 - \mathbb{P}(Q_t = \infty) \\
&= \mathbb{P}(T > t)
\end{aligned}$$

□

Ainsi, dans ce cas, le temps d'absorption en \mathbb{Z} de X^* est un temps à la stationnarité et sa finitude p.s. est équivalente à l'existence d'un temps fort de stationnarité fini (pour la loi initial μ_0).

On est enfin en mesure d'énoncer et de montrer le résultat principal de cette section :

Théorème 2.1.5. *Il existe un temps fort de stationnarité fini pour le processus X , pour toute distribution initiale, si et seulement si :*

$$\sum_{i=1}^{\infty} \mu(i+1) \sum_{j=1}^i \frac{1}{\mu(j)b_j} < \infty \tag{2.8}$$

et

$$\sum_{i=1}^{\infty} \mu(-i-1) \sum_{j=1}^i \frac{1}{\mu(-j)a_{-j}} < \infty \tag{2.9}$$

Démonstration. • Condition nécessaire :

On se place dans les conditions de la proposition 2.1.4, c'est-à-dire $\mu_0 = \mu|_{\llbracket -\infty, 0 \rrbracket}$ et $X_0^* = \llbracket -\infty, 0 \rrbracket$ p.s.. On a vu que dans ce cas, X^* est optimal, et donc s'il existe un temps fort de stationnarité alors X^* est absorbé en temps fini p.s.. Or, ceci est vrai si et seulement si la série

$$\sum_{i=1}^{\infty} \prod_{j=1}^i \frac{\tilde{a}_j}{\tilde{b}_j} \sum_{k=1}^i \prod_{l=1}^k \frac{\tilde{b}_{l-1}}{\tilde{a}_l}$$

converge. En effet, il s'agit du critère d'explosion d'un processus de vie et de mort sur \mathbb{N} (voir par exemple [3], section 8.1), que l'on étend ici à une chaîne sur \mathbb{Z} à valeur dans \mathbb{N} infiniment souvent, de la même manière que pour (2.7). Cette condition se reformule ainsi :

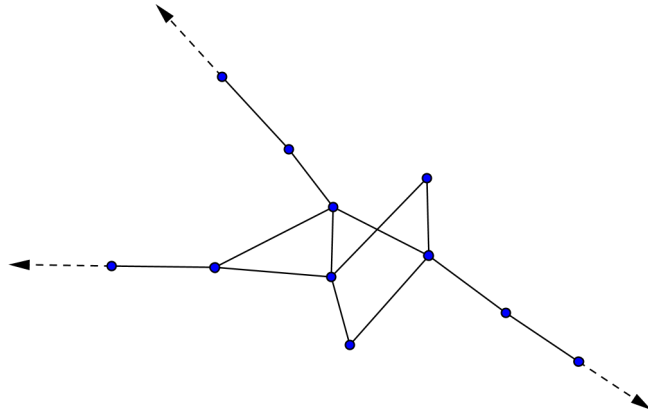


FIGURE 2.1 – Exemple de graphe considéré

$$\begin{aligned}
& \sum_{i=1}^{\infty} \prod_{j=1}^i \frac{b_j \mu(\llbracket -\infty, j-1 \rrbracket)}{a_{j+1} \mu(\llbracket -\infty, j+1 \rrbracket)} \sum_{k=1}^i \prod_{l=1}^k \frac{a_l}{b_l} < \infty \\
& \Leftrightarrow \sum_{i=1}^{\infty} \prod_{j=1}^i \frac{b_j}{a_{j+1}} \sum_{k=1}^i \prod_{l=1}^k \frac{a_l}{b_l} < \infty \\
& \Leftrightarrow \sum_{i=1}^{\infty} \mu(i+1) \sum_{j=1}^i \frac{1}{\mu(j)b_j} < \infty
\end{aligned}$$

où on a utilisé le fait que $\mu(\llbracket -\infty, j \rrbracket)$ est borné pour $j \in \mathbb{N}$. Ceci prouve que (2.8) est une condition nécessaire. On traite (2.9) de la même manière, en prenant $\mu_{\llbracket 0, +\infty \rrbracket}$ comme loi initiale pour le processus X , $X_0^* = \llbracket 0, +\infty \rrbracket$ presque sûrement et en considérant la borne gauche du processus dual X^* .

- Condition suffisante :

Inversement, si le processus X^* partant de $\llbracket -\infty, 0 \rrbracket$ (resp. $\llbracket 0, \infty \rrbracket$) est absorbé en temps fini p.s., alors d'après la Proposition 2.1.3, il existe un dual qui le sera aussi, pour toute loi initiale μ_0 ce qui conclut la preuve. □

2.2 Marche aléatoire sur un graphe

Le cas précédent peut être considéré comme une marche aléatoire sur un graphe, où Z est vu comme un graphe constitué uniquement de deux branches infinies reliées en 0. On cherche maintenant à étendre le résultat obtenu dans la section précédente à des graphes comportant un nombre fini de branches infinies, plus un nombre fini de sommets, comme représenté dans la figure 2.1

2.2.1 Définition du cadre et notations

On considère donc un générateur \mathcal{L} sur un ensemble dénombrable G , irréductible et tel que pour tout $x \in G$ l'ensemble $\{y \in G : \mathcal{L}_{x,y} > 0 \text{ ou } \mathcal{L}_{y,x} > 0\}$ est fini. Soit $(X_t)_{t \geq 0}$ un processus minimal de générateur \mathcal{L} et μ_0 sa loi initiale. On munit alors G d'une structure de graphe simple, non-orienté, à poids et connexe, où deux sommets distincts p et $q \in G$ sont reliés par une arête si et seulement si $\mathcal{L}_{p,q} \neq 0$ ou $\mathcal{L}_{q,p} \neq 0$ (chaque sommet est donc de degré fini).

On suppose que \mathcal{L} est non-explosif, récurrent positif et réversible, et on note μ sa mesure stationnaire. On rappelle que μ vérifie (21) :

$$\mu(q)\mathcal{L}_{q,p} = \mu(p)\mathcal{L}_{p,q}, \quad \forall p, q \in G$$

et en particulier, $\mathcal{L}_{p,q} \neq 0$ implique $\mathcal{L}_{q,p} \neq 0$. Cette égalité ainsi que $\sum_q \mathcal{L}_{p,q} = 0$ seront à la base de tous les calculs qui suivent. On suppose de plus que :

$$\sum_{x \in G} \mu(x)\mathcal{L}_x < \infty$$

Commençons par poser le vocabulaire et les notations qui seront utilisées dans la suite.

Définition 2.2.1. (i) Pour toute partie $Q \subset G$, on notera Q^c le complémentaire de Q dans G .

(ii) Deux sommets $p, q \in G$ sont dits voisins si ils sont reliés par une arête.

Pour $Q \subset G$ et $q \in Q$, on note $V_Q(q)$ l'ensemble des voisins de q dans Q , $V(q) := V_G(q)$ et $V(Q) := \{q' \in V_{Q^c}(q), q \in Q\}$.

(iii) Le degré d'un sommet est le nombre de ses voisins.

(iv) Un chemin entre deux sommets $p, q \in G$ est une suite de sommets $p = p_1, \dots, p_n = q$ telle que p_i et p_{i+1} soient voisins, pour tout $1 \leq i < n$, et que tous les p_i soient distincts.

(v) Un sous-ensemble Q de G est dit connexe si entre n'importe quel couple de sommets de Q il existe un chemin inclus dans Q . Il est dit convexe si tout chemin entre deux sommets de Q est inclus dans Q .

(vi) Si $Q' \subset Q \subset G$, Q' est une composante connexe de Q si Q' est connexe et maximal (pour l'inclusion) parmi les sous-ensembles connexes de Q . On note $\mathcal{C}(Q)$ l'ensemble des composantes connexes de Q .

(vii) On appelle centre de G et on note C_G la réunion de tous les chemins entre deux sommets de degrés différents de deux (c'est-à-dire l'enveloppe convexe de l'ensemble de ces sommets).

(viii) On appelle branche infinie toute composante connexe (nécessairement infinie) de $G \setminus C_G$. On note Q_i , $i \in I$ les différentes branches infinies de G .

(ix) A chaque branche infinie Q_i , on adjoint un point à l'infini noté Δ_i . On note $\bar{Q}_i = Q_i \cup \{\Delta_i\}$.

(x) Dans le cas où C_G est non vide, si l'on munit \mathbb{N} d'une structure de graphe où il existe une arête entre n et $m \in \mathbb{N}$ si et seulement si $|n - m| = 1$, alors chaque Q_i est isomorphe à \mathbb{N} en tant que graphe et on note $\varphi_i : \mathbb{N} \cup \{\infty\} \rightarrow \bar{Q}_i$ l'unique isomorphisme correspondant, complété en posant $\varphi_i(\infty) = \Delta_i$. On utilisera également la notation suivante :

$$\varphi_i(\llbracket p, \infty \rrbracket) := \bigcup_{q \geq p} \{\varphi_i(q)\} \cup \{\Delta_i\}$$

(xi) On note, comme en (20) :

$$\mu^i(n) := \prod_{k=0}^{n-1} \frac{\mathcal{L}_{\varphi_i(k), \varphi_i(k+1)}}{\mathcal{L}_{\varphi_i(k+1), \varphi_i(k)}} = \frac{\mu(\varphi_i(n))}{\mu(\varphi_i(0))}$$

On suppose donc que C_G est fini, ainsi que le nombre de branches infinies ($I = \llbracket 1, \dots, N \rrbracket$ avec $N \in \mathbb{N}$), et non vide (le cas où C_G est vide correspond à $G = \mathbb{Z}$, qui a été traité dans la section précédente). Dans ce contexte on pose :

$$\bar{G} = G \cup \bigcup_{i \in I} \{\Delta_i\}$$

et on définit une métrique \bar{d} sur \bar{G} de la manière suivante :

— Si $p, q \in Q_i$, $i \in I$:

$$\bar{d}(p, q) = \left| \sum_{n=\varphi_i^{-1}(p)+1}^{\varphi_i^{-1}(q)} \frac{1}{n^2} \right|$$

et :

$$\bar{d}(p, \Delta_i) = \bar{d}(\Delta_i, p) = \sum_{n=\varphi_i^{-1}(p)+1}^{+\infty} \frac{1}{n^2}$$

— Si $p, q \in C_G$:

$$\bar{d}(p, q) = \inf\{\text{card}(\gamma), \gamma \text{ chemin entre } p \text{ et } q\} - 1$$

c'est-à-dire la longueur du plus court chemin (en particulier, si p et q sont voisins, alors $\bar{d}(p, q) = 1$).

— Si $i \in I$ et (p_i, q_i) est l'unique paire voisine dans $Q_i \times C_G$, on pose :

$$\bar{d}(q, p) = \bar{d}(p, q) := 1 + \bar{d}(p, p_i) + \bar{d}(q_i, q), \quad \forall p \in \bar{Q}_i, \quad \forall q \in C_G$$

et :

$$\bar{d}(q, p) = \bar{d}(p, q) := \bar{d}(p, q_i) + \bar{d}(q_i, q), \quad \forall p \in \bar{Q}_i \quad \forall q \in \bar{Q}_j \quad (i \neq j)$$

Remarque 2.2.2. Si $p_1, p_2, p_3 \in Q_i$, $i \in I$, sont trois sommets d'une même branche infinie alors $\bar{d}(p_1, \Delta_i) < \bar{d}(p_2, \Delta_i) < \bar{d}(p_3, \Delta_i)$, si et seulement si l'unique chemin entre p_1 et p_3 contient p_2 .

Il est facile de vérifier que G muni de cette métrique est compacte : la topologie induite sur chacune des branches est celle du compactifié d'Alexandrov.

Soit G^* l'ensemble des compacts connexes non vides de \bar{G} , muni de la distance de Hausdorff :

$$d^*(Q, Q') := \max\left\{\sup_{p \in Q} \inf_{p' \in Q'} \bar{d}(p, p'), \sup_{p' \in Q'} \inf_{p \in Q} \bar{d}(p, p')\right\}$$

C'est un sous-espace fermé de l'espace des parties compactes (muni de la distance de Hausdorff). On rappelle que ce dernier est compact et donc G^* aussi.

Décrivons un peu ces deux espaces métriques.

En notant $B_d(x, \varepsilon)$ (respectivement $\bar{B}_d(x, \varepsilon)$) la boule ouverte (respectivement fermée) de centre x et de rayon ε pour une métrique d , on a :

$$B_{\bar{d}}(p, 1/2) = \bar{B}_{\bar{d}}(p, 1/2) = \{p\}, \quad \forall p \in C_G$$

$$B_{\bar{d}}(\varphi_i(p), \frac{1}{(p+1)^2}) = \bar{B}_{\bar{d}}(\varphi_i(p), \frac{1}{2(p+1)^2}) = \{p\}, \quad \forall p \in \mathbb{N}, \quad \forall i \in I$$

$$B_{\bar{d}}(\Delta_i, \varepsilon) = \{\varphi_i(n); \sum_{j=n+1}^{+\infty} \frac{1}{j^2} < \varepsilon\} \quad \forall i \in I, \quad \forall \varepsilon < \sum_{j=0}^{+\infty} \frac{1}{j^2}$$

$$\bar{B}_{\bar{d}}(\Delta_i, \varepsilon) = \{\varphi_i(n); \sum_{j=n+1}^{+\infty} \frac{1}{j^2} \leq \varepsilon\} \quad \forall i \in I, \forall \varepsilon < \sum_{j=0}^{+\infty} \frac{1}{j^2}$$

Les ensembles ci-dessus sont donc les ouverts élémentaires de la topologie induite par la métrique \bar{d} . En particulier chaque singleton est fermé (comme complémentaire d'une réunion dénombrable d'ouverts), et chaque singleton de G est aussi ouvert. De plus, les Δ_i sont les seuls points d'accumulations de \bar{G} .

Les éléments de G^* sont des compacts (donc des fermés) de \bar{G} , c'est-à-dire des réunions finies de singletons de G et de voisinages de Δ_i tels que décrits précédemment.

Les affirmations précédentes sont faciles à vérifier à partir des définitions. Le résultat suivant n'est pas beaucoup plus difficile :

Lemme 2.2.3. *Soit $Q \in G^*$, et $I_0 := \{i \in I; \Delta_i \in Q\}$. Alors, pour tout $\eta > 0$, il existe $\varepsilon > 0$ tel que pour tout $Q' \in B_{\bar{d}^*}(Q, \varepsilon)$:*

$$Q' \setminus \left(\bigcup_{i \in I_0} B_{\bar{d}}(\Delta_i, \eta) \right) = Q \setminus \left(\bigcup_{i \in I_0} B_{\bar{d}}(\Delta_i, \eta) \right)$$

Autrement dit, un voisin « suffisamment proche » de Q doit coïncider avec Q en dehors de voisinages des $\Delta_i \in G$. En particulier, si $Q \in G^*$ est inclus dans G , alors Q est un point isolé (dans le sens où $\{Q\}$ est un voisinage ouvert de Q).

Démonstration. Il suffit de prendre :

$$\varepsilon < \inf \{ \bar{d}(p, q); p, q \notin \bigcup_{i \in I_0} B_{\bar{d}}(\Delta_i, \eta), p \in Q \}$$

□

Pour tout $i \in I \cup \{0\}$, on pose :

$$G_i^* := \{Q \in G^*; |\{j \in I; \Delta_j \in Q\} = i\} \quad (2.10)$$

l'ensemble des compacts connexes non-vides de G qui contiennent exactement i points Δ_j .

2.2.2 Générateur dual

On définit un générateur \mathcal{L}^* sur G^* par :

$$\forall Q \neq Q' \in G^*, \quad \mathcal{L}_{Q, Q'}^* := \begin{cases} \left(\frac{\mu(Q')}{\mu(Q)} \right) \sum_{q \in Q} \mathcal{L}_{p, q} & \text{si } Q' = Q \cup \{p\}, p \notin Q \\ \left(\frac{\mu(Q')}{\mu(Q)} \right) \sum_{q \notin Q} \mathcal{L}_{p, q} & \text{si } Q' \in \mathcal{C}(Q \setminus \{p\}), p \in Q \\ 0 & \text{sinon} \end{cases} \quad (2.11)$$

(où on rappelle que $\mathcal{C}(Q)$ désigne l'ensemble des parties connexes de Q).

On remarque que les coefficients ci-dessus sont bien définis, puisque si $Q' \in \mathcal{C}(Q \setminus \{p\}) \cap \mathcal{C}(Q \setminus \{q\})$ avec $p, q \in Q$, alors nécessairement $p, q \notin Q'$ et $Q' \cup \{p\}, Q' \cup \{q\}$ sont connexes, donc $p = q$.

La dynamique d'un processus minimal X^* de générateur \mathcal{L}^* est à comprendre ainsi : à chaque temps de saut, on rajoute à X^* un point voisin, ou on lui enlève un point voisin de son complémentaire. Lorsque le fait d'enlever un point sépare X^* en plusieurs parties

connexes, on choisit l'une de ces parties avec une probabilité proportionnelle à son poids relativement à μ .

De là, on déduit que le nombre de transitions possibles est borné : pour un $Q \in G^*$ donné, le nombre de $Q \cup \{p\}$ avec $p \in V(Q)$ est majoré par $|C_G| + N$ et le nombre de $Q' \in \mathcal{C}(Q \setminus \{p\})$ avec $p \in V(Q')$ est borné (grossièrement) par $|C_G|^2$.

Lemme 2.2.4. \mathcal{L}^* est un générateur stratifié sur G^* , et (G_0^*, \dots, G_N^*) en est une stratification.

Démonstration. Les G_i^* sont disjoints et les coefficients de \mathcal{L}^* empêchent toute transition d'un ensemble G_i^* à un ensemble G_j^* , si $j > i$. Les deux premières conditions de la définition 1.2.1 sont donc remplies.

Soit X^* une trajectoire d'un processus minimal associé au générateur \mathcal{L}^* . On appelle T_n , $n \in \mathbb{N}$ la chaîne des temps de sauts de X^* , $(Y_n)_{n \in \mathbb{N}}$ sa chaîne de sauts et $T = \lim T_n$ son temps d'explosion, que l'on suppose fini. Y passe donc par un nombre infini d'états et visite chacun d'eux un nombre fini de fois. Comme G^* est compact, il en résulte que Y admet au moins un point d'accumulation, qui est un point d'accumulation de G^* . S'il est unique alors Y_n converge vers ce point. Supposons donc par l'absurde qu'il en existe deux distincts, que l'on appelle E_1 et E_2 , limites respectives de deux sous-suites $(Y_{\phi_1(n)})_n$ et $(Y_{\phi_2(n)})_n$ de $(Y_n)_n$. L'idée de ce qui suit est de montrer qu'on ne peut pas passer infiniment souvent d'une sous-suite à l'autre.

Puisque E_1 et E_2 sont deux compacts connexes de \bar{G} et $E_1 \neq E_2$, il existe $p \in G$ tel que $p \in E_1$ et $p \notin E_2$ (ou l'inverse).

D'après le lemme 2.2.3, il existe $n_1, n_2 \in \mathbb{N}$ tels que $p \in Y_{\phi_1(n)}$, $\forall n > n_1$ et $p \notin Y_{\phi_2(n)}$, $\forall n > n_2$. Après le temps $\max\{\phi_1(n_1), \phi_2(n_2)\}$ la chaîne Y effectue donc une infinité de transition d'un état Q contenant un voisin de p à l'état $Q \cup \{p\}$. Le taux d'une telle transition étant majoré par :

$$\left(1 + \sup_{q \in V(p)} \frac{\mu(p)}{\mu(q)} \sum_{q' \in V(p)} \mathcal{L}_{p,q'} \right) < \infty$$

c'est impossible presque sûrement.

La suite $(Y_n)_n$ admet donc une limite unique p.s. $E \in G_k^*$, $k \in I \cup \{0\}$.

Soit $I_0 := \{i \in I; \Delta_i \in E\}$. D'après le lemme 2.2.3, pour tout $\eta > 0$, il existe n_1 tel que pour tout $n > n_1$:

$$Y_n \setminus \left(\bigcup_{i \in I_0} B_{\bar{d}}(\Delta_i, \eta) \right) = E \setminus \left(\bigcup_{i \in I_0} B_{\bar{d}}(\Delta_i, \eta) \right)$$

En particulier, si $i \notin I_0$, alors $\Delta_i \notin Y_n$. De plus, il existe $n > n_1$ tel que $Y_n \neq E$, et puisque Y_n est connexe, il existe $i \in I_0$ tel que $\Delta_i \notin Y_n$. Le nombre de Δ_i dans Y_n est donc strictement plus petit que dans E , c'est-à-dire $Y_n \in G_j^*$ avec $j < k$.

Maintenant, le nombre de transitions du processus minimal d'un G_i^* vers un G_j^* , $i \neq j$, est fini p.s. (puisque elles ne se font que si $j < i$) et donc il en existe une dernière avant le temps d'explosion. Soit T_{n_2} le temps de cette dernière transition et G_j^* , $j < k$, l'espace d'arrivée ($Y_{n_2} \in G_j^*$), alors, pour tout $t \in [T_{n_2}, T[$, $X_t \in G_j^*$. La troisième condition est donc aussi satisfaite. \square

On définit le noyau $\Lambda(\cdot, \cdot)$ par :

$$\forall Q \in G^*, p \in G, \quad \Lambda(Q, p) = \frac{\mu(p)\delta_p(Q)}{\mu(Q)} \quad (2.12)$$

de sorte que Λ est continu en la première variable et que \mathcal{L} satisfait bien la condition de Λ -intégrabilité.

Lemme 2.2.5. \mathcal{L} et \mathcal{L}^* vérifient $\mathcal{L}^*\Lambda = \Lambda\mathcal{L}$.

Démonstration. Soient $Q \in G^*$ et $p \in G$. Si $p \notin Q$, alors :

$$\begin{aligned} \sum_{Q' \in G^*} \mathcal{L}_{Q, Q'}^* \Lambda(Q', p) &= \sum_{q \in G} \mathcal{L}_{Q, Q \cup \{q\}}^* \frac{\mu(p)\delta_p(Q \cup \{q\})}{\mu(Q \cup \{q\})} \\ &= \mathcal{L}_{Q, Q \cup \{p\}}^* \frac{\mu(p)}{\mu(Q \cup \{p\})} \\ &= \frac{\mu(p)}{\mu(Q)} \sum_{q \in Q} \mathcal{L}_{p, q} \\ &= \sum_{q \in Q} \frac{\mu(q)}{\mu(Q)} \mathcal{L}_{q, p} \\ &= \sum_{q \in Q} \Lambda(Q, q) \mathcal{L}_{q, p} \end{aligned}$$

D'autre part, si $p \in Q$, alors :

$$\begin{aligned} \sum_{Q' \in G^*} \mathcal{L}_{Q, Q'}^* \Lambda(Q', p) &= \sum_{q \neq Q} \mathcal{L}_{Q, Q \cup \{q\}}^* \frac{\mu(p)}{\mu(Q \cup \{q\})} + \sum_{q \neq p \in Q} \sum_{Q' \in \mathcal{C}(Q \setminus \{q\})} \mathcal{L}_{Q, Q'}^* \frac{\mu(p)\delta_p(Q')}{\mu(Q')} \\ &\quad + \mathcal{L}_{Q, Q}^* \Lambda(Q, p) \\ &= \sum_{q \neq Q} \sum_{q' \in Q} \frac{\mu(p)}{\mu(Q)} \mathcal{L}_{q, q'} + \sum_{q \neq p \in Q} \sum_{q' \neq Q} \frac{\mu(p)}{\mu(Q)} \mathcal{L}_{q, q'} \\ &\quad - \frac{\mu(p)}{\mu(Q)} \left(\sum_{q \neq Q} \left(1 + \frac{\mu(q)}{\mu(Q)} \right) \sum_{q' \in Q} \mathcal{L}_{q, q'} + \sum_{q \in Q} \left(1 - \frac{\mu(q)}{\mu(Q)} \right) \sum_{q' \neq Q} \mathcal{L}_{q, q'} \right) \\ &= \frac{\mu(p)}{\mu(Q)} \left(\sum_{q \neq Q, q' \in Q} \frac{\mu(q)}{\mu(Q)} \mathcal{L}_{q, q'} - \sum_{q \in Q, q' \neq Q} \frac{\mu(q)}{\mu(Q)} \mathcal{L}_{q, q'} - \sum_{q' \neq Q} \mathcal{L}_{p, q'} \right) \\ &= \frac{\mu(p)}{\mu(Q)} \left(\sum_{q \neq Q, q' \in Q} \frac{\mu(q')}{\mu(Q)} \mathcal{L}_{q', q} - \sum_{q \in Q, q' \neq Q} \frac{\mu(q)}{\mu(Q)} \mathcal{L}_{q, q'} - \sum_{q' \neq Q} \mathcal{L}_{p, q'} \right) \\ &= - \sum_{q' \neq Q} \frac{\mu(p)}{\mu(Q)} \mathcal{L}_{p, q'} \\ &= \sum_{q \in Q} \frac{\mu(p)}{\mu(Q)} \mathcal{L}_{p, q} \\ &= \sum_{q \in Q} \frac{\mu(q)}{\mu(Q)} \mathcal{L}_{q, p} \end{aligned}$$

On a donc montré que $(\mathcal{L}^*\Lambda)_{Q, p} = (\Lambda\mathcal{L})_{Q, p}$. □

Grâce aux deux lemmes précédents et au théorème 1.2.7 on déduit que quelle que soit la loi initiale μ_0 de X , il existe un processus stratifié de générateur \mathcal{L}^* qui soit Λ -lié à X (en prenant comme loi initiale pour X^* , par exemple, $\mu_0^*(A) = \sum_{x \in \mathcal{S}} \mu_0(x) \mathbf{1}_{A=\{x\}}$ pour tout $A \in \mathcal{S}^*$). Dans ce cas, par définition de Λ , X^* est aussi un dual de stationnarité forte pour X

On remarque facilement qu'il existe $N + 1$ états absorbants pour un processus X^* de générateur \mathcal{L}^* : chacun des singletons $\{\Delta_i\}$ et \bar{G} tout entier. Cependant, en entrelaçant X et X^* par Λ , on a :

$$\begin{aligned}
0 &= \mathbb{P}(X_t = \Delta_i \mid X_0^* \neq \{\Delta_i\}) \\
&\geq \mathbb{P}(X_t = \Delta_i \mid X_t^* = \{\Delta_i\}, X_0^* \neq \{\Delta_i\}) \mathbb{P}(X_t^* = \{\Delta_i\} \mid X_0^* \neq \{\Delta_i\}) \\
&\geq \mathbb{P}(X_t = \Delta_i \mid X_t^* = \{\Delta_i\}) \mathbb{P}(X_t^* = \{\Delta_i\} \mid X_0^* \neq \{\Delta_i\}) \\
&= \mathbb{P}(X_t^* = \{\Delta_i\} \mid X_0^* \neq \{\Delta_i\})
\end{aligned} \tag{2.13}$$

car X est non-explosif. On en conclut que le seul état absorbant pouvant être atteint en temps fini est \bar{G} .

2.2.3 Comportement du processus dual

Il nous reste donc à étudier le temps d'absorption de X^* . Plus précisément, on va étudier le comportement des bornes de X^* sur chaque branche infinie en les comparant à des processus de vie et de mort sur \mathbb{N} .

Pour tout $i \in I$, pour tout $p \in \mathbb{N}$, on pose :

$$G_p^i := \bar{G} \setminus \varphi_i(\llbracket p + 1, \infty \rrbracket) = \bigcup_{j \neq i} \bar{Q}_j \cup C_G \cup \varphi_i(\llbracket 0, p \rrbracket) \in G^* \tag{2.14}$$

(c'est-à-dire l'élément de G^* contenant \bar{G} tout entier sauf les points de \bar{Q}_i au-delà de $\varphi_i(p)$), et :

$$\begin{cases} \mathcal{L}_{p,p+1}^i := \left(\frac{\mu(G_{p+1}^i)}{\mu(G_p^i)} \right) \mathcal{L}_{\varphi_i(p+1), \varphi_i(p)} = \mathcal{L}_{G_p^i, G_{p+1}^i}^* \\ \mathcal{L}_{p+1,p}^i := \left(\frac{\mu(G_p^i)}{\mu(G_{p+1}^i)} \right) \mathcal{L}_{\varphi_i(p+1), \varphi_i(p+2)} = \mathcal{L}_{G_{p+1}^i, G_p^i}^* \\ \mathcal{L}_{p,p}^i = -\mathcal{L}_{p,p+1}^i - \mathcal{L}_{p,p-1}^i \end{cases}$$

On définit ainsi des matrices $(\mathcal{L}_{p,q}^i)_{p,q \in \mathbb{N}}$ de générateurs de processus de vie et de mort sur \mathbb{N} .

Lemme 2.2.6. *Il existe $\lambda > 0$ tel que, pour tout $i \in I$ et toute mesure de probabilité μ_0^* sur G^* , il existe une famille (X^*, Y, Z^i) telle que :*

- Le couple (X^*, Z^i) est markovien et ses marginales X^* et Z^i sont des processus de Markov, de générateurs respectifs \mathcal{L}^* et \mathcal{L}^i .
- Y est une variable aléatoire de loi exponentielle de paramètre λ .
- Y et Z^i sont indépendantes.
- X^* a pour loi initiale μ_0^* .
- Pour tout $t > s > 0$:

$$\mathbb{P}(X_t^* \cap \varphi_i(\llbracket Z_t^i, \infty \rrbracket) \neq \emptyset \mid t < Y, X_0^* \notin G_0^i) = 1$$

$$\mathbb{P}(X_s^* \cap C_G \subset X_t^* \mid t < Y) = 1$$

Démonstration. Nous allons commencer par distinguer différents types de transition pour le processus X^* . Le premier type de transition est celui qui ne modifie X^* que d'un point situé sur Q_i . Les transitions concernées sont le passage d'un état Q à l'état $Q \cup \{\varphi_i(p)\}$, ou à l'état $Q \setminus \{\varphi_i(p-1)\}$, où $p = \sup\{q \in \mathbb{N}; \varphi_i(q) \in Q\} + 1$. Ces transitions nous « éloignent ou nous rapprochent d'une étape » d'un état contenant le point Δ_i . On notera donc $A_i^+(Q) = Q \cup \{\varphi_i(p)\}$, $A_i^-(Q) = Q \setminus \{\varphi_i(p-1)\}$ et $A_i(Q) = \{A_i^+(Q), A_i^-(Q)\}$, pour tout $Q \in G^*$.

Le deuxième type de transition est celui qui est susceptible d'enlever plusieurs points d'un coup. Pour cela, il faut $p \in Q \in G^*$ tels que $Q \setminus \{p\}$ ne soit pas connexe et $p \in V(Q^c)$. Ceci n'est possible que si $p \in C_G$. On pose donc naturellement, pour tout $Q \in G^*$:

$$B(Q) := \{Q' \in \mathcal{C}(Q \setminus \{p\}); p \in Q \cap C_G\}$$

Dans le troisième type on regroupe toutes les autres transitions, celles qui n'influent pas dans notre problématique :

$$C_i(Q) = \{Q' \in G^* : \mathcal{L}_{Q,Q'} > 0, Q' \notin A_i(Q) \cup B(Q)\}$$

pour tout $Q \in G^*$. Ces transitions sont celles qui enlèvent un point de Q_j , $j \neq i$ ou qui ajoutent un point qui n'est pas dans Q_i .

On parlera de transition de type A_i^+ (resp. A_i^- , resp. B , resp. C) pour une transition d'un état $Q \in G^*$ à un état $Q' = A_i^+(Q)$ (resp $Q' = A_i^-$, resp. $Q' \in B(Q)$, resp. $Q' \in C_i(Q)$).

On remarque que l'ensemble $\{\mathcal{L}_{Q,Q'} : Q \in G^*, Q' \in B(Q)\}$ est borné. En effet, si $Q \in G^*$ est tel que $B(Q)$ est non vide, alors il contient un point de C_G , donc :

$$\mu(Q) \geq \inf\{\mu(p); p \in C_G\} =: a_1$$

et si $p_0 \in Q \cap C_G$, alors :

$$\sum_{q \neq Q} \mathcal{L}_{q,p_0} \leq \sup_{p \in C_G} \sum_{q \in V(C_G)} \mathcal{L}_{q,p} =: a_2$$

ce qui donne pour tout $Q' \in \mathcal{C}(Q \setminus \{p_0\})$:

$$\begin{aligned} \mathcal{L}_{Q,Q'}^* &= \frac{\mu(Q')}{\mu(Q)} \sum_{q \notin Q} \mathcal{L}_{q,p_0} \\ &\leq \frac{1}{a_1} a_2 \end{aligned}$$

L'ensemble $\{|B(Q)|; Q \in G^*\}$ étant également borné par une constante a_3 (grossièrement, $N + |C_G|$), on en déduit :

$$\sum_{Q' \in B(Q)} \mathcal{L}_{Q,Q'}^* \leq \frac{a_2 a_3}{a_1} =: \lambda$$

cette constante ne dépendant pas de Q

On observe enfin que pour tout $p \in \mathbb{N}$, si $\varphi_i(p) \in Q$ et $\varphi_i(p+1) \notin Q$, $Q \in G^*$, alors $\mathcal{L}_{Q,A_i^+(Q)}^* \geq \mathcal{L}_{p,p+1}^i$ et $\mathcal{L}_{Q,A_i^-(Q)}^* \leq \mathcal{L}_{p,p-1}^i$.

Nous allons maintenant, grâce aux remarques précédentes, construire la famille (X^*, Y, Z^i) en empêchant toute transition de type B avant le temps Y , et en forçant Z^i à descendre en cas de transition de type A_i^- ou en imposant une transition de type A_i^+ si Z^i monte, lorsque cela est nécessaire. Bien sûr, toute cette construction se fait uniquement lorsque X_0^* contient au moins un point de Q_i .

Construction de la famille (X^*, Y, Z^i) : Dans le but de se ramener à un processus de Markov, on va construire (X^*, N, Z^i) , où N est un processus de Poisson d'intensité λ ; Y est alors le premier temps de saut de N . On commence par simuler X^* selon la loi μ_0^* . Si $X_0^* \cap Q_i = \emptyset$, on arrête là la construction (on laisse par exemple X^* et Z^i évoluer indépendamment selon leur générateur respectif et on tire Y indépendamment). Sinon, on pose $Z_0^i = \sup\{p \in \mathbb{N} : \varphi_i(p) \in X_0^*\}$ et $T_0 = 0$.

On suppose (X^*, N, Z^i) construit jusqu'au temps T_n , $n \geq 0$, et $Z_{T_n}^i = p$ et $X_{T_n}^* = Q$. Si $N_t \geq 1$, la construction est terminée est on laisse X^* et Z évoluer indépendamment. Sinon, on distingue alors deux cas :

— *Premier cas* : $\varphi_i(Z_{T_n}^i) \notin A_i^-(X_{T_n}^*)$

C'est le cas où $Z_{T_n}^i$ se trouve « au bout » de $X_{T_n}^*$, et il faut donc faire attention à ce qu'il n'en sorte pas au temps T_{n+1} . On tire alors une variable exponentielle ϵ_n^1 de paramètre λ et une autre, ϵ_n^2 , indépendante de ϵ_n^1 et de paramètre :

$$\Lambda_n^2 := \sum_{Q' \in C_i(Q)} \mathcal{L}_{Q, Q'}^* + \mathcal{L}_{Q, A^+(Q)}^* + \mathcal{L}_{p, p-1}^i$$

On pose $T_{n+1} = T_n + \epsilon_n^1 \wedge \epsilon_n^2$. Si $\epsilon_n^1 < \epsilon_n^2$, on impose $N_{T_{n+1}} = 1$ et on choisit $X_{T_{n+1}}^* = Q'$ avec probabilité $\frac{\mathcal{L}_{Q, Q'}^*}{\lambda}$, pour tout $Q' \in B(Q)$. La somme de ces probabilités peut être strictement inférieure à 1, et si aucun de ces états n'est choisi, on reste en Q .

Si $\epsilon_n^1 > \epsilon_n^2$, on laisse $N_{T_{n+1}}$ en 0, et on pose :

$$(X_{T_{n+1}}^*, Z_{T_{n+1}}^i) = \begin{cases} (Q', p) & \text{avec probabilité } \frac{\mathcal{L}_{Q, Q'}^*}{\Lambda_n^2}, Q' \in C_i(Q) \\ (Q', p+1) & \text{avec probabilité } \frac{\mathcal{L}_{p, p+1}^i}{\Lambda_n^2}, Q' = A_i^+(Q) \\ (Q', p) & \text{avec probabilité } \frac{\mathcal{L}_{Q, Q'}^* - \mathcal{L}_{p, p+1}^i}{\Lambda_n^2}, Q' = A_i^+(Q) \\ (Q, p-1) & \text{avec probabilité } \frac{\mathcal{L}_{p, p-1}^i - \mathcal{L}_{Q, A_i^-(Q)}^*}{\Lambda_n^2} \\ (Q', p-1) & \text{avec probabilité } \frac{\mathcal{L}_{p, p-1}^i}{\Lambda_n^2}, Q' = A_i^-(Q) \end{cases}$$

— *Deuxième cas* : $\varphi_i(Z_{T_n}^i) \in A_i^-(X_{T_n}^*)$

Dans ce cas, il n'y a pas besoin de s'assurer que « le bout » de X^* reste au-dessus de Z^i , mais seulement de vérifier si une transition de type B a lieu. On tire donc une variable exponentielle ϵ_n^1 de paramètre λ et une autre, ϵ_n^2 , indépendante de ϵ_n^1 et de paramètre :

$$\Lambda_n^2 := \sum_{Q' \in C_i(Q) \cup A_i(Q)} \mathcal{L}_{Q, Q'}^* + \mathcal{L}_{p, p-1}^i + \mathcal{L}_{p, p-1}^i$$

On pose $T_{n+1} = T_n + \epsilon_n^1 \wedge \epsilon_n^2$. Si $\epsilon_n^1 < \epsilon_n^2$, on procède exactement comme dans le premier cas. Sinon, on choisit :

$$(X_{T_{n+1}}^*, Z_{T_{n+1}}^i) = \begin{cases} (Q', p) & \text{avec probabilité } \frac{\mathcal{L}_{Q, Q'}^*}{\Lambda_n^2}, Q' \in C_i(Q) \cup A_i(Q) \\ (Q, p+1) & \text{avec probabilité } \frac{\mathcal{L}_{p, p+1}^i}{\Lambda_n^2} \\ (Q, p-1) & \text{avec probabilité } \frac{\mathcal{L}_{p, p-1}^i}{\Lambda_n^2} \end{cases}$$

Par construction, on a bien $X_t^* \cap \varphi_i(\llbracket Z_t^i, \infty \rrbracket) \neq \emptyset$ si $t < Y$, et Y et Z^i indépendantes, jusqu'au premier temps d'explosion $\tau_1 := \lim T_n$. On vérifie comme dans le Lemme 2.1.3 que les processus ont bien les générateurs recherchés. Partant du premier temps d'explosion, on recommence la construction de la même manière. \square

Lemme 2.2.7. *Pour tout $i \in I$, si Z^i est un processus de générateur \mathcal{L}^i , alors le temps d'explosion de Z^i est fini presque sûrement si et seulement si la condition (22) :*

$$\sum_{j=1}^{\infty} \mu^i(j+1) \sum_{k=1}^j \frac{1}{\mu^i(k) \mathcal{L}_{\varphi_i(k), \varphi_i(k+1)}} < \infty \quad (2.15)$$

est remplie. Sinon, il est non-explosif.

Démonstration. On rappelle que le critère d'explosion pour Z^i est :

$$\sum_{i=1}^{\infty} \prod_{j=1}^i \frac{\mathcal{L}_{j,j-1}^i}{\mathcal{L}_{j,j+1}^i} \sum_{k=1}^i \prod_{l=1}^k \frac{\mathcal{L}_{j-1,j}^i}{\mathcal{L}_{j,j-1}^i} < \infty$$

(cf [3], section 8.1). C'est-à-dire :

$$\begin{aligned} \infty &> \sum_{n=1}^{\infty} \prod_{j=1}^n \frac{\left(\frac{\mu(G_{j-1}^i)}{\mu(G_j^i)} \right) \mathcal{L}_{\varphi_i(j), \varphi_i(j+1)}}{\left(\frac{\mu(G_{j+1}^i)}{\mu(G_j^i)} \right) \mathcal{L}_{\varphi_i(j+1), \varphi_i(j)}} \sum_{k=1}^n \prod_{l=1}^k \frac{\left(\frac{\mu(G_j^i)}{\mu(G_{j-1}^i)} \right) \mathcal{L}_{\varphi_i(j), \varphi_i(j-1)}}{\left(\frac{\mu(G_{j-1}^i)}{\mu(G_j^i)} \right) \mathcal{L}_{\varphi_i(j), \varphi_i(j+1)}} \\ &\geq \sum_{n=1}^{\infty} \prod_{j=1}^n \frac{\mu(G_{j-1}^i) \mathcal{L}_{\varphi_i(j), \varphi_i(j+1)}}{\mu(G_{j+1}^i) \mathcal{L}_{\varphi_i(j+1), \varphi_i(j)}} \sum_{k=1}^n \prod_{l=1}^k \frac{\mathcal{L}_{\varphi_i(j), \varphi_i(j-1)}}{\mathcal{L}_{\varphi_i(j), \varphi_i(j+1)}} \\ &\geq \sum_{n=1}^{\infty} \prod_{j=1}^n \frac{\mu(G_0^i) \mathcal{L}_{\varphi_i(j), \varphi_i(j+1)}}{\mathcal{L}_{\varphi_i(j+1), \varphi_i(j)}} \sum_{k=1}^n \prod_{l=1}^k \frac{\mathcal{L}_{\varphi_i(j), \varphi_i(j-1)}}{\mathcal{L}_{\varphi_i(j), \varphi_i(j+1)}} \\ &\geq \frac{\mu(G_0^i)}{\mu(\varphi_i(0))} \sum_{n=1}^{\infty} \mu(\varphi_i(n+1)) \sum_{k=1}^n \frac{1}{\mu(\varphi_i(k)) \mathcal{L}_{\varphi_i(k), \varphi_i(k+1)}} \\ &= \frac{\mu(G_0^i)}{\mu(\varphi_i(0))} \sum_{n=1}^{\infty} \mu^i(n+1) \sum_{k=1}^n \frac{1}{\mu^i(k) \mathcal{L}_{\varphi_i(k), \varphi_i(k+1)}} \end{aligned}$$

où les G_j^i sont les sous-ensembles de G^* définis en (2.14). \square

L'intérêt des générateurs \mathcal{L}^i est donc de fournir une condition suffisante à l'explosion d'un processus X^* de générateur \mathcal{L}^* le long de chacune des branches Q_i . On obtient une condition nécessaire de manière similaire en définissant, pour tout $i \in I$, pour tout $n \in \mathbb{N}$ et $p \geq n$:

$$\begin{cases} \mathcal{L}_{p,p+1}^{i,n} := \left(1 + \frac{\mu(\varphi_i(p+1))}{\mu(\varphi_i(\llbracket n, p \rrbracket))} \right) \mathcal{L}_{\varphi_i(p+1), \varphi_i(p)} = \mathcal{L}_{\varphi_i(\llbracket n, p \rrbracket), \varphi_i(\llbracket n, p+1 \rrbracket)}^* \\ \mathcal{L}_{p,p-1}^{i,n} := \left(1 - \frac{\mu(\varphi_i(p))}{\mu(\varphi_i(\llbracket n, p \rrbracket))} \right) \mathcal{L}_{\varphi_i(p-1), \varphi_i(p)} = \mathcal{L}_{\varphi_i(\llbracket n, p \rrbracket), \varphi_i(\llbracket n, p-1 \rrbracket)}^* \\ \mathcal{L}_{p,p}^{i,n} = -\mathcal{L}_{p,p+1}^{i,n} - \mathcal{L}_{p,p-1}^{i,n} \end{cases}$$

Les matrices $(\mathcal{L}_{p,q}^{i,n})_{p,q \geq n}$ définissent des générateurs de processus de vie et de mort sur $\llbracket n, +\infty \rrbracket$. Exactement comme précédemment, on a les deux lemmes suivants :

Lemme 2.2.8. Pour chaque $i \in I$, et pour toute mesure de probabilité μ_0^* sur G^* , il existe un processus X^* de générateur \mathcal{L}^* et de loi initiale μ_0^* et une suite $(Z^{i,n})_{n \in \mathbb{N}}$ de processus de vie et de mort tels que, pour tout $n \in \mathbb{N}$, $Z^{i,n}$ a pour générateur $\mathcal{L}^{i,n}$ et :

$$\mathbb{P}(X_t^* \cap \varphi_i(\llbracket Z_t^{i,n}, \infty \rrbracket) \subset \{\varphi_i(Z_t^{i,n})\} \mid \Delta_i \notin X_0^*, t < T_n) = 1$$

pour tout $t > 0$, avec $T_n := \inf\{s > 0; X_s^* \cap G_n^i = \emptyset\}$.

Démonstration. Il suffit d'adapter légèrement la preuve du lemme 2.1.3. Cette fois on laisse X^* et chacun des $Z^{i,n}$ évoluer indépendamment jusqu'au temps

$$\tau := \inf\{t > 0; \exists n \in \mathbb{N} : Z_t^{i,n} \in X_t^*\}$$

puis on applique la même méthode pour forcer X^* à rester dans $G_{Z^{i,n}}^i$. Lorsque X^* sort entièrement de G_n^i , les taux de transitions sont tels qu'il devient impossible de le coupler ainsi avec $Z^{i,n}$ \square

Lemme 2.2.9. Si $Z^{i,n}$ est un processus de générateur $\mathcal{L}^{i,n}$, avec $i \in I$, $n \in \mathbb{N}$ alors le temps d'explosion de $Z^{i,n}$ est fini presque sûrement si et seulement si :

$$\sum_{j=n+1}^{\infty} \mu^i(j+1) \sum_{k=n+1}^j \frac{1}{\mu^i(k) \mathcal{L}_{\varphi_i(k), \varphi_i(k+1)}} < \infty \quad (2.16)$$

Théorème 2.2.10. Soit X^* un processus de générateur \mathcal{L}^* . S'il existe $i \in I$ tel que la condition (22) n'est pas remplie, alors pour tout $t > 0$:

$$\mathbb{P}(\Delta_i \in X_t^* \mid \Delta_i \notin X_0^*) = 0$$

Réciproquement, si pour tout $i \in I$ la condition (22) est remplie alors, presque sûrement, il existe $T < \infty$ tel que $X_t^* = \bar{G}$, $\forall t > T$.

Démonstration. La première partie se démontre à partir des lemmes 2.2.8 et 2.2.9. En effet, soit $i \in I$ tel que la condition (22) ne soit pas remplie. On constate immédiatement que pour tout $n \in \mathbb{N}$, la condition (2.16) ne l'est pas non plus. Soit donc une famille de processus $(\tilde{X}^*, (Z^{i,n})_{n \in \mathbb{N}})$ comme dans le lemme 2.2.8, \tilde{X}^* ayant même loi que X^* . On a alors, pour tout $t > 0$:

$$\begin{aligned} \mathbb{P}(\Delta_i \in X_t^* \mid \Delta_i \notin X_0^*) &= \mathbb{P}(\Delta_i \in \tilde{X}_t^* \mid \Delta_i \notin \tilde{X}_0^*) \\ &\leq \mathbb{P}\left(\bigcap_{n \in \mathbb{N}} \{\tilde{X}_t^* \cap \varphi_i(\llbracket Z_t^{i,n} + 1, \infty \rrbracket) \neq \emptyset\} \mid \Delta_i \notin \tilde{X}_0^*\right) \\ &\leq \mathbb{P}\left(\bigcap_{n \in \mathbb{N}} \{t \geq T_n\} \mid \Delta_i \notin \tilde{X}_0^*\right) \\ &= \mathbb{P}(\exists s \leq t; \tilde{X}_s^* = \{\Delta_i\} \mid \Delta_i \notin \tilde{X}_0^*) \\ &= \mathbb{P}(\tilde{X}_t^* = \{\Delta_i\} \mid \Delta_i \notin \tilde{X}_0^*) \end{aligned}$$

et cette dernière probabilité vaut 0 d'après (2.13).

La deuxième partie se démontre grâce aux lemmes 2.2.6 et 2.2.7. Remarquons que puisque X^* a la loi d'un dual de X , lequel est récurrent positif, X^* va atteindre C_G en temps fini presque sûrement. On peut donc supposer sans perte de généralité que le processus commence à cet instant, et donc que X_0^* contient au moins un point de C_G .

Lorsque le processus X^* atteint un état contenant un Δ_i (nécessairement après un temps d'explosion), celui-ci ne peut sortir de X^* qu'en même temps que tout le reste

de la branche Q_i et que le point p_i de C_G voisin de Q_i . Cet événement correspond à une transition de type B (cf démonstration du lemme 2.2.6). L'idée est donc de montrer que le processus peut atteindre successivement chacun des Δ_i sans perdre de points de C_G (et donc sans perdre les Δ_i déjà gagnés). Pour cela, on minore, uniformément sur les conditions initiales, les probabilités d'atteindre chacun des Δ_i avant une transition de type B . Cette minoration se fait en deux temps : d'abord, on minore la probabilité d'atteindre $\varphi_i(0)$ en suivant un chemin donné, pour ensuite coupler notre processus avec Y et Z^i comme dans le lemme 2.2.6.

Pour chaque $i \in I$, on note $f_i : G^* \rightarrow G^*$ la fonction définie par :

$$f_i(Q) = Q \cap (C_G \cup \{\varphi_i(0)\})$$

Un moyen de garantir l'absence de transition de type B sur un intervalle de temps $[0, t]$ est de demander que $f_i(X^*)$ soit croissante (au sens de l'inclusion) sur cet intervalle.

Soient $t > 0$, $i \in I$, $C \subset C_G \cup \{\varphi_i(0)\}$ tel que $C \not\subseteq \{\varphi_i(0)\}$ (donc C contient au moins un point de C_G), $Q \in G^*$ tel que $f_i(Q) = C$ et (q_0, \dots, q_{n_0}) un chemin avec $q_0 \in C$, $q_1, \dots, q_{n_0-1} \in C_G \setminus C$ et $q_{n_0} = \varphi_i(0)$. Si $n_0 = 0$, c'est que $\varphi_i(0) \in Q$. On note $C^n = C \cup \{q_1\} \cup \dots \cup \{q_n\}$, $0 \leq n \leq n_0$. On note également T_k le k -ème temps de saut de X^* , $k \in \mathbb{N}$, avec $T_0 = 0$, $(Y_k)_{k \in \mathbb{N}}$ la chaîne de sauts sous-jacente : $X_{T_k}^* = Y_k$ pour tout $k \in \mathbb{N}$ et $T_{\phi(k)}$ la sous-suite des temps de sauts de $f_i(X)$. Les processus $(T_{\phi(k)})$ et $(f_i(Y_{\phi(k)}))$, traduisant l'évolution de X^* sur $C_G \cup \{\varphi_i(0)\}$, ne sont pas markoviens, mais on peut aisément minorer les probabilités de transition à l'aide du processus X^* . Le premier objectif est de minorer $\mathbb{P}(f_i(X_t^*) = C^n, f_i(X^*) \text{ croissante sur } [0, t] \mid X_0^* = Q)$ par une constante ne dépendant pas de Q . On commence grossièrement :

$$\begin{aligned} & \mathbb{P}(f_i(X_t^*) = C^n, f_i(X^*) \text{ croissante sur } [0, t] \mid X_0^* = Q) \\ & \geq \mathbb{P} \left(\bigcap_{k=1}^{n_0} \{T_{\phi(k)} - T_{\phi(k-1)} \leq t/n_0, f_i(Y_{\phi(k)}) = C^k, T_{\phi(n_0+1)} - T_{\phi(n_0)} > t\} \mid X_0^* = Q \right) \\ & = \prod_{k=1}^{n_0} \mathbb{P}(\epsilon_k \in [0, t/n_0]) p_k \times \mathbb{P}(\epsilon_{n_0+1} > t/n_0) \end{aligned} \quad (2.17)$$

où pour tout $k \leq n_0 + 1$:

$$\epsilon_k \sim \mathcal{L}(T_{\phi(k)} - T_{\phi(k-1)} \mid \bigcap_{n=1}^{k-1} \{f_i(Y_{\phi(n)}) = C^n, T_{\phi(n)} - T_{\phi(n-1)} \in [0, t/n_0], X_0^* = Q\})$$

et pour tout $k \leq n_0$:

$$p_k = \mathbb{P}(f_i(Y_{\phi(k)}) = C^k \mid \bigcap_{n=0}^{k-1} \{f_i(Y_{\phi(n)}) = C^n, T_{\phi(n+1)} - T_{\phi(n)} \in [0, t/n_0], X_0^* = Q\})$$

(noter que dans la probabilité conditionnelle ci-dessus, $Y_0 = C^0$ n'est pas nécessaire car impliqué par $X_0^* = Q$, et n'a été laissé que par commodité de notation).

Si en un temps $T_{\phi(k)}$ le processus X^* passe d'un état Q' à Q'' alors on a $Q'' = Q' \cup \{p\}$, $p \in f_i(V(Q))$ ou $Q'' \in \mathcal{C}(Q \setminus \{p\})$, $p \in f_i(Q')$, et on vérifie facilement à partir des coefficients de \mathcal{L}^* que :

$$\left(\min_{q \in V(f_i(G))} \mu(q) \right) \left(\min_{\substack{q \in V(C_G) \\ q' \in V(q)}} \mathcal{L}_{q,q'} \right) \leq \mathcal{L}_{Q',Q''}^* \leq \frac{\max_{q \in V(C_G)} |\mathcal{L}_{q,q}|}{\min_{q' \in C_G} \mu(q')}$$

On appelle $\alpha_1 < \infty$ le membre de gauche et $\alpha_2 > 0$ celui de droite. On vérifie également que pour un Q' donné, le nombre de Q'' possible de la forme $Q' \cup \{p\}$ est majoré par $|C_G|$,

et $|\mathcal{C}(Q' \setminus \{p\})|$ avec $p \in f_i(Q')$ est majoré par $(|C_G| + 1)^2$. En conditionnant par $X_{T_{\phi(k)-1}}^*$, on obtient donc :

$$\begin{aligned} p_k &\geq \inf \{ \mathbb{P}(X_{T_{\phi(k)}}^* = Q' \cup \{p\} \mid X_{T_{\phi(k)-1}}^* = Q'), f_i(Q') \neq \emptyset, p \in V(Q') \cap V(C_G) \} \\ &= \inf \left\{ \frac{\mathcal{L}_{Q', Q' \cup \{p\}}^*}{\sum_{\substack{Q'' \in G^* \\ f_i(Q'') \neq f_i(Q')}} \mathcal{L}_{Q', Q''}^*}, f_i(Q') \neq \emptyset, p \in V(Q') \cap V(C_G) \right\} \\ &\geq \frac{\alpha_2}{(|C_G| + (|C_G| + 1)^2)\alpha_1} =: \alpha_0 > 0 \end{aligned}$$

et, conditionnellement à $X_{T_{\phi(k-1)}}^*$ on peut encadrer $T_{\phi(k)} - T_{\phi(k-1)}$ par deux variables exponentielles E_1 (inférieurement) et E_2 (supérieurement), de paramètres respectifs :

$$\begin{aligned} \sup \left\{ \sum_{\substack{Q'' \in G^* \\ f_i(Q'') \neq f_i(Q')}} \mathcal{L}_{Q', Q''}^* : Q' \in G^*, f_i(Q') \neq \emptyset \right\} &\leq (|C_G| + (|C_G| + 1)^2)\alpha_1 \\ \inf \left\{ \sum_{\substack{Q'' \in G^* \\ f_i(Q'') \neq f_i(Q')}} \mathcal{L}_{Q', Q''}^* : Q' \in G^*, f_i(Q') \neq \emptyset \right\} &\geq \alpha_2 \end{aligned}$$

Autrement dit :

$$\mathbb{P}(\epsilon_k \in [0, t/n_0]) \geq \mathbb{P}(E_2 \in [0, t/n_0]) \geq 1 - \exp(-\alpha_2 t/n_0)$$

et :

$$\mathbb{P}(\epsilon_{n_0+1} > t/n_0) \geq \mathbb{P}(E_1 > t/n_0) \geq \exp(-(|C_G| + (|C_G| + 1)^2)\alpha_1 t/n_0)$$

En revenant à (2.17) et en utilisant le fait que $n_0 \leq |C_G|$, on voit qu'on a minoré la probabilité d'atteindre $\varphi_i(0)$ en temps t et avant une transition de type B par une constante ε_t ne dépendant ni de i , ni de la configuration initiale C . Nous avons donc réalisé la première étape de la minoration.

Partant de l'état atteint lors de cette première étape, on construit (X^*, Y, Z^i) comme dans le lemme 2.2.6. La probabilité d'atteindre un état contenant Δ_i avant une transition de type B , en temps t est minorée par $\mathbb{P}(Z_t^i = \infty, Y > t \mid Z_0^i = \varphi_i(0))$ d'après la loi du couple (X, Z^i) , et cette probabilité est strictement positive car Z^i et Y sont indépendants. En réalisant les deux étapes consécutivement (chacune en temps $t/2$ par exemple), on obtient donc :

$$\mathbb{P}(\Delta_i \in X_t^*, f_i(X^*) \text{ croissante sur } [0, t] \mid X_0^*, X_0^* \cap C_G \neq \emptyset) \geq \varepsilon_t^i$$

constante ne dépendant que de i et t . Par markoviannité et homogénéité, il vient :

$$\begin{aligned} &\mathbb{P}(\Delta_i \in X_{Nt}^*, \forall i \in I \mid X_0^* \cap C_G \neq \emptyset) \\ &\geq \prod_{i=1}^N \inf_{\substack{Q \in G^* \\ Q \cap C_G \neq \emptyset}} \mathbb{P}(\Delta_i \in X_{it}^*, f_i(X^*) \text{ croissante sur } [(i-1)t, it] \mid X_{(i-1)t}^* = Q) \\ &\geq \prod_{i=1}^N \inf_{\substack{Q \in G^* \\ Q \cap C_G \neq \emptyset}} \mathbb{P}(\Delta_i \in X_t^*, f_i(X^*) \text{ croissante sur } [0, t] \mid X_0^* = Q) \\ &\geq \prod_{i \in I} \varepsilon_t^i > 0 \end{aligned}$$

L'ensemble G_N^* des états contenant tous les Δ_i (défini en (2.10)) est donc récurrent positif. Il ne reste plus qu'à conclure en remarquant que d'un état $Q \in G_N^*$, on peut passer à \bar{G} en un nombre fini de sauts : si q_1, \dots, q_{n_0} sont tels que $\{q_i, i \leq n_0\} = C_G \setminus Q$ et $q_i \in V(Q \cup \{q_1\} \cup \dots \cup \{q_{i-1}\})$, alors

$$\mathbb{P}(X_{T_i}^* = Q \cup \{q_1\} \cup \dots \cup \{q_i\}, \forall i \leq n_0) > \varepsilon$$

où ε est une constante strictement positive indépendante de Q et n_0 .

\bar{G} est donc récurrent positif, et le processus X^* est absorbé en temps fini presque sûrement. \square

2.2.4 Preuve du théorème 0.4.1

En utilisant les informations que l'on vient d'obtenir sur la dynamique du processus dual X^* , nous sommes déjà en mesure de montrer que la condition 22 dans le théorème 0.4.1 suffit à l'existence d'un temps fort de stationnarité pour toute loi initiale. Un dernier lemme technique nous permettra de montrer que cette condition est nécessaire :

Lemme 2.2.11. *Si X^* est un processus stratifié associé au générateur \mathcal{L}^* , avec $\mathbb{E} \left(\frac{1}{\mu(X_0^*)} \right) < +\infty$, alors $\mathbb{E} \left(\frac{1}{\mu(X_t^*)} \right) < +\infty$ pour tout $t > 0$.*

Démonstration. Soient $A_n := \bigcup_{i \in I} \varphi_i(\llbracket n, \infty \rrbracket)$ et

$$\tau_n := \inf\{t > 0; X_t^* \subset A_n\}$$

On remarque que la suite $(A_n)_{n \in \mathbb{N}}$ est strictement décroissante et tend vers $A_\infty = \{\Delta_i, i \in I\}$ tandis que la suite $(\tau_n)_{n \in \mathbb{N}}$ est strictement croissante (à partir du premier rang n tel que $\tau_n \neq 0$) et tend vers l'infini (car l'ensemble A_∞ ne peut pas être atteint en temps fini). Notamment, pour tout $t > 0$:

$$\mathbb{P}(X_{t \wedge \tau_n}^* \not\subset A_{n+1} \mid X_0^* \not\subset A_{n+1}) = 1$$

De plus, si $X_0^* \not\subset A_{n+1}$, alors pour tout $n \in \mathbb{N}$, on a :

$$\frac{1}{\mu(X_{t \wedge \tau_n}^*)} \leq \sup_{q \in G \setminus A_{n+1}} \frac{1}{\mu(q)} =: K_n$$

On pose alors, pour tout $Q \in G^*$, pour tout $n \in \mathbb{N}$:

$$g_n(Q) = \frac{1}{\mu(Q)} \wedge K_n$$

Ces fonctions sont continues bornées et pour tout n , g_n coïncide avec $\frac{1}{\mu}$ sur l'ensemble $\{Q \in G^*; Q \not\subset A_{n+1}\}$. Utilisant le fait que

$$Q \not\subset A_n \Rightarrow Q' \not\subset A_{n+1}, \forall Q' \in G^* \text{ tel que } \mathcal{L}_{Q, Q'}^* > 0$$

il vient, pour tout $n \in \mathbb{N}$, pour tout $Q \notin A_n$:

$$\begin{aligned}
\mathcal{L}^* \frac{1}{\mu}(Q) &= \mathcal{L}^* g_n(Q) \\
&= \sum_{Q' \neq Q} \mathcal{L}^*_{Q,Q'} \left(\frac{1}{\mu(Q')} - \frac{1}{\mu(Q)} \right) \\
&= \sum_{q \notin Q} \left(\frac{\mu(Q \cup \{q\})}{\mu(Q)} \left(\frac{1}{\mu(Q \cup \{q\})} - \frac{1}{\mu(Q)} \right) \sum_{p \in Q} \mathcal{L}_{q,p} \right) \\
&\quad + \sum_{p \in Q} \sum_{Q' \in \mathcal{C}(Q \setminus \{p\})} \left(\frac{\mu(Q')}{\mu(Q)} \left(\frac{1}{\mu(Q')} - \frac{1}{\mu(Q)} \right) \sum_{q \notin Q} \mathcal{L}_{p,q} \right) \\
&= \sum_{q \notin Q} \frac{-\mu(q)}{\mu(Q)^2} \sum_{p \in Q} \mathcal{L}_{q,p} + \sum_{p \in Q} \sum_{Q' \in \mathcal{C}(Q \setminus \{p\})} \frac{\mu(Q) - \mu(Q')}{\mu(Q)^2} \sum_{q \notin Q} \mathcal{L}_{p,q} \\
&= \sum_{p \in Q} \frac{|\mathcal{C}(Q \setminus \{p\})| - 1}{\mu(Q)} \sum_{q \notin Q} \mathcal{L}_{p,q}
\end{aligned}$$

Les termes de cette somme sont nuls si $p \notin C_G$, on en déduit facilement que cette somme est majorée par une constante K qui ne dépend pas de n . Donc, pour tout $n \in \mathbb{N}$ tel que $X_0^* \notin A_{n+1}$ presque sûrement, le processus :

$$\begin{aligned}
M_t &:= \frac{1}{\mu(X_{t \wedge \tau_n}^*)} - \frac{1}{\mu(X_0^*)} - \int_0^{t \wedge \tau_n} \mathcal{L}^* \frac{1}{\mu}(X_s^*) ds \\
&= g_n(X_{t \wedge \tau_n}^*) - g_n(X_0^*) - \int_0^{t \wedge \tau_n} \mathcal{L}^* g_n(X_s^*) ds
\end{aligned}$$

est une martingale, d'où :

$$\begin{aligned}
\mathbb{E} \left(\frac{1}{\mu(X_{t \wedge \tau_n}^*)} \right) &= \mathbb{E} \left(\frac{1}{\mu(X_0^*)} \right) + \mathbb{E} \left(\int_0^{t \wedge \tau_n} \mathcal{L}^* \frac{1}{\mu}(X_s^*) ds \right) \\
&\leq \mathbb{E} \left(\frac{1}{\mu(X_0^*)} \right) + K \mathbb{E}(t \wedge \tau_n) \\
&\leq \mathbb{E} \left(\frac{1}{\mu(X_0^*)} \right) + Kt
\end{aligned}$$

On conclut par convergence dominée que $\mathbb{E} \left(\frac{1}{\mu(X_t^*)} \right) \leq \mathbb{E} \left(\frac{1}{\mu(X_0^*)} \right) + Kt$. \square

Preuve du théorème 0.4.1. Si la condition (22) est remplie pour tout $i \in I$, alors de manière classique on construit un temps fort de stationnarité fini à partir du temps d'explosion de X^* , dual de X ayant pour générateur \mathcal{L}^* .

Réciproquement, supposons qu'il existe $i \in I$ tel que cette condition ne soit pas remplie. Supposons que la loi initiale μ_0 de X est à support fini. Soit X^* un dual de stationnarité forte de X , de générateur \mathcal{L}^* . On utilise la fonction de séparation \mathfrak{s} , et le théorème de convergence dominée avec le Lemme 2.2.11 :

$$\begin{aligned}
\mathfrak{s}(t) &= \sup_{q \in G} \left(1 - \frac{\mu_t(q)}{\mu(q)} \right) \\
&= \sup_{q \in G} \mathbb{E} \left(1 - \frac{\Lambda(X_t^*, q)}{\mu(q)} \right) \\
&= 1 - \inf_{q \in G} \mathbb{E} \left(\frac{\delta_q(X_t^*)}{\mu(X_t^*)} \right) \\
&\geq 1 - \lim_{q \rightarrow \infty} \mathbb{E} \left(\frac{\delta_{\varphi_i(q)}(X_t^*)}{\mu(X_t^*)} \right) \\
&= 1 - \mathbb{E} \left(\lim_{q \rightarrow \infty} \frac{\delta_{\varphi_i(q)}(X_t^*)}{\mu(X_t^*)} \right) \\
&= 1
\end{aligned}$$

Or, on sait que si T est un temps fort de stationnarité pour X , alors pour tout $t > 0$, $\mathbb{P}(T > t) \geq \mathfrak{s}(t)$. Donc $T = \infty$ presque sûrement. \square

Chapitre 3

Dualité pour une marche aléatoire non-réversible sur un graphe

On considère ici le cadre de la section 2.2, que l'on généralise légèrement. On réutilisera en particulier le vocabulaire établi dans la définition 2.2.1. On a toujours un générateur \mathcal{L} sur un ensemble dénombrable G , admettant une unique mesure invariante μ , qu'on ne suppose plus nécessairement réversible. Ceci implique quelques petits changements, notamment la mesure invariante μ vérifie maintenant une version plus faible de (21) :

$$\forall p \in G, \quad \sum_{q \neq p \in G} \mu(p) \mathcal{L}_{p,q} - \mu(q) \mathcal{L}_{q,p} = 0 \quad (3.1)$$

Plus généralement, pour toute partie $Q \subset G$, on a l'égalité :

$$\sum_{p \in Q} \sum_{q \notin Q} \mu(q) \mathcal{L}_{q,p} = \sum_{p \notin Q} \sum_{q \in Q} \mu(q) \mathcal{L}_{q,p}$$

On munit cette fois G d'une structure de graphe orientée, pour lequel il existe une arête allant de p à q , $p, q \in G$, si et seulement si $\mathcal{L}_{p,q} > 0$. Pour tout $p \in G$, on définit l'ensemble $V(p) = \{q \in G : \mathcal{L}_{p,q} \neq 0\}$, que l'on suppose fini, et on pose $V(Q) = \bigcup_{p \in Q} V(p)$, $\forall Q \subset G$. Attention, on ne demande pas cette fois que les ensembles $\{q \in G : p \in V(q)\}$ soient finis pour tout $p \in G$, et on ne pose pas d'autres conditions sur la structure du graphe obtenu (notamment pour les branches infinies).

Pour toute partie compacte Q de \bar{G} , on décompose $V(Q)$ en posant :

$$\begin{aligned} \mathring{Q} &:= V(Q^c)^c = \{p \in Q : \mathcal{L}_{q,p} = 0, \forall q \in Q^c\} \\ \text{Fr}_i(Q) &:= V(Q^c) \cap Q = Q \setminus \mathring{Q} = \{p \in Q : \exists q \in Q^c, \mathcal{L}_{q,p} > 0\} \\ \text{Fr}_e(Q) &:= \text{Fr}_i(Q^c) = \{p \in Q^c : \exists q \in Q, \mathcal{L}_{q,p} > 0\} \\ \text{Fr}(Q) &:= \text{Fr}_i(Q) \cup \text{Fr}_e(Q) \end{aligned}$$

de sorte que $V(Q) = \mathring{Q} \sqcup \text{Fr}_i(Q) \sqcup \text{Fr}_e(Q)$. On appellera \mathring{Q} l'intérieur de Q , $\text{Fr}_i(Q)$ sa frontière intérieure, $\text{Fr}_e(Q)$ sa frontière extérieure et $\text{Fr}(Q)$ sa frontière.

On va ensuite élargir l'espace d'état du dual, et G^* désignera maintenant l'ensemble des parties compactes de \bar{G} ayant un nombre fini de composantes connexes, muni de la distance de Hausdorff. On rappelle que l'ensemble des parties compactes de \bar{G} est compact car \bar{G} l'est, mais en revanche G^* n'est pas fermé dès que G n'est pas fini. Étant donné une structure de graphe orienté sur G on peut se demander comment choisir les « voisins potentiels » des parties $Q \in G^*$, c'est-à-dire les $Q' \in G^*$ pour lesquels on autorise $L_{Q,Q'}^* > 0$, de manière à ce que l'équation (3.2) ait une solution quelque soit \mathcal{L} compatible avec la structure de G . Pour tout $Q \in G^*$ on définit donc un ensemble $\mathcal{V}(Q) \subset G^* \setminus \{Q\}$ qui constituera l'ensemble des voisins potentiels de Q .

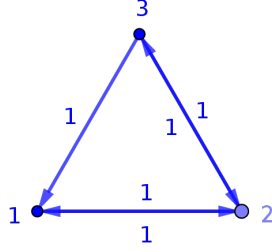


FIGURE 3.1 – Exemple de marche aléatoire non-réversible

Exemple 3.0.1. Dans la partie précédente, on a fait le choix de prendre :

$$\forall Q \in G^*, \quad \mathcal{V}(Q) = \left\{ Q \cup \{p\}, p \in \text{Fr}_e(Q) \right\} \cup \left\{ Q \in \mathcal{C}(Q \setminus \{p\}), p \in \text{Fr}_i(Q) \right\}$$

Ce choix est toujours valable dans le cas présent, et il existe un unique générateur L^* (que l'on explicitera plus tard) sur G^* , tel que $L_{Q,Q'}^* = 0$ si $Q' \notin \mathcal{V}(Q) \cup \{Q\}$, et satisfaisant la relation $L^* \Lambda = \Lambda \mathcal{L}$. Bien sûr, si \mathcal{L} est réversible ce générateur coïncide avec le générateur \mathcal{L}^* défini dans la partie précédente.

Soit Λ le noyau de transition de G^* à G défini comme en (2.12). Du fait de la possible irréversibilité de μ , le générateur \mathcal{L}^* défini par (2.11) ne satisfait plus nécessairement la relation d'entrelacement

$$\mathcal{L}^* \Lambda = \Lambda \mathcal{L} \tag{3.2}$$

comme on peut le voir dans l'exemple suivant.

Exemple 3.0.2. Considérons le processus de Markov X sur $\{1, 2, 3\}$ dont la Q -matrice est :

$$\mathcal{L} = \begin{pmatrix} -1 & 1 & 0 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix}$$

(voir figure 3.1). Une rapide résolution de l'équation matricielle $\mu \mathcal{L} = 0$ donne :

$$\mu = \begin{pmatrix} 1/2 \\ 1/3 \\ 1/6 \end{pmatrix}$$

et cette mesure invariante n'est clairement pas réversible puisque $\mu(1)\mathcal{L}_{1,3} = 0$ tandis que $\mu(3)\mathcal{L}_{3,1} = 1/6$. Si on définit un générateur \mathcal{L}^* par (2.11), on a alors notamment :

$$\mathcal{L}_{\{1\},Q}^* = \begin{cases} 5/3 & \text{si } Q' = \{1, 2\} \\ 4/3 & \text{si } Q' = \{1, 3\} \\ -3 & \text{si } Q' = \{1\} \\ 0 & \text{sinon} \end{cases}$$

ce qui implique que $[\mathcal{L}^* \Lambda]_{\{1\},3} = \mathcal{L}_{\{1\},\{1,3\}}^* \Lambda(\{1, 3\}, 3) = 1/4$. D'autre part, on a $[\Lambda \mathcal{L}]_{\{1\},3} = 0$, ce qui montre que la relation (3.2) n'est pas satisfaite.

Nous allons voir dans cette section comment généraliser la définition du générateur dual afin de préserver la relation d'entrelacement. Plus généralement, nous verrons quels sont les duaux qui la satisfont. Nous en présenterons certains avec des dynamiques légèrement différentes, et sans doute plus efficaces que celui considéré jusqu'à maintenant, dans le sens où ils devraient être absorbés plus rapidement.

On considère un générateur L^* sur G^* , tel que $L^*\Lambda = \Lambda\mathcal{L}$. Pour $Q \in G^*$, en considérant l'ensemble $\mathcal{V}(Q)$ de ses voisins potentiels comme fixé, on va étudier l'ensemble des coefficients $L_{Q,Q'}^*$, $Q' \in \mathcal{V}(Q)$ en considérant l'équation $[L^*\Lambda]_{Q,p} = [\Lambda\mathcal{L}]_{Q,p}$ pour différents $p \in G$. Dans un premier temps, remarquons que si p est tel que pour tout $q \in Q$, $\mathcal{L}_{q,p} = 0$ (c'est-à-dire $p \in V(Q)^c$), alors on a d'une part :

$$[\Lambda\mathcal{L}]_{Q,p} = \sum_{q \in Q} \frac{\mu(q)}{\mu(Q)} \mathcal{L}_{q,p} = 0$$

et d'autre part :

$$[L^*\Lambda]_{Q,p} = \sum_{\substack{Q' \in \mathcal{V}(Q) \\ p \in Q'}} L_{Q,Q'}^* \frac{\mu(p)}{\mu(Q')}$$

On en déduit que $L_{Q,Q'}^* = 0$ pour tout $Q' \in \mathcal{V}(Q)$ tel que $p \in Q'$. On peut donc supposer que $\mathcal{V}(Q) \subset \{Q' \in G^*, Q' \subset V(Q)\}$.

Pour tout $Q' \in \mathcal{V}(Q)$, soit $\alpha_{Q'}^Q \in \mathbb{R}_+$ tel que $L_{Q,Q'}^* = \frac{\mu(Q')}{\mu(Q)} \alpha_{Q'}^Q$. Lorsqu'il n'y a pas d'ambiguïté sur la dépendance en Q , on omettra l'exposant.

— Pour tout $p \in \text{Fr}_e(Q)$, on a :

$$\begin{aligned} [L^*\Lambda]_{Q,p} &= \sum_{\substack{Q' \in \mathcal{V}(Q) \\ p \in Q'}} L_{Q,Q'}^* \frac{\mu(p)}{\mu(Q')} \\ &= \frac{\mu(p)}{\mu(Q)} \sum_{\substack{Q' \in \mathcal{V}(Q) \\ p \in Q'}} \alpha_{Q'} \end{aligned}$$

et :

$$[\Lambda\mathcal{L}]_{Q,p} = \sum_{q \in Q} \frac{\mu(q)}{\mu(Q)} \mathcal{L}_{q,p}$$

donc on obtient un premier système de $|\text{Fr}_e(Q)|$ équations :

$$\forall p \in \text{Fr}_e(Q), \quad \sum_{\substack{Q' \in \mathcal{V}(Q) \\ p \in Q'}} \alpha_{Q'} = \sum_{q \in Q} \frac{\mu(q)}{\mu(p)} \mathcal{L}_{q,p} \quad (3.3)$$

— Pour $p \in Q$ on a :

$$\begin{aligned} [L^*\Lambda]_{Q,p} &= \sum_{\substack{Q' \in \mathcal{V}(Q) \\ p \in Q'}} L_{Q,Q'}^* \frac{\mu(p)}{\mu(Q')} + L_{Q,Q}^* \frac{\mu(p)}{\mu(Q)} \\ &= \sum_{\substack{Q' \in \mathcal{V}(Q) \\ p \in Q'}} L_{Q,Q'}^* \frac{\mu(p)}{\mu(Q')} - \sum_{Q' \in \mathcal{V}(Q)} L_{Q,Q'}^* \frac{\mu(p)}{\mu(Q)} \\ &= \frac{\mu(p)}{\mu(Q)} \left(\sum_{\substack{Q' \in \mathcal{V}(Q) \\ p \in Q'}} \alpha_{Q'} - \sum_{Q' \in \mathcal{V}(Q)} \frac{\alpha_{Q'} \mu(Q')}{\mu(Q)} \right) \end{aligned}$$

et :

$$\begin{aligned} [\Lambda\mathcal{L}]_{Q,p} &= \sum_{q \in Q} \frac{\mu(q)}{\mu(Q)} \mathcal{L}_{q,p} \\ &= - \sum_{q \in Q^c} \frac{\mu(q)}{\mu(Q)} \mathcal{L}_{q,p} \end{aligned}$$

cette dernière somme étant nulle si $p \in \mathring{Q}$. On obtient une deuxième série d'équations :

$$\forall p \in \text{Fr}_i(Q), \quad \sum_{\substack{Q' \in \mathcal{V}(Q) \\ p \in Q'}} \alpha_{Q'} - \sum_{Q' \in \mathcal{V}(Q)} \frac{\alpha_{Q'} \mu(Q')}{\mu(Q)} = - \sum_{q \in Q^c} \frac{\mu(q)}{\mu(p)} \mathcal{L}_{q,p} \quad (3.4)$$

$$\forall p \in \mathring{Q}, \quad \sum_{\substack{Q' \in \mathcal{V}(Q) \\ p \in Q'}} \alpha_{Q'} - \sum_{Q' \in \mathcal{V}(Q)} \frac{\alpha_{Q'} \mu(Q')}{\mu(Q)} = 0 \quad (3.5)$$

A la fin de ce chapitre nous ferons deux conjectures sur l'existence de solutions à ces équations, mais nous ne pouvons bien sûr pas toutes les décrire de manière exhaustive. Nous allons donc focaliser notre attention sur un ensemble particulier de ces générateurs. Considérant un processus ayant pour générateur L^* , quand celui-ci saute d'un état $Q \in G^*$ à un état $Q' \in \mathcal{V}$, on peut considérer qu'il « s'agrandit » si $Q' \supset Q$, qu'il « se rétrécit » si $Q' \subset Q$ et qu'il « se déplace » sinon. Le déplacement ne nous semble pas vraiment utile dans la quête de temps forts de stationnarité, c'est pourquoi nous nous restreindrons ici aux cas où il n'existe pas. Plus précisément, on pose $\mathcal{V}^+(Q) := \{Q' \in \mathcal{V}(Q); Q \subset Q'\}$ et $\mathcal{V}^-(Q) := \{Q' \in \mathcal{V}(Q); Q \supset Q'\}$, et on suppose que $\mathcal{V}(Q) = \mathcal{V}^+(Q) \cup \mathcal{V}^-(Q)$. On peut cependant imaginer d'éventuels développements futurs de cet entrelacement, pour lesquels il pourrait être intéressant d'autoriser un déplacement aléatoire du dual (pour « explorer » l'espace par exemple, et jouant alors un rôle similaire à celui du mouvement brownien dans la méthode du gradient stochastique) et où cette restriction ne serait alors pas pertinente.

Avec cette restriction, (3.3) se réécrit :

$$\forall p \in \text{Fr}_e(Q), \quad \sum_{\substack{Q' \in \mathcal{V}^+(Q) \\ p \in Q'}} \alpha_{Q'} = \sum_{q \in Q} \frac{\mu(q)}{\mu(p)} \mathcal{L}_{q,p} \quad (3.3')$$

Et en soustrayant (3.5) à (3.4) :

$$\forall p \in \text{Fr}_i(Q), p' \in \mathring{Q}, \quad \sum_{\substack{Q' \in \mathcal{V}^-(Q) \\ p \in Q'}} \alpha_{Q'} - \sum_{\substack{Q' \in \mathcal{V}^-(Q) \\ p' \in Q'}} \alpha_{Q'} = - \sum_{q \in Q^c} \frac{\mu(q)}{\mu(p)} \mathcal{L}_{q,p} \quad (3.4')$$

Les deux propositions suivantes sont des outils simples pour construire des générateurs duaux à partir de générateurs déjà existants, la première par combinaison convexe de deux générateurs, et la deuxième en remplaçant un voisin potentiel d'un éléments $Q \in G^*$ par une partition de ce voisin.

Proposition 3.0.3. *Soit L_1^* et L_2^* deux générateurs satisfaisant (3.2), et $\alpha \in [0, 1]$. Alors le générateur $L^* := \alpha L_1^* + (1 - \alpha)L_2^*$ satisfait aussi (3.2).*

Démonstration. La preuve est immédiate puisque l'on a :

$$\begin{aligned} L^* \Lambda &= (\alpha L_1^* + (1 - \alpha)L_2^*) \Lambda \\ &= \alpha L_1^* \Lambda + (1 - \alpha)L_2^* \Lambda \\ &= \alpha \Lambda \mathcal{L} + (1 - \alpha) \Lambda \mathcal{L} \\ &= \Lambda \mathcal{L} \end{aligned}$$

□

Proposition 3.0.4. Soit L^* un générateur satisfaisant (3.2), $Q_0 \in G^*$ et $Q'_0 \in \mathcal{V}(Q'_0)$. Si $\{A_1, \dots, A_k\}$, $k \in \mathbb{N}$ est une partition de Q'_0 , alors le générateur \tilde{L}^* défini par :

$$\forall Q \neq Q' \in G^*, \quad \tilde{L}^*_{Q,Q'} = \begin{cases} \frac{\mu(A_i)}{\mu(Q'_0)} L^*_{Q_0,Q'_0}, & \text{si } Q = Q_0, Q' = A_i, i = 1, \dots, k \\ 0, & \text{si } Q = Q_0, Q' = Q'_0 \\ L^*_{Q,Q'}, & \text{sinon} \end{cases}$$

est une autre solution de (3.2).

Démonstration. Pour $Q \neq Q_0 \in G^*$ et $p \in V(Q)$, il est clair que l'égalité $[\tilde{L}^* \Lambda]_{Q,p} = [\Lambda \mathcal{L}]_{Q,p}$ est satisfaite, de même que pour $Q = Q_0$ et $p \notin Q_0$. Il suffit donc de la vérifier pour $Q = Q_0$ et $p \in Q_0$:

$$\begin{aligned} [\tilde{L}^* \Lambda]_{Q,p} &= \sum_{Q' \ni p} \tilde{L}^*_{Q,Q'} \frac{\mu(p)}{\mu(Q')} \\ &= \sum_{\substack{Q' \neq Q'_0 \in \mathcal{V}(Q) \\ p \in Q'}} L^*_{Q,Q'} \frac{\mu(p)}{\mu(Q')} + \sum_{i=1}^k \left(\frac{\delta_p(A_i) \mu(A_i)}{\mu(Q'_0)} L^*_{Q_0,Q'_0} \right) \frac{\mu(p)}{\mu(A_i)} \\ &= \sum_{\substack{Q' \neq Q'_0 \in \mathcal{V}(Q) \\ p \in Q'}} L^*_{Q,Q'} \frac{\mu(p)}{\mu(Q')} + L^*_{Q_0,Q'_0} \frac{\mu(p)}{\mu(Q'_0)} \\ &= \sum_{\substack{Q' \in \mathcal{V}(Q) \\ p \in Q'}} L^*_{Q,Q'} \frac{\mu(p)}{\mu(Q')} \\ &= [L^* \Lambda]_{Q,p} \\ &= [\Lambda \mathcal{L}]_{Q,p} \end{aligned}$$

□

Cette proposition a une interprétation probabiliste assez simple. Si X est un processus de générateur \mathcal{L} et X^* un dual de stationnarité forte pour X , de générateur L^* (satisfaisant (3.2)), alors lorsque X^* saute de Q_0 à Q'_0 , on peut décider de le faire sauter de nouveau instantanément à l'ensemble A_i , $i = 1, \dots, k$ qui contient X_T , et le nouveau processus \tilde{X} obtenu a pour générateur \tilde{L}^* et est toujours un dual pour X . En effet si on pose $T = \inf\{t \geq 0 : X_{T-}^* = Q, X_T^* = Q'\}$, alors X^* et \tilde{X}^* sont identiques jusqu'au temps T et ce dernier est donc entrelacé avec X . Au temps T on a, pour $i = 1, \dots, k$ et $p \in A_i$:

$$\begin{aligned} \mathbb{P}(X_T = p \mid \tilde{X}_T^* = A_i) &= \frac{\mathbb{P}(\tilde{X}_T^* = A_i \mid X_T = p) \mathbb{P}(X_T = p)}{\mathbb{P}(\tilde{X}_T^* = A_i)} \\ &= \frac{1 \times \frac{\mu(p)}{\mu(Q'_0)}}{\mathbb{P}(X_T \in A_i)} \\ &= \frac{\mu(p)}{\frac{\mu(Q'_0)}{\mu(A_i)}} = \frac{\mu(p)}{\mu(A_i)} \\ &= \frac{\mu(p)}{\mu(Q'_0)} \end{aligned}$$

Le processus $(\tilde{X}_{T+t}^*)_{t \geq 0}$ évolue ensuite selon le générateur L^* et sa loi initiale est entrelacée avec celle de $(X_{T+t})_{t \geq 0}$. On peut donc coupler ces deux processus de manière à conserver la relation d'entrelacement jusqu'au prochain temps d'atteinte de Q_0 et on itère alors le raisonnement précédent.

Sur ce même principe, on peut en fait construire plein d'autres duaux de stationnarité forte de la manière suivante : soient L^* un générateur satisfaisant (3.2), X^* un processus de Markov fini associé à L^* , T un temps d'arrêt pour X^* et $(A_i, p_i)_{i=1, \dots, N}$, $N \in \mathbb{N}$ une famille de variable aléatoires X_T^* -mesurables (N pouvant être aléatoire également) satisfaisant presque sûrement :

$$\begin{aligned} \forall i = 1, \dots, N, \quad & A_i \subset X_T^*, \quad p_i \in [0, 1] \\ \forall x \in X_T^*, \quad & \sum_{i=1}^N p_i \delta_x(A_i) = 1 \end{aligned}$$

Alors le processus $Y = (Y_t)_{t \geq 0}$ défini par :

- $Y_{[0, T[} = X_{[0, T[}^*$
- Connaissant X_T et X_T^* , on pose $Y_T = A_i$ avec probabilité $p_i \delta_{X_T}(A_i)$
- $(Y_{T+t})_{t \geq 0}$ est un processus minimal de générateur L^*

est un dual de stationnarité forte pour X .

Nous allons maintenant donner des exemples de $\mathcal{V}(Q)$ et les L^* qui vont avec.

Exemple 3.0.5. On prend pour tout $Q \in G^*$, $\mathcal{V}_1^+(Q) = \{Q \cup \{p\}, p \in \text{Fr}_e(Q)\}$ et $\mathcal{V}_1^-(Q) = \{Q \setminus \{p\}, p \in \text{Fr}_i(Q)\}$. C'est en quelque sorte le cas le plus simple, et l'on a de manière immédiate :

$$\begin{aligned} \forall p \in \text{Fr}_e(Q), \quad & \alpha_{Q \cup \{p\}} = \sum_{q \in Q} \frac{\mu(q)}{\mu(p)} \mathcal{L}_{q,p} \\ \forall p \in \text{Fr}_i(Q), \quad & \alpha_{Q \setminus \{p\}} = \sum_{q \in Q^c} \frac{\mu(q)}{\mu(p)} \mathcal{L}_{q,p} \end{aligned}$$

On comprend ici l'intérêt d'avoir imposé aux éléments de G^* d'avoir un nombre fini de composantes connexes. En effet, considérons par exemple le cas de la file d'attente $M/M/\infty$ sur \mathbb{N} , de taux de service et d'arrivée 1 (c'est-à-dire, $\mathcal{L}_{x,x+1} = 1$ et $\mathcal{L}_{x,x-1} = x$ pour tout $x \in \mathbb{N}$, voir Kendall [48]). Ce processus admet une unique mesure invariante réversible (une loi de Poisson de paramètre 1), et l'on peut donc définir les coefficients α comme dans l'exemple précédent, en tenant compte de la réversibilité de la mesure :

$$\forall p \in \text{Fr}_e(Q), \quad \alpha_{Q \cup \{p\}} = \sum_{q \in Q} \mathcal{L}_{p,q}$$

et similairement pour $\alpha_{Q \setminus \{p\}}$. Si on les définit ainsi pour l'ensemble $Q = 2\mathbb{N}$ par exemple, alors pour tout p impair on obtient :

$$L_{Q, Q \cup \{p\}}^* \geq \alpha_{Q, Q \cup \{p\}} = p + 1$$

On a donc $L_Q^* = \infty$, c'est-à-dire que l'état Q n'est pas stable. De manière générale si l'on tente de prolonger la définition ci-dessus à un Q tel que $\text{Fr}(Q)$ soit infini, alors on n'a pas la garantie d'avoir $L_Q^* < \infty$. Si G est infini, on ne peut donc définir le dual que jusqu'au temps où le nombre de ses composantes connexes explose. Si on applique la proposition 3.0.4 à l'exemple de générateur que l'on vient de présenter, en considérant pour chaque Q' la partition formée par l'ensemble de ses composantes connexes, on obtient le dual défini dans la partie 2.2 qui, lui, présente l'avantage de rester tout le temps connexe, et donc de pouvoir être défini tout le temps (ou au moins jusqu'à son temps de double explosion, cf. définition 1.2.3).

Exemple 3.0.6. On considère un exemple qui sera utilisé dans la section suivante. Dans l'exemple précédent, le dual ne peut gagner ou perdre qu'un point à la fois. On va maintenant le forcer à « gagner autant de points de $\text{Fr}_e(Q)$ que possible », ou à « perdre autant de points de $\text{Fr}_i(Q)$ que possible », d'un coup.

Soit $Q \in G^*$, $m_e = |\text{Fr}_e(Q)|$ et $m_i = |\text{Fr}_i(Q)|$. On numérote $p_e^1, \dots, p_e^{m_e}$ les éléments de $\text{Fr}_e(Q)$ et $p_i^1, \dots, p_i^{m_i}$ ceux de $\text{Fr}_i(Q)$ (ici aussi, on omettra généralement les exposants quand il n'y a pas d'ambiguïté) de manière à ce que :

$$\begin{aligned} \forall i < m_e, \quad \sum_{q \in Q} \frac{\mu(q)}{\mu(p_e^i)} \mathcal{L}_{q, p_e^i} &\geq \sum_{q \in Q} \frac{\mu(q)}{\mu(p_e^{i+1})} \mathcal{L}_{q, p_e^{i+1}} \\ \forall i < m_i, \quad \sum_{q \notin Q} \frac{\mu(q)}{\mu(p_i^i)} \mathcal{L}_{q, p_i^i} &\geq \sum_{q \in Q} \frac{\mu(q)}{\mu(p_i^{i+1})} \mathcal{L}_{q, p_i^{i+1}} \end{aligned}$$

On définit ensuite :

$$\begin{aligned} \forall i \leq m_e, \quad \partial_i^+ Q &= Q \cup \bigcup_{k=1}^i \{p_e^k\} \\ \forall i \leq m_i, \quad \partial_i^- Q &= Q \setminus \bigcup_{k=1}^i \{p_i^k\} \end{aligned}$$

Et on pose $\mathcal{V}_2^+(Q) = \{\partial_i^+ Q, i = 1, \dots, n\}$ et $\mathcal{V}_2^-(Q) = \{\partial_i^- Q, i = 1, \dots, m_i\}$. En prenant $p = p_e^{m_e}$ dans (3.3'), on obtient :

$$\alpha_{\partial_{m_e}^+ Q} = \sum_{q \in Q} \frac{\mu(q)}{\mu(p_e^{m_e})} \mathcal{L}_{q, p_e^{m_e}}$$

puis successivement chacun des p_e^i par ordre décroissant :

$$\forall i < m_e, \quad \alpha_{\partial_i^+ Q} = \sum_{q \in Q} \frac{\mu(q)}{\mu(p_e^i)} \mathcal{L}_{q, p_e^i} - \sum_{q \in Q} \frac{\mu(q)}{\mu(p_e^{i+1})} \mathcal{L}_{q, p_e^{i+1}}$$

De la même manière, en prenant successivement $p = p_i^i$ par ordre décroissant et $p' \in \mathring{Q}$ quelconque dans (3.4'), on obtient :

$$\begin{aligned} \alpha_{\partial_{m_i}^- Q} &= \sum_{q \notin Q} \frac{\mu(q)}{\mu(p_i^{m_i})} \mathcal{L}_{q, p_i^{m_i}} \\ \forall i < m_i, \quad \alpha_{\partial_i^- Q} &= \sum_{q \notin Q} \frac{\mu(q)}{\mu(p_i^i)} \mathcal{L}_{q, p_i^i} - \sum_{q \notin Q} \frac{\mu(q)}{\mu(p_i^{i+1})} \mathcal{L}_{q, p_i^{i+1}} \end{aligned}$$

Vérifions que l'équation (3.5) est également satisfaite. Pour tout $Q' \in \mathcal{V}(Q)$, on a $\mathring{Q} \subset Q'$, donc pour tout $p \in \mathring{Q}$:

$$\begin{aligned} \sum_{\substack{Q' \in \mathcal{V}(Q) \\ p \in Q'}} \alpha_{Q'} - \sum_{Q' \in \mathcal{V}(Q)} \frac{\alpha_{Q'} \mu(Q')}{\mu(Q)} &= \sum_{Q' \in \mathcal{V}(Q)} \alpha_{Q'} \left(1 - \frac{\mu(Q')}{\mu(Q)}\right) \\ &= \sum_{i=1}^{m_i} \alpha_{\partial_i^- Q} \frac{\mu(Q \setminus \partial_i^- Q)}{\mu(Q)} - \sum_{i=1}^{m_e} \alpha_{\partial_i^+ Q} \frac{\mu(\partial_i^+ Q \setminus Q)}{\mu(Q)} \\ &= \sum_{i=1}^{m_i} \sum_{q \notin Q} \frac{\mu(q)}{\mu(p_i^i)} \mathcal{L}_{q, p_i^i} \frac{\mu(p_i^i)}{\mu(Q)} - \sum_{i=1}^{m_e} \sum_{q \in Q} \frac{\mu(q)}{\mu(p_e^i)} \mathcal{L}_{q, p_e^i} \frac{\mu(p_e^i)}{\mu(Q)} \\ &= \sum_{p \in Q} \sum_{q \notin Q} \frac{\mu(q)}{\mu(Q)} \mathcal{L}_{q, p} - \sum_{p \notin Q} \sum_{q \in Q} \frac{\mu(q)}{\mu(Q)} \mathcal{L}_{q, p} \\ &= 0 \end{aligned}$$

Comme dans l'exemple précédent, un tel dual n'est défini que jusqu'au temps d'explosion du nombre de ses composantes connexes.

Il pourrait être intéressant de comparer la vitesse d'accroissement des duals présentés dans ces deux exemples : en étudiant rapidement les taux de sauts de chacun, on constate assez facilement (nous n'en donnerons pas la preuve ici) que le deuxième s'agrandit plus vite car il peut ajouter plusieurs points en un seul saut, mais il se rétrécit aussi plus vite par symétrie. Lorsque le dual tend à s'agrandir et à remplir tout l'espace, il semble donc naturel de supposer que ceci se produit plus rapidement dans le deuxième cas. Pour accélérer encore cette dynamique, on peut aussi choisir de construire un troisième dual, avec les ensembles $\mathcal{V}_1^-(Q)$ du premier et $\mathcal{V}_2^+(Q)$ du second, et les coefficients correspondants dans les deux exemples pour le générateur. Ce dual ajoute plusieurs points d'un coup mais n'en perd qu'un à la fois, et combine donc les avantages de chacun des deux duals présentés. Remarquons que cette troisième alternative est l'équivalent en temps continu et à espace d'état infini de l'algorithme glouton de Diaconis et Fill [23] (section 3.4), à l'exception de la question de la loi initiale, qui a été éludée ici mais qui peut être traitée exactement de la même manière que dans le cas discret. Cette option semble donc plus avantageuse en terme de rapidité de convergence. Cependant, le dual proposé dans l'exemple 3.0.6 présente l'avantage, que l'on développera dans la section suivante, d'approximer une diffusion multidimensionnelle sur un graphe continu dont toutes les coordonnées sont dirigées par un même mouvement brownien.

Il serait également intéressant d'adapter les calculs du chapitre précédent au cas non-réversible, en utilisant les outils présentés ici.

Nous présentons maintenant quelques résultats non encore démontrés, mais qui ont vocation à donner une intuition et à aiguiller les calculs lors de la recherche de processus duals.

Conjecture 3.0.7. *L'équation (3.2) admet (au moins) une solution L^* quelque soit \mathcal{L} , si et seulement si pour tout $Q \in G^*$, $\mathcal{V}(Q) \cup \{Q\}$ sépare les points de $\text{Fr}(Q)$ et l'ensemble \mathring{Q} au sens de Kolmogorov, c'est-à-dire pour tout $p \in V(Q)$ et tout $p' \in \text{Fr}(Q)$, il existe $U \in \mathcal{V}(Q) \cup \{Q\}$ tels que $p \in U$ et $p' \notin U$ ou inversement (axiome de séparation T_0 , voir par exemple Willard [91]).*

NB : on ne demande pas que les points de \mathring{Q} entre eux soit séparés par $\mathcal{V}(Q)$.

Remarque 3.0.8. *L'hypothèse de séparation s'énonce de manière équivalente par : la topologie engendrée par les éléments de $\mathcal{V}(Q)$ contient \mathring{Q} et les singletons $\{p\}$, $p \in \text{Fr}(Q)$.*

Démonstration partielle de la conjecture 3.0.7. Ce résultat doit pouvoir se prouver en mettant en relation les combinaisons linéaires des différentes équations et les opérations d'union/intersection d'ensembles de $\mathcal{V}(Q)$.

Une implication se montre facilement. En effet, si l'on considère deux points $p, q \in \text{Fr}_e(Q)$ et que l'on suppose qu'ils ne sont pas séparés, alors on a :

$$\sum_{\substack{Q' \in \mathcal{V}(Q) \\ p \in Q'}} \alpha_{Q'} = \sum_{\substack{Q' \in \mathcal{V}(Q) \\ q \in Q'}} \alpha_{Q'}$$

et en considérant l'équation (3.3) successivement avec p et p' on obtient :

$$\sum_{q \in Q} \frac{\mu(q)}{\mu(p)} \mathcal{L}_{q,p} = \sum_{q \in Q} \frac{\mu(q)}{\mu(p')} \mathcal{L}_{q,p'}$$

ce qui n'est pas vrai pour tout \mathcal{L} , et l'équation (3.2) n'a donc pas forcément de solution. On traite similairement le cas $p, p' \in \text{Fr}_i(Q)$ en remplaçant (3.3) par (3.4), et le cas

$p \in \text{Fr}_i(Q)$, $p' \in \mathring{Q}$ en considérant l'équation (3.4') avec p et p' . Enfin, deux points $p \in Q$ et $p' \in \text{Fr}_e(Q)$ sont trivialement séparés par Q ($p \in Q$ et $p' \notin Q$), ce qui achève de montrer que si deux points $p \in V(Q)$ et $p' \in \text{Fr}(Q)$ ne sont pas séparés alors il existe des \mathcal{L} pour lesquels (3.2) n'admet pas de solution. \square

De la conjecture 3.0.7 on déduit aisément deux corollaires.

Corollaire 3.0.9. *Pour que (3.2) ait une solution quelque soit \mathcal{L} , il faut que $|\mathcal{V}(Q)| \geq |\text{Fr}(Q)|$ pour tout $Q \in G^*$. De plus, il est possible de choisir $\mathcal{V}(Q)$ avec $|\mathcal{V}(Q)| = |\text{Fr}(Q)|$ pour tout $Q \in G^*$ et tel (3.2) ait une solution. Dans ce cas, la solution est unique.*

Le second corollaire est utile pour comprendre la dynamique d'un processus dual. Il découle immédiatement de la conjecture et on en donne également ici une preuve directe assez simple.

Corollaire 3.0.10. *S'il existe une solution à (3.2) quelque soit \mathcal{L} , alors pour tout $Q \in G^*$ et tout $p \in \text{Fr}_e(Q)$, $q \in \text{Fr}_i(Q)$, il existe $Q', Q'' \in \mathcal{V}(Q)$ tels que $p \in Q'$ et $q \notin Q''$.*

Démonstration. Soit $Q \in G^*$ et $p \in \text{Fr}_e(Q)$. Par définition de la frontière extérieure, il existe $q \in Q$ tel que $\mathcal{L}_{q,p} > 0$, donc $\sum_{q \in Q} \frac{\mu(q)}{\mu(p)} \mathcal{L}_{p,q} > 0$ et d'après (3.3) il existe $Q' \in \mathcal{V}(Q)$ tel que $\alpha_{Q'} > 0$.

Soit maintenant $q \in \text{Fr}_i(Q)$. Si on suppose par l'absurde que $q \in Q'$ pour tout $Q' \in \mathcal{V}(Q)$, alors on retranchant (3.5) à (3.4), on obtient, pour un $p \in \mathring{Q}$ quelconque :

$$\sum_{Q' \in \mathcal{V}(Q)} \alpha_{Q'} - \sum_{\substack{Q' \in \mathcal{V}(Q) \\ p \in Q'}} \alpha_{Q'} = - \sum_{p \in Q^c} \frac{\mu(p)}{\mu(q)} \mathcal{L}_{p,q} < 0$$

c'est à dire :

$$\sum_{\substack{Q' \in \mathcal{V}(Q) \\ p \notin Q'}} \alpha_{Q'} < 0$$

ce qui est impossible. \square

Pour les deux exemples de duaux présentés dans ce chapitre, ainsi que pour d'autres exemples envisagés lors de ces travaux, il est apparu au fil des calculs qu'en résolvant les équations (3.3') et (3.4') (ou de manière équivalente les équations (3.3) et (3.4)), la solution trouvée satisfaisait automatiquement l'équation (3.5). On en a fait la conjecture suivante.

Conjecture 3.0.11. *Si les $\mathcal{V}(Q)$, $Q \in G^*$ sont tels qu'il existe une solution à (3.2) quelque soit \mathcal{L} , et si un générateur L^* satisfait (3.3') et (3.4'), alors il satisfait aussi (3.5), et par conséquent est solution de (3.2).*

Deuxième partie

Dualité pour une diffusion sur un graphe quantique

Chapitre 4

Diffusion sur un graphe

4.1 Cadre théorique

On considère un graphe fini connexe et non-orienté $G := (V, E)$. On le suppose simple et sans boucle, de sorte que l'ensemble des arêtes E peut être vu comme un ensemble de paires de l'ensemble des sommets V . Ces hypothèses ne sont en fait pas restrictives, comme on le verra dans la remarque 4.1.2. Soit \mathbb{G} le graphe continu et compact où chaque arête $e \in E$ est remplacée par un ensemble R_e homéomorphe au segment réel $[0, 1]$. Naturellement, les deux extrémités de R_e sont les deux éléments de e . Pour tout $e \in E$, on fixe un homéomorphisme $x_e : R_e \rightarrow [0, l_e]$, où $l_e > 0$ sera vu comme la **longueur** de R_e . En particulier, on a $e = \{e^-, e^+\}$ où $e^- := x_e^{-1}(0)$ et $e^+ := x_e^{-1}(l_e)$, et donc le choix de l'application coordonnées $(x_e)_{e \in E}$ fournit une orientation arbitraire de E , dont les arêtes e peuvent être vues comme les couples (e^-, e^+) . Certaines notions dont nous auront besoin reposent sur les **arêtes orientées** et nous noterons leur ensemble \vec{E} :

$$\vec{E} := \{(e^-, e^+), (e^+, e^-) : e \in E\} \subset V \times V$$

On utilisera comme indicateur de l'orientation d'une arête la notation $i_{\vec{e}}$, définie par :

$$\forall \vec{e} \in \vec{E}, \quad i_{\vec{e}} = \begin{cases} 1 & \text{si } \vec{e} = (e^-, e^+), e \in E \\ -1 & \text{si } \vec{e} = (e^+, e^-), e \in E \end{cases}$$

Pour chaque $e \in E$, l'application x_e munit R_e d'une structure différentielle : on note $\mathcal{C}^2(R_e)$ l'ensemble des fonctions $f : R_e \rightarrow \mathbb{R}$, telles que $f \circ x_e^{-1}$ est deux fois continûment différentiable sur $[0, l_e]$. L'**opérateur différentiel de premier ordre** ∂_e est défini sur $\mathcal{C}^2(R_e)$ par :

$$\forall f \in \mathcal{C}^2(R_e), \forall x \in R_e, \quad \partial_e f(x) := (f \circ x_e^{-1})'(x_e(x))$$

où le prime dans le membre de droite correspond à la dérivation usuelle sur $[0, l_e]$. La notation prime sera aussi utilisée sur V pour désigner la différentiation dans la direction d'une arête adjacente. Plus précisément, pour $x \in V$ soit $E_x := \{e \in E : x \in e\}$ l'ensemble des arêtes adjacentes à x , et posons :

$$\forall x \in V, \forall e \in E_x, \forall f \in \mathcal{C}^2(R_e), \quad f'(x, e) = \begin{cases} \partial_e f(e^-) & , \text{ si } x = e^- \\ -\partial_e f(e^+) & , \text{ si } x = e^+ \end{cases} \quad (4.1)$$

De la même manière, on utilisera la notation $\vec{E}_x = \{\vec{e} \in \vec{E}; \vec{e} = (x, y), y \in V\}$, pour désigner l'ensemble des arêtes orientées partant de $x \in V$.

La famille $(x_e)_{e \in E}$ munit également le graphe \mathbb{G} d'une structure d'espace métrique. On définit une distance d_e sur chaque arête R_e , $e \in E$, par :

$$\forall x, y \in R_e, \quad d_e(x, y) = |x_e(x) - x_e(y)|$$

et une distance globale d sur \mathbb{G} comme la longueur du plus court chemin entre deux points $x, y \in \mathbb{G}$, c'est-à-dire le minimum de $\sum_{i=1}^n d_{e_i}(x_{i-1}, x_i)$ où $n \in \mathbb{N}$, les $e_i \in E$ et les $x_i \in \mathbb{G}$ sont tels que $x_0 = x$, $x_n = y$ et $x_{i-1}, x_i \in R_{e_i}$ pour tout $i = 1, \dots, n$.

Pour introduire un opérateur de diffusion \mathcal{L} sur l'espace d'état \mathbb{G} , nous avons besoin de quelques définitions supplémentaires. Pour chaque $e \in E$, on se donne un **coefficient de drift** $b_e \in \mathcal{C}(R_e)$. Un **générateur de diffusion** \mathcal{L}_e est défini localement sur R_e via :

$$\forall f \in \mathcal{C}^2(R_e), \quad \mathcal{L}_e[f] := \partial_e^2 f + b_e \partial_e f \quad (4.2)$$

Remarque 4.1.1. *Il pourrait être tentant de considérer des opérateurs plus généraux de la forme :*

$$\forall f \in \mathcal{C}^2(R_e), \quad \mathcal{L}_e[f] := a_e \partial_e^2 f + b_e \partial_e f \quad (4.3)$$

où le **coefficient de diffusion** $a_e \in \mathcal{C}(R_e)$ est strictement positif. En fait, ce cadre ne serait pas vraiment plus général.

Pour s'en convaincre, considérons le \mathcal{C}^1 -difféomorphisme :

$$\varphi_e : [0, l_e] \ni x \mapsto \int_0^x \frac{1}{\sqrt{a_e(x_e^{-1}(y))}} dy \in [0, \tilde{l}_e]$$

où $\tilde{l}_e = \int_0^{l_e} 1/\sqrt{a_e(x_e^{-1}(y))} dy$. L'application $\tilde{x}_e := \varphi_e \circ x_e$ est alors un homéomorphisme entre R_e et $[0, \tilde{l}_e]$, et un élément de $\mathcal{C}^2(R_e)$. Considérons l'opérateur différentiel $\tilde{\partial}_e$ construit en conséquence. Alors on a :

$$\forall f \in \mathcal{C}^2(R_e), \quad \mathcal{L}_e[f] = \tilde{\partial}_e^2 f + \tilde{b}_e \tilde{\partial}_e f \quad (4.4)$$

où $\tilde{b}_e := \mathcal{L}_e[\tilde{x}_e]$. En effet, si on considère le **carré du champ** $\Gamma_{\mathcal{L}_e}$ associé à \mathcal{L}_e :

$$\begin{aligned} \forall f \in \mathcal{C}^2(R_e), \quad \Gamma_{\mathcal{L}_e}[f, f] &:= \frac{1}{2}(\mathcal{L}_e[f^2] - 2f\mathcal{L}_e[f]) \\ &= a_e(\partial_e[f])^2 \end{aligned}$$

Prenant en compte que pour toute fonction régulière $g : [0, \tilde{l}_e] \rightarrow \mathbb{R}$, on a la **formule de changement de variable** (voir par exemple la section 1.11.1 de Bakry, Gentil et Ledoux [5])

$$\mathcal{L}_e[g \circ \tilde{x}_e] = g'(\tilde{x}_e)\mathcal{L}_e[\tilde{x}_e] + g''(\tilde{x}_e)\Gamma_{\mathcal{L}_e}[\tilde{x}_e, \tilde{x}_e]$$

on obtient (4.4), puisque φ_e a été défini de sorte à ce que :

$$\Gamma_{\mathcal{L}_e}[\tilde{x}_e, \tilde{x}_e] \equiv 1$$

Ces considérations nous ramènent à la situation (4.2), et montre également pourquoi il est important d'autoriser la dépendance en e de la longueur l_e dans notre contexte initial : alors que dans la formulation (4.3) il est possible d'imposer que $l_e = 1$ (pour tout $e \in E$), une fois la transformation appliquée à (4.4), on peut arriver à des longueurs \tilde{l}_e qui ne sont plus constantes. Mais comme il apparaîtra clairement dans les calculs qui suivent, il est plus simple de travailler avec des longueurs variables qu'avec des coefficients de diffusion variables.

Pour assembler tous les \mathcal{L}_e en un unique opérateur \mathcal{L} , on doit décrire comment le processus sous-jacent se comporte sur V . Ceci est fait en imposant des conditions de Neumann sur V aux fonctions sur lesquelles \mathcal{L} va agir.

Donc pour chaque $x \in V$, on se donne une mesure de probabilité $p(x, \cdot)$ sur E telle que pour tout $e \in E$, on ait $p(x, e) > 0$ si et seulement si $e \in E_x$, c'est-à-dire $x \in e$. On note $p := (p(x, e))_{x \in V, e \in E}$, qui est un noyau markovien de V vers E respectant la structure de G (i.e. une matrice d'incidence à poids). Il sera pratique d'interpréter p comme un noyau de Markov sur V (une matrice d'adjacence à poids), appelé **noyau de transition**, donné par :

$$\forall x, y \in V, \quad p(x, y) := p(x, \{x, y\})$$

Soit \mathcal{F} l'ensemble des fonctions réelles f sur \mathbb{G} dont la restriction à R_e appartient à $\mathcal{C}^2(R_e)$ (notons que ceci implique que f est globalement continue sur \mathbb{G}) satisfaisant les **conditions de Neumann** :

$$\forall x \in V, \quad \sum_{e \in E_x} p(x, e) f'(x, e) = 0 \quad (4.5)$$

Un générateur de diffusion \mathcal{L} est alors défini globalement sur \mathcal{F} par :

$$\forall f \in \mathcal{F}, \forall x \in \mathbb{G}, \quad \mathcal{L}[f](x) := \begin{cases} \mathcal{L}_e[f|_e](x) & , \text{ si } x \in R_e \setminus V, e \in E \\ \sum_{e \in E_x} p(x, e) \mathcal{L}_e[f|_e](x) & , \text{ si } x \in V \end{cases} \quad (4.6)$$

où l'on a adopté la convention suivante, appliquée dans toute la suite : pour $e \in E$, $f|_e$ représente la restriction à R_e d'une fonction $f \in \mathcal{F}$ globalement définie sur \mathbb{G} , tandis que $(f_e)_{e \in E}$ représente un élément générique de $\prod_{e \in E} \mathcal{C}^2(R_e)$, en général sans $f \in \mathcal{F}$ correspondant (comme par exemple pour $(b_e)_{e \in E}$ utilisé dans (4.2)). La définition ci-dessus de $\mathcal{L}[f]$ sur V est assez arbitraire, puisque ces valeurs ne jouent pas de rôle dans ce qui suit. En fait, il n'est pas possible de choisir $\mathcal{L}[f]$ sur V de sorte que $\mathcal{L}[f]$ soit continue sur \mathbb{G} pour tout $f \in \mathcal{F}$ (sauf dans le cas où G a au plus une arête), autrement dit \mathcal{L} n'est pas un endomorphisme de \mathcal{F} .

Remarque 4.1.2. *Les conditions de Neumann sur les sommets sont en fait une généralisation de la dérivabilité d'une fonction en un point. En effet, considérons une arête $e \in E$ et un point $x \in R_e \setminus \{e^+, e^-\}$. On définit un nouveau graphe $\tilde{G} = (\tilde{V}, \tilde{E})$, par $\tilde{V} = V \cup \{x\}$ et $\tilde{E} = E \setminus \{e\} \cup \{e_1, e_2\}$, avec $e_1 = \{e^-, x\}$, $e_2 = \{e^+, x\}$. Alors le graphe continu \mathbb{G} peut être défini comme la réunion des $R_{e'}$, $e' \neq e \in E$ et de $R_{e_1} = x_e^{-1}([0, x_e(x)])$, $R_{e_2} = x_e^{-1}([x_e(x), l_e])$, ces deux dernières arêtes étant homéomorphes à $[0, x_e(x)]$ et $[0, l_e - x_e(x)]$ respectivement. De plus, pour tout $f \in \mathcal{F}$, la restriction de f à R_{e_1} et R_{e_2} est bien \mathcal{C}^2 et f satisfait les conditions de Neumann pour chaque sommet, en rajoutant aux coefficients précédents les coefficients :*

$$\begin{aligned} p(x, e_1) &= p(x, e_2) = 1/2 \\ p(e^-, e_1) &= p(e^-, e), \quad p(e^+, e_2) = p(e^+, e) \end{aligned}$$

Si on remplace \mathcal{L}_e par $\mathcal{L}_{e_1} := \mathcal{L}_{e|_{e_1}}$ et $\mathcal{L}_{e_2} := \mathcal{L}_{e|_{e_2}}$, alors la solution du problème de martingale associée au nouveau générateur $\tilde{\mathcal{L}}$, défini comme précédemment mais avec les nouveaux \mathcal{L}_e , est égale en loi à celle associée à \mathcal{L} .

Ceci justifie également le choix de ne considérer que des graphes simples et sans boucles : si \mathbb{G}' est un graphe continu défini à partir d'un graphe G' quelconque, alors on peut toujours se ramener à un graphe simple et sans boucle en rajoutant des points sur les arêtes comme on vient de faire.

Remarque 4.1.3. *Heuristiquement, une diffusion générée par \mathcal{L} évolue selon le générateur \mathcal{L}_e sur $R_e \setminus \{e^-, e^+\}$, pour $e \in E$, et quand elle atteint un sommet $x \in V$, elle choisit d'aller dans la direction $e \in E$ avec la probabilité $p(x, e)$. Il n'est pas facile de justifier rigoureusement cette affirmation, puisque quand la diffusion tape un sommet $x \in V$, elle va ensuite le taper infiniment souvent sur tout intervalle de temps positif. Il est possible de surmonter cette difficulté par une approximation utilisant la théorie des excursions (voir par exemple le chapitre 12 de Revuz et Yor [75]) et de « choisir les excursions suivantes » suivant une mesure de Poisson (dont la mesure d'intensité sur les trajectoires est infinie, reflétant qu'il y a beaucoup de petites excursions).*

Mais toutes ces considérations seront cachées dans la formulation suivante du problème de martingale.

Freidlin et Wentzell ont prouvé dans [40] (voir aussi Freidlin et Sheu [39]) que le **problème de martingale** associé à \mathcal{L} est **bien posé** : pour toute distribution initiale m_0 sur \mathbb{G} , il existe un unique (en loi) processus stochastique $X := (X_t)_{t \geq 0}$ sur \mathbb{G} dont la loi initiale $\mathcal{L}(X_0)$ est égale à m_0 et tel que pour tout $f \in \mathcal{F}$, le processus $\mathcal{M}^f := (\mathcal{M}_t^f)_{t \geq 0}$ défini par :

$$\forall t \geq 0, \quad \mathcal{M}_t^f := f(X_t) - f(X_0) - \int_0^t \mathcal{L}[f](X_s) ds$$

est une martingale. Le processus X solution de ce problème de martingale est une **diffusion**, c'est-à-dire un processus de Markov à trajectoires continues.

En général \mathcal{L} n'est pas **réversible**, *i.e.* il n'existe pas de mesure de probabilité μ sur \mathbb{G} telle que \mathcal{L} soit symétrique dans $\mathbb{L}^2(\mu)$. Toutefois la positivité des mesures de probabilité $p(x, \cdot)$ sur E_x , pour $x \in V$, ainsi que la connexité de G , assure qu'il existe une unique **mesure de probabilité invariante** μ pour \mathcal{L} , c'est-à-dire telle que :

$$\forall f \in \mathcal{F}, \quad \mu[\mathcal{L}[f]] = 0$$

La signification probabiliste de μ est que quand X a pour loi initiale μ , alors pour tout $t \geq 0$, la loi de X_t est toujours μ .

4.2 Réversibilité et mesure invariante

Pour tout $e \in E$, soit λ_e l'image par x_e^{-1} de la mesure de Lebesgue sur $[0, l_e]$. On définit la fonction β_e sur R_e^2 par :

$$\forall x, y \in R_e, \quad \beta_e(x, y) := \begin{cases} \int_{[x, y]} b_e(t) \lambda_e(dt) & , \text{ si } x_e(x) \leq x_e(y) \\ - \int_{[x, y]} b_e(t) \lambda_e(dt) & , \text{ sinon} \end{cases}$$

où $[x, y] := x_e^{-1}([\min(x_e(x), x_e(y)), \max(x_e(x), x_e(y))])$ (une telle notation sera utilisée dans tout le papier, sans autre préavis).

On considère la Q-matrice $\kappa := (\kappa(x, y))_{x, y \in V}$ définie par :

$$\forall x \neq y \in V, \quad \kappa(x, y) := \begin{cases} \frac{p(x, y)}{\int_{[x, y]} e^{\beta_e(z, x)} \lambda_e(dz)} & , \text{ si } \{x, y\} = e \in E \\ 0 & , \text{ sinon} \end{cases}$$

(les valeurs sur la diagonale sont imposées par le fait que les sommes sur chaque ligne de κ doivent s'annuler), où p est le noyau de transition. C'est le générateur d'un processus

de Markov fini ergodique sur V , donc il a une unique mesure de probabilité invariante α , autrement dit une mesure sur V de masse totale 1 et telle que :

$$\forall x \in V, \quad \sum_{y \in V} \alpha(y) \kappa(y, x) = 0 \quad (4.7)$$

Remarquons que puisque $\kappa(x, x) = -\sum_{y \neq x} \kappa(x, y)$, cette équation peut aussi s'écrire sous la forme suivante, plus utile :

$$\forall x \in V, \quad \sum_{\vec{e} \in \vec{E}_x} \gamma(\vec{e}) = 0 \quad (4.8)$$

où l'on a défini :

$$\forall \vec{e} = (x, y) \in \vec{E}, \quad \gamma(\vec{e}) := \alpha(y) \kappa(y, x) - \alpha(x) \kappa(x, y) \quad (4.9)$$

La caractérisation suivante de la réversibilité est prouvée dans [64] sous une forme légèrement différente :

Proposition 4.2.1. *Le générateur \mathcal{L} est réversible si et seulement si κ est réversible, c'est-à-dire si :*

$$\forall x \neq y \in V, \quad \alpha(x) \kappa(x, y) = \alpha(y) \kappa(y, x) \quad (4.10)$$

Dans la formulation de [64], les valeurs de $\kappa(x, y)$ ne sont pas exactement les mêmes, mais comme elles ne sont multipliées que par une quantité dépendant symétriquement de x et y , les équations (4.10) sont multipliées par certaines constantes et le résultat reste le même. La mesure invariante est aussi donnée dans [64] dans le cas réversible, mais le résultat suivant en fournit une généralisation. On définit, pour tout $e \in E$ et $x \in R_e$:

$$\tilde{\mu}_e(x) := \alpha(e^+) \kappa(e^+, e^-) \int_{[e^-, x]} e^{\beta_e(y, x)} \lambda_e(dy) + \alpha(e^-) \kappa(e^-, e^+) \int_{[x, e^+]} e^{\beta_e(y, x)} \lambda_e(dy) \quad (4.11)$$

et par un abus de notation usuel, on écrira aussi $\tilde{\mu}_e$ pour la mesure sur \mathbb{G} dont la densité sur R_e par rapport à λ_e est donnée par (4.11) et s'annulant en-dehors de R_e .

Proposition 4.2.2. *La mesure invariante de \mathcal{L} est donnée par :*

$$\mu = \frac{1}{Z} \sum_{e \in E} \tilde{\mu}_e \quad (4.12)$$

où Z est la constante de normalisation.

Démonstration. On commence la preuve en explicitant deux propriétés importantes sur $\tilde{\mu}_e$, $e \in E$:

$$\begin{aligned} \forall e \in E, x \in R_e, \quad \partial_e \tilde{\mu}_e(x) &= \alpha(e^+) \kappa(e^+, e^-) - \alpha(e^-) \kappa(e^-, e^+) + b_e(x) \tilde{\mu}_e(x) \\ &= \gamma(e^-, e^+) + b_e(x) \tilde{\mu}_e(x) \end{aligned} \quad (4.13)$$

$$\forall v \in V, e \in E_v, \quad \tilde{\mu}_e(v) = \alpha(v) p(v, e) \quad (4.14)$$

Combinées avec (4.8) et la définition de \mathcal{F} , ces deux égalités seront la clé des calculs qui

suivent. Soit $f \in \mathcal{F}$, une simple intégration par partie donne :

$$\begin{aligned}
Z\mu[\mathcal{L}[f]] &= Z \int_{\mathbb{G}} \mathcal{L}[f] d\mu \\
&= \sum_{e \in E} \int_{R_e} \mathcal{L}_e[f|_e](x) \tilde{\mu}_e(x) \lambda_e(dx) \\
&= \sum_{e \in E} \int_{R_e} \left(\partial_e^2 f|_e(x) + b_e(x) \partial_e f|_e(x) \right) \tilde{\mu}_e(x) \lambda_e(dx) \\
&= \sum_{e \in E} \left([\partial_e f|_e(x) \tilde{\mu}_e(x)]_{e^-}^{e^+} + \int_{R_e} \left(b_e(x) \tilde{\mu}_e(x) - \partial_e \tilde{\mu}_e(x) \right) \partial_e f|_e(x) \lambda_e(dx) \right) \\
&= - \sum_{e \in E} \left(f'(e^+, e) \tilde{\mu}_e(e^+) + f'(e^-, e) \tilde{\mu}_e(e^-) \right) \\
&\quad + \sum_{e \in E} \int_{R_e} \left(b_e(x) \tilde{\mu}_e(x) - \partial_e \tilde{\mu}_e(x) \right) \partial_e f|_e(x) \lambda_e(dx)
\end{aligned}$$

La première somme dans la dernière expression peut être réarrangée en :

$$- \sum_{v \in V} \sum_{e \in E_v} f'(v, e) \tilde{\mu}_e(v)$$

qui par (4.14) est égal à :

$$- \sum_{v \in V} \alpha(v) \sum_{e \in E_v} f'(v, e) p(v, e) = 0$$

Pour la deuxième somme, on utilise (4.13) pour obtenir :

$$\begin{aligned}
&\sum_{e \in E} \int_{R_e} \left(b_e(x) \tilde{\mu}_e(x) - \partial_e \tilde{\mu}_e(x) \right) \partial_e f|_e(x) \lambda_e(dx) \\
&= \sum_{e \in E} \left(\alpha(e^-) \kappa(e^-, e^+) - \alpha(e^+) \kappa(e^+, e^-) \right) \int_{R_e} \partial_e f|_e(x) \lambda_e(dx) \\
&= \sum_{e \in E} \gamma(e^+, e^-) \left(f(e^+) - f(e^-) \right) \\
&= \sum_{x \in V} \sum_{y \neq x \in V} \gamma(x, y) f(x) \\
&= \sum_{x \in V} f(x) \sum_{y \neq x \in V} \gamma(x, y)
\end{aligned}$$

qui vaut 0 par (4.8). □

Remarque 4.2.3. Si $\alpha(e^+) \kappa(e^+, e^-) = \alpha(e^-) \kappa(e^-, e^+)$ pour un certain $e \in E$ (c'est-à-dire, $\gamma(e^-, e^+) = 0$), la densité $\tilde{\mu}_e$ se simplifie comme suit :

$$\begin{aligned}
\tilde{\mu}_e(x) &= \alpha(e^-) \kappa(e^-, e^+) \int_{R_e} e^{\beta_e(y, x)} \lambda_e(dy) \\
&= \alpha(e^-) \frac{p(e^-, e)}{\int_{R_e} e^{\beta_e(z, e^-)} \lambda_e(dz)} \int_{R_e} e^{\beta_e(y, x)} \lambda_e(dy) \\
&= \alpha(e^-) \frac{p(e^-, e)}{\int_{R_e} e^{\beta_e(z, e^-)} \lambda_e(dz)} e^{\beta_e(e^-, x)} \int_{R_e} e^{\beta_e(y, e^-)} \lambda_e(dy) \\
&= \alpha(e^-) p(e^-, e) e^{\beta_e(e^-, x)}
\end{aligned}$$

qui est invariant et réversible pour \mathcal{L}_e . Si $\alpha(e^+) \kappa(e^+, e^-) = \alpha(e^-) \kappa(e^-, e^+)$ pour tout $e \in E$, on retrouve la mesure invariante donnée dans [64] dans le cas réversible. De ce fait, les coefficients $\gamma(\vec{e})$, $\vec{e} \in \vec{E}$ peuvent être vus comme le manque de réversibilité de μ .

4.3 Espace d'état dual et formule de Stokes

Ensuite, nous avons besoin d'une notion de sous-ensembles convenables de \mathbb{G} sur lesquels se fera l'intégration. Les calculs qui suivent nous conduisent à choisir \mathcal{G} , l'ensemble de tous les ensembles parfaits (*i.e.* fermés et sans point isolé) D de \mathbb{G} , non vides et dont la frontière ∂D consiste en un nombre fini de points. En vue d'en faire un espace complet pour la distance introduite ci-après, on définit également :

- \mathcal{G}_1 l'ensemble des parties dénombrables de $\mathbb{G} \times [0, 1]$ de la forme $\{(x_i, \alpha_i), i \in \mathbb{N}\}$ tels que les x_i sont distincts deux à deux et que $\sum_{i \in \mathbb{N}} \alpha_i = 1$. Lorsqu'il existe $i \in \mathbb{N}$ tel que $\alpha_i = 1$, l'ensemble précédent sera identifié avec le singleton $\{x_i\}$. Plus largement, lorsque la valeur des coefficients ne nous intéresse pas, nous considérerons cet ensemble comme une union (au plus) dénombrable de singletons.
- \mathcal{G}_2 l'ensemble des ensembles parfaits non vides de \mathbb{G} dont la frontière est infinie dénombrable.
- $\bar{\mathcal{G}} := \mathcal{G} \sqcup \mathcal{G}_1 \sqcup \mathcal{G}_2$.

Dans la suite, \mathcal{G} aura vocation à être l'espace d'état d'un processus dual d'une diffusion sur \mathbb{G} , et $\mathcal{G}_1 \cup \mathcal{G}_2$ constitue la frontière de cet ensemble. On montrera avec la proposition 4.4.4 que \mathcal{G}_1 est inaccessible, ce qui n'est pas aussi évident pour \mathcal{G}_2 .

On définit un noyau de transition Λ de $\bar{\mathcal{G}}$ dans \mathbb{G} par :

$$\forall D \in \bar{\mathcal{G}}, \forall A \in \mathcal{B}(\mathbb{G}), \quad \Lambda(D, A) := \begin{cases} \sum_{(x, \alpha) \in D} \alpha \delta_x(A), & \text{si } D \in \mathcal{G}_1 \\ \frac{\mu(D \cap A)}{\mu(D)}, & \text{sinon} \end{cases} \quad (4.15)$$

où $\mathcal{B}(\mathbb{G})$ désigne l'ensemble des parties mesurable de \mathbb{G} . Notons que si $D \in \mathcal{G} \cup \mathcal{G}_2$, alors $\mu(D) > 0$ et ce noyau est donc bien défini. Pour tout $D \in \bar{\mathcal{G}}$, $\Lambda(D, \cdot)$ est une mesure de probabilité sur \mathbb{G} et on considère l'application $d_{\mathcal{G}} : \bar{\mathcal{G}} \times \bar{\mathcal{G}} \rightarrow \mathbb{R}_+$ définie par :

$$\forall A, B \in \bar{\mathcal{G}}, \quad d_{\mathcal{G}}(A, B) := \rho(\Lambda(A, \cdot), \Lambda(B, \cdot))$$

où on rappelle que ρ désigne la distance de Prokhorov sur $\mathcal{P}(\mathbb{G})$. L'application $d_{\mathcal{G}}$ définit une distance sur $\bar{\mathcal{G}}$, pour laquelle ce dernier est complet et nous le considérerons donc muni de la topologie induite. La principale différence topologique avec la distance de Hausdorff tient dans le fait qu'une suite de $\bar{\mathcal{G}}$ du type $(A \sqcup B_n)_{n \in \mathbb{N}}$, avec $A, B \in \mathcal{G} \cup \mathcal{G}_2$ et $\mu(B_n) \rightarrow 0$, converge vers A , tandis qu'avec la distance de Hausdorff elle ne converge pas dans $\bar{\mathcal{G}}$. Une suite $(A_n)_{n \in \mathbb{N}}$ de $\bar{\mathcal{G}}$ converge vers $A \in \bar{\mathcal{G}}$ si et seulement si, pour toute fonction $f : \mathbb{G} \rightarrow \mathbb{R}$ continue bornée, $\Lambda(A_n, f) \rightarrow \Lambda(A, f)$. Cela correspond à deux cas distincts :

- Soit $\mu(A) > 0$ et $\mu(A_n \Delta A) \rightarrow 0$ (le symbole Δ désignant l'opération de différence symétrique)
- Soit $\mu(A) = 0$, $\mu(A_n) \rightarrow 0$ et pour tout $(x, \alpha) \in A$ la suite $(C_n)_{n \in \mathbb{N}}$ des composantes connexes de A_n contenant x satisfait $\frac{\mu(C_n)}{\mu(A_n)} \rightarrow \alpha$.

On comprend mieux avec ce deuxième cas pourquoi on doit imposer une pondération des singletons pour les éléments de \mathcal{G}_1 . En effet, la convergence (au sens de Hausdorff) d'une suite $(A_n)_{n \in \mathbb{N}}$ de $\mathcal{G} \cup \mathcal{G}_2$ vers une union dénombrable de singletons n'implique pas nécessairement la convergence faible de la suite de mesure $(\Lambda(A_n, \cdot))_{n \in \mathbb{N}}$ vers une limite unique, même si toutes les valeurs d'adhérence auront pour même support cette union de singletons.

Observons enfin que pour tout $A, B \in \bar{\mathcal{G}}$, on a :

$$d_{\mathcal{G}}(A, B) \leq \mu(A \Delta B) \quad (4.16)$$

À $D \in \mathcal{G}$ et $x \in \partial D$, on veut associer l'ensemble $\vec{\partial}D(x) \subset \vec{E}$ des **arêtes entrantes** en x . On commence par considérer sa version non-orientée :

$$\partial D(x) := \left\{ e \in E : x \in \overline{R_e \cap (D \setminus \{x\})} \right\} \quad (4.17)$$

(où la barre représente l'opération de fermeture), à savoir l'ensemble des arêtes e telles que x peut être approché le long de R_e par des points à l'intérieur de D . L'ensemble $\vec{\partial}D(x) \subset \vec{E}$ est l'orientation de $\partial D(x)$ en direction de l'intérieur de D . Son sens heuristique devrait être clair, mais sa définition rigoureuse demande un peu de soin. On distingue deux cas :

- Quand $x \in \partial D$ appartient à V , on définit :

$$\vec{\partial}D(x) \subset \vec{E} := \{ \vec{e} \in \vec{E}_x : e \in \partial D(x) \}$$

- Quand $x \in \partial D \setminus V$, il existe un unique $e \in E$ tel que $x \in R_e$, et on a $\partial D(x) = \{e\}$. Si on suppose que x n'est pas un point isolé de D , alors chaque voisinage de x dans D intersecte soit (x, e^+) soit (e^-, x) (et ces alternatives s'excluent mutuellement, sinon on n'aurait pas $x \in \partial D$, par définition de \mathcal{G}). Dans le premier cas, on définit $\vec{\partial}D(x) = \{(e^-, e^+)\}$ dans le deuxième $\vec{\partial}D(x) = \{(e^+, e^-)\}$.

On étend ensuite les dérivées directionnelles définies en (4.1) :

$$\forall \vec{e} \in \vec{E}, \forall x \in R_e, \forall f \in \mathcal{F}, \quad f'(x, \vec{e}) := i_{\vec{e}} \partial_e f|_e(x)$$

(avec un léger abus de notation, e est l'arête non-orientée correspondant à l'arête orientée \vec{e} , cette convention sera appliquée dans la suite). En particulier, (4.1) peut être réécrit sous la forme :

$$\forall \{x, y\} \in E, \quad f'(x, \{x, y\}) = f'(x, (x, y))$$

Considérons ensuite :

$$\forall e \in E, x \in R_e \quad \mu_e(x) = \frac{\tilde{\mu}_e(x)}{Z}$$

Pour $e \in E$, μ_e est la version continue de la densité par rapport à λ_e de la restriction de μ à R_e . Notons que pour $x \in V$ et $e \in E_x$, on a :

$$\mu_e(x) = \frac{\alpha(x) p(x, e)}{Z} \quad (4.18)$$

donc sauf si pour tout $x \in V$, $p(x, \cdot)$ est la distribution uniforme sur E_x , il n'existe pas de fonction $\underline{\mu}$ sur \mathbb{G} telle que $(\mu_e)_{e \in E}$ est de la forme $(\underline{\mu}|_e)_{e \in E}$, car $\underline{\mu}$ devrait alors prendre plusieurs valeurs sur V .

Finalement, on définit la μ -**dérivée intérieure (normale)** de $f \in \mathcal{F}$ en $x \in \partial D$, pour $D \in \mathcal{G}$, par :

$$\partial_D f(x) := \sum_{\vec{e} \in \vec{\partial}D(x)} \mu_e(x) f'(x, \vec{e})$$

et sa μ -**irréversibilité** en $x \in \partial D$, pour $D \in \mathcal{G}$:

$$\gamma_D f(x) := \sum_{\vec{e} \in \vec{\partial}D(x)} \frac{\gamma(\vec{e})}{Z} f(x)$$

Prenant en compte les conditions de Neumann imposées aux éléments de \mathcal{F} , on calcule que :

$$\forall D \in \mathcal{G}, \forall x \in \partial D, \forall f \in \mathcal{F}, \quad \partial_D f(x) = -\partial_{\overline{\mathbb{G} \setminus D}} f(x) \quad (4.19)$$

(on aura noté que $\overline{\mathbb{G} \setminus D}$ appartient à \mathcal{G} pour $D \in \mathcal{G}$, car cette opération conserve le nombre de points sur la frontière, ou le réduit à 0 si D est une union dénombrable de singletons), et par continuité de f et (4.8) :

$$\forall D \in \mathcal{G}, \forall x \in \partial D, \forall f \in \mathcal{F}, \quad \gamma_D f(x) = -\gamma_{\overline{\mathbb{G} \setminus D}} f(x) \quad (4.20)$$

Après tous ces prérequis, on obtient une formule de Stokes :

Théorème 4.3.1. *On a pour tout $D \in \mathcal{G}$ et tout $f \in \mathcal{F}$,*

$$\int_D \mathcal{L}[f] d\mu = \sum_{x \in \partial D} (-\partial_D f(x) + \gamma_D f(x))$$

On notera que comme conséquence de (4.19) et (4.20), on retrouve que pour tout $f \in \mathcal{F}$,

$$\begin{aligned} \int \mathcal{L}[f] d\mu &= \int_D \mathcal{L}[f] d\mu + \int_{\overline{\mathbb{G} \setminus D}} \mathcal{L}[f] d\mu \\ &= \int_D \mathcal{L}[f] d\mu + \int_{\overline{\mathbb{G} \setminus D}} \mathcal{L}[f] d\mu \\ &= 0 \end{aligned}$$

Remarque 4.3.2. *Dans le membre de droite de l'égalité du théorème 4.3.1, on a utilisé une somme sur la frontière topologique ∂D de D , par analogie avec la formule de Stokes usuelle. Mais conceptuellement il serait mieux de définir la frontière de D comme l'ensemble :*

$$\partial D := \{(x, \vec{e}) \in \partial D \times \vec{E} : \vec{e} \in \vec{\partial} D(x)\}$$

et de considérer la mesure ν sur $\mathbb{G} \times \vec{E}$ donnée par :

$$\nu := \sum_{(x, \vec{e}) \in \mathbb{G} \times \vec{E}} \mu_e(x) \delta_{(x, \vec{e})} \quad (4.21)$$

où δ_y est la mesure de Dirac en y , pour tout objet y . On peut alors écrire lorsque μ est réversible :

$$\int_D \mathcal{L}[f] d\mu = - \int_{\partial D} f' d\nu$$

On remarque que ν donne un poids infini à n'importe quel sous-ensemble (mesurable) de $\mathbb{G} \times \vec{E}$ qui n'est pas fini, ceci expliquant le choix de \mathcal{G} .

Preuve du théorème 4.3.1. On commence avec le cas simple où D est un sous-ensemble fermé et connexe d'une arête R_e , $e \in E$, d'intérieur non-vide, c'est-à-dire $D = [x_1, x_2] \subset R_e$ avec $x_e(x_1) < x_e(x_2)$. On rappelle que :

$$\partial_e \mu_e = \partial_e \frac{\tilde{\mu}_e}{Z} = \frac{\gamma(e^-, e^+)}{Z} + b_e \mu_e$$

et on observe alors que :

$$\mathcal{L}_e = \frac{1}{\mu_e} \left(\partial_e[\mu_e \partial_e \cdot] - \frac{\gamma(e^-, e^+)}{Z} \partial_e \right)$$

donc dans ce cas, en utilisant le théorème de transport, on obtient :

$$\begin{aligned} \int_D \mathcal{L}[f] d\mu &= \int_{[x_1, x_2]} \mathcal{L}_e[f|_e](y) \mu_e(y) \lambda_e(dy) \\ &= \int_{[x_1, x_2]} \partial_e[\mu_e \partial_e f|_e](y) \lambda_e(dy) - \int_{[x_1, x_2]} \frac{\gamma(e^-, e^+)}{Z} \partial_e f(y) \lambda_e(dy) \\ &= \mu_e(x_2) \partial_e f|_e(x_2) - \mu_e(x_1) \partial_e f|_e(x_1) - \frac{\gamma(e^-, e^+)}{Z} (f(x_2) - f(x_1)) \\ &= -\mu_e(x_2) f'(x_2, (e^+, e^-)) - \mu_e(x_1) f'(x_1, (e^-, e^+)) \\ &\quad + \frac{\gamma(e^+, e^-)}{Z} f(x_2) + \frac{\gamma(e^-, e^+)}{Z} f(x_1) \\ &= -\partial_D f(x_2) - \partial_D f(x_1) + \gamma_D f(x_2) + \gamma_D f(x_1) \end{aligned}$$

et dans ce cas particulier le théorème est prouvé. Dans le cas général, on veut écrire n'importe quel $D \in \mathcal{G}$ comme une réunion d'ensembles de la forme précédente. Pour $D \in \mathcal{G}$, soit $(D_i)_{i \in I}$ la famille (finie) des ensembles qui sont une composante connexe de $D \cap R_e$ pour un $e \in E$. Alors, pour tout $i \in I$, il existe $x_1^i, x_2^i \in \partial D \cup V$ et $e_i \in E$ tels que $D_i = [x_1^i, x_2^i] \subset R_{e_i}$ avec $x_e(x_1^i) < x_e(x_2^i)$. On obtient :

$$D = \bigcup_{i \in I} D_i$$

Cette union n'est pas disjointe, mais on remarque cependant que $\bigcup_{i \neq j} D_i \cap D_j$ n'a qu'un nombre fini de points, tous dans V . On a aussi $\partial D \subset \{x_j^i : i \in I, j = 1, 2\} = \bigcup_{e \in E} \partial(D \cap R_e)$ et :

$$\begin{aligned} \int_D \mathcal{L}[f] d\mu &= \sum_{i \in I} \int_{x_1^i}^{x_2^i} \mathcal{L}[f] d\mu \\ &= \sum_{i \in I} \left(-\mu_{e_i}(x_2^i) f'(x_2^i, (e_i^+, e_i^-)) - \mu_{e_i}(x_1^i) f'(x_1^i, (e_i^-, e_i^+)) \right. \\ &\quad \left. + \frac{\gamma(e_i^-, e_i^+)}{Z} f(x_2^i) + \frac{\gamma(e_i^+, e_i^-)}{Z} f(x_1^i) \right) \end{aligned}$$

donc on veut regrouper les termes dans la somme ci-dessus de sorte à faire apparaître clairement les $\partial_D f(x)$ et $\gamma_D f(x)$ pour chaque $x \in \partial D$ et à annuler le reste. On considère $x \in \bigcup_{i \in I} \{x_1^i, x_2^i\}$. On peut distinguer trois cas :

- si $x \notin V$, alors $x \in \mathring{R}_e$ pour un certain $e \in E$, $x \in \partial D$ et on a $\mu_e(x) f'(x, \vec{e}) = \partial_D f(x)$ et $\frac{\gamma(\vec{e})}{Z} f(x) = \gamma_D f(x)$ (où \vec{e} est la seule arête orientée dans $\vec{\partial} D(x)$). De plus, il existe une unique paire $(i, j) \in I \times \{1, 2\}$ telle que $x = x_j^i$.
- si $x \in V \cap \mathring{D}$: pour tout $i \in I$, si $x_1^i = x$ ou $x_2^i = x$, alors $e_i \in E_x$. Réciproquement, pour tout $e \in E_x$ il existe un unique indice $i \in I$ tel que $e_i = e$ et soit $x_1^i = x (= e^-)$ soit $x_2^i = x (= e^+)$. En utilisant (4.18) et la condition de Neumann en x on obtient :

$$\begin{aligned} \sum_{i \in I: x_1^i = x} \mu_e(x_1^i) f'(x_1^i, (e_i^-, e_i^+)) + \sum_{i \in I: x_2^i = x} \mu_e(x_2^i) f'(x_2^i, (e_i^+, e_i^-)) &= \sum_{e \in E_x} \mu_e(x) f'(x, e) \\ &= 0 \end{aligned}$$

et par (4.8) :

$$\left(\sum_{i \in I: x_1^i = x} \gamma(e_i^-, e_i^+) + \sum_{i \in I: x_2^i = x} \gamma(e_i^+, e_i^-) \right) \frac{f(x)}{Z} = 0$$

- si $x \in V \cap \partial D$: pour tout $\vec{e} \in \vec{\partial}D(x)$ il existe un unique $i \in I$ tel que $e_i = e$ et soit $x_1^i = x$ (si $\vec{e} = (e^-, e^+)$) soit $x_2^i = x$ (si $\vec{e} = (e^+, e^-)$). Réciproquement, pour tout indice $i \in I$, si $x_1^i = x$ alors $(e_i^-, e_i^+) \in \vec{\partial}D(x)$, et si $x_2^i = x$, alors $(e_i^+, e_i^-) \in \vec{\partial}D(x)$. On en déduit que :

$$\begin{aligned} \sum_{i \in I: x_1^i = x} \mu_e(x_1^i) f'(x_1^i, (e_i^-, e_i^+)) + \sum_{i \in I: x_2^i = x} \mu_e(x_2^i) f'(x_2^i, (e_i^+, e_i^-)) &= \sum_{\vec{e} \in \vec{\partial}D(x)} \mu_e(x) f'(x, \vec{e}) \\ &= \partial_D f(x) \end{aligned}$$

et :

$$\begin{aligned} \left(\sum_{i \in I: x_1^i = x} \gamma(e_i^-, e_i^+) + \sum_{i \in I: x_2^i = x} \gamma(e_i^+, e_i^-) \right) \frac{f(x)}{Z} &= \sum_{\vec{e} \in \vec{\partial}D(x)} \frac{\gamma(\vec{e})}{Z} f(x) \\ &= \gamma_D f(x) \end{aligned}$$

Ces trois cas étant exhaustif, on peut décomposer I ainsi :

$$I = \bigcup_{j=1,2} \left(\bigcup_{x \in \partial D \setminus V} \{i \in I : x_j^i = x\} \cup \bigcup_{x \in \partial D \cap V} \{i \in I : x_j^i = x\} \cup \bigcup_{x \in \bar{D} \cap V} \{i \in I : x_j^i = x\} \right)$$

et on obtient :

$$\begin{aligned} - \sum_{i \in I} (\mu_{e_i}(x_2^i) f'(x_2^i, (e_i^+, e_i^-)) + \mu_{e_i}(x_1^i) f'(x_1^i, (e_i^-, e_i^+))) &= - \sum_{x \in \partial D} \partial_D f(x) \\ \frac{1}{Z} \sum_{i \in I} (\gamma(e_i^-, e_i^+) f(x_2^i) + \gamma(e_i^+, e_i^-) f(x_1^i)) &= \sum_{x \in \partial D} \gamma_D f(x) \end{aligned}$$

ce qui conclut la preuve. □

4.4 Générateur dual

On veut construire des processus de diffusion $(D_t)_{t \geq 0}$ ayant \mathcal{G} comme espace d'état, et qui sont entrelacés par dualité avec X . Pour aller dans cette direction on considère un cadre inspiré de l'approche introduite dans [18] pour des diffusions elliptiques sur des variétés différentielles.

On commence par définir un générateur markovien \mathfrak{L} sur un plus petit sous-ensemble \mathcal{D} de \mathcal{G} , similairement au fait que \mathcal{L} est naturellement bien défini seulement sur $\mathbb{G} \setminus V$. On prend pour \mathcal{D} l'ensemble des $D \in \mathcal{G}$ dont la frontière n'intersecte pas V . On définit ensuite des observables convenables sur lesquelles on va faire agir \mathfrak{L} . Le rôle des **observables élémentaires** est joué par les applications :

$$F_f : \mathcal{D} \ni D \mapsto F_f(D) := \int_D f \, d\mu \quad (4.22)$$

associées aux fonctions $f \in \mathcal{F}$.

L'action de \mathfrak{L} sur les observables élémentaires ci-dessus est donné par :

$$\forall f \in \mathcal{F}, \forall D \in \mathcal{D}, \quad \mathfrak{L}[F_f](D) = \int_D \mathcal{L}[f] d\mu + 2 \frac{\underline{\mu}(\partial D)}{\mu(D)} \int_{\partial D} f d\underline{\mu} \quad (4.23)$$

où $2\underline{\mu}$ est la mesure sur \mathbb{G} obtenu par projection sur le premier facteur de la mesure ν définie en (4.21). On note que par définition de \mathcal{D} , on a $\mu(D) > 0$ pour tout $D \in \mathcal{D}$, donc le membre de droite ci-dessus est bien défini. Pour $D \in \mathcal{G} \setminus \mathcal{D}$, on pose de manière arbitraire

$$\forall f \in \mathcal{F}, \quad \mathfrak{L}[F_f](D) = \lim_{\substack{D' \rightarrow D \\ D' \in \mathcal{D}}} \mathfrak{L}[F_f](D')$$

Remarque 4.4.1. Dans (4.23), le terme $\int_{\partial D} f d\underline{\mu}$ est juste une somme finie, et comme la frontière des éléments de \mathcal{D} n'intersecte pas V , il peut être réécrit $\sum_{x \in \partial D} f(x) \underline{\mu}(x)$, où $\underline{\mu}$ est la fonction définie sur $\mathbb{G} \setminus V$ par

$$\forall x \in \mathbb{G} \setminus V, \quad \underline{\mu}(x) := \mu_{e(x)}(x)$$

où $e(x) \in E$ et l'unique arête telle que $x \in R_{e(x)}$.

Sur V , on pose :

$$\forall x \in V, \quad \underline{\mu}(x) = \frac{1}{2} \sum_{e \in E_x} \mu_e(x)$$

Mais ce choix est arbitraire pour la définition de \mathfrak{L} ; à la place de toutes les arêtes adjacentes, on aurait pu considérer seulement les arêtes entrantes ou sortantes de D (introduisant alors une dépendance supplémentaire en D pour $\underline{\mu}$). Cependant, aucune de ces alternatives n'est vraiment satisfaisante, d'où la définition de \mathcal{D} . Similairement à $\mathcal{L}[f]$ sur \mathbb{G} , il n'est pas possible en général d'étendre $\mathfrak{L}[F_f]$ en une fonction continue sur \mathcal{G} , pour tout $f \in \mathcal{F}$. En fait, sauf hypothèses particulières sur tous les sommets du graphe (que nous n'explicitons pas ici), $\mathfrak{L}[F_f]$ n'est pas continue sur \mathbb{G} , pour toute fonction $f \in \mathcal{F}$ non identiquement nulle.

On introduit une forme bilinéaire symétrique $\Gamma_{\mathfrak{L}}$ (qui sera la carré du champ associé à \mathfrak{L}) sur de telles fonctionnelles, via :

$$\forall f, g \in \mathcal{F}, \forall D \in \mathcal{D}, \quad \Gamma_{\mathfrak{L}}[F_f, F_g](D) = \left(\int_{\partial D} f d\underline{\mu} \right) \left(\int_{\partial D} g d\underline{\mu} \right) \quad (4.24)$$

Comme les processus de Markov à valeurs dans \mathcal{G} auxquels on s'intéresse seront à trajectoires continues, nous sommes naturellement amenés à la définition suivante (voir encore section 1.11.1 de Bakry, Gentil et Ledoux [5]).

On considère \mathfrak{D} l'algèbre constituée des **observables** de la forme $\mathfrak{F} := \mathfrak{f}(F_{f_1}, \dots, F_{f_n})$, où $n \in \mathbb{Z}_+$, $f_1, \dots, f_n \in \mathcal{F}$ et $\mathfrak{f} : \mathcal{R} \rightarrow \mathbb{R}$ est une application \mathcal{C}^2 , avec \mathcal{R} un ouvert de \mathbb{R}^n contenant l'image de \mathcal{D} par $(F_{f_1}, \dots, F_{f_n})$. Pour une telle fonctionnelle \mathfrak{F} , on pose :

$$\mathfrak{L}[\mathfrak{F}] = \sum_{j=1}^n \partial_j \mathfrak{f}(F_{f_1}, \dots, F_{f_n}) \mathfrak{L}[F_{f_j}] + \sum_{k,l=1}^n \partial_k \mathfrak{f}(F_{f_1}, \dots, F_{f_n}) \Gamma_{\mathfrak{L}}[F_{f_k}, F_{f_l}] \quad (4.25)$$

À deux éléments \mathfrak{D} , $\mathfrak{F} := \mathfrak{f}(F_{f_1}, \dots, F_{f_n})$ et $\mathfrak{G} := \mathfrak{g}(F_{g_1}, \dots, F_{g_m})$, on associe également :

$$\Gamma_{\mathfrak{L}}[\mathfrak{F}, \mathfrak{G}] := \sum_{l=1}^n \sum_{k=1}^m \partial_l \mathfrak{f}(F_{f_1}, \dots, F_{f_n}) \partial_k \mathfrak{g}(F_{g_1}, \dots, F_{g_m}) \Gamma_{\mathfrak{L}}[F_{f_l}, F_{g_k}]$$

Remarque 4.4.2. *A priori les définitions ci-dessus sont ambiguës, car elles semblent dépendre de l'écriture de $\mathfrak{F} \in \mathfrak{D}$ sous la forme $\mathfrak{f}(F_{f_1}, \dots, F_{f_n})$ et similairement pour \mathfrak{G} . Pour voir qu'elles sont en fait bien définies, noter que :*

$$\forall \mathfrak{F}, \mathfrak{G} \in \mathfrak{D}, \quad \Gamma_{\mathfrak{L}}[\mathfrak{F}, \mathfrak{G}] = \frac{1}{2} (\mathfrak{L}[\mathfrak{F}\mathfrak{G}] - \mathfrak{F}\mathfrak{L}[\mathfrak{G}] - \mathfrak{G}\mathfrak{L}[\mathfrak{F}])$$

Cette propriété implique que si \mathfrak{f} est un polynôme à n variables, alors pour tout $\mathfrak{F} := \mathfrak{f}(f_1, \dots, f_n)$, avec $f_1, \dots, f_n \in \mathcal{F}$, l'objet $\mathfrak{L}[\mathfrak{F}]$ est uniquement déterminé. Ceci repose sur une récurrence sur le degré de \mathfrak{f} , initialisée par (4.23) et (4.24). Le cas général des fonctions régulières \mathfrak{f} se déduit de leur approximation sur tout compact par des applications polynomiales.

Une de nos principales préoccupations est de prouver que le problème de martingale associé à $(\mathfrak{L}, \mathfrak{D})$ est bien posé. Pour justifier cette tâche au préalable, on vérifie que \mathfrak{L} est entrelacé avec \mathcal{L} par le noyau de transition Λ , défini en (4.15). En suivant les preuves des propositions 20 et 23 de Miclo [65], on pourra alors montrer, si le problème de martingale est bien posé, que l'on peut construire un couple (X, X^*) de manière à ce que les marginales X et X^* aient pour générateurs respectifs \mathcal{L} et \mathfrak{L} , et que X^* soit un dual de stationnarité forte pour X .

Lemme 4.4.3. *\mathcal{L} et \mathfrak{L} sont Λ -liés, c'est-à-dire :*

$$\forall f \in \mathcal{F}, \quad \mathfrak{L}[\Lambda[f]] = \Lambda[\mathcal{L}[f]] \quad (4.26)$$

Démonstration. Par définition de Λ , on a, pour tout $f \in \mathcal{F}$ et tout $D \in \mathfrak{D}$:

$$\begin{aligned} \Lambda[f](D) &= \frac{\int_D f d\mu}{\mu(D)} \\ &= \frac{F_f}{F_1}(D) \\ &= \mathfrak{f}(F_f, F_1)(D) \end{aligned}$$

où $\mathbf{1}$ est la fonction constante égale à 1 et :

$$\begin{aligned} \mathfrak{f} : \mathbb{R} \times \mathbb{R}^* &\rightarrow \mathbb{R} \\ (x, y) &\mapsto \frac{x}{y} \end{aligned}$$

L'application $\mathfrak{F} := \mathfrak{f}(F_f, F_1)$ est dans \mathfrak{D} et :

$$\mathfrak{L}[\mathfrak{F}](D) = \frac{\mathfrak{L}[F_f]}{F_1}(D) - \frac{F_f \mathfrak{L}[F_1]}{F_1^2}(D) - 2 \frac{\Gamma_{\mathfrak{L}}[F_f, F_1]}{F_1^2}(D) + 2 \frac{F_f \Gamma_{\mathfrak{L}}[F_1, F_1]}{F_1^3}(D)$$

où :

$$\begin{aligned} \mathfrak{L}[F_1](D) &= 2 \frac{\mu(\partial D)^2}{F_1(D)} \\ \Gamma_{\mathfrak{L}}[F_f, F_1](D) &= \mu(\partial D) \int_{\partial D} f d\mu \\ \Gamma_{\mathfrak{L}}[F_1, F_1](D) &= \mu(\partial D)^2 \end{aligned}$$

donc au final :

$$\begin{aligned} \mathfrak{L}[\mathfrak{F}](D) &= \frac{\int_D \mathcal{L}[f] d\mu}{\mu(D)} \\ &= \Lambda[\mathcal{L}[f]](D) \end{aligned}$$

Maintenant, si D est dans $\mathcal{G} \setminus \mathcal{D}$, alors :

$$\begin{aligned} \mathfrak{L}[\mathfrak{F}](D) &= \lim_{\substack{D' \rightarrow D \\ D' \in \mathcal{D}}} \mathfrak{L}[\mathfrak{F}](D') \\ &= \lim_{\substack{D' \rightarrow D \\ D' \in \mathcal{D}}} \Lambda[\mathcal{L}[f]](D') \\ &= \Lambda[\mathcal{L}[f]](D) \end{aligned}$$

ce qui conclut la preuve. \square

Le résultat suivant justifie, d'une part, qu'un processus de générateur \mathfrak{L} , s'il existe, est absorbé en temps fini (ce qui, rappelons-le, est le but recherché), et d'autre part que l'ensemble \mathcal{G}_1 constitue une frontière inaccessible du domaine \mathcal{G} .

Proposition 4.4.4. *Soient μ_0 une mesure de probabilité sur \mathcal{G} et $(Y_t)_{t \in [0, T]}$ une solution du problème de martingale associé à (μ_0, \mathfrak{L}) jusqu'au temps T de sortie de \mathcal{G} , telle que :*

$$\forall t \geq 0, \quad \int_0^{t \wedge T} \mathbf{1}_{\{Y_s \in \mathcal{G} \setminus \mathcal{D}\}} ds = 0$$

Alors jusqu'au temps T , $(\mu(Y_t))_{t \geq 0}$ est un processus de Bessel de dimension 3, changé de temps et stoppé en 1.

Démonstration. Soit $f \in \mathcal{C}^2(\mathbb{R})$. Comme $\mu = F_1$, où $\mathbf{1}$ est la fonction constante égale à un, il existe une martingale locale $(M_t)_{t \geq 0}$ telle que :

$$\forall t > 0, \quad f(\mu(Y_{t \wedge T})) = f(\mu(Y_0)) + \int_0^{t \wedge T} \mathfrak{L}[f \circ \mu](Y_s) ds + M_{t \wedge T}$$

Par définition de \mathfrak{L} (voir (4.25)) on a :

$$\mathfrak{L}(f \circ \mu) = f'(\mu)\mathfrak{L}(\mu) + f''(\mu)\Gamma_{\mathfrak{L}}(\mu, \mu)$$

où pour tout $D \in \mathcal{D}$:

$$\begin{aligned} \mathfrak{L}(\mu)(D) &= \int_D \mathcal{L}(\mathbf{1}) d\mu + 2 \frac{\mu(\partial D)}{\mu(D)} \int_{\partial D} \mathbf{1} d\mu \\ &= 2 \frac{\mu(\partial D)^2}{\mu(D)} \end{aligned}$$

et :

$$\begin{aligned} \Gamma_{\mathfrak{L}}(\mu, \mu) &= \left(\int_{\partial D} \mathbf{1} d\mu \right)^2 \\ &= \mu(\partial D)^2 \end{aligned}$$

Donc, si on pose :

$$\hat{L} = \frac{1}{2} \partial^2 + \frac{1}{x} \partial$$

le générateur d'un processus de Bessel de dimension 3, on obtient :

$$\forall t > 0, \quad f(\mu(Y_{t \wedge T})) = f(\mu(Y_0)) + 2 \int_0^{t \wedge T} \mu(\partial Y_s)^2 \hat{L}[f](\mu(Y_s)) ds + M_{t \wedge T}$$

On définit τ et ς par :

$$\begin{aligned}\tau &:= \inf\{t \geq 0 : Y_t = \mathbb{G}\} \wedge T \\ \varsigma &:= 2 \int_0^\tau \underline{\mu}(\partial Y_s)^2 ds\end{aligned}$$

On veut définir le changement de temps $(\theta_t)_{t \in [0, \varsigma]}$ par :

$$\forall t \in [0, \varsigma], \quad 2 \int_0^{\theta_t} \underline{\mu}(\partial Y_s)^2 ds = t$$

mais le problème est que le changement de variable $s \mapsto \theta_s$ n'est pas \mathcal{C}^1 car $\mathcal{G} \ni D \mapsto \underline{\mu}(\partial D)$ n'est pas continue sur \mathcal{D} . Pour résoudre ce problème, on approche cette fonction par une suite de fonctions continues. On définit donc, pour tout $e \in E, x \in R_e$ et $n \in \mathbb{N}$:

$$\underline{\mu}_n(\{x\}) = n(d(x, e^-) \mathbf{1}_{\{d(x, e^-) \leq 1/n\}} + d(e^+, x) \mathbf{1}_{\{d(x, e^+) \leq 1/n\}}) \underline{\mu}(x)$$

et pour tout $D \in \mathcal{G}$:

$$\underline{\mu}_n(\partial D) = \sum_{x \in \partial D} \underline{\mu}_n(\{x\})$$

de sorte que $\lim_{n \rightarrow +\infty} \underline{\mu}_n(\partial D) = \underline{\mu}(\partial D)$ pour tout $D \in \mathcal{D}$. On définit également, pour tout $k, n \in \mathbb{N}$:

$$\begin{aligned}\tau^k &:= \inf\{t \geq 0 : \mu(Y_t) \leq 1/k\} \wedge \tau \\ \varsigma^k &:= 2 \int_0^{\tau^k} \underline{\mu}(\partial Y_s)^2 ds \\ \varsigma_n &:= 2 \int_0^{\tau} \underline{\mu}_n(\partial Y_s)^2 ds \\ \varsigma_n^k &:= \varsigma_n \wedge \varsigma^k\end{aligned}$$

et :

$$\forall t \in [0, \varsigma_n], \quad 2 \int_0^{\theta_t^n} \underline{\mu}_n(\partial Y_s)^2 ds = t$$

Par convergence dominée et un changement de variable ($\theta_{u \wedge \varsigma_n^k}^n = s$) on a :

$$\begin{aligned}\mathfrak{f}(\mu(Y_{\theta_{t \wedge \varsigma^k}})) &= \lim_{n \rightarrow \infty} \mathfrak{f}(\mu(Y_{\theta_{t \wedge \varsigma_n^k}^n})) \\ &= \mathfrak{f}(\mu(Y_0)) + 2 \lim_{n \rightarrow \infty} \int_0^{\theta_{t \wedge \varsigma_n^k}^n} \underline{\mu}_n(\partial Y_s)^2 \hat{L}[\mathfrak{f}](\mu(Y_s)) ds + M_{\theta_{t \wedge \varsigma_n^k}^n} \\ &= \mathfrak{f}(\mu(Y_0)) + \lim_{n \rightarrow \infty} \int_0^t \hat{L}[\mathfrak{f}](\mu(Y_{\theta_{u \wedge \varsigma_n^k}^n})) du + M_{\theta_{t \wedge \varsigma_n^k}^n} \\ &= \mathfrak{f}(\mu(Y_0)) + \int_0^{t \wedge \varsigma^k} \hat{L}[\mathfrak{f}](\mu(Y_{\theta_u})) du + M_{\theta_{t \wedge \varsigma^k}}\end{aligned}$$

c'est-à-dire, le processus $(\mu(Y_{\theta_t}))_{t \geq 0}$ est solution du problème de martingale associé à \hat{L} jusqu'au temps ς^k . Puisque \hat{L} est le générateur d'un processus de Bessel, on a $\lim_{k \rightarrow +\infty} \varsigma_k = \varsigma$, et comme sur l'ensemble $\{\tau < \infty\} \cap \{\tau < T\}$, on a $Y_{\theta_\varsigma} = Y_\tau = \mathbb{G}$, c'est donc un processus de Bessel de dimension 3, stoppé en 1, jusqu'au temps T . \square

Une conséquence pratique de cette proposition est que si $(D_t)_{t \geq 0}$ est un processus de générateur \mathfrak{L} , alors $\mu(D_t) > 0$ pour tout $t > 0$ (presque sûrement).

Conclusion

Nous avons donc construit un opérateur \mathfrak{L} , agissant sur une classe \mathfrak{F} de fonctions sur un ensemble \mathcal{G} de sous-parties du graphe. De la même manière qu'un processus de diffusion classique sur \mathbb{R}^n passe un temps nul sur n'importe quelle partie de mesure nulle, il semble raisonnable de supposer que s'il existe un processus de Markov $(D_t)_{t \geq 0}$ associé à cet opérateur, alors les extrémités de celui-ci passent un temps nul sur chaque sommet et donc que le processus passe un temps nul en \mathcal{D} . Pour cette raison, les valeurs de $\mathfrak{L}f$ sur \mathcal{D} , pour tout $f \in \mathfrak{F}$, sont définies de manière arbitraire.

Si l'on suppose donc qu'un tel processus $(D_t)_{t \geq 0}$ existe, au moins jusqu'au temps T de sortie de \mathcal{G} , nous avons alors montré qu'il est un dual de stationnarité forte pour la diffusion associée à \mathfrak{L} . De plus, l'ensemble \mathcal{G}_1 constitue une frontière entrante pour ce processus, et le temps T est donc le temps d'atteinte de \mathcal{G}_2 , l'ensemble des parties compactes sans points isolés ayant une infinité de composantes connexes.

Il reste donc à montrer l'existence de ce processus, et l'inaccessibilité de la frontière \mathcal{G}_2 . Dans les chapitres suivants, c'est à ces questions que nous nous intéresserons, ainsi qu'à celle de l'unicité du processus (en étendant cependant la définition de \mathfrak{L} à une classe de fonctions plus large que \mathfrak{F}).

Chapitre 5

Étude d'un cas particulier

Ce chapitre est dédié à l'étude du dual d'un mouvement brownien sur un graphe en étoile dont les arêtes ont la même longueur (voir figure 5.1), et a pour but de donner des pistes pour étudier l'existence et l'unicité d'une solution au problème de martingale dans le cas général. Dans ce cadre particulier, on se ramène à l'intervalle $[-1, 1]$ avec une singularité en 0, et le processus dual est une diffusion sur un compact de $[-1, 1]^2$, dont la dérive présente une singularité. Ce genre de processus a été étudié par Takanobu [84],[85] notamment, au moyen d'une équation différentielle stochastique. Dans un premier temps nous allons donc établir un lien formel entre cette EDS et un problème de martingale.

5.1 Diffusion avec une dérive singulière

Le processus qui sera au centre de notre étude dans la section 5.2 est une diffusion sur \mathbb{R}^2 avec une barrière perméable (les axes $x = 0$ et $y = 0$). Ce type de processus peut-être vu comme la version multi-dimensionnelle du mouvement brownien biaisé (voir par exemple Lejay [51]), et se rapproche également des processus de diffusion sur une région de \mathbb{R}^n avec des conditions au bord (réflexion, absorption ou élasticité). Cependant, contrairement à ces derniers types de processus, sur lesquels la littérature est abondante (Wentzell [90], Stroock et Varadhan [81], Takanobu et Watanabe [86], etc.), les diffusions avec une frontière perméables n'ont, à notre connaissance, pas été beaucoup étudiées. On peut, outre le résultat d'existence et d'unicité qui nous intéresse, mentionner une étude analytique de Portenko [72] sur le semi-groupe d'une telle diffusion et une approche par un problème de sous-martingale de Anulova [4].

Le résultat qui nous sera utile ici est dû à Takanobu [84] et concerne l'existence et l'unicité de la solution d'une équation différentielle stochastique multi-dimensionnelle avec un temps local. Bien que l'équivalence de cette EDS avec un problème de martingale ait été

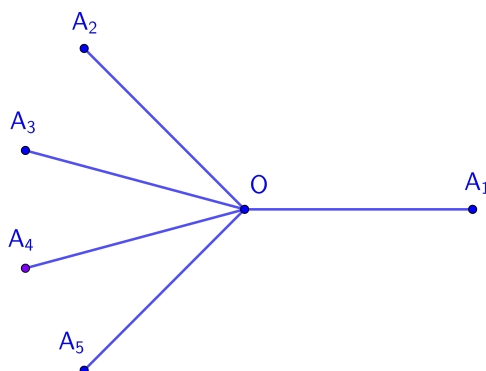


FIGURE 5.1 – Exemple de graphe en étoile à cinq branches

évoquée par ce même auteur, elle n'a pas, à notre connaissance et à ce jour, été démontrée formellement. Notre tâche ici sera donc de démontrer cette équivalence afin de pouvoir exploiter les résultats obtenus par Takanobu. Nous considérons ici un cas particulier des EDS étudiées par celui-ci, mais néanmoins plus général que celui que nous utiliserons dans la suite. Le résultat énoncé doit encore pouvoir s'étendre assez facilement.

Les notations de cette section sont indépendantes, et nous en reprendrons notamment certaines de Anulova [4].

On se place sur \mathbb{R}^n , $n \in \mathbb{N}$, que l'on identifie naturellement avec $\mathbb{R} \times \mathbb{R}^{n-1}$, et on note \mathcal{C}^0 , \mathcal{C}^1 , \mathcal{C}^2 pour $\mathcal{C}^0(\mathbb{R}^n)$, $\mathcal{C}^1(\mathbb{R}^n)$, $\mathcal{C}^2(\mathbb{R}^n)$. Pour $f \in \mathcal{C}^0$ ou $\mathcal{C}^0(\mathbb{R}^{n-1})$, on note $\partial_i f$ la i -ième dérivée partielle de f lorsqu'elle existe, et $\partial_{i,j} f = \partial_i \partial_j f$, $\partial_i^2 f = \partial_{i,i} f$, $i, j = 1, \dots, n$.

On décompose \mathbb{R}^n en deux demi-espaces et leur frontière : $\Lambda_{\pm} = \mathbb{R}_{\pm}^* \times \mathbb{R}^{n-1}$ et $\Gamma = \{0\} \times \mathbb{R}^{n-1}$, de sorte que l'on a $\mathbb{R}^n = \Lambda_+ \cup \Lambda_- \cup \Gamma$ avec l'identification $\mathbb{R}^n = \mathbb{R} \times \mathbb{R}^{n-1}$ déjà mentionnée. On définit également $\bar{\Lambda}_{\pm} = \Lambda_{\pm} \cup \Gamma = \mathbb{R}_{\pm} \times \mathbb{R}^{n-1}$

Si $f : \mathbb{R} \rightarrow \mathbb{R}$ admet une limite à droite (resp. à gauche) au point a , on notera $f(a^+) = \lim_{x \rightarrow a^+} f(x)$ (resp. $f(a^-) = \lim_{x \rightarrow a^-} f(x)$). En particulier, si $f \in \mathcal{C}^0(\mathbb{R}^n)$ est telle que sa restriction à $\bar{\Lambda}_{\pm}$ est \mathcal{C}^1 , $\partial_1 f(0^{\pm}, y)$ désigne la dérivée à droite ou à gauche de la première application partielle de f , $x \mapsto f(x, y)$, en 0, pour tout $y \in \mathbb{R}^{n-1}$.

On définit également $\hat{\mathcal{C}}$ comme l'ensemble des fonctions $f : \mathbb{R}^n \rightarrow \mathbb{R}$ satisfaisant :

- $f|_{\bar{\Lambda}_{\pm}} \in \mathcal{C}^2(\bar{\Lambda}_{\pm})$ (et en particulier, f est continue sur \mathbb{R}^n)
- Pour tout $i = 2, \dots, n$, $(\partial_i f)|_{\bar{\Lambda}_{\pm}} \in \mathcal{C}^1(\bar{\Lambda}_{\pm})$ (on suppose que ces dérivées partielles existent et elles sont alors continues sur \mathbb{R}^n)
- Pour tout $i, j = 2, \dots, n$, $\partial_{i,j} f \in \mathcal{C}^0(\mathbb{R}^n)$

On notera par exemple que la troisième condition ne découle pas de la deuxième : cette dernière signifie que les dérivées partielles $\partial_i f$, $i = 2, \dots, n$, sont bien définies sur \mathbb{R}^n , et sont continûment différentiables sur chacun des demi-espaces fermés $\bar{\Lambda}_+$ et $\bar{\Lambda}_-$. En revanche, la valeur de leurs différentielles sur Γ n'est pas nécessairement la même selon que l'on différentie sur l'un ou l'autre de ces demi-espaces.

Heuristiquement, les éléments de $\hat{\mathcal{C}}$ sont ceux de \mathcal{C}^2 , pour lesquels on a autorisé la première application partielle à ne pas être \mathcal{C}^2 en 0, mais simplement deux fois dérivable à droite et à gauche. En particulier, $\partial_{i,j} f(x, y)$ est bien défini pour tout $i, j = 1, \dots, n$ et tout $(x, y) \in \mathbb{R} \times \mathbb{R}^{n-1}$, sauf si $x = 0$ et i ou $j = 1$. Dans ce dernier cas, on peut remplacer x par 0^{\pm} , et les applications $\mathbb{R}^{n-1} \ni y \mapsto \partial_{i,j} f(0^{\pm}, y)$ sont continues.

On se donne ensuite des paramètres $\sigma = (\sigma_{i,j}(\cdot))_{i,j \in \llbracket 1, n \rrbracket} : \mathbb{R}^n \rightarrow \mathcal{M}_n(\mathbb{R})$, $a = \sigma \sigma^t = (a_{i,j}(\cdot))_{i,j \in \llbracket 1, n \rrbracket} : \mathbb{R}^n \rightarrow \mathcal{S}_n^+$ (l'ensemble des matrices $n \times n$ symétriques et positives), $b = (b_i(\cdot))_{i \in \llbracket 1, n \rrbracket} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ et $c > 0$, avec σ , a et b continus et $a_{i,i}(\cdot) > 0$ pour tout $i \in \llbracket 1, n \rrbracket$, et on introduit les opérateurs \mathcal{L} et B , dont l'action sur $f \in \hat{\mathcal{C}}$ est définie par

$$\begin{aligned} \forall z \in \Lambda_+ \cup \Lambda_-, \quad \mathcal{L}f(z) &= \left(\frac{1}{2} \sum_{i,j=1}^n a_{i,j}(z) \partial_{i,j} f(z) + \sum_{i=1}^n b_i(z) \partial_i f(z) \right) \\ \forall z = (0, y) \in \Gamma, \quad \mathcal{L}f(z) &= \alpha[f](y) \\ \forall z = (x, y) \in \mathbb{R}^n, \quad Bf(z) &= \partial_1 f(0^+, y) - c \partial_1 f(0^-, y) \end{aligned}$$

où, pour simplifier, pour tout $f \in \hat{\mathcal{C}}$ et tout $y \in \mathbb{R}^{n-1}$, $\alpha[f](y)$ est une combinaison convexe de $\lim_{x \rightarrow 0^+} \mathcal{L}f(x, y)$ et de $\lim_{x \rightarrow 0^-} \mathcal{L}f(x, y)$, dépendant continûment de y et indépendante de f . En particulier, si f est \mathcal{C}^2 , alors $\mathcal{L}f$ est continu.

Enfin, on définit l'ensemble

$$\mathcal{C}^B := \{f \in \hat{\mathcal{C}} : Bf \equiv 0\}$$

Théorème 5.1.1. Soient $Z = (X_t, Y_t)_{t \geq 0}$ un processus sur \mathbb{R}^n à trajectoires continues bornées, avec $Y = (Y^2, \dots, Y^n) \in \mathbb{R}^{n-1}$, et pour tout $f \in \hat{\mathcal{C}}$ le processus $\tilde{M}^f = (\tilde{M}_t^f)_{t \geq 0}$ défini par

$$\forall t \geq 0, \quad \tilde{M}_t^f = f(Z_t) - f(Z_0) - \int_0^t \mathcal{L}f(Z_s) ds$$

Si X est une semi-martingale, on note $\{L_t^x, (x, t) \in \mathbb{R} \times \mathbb{R}^+\}$ ses temps locaux (cf Revuz et Yor [75]) et $M^f = (M_t^f)_{t \geq 0}$ le processus défini pour tout $f \in \hat{\mathcal{C}}$ par

$$\forall t \geq 0, \quad M_t^f = \tilde{M}_t^f - \int_0^t Bf(Z_s) dL_s^0$$

Alors les quatre assertions suivantes sont équivalentes :

- (i) X est une \mathcal{F}^Z -semi-martingale, et M^f est une martingale pour tout $f \in \mathcal{C}^2$
- (ii) \tilde{M}^f est une martingale pour tout $f \in \mathcal{C}^B$
- (iii) X est une \mathcal{F}^Z -semi-martingale et M^f est une martingale pour tout $f \in \hat{\mathcal{C}}$
- (iv) Z est une solution (faible) de l'équation différentielle stochastique suivante :

$$\begin{cases} dX_t = \sum_{j=1}^n \sigma_{1,j}(Z_t) dB_t^j + b_1(Z_t) dt + (1-c) dL_t^0 \\ dY_t^i = \sum_{j=1}^n \sigma_{i,j}(Z_t) dB_t^j + b_i(Z_t) dt, \quad \forall i = 2, \dots, n \end{cases} \quad (\mathbf{E})$$

dans le sens où X est une \mathcal{F}^Z -semi-martingale et il existe un mouvement brownien $B = (B^j)_{j \in \llbracket 1, n \rrbracket}$ sur \mathbb{R}^n (éventuellement défini sur un espace de probabilité élargi) satisfaisant l'équation (\mathbf{E}) .

Démonstration. Les implications $(\text{iii}) \Rightarrow (\text{i})$ et $(\text{iii}) \Rightarrow (\text{ii})$ sont immédiates, et l'implication $(\text{iv}) \Rightarrow (\text{i})$ découle de manière classique d'une application du lemme d'Itô à Z solution de (\mathbf{E}) , pour $f \in \mathcal{C}^2$. Nous allons donc montrer les implications $(\text{i}) \Rightarrow (\text{iii})$, $(\text{ii}) \Rightarrow (\text{i})$ et $(\text{i}) \Rightarrow (\text{iv})$.

Dans cette optique, commençons par remarquer que la définition de M^f peut être étendue au cas où X n'est pas une semi-martingale, à condition que f soit dans \mathcal{C}^B . Il suffit pour cela de poser (arbitrairement) $\int_0^t Bf(X_s) dL_s^0 = 0$, même si la mesure dL_s^0 n'est pas définie. On définit alors dans ce cas $M^f = \tilde{M}^f$. On observe également que \mathcal{C}^2 et \mathcal{C}^B sont deux anneaux, et en particulier sont stables par somme et produit. Sous l'hypothèse (i) (resp. (ii)), si $f, g \in \mathcal{C}^2$ (resp. \mathcal{C}^B), alors $M^f, M^g, M^{f^2}, M^{f+g}, M^{fg}$ et M^{g^2} sont donc des martingales. Calculons le crochet de M^f et M^g (ce calcul est inspiré d'un document de cours de Miclo [63]).

Soit $t \geq 0$. On pose

$$N_t = \int_0^t \mathcal{L}f(Z_s) ds + \int_0^t Bf(Z_s) dL_s^0$$

et on a alors

$$\begin{aligned} (f(Z_0) + M_t^f)^2 &= (f(Z_t) - N_t)^2 \\ &= f^2(Z_t) - 2f(Z_t)N_t + N_t^2 \end{aligned}$$

Par définition de M^{f^2} , on a

$$f^2(Z_t) = f^2(Z_0) + \int_0^t \mathcal{L}f^2(Z_s) ds + \int_0^t Bf^2(Z_s) dL_s^0 + M_t^{f^2}$$

et par la formule d'intégration par parties pour l'intégrale de Stieljes

$$\begin{aligned} N_t^2 &= 2 \int_0^t N_s dN_s \\ f(Z_t)N_t &= \int_0^t f(Z_s) dN_s + \int_0^t N_s df(Z_s) \\ &= \int_0^t f(Z_s) dN_s + \int_0^t N_s dM_s^f + \int_0^t N_s dN_s \end{aligned}$$

En remarquant que $Bf^2 = 2fBf$ (car f est continue), on obtient

$$(M_t^f)^2 = -2f(Z_0)M_t^f + \int_0^t \Gamma[f, f](Z_s) ds + M_t^{f^2} - 2 \int_0^t N_s dM_s^f$$

où Γ est l'opérateur carré du champ associé à \mathcal{L} , défini par $\Gamma[f, g] = \mathcal{L}[fg] - f\mathcal{L}g - g\mathcal{L}f$. Le processus

$$(M^f)^2 - \int_0^\cdot \Gamma[f, f](Z_s) ds$$

est donc une martingale, d'où l'on déduit que

$$\langle M^f, M^f \rangle_t = \int_0^t \Gamma[f, f](Z_s) ds$$

En appliquant ce résultat à M^g et M^{f+g} , et par définition du crochet on trouve

$$\begin{aligned} \langle M^f, M^g \rangle_t &= \frac{1}{2} (\langle M^f + M^g, M^f + M^g \rangle_t - \langle M^f, M^f \rangle_t - \langle M^g, M^g \rangle_t) \\ &= \frac{1}{2} (\langle M^{f+g}, M^{f+g} \rangle_t - \langle M^f, M^f \rangle_t - \langle M^g, M^g \rangle_t) \\ &= \frac{1}{2} \int_0^t (\Gamma[f+g, f+g](Z_s) - \Gamma[f, f](Z_s) - \Gamma[g, g](Z_s)) ds \\ &= \int_0^t \Gamma[f, g](Z_s) ds \end{aligned}$$

Une première conséquence utile de ce crochet s'obtient en considérant le cas $f = g = f_1 : \mathbb{R} \times \mathbb{R}^{n-1} \ni (x, y) \mapsto x$. Sous l'hypothèse (i), la fonction f_1 appartenant à \mathcal{C}^2 , le processus $X = f_1(Z)$ est bien une semi-martingale et ses temps locaux sont bien définis. Le crochet de X est donné par :

$$\begin{aligned} \langle X, X \rangle_t &= \langle M^{f_1}, M^{f_1} \rangle_t \\ &= \int_0^t \Gamma[f_1, f_1](Z_s) ds \\ &= \int_0^t a_{1,1}(Z_s) ds \end{aligned}$$

En utilisant les temps locaux $\{L_s^x(X), x \in \mathbb{R}, s \in \mathbb{R}_+\}$ de X et la formule des temps d'occupation (cf Revuz et Yor [75], chapitre VI, exercice 1.15), on a donc :

$$\int_0^t \mathbf{1}_{\{X_s=0\}} ds = \int_{\mathbb{R}} \frac{\mathbf{1}_{\{x=0\}}}{a_{1,1}(x, Y_s)} L_s^x(X) dx = 0$$

Le même résultat s'obtient sous l'hypothèse (ii) en utilisant $\tilde{f}_1 : \mathbb{R} \times \mathbb{R}^{n-1} \ni (x, y) \mapsto x\mathbf{1}_{\{x \geq 0\}} - \frac{1}{c}x\mathbf{1}_{\{x < 0\}}$ à la place de f_1 . Ceci justifie que le temps passé par X en 0 est de mesure

(de Lebesgue) nulle, et donc que le processus M^f n'est pas impacté par une modification des valeurs de $\mathcal{L}f(0, y)$, $y \in \mathbb{R}^{n-1}$ (puisque l'intégrale du générateur ne change pas). Celles-ci peuvent donc être choisies arbitrairement, tant que $\Gamma[f, f](0, y)$ est non-nul pour tout $y \in \mathbb{R}^{n-1}$, au moins pour certains $f \in \hat{\mathcal{C}}$ (par exemple, pour $f = f_1$ ou \tilde{f}_1 sous l'hypothèse (i) ou (ii) respectivement, mais d'autres choix sont également valables).

Le calcul du crochet va maintenant nous permettre de montrer les trois implications. Commençons par la plus simple.

Preuve de (i) \Rightarrow (iv). Pour tout $i = 1, \dots, n$, on définit la fonction f_i comme la i -ème application coordonnée sur \mathbb{R}^n , c'est-à-dire $f_i(x_1, \dots, x_n) = x_i$. Ces fonctions appartiennent à \mathcal{C}^2 et $\forall i, j \in \llbracket 1, n \rrbracket$, $\Gamma[f_i, f_j] = a_{i,j}$. Sous l'hypothèse (i), d'après le calcul du crochet qui précède et le théorème de représentation des semi-martingales (Ikeda et Watanabe [44], chapitre 2, théorème 7.1'), il existe un brownien n -dimensionnel $B = (B^j)_{j \in \llbracket 1, n \rrbracket}$, éventuellement défini sur un espace de probabilité élargi, tel que

$$M_t^{f_i} = \int_0^t \sigma_{i,j}(Z_s) dB_s^j, \quad \forall t \geq 0, \quad i = 1, \dots, n$$

Puisque de plus, $\mathcal{L}f_i = b_i$, $\forall i = 1, \dots, n$, on conclut que Z est solution de (E).

Preuve de (i) \Rightarrow (iii). On suppose (i) et on choisit une fonction $f \in \hat{\mathcal{C}}$. On pose

$$\begin{aligned} \beta_1(y) &= \partial_1 f(0^+, y) - \partial_1 f(0^-, y) \\ \beta_2(y) &= \partial_1^2 f(0^+, y) - \partial_1^2 f(0^-, y) \end{aligned}$$

pour tout $y \in \mathbb{R}^{n-1}$. Pour tout $\varepsilon > 0$, on définit successivement les fonctions

$$\begin{aligned} k^{(\varepsilon)} : \mathbb{R} \times \mathbb{R}^{n-1} &\rightarrow \mathbb{R} \\ (x, y) &\mapsto \beta_1(y) + x \left(\beta_2(y) + 6 \frac{\beta_1(y)}{\varepsilon} \right) \\ F^{(\varepsilon)} : \mathbb{R} \times \mathbb{R}^{n-1} &\rightarrow \mathbb{R} \\ (x, y) &\mapsto \frac{x(\varepsilon - x)^3}{\varepsilon^3} k^{(\varepsilon)}(x, y) \mathbf{1}_{\{0 < x < \varepsilon\}} \\ f^{(\varepsilon)} &= f + F^{(\varepsilon)} \end{aligned}$$

Pour tout $\varepsilon > 0$, la fonction $f^{(\varepsilon)}$ est évidemment \mathcal{C}^2 au voisinage de tout point $(x, y) \in \mathbb{R} \times \mathbb{R}^{n-1}$ tel que $x \notin \{0, \varepsilon\}$. De plus, de manière immédiate, $F^{(\varepsilon)}$ et toutes ses dérivées partielles à l'ordre 1 et 2 s'annulent en (ε, y) , pour tout $y \in \mathbb{R}^{n-1}$, c'est-à-dire :

$$F^{(\varepsilon)}(\varepsilon, y) = \partial_i F^{(\varepsilon)}(\varepsilon, y) = \partial_i \partial_j F^{(\varepsilon)}(\varepsilon, y) = 0$$

pour $i, j = 1, \dots, n$, ainsi que $F^{(\varepsilon)}$ et toutes ses dérivées partielles, à l'exception de celles selon la première variable, en $(0, y)$, $y \in \mathbb{R}^{n-1}$:

$$F^{(\varepsilon)}(0, y) = \partial_i F^{(\varepsilon)}(0, y) = \partial_i \partial_j F^{(\varepsilon)}(0, y) = 0$$

pour $i, j = 2, \dots, n$. Autrement dit, $F^{(\varepsilon)}$ est défini de telle manière que $f^{(\varepsilon)} \in \hat{\mathcal{C}}$. Calculons ses dérivées partielles selon la première variable. Pour tout $(x, y) \in]0, \varepsilon[\times \mathbb{R}^{n-1}$, on a

$$\begin{aligned} \partial_1 F^{(\varepsilon)}(x, y) &= \frac{(\varepsilon - x)^2}{\varepsilon^3} \left((4x - \varepsilon) k^{(\varepsilon)}(x, y) + x(x - \varepsilon) \partial_1 k^{(\varepsilon)}(x, y) \right) \\ \partial_1^2 F^{(\varepsilon)}(x, y) &= \frac{(\varepsilon - x)}{\varepsilon^3} \left(6(2x - \varepsilon) k^{(\varepsilon)}(x, y) + (x - \varepsilon)(4x - \varepsilon) \partial_1 k^{(\varepsilon)}(x, y) \right) \end{aligned}$$

donc en faisant tendre x vers 0 :

$$\begin{aligned}\partial_1 F^{(\varepsilon)}(0^+, y) &= -k^{(\varepsilon)}(0, y) = -\beta_1(y) \\ \partial_1^2 F^{(\varepsilon)}(0^+, y) &= \frac{6}{\varepsilon} k^{(\varepsilon)}(0, y) - \partial_1 k^{(\varepsilon)}(0, y) = -\beta_2(y) \\ \partial_{1,i} F^{(\varepsilon)}(0^+, y) &= \partial_{i,1} F^{(\varepsilon)}(0^+, y) = -\partial_i k^{(\varepsilon)}(0, y) = -\partial_{i-1} \beta_1(y), \quad i = 2, \dots, n\end{aligned}$$

Ce qui donne par exemple pour ∂_1 :

$$\begin{aligned}\partial_1 f^{(\varepsilon)}(0^+, y) &= \partial_1 f(0^+, y) - \beta_1(y) = \partial_1 f(0^-, y) \partial_1 f^{(\varepsilon)}(0^-, y) \\ \lim_{x \rightarrow 0^+} \frac{f^{(\varepsilon)}(x, y) - f_\varepsilon(0, y)}{y} &= \partial_1 f(0^+, y) - \beta_1(y) = \partial_1 f(0^\pm, y) \\ \lim_{x \rightarrow 0^-} \frac{f^{(\varepsilon)}(x, y) - f_\varepsilon(0, y)}{y} &= \partial_1 f(0^\pm, y)\end{aligned}$$

et de même pour les dérivées partielles d'ordre deux. On en déduit que toutes ces dérivées partielles sont bien définies et continues, et les fonction $f^{(\varepsilon)}$ sont donc dans \mathcal{C}^2 pour tout ε . Elles sont de plus bornées uniformément en ε , de même que toutes leurs dérivées partielles d'ordre 1 et 2 à l'exception de $\partial_1^2 f^{(\varepsilon)}$. Pour ce qui est de cette dernière, on peut tout de même isoler dans $\partial_1^2 F^{(\varepsilon)}$ la partie qui n'est pas bornée uniformément en ε , ce qui nous simplifiera les calculs par la suite. Pour tout $x \neq 0 \in \mathbb{R}$ et $y \in \mathbb{R}^{n-1}$ on note

$$\partial_1^2 F^{(\varepsilon)}(x, y) = 6\beta_1(y)Q(x, \varepsilon)\mathbf{1}_{\{0 < x < \varepsilon\}} + r_\varepsilon(x, y)$$

où $\sup_{x, \varepsilon} r_\varepsilon(x) < +\infty$ et Q est définie par :

$$Q(x, \varepsilon) = \frac{(x - \varepsilon)(2x - \varepsilon)}{\varepsilon^3} \begin{cases} \geq 0 & \text{si } 0 < x \leq \frac{\varepsilon}{2} \\ \leq 0 & \text{si } \frac{\varepsilon}{2} \leq x \leq \varepsilon \end{cases}$$

Pour tout $\varepsilon > 0$, le processus $M^{f^{(\varepsilon)}}$ est une martingale, et :

$$M_t^{f^{(\varepsilon)}} = f^{(\varepsilon)}(Z_t) - f^{(\varepsilon)}(Z_0) - \int_0^t (\mathcal{L}f(Z_s) + \mathcal{L}F^{(\varepsilon)}(Z_s))ds - \int_0^t Bf^{(\varepsilon)}(Z_s)dL_s^0(X)$$

avec $f^{(\varepsilon)}(Z_s) - f^{(\varepsilon)}(Z_0)$ qui est borné et converge vers $f(Z_s) - f(Z_0)$ quand $\varepsilon \rightarrow 0^+$, et $Bf^{(\varepsilon)}(0, Y_s) = \partial_1 f(0^-, Y_s)(1 - c)$. Il ne reste donc qu'à étudier la convergence de $\int_0^t \mathcal{L}F^{(\varepsilon)}(Z_s)ds$ lorsque $\varepsilon \rightarrow 0^+$.

Pour cela, on va utiliser les temps locaux de X , par l'intermédiaire de la formule des temps d'occupation. On rappelle que le crochet de X est donné par :

$$\langle X, X \rangle_t = \int_0^t a_{1,1}(Z_s)ds$$

donc

$$\begin{aligned}\int_0^t \mathcal{L}F^{(\varepsilon)}(Z_s)ds &= \int_0^t \frac{\mathcal{L}F^{(\varepsilon)}(Z_s)}{a_{1,1}(Z_s)} d\langle X, X \rangle_s \\ &= \int_{\mathbb{R}} \int_0^t \frac{\mathcal{L}F^{(\varepsilon)}(x, Y_s)}{a_{1,1}(x, Y_s)} dL_s^x(X) dx \\ &= \int_0^\varepsilon 6Q(x, \varepsilon) \int_0^t \beta_1(Y_s) dL_s^x(X) dx + \tilde{r}_\varepsilon(Z)\end{aligned}$$

où \tilde{r}_ε est un reste qui tend vers 0, et qui contient l'intégrale de la partie bornée r_ε de $\partial_1^2 F^{(\varepsilon)}$, ainsi que l'intégrale de toutes les autres dérivées partielles de $F^{(\varepsilon)}$. On note

$$\tilde{L}_t^x = 6 \int_0^t \beta_1(Y_s) dL_s^x(X)$$

ce qui donne

$$\lim_{\varepsilon \rightarrow 0^+} \int_0^t \mathcal{L}F^{(\varepsilon)}(Z_s) ds = \lim_{\varepsilon \rightarrow 0^+} \int_0^{\frac{\varepsilon}{2}} Q(x, \varepsilon) \tilde{L}_t^x dx + \lim_{\varepsilon \rightarrow 0^+} \int_{\frac{\varepsilon}{2}}^\varepsilon Q(x, \varepsilon) \tilde{L}_t^x dx$$

à condition que ces limites existent et que l'une au moins des deux limites du second membre soit finie. La fonction $x \mapsto Q(x, \varepsilon)$ est de signe constant dans chacune de ces deux intégrales, et on peut donc les encadrer comme suit :

$$\begin{aligned} \inf_{x \in [0, \frac{\varepsilon}{2}]} \tilde{L}_t^x \int_0^{\frac{\varepsilon}{2}} Q(x, \varepsilon) dx &\leq \int_0^{\frac{\varepsilon}{2}} Q(x, \varepsilon) \tilde{L}_t^x dx \leq \sup_{x \in [0, \frac{\varepsilon}{2}]} \tilde{L}_t^x \int_0^{\frac{\varepsilon}{2}} Q(x, \varepsilon) dx \\ \sup_{x \in [\frac{\varepsilon}{2}, \varepsilon]} \tilde{L}_t^x \int_{\frac{\varepsilon}{2}}^\varepsilon Q(x, \varepsilon) dx &\leq \int_{\frac{\varepsilon}{2}}^\varepsilon Q(x, \varepsilon) \tilde{L}_t^x dx \leq \inf_{x \in [\frac{\varepsilon}{2}, \varepsilon]} \tilde{L}_t^x \int_{\frac{\varepsilon}{2}}^\varepsilon Q(x, \varepsilon) dx \end{aligned}$$

Or :

$$\begin{aligned} \int_0^{\frac{\varepsilon}{2}} Q(x, \varepsilon) dx &= \frac{1}{3} \\ \int_{\frac{\varepsilon}{2}}^\varepsilon Q(x, \varepsilon) dx &= \frac{-1}{6} \end{aligned}$$

et $x \mapsto \tilde{L}_t^x$ est continue à droite, car l'application $(x, t) \mapsto L_t^x$ est continue à droite en x et croissante en t (cf Bray, [9], théorème 3). On a donc

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0^+} \int_0^t \mathcal{L}F^{(\varepsilon)}(Z_s) ds &= \lim_{\varepsilon \rightarrow 0^+} \frac{\tilde{L}_t^x}{3} - \lim_{\varepsilon \rightarrow 0^+} \frac{\tilde{L}_t^x}{6} \\ &= \frac{\tilde{L}_t^0}{6} \\ &= \int_0^t \beta_1(Y_s) dL_s^0 \end{aligned}$$

Maintenant, pour tout $t > s > 0$, la fonction aléatoire

$$\varepsilon \mapsto F^{(\varepsilon)}(Z_t) - f^{(\varepsilon)}(Z_s) - \int_s^t \mathcal{L}F^{(\varepsilon)}(Z_u) du$$

est bornée p.s., et donc par le théorème de convergence dominée

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0^+} M_t^{f^{(\varepsilon)}} &= f(Z_t) - f(Z_0) - \int_0^t \mathcal{L}f(Z_u) du - \int_0^t (\beta_1(Y_u) + \partial_1 f(0^-, Y_u)(1-c)) dL_u^0 \\ &= f(Z_t) - f(Z_0) - \int_0^t \mathcal{L}f(Z_u) du - \int_0^t Bf(0, Y_u) dL_u^0 \\ &= M_t^f \end{aligned}$$

définit bien une martingale $M^f = (M_t^f)_{t \geq 0}$.

Preuve de (ii)⇒(i). La preuve est identique à celle de l'implication précédente, à deux détails près : pour approcher une fonction $f \in \mathcal{C}^2$ par une famille de fonction $(f^{(\varepsilon)})_{\varepsilon>0} \subset \hat{\mathcal{C}}$, on n'a pas besoin de se soucier de la continuité des applications $x \mapsto \partial_1^2 f^{(\varepsilon)}(x, y)$ en 0, ce qui simplifie les calculs. En revanche, il faut être plus précautionneux pour montrer la convergence de $\int_0^\varepsilon \mathcal{L}F^{(\varepsilon)}(Z_s)ds$, car on ne peut pas ici utiliser directement les temps locaux de X sans avoir montré au préalable que ce dernier est une semi-martingale.

On suppose donc (ii) et on choisit une fonction $f \in \mathcal{C}^2$. On pose $\beta(y) = Bf(0, y) = \partial_1 f(0, y)(1 - c)$ pour tout $y \in \mathbb{R}^{n-1}$. Pour tout $\varepsilon > 0$, on définit la fonction

$$\begin{aligned} F^{(\varepsilon)} : \mathbb{R} \times \mathbb{R}^{n-1} &\rightarrow \mathbb{R} \\ (x, y) &\mapsto \frac{x(x - \varepsilon)^3}{\varepsilon^3} \beta(y) \mathbf{1}_{\{0 < x < \varepsilon\}} \end{aligned}$$

et on pose

$$f^{(\varepsilon)} = f + F^{(\varepsilon)}$$

Comme précédemment, on vérifie facilement que $f^{(\varepsilon)} \in \mathcal{C}^B$ pour tout $\varepsilon > 0$ et que $F^{(\varepsilon)}$ ainsi que toutes ses dérivées partielles d'ordre 1 et 2, à l'exception de $\partial_1^2 F^{(\varepsilon)}$, sont bornées, uniformément en ε . Pour ce qui est de $\partial_1^2 F^{(\varepsilon)}$, pour tout $x \neq 0 \in \mathbb{R}$ et $y \in \mathbb{R}^{n-1}$ on a

$$\begin{aligned} \partial_1^2 F^{(\varepsilon)}(x, y) &= \frac{6(x - \varepsilon)^2 + 6x(x - \varepsilon)}{\varepsilon^3} \beta(y) \mathbf{1}_{\{0 < x < \varepsilon\}} \\ &= \frac{6(x - \varepsilon)(2x - \varepsilon)}{\varepsilon^3} \beta(y) \mathbf{1}_{\{0 < x < \varepsilon\}} \\ &= 6\beta(y)Q(x, \varepsilon) \end{aligned}$$

avec

$$Q(x, \varepsilon) = \frac{(x - \varepsilon)(2x - \varepsilon)}{\varepsilon^3} \begin{cases} \geq 0 & \text{si } 0 < x \leq \frac{\varepsilon}{2} \\ \leq 0 & \text{si } \frac{\varepsilon}{2} \leq x \leq \varepsilon \end{cases}$$

Pour construire la famille des temps locaux de X , il faut commencer par montrer que celui-ci est une semi-martingale. Puisque $X = f_1(Z)$ et que la fonction f_1 n'appartient pas à \mathcal{C}^B , nous allons recourir temporairement au processus $\tilde{X} = \tilde{f}_1(Z)$, avec $\tilde{f}_1(x, y) = x\mathbf{1}_{\{x \geq 0\}} - \frac{1}{c}x\mathbf{1}_{\{x < 0\}}$. Ce processus, lui, est bien une semi-martingale puisque \tilde{f}_1 appartient à \mathcal{C}^B , et on va donc utiliser ses temps locaux $\{L_s^x(\tilde{X}), x \in \mathbb{R}, s \in \mathbb{R}_+\}$, de la même manière que ceux de X dans la preuve de l'implication précédente. Le crochet de \tilde{X} est donné par :

$$\begin{aligned} \langle \tilde{X}, \tilde{X} \rangle_t &= \langle M^{\tilde{f}_1}, M^{\tilde{f}_1} \rangle_t \\ &= \int_0^t \Gamma[\tilde{f}_1, \tilde{f}_1](X_s) ds \\ &= \int_0^t a_{1,1}(X_s) (\mathbf{1}_{\{X_s \geq 0\}} - \frac{1}{c} \mathbf{1}_{\{X_s < 0\}}) ds \end{aligned}$$

et \tilde{X} coïncide avec X sur \mathbb{R}_+ , donc en posant

$$\tilde{L}_t^x = 6 \int_0^t \beta(Y_s) dL_s^x(\tilde{X})$$

on trouve comme précédemment

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0^+} \int_0^t \mathcal{L}F^{(\varepsilon)}(Z_s) ds &= \lim_{\varepsilon \rightarrow 0^+} \int_0^{\frac{\varepsilon}{2}} Q(x, \varepsilon) \tilde{L}_t^x dx + \lim_{\varepsilon \rightarrow 0^+} \int_{\frac{\varepsilon}{2}}^\varepsilon Q(x, \varepsilon) \tilde{L}_t^x dx \\ &= \int_0^t \beta(Y_s) d\tilde{L}_s^0 \end{aligned}$$

Pour tout $t > s > 0$, la fonction aléatoire

$$\varepsilon \mapsto F^{(\varepsilon)}(Z_t) - f^{(\varepsilon)}(Z_s) - \int_s^t \mathcal{L}F^{(\varepsilon)}(Z_u)du$$

est bornée p.s., et donc pour tout $t \geq 0$ on a la convergence

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0^+} M_t^{f^{(\varepsilon)}} &= f(Z_t) - f(Z_0) - \int_0^t \mathcal{L}f(Z_s)ds - \int_0^t \beta(Y_s)d\tilde{L}_s^0 \\ &=: \tilde{M}_t^f \end{aligned}$$

et par le théorème de convergence dominée, le processus $\tilde{M}^f = (\tilde{M}_t^f)_{t \geq 0}$ est une martingale. En particulier, $f_1(Z) = X$ est une semi-martingale, dont on peut définir les temps locaux $\{L_t^x, (x, t) \in \mathbb{R} \times \mathbb{R}^+\}$, et en reprenant les calculs ci-dessus, on vérifie facilement que la mesure dL_t^x coïncide avec $d\tilde{L}_t^x$, pour tout $x \geq 0$, donc $\tilde{M}^f = M^f$. Ceci achève la preuve de l'implication. \square

Remarque 5.1.2. *Pour simplifier, et parce que nous n'aurons besoin que de ça ensuite, on a choisi ici de supposer que les trajectoires de Z sont bornées. Moyennant un argument de localisation sur des compacts, on peut bien sûr généraliser le résultat en s'affranchissant de cette hypothèse. D'autres généralisations sont également possibles sans trop d'efforts, pour montrer l'équivalence entre des équations différentielles stochastiques se rapprochant plus de celle étudiée par Takanobu [84] et le problème de sous-martingale de Anulova [4].*

De même, on a fait des hypothèses assez fortes sur α , mais on a vu dans la démonstration qu'il suffirait de supposer que $\alpha[f](y) \neq 0$ pour tout $y \in \mathbb{R}^{n-1}$, pour certaines fonctions $f \in \hat{\mathcal{C}}$ bien choisies.

5.2 Graphe à deux branches sans dérive

On reprend ici les notations du chapitre 4, et on considère le cas simple où $\mathbb{G} = [l^-, l^+]$, $l^- < 0 < l^+$, avec trois sommets en l^- , l^+ et 0 et deux arêtes $R_{e_1} = [l^-, 0]$ et $R_{e_2} = [0, l^+]$, et b est continue sur \mathbb{G} . On pose $\alpha = p(0, e_1) = 1 - p(0, e_2) \in]0, 1[$. On définit I comme l'ensemble des sous-intervalles fermés d'intérieur non vide de \mathbb{G} , que l'on considère comme un sous-ensemble de \mathbb{G}^2 :

$$I := \{(x, y) \in [l^-, l^+]^2 : x < y\}$$

et \bar{I} comme la fermeture topologique de I , c'est-à-dire l'ensemble des sous-intervalles fermés de \mathbb{G} :

$$\bar{I} = \{(x, y) \in [l^-, l^+]^2 : x \leq y\}$$

Pour tout $f \in \mathcal{F}$ et tout $(x, y) \in I$ tel que $x, y \neq 0$, on vérifie grâce au théorème 4.3.1 et à la réversibilité de μ (le calcul dans le cas général sera détaillé au chapitre 6) que le générateur \mathfrak{L} défini dans le chapitre précédent satisfait :

$$\mathfrak{L}[F_f](x, y) = \left((\mathfrak{d}_y - \mathfrak{d}_x)^2 - b(x)\mathfrak{d}_x - b(y)\mathfrak{d}_y + 2 \frac{\mu(x) + \mu(y)}{\mu([x, y])} (\mathfrak{d}_y - \mathfrak{d}_x) \right) [F_f](x, y)$$

où les notations \mathfrak{d}_x et \mathfrak{d}_y sont utilisées ici pour désigner les dérivées partielles usuelles sur $I \cap]l^-, l^+[$ par rapport à la première et à la deuxième variable respectivement, et 0 sur $I \setminus]l^-, l^+[$:

$$\mathfrak{d}_x = \mathbf{1}_{\{x \neq l^-\}} \partial_1, \quad \mathfrak{d}_y = \mathbf{1}_{\{y \neq l^+\}} \partial_2$$

On reconnaît donc un générateur de diffusion sur \mathbb{R}^2 , à la différence que les fonctions dans son domaine ne sont pas régulières. En effet les dérivées partielles des applications

F_f , $f \in \mathcal{F}$ ne sont généralement pas définies sur $\{0\} \times [l^-, l^+] \cup [l^-, l^+] \times \{0\}$, mais les applications partielles $G_f^y : x \mapsto F_f(x, y)$, $y \in [l^-, l^+]$ et $H_f^x : y \mapsto F_f(x, y)$, $x \in [l^-, l^+]$ satisfont en 0 la même condition de Neumann que les éléments de \mathcal{F} , avec les coefficients α et $(1 - \alpha)$ inversés. Plus explicitement, soit $y \in]0, l^+[$, $f \in \mathcal{F}$ et $g = G_f^y$, on a alors

$$g'(0^+) = \lim_{x \rightarrow 0^+} \frac{-1}{x} \left(\int_0^x f(t) d\mu(t) \right) = -f(0)\mu_{e_2}(0)$$

et d'après 4.18

$$g'(0^-) = -f(0)\mu_{e_1}(0) = -f(0)\frac{\alpha}{1-\alpha}\mu_{e_2}(0)$$

Le même résultat s'obtient pour la deuxième application partielle, et donc en utilisant la notation des dérivées directionnelles $g'(0, e_i)$, définies en (4.1), pour $g = G_f^y$ ou H_f^x , on a la relation suivante

$$(1 - \alpha)g'(0, e_1) + \alpha g'(0, e_2) = 0 \quad (5.1)$$

On veut alors étendre la définition de l'opérateur \mathfrak{L} à un domaine $\mathcal{D}(\mathfrak{L})$ contenant les fonctions F_f , et d'autres fonctions « régulières » sur I (en dehors des axes $x = 0$ et $y = 0$) dont les applications partielles satisfont la condition de Neumann (5.1) en 0.

Afin de pouvoir travailler avec la singularité en 0 sans avoir recours systématiquement aux dérivées directionnelles des applications partielles, nous utiliserons la notation introduite dans la section précédente, à savoir $f(a^\pm)$ pour la limite à droite ou à gauche d'une fonction f en $a \in \mathbb{R}$.

On définit alors $\mathcal{D}(\mathfrak{L})$ comme l'ensemble des fonctions $g : I \rightarrow \mathbb{R}$ continues telles que :

- $g|_{e_i \times e_j} \in \mathcal{C}^2(R_{e_i} \times R_{e_j})$, $i, j = 1, 2$ (où $g|_{e_i \times e_j}$ désigne la restriction de g à l'ensemble $\{(x, y) \in R_{e_i} \times R_{e_j} : x < y\}$)
- $\partial_1[g|_{e_i \times \mathbb{G}}]$, $\partial_1^2[g|_{e_i \times \mathbb{G}}]$, $\partial_2[g|_{\mathbb{G} \times e_i}]$ et $\partial_2^2[g|_{\mathbb{G} \times e_i}]$ sont bien définies et continues pour $i = 1, 2$.
- les dérivées partielles satisfont la condition de Neumann suivante :

$$\begin{cases} (1 - \alpha)\partial_1 g(0^-, y) - \alpha\partial_1 g(0^+, y) = 0, & \forall y \in [0, l^+] \\ (1 - \alpha)\partial_2 g(x, 0^-) - \alpha\partial_2 g(x, 0^+) = 0, & \forall x \in [l^-, 0] \end{cases}$$

Une explication sur la régularité des fonctions dans $\mathcal{D}(\mathfrak{L})$: ces fonctions et leurs dérivées partielles d'ordre 1 et 2 sont continues, à part au voisinage de l'axe $x = 0$ (resp. $y = 0$) lorsqu'on dérive par rapport à la première (resp. deuxième) variable. De chaque côté de ces axes, les dérivées partielles admettent quand même des limites.

On pose alors, pour tout $g \in \mathcal{D}(\mathfrak{L})$ et tout $(x, y) \in I$:

$$b^*(x, y) = 2 \frac{\mu(x) + \mu(y)}{\mu([x, y])}$$

$$\mathfrak{L}[g](x, y) := ((\mathfrak{d}_2 - \mathfrak{d}_1)^2 - b(x)\mathfrak{d}_1 - b(y)\mathfrak{d}_2 + b^*(x, y)(\mathfrak{d}_2 - \mathfrak{d}_1)) [g](x, y)$$

si $x, y \neq 0$ et :

$$\begin{aligned} \forall y \in]0, l^+], \quad \mathfrak{L}[g](0, y) &= \lim_{x \rightarrow 0^+} \mathfrak{L}[g](x, y) \\ \forall x \in [l^-, 0[, \quad \mathfrak{L}[g](x, 0) &= \lim_{y \rightarrow 0^-} \mathfrak{L}[g](x, y) \end{aligned}$$

et cette définition est consistante avec celle que l'on avait jusqu'à maintenant si $g = F_f$, $f \in \mathcal{F}$ ou plus généralement si $g = \mathfrak{F} \in \mathfrak{D}$.

Remarque 5.2.1. On rappelle que l'on a défini, dans le chapitre 4, la fonction $\mathfrak{L}f$ en D comme sa limite en D' lorsque D' tend vers D en croissant, d'où le choix de la limite à droite en x et à gauche en y .

Nous verrons dans la section 6.2 comment généraliser la définition du générateur et de son domaine sur \mathcal{G} lorsque le graphe est plus complexe.

Proposition 5.2.2. Pour toute mesure de probabilité $\mu_0^* \in \mathcal{P}(\bar{I})$, il existe un unique (en loi) processus continu $Z^* = (X_t^*, Y_t^*)_{t \geq 0}$, de loi initiale μ_0^* , tel que $Z_t^* \in I$ pour tout $t > 0$ et solution du problème de martingale pour $(\mathfrak{L}, \mathcal{D}(\mathfrak{L}))$.

Pour démontrer cette proposition, nous utiliserons le lemme suivant.

Lemme 5.2.3. Pour toute fonction $\tilde{b} : \mathbb{R}^2 \rightarrow \mathbb{R}$ mesurable bornée, les équations différentielles stochastiques suivantes

$$\begin{cases} dX_t^* = 2dB_t - b(X_t)dt + \tilde{b}(X_t^*, Y_t^*)dt - \frac{1}{\alpha}dL_t^0(X^*) \\ dY_t^* = -2dB_t - b(Y_t)dt - \tilde{b}(X_t^*, Y_t^*)dt \end{cases} \quad (\text{E1})$$

$$\begin{cases} dX_t^* = 2dB_t - b(X_t)dt + \tilde{b}(X_t^*, Y_t^*)dt \\ dY_t^* = -2dB_t - b(Y_t)dt - \tilde{b}(X_t^*, Y_t^*)dt - \frac{1}{\alpha}dL_t^0(Y^*) \end{cases} \quad (\text{E2})$$

(où $L_t^0(X^*)$ et $L_t^0(Y^*)$ désignent les temps locaux en 0 des semi-martingales X^* et Y^* respectivement) admettent une solution respective, unique en loi.

Démonstration. On montre le résultat pour (E1), la preuve pour (E2) étant symétrique.

On introduit l'équation suivante, qui n'est rien d'autre que l'équation (E1) sans le coefficient de dérive \tilde{b} :

$$\begin{cases} dX_t^* = 2dB_t - b(X_t)dt - \frac{1}{\alpha}dL_t^0(X^*) \\ dY_t^* = -2dB_t - b(Y_t)dt \end{cases} \quad (\text{E1}')$$

D'après Takanobu [84], cette EDS admet une unique solution $Z^* = (X^*, Y^*)$. Une simple application du théorème de Girsanov (cf Ikeda et Watanabe, chapitre 4.4) permet de conclure que Z^* est aussi l'unique solution de (E1), après un changement de mesure de probabilité bien choisi sur (Ω, \mathcal{A}) . \square

On observera ici l'importance d'avoir supposé que le coefficient de dérive du processus primal, b , est continu. En effet, dans les EDS considérées ici, le terme de diffusion est donné par un mouvement brownien unidimensionnel et le théorème de Girsanov ne permet donc de traiter que des coefficients de dérivées proportionnels dans les deux équations du système, comme c'est le cas avec les termes $\pm \tilde{b}(X_t^*, Y_t^*)$. On ne peut donc pas se ramener à un système de la forme (E1') sans les termes $b(X_t^*)$ et $b(Y_t^*)$, et ceux-ci doivent donc être continus pour rentrer dans le cadre du résultat d'unicité de la solution de Takanobu.

Démonstration de la proposition 5.2.2. Le lemme 5.2.3 ne permet pas de prouver directement le résultat désiré, car le coefficient de dérive b^* n'est pas borné, et chacune de deux EDS ne fait intervenir qu'un seul temps local, alors qu'il nous en faudrait deux. L'idée de ce qui suit est donc d'étendre le résultat de ce lemme en plusieurs étapes :

- (a) On commencera par construire le processus solution du problème de martingale sur des intervalles de temps où une seule des deux coordonnées peut passer en 0, et donc où un seul temps local intervient. En recollant ces intervalles, le processus est construit jusqu'au premier temps de sortie d'un sous-ensemble d'un compact de I .

- (b) On vérifie ensuite que le processus construit est bien l'unique solution du problème de martingale
- (c) Dans un troisième temps, en faisant tendre vers \bar{I} les ensemble sur lesquels le processus est construit, on montre que le temps de sortie de ces ensemble tend vers le premier moment où l'une des coordonnées atteint un bord de $[l^-, l^+]$, grâce à la proposition 4.4.4. On montre que le processus obtenu est bien solution du problème de martingale.
- (d) Enfin, la dernière étape consiste à étendre la condition initiale à \bar{I} . Cette étape est identique à Miclo [65] et ne sera pas détaillée.

Étape (a). Pour tout $n \in \mathbb{N}$, on définit le sous-ensemble I_n de I :

$$I_n = \{(x, y) \in]l^-, l^+[^2 : y - x \geq 1/n\}$$

Sur ces ensembles, le coefficient de dérive b^* est borné, et on définit la suite de fonctions sur I $(b_n)_{n \in \mathbb{N}}$ par :

$$\forall n \in \mathbb{N}, \quad b_n = b^* \wedge \sup_{(x,y) \in I_n} b^*(x, y)$$

On va alors construire par itération le processus $Z^* = (X^*, Y^*)$, unique solution du problème de martingale associé à $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$, jusqu'au premier temps de sortie de I_n , que l'on note τ_n :

$$\tau_n := \inf\{t \geq 0 : Y_t^* - X_t^* \leq 1/n \text{ ou } X_t^* = l^- \text{ ou } Y_t^* = l^+\}$$

On note P_z^1 et P_z^2 les lois des uniques solutions de (E1) et (E2) respectivement, partant de $z \in I$, avec $\tilde{b} = b_n$ (la dépendance de ces lois en n étant implicite). On pose $\tau_n^0 = 0$ et $Z_0^* = (x_0, y_0) \in I$. Par récurrence, on suppose le processus $Z^* = (X^*, Y^*)$ construit jusqu'au temps τ_n^{2i} , $i \in \mathbb{N}$, et on définit

$$(Z_{\tau_n^{2i}+t}^*)_{t \geq 0} = (X_{\tau_n^{2i}+t}^*, Y_{\tau_n^{2i}+t}^*)_{t \geq 0}$$

comme un processus sur \mathbb{R}^2 dont la loi conditionnelle sachant $\mathcal{F}_{\tau_n^{2i}}^{Z^*}$ est égale à $P_{Z_{\tau_n^{2i}}}^1$, jusqu'au temps :

$$\tau_n^{2i+1} := \inf\{t \geq \tau_n^{2i} : Y_t^* - X_t^* \leq 1/n \text{ ou } Y_t^* \leq 1/2n \text{ ou } X_t^* = l^- \text{ ou } Y_t^* = l^+\}$$

avec la convention $\inf \emptyset = +\infty$. On procède de même après τ_n^{2i+1} , en définissant

$$(Z_{\tau_n^{2i+1}+t}^*)_{t \geq 0} = (X_{\tau_n^{2i+1}+t}^*, Y_{\tau_n^{2i+1}+t}^*)_{t \geq 0}$$

comme un processus sur \mathbb{R}^2 , dont la loi conditionnelle sachant $\mathcal{F}_{\tau_n^{2i+1}}^{Z^*}$ est égale à $P_{Z_{\tau_n^{2i+1}}}^2$, jusqu'au temps :

$$\tau_n^{2i+2} := \inf\{t \geq \tau_n^{2i+1} : Y_t^* - X_t^* \leq 1/n \text{ ou } X_t^* \geq -1/2n \text{ ou } X_t^* = l^- \text{ ou } Y_t^* = l^+\}$$

En faisant tendre i vers $+\infty$, on a ainsi construit le processus $Z^* = (X^*, Y^*)$ jusqu'au temps τ_n .

Étape (b). Soit $f \in \mathcal{D}(\mathcal{L})$, on pose

$$\forall t > 0 \quad M_t^f = f(Z_t^*) - f(Z_0^*) - \int_0^t \mathcal{L}f(Z_s^*) ds$$

et on veut vérifier que le processus $(M_{t \wedge \tau_n}^f)_{t \geq 0}$ est une martingale. Soit $i \in 2\mathbb{N}$ un entier pair, alors par construction, pour tout $t \in [\tau_n^i, \tau_n^{i+1}]$ on a $Y_t^* \neq 0$. Il existe donc une fonction $\tilde{f} : I \rightarrow \mathbb{R}$, qui coïncide avec f et appartenant à l'ensemble \mathcal{C}^B défini dans la section 5.1, avec $c = \frac{1-\alpha}{\alpha}$ dans la définition de B (autrement dit \tilde{f} est égale à la fonction f , régularisée autour de $[l^-, l^+] \times \{0\}$). Par le théorème 5.1.1, $(M_{t \wedge \tau_n^{i+1}}^f - M_{t \wedge \tau_n^i}^f)_{t \geq 0}$ est donc une martingale. En traitant similairement le cas $i \in \mathbb{N}$ impair, on obtient que $(M_{t \wedge \tau_n}^f)_{t \geq 0}$ est une martingale locale. Comme elle est bornée, c'est une vraie martingale.

Par les mêmes arguments, si \tilde{Z}^* est un deuxième processus partant de z tel que le processus $(\tilde{M}_{t \wedge \tilde{\tau}_n}^f)_{t \geq 0}$ (avec des notations évidentes pour \tilde{M}^f et $\tilde{\tau}_n$) est une martingale pour tout $f \in \mathcal{D}(\mathfrak{L})$, alors la loi de \tilde{Z}^* sur l'intervalle $[\tilde{\tau}_n^i, \tilde{\tau}_n^{i+1}]$ coïncide avec celle de Z^* sur $[\tau_n^i, \tau_n^{i+1}]$, pour tout $i \in \mathbb{N}$, et ces deux processus ont donc la même loi jusqu'au temps τ_n .

Étape (c). En faisant tendre n vers $+\infty$, on étend la construction de Z^* jusqu'au temps $\tau := \lim_{n \rightarrow +\infty} \tau_n$. Par la proposition 4.4.4, on a :

$$\tau := \inf\{t \geq 0 : X_t^* = l^- \text{ ou } Y_t^* = l^+\}$$

et le processus $(M_{t \wedge \tau}^f)_{t \geq 0}$ est une martingale locale pour tout $f \in \mathcal{D}(\mathfrak{L})$. Il ne reste donc plus qu'à vérifier que c'est en fait une vraie martingale. Il suffit pour cela de vérifier que $\sup_{t \geq 0} M_t^f < \infty$ presque sûrement. La fonction f étant bornée ainsi que ses dérivées et la fonction μ , cela revient en fait à vérifier que le processus

$$t \mapsto \frac{1}{\mu(Z_t^*)}$$

est borné presque sûrement. Cette partie est identique à la preuve du lemme 11 de Miclo [65] et nous ne la reproduisons pas ici.

Enfin, à partir du temps τ , l'une des deux dérivées partielles (celle en x si $X_\tau^* = l^-$, celle en y si $Y_\tau^* = l^+$) « disparaît » dans l'expression du générateur, et celui-ci est alors un générateur d'une diffusion classique sur \mathbb{G} , absorbée en l^- si $Y_\tau = l^+$ et en l^+ si $X_\tau = l^-$. Le caractère bien-posé du problème de martingale est donc donné par Freidlin et Wentzell [40] (cf section 4.1).

Étape (d). Nous avons montré à ce stade que pour tout $z \in I$, il existe un processus Z^* partant de z , solution du problème de martingale associé à $(\mathfrak{L}, \mathcal{D})$, que la loi de Z^* est uniquement déterminée, et que $Z_t^* \in I$ pour tout $t > 0$. Il ne reste donc qu'à étendre ce résultat à $z \in \bar{I}$. Pour cette partie, il suffit de suivre la preuve de la proposition 12 de Miclo [65]. \square

Nous avons donc montré que le problème de martingale associé à $(\mathfrak{L}, \mathcal{D}(\mathfrak{L}))$ est bien posé, et puisque $\mathfrak{D} \subset \mathcal{D}(\mathfrak{L})$, celui associé à $(\mathfrak{L}, \mathfrak{D})$ admet au moins une solution. Pour montrer l'unicité de cette solution grâce à la précédente, il faudrait en fait montrer que \mathfrak{D} est bp-dense dans $\mathcal{D}(\mathfrak{L})$ (c'est-à-dire, pour tout $f \in \mathcal{D}(\mathfrak{L})$, il existe une suite $(f_n)_{n \in \mathbb{N}}$ de \mathfrak{D} telle que $f_n(x, y) \rightarrow f(x, y)$ pour tout $(x, y) \in I$ et $\sup_n \|f_n\|_\infty < \infty$, cf Ethier et Kurtz [29]), ce qui est loin d'être évident. Par exemple, si l'on considère le cas simple $\alpha = 1/2$, il semble difficile d'approcher la fonction $f_1 : (x, y) \rightarrow x$ par des éléments de \mathfrak{D} . Peut-être faudrait-il donc élargir ce dernier. Néanmoins, ce résultat reste satisfaisant puisque le résultat du lemme 4.4.3 est bien sûr toujours valable, et donc le processus Z^* obtenu est un dual de stationnarité forte pour X .

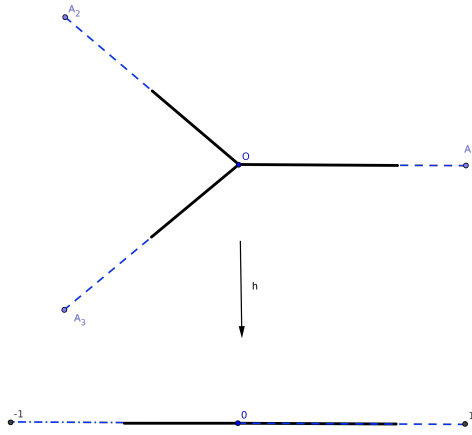


FIGURE 5.2 – Projection par h d'un état du dual sur le segment $[-1, 1]$

5.3 Adaptation au graphe en étoile

On généralise un petit peu en considérant le cas où \mathbb{G} est constitué d'un nœud central, et d'un nombre fini d'arêtes, toutes connectées à ce nœud et de longueur 1. Autrement dit, on a

$$V = \{O, A_1, \dots, A_n\}, \quad E = \{e_1, \dots, e_n\}, \quad n \in \mathbb{N}$$

avec $e_i = \{O, A_i\}$ et $l_{e_i} = 1$, $i = 1, \dots, n$, comme illustré en figure 5.1 avec $n = 5$. Observons notamment que dans ce cas la mesure invariante μ est réversible. On suppose pour simplifier que $b_{e_1} \equiv 0$ pour tout $i \in \llbracket 1, n \rrbracket$ (le processus primal est donc un mouvement brownien sur \mathbb{G}), et on définit les paramétrisations de telle sorte que $x_{e_i}(O) = 0$, $i = 1, \dots, n$, que $x_{e_1}(A_1) = 1$ et que $x_{e_i}(A_i) = -1$ pour tout $i \in \llbracket 2, n \rrbracket$. On ne fait en revanche aucune hypothèse sur les coefficients $p(O, e_i)$, $i = 1, \dots, n$, et la mesure invariante μ n'est donc pas globalement uniforme, mais seulement uniforme sur chaque arête, celles-ci ayant chacune un poids égal au coefficient correspondant, c'est-à-dire :

$$\forall i \in \llbracket 1, n \rrbracket, \quad \mu(R_{e_i}) = p(O, e_i)$$

On définit la fonction $\tilde{h} : \mathbb{G} \rightarrow [-1, 1]$ par :

$$\forall x \in \mathbb{G}, \quad \tilde{h}(x) := \sum_{i=1}^n x_{e_i}(x) \mathbf{1}_{x \in R_{e_i}}$$

Cette fonction n'est pas bijective, mais induit une fonction réciproque

$$\begin{aligned} \tilde{h}^{-1} : \mathcal{P}([-1, 1]) &\rightarrow \mathcal{P}(\mathbb{G}) \\ A &\mapsto \{x \in \mathbb{G} : \tilde{h}(x) \in A\} \end{aligned}$$

Comme précédemment, I désigne l'ensemble des intervalles fermés d'intérieur non-vides de $[-1, 1]$, et \mathcal{H} l'image de I par \tilde{h}^{-1} . On définit alors la fonction $h : \mathcal{H} \rightarrow I$ par :

$$\forall A \in \mathcal{H}, \quad h(A) = \{\tilde{h}(x), x \in A\}$$

Cette fonction est bijective, et $h^{-1} = \tilde{h}^{-1}$. Les éléments de \mathcal{H} sont donc les ensembles A de \mathbb{G} , tels que pour tout $i, j = 2, \dots, n$, $A \cap (R_{e_1} \cup R_{e_i})$ est connexe et $x_{e_i}(A \cap R_{e_i}) = x_{e_j}(A \cap R_{e_j})$. L'image d'un tel élément est l'intervalle $x_{e_1}(A \cap R_{e_1}) \cup x_{e_2}(A \cap R_{e_2})$. Cette projection du graphe sur un segment est illustrée par la figure 5.2

Enfin, on transporte les mesures μ et $\underline{\mu}$ sur $[-1, 1]$, et on les notera de la même manière, en posant :

$$\begin{aligned} \forall A \in I, \quad \mu(A) &= \mu(h^{-1}(A)) \\ \forall A \subset [-1, 1], \quad \underline{\mu}(A) &:= \underline{\mu}(\tilde{h}^{-1}(A)) \end{aligned}$$

Notre but est alors de construire un processus à valeurs dans \mathcal{H} , comme l'image d'un processus de I par h^{-1} .

Soient $(x, y) \in I$ et $f \in \mathcal{F}$. Comme dans la section précédente, on a :

$$\mathfrak{L}[F_f](h^{-1}(x, y)) = \left((\mathfrak{d}_2 - \mathfrak{d}_1)^2 + 2 \frac{\underline{\mu}(\{x, y\})}{\mu([x, y])} (\mathfrak{d}_2 - \mathfrak{d}_1) \right) [F_f \circ h^{-1}](x, y)$$

avec

$$\mathfrak{d}_1 = \mathbf{1}_{\{x \neq -1\}} \hat{\mathfrak{d}}_1, \quad \mathfrak{d}_2 = \mathbf{1}_{\{y \neq 1\}} \hat{\mathfrak{d}}_2$$

On définit donc le générateur $\tilde{\mathfrak{L}}$ sur I par

$$\tilde{\mathfrak{L}}[g](x, y) = \left((\mathfrak{d}_2 - \mathfrak{d}_1)^2 + 2 \frac{\underline{\mu}(\{x, y\})}{\mu([x, y])} (\mathfrak{d}_2 - \mathfrak{d}_1) \right) [g](x, y)$$

pour tout $(x, y) \in I$ tel que $x, y \neq 0$, et

$$\begin{aligned} \forall y \in]0, 1], \quad \tilde{\mathfrak{L}}[g](0, y) &= \lim_{x \rightarrow 0^+} \tilde{\mathfrak{L}}[g](x, y) \\ \forall x \in [-1, 0[, \quad \tilde{\mathfrak{L}}[g](x, 0) &= \lim_{y \rightarrow 0^-} \tilde{\mathfrak{L}}[g](x, y) \end{aligned}$$

pour tout g appartenant à l'ensemble $\mathcal{D}(\tilde{\mathfrak{L}})$ défini comme dans la section 5.2. On rappelle que g appartient à $\mathcal{D}(\tilde{\mathfrak{L}})$ si g est « suffisamment régulière » (notamment, \mathcal{C}^2 en-dehors des axes $x = 0$ et $y = 0$) et si ses dérivées partielles satisfont les conditions de Neumann suivantes :

$$\begin{aligned} \forall x < 0, \quad \mathfrak{p}(O, e_1) \hat{\partial}_2 g(x, O^-) &= (1 - \mathfrak{p}(O, e_1)) \hat{\partial}_2 g(x, 0^+) \\ \forall y > 0, \quad \mathfrak{p}(O, e_1) \hat{\partial}_1 g(0^-, y) &= (1 - \mathfrak{p}(O, e_1)) \hat{\partial}_1 g(0^+, y) \end{aligned}$$

Enfin on étend le générateur \mathfrak{L} , de la manière suivante : si $\mathcal{D}(\mathfrak{L})$ désigne l'ensemble des fonctions $f : \mathcal{H} \rightarrow \mathbb{R}$ telles que $f \circ h^{-1} \in \mathcal{D}(\tilde{\mathfrak{L}})$, alors on pose pour tout $f \in \mathcal{D}(\mathfrak{L})$ et tout $A \in \mathcal{H}$:

$$\mathfrak{L}[f](A) = \tilde{\mathfrak{L}}[f \circ h^{-1}](h(A))$$

On peut vérifier que $\mathfrak{D} \subset \mathcal{D}(\mathfrak{L})$ et que la définition de \mathfrak{L} sur $\mathcal{D}(\mathfrak{L})$ est cohérente avec celle sur \mathfrak{D} .

Proposition 5.3.1. *Pour tout $A \in \mathcal{H}$, il existe un unique (en loi) processus $D = (D_t)_{t \geq 0}$ solution du problème de martingale associé à $(\mathfrak{L}, \mathcal{D}(\mathfrak{L}))$ et tel que $D_0 = A$ presque sûrement.*

Démonstration. La preuve découle immédiatement du caractère bien posé du problème de martingale associé à $(\tilde{\mathfrak{L}}, \mathcal{D}(\tilde{\mathfrak{L}}))$, démontré dans la section 5.2. En effet, si $Z = (Z_t)_{t \geq 0}$ est associé à $\tilde{\mathfrak{L}}$ et $D = h^{-1}(Z)$, alors pour tout $f \in \mathcal{D}(\mathfrak{L})$ on a

$$\begin{aligned} \forall t \geq 0 \quad M_t^f &:= f(D_t) - f(D_0) - \int_0^t \mathfrak{L}[f](D_s) ds \\ &= f \circ h^{-1}(Z_t) - f \circ h^{-1}(Z_0) - \int_0^t \tilde{\mathfrak{L}}[f \circ h^{-1}](Z_s) ds \end{aligned}$$

et le processus $M^f = (M_t^f)_{t \geq 0}$ est une martingale. Réciproquement, si D est solution du problème de martingale associé à $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$, alors $Z = h(D)$ est solution de celui associé à $(\tilde{\mathcal{L}}, \mathcal{D}(\tilde{\mathcal{L}}))$, par le même argument. \square

Ce résultat n'est pas encore suffisant pour construire un dual de stationnarité pour toute condition initiale. En effet, suivant la méthode de Miclo, on peut faire partir le processus D d'une condition initiale A de mesure aussi petite que l'on veut, et en passant à la limite on peut autoriser D à partir d'un singleton $\{z_0\} \subset R_{e_1}$, ou d'une réunion de singletons $\bigcup_{i=2}^n \{x_{e_i}(z_0)\}$, $z_0 \in [-1, 0]$, mais pas d'un singleton de R_{e_i} , $i = 2, \dots, n$, autre que $\{O\}$. Ce problème est résolu en remarquant que tout ce qui a été introduit dans cette section dépend d'un choix de paramétrisation sur \mathbb{G} , qui est arbitraire. En échangeant successivement le rôle de A_i , $i = 1, \dots, n$, avec celui de A_1 , on définit les objets h_i et \mathcal{H}_i de la même manière que h et \mathcal{H} respectivement (en particulier, $h_1 = h$ et $\mathcal{H}_1 = \mathcal{H}$). Notons que pour tout $i \neq j \in \llbracket 1, n \rrbracket$, $\mathcal{H}_i \cap \mathcal{H}_j = \{O\}$. On peut ainsi définir le générateur \mathcal{L} sur $\bigcup_{i=1}^n \mathcal{H}_i$ par

$$\mathcal{L}[f](A) = \sum_{i=1}^n \mathbf{1}_{\{A \in \mathcal{H}_i\}} \tilde{\mathcal{L}}[f \circ h_i^{-1}](h_i(A))$$

pour tout $A \in \bigcup_{i=1}^n \mathcal{H}_i$ et tout f tel que $f \circ h_i^{-1} \in \mathcal{D}(\tilde{\mathcal{L}})$ pour tout $i \in \llbracket 1, n \rrbracket$. On obtient finalement le résultat suivant :

Corollaire 5.3.2. *Pour tout $\mu \in \mathcal{P}(\mathbb{G})$, il existe un couple (X, X^*) tel que X soit un processus de diffusion sur \mathbb{G} , de générateur \mathcal{L} et de loi initiale μ et que X^* soit l'unique solution du problème de martingale associé à $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ et un dual de stationnarité forte pour X .*

Démonstration. Le caractère bien posé du problème de martingale associé à $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ découle de ce qui précède. Le fait de pouvoir construire une solution X^* comme le dual d'un processus X de générateur \mathcal{L} découle du lemme 4.4.3 et se montre comme la proposition 23 de Miclo [65]. \square

Remarque 5.3.3. *Il n'y a pas de difficulté à ajouter une dérive b , à condition de supposer celle-ci continue sur \mathbb{G} , nulle en O et identique sur chaque arête (c'est-à-dire, $b_{e_i} \circ x_{e_i}^{-1} = b_{e_j} \circ x_{e_j}^{-1}$ dès que e_i et e_j sont paramétrées de la même manière, $i, j = 1, \dots, n$). En revanche, sans supposer la nullité en O la continuité ne pourra pas être assurée pour n'importe quel choix de paramétrisation, et on ne peut donc pas appliquer le résultat de la section 5.2 pour montrer que le problème de martingale est bien posé pour toute condition initiale. De même, si la dérive n'est pas identique sur chaque arête, il ne sera pas possible de construire le processus en bijection avec un processus sur I .*

Chapitre 6

Approche par des processus de Markov finis

Le générateur présenté dans la section précédente peut être vu localement comme le générateur d'un processus de diffusion multi-dimensionnel. Si $(D_t)_{t \geq 0}$ est un processus sur \mathcal{G} associé à ce générateur, alors toutes ses extrémités se comportent comme des diffusions sur les arêtes, dirigées par un même mouvement brownien, jusqu'à ce que l'une d'elles rencontre un obstacle (un sommet ou une autre extrémité). Le comportement du processus à ce moment-là n'est pas aisément descriptible et c'est pourquoi nous aurons recours dans cette section à une approximation par des chaînes de Markov. Cette approche a également l'avantage de pouvoir utiliser les résultats obtenus dans la première partie de cette thèse pour montrer directement l'entrelacement avec X , la solution du problème de martingale associé à \mathcal{L} . Pour cela, nous allons en premier lieu définir une suite de processus de saut $(X^n)_{n \in \mathbb{N}} := ((X_t^n)_{t \geq 0})_{n \in \mathbb{N}}$ convergeant en loi vers X .

Observons que contrairement aux approximations de diffusions que l'on trouve le plus souvent (pour une des constructions classiques du mouvement brownien par exemple), les chaînes de Markov utilisées sont ici à temps continu. Cette démarche se rapproche donc plus de celle que l'on retrouve dans les travaux de Kushner (voir par exemple Kushner et Dupuis [49]).

6.1 Approximation d'une diffusion

Une première difficulté dans cette démarche est liée à la taille des arêtes : si celles-ci avaient toutes la même longueur, on pourrait les diviser en une suite de segments de plus en plus petits et tous de même taille, de sorte que chaque X^n soit un processus de Markov fini avec une taille de saut fixe. C'est également possible si $\frac{l_e}{l_{e'}} \in \mathbb{Q}$ pour tout $e, e' \in E$, mais cela ne marche pas dans le cas général. Le lemme suivant permet de surmonter cette difficulté.

Lemme 6.1.1. *Il existe des familles $(h_n)_{n \in \mathbb{N}}$ et $(h_n^e)_{e \in E, n \in \mathbb{N}}$ de nombres réels strictement positifs telles que :*

(1) $\lim_{n \rightarrow \infty} h_n = 0$.

(2) *Il existe une constante $M > 0$ telle que pour tout $e \in E$, $n \in \mathbb{N}$, $|h_n^e - h_n| < M h_n^{2 + \frac{1}{|E|}}$.*

(3) *pour tout $e \in E$ et tout $n \in \mathbb{N}$, h_n^e divise l_e .*

Remarque 6.1.2. *La condition (2), implique notamment, pour tout $e \in E$, $h_n^e = h_n + o(h_n^2)$, au sens des notations de Landau. Le cardinal de $|E|$ étant fini, cela reste vrai si on remplace h_n^e par $\min_e h_n^e$ ou $\max_e h_n^e$.*

Démonstration. Fixons arbitrairement une arête de référence $e_0 \in E$. Le théorème d'approximation simultanée de Dirichlet (cf. [79], chapitre II, théorème 1B) permet d'approcher uniformément les ratios $\alpha_e := \frac{l_e}{l_{e_0}}$, $e \in E$, par des rationnels de même dénominateur. Plus précisément, il existe des familles $(q_n)_{n \in \mathbb{N}}$ et $(p_n^e)_{e \in E, n \in \mathbb{N}}$ d'entiers strictement positifs telles que $\lim_{n \rightarrow \infty} q_n = +\infty$ et, pour tout $e \in E$ et $n \in \mathbb{N}$:

$$\left| \alpha_e - \frac{p_n^e}{q_n} \right| < \frac{1}{q_n^{1 + \frac{1}{|E|}}} \quad (6.1)$$

En multipliant de chaque coté par $\frac{l_{e_0}}{p_n^e}$, cette inégalité s'exprime de manière équivalente par :

$$\left| \frac{l_e}{p_n^e} - \frac{l_{e_0}}{q_n} \right| < \frac{l_{e_0}}{p_n^e q_n^{1 + \frac{1}{|E|}}} \quad (6.2)$$

D'autre part, de (6.1) on déduit une minoration sur p_n^e :

$$p_n^e > q_n \alpha_e - \frac{1}{q_n^{\frac{1}{|E|}}}$$

Comme $q_n \rightarrow \infty$, le minorant est non nul pour tout $e \in E$ à partir d'un certain rang n_0 (on supposera sans perte de généralité que $n_0 = 0$), et on l'injecte dans (6.2) pour obtenir :

$$\begin{aligned} \left| \frac{l_e}{p_n^e} - \frac{l_{e_0}}{q_n} \right| &< \frac{l_{e_0}}{\left(q_n \alpha_e - \frac{1}{q_n^{\frac{1}{|E|}}} \right) q_n^{1 + \frac{1}{|E|}}} \\ &= \left(\frac{l_{e_0}}{q_n} \right)^{2 + \frac{1}{|E|}} \left(\frac{1}{l_{e_0}^{1 + \frac{1}{|E|}} \left(\alpha_e - \frac{1}{q_n^{1 + \frac{1}{|E|}}} \right)} \right) \end{aligned}$$

On pose pour tout $n \in \mathbb{N}$, $e \in E$, $h_n := \frac{l_{e_0}}{q_n}$ et $h_n^e := \frac{l_e}{p_n^e}$. On a bien $h_n \rightarrow 0$ et h_n^e divise l_e , pour tout $e \in E$, et en choisissant :

$$M := \sup_{e \in E, n \in \mathbb{N}} \left(\frac{1}{l_{e_0}^{1 + \frac{1}{|E|}} \left(\alpha_e - \frac{1}{q_n^{1 + \frac{1}{|E|}}} \right)} \right) < \infty$$

dans la condition (2), on conclut la démonstration du lemme. \square

On choisit donc des suites $(h_n)_{n \in \mathbb{N}}$ et $(h_n^e)_{e \in E, n \in \mathbb{N}}$ satisfaisant les conditions du lemme 6.1.1. Pour chaque $n \in \mathbb{N}$, on pose :

$$\begin{aligned} V_n &:= \bigcup_{e \in E} \left\{ x_e^{-1}(k h_n^e), k = 0, \dots, \frac{l_e}{h_n^e} \right\} \\ E_n &:= \bigcup_{e \in E} \left\{ \{x_e^{-1}(k h_n^e), x_e^{-1}((k+1) h_n^e)\}, k = 0, \dots, \frac{l_e}{h_n^e} - 1 \right\} \end{aligned}$$

de sorte que $G_n := (V_n, E_n)$ est un graphe discret dont les sommets sont les sommets de \mathbb{G} et des points régulièrement espacés sur chaque arête. On définit un noyau de probabilité

$p_n := (p_n(x, y))_{x, y \in V_n}$ de V_n dans V_n par :

$$\forall x, y \in V_n \quad p_n(x, y) := \begin{cases} p(x, e) & \text{si } e \in E, x \in V, x, y \in R_e \text{ et } \{x, y\} \in E_n \\ 1/2 & \text{si } x \notin V \text{ et } \{x, y\} \in E_n \\ 0 & \text{sinon} \end{cases}$$

Enfin, on définit une suite $(\mathcal{L}^n)_{n \in \mathbb{N}}$ de générateurs de processus de Markov finis sur \mathbb{G} , où pour tout $n \in \mathbb{N}$, \mathcal{L}^n a pour espace d'état V_n et :

$$\forall e \in E, \forall x \neq y \in V_n \cap R_e, \quad \mathcal{L}_{x,y}^n = \begin{cases} \frac{2p_n(x, y)}{h_n \int_{[x,y]} e^{\beta_e(t,x)} \lambda_e(dt)}, & \text{si } \{x, y\} \in E_n \\ 0 & \text{sinon} \end{cases} \quad (6.3)$$

où $(\beta_e)_{e \in E}$ est la famille de fonctions définie en début de section 4.2. L'expression de ce générateur n'étant pas facilement manipulable, le lemme suivant en donne une approximation suffisante pour montrer qu'il converge vers le générateur d'une diffusion sur \mathbb{G} . Le terme $\mathcal{L}_{x,y}^n$ dépendant à la fois de x, y et n (x et y dépendant eux-mêmes de n à travers l'hypothèse $x, y \in V_n$), dans le lemme suivant on utilisera la notation $o_n(\cdot)$ dans le sens suivant :

$$f(x, y, n) = o_n(g(n)) \quad \text{si et seulement si} \quad \lim_{n \rightarrow \infty} \frac{\sup_{x,y} |f(x, y, n)|}{g(n)} = 0$$

avec le même abus d'utilisation du symbole « = » que pour les notations de Landau. Toujours dans le but d'alléger les calculs, lorsque $x \in R_e$, $e \in E$, si le contexte est clair on identifiera x et $x_e(x)$. On notera également :

$$\forall h \in \mathbb{R}, \quad x \pm h\vec{e} := i_{\vec{e}} x_e^{-1}(x_e(x) \pm h)$$

dès lors que $x_e(x) \pm h \in [0, l_e]$. Si $x \in \overset{\circ}{R}_e$ et $\vec{e} = (e^-, e^+)$, on abrégera cette notation en $x \pm h$.

Lemme 6.1.3. *Pour tout $n \in \mathbb{N}$, $\vec{e} \in \vec{E}$, $x \in R_e \cap V_n$ et $y = x + h_n^e \vec{e}$, on a :*

$$\mathcal{L}_{x,y}^n = p_n(x, y) \left(\frac{2}{h_n^2} + i_{\vec{e}} \frac{b_e(x)}{h_n} + o_n \left(\frac{1}{h_n} \right) \right)$$

Démonstration. Soit $\vec{e} \in \vec{E}$, $n \in \mathbb{N}$ et $k \in \llbracket 0, \frac{l_e}{h_n^e} - 1 \rrbracket$. On suppose que $\vec{e} = (e^-, e^+)$, l'autre cas se traitant similairement, et on pose $x = x_e^{-1}(k h_n^e)$ et $y = x + h_n^e \vec{e} = x_e^{-1}((k+1)h_n^e)$. On a alors :

$$\mathcal{L}_{x,y}^n = \frac{2p_n(x, y)}{h_n f_x(y)}$$

où $f_x : [0, l_e] \rightarrow \mathbb{R}$ est la fonction définie par :

$$\forall z \in [0, l_e], \quad f_x(z) = \int_x^z e^{\beta_e(t,x)} dt$$

avec ici :

$$\beta_e(t, x) = - \int_x^t b_e(u) du$$

Dans le but d'utiliser la formule de Taylor-Lagrange à l'ordre 2 en x , on calcule les dérivées de cette fonction :

$$\begin{aligned} f'_x(z) &= e^{\beta_e(z,x)} \\ f''_x(z) &= -b_e(z)f'_x(z) \\ f_x^{(3)}(z) &= \left(-\frac{d}{dz}b_e(z) + b_e(z)^2 \right) f'_x(z) \end{aligned}$$

On a donc, par un développement limité à l'ordre 2 en x :

$$f_x(z) = (z-x) \left(1 - b_e(x) \frac{(z-x)}{2} + f_x^{(3)}(x) \frac{(z-x)^2}{6} + o((z-x)^2) \right)$$

(ici, le o est pris pour $z \rightarrow x$). Si on pose $g_x(z) = \frac{z-x}{f_x(z)}$, $z \in [0, l_e]$, alors en utilisant le développement limité de la fonction $t \rightarrow \frac{1}{1+t}$, on obtient :

$$g_x(z) = 1 + b_e(x) \frac{z-x}{2} + \left(\frac{b_e(x)^2}{4} - \frac{f_x^{(3)}(x)}{6} \right) (z-x)^2 + o((z-x)^2)$$

La fonction g_x est donc de classe \mathcal{C}^2 et sa dérivée seconde :

$$g_x''(z) = \frac{b_e(x)^2}{2} - \frac{f_x^{(3)}(x)}{3}$$

est bornée. Pour tout $z \in [0, l_e]$ il existe donc $\xi \in [x, z]$ tel que :

$$g_x(z) = 1 + b_e(x) \frac{z-x}{2} + g_x''(\xi) \frac{(z-x)^2}{2}$$

En appliquant ceci à $z = x_e(y)$, on en déduit qu'il existe $\xi \in x_e([x, y])$ tel que :

$$\begin{aligned} \mathcal{L}_{x,y}^n &= \frac{2p_n(x,y)}{h_n h_n^e} \left(1 + b_e(x) \frac{h_n^e}{2} + g''(\xi) \frac{h_n^{e2}}{2} \right) \\ &= \frac{p_n(x,y)}{h_n^2} (2 + b_e(x) h_n) + o_n \left(\frac{1}{h_n} \right) \end{aligned}$$

où on a utilisé le fait que $\frac{1}{h_n^e} - \frac{1}{h_n} = o_n(1)$. □

Dans le but d'obtenir des résultats de convergence de la suite $(\mathcal{L}^n)_{n \in \mathbb{N}}$ vers \mathcal{L} , on prolonge chacun de ces \mathcal{L}^n en un générateur sur \mathbb{G} , que l'on notera de la même manière. Pour cela, à chaque étape on va simplement identifier chaque point $z \in \mathbb{G}$ avec le point $x \in V_n$ dont il est le plus proche. Plus formellement, pour tout $n \in \mathbb{N}$ on définit une partition $\{I_n(x), x \in V_n\}$ de \mathbb{G} par :

$$\forall x \in V_n, \quad I_n(x) := \bigcup_{e \in E} \left\{ z \in R_e : x_e(x) - x_e(z) \in \left[-\frac{h_n^e}{2}, \frac{h_n^e}{2} \right] \right\}$$

et pour tout $x \in V_n$, pour tout $z \in I_n(x)$, on pose :

$$\forall y \neq z \in \mathbb{G}, \quad \mathcal{L}_{z,y}^n = \begin{cases} \mathcal{L}_{x,y}^n & \text{si } y \in V_n \setminus \{x\} \\ 0 & \text{sinon} \end{cases}$$

Soit μ_0 une mesure de probabilité sur \mathbb{G} . Pour tout $n \in \mathbb{N}$, on définit une version discrète μ_0^n de μ_0 par :

$$\mu_0^n := \sum_{x \in V_n} \mu_0(I_n(x)) \delta_x$$

et $X^n := (X_t^n)_{t \geq 0}$ un processus de loi initiale μ_0^n et de générateur \mathcal{L}^n . Alors

$$\mathbb{P}(X_t^n \in V_n, \forall t > 0) = 1$$

et on a :

Lemme 6.1.4. *La famille $(X^n)_{n \in \mathbb{N}}$ est relativement compacte pour la topologie de la convergence faible. Plus précisément, la famille $(\mathcal{L}(X^n))$ est relativement compacte dans $\mathcal{P}(D_{\mathbb{G}}[0, \infty))$ muni de la distance de Prokhorov (voir introduction).*

Démonstration. La preuve de ce lemme consiste à vérifier les deux conditions suivantes (cf [29], chapitre 3, lemme 7.4) :

(a) Pour tout $\eta > 0$ et $t > 0$, il existe un compact $\Gamma_{\eta,t} \subset \mathbb{G}$ tel que :

$$\liminf_{n \rightarrow \infty} \mathbb{P}(X_t^n \in \Gamma_{\eta,t}^\eta) \geq 1 - \eta$$

où $\Gamma_{\eta,t}^\eta$ désigne le η -élargi de $\Gamma_{\eta,t}$ (c'est-à-dire, l'ensemble $\{x \in \mathbb{G} : d(x, \Gamma_{\eta,t}) < \eta\}$)

(b) Pour tout $\eta > 0$ et $T > 0$, il existe $\delta > 0$ tel que :

$$\limsup_{n \rightarrow \infty} \mathbb{P}(w(X^n, \delta, T) \geq \eta) \leq \eta$$

On rappelle la définition de w (cf [29], chapitre 3, (6.2)) : pour tout $x \in D_{\mathbb{G}}(0, \infty)$ (l'espace des fonctions $[0, \infty) \rightarrow \mathbb{G}$ càdlàg), et tout $\delta, T > 0$,

$$w(x, \delta, T) := \inf_{\{t_i\}} \max_i \sup_{s, t \in [t_i, t_{i+1}[} \{d(x(t), x(s))\}$$

où les $\{t_i\}$ sont pris parmi les partitions de la forme $0 = t_0 < t_1 < \dots < t_{k-1} < T \leq t_k$ avec $\min(t_i - t_{i-1}) > \delta$ et $k \geq 1$.

La condition (a) est trivialement satisfaite puisque \mathbb{G} est compacte, et nous passons donc directement à la condition (b), qui est un peu plus technique. Pour cela on va montrer plus précisément :

$$\forall \eta, T > 0, \quad \lim_{\delta \rightarrow 0} \lim_{n \rightarrow \infty} \mathbb{P}(w(X^n, \delta, T) \leq \eta) = 1 \quad (6.4)$$

Dans ce qui suit, on fixe $n \in \mathbb{N}$ et $\delta, T, \eta > 0$. Dire qu'un $\omega \in \Omega$ appartient à l'événement $\{w(X^n, \delta, T) \leq \eta\}$ signifie qu'il existe une partition $\{t_i, i = 1, \dots, k\}$ de $[0, T]$ (ou d'un intervalle un peu plus grand), dont le pas est plus grand que δ , et pour laquelle l'amplitude maximale de la trajectoire $X^n(\omega)$ sur chaque intervalle $[t_i, t_{i+1}[$ est plus petite que η . Pour un processus de saut dont la taille des sauts est plus grande que η , cela correspond par exemple à l'ensemble des trajectoires pour lesquelles les temps de sauts sont plus grands que δ (jusqu'au temps T).

On va donc minorer grossièrement la probabilité de l'événement $\{w(X^n, \delta, T) \leq \eta\}$ en choisissant une partition particulière de $[0, T]$. On définit par récurrence une suite de temps d'arrêts $(T_k)_{k \in \mathbb{N}}$, et une variable aléatoire K à valeurs entières, par $T_0 = 0$ et :

— Pour tout $k \in \mathbb{N}$, $T_{k+1} := \inf \{t > T_k, d(X_{T_k}^n, X_t^n) > \frac{\eta}{2}\}$

— $K := \inf \{k \in \mathbb{N}, T_k \geq T\}$.

Par l'inégalité triangulaire, on a :

$$\forall k \in \mathbb{N}, \quad \sup_{T_k \leq t, s < T_{k+1}} \{d(X_t^n, X_s^n)\} \leq \eta$$

donc une condition suffisante (mais non nécessaire) pour $w(X^n, \delta, T) \leq \eta$ est alors que $T_i - T_{i-1} > \delta$ pour tout $i = 1, \dots, K$. Remarquons que dans ce cas on a nécessairement $K \leq T/\delta$.

On pose

$$\varepsilon := \min(\eta/4, \min\{l_e, e \in E\})$$

Pour parcourir une distance $\eta/2$ (*i.e.* atteindre un point situé à une distance au moins $\eta/2$ de sa position initiale) entre les temps T_i et T_{i+1} , $i \in \mathbb{N}$, le processus X^n , doit alors parcourir une distance au moins égale à ε sur une même arête. En effet, si on suppose le contraire, alors en particulier il ne parcourt pas une arête entière, donc $X_{T_{i+1}}$ et X_{T_i} se trouvent soit sur une même arête, soit sur deux arêtes voisines. Dans les deux cas, la distance entre $X_{T_{i+1}}$ et X_{T_i} est strictement inférieure à deux fois la distance maximale parcourue sur une arête, elle-même inférieure à $\eta/4$, ce qui contredit la définition des T_k .

On encadre grossièrement les taux de sauts de X^n sur chaque arête en posant :

$$\begin{aligned} \forall e \in E, \quad M^e &:= \frac{1}{h_n \int_0^{h_n^e} e^{-Bt} dt} \geq \max_{\substack{y, z \in R_e \\ \{y, z\} \in E_n}} \mathcal{L}_{y, z}^n \\ \forall e \in E, \quad m^e &:= \frac{1}{h_n \int_0^{h_n^e} e^{Bt} dt} \leq \min_{\substack{y, z \in R_e \\ \{y, z\} \in E_n}} \mathcal{L}_{y, z}^n \end{aligned}$$

avec $B := \max_{e \in E} \sup_{t \in R_e} |b_e(t)|$. Pour tout $e \in E$, soit Z^e un processus de vie et de mort sur \mathbb{N} , partant de 0, de taux de naissance $b^0 = 2M^e$ et $b^n := M^e$, $\forall n \in \mathbb{N}^*$ et de taux de décès $a^n := m^e$, $\forall n \in \mathbb{N}^*$, ($m_0 = 0$). Heuristiquement, sur chaque arête R_e , $e \in E$, X^n s'éloigne plus lentement de sa position initiale que Z^e de 0 (en loi). Partant de $X_{T_i} \in R_e$, on peut alors coupler $(X_{T_i+t}^n)_{t \geq 0}$ avec une copie de chaque $Z^{e'}$, $e' \in E$, toutes indépendantes entre elles, de manière à contrôler le déplacement sur chaque arête. Si on définit, pour tout $e \in E$:

$$\sigma_e^\varepsilon := \inf\{t > 0, Z_t^e h_n^e \geq \varepsilon\}$$

alors un tel couplage donne :

$$\begin{aligned} \mathbb{P}(T_{i+1} - T_i > \delta \mid X_{T_i}, X_{T_i} \in R_e) &\geq \mathbb{P}(\sigma_e^\varepsilon > \delta)^2 \prod_{e' \neq e \in E} \mathbb{P}(\sigma_{e'}^\varepsilon > \delta) \\ &\geq \left(\min_{e' \in E} \mathbb{P}(\sigma_{e'}^\varepsilon > \delta)\right)^{|E|} \end{aligned}$$

Avec ce qui précède, on peut donc écrire :

$$\begin{aligned} \mathbb{P}(w(X^n, \delta, T) \leq \eta) &\geq \mathbb{P}(T_i - T_{i-1} > \delta, \forall i = 1, \dots, K) \\ &\geq \left(\min_{e \in E} \mathbb{P}(\sigma_e^\varepsilon > \delta)\right)^{|E|T/\delta} \end{aligned} \quad (6.5)$$

Pour montrer (6.4), on va alors minorer $\mathbb{P}(\sigma_e^\varepsilon > \delta)$ de manière appropriée pour tout $e \in E$.

Fixant $e \in E$, posons $N = \left\lceil \frac{\varepsilon}{h_n^e} \right\rceil$, de sorte que pour tout $t \geq 0$, $Z_t^e h_n^e \geq \varepsilon$ si et seulement si $Z_t^e \geq N$. σ_e^ε est donc le temps d'atteinte de N . Par l'inégalité de Markov, on a :

$$\begin{aligned} \mathbb{P}(\sigma_e^\varepsilon > \delta) &= 1 - \mathbb{P}(\sigma_e^\varepsilon \leq \delta) \\ &= 1 - \mathbb{P}(\exp(-\sigma_e^\varepsilon/\delta) \geq e^{-1}) \\ &\geq 1 - e \mathbb{E}(\exp(-\sigma_e^\varepsilon/\delta)) \end{aligned} \quad (6.6)$$

Pour calculer $\mathbb{E}(\exp(-\sigma_e^\varepsilon/\delta))$, on s'inspire de la méthode utilisée par exemple dans [3] pour calculer $\mathbb{E}(\sigma_e^\varepsilon)$. On pose $\tau_0 = 0$, et pour tout $k = 1, \dots, N$, on définit par récurrence

$$\tau_k := \inf\{t > 0, Z_t^e = k\} - \tau_{k-1},$$

le premier temps de passage de l'état $k - 1$ à k . Tous les τ_i sont indépendants (par markovianité) et :

$$\begin{aligned} \mathbb{E}(\exp(-\sigma_e^\varepsilon/\delta)) &= \mathbb{E}\left(\exp\left(-\sum_{i=1}^N \tau_i/\delta\right)\right) \\ &= \prod_{i=1}^N \mathbb{E}(\exp(-\tau_i/\delta)) \end{aligned}$$

On rappelle que si Y est une variable aléatoire de loi exponentielle de paramètre $\theta > 0$, alors :

$$\mathbb{E}(\exp(-Y/\delta)) = \frac{\theta}{1/\delta + \theta}$$

donc, posant $v_i = \mathbb{E}(\exp(-\tau_i/\delta))$, $i = 1, \dots, N$ on a :

$$v_1 = \frac{2M^e}{1/\delta + 2M^e}$$

(e étant fixé, dans toute la suite on omettra l'exposant) et pour $i > 1$, si Y est une variable aléatoire de loi exponentielle de paramètre $m + M$, indépendante des τ_i :

$$\begin{aligned} v_i &= \frac{M}{m + M} \mathbb{E}(\exp(-Y/\delta)) + \frac{m}{m + M} \mathbb{E}\left(\exp\left(-\frac{Y + \tau_{i-1} + \tau_i}{\delta}\right)\right) \\ &= \frac{M}{m + M} \frac{m + M}{1/\delta + m + M} + \frac{m}{m + M} \frac{m + M}{1/\delta + m + M} v_{i-1} v_i \\ &= \frac{M}{1/\delta + m + M} + \frac{m}{1/\delta + m + M} v_{i-1} v_i \end{aligned}$$

et donc :

$$v_i = \frac{M/m}{\frac{1/\delta + m + M}{m} - v_{i-1}}$$

Pour étudier la suite $(v_i)_i$, on pose $\alpha := m/M$, $\beta := \frac{1/\delta + m + M}{m}$, $w_1 = 1$, $w_2 = \beta - v_1$ et pour tout $i > 2$, $w_i = \beta w_{i-1} - \alpha w_{i-2}$. On vérifie alors par récurrence que l'on a $v_i = \frac{\alpha w_{i-1}}{w_i}$, $i \geq 2$, et donc :

$$\begin{aligned} \mathbb{E}(\exp(-\sigma_e^\varepsilon/\delta)) &= v_1 \prod_{i=2}^N \frac{\alpha w_{i-1}}{w_i} \\ &= \frac{v_1 \alpha^{N-1}}{w_N} \end{aligned}$$

Une étude classique de la suite vectorielle $((w_{i+1}, w_i))_i$ donne :

$$w_N = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} \beta & -\alpha \\ 1 & 0 \end{pmatrix}^{N-2} \begin{pmatrix} w_2 \\ w_1 \end{pmatrix}$$

Si on note :

$$\lambda_{\pm} := \frac{\beta \pm \sqrt{\beta^2 - 4\alpha}}{2}$$

les deux valeurs propres λ_+ et λ_- de la matrice ci-dessus alors on a :

$$w_{N+1} = \frac{(\beta - v_1) (\lambda_+^N - \lambda_-^N) + \lambda_-^N \lambda_+ - \lambda_+^N \lambda_-}{\lambda_+ - \lambda_-}$$

Intéressons-nous maintenant au comportement asymptotique de cet objet lorsque l'on fait $n \rightarrow +\infty$. Rappelons que :

$$M = \frac{1}{h_n \int_0^{h_n^e} e^{-Bt} dt} = \frac{1}{h_n (1 - e^{-B h_n^e})} \quad \text{et} \quad m = \frac{1}{h_n \int_0^{h_n^e} e^{Bt} dt} = \frac{1}{h_n (e^{B h_n^e} - 1)}$$

donc :

$$\begin{aligned} \alpha &= \frac{M}{m} = e^{B h_n^e} = 1 + B h_n + o(h_n) & \text{et} & \quad \beta = 1 + \alpha + \frac{1}{\delta m} \\ & & & = 1 + e^{B h_n^e} + \frac{h_n (e^{B h_n^e} - 1)}{\delta B} \\ & & & = 2 + B h_n + \left(\frac{B^2}{2} + \frac{1}{\delta} \right) h_n^2 + o(h_n^2) \end{aligned}$$

Selon les besoins, on utilisera alternativement les formes exactes ou les développements limités de α et β . En utilisant les formules usuelles de développements limités, on trouve également :

$$\begin{aligned} v_1 &= \frac{2M}{1/\delta + 2M} = 1 + o(h_n) \\ \beta^2 - 4\alpha &= (1 - \alpha)^2 + \frac{1}{\delta^2 m^2} + 2 \frac{1 + \alpha}{\delta m} \\ &= (1 - e^{B h_n^e})^2 + \frac{h_n^2 (e^{B h_n^e} - 1)^2}{\delta^2 B^2} + 2 \frac{(e^{2B h_n^e} - 1) h_n^e}{\delta B} \\ &= (B^2 + 4/\delta) h_n^2 (1 + 2B h_n + o(h_n)) \\ \sqrt{\beta^2 - 4\alpha} &= \sqrt{B^2 + 4/\delta} h_n (1 + B h_n) + o(h_n^2) \\ \lambda_- &= 1 + \left(B - \sqrt{B^2 + 4/\delta} \right) h_n/2 + o(h_n) \\ \lambda_+ &= 1 + \left(B + \sqrt{B^2 + 4/\delta} \right) h_n/2 + o(h_n) \\ \lambda_-^N &= \exp(N \ln(\lambda_-)) \\ &= \exp \left(N \left(\left(B - \sqrt{B^2 + 4/\delta} \right) h_n/2 + o(h_n) \right) \right) \\ &= e^{\frac{NBh_n}{2}} e^{-\frac{N\sqrt{B^2+4/\delta}}{2}} e^{No(h_n)} \\ \lambda_+^N &= e^{\frac{NBh_n}{2}} e^{\frac{N\sqrt{B^2+4/\delta}}{2}} e^{No(h_n)} \end{aligned}$$

puis :

$$\lambda_+^N + \lambda_-^N = 2e^{\frac{NBh_n}{2} + No(h_n)} \cosh\left(\frac{N\sqrt{B^2 + 4/\delta}}{2}\right)$$

$$\lambda_+^N - \lambda_-^N = 2e^{\frac{NBh_n}{2} + No(h_n)} \sinh\left(\frac{N\sqrt{B^2 + 4/\delta}}{2}\right)$$

$$\lambda_-^N \lambda_+ - \lambda_+^N \lambda_- = (\lambda_+^N + \lambda_-^N) \sqrt{B^2 + 4/\delta} h_n/2 + (\lambda_+^N - \lambda_-^N) (1 + Bh_n/2) + o(h_n)$$

De plus, puisque $N = \left\lceil \frac{\varepsilon}{h_n^\varepsilon} \right\rceil$, on a $Nh_n = \varepsilon + o(1)$. Mettant tout ceci bout-à-bout, on obtient :

$$\begin{aligned} w_{N+1} &= \frac{Bh_n/2 (\lambda_+^N - \lambda_-^N) - \sqrt{B^2 + 4/\delta} (\lambda_+^N + \lambda_-^N) h_n/2}{\sqrt{B^2 + 4/\delta} h_n} + o(1) \\ &= \frac{e^{\varepsilon B/2} \left(B \sinh\left(\varepsilon \frac{\sqrt{B^2 + 4/\delta}}{2}\right) + \sqrt{B^2 + 4/\delta} \cosh\left(\varepsilon \frac{\sqrt{B^2 + 4/\delta}}{2}\right) \right)}{\sqrt{B^2 + 4/\delta}} + o(1) \\ \alpha^N &= e^{\varepsilon B} \end{aligned}$$

et donc finalement :

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{E}(\exp(-\sigma_e^\varepsilon/\delta)) &= \frac{e^{\varepsilon B/2} \sqrt{B^2 + 4/\delta}}{B \sinh\left(\varepsilon \frac{\sqrt{B^2 + 4/\delta}}{2}\right) + \sqrt{B^2 + 4/\delta} \cosh\left(\varepsilon \frac{\sqrt{B^2 + 4/\delta}}{2}\right)} \\ &\leq e^{\varepsilon B/2} \exp\left(-\varepsilon \frac{\sqrt{B^2 + 4/\delta}}{2}\right) \\ &= o(\delta) \end{aligned}$$

Par (6.6) on conclut que :

$$\mathbb{P}(\sigma_e^\varepsilon > \delta)^{|E|T/\delta} \geq (1 + o(\delta))^{|E|T/\delta} \xrightarrow{\delta \rightarrow 0} 1$$

et donc par (6.5) il existe $\delta > 0$ tel que $\mathbb{P}(w(X^n, \delta, T) \geq \eta) \leq \eta$. \square

On peut maintenant montrer la convergence des solutions des problèmes de martingale associés aux \mathcal{L}^n .

Proposition 6.1.5. *La suite de processus $(X^n)_{n \in \mathbb{N}}$ converge en loi vers un processus X de générateur \mathcal{L} et de loi initiale μ_0 .*

Démonstration. D'après le lemme 6.1.4, la famille $(X^n)_{n \in \mathbb{N}}$ est relativement compacte et admet donc au moins une sous-suite $X^{\varphi(n)}$ convergeant vers un processus $X := (X_t)_{t \geq 0}$. Puisque μ_0^n converge faiblement vers μ_0 , la loi initiale de X est nécessairement donnée par μ_0 . Si $\mathcal{D}(\mathcal{L})$ désigne l'ensemble des fonctions $f \in \mathcal{F}$ telles que $\mathcal{L}f \in \mathcal{C}(\mathbb{G})$, il suffit alors de vérifier que pour tout $f \in \mathcal{D}(\mathcal{L})$, il existe une suite $(f_n)_{n \in \mathbb{N}}$ de fonctions sur \mathbb{G} telle que :

$$\|f_n - f\|_\infty \rightarrow 0 \quad \text{et} \quad \|\mathcal{L}^n f_n - \mathcal{L}f\|_\infty \rightarrow 0 \tag{6.7}$$

pour pouvoir appliquer le lemme 5.1, chapitre 4 de [29], et déduire que X à pour générateur \mathcal{L} . Comme la solution du problème de martingale pour $(\mathcal{L}, \mathcal{D}(\mathcal{L}))$ est unique en loi (cf [40]), on en conclut alors que $(X^n)_{n \in \mathbb{N}}$ converge en loi vers X (unicité de la valeur d'adhérence)

Il ne reste donc qu'à trouver une suite $(f_n)_{n \in \mathbb{N}}$ satisfaisant (6.7). On prend pour cela :

$$\forall n \in \mathbb{N}, x \in V_n, z \in I_n(x), \quad f_n(z) = f(x)$$

La fonction f étant continue sur un compact, elle est uniformément continue et donc $(f_n)_{n \in \mathbb{N}}$ converge uniformément vers f . Montrons que $\mathcal{L}^n f_n$ converge uniformément vers $\mathcal{L}f$. Soit $z \in \mathbb{G}$ et pour tout $n \in \mathbb{N}$, $x_n(z) \in V_n$ tel que $z \in I_n(x_n(z))$. La suite de fonction $(x_n)_{n \in \mathbb{N}}$ tend uniformément vers la fonction identité de \mathbb{G} , donc d'après le lemme 6.1.3, la condition de Neumann sur f et la continuité uniforme de $\mathcal{L}f$ sur \mathbb{G} , on a (en notant simplement x_n pour $x_n(z)$) :

$$\begin{aligned} \mathcal{L}^n f_n(z) &= \sum_{\vec{e} \in \vec{E}_{x_n}} \mathcal{L}_{x_n, x_n + h_n^e \vec{e}}^n (f(x_n + h_n^e \vec{e}) - f(x_n)) \\ &= \sum_{\vec{e} \in \vec{E}_x} \frac{p_n(x_n, x_n + h_n^e \vec{e})}{h_n^2} (2 + i_{\vec{e}} b_e(x_n) h_n) \left(f'(x_n, \vec{e}) h_n + \partial_e^2 f(x_n) \frac{h_n^2}{2} \right) + o_n(1) \\ &= \sum_{\vec{e} \in \vec{E}_{x_n}} p_n(x_n, x_n + h_n^e \vec{e}) \left(\frac{2f'(x_n, \vec{e})}{h_n} + b_e(x_n) \partial_e f(x_n) + \partial_e^2 f(x_n) \right) + o_n(1) \\ &= \sum_{\vec{e} \in \vec{E}_{x_n}} p_n(x_n, x_n + h_n^e \vec{e}) (b_e(x_n) \partial_e f(x_n) + \partial_e^2 f(x_n)) + o_n(1) \\ &= \mathcal{L}f(x_n) + o_n(1) \\ &= \mathcal{L}f(z) + o_n(1) \end{aligned}$$

qui est exactement le résultat attendu. □

Remarque 6.1.6. *Ce résultat s'applique en fait à n'importe quel famille de générateur $(\tilde{\mathcal{L}}^n)_{n \in \mathbb{N}}$ présentant le même développement asymptotique que celui de la famille $(\mathcal{L}^n)_{n \in \mathbb{N}}$ donné dans le lemme 6.1.3. L'intérêt de ce choix spécifique tient en fait dans la proximité entre ces générateurs et le générateur κ présenté dans la section 4.2, qui permet d'assurer la convergence des lois stationnaires discrètes vers la loi stationnaire du processus limite, dans la proposition suivante.*

Pour tout $n \in \mathbb{N}$, soit μ^n la loi stationnaire de X^n . On notera de la même manière la fonction $\mu^n : \mathbb{G} \rightarrow [0, 1]$ définie par :

$$\begin{aligned} \forall x \in V_n \setminus V, z \in I_n(x), \quad \mu^n(z) &:= \mu^n(\{x\}) \\ \forall x \in V, \vec{e} \in \vec{E}_x, z \in I_n(x) \cap \mathring{R}_e, \quad \mu^n(z) &:= \mu^n(\{x + h_n^e \vec{e}\}) \\ \forall x \in V, \quad \mu^n(x) &= \mu^n(\{x\}) \end{aligned}$$

Proposition 6.1.7. *La suite de mesure $(\mu^n)_{n \in \mathbb{N}}$ converge faiblement vers μ , la mesure stationnaire de \mathcal{L} . Plus précisément, on a pour la suite de fonctions correspondantes :*

$$\begin{aligned} \lim_{n \rightarrow \infty} \sup_{x \in \mathbb{G} \setminus V} \left| \frac{\mu^n(x)}{h_n} - \mu(x) \right| &= 0 \\ \lim_{n \rightarrow \infty} \sup_{x \in V, e \in E_x} \left| \frac{2p(x, e) \mu^n(x)}{h_n} - \mu_e(x) \right| &= 0 \end{aligned}$$

Remarque 6.1.8. La deuxième limite implique notamment :

$$\begin{aligned}\underline{\mu}(x) &= \frac{1}{2} \sum_{e \in E_x} \mu_e(x) \\ &= \sum_{e \in E_x} \frac{p(x, e) \mu^n(x)}{h_n} + o_n(1) \\ &= \frac{\mu^n(x)}{h_n} + o_n(1)\end{aligned}$$

La fonction $\frac{\mu^n}{h_n}$ converge uniformément vers $\underline{\mu}$, qui est discontinue en chaque point de V (sauf sur les sommets $x \in V$ de degré 2 et tels que $p(x, e) = 1/2$ pour $e \in E_x$). C'est la raison pour laquelle on n'a pas défini $\mu^n(z) = \mu^n\{x\}$ pour $z \in I_n(x)$, $x \in V$: cette fonction convergerait simplement mais pas uniformément sur $\mathbb{G} \setminus V$.

Démonstration. L'astuce dans cette preuve consiste à remarquer que toute diffusion sur \mathbb{G} , définie à partir de la structure de graphe de G , peut être alternativement définie à partir de n'importe quel G_n . En effet, si pour tout $e \in E$ et tout $x, y \in R_e$ tels que $\{x, y\} \in E_n$, on définit un opérateur $\mathcal{L}_{\{x, y\}}$ sur $\mathcal{C}^\infty([x, y])$ par :

$$\forall f \in \mathcal{C}^\infty([x, y]), \quad \mathcal{L}_{\{x, y\}} f = (\mathcal{L}_e \tilde{f})|_{[x, y]}$$

où $\tilde{f} \in \mathcal{C}^\infty(R_e)$ est un prolongement arbitraire de f sur R_e , alors on a :

$$\forall f \in \mathcal{F}, \forall z \in \mathbb{G}, \quad \mathcal{L}[f](z) = \begin{cases} \mathcal{L}_{\{x, y\}}[f|_{[x, y]}](z), & \text{si } z \in]x, y[, \{x, y\} \in E_n \\ \sum_{\substack{e=\{x, z\} \\ e \in E_n}} p_n(z, x) \mathcal{L}_e[f|_e](z), & \text{si } z \in V_n \end{cases}$$

De plus, toute fonction $f \in \mathcal{F}$ satisfait la condition de Neumann en chaque sommet $x \in V_n$, pour la famille $(p(x, y))_{y \in V_n}$, et sa restriction sur chaque $[x, y]$ est évidemment \mathcal{C}^∞ . On a donc une définition alternative de L , et donc de la mesure invariante associée. Pour tout $n \in \mathbb{N}$, on définit alors le même générateur que dans la section 4.2, mais sur V_n au lieu de V :

$$\forall e \in E, \forall x \neq y \in V_n \cap R_e, \quad \kappa_n(x, y) := \begin{cases} \frac{p_n(x, y)}{\int_{[x, y]} e^{\beta_e(t, x)} \lambda_e(dt)}, & \text{si } \{x, y\} \in E_n \\ 0 & \text{sinon} \end{cases}$$

On remarque que $\kappa_n = \frac{h_n}{2} \mathcal{L}^n$, donc κ_n et \mathcal{L}^n ont la même mesure invariante μ^n . On obtient donc l'équation (4.18), où la mesure α est remplacé par μ^n et où le terme $p_n(x, e) = 1/2$ est incorporé dans la constante Z_n , qui ne dépend que de $n \in \mathbb{N}$:

$$\forall e \in E, x \in \mathring{R}_e \cap V_n, \quad \mu_e(x) = \frac{\mu^n(x)}{Z_n} \quad (6.8)$$

$$\forall x \in V, e \in E_x, \quad \mu_e(x) = \frac{2p(x, e) \mu^n(x)}{Z_n} \quad (6.9)$$

donc $\lim_{n \rightarrow \infty} \mu^n(x) = 0$ pour tout $x \in \mathbb{G} \setminus V$, et :

$$\begin{aligned}\lim_{n \rightarrow \infty} Z_n \sum_{e \in E} \sum_{x \in \mathring{R}_e \cap V_n} \mu_e(x) &= \lim_{n \rightarrow \infty} \sum_{x \in \mathbb{G} \setminus V} \mu^n(x) \\ &= 1 - \lim_{n \rightarrow \infty} \sum_{x \in V} \mu^n(x) \\ &= 1\end{aligned}$$

Par la convergence des sommes de Riemann, on a aussi :

$$\begin{aligned} \lim_{n \rightarrow \infty} h_n \sum_{e \in E} \sum_{x \in \mathring{R}_e \cap V_n} \mu_e(x) &= \lim_{n \rightarrow \infty} \sum_{e \in E} \sum_{x \in \mathring{R}_e \cap V_n} \mu_e(x) h_n^e \\ &= \int_{\mathcal{G} \setminus V} d\mu \\ &= 1 \end{aligned}$$

En faisant la différence de ces deux limites, on obtient :

$$\lim_{n \rightarrow \infty} (Z_n - h_n) \sum_{e \in E} \sum_{x \in \mathring{R}_e \cap V_n} \mu_e(x) = 0$$

Or $\mu_e(x)$ est borné en $e \in E$ et en $x \in R_e$, donc on a :

$$\sum_{e \in E} \sum_{x \in \mathring{R}_e \cap V_n} \mu_e(x) = O(|V_n|) = O\left(\frac{1}{h_n}\right)$$

d'où l'on conclut :

$$Z_n = h_n + o(h_n)$$

Soit $z \in \mathring{R}_e$, $e \in E$. Pour tout $n \in \mathbb{N}$ on pose $x_n \in V_n$ tel que $z \in I_n(x_n)$. La continuité uniforme de μ_e et l'égalité précédente, injectée dans (6.8), donne alors :

$$\begin{aligned} \mu(z) &= \mu_e(x_n) + o_n(1) \\ &= \frac{\mu^n(x_n)}{h_n} + o_n(1) \\ &= \frac{\mu^n(z)}{h_n} + o_n(1) \end{aligned}$$

et la deuxième limite est immédiate en utilisant (6.9). □

6.2 Approximation du dual d'une diffusion

On passe maintenant à la construction d'un dual discret convergeant vers un processus de générateur \mathfrak{L} . La définition de ce dernier sur les fonctionnelles de la forme F_f , $f \in \mathcal{F}$, bien que pratique à manipuler, présente malheureusement l'inconvénient de l'absence de continuité de $\mathfrak{L}[F_f]$, pour toute fonction $f \in \mathcal{F}$ non identiquement nulle. Cette absence de continuité serait toujours vraie si on avait muni \mathcal{G} de la topologie induite par la distance de Hausdorff, et empêche d'utiliser le lemme 5.1, chapitre 4 de [29] comme dans la section précédente. Nous allons donc modifier un peu le domaine du générateur, et pour cela définir quelques notions supplémentaires.

6.2.1 Redéfinition du générateur dual

Dans le chapitre 5, on a exprimé une fonction sur un sous-ensemble de \mathcal{G} comme fonction des bornes de ce sous-ensemble, ce qui a permis de donner une expression du générateur en terme de dérivées partielles. Ceci n'a été possible que parce que dans ce cas le nombre de bornes du processus était décroissant (deux jusqu'au moment où l'une des deux est absorbée, puis une, puis zéro), ce qui n'est pas le cas en général. Nous allons donc dans cette section définir une forme de dérivées partielles pour une fonction sur \mathcal{G} , que l'on utilise ensuite pour donner une nouvelle expression du générateur.

Pour tout $D \in \mathcal{G}$, on reprend la notation $\mathfrak{d}D$ définie dans la remarque 4.3.2, et on note $\mathfrak{d}^c D := \mathfrak{d}\overline{D}^c$ l'ensemble des couples $(x, \vec{e}) \in \mathbb{G} \times \vec{E}$ où $x \in \partial D$ et \vec{e} est une arête sortante de D en x . Remarquons que $\mathfrak{d}D$ (et donc $\mathfrak{d}^c D$ aussi) caractérise D de manière unique (si $\mathfrak{d}D = \emptyset$ alors $D = \mathbb{G}$), et on notera \mathcal{A} l'ensemble des $A \subset \mathbb{G} \times \vec{E}$ finis, tels que $\mathfrak{d}D = A$ pour un certain $D \in \mathcal{G}$ (forcément unique), que l'on notera $D(A)$. Si $A \in \mathcal{A}$, alors il existe également $D \in \mathcal{G}$ tel que $\mathfrak{d}^c D = A$, et on le notera $D^c(A)$ (on remarque que si $A \neq \emptyset$, on a $D^c(A) = \overline{D(A)}^c$). On notera qu'il existe une bijection naturelle entre $\mathfrak{d}D$ et $\mathfrak{d}^c D$ si $D \in \mathcal{D}$, mais que ces deux ensembles ne sont pas nécessairement de même cardinal si $D \in \mathcal{G} \setminus \mathcal{D}$: en un point $x \in \partial D \cap V$, il n'y a pas nécessairement le même nombre d'arêtes entrantes et sortantes, et chaque arête $e \in E_x$ correspond soit à une arête (orientée) entrante soit à une arête sortante. Dans la suite, on notera $\mathfrak{d}^{(c)} D$ et $D^{(c)}(A)$ pour indiquer que les définitions introduites sont valables que l'on considère, au choix, $\mathfrak{d}D$ et $D(A)$, ou $\mathfrak{d}^c D$ et $D^c(A)$.

On munit l'ensemble des parties finies de $\mathbb{G} \times \vec{E}$ d'une topologie induisant la notion de convergence suivante : pour tout $A \in \mathcal{P}(\mathbb{G} \times \vec{E})$ fini, une suite $(A_n)_{n \in \mathbb{N}} \subset \mathcal{P}(\mathbb{G} \times \vec{E})$ converge vers A si et seulement si il existe $n_0 \in \mathbb{N}$ tel que $|A| = |A_n|$ pour tout $n \geq n_0$, et pour tout $y = (x, \vec{e}) \in A$ il existe $(x_n)_{n \in \mathbb{N}}$ avec $(x_n, \vec{e}) \in A_n$, $\forall n \geq n_0$ et $x_n \rightarrow x$. Notamment, les notions de fermeture et de continuité sont définies par cette convergence. Une telle topologie peut par exemple être induite par la distance qui à deux parties finies A et B de $\mathbb{G} \times \vec{E}$ associe :

- 1 si il existe $\vec{e} \in \vec{E}$ tel que $|\{(x, \vec{e}) \in A\}| \neq |\{(x, \vec{e}) \in B\}|$
- $\min(1, \inf \|\{(d(x_i, y_i))_{i=1, \dots, |A|}\}\|)$, où $\|\cdot\|$ est la norme euclidienne sur $\mathbb{R}^{|A|}$ et où l'infimum est pris sur toutes les indexations possibles $(x_i, \vec{e}_i)_{i=1, \dots, |A|}$ et $(y_i, \vec{e}_i)_{i=1, \dots, |A|}$ des ensembles A et B .

Soit $\bar{\mathcal{A}}$ la fermeture de \mathcal{A} dans $\mathcal{P}(\mathbb{G} \times \vec{E})$. Il semble important à cette étape de donner des précisions sur cet ensemble. Soit $A := \{(x_i, \vec{e}_i), i = 1, \dots, n\} \in \bar{\mathcal{A}} \setminus \mathcal{A}$ alors il existe au moins un indice $i \in \llbracket 1, n \rrbracket$ tel que l'une des trois conditions suivantes soit remplie :

- $x_i \in V$ et $-\vec{e}_i \in \vec{E}_{x_i}$ (c'est-à-dire qu'il existe $y \in V$ tel que $(y, x_i) = \vec{e}_i$).
- il existe $j \in \llbracket 1, n \rrbracket$ tel que $x_j = x_i$ et $\vec{e}_i = -\vec{e}_j$.
- $x_i \in V$ et pour tout $\vec{e} \in \vec{E}_{x_i}$ il existe $j \in \llbracket 1, n \rrbracket$ tel que $\vec{e}_j = \vec{e}$.

Réciproquement, si l'une de ces conditions est remplie pour un certain $A = \{(x_i, \vec{e}_i), i = 1, \dots, n\} \subset \mathbb{G} \times \vec{E}$ et un certain indice $i = 1, \dots, n$, alors $A \notin \mathcal{A}$. En revanche, cela n'implique pas nécessairement $A \in \bar{\mathcal{A}}$.

Il est relativement aisé de vérifier que l'application $A \in \mathcal{A} \mapsto D^{(c)}(A) \in \mathcal{G}$ admet un prolongement presque continu sur $\bar{\mathcal{A}}$, que l'on notera de la même manière. Le seul cas particulier de ce prolongement, est celui où $A \in \bar{\mathcal{A}}$ contient le point $(p, (e^-, e^+)) \in \mathbb{G} \times \vec{E}$ si et seulement si il contient aussi le point $(p, (e^+, e^-))$. Dans ce cas, on peut vérifier que si $(A_n)_{n \in \mathbb{N}} \subset \mathcal{A}$ converge vers A , alors $(D^{(c)}(A_n))_{n \in \mathbb{N}}$ converge soit vers \mathbb{G} soit vers un ensemble fini. Cependant, \mathcal{G} ayant vocation à être l'espace d'état d'un processus pour lequel, d'après la proposition 4.4.4, l'ensemble des parties finies est une frontière inaccessible, il semble raisonnable de poser dans ce cas $D^{(c)}(A) = \mathbb{G}$, et $D^{(c)}(\cdot)$ admet donc des points de discontinuité sur $\bar{\mathcal{A}}$, par lesquels on peut vérifier que l'on ne sera pas concerné dans les calculs qui suivent.

Exemple 6.2.1. On a représenté dans la figure 6.1 un ensemble D (en pointillés) et sa frontière. L'orientation des arêtes est représentée par le sens des vecteurs (par exemple, l'arête \vec{e}_1 est celle allant de O à b). La configuration de départ est celle de l'image de gauche, et on a

$$\mathfrak{d}D = \{(d, -\vec{e}_1), (e, \vec{e}_1), (f, -\vec{e}_1), (g, -\vec{e}_2), (h, -\vec{e}_3)\}$$

On peut remarquer sur cet exemple que le choix de cette frontière et d'une arête entrante (par exemple $-\vec{e}_1$ en d) détermine l'ensemble D de manière unique. En effet, si on avait

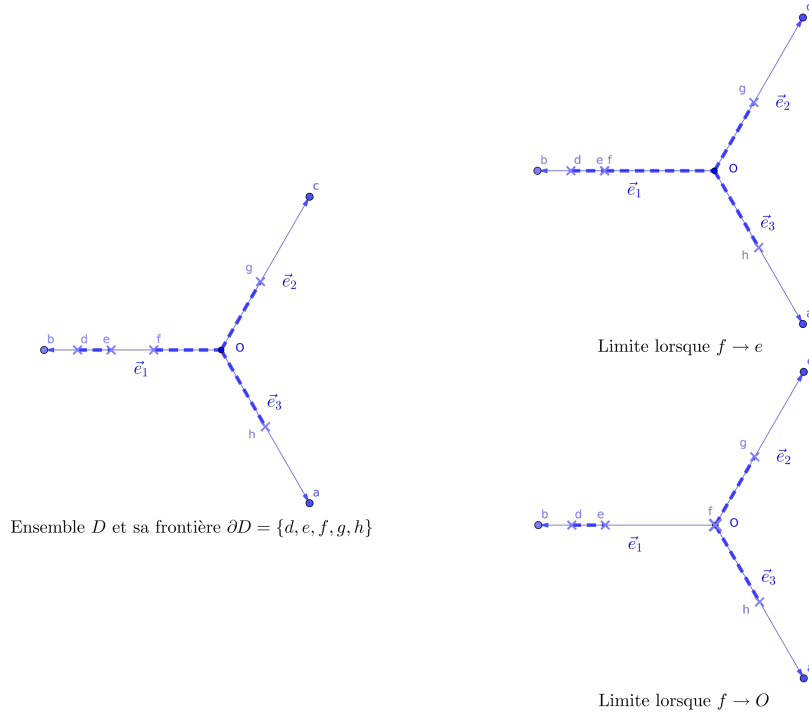


FIGURE 6.1

choisi \vec{e}_1 comme arête entrante en d , alors toutes les arêtes entrantes auraient été inversées et l'on aurait obtenu le complémentaire de l'ensemble D représenté ici.

À droite, on a fixé tous les points de la frontière à l'exception de f , que l'on a fait varier pour illustrer deux des trois cas de convergence vers un élément de $\bar{\mathcal{A}} \setminus \mathcal{A}$ (ces trois cas n'étant bien sûr pas mutuellement exclusifs). Dans le premier cas, on aboutit à la situation où il existe deux éléments $((e, \vec{e}_1)$ et $(f, -\vec{e}_1)$) dont les premières coordonnées sont égales et les deuxièmes sont les deux arêtes orientées correspondant à une même arête non-orientée (e et f ne sont donc plus à la frontière de D).

La deuxième situation illustrée ici est celle où un point de la frontière (f) atteint un sommet du graphe de telle sorte que l'arête entrante qui lui est associée ($-\vec{e}_1$) ne soit plus une arête entrante en ce point, bien que ce dernier soit toujours un point de la frontière (contrairement au cas précédent). On aurait pu illustrer le troisième cas en faisant tendre d vers b par exemple.

Dans tous les cas, on constate que l'ensemble $D(A_0)$ lorsque $A_0 \in \bar{\mathcal{A}} \setminus \mathcal{A}$ peut être déterminé de manière unique comme limite d'élément de la forme $D(A)$ avec $A \rightarrow A_0$, mais ne vérifie pas $\mathfrak{d}D(A) = A$.

Soit un ensemble $D \in \mathcal{G}$, et $\varepsilon := (\varepsilon_y, y)_{y \in \mathfrak{d}^{(c)}D}$, où $(\varepsilon_y)_{y \in \mathfrak{d}^{(c)}D}$ est une famille de nombres réels satisfaisant :

— $y + \varepsilon_y := (x + \varepsilon_y \vec{e}, \vec{e})$ existe pour tout $y = (x, \vec{e}) \in \mathfrak{d}^{(c)}D$.

— $\{y + \varepsilon_y, y \in \mathfrak{d}^{(c)}D\} \in \bar{\mathcal{A}}$.

On généralise la notion de ε -élargi de l'ensemble D par :

$$D^\varepsilon := D^{(c)}\left(\{y + \varepsilon_y, y \in \mathfrak{d}^{(c)}D\}\right)$$

Cela correspond en fait à l'ensemble D dont chaque bord a été « décalé » de ε_y dans la direction sortante (ou entrante). Dans la suite, lorsqu'on écrira D^ε on supposera que D et ε satisfont bien les conditions que l'on a imposées dans la définition. Dans le cas particulier où il existe $y \in \mathfrak{d}^{(c)}D$ tel que $\varepsilon_y = \eta \in \mathbb{R}$ et $\varepsilon_{y'} = 0$ pour tout $y' \neq y \in \mathfrak{d}^{(c)}D$, on

notera :

$$D^{\eta \cdot y} := D^\varepsilon$$

ce qui définit une notion d'élargi dans une direction particulière. On remarque que même lorsque $D^{\eta \cdot y}$ est bien défini, on n'a pas nécessairement $y + \eta \in \mathfrak{d}^{(c)} D^{\eta \cdot y}$: par exemple, si on note $y = (x, \vec{e})$, il peut exister $x' \in \partial D \setminus V$ tel que $x + \eta \vec{e} = x'$, et ce point n'appartient donc pas à la frontière de $D^{\eta \cdot y}$.

On introduit alors une notion de dérivées partielles pour une fonction continue $f : \mathcal{G} \rightarrow \mathbb{R}$, par :

$$\begin{aligned} \forall D \in \mathcal{G}, \forall y \in \mathfrak{d}^{(c)} D, \quad \partial_y f(D) &:= \lim_{\eta \rightarrow 0} \frac{f(D^{\eta \cdot y}) - f(D)}{\eta} \\ \forall D \in \mathcal{G}, \forall y \in \mathfrak{d}^{(c)} D, \quad \partial_y^2 f(D) &:= \lim_{\eta \rightarrow 0} \frac{\partial_{y+\eta} f(D^{\eta \cdot y}) - \partial_y f(D)}{\eta} \end{aligned}$$

dès que ces limites existent et où celles-ci sont prises sur les $\eta \in \mathbb{R}$ tels que $D^{\eta \cdot y}$ a un sens (par exemple, si $y = (x, \vec{e}) \in \mathfrak{d} D$ est tel que $x \in V$, alors $D^{\eta \cdot y}$ n'a de sens que pour $\eta > 0$). Par analogie avec les fonctions sur \mathbb{R}^n , on dira que f est \mathcal{C}^1 ou \mathcal{C}^2 sur \mathcal{G} si ses dérivées partielles existent et sont continues, dans le sens où, pour tout $A \in \bar{\mathcal{A}}$, $y = (x, \vec{e}) \in A$, $(A_n)_{n \in \mathbb{N}} \subset \mathcal{A}$ et $(y_n)_{n \in \mathbb{N}} = (x_n, \vec{e})_{n \in \mathbb{N}}$ tels que $A_n \rightarrow A$, $y_n \in A_n$, $\forall n \in \mathbb{N}$ et $x_n \rightarrow x$, on a :

$$\lim_{n \rightarrow \infty} \partial_{y_n} f(D^{(c)}(A_n)) \begin{cases} \text{existe} \\ = \partial_y f(D^{(c)}(A)), \end{cases} \text{ si } A \in \mathcal{A}$$

et similairement pour ∂^2 .

Remarque 6.2.2. On prendra garde au fait que la fonction $\partial_y f$ n'est définie que sur l'ensemble $\{D \in \mathcal{G} : y \in \mathfrak{d} D \cup \mathfrak{d}^c D\}$, raison pour laquelle on ne peut pas définir formellement la dérivée seconde par $\partial_y^2 = \partial_y \partial_y$. Dans la suite, on fera toutefois cet abus de notation, cela n'entraînant pas de risque de confusion puisque l'opérateur $\partial_y \partial_y$ en lui-même n'est pas défini.

Il est facile de vérifier que les formules de Taylor s'appliquent, et on l'explicite ici à l'ordre 2 (avec l'abus de notation que l'on vient de mentionner) dans le cas où f est \mathcal{C}^2 :

$$\begin{aligned} \forall D \in \mathcal{G}, \forall \varepsilon = (\varepsilon_y, y)_{y \in \mathfrak{d}^{(c)} D}, \quad f(D^\varepsilon) - f(D) \\ = \sum_{y \in \mathfrak{d} D} \partial_y f(D) \varepsilon_y + \sum_{y, z \in \mathfrak{d} D} \partial_y \partial_z f(D) \frac{\varepsilon_y \varepsilon_z}{2} + o(\|(\varepsilon_y)_y\|^2) \end{aligned}$$

(si D^ε est bien défini), où $\|\cdot\|$ désigne la norme euclidienne sur $\mathbb{R}^{|\mathfrak{d}^{(c)} D|}$.

Après ces préliminaires, on peut maintenant utiliser cette nouvelle notion de dérivées partielles pour élargir la définition du générateur \mathfrak{L} . Pour toute fonction $f \in \mathcal{C}^2(\mathcal{G})$ et pour tout $D \in \mathcal{D}$, on pose :

$$\mathfrak{L}[f](D) = \sum_{y, z \in \mathfrak{d} D} \partial_y \partial_z f(D) - \sum_{y = (x, \vec{e}) \in \mathfrak{d} D} \partial_y f(D) \left(2 \frac{\underline{\mu}(\partial D)}{\underline{\mu}(D)} - i_{\vec{e}} b_e(x) + 2 \frac{\underline{\mu}'(x, \vec{e})}{\underline{\mu}(x)} \right) \quad (6.10)$$

Cette définition est compatible avec celle donnée au chapitre 4. En effet, pour tout $f \in \mathcal{F}$, $D \in \mathcal{D}$ et $y = (x, \vec{e}), z \in \mathfrak{d} D$, on a :

$$\begin{aligned} \partial_y [F_f](D) &= -f(x) \underline{\mu}(x) \\ \partial_z \partial_y [F_f](D) &= \begin{cases} 0 & \text{si } y \neq z \\ -f'(x, \vec{e}) \underline{\mu}(x) - f(x) \underline{\mu}'(x, \vec{e}) & \text{sinon} \end{cases} \\ \underline{\mu}'(x, \vec{e}) &= \frac{i_{\vec{e}} \partial_e \tilde{\mu}_e(x)}{Z} = i_{\vec{e}} b_e(x) \underline{\mu}(x) + \frac{\gamma_{\vec{e}}}{Z} \end{aligned}$$

cette dernière égalité découlant de (4.13), donc :

$$\begin{aligned}\mathfrak{L}[F_f](D) &= \sum_{(x,\vec{e}) \in \partial D} \left(-f'(x, \vec{e}) \underline{\mu}(x) + f(x) \underline{\mu}'(x, \vec{e}) + f(x) \underline{\mu}(x) \left(2 \frac{\underline{\mu}(\partial D)}{\underline{\mu}(D)} - i_{\vec{e}} b_e(x) \right) \right) \\ &= \sum_{x \in \partial D} \left(-\partial_D f(x) + \gamma_D f(x) \right) + 2 \frac{\underline{\mu}(\partial D)}{\underline{\mu}(D)} \int_{\partial D} f d\underline{\mu}\end{aligned}$$

On remarque que pour toute fonction $f \in \mathcal{C}^2(\mathcal{G})$, pour tout $D \in \mathbb{G}$ et $x \in \partial D \setminus V$, si on note $y := (x, \vec{e}) \in \mathfrak{d}D$ et $y' := (x, \vec{e}') \in \mathfrak{d}^c D$, alors on a :

$$\partial_y f(D) + \partial_{y'} f(D) = 0$$

Les deux applications ∂_y et $\partial_{y'}$ jouent donc le rôle des dérivées à droite et à gauche pour des fonctions d'une variable réelle. De la même manière que pour les fonctions sur \mathbb{G} , on généralise l'égalité ci-dessus aux sommets de \mathbb{G} , en définissant $\mathcal{D}(\mathfrak{L})$ comme l'ensemble des fonctions $f \in \mathcal{C}^2(\mathcal{G})$ telles que pour tout $D \in \mathcal{G}$ et tout $x \in \partial D \cap V$:

$$\sum_{y=(x,\vec{e}) \in \partial D} \partial_y f(D) \sum_{(x,\vec{e}') \in \mathfrak{d}^c D} p(x, e) + \sum_{y=(x,\vec{e}') \in \mathfrak{d}^c D} \partial_{y'} f(D) \sum_{(x,\vec{e}) \in \partial D} p(x, e) = 0 \quad (6.11)$$

On peut vérifier facilement que pour tout $f \in \mathcal{C}(\mathbb{G})$, la fonctionnelle F_f est un élément de $\mathcal{D}(\mathfrak{L})$.

La question de l'existence et de l'unicité de la solution du problème de martingale associé à ce générateur est déjà partiellement résolue. En effet, dire que $(D_t)_{t \in [0, \tau]}$ est une telle solution jusqu'au temps $\tau := \tau_1 \wedge \tau_2$, où :

$$\begin{aligned}\tau_1 &:= \inf\{t > 0 : \partial D_t \cap V \neq \emptyset\} \\ \tau_2 &:= \inf\{t > 0 : |\partial D_t| \neq |\partial D_{t-}|\}\end{aligned}$$

est équivalent à dire que $(\partial D_t)_{t \in [0, \tau]}$ est une diffusion multidimensionnelle, dont le générateur découle de manière relativement immédiate de l'expression de \mathfrak{L} . Le processus ∂D est uniquement déterminé jusqu'au temps τ , et on en déduit D par continuité des trajectoires. Autrement dit, on peut définir un processus D de générateur \mathfrak{L} jusqu'au moment où l'un de ses bords rencontre un sommet ou un autre bord. Si $\tau_2 < \tau_1$, on peut faire repartir le processus de D_{τ_2} , et par recollement des trajectoires il n'est pas difficile de définir D jusqu'au temps τ_1 .

Un problème se pose en revanche au temps τ_1 . Si on essaie de faire repartir D de $D_{\tau_1} \in \mathcal{G} \setminus \mathcal{D}$, alors l'une de ses extrémités se trouve en un sommet du graphe. Elle va alors « choisir » une direction, et en revenir une infinité de fois en un temps arbitrairement court (comme un brownien partant de 0). Ce phénomène peut alors donner lieu à une augmentation du nombre de composantes connexes, et nous ne savons pas, pour l'instant, montrer de manière directe que D ne va pas atteindre \mathcal{G}_2 à ce moment-là.

6.2.2 Mise en place de la convergence

Afin de contourner le problème que nous venons d'évoquer, nous allons maintenant recourir à une approximation par des processus de Markov discrets pour montrer l'existence d'une solution au problème de martingale.

Pour tout $n, N \in \mathbb{N}$, on définit \mathcal{D}_n^N l'ensemble des parties $Q \subset V_n$ ayant au maximum N composantes connexes, satisfaisant $\text{Fr}(Q) \cap V = \emptyset$ et telles que pour tout $p \in \text{Fr}_i(Q)$ (resp. $\text{Fr}_e(Q)$), on a $p \in \text{Fr}_e(Q \setminus \{p\})$ (resp. $\text{Fr}_i(Q \cup \{p\})$). Autrement dit, chaque point de la frontière de $Q \in \mathcal{D}_n^N$ est à l'intérieur d'une arête et a donc deux voisins dans V_n , l'un appartenant à Q et l'autre non. En particulier, on a $|\text{Fr}_e(Q)| = |\text{Fr}_i(Q)|$. Dans un premier temps, on considérera $N \in \mathbb{N}$ fixé et on notera simplement \mathcal{D}_n . La suite $(\mathcal{D}_n)_{n \in \mathbb{N}}$ est une discrétisation d'une sous-partie de l'ensemble \mathcal{G}^N des éléments de \mathcal{G} ayant au plus N composantes connexes, et on identifiera tout ensemble $Q \subset V_n$ avec l'ensemble $\bigcup_{p \in Q} \overline{I_n(p)}$, de sorte que l'on peut considérer \mathcal{D}_n comme un sous-ensemble de \mathcal{G} . On fera cependant attention à la distinction entre $\text{Fr}(Q)$ et ∂Q , les points de ce dernier étant des extrémités d'intervalles de la forme $I_n(p)$, et donc situés à mi-chemin entre un point de $\text{Fr}_e(Q)$ et un point de $\text{Fr}_i(Q)$ (en particulier, $V_n \cap \partial Q = \emptyset$).

Soit $(g_n)_{n \in \mathbb{N}}$ la suite de fonctions définie par :

$$\forall n \in \mathbb{N}, \quad g_n : \begin{array}{l} \mathcal{G}^N \rightarrow \mathcal{P}(V_n) \\ D \mapsto D \cap V_n \end{array}$$

Avec l'identification de $D \cap V_n$ comme élément de \mathcal{G} , la suite $(g_n)_{n \in \mathbb{N}}$ converge uniformément vers la fonction identité de \mathcal{G}^N .

On définit enfin le générateur dual de \mathcal{L}^n comme celui de l'exemple 3.0.6, et on le note \mathfrak{L}^n .

Proposition 6.2.3. *Pour toute fonction $f \in \mathcal{D}(\mathfrak{L})$ telle que $\mathfrak{L}[f]$ est continue, la suite $(f \circ g_n)_{n \in \mathbb{N}}$ converge uniformément vers f et la suite $(\mathfrak{L}^n[f \circ g_n])_{n \in \mathbb{N}}$ converge vers $\mathfrak{L}f$ uniformément sur tout $g_k^{-1}(\mathcal{D}_k)$, $k \geq 1$ et est bornée uniformément sur \mathcal{G}^N .*

Démonstration. La preuve qui suit ne présente pas d'astuce particulière. La seule difficulté est d'ordre technique, et consiste à exprimer correctement les développements limités des coefficients de la suite de générateur $(\mathfrak{L}^n)_{n \in \mathbb{N}}$.

Soient $f \in \mathcal{D}(\mathfrak{L})$ et la suite $(f_n)_{n \in \mathbb{N}}$ définie par $f_n = f \circ g_n$, $\forall n \in \mathbb{N}$. Puisque $(g_n)_{n \in \mathbb{N}}$ converge uniformément vers la fonction identité et que la fonction f est (uniformément) continue, la suite $(f_n)_{n \in \mathbb{N}}$ converge donc uniformément vers f .

Soient $k \in \mathbb{N}$, $D \in g_k^{-1}(\mathcal{D}_k)$, $n \in \mathbb{N}$ et $Q = g_n(D)$. On remarque que si n est assez grand (la valeur limite pouvant être strictement supérieure à k), on a $Q \in \mathcal{D}_n$, et on supposera donc que tel est le cas. Soient $(p_e^i)_{1 \leq i \leq m}$ et $(p_i^i)_{1 \leq i \leq m}$ l'indexation de ses frontières extérieure et intérieure telle que décrite dans l'exemple 3.0.6 (on a ici $m = m_e = m_i$), et $(\partial_i^+ Q)_{1 \leq i \leq m}$ et $(\partial_i^- Q)_{1 \leq i \leq m}$ les ensembles correspondants, de sorte que l'on a, on rappelle :

$$\begin{aligned} \mathfrak{L}_{Q, \partial_m^+ Q}^n &= \frac{\mu^n(\partial_m^+ Q)}{\mu^n(Q)} \sum_{q \in Q} \frac{\mu^n(q)}{\mu^n(p_e^m)} \mathcal{L}_{q, p_e^m}^n \\ \forall i < m, \quad \mathfrak{L}_{Q, \partial_i^+ Q}^n &= \frac{\mu^n(\partial_i^+ Q)}{\mu^n(Q)} \left(\sum_{q \in Q} \frac{\mu^n(q)}{\mu^n(p_e^i)} \mathcal{L}_{q, p_e^i}^n - \sum_{q \in Q} \frac{\mu^n(q)}{\mu^n(p_e^{i+1})} \mathcal{L}_{q, p_e^{i+1}}^n \right) \\ \mathfrak{L}_{Q, \partial_m^- Q}^n &= \frac{\mu^n(\partial_m^- Q)}{\mu^n(Q)} \sum_{q \notin Q} \frac{\mu^n(q)}{\mu^n(p_i^m)} \mathcal{L}_{q, p_i^m}^n \\ \forall i < m, \quad \mathfrak{L}_{Q, \partial_i^- Q}^n &= \frac{\mu^n(\partial_i^- Q)}{\mu^n(Q)} \left(\sum_{q \notin Q} \frac{\mu^n(q)}{\mu^n(p_i^i)} \mathcal{L}_{q, p_i^i}^n - \sum_{q \notin Q} \frac{\mu^n(q)}{\mu^n(p_i^{i+1})} \mathcal{L}_{q, p_i^{i+1}}^n \right) \end{aligned}$$

Ces quantités sont positives, $\partial_i^+ Q \cup \{p_e^{i+1}\} = \partial_{i+1}^+ Q$ pour tout $i < m$ et $\partial_i^- Q \setminus \{p_i^{i+1}\} = \partial_{i+1}^- Q$ pour tout $i < m$ (avec $\partial_0^+ Q = \partial_0^- Q = D$). De plus, par définition de \mathcal{D}_n , chacun des points

de la frontière (intérieure et extérieure) de Q ne peut avoir qu'un voisin dans Q et un dans Q^c . Autrement dit, les sommes dans les expressions ci-dessus ne comportent qu'un terme. Pour tout $i \leq m$, on notera q_e^i (resp. q_i^i) l'unique voisin de p_e^i (resp. p_i^i) dans Q (resp. Q^c) et \vec{e}_i (resp. \vec{e}'_i) l'unique arête orientée telle que $p_e^i + h_n^{e_i} \vec{e}_i = q_e^i$ (resp. $q_i^i + h_n^{e'_i} \vec{e}'_i = p_i^i$). On remarque qu'il existe une permutation ϕ sur $\llbracket 1, m \rrbracket$ telle que pour tout $i \leq m$, on a $p_e^i = q_i^{\phi(i)}$, $q_e^i = p_i^{\phi(i)}$ et $\vec{e}_i = \vec{e}'_{\phi(i)}$ et que les \vec{e}_i sont les arêtes entrantes de Q .

On cherche alors à calculer :

$$\mathfrak{L}^n[f_n](D) = \sum_{i=1}^m \mathfrak{L}_{Q, \partial_i^+ Q}^n \left(f(\partial_i^+ Q) - f(Q) \right) + \sum_{i=1}^m \mathfrak{L}_{Q, \partial_i^- Q}^n \left(f(\partial_i^- Q) - f(Q) \right)$$

et on s'intéresse ici à la première somme, la deuxième se traitant de manière similaire. En reprenant la notation $\alpha_{Q'}$ de l'exemple 3.0.6, on peut écrire :

$$\sum_{i=1}^m \mathfrak{L}_{Q, \partial_i^+ Q}^n \left(f(\partial_i^+ Q) - f(Q) \right) = \sum_{i=1}^m \left(f(\partial_i^+ Q) - f(Q) \right) \frac{\mu^n(\partial_i^+ Q)}{\mu^n(Q)} \alpha_{\partial_i^+ Q}$$

Nous allons donc examiner ces objets de plus près, afin de pouvoir déterminer leur comportement asymptotique lorsque l'on fait $n \rightarrow \infty$ (on fait remarquer que Q ici dépend implicitement de n). On conserve la notation $o_n(g(n))$ pour désigner une fonction dominée uniformément (sur $g_i^{-1}(\mathcal{D}_i)$) par $g(n)$.

D'après la proposition 6.1.7, on a :

$$\begin{aligned} \mu^n(\partial_m^+ Q) &= \mu^n(Q) + \mu^n(\text{Fr}_e(Q)) \\ &= \mu^n(Q) + \underline{\mu}(\text{Fr}_e(Q)) h_n + o_n(h_n) \\ &= \mu^n(Q) + (\underline{\mu}(\partial Q) + o_n(1)) h_n + o_n(h_n) \\ &= \mu^n(Q) + \underline{\mu}(\partial Q) h_n + o_n(h_n) \\ \mu^n(Q) &= \mu(Q) + o_n(1) \\ \frac{\mu^n(\partial_m^+ Q)}{\mu^n(Q)} &= 1 + \frac{\underline{\mu}(\partial Q)}{\mu(Q)} h_n + o_n(h_n) \\ \forall i < m, \quad \frac{\mu^n(\partial_i^+ Q)}{\mu^n(Q)} &= 1 + o_n(1) \end{aligned}$$

De plus, pour tout $\vec{e} \in \vec{E}$ si $p, q \in \mathring{R}_e$ sont tels que $q = p + h_n^e \vec{e}$, alors par la proposition 6.1.7 on a :

$$\begin{aligned} \frac{\mu^n(q)}{\mu^n(p)} &= \frac{p_n(p, q) \mu_e(q)}{p_n(q, p) \mu_e(p)} \\ &= \frac{\mu_e(p) + \mu'_e(p, \vec{e}) h_n + o_n(h_n)}{\mu_e(p)} \\ &= \left(1 + \frac{\mu'_e(p, \vec{e})}{\mu_e(p)} h_n \right) + o_n(h_n) \end{aligned}$$

Prenant en compte le développement de $\mathcal{L}_{q,p}^n$ donné par le lemme 6.1.3, on obtient, pour les mêmes $p, q \in \mathring{R}_e$:

$$\begin{aligned} \frac{\mu^n(q)}{\mu^n(p)} \mathcal{L}_{q,p}^n &= \left(1 + \frac{\mu'_e(p, \vec{e})}{\mu_e(p)} h_n \right) \left(\frac{1}{h_n^2} - i_{\vec{e}} \frac{b_e(q)}{2h_n} \right) + o_n \left(\frac{1}{h_n} \right) \\ &= \frac{1}{h_n^2} + \left(\frac{\mu'_e(p, \vec{e})}{\mu_e(p)} - \frac{i_{\vec{e}} b_e(q)}{2} \right) \frac{1}{h_n} + o_n \left(\frac{1}{h_n} \right) \end{aligned}$$

On obtient :

$$\begin{aligned} \sum_{q \in Q} \frac{\mu^n(q)}{\mu^n(p_{\mathbf{e}}^i)} \mathcal{L}_{q, p_{\mathbf{e}}^i}^n &= \frac{\mu^n(q_{\mathbf{e}}^i)}{\mu^n(p_{\mathbf{e}}^i)} \mathcal{L}_{q_{\mathbf{e}}^i, p_{\mathbf{e}}^i}^n \\ &= \frac{1}{h_n^2} + \left(\frac{\underline{\mu}'(q_{\mathbf{e}}^i, \vec{e}_i)}{\underline{\mu}(q_{\mathbf{e}}^i)} - \frac{i_{\vec{e}_i} b_{e_i}(q_{\mathbf{e}}^i)}{2} \right) \frac{1}{h_n} + o_n \left(\frac{1}{h_n} \right) \end{aligned}$$

et donc :

$$\begin{aligned} \alpha_{\partial_i^+ Q} &= \sum_{q \in Q} \frac{\mu^n(q)}{\mu^n(p_{\mathbf{e}}^i)} \mathcal{L}_{q, p_{\mathbf{e}}^i}^n - \sum_{q \in Q} \frac{\mu^n(q)}{\mu^n(p_{\mathbf{e}}^{i+1})} \mathcal{L}_{q, p_{\mathbf{e}}^{i+1}}^n \\ &= \left(\frac{\underline{\mu}'(q_{\mathbf{e}}^i, \vec{e}_i)}{\underline{\mu}(q_{\mathbf{e}}^i)} - \frac{i_{\vec{e}_i} b_{e_i}(q_{\mathbf{e}}^i)}{2} - \frac{\underline{\mu}'(q_{\mathbf{e}}^{i+1}, \vec{e}_{i+1})}{\underline{\mu}(q_{\mathbf{e}}^{i+1})} + \frac{i_{\vec{e}_{i+1}} b_{e_{i+1}}(q_{\mathbf{e}}^{i+1})}{2} \right) \frac{1}{h_n} + o_n \left(\frac{1}{h_n} \right) \end{aligned}$$

On trouve alors, pour tout $0 < i < m$:

$$\begin{aligned} \mathfrak{L}_{Q, \partial_m^+ Q}^n &= \frac{1}{h_n^2} + \left(\frac{\underline{\mu}(\partial Q)}{\mu(Q)} + \frac{\underline{\mu}'(q_{\mathbf{e}}^m, \vec{e}_m)}{\underline{\mu}(q_{\mathbf{e}}^m)} - \frac{i_{\vec{e}_m} b_{e_m}(q_{\mathbf{e}}^m)}{2} \right) \frac{1}{h_n} + o_n \left(\frac{1}{h_n} \right) \\ \mathfrak{L}_{Q, \partial_i^+ Q}^n &= \left(\frac{\underline{\mu}'(q_{\mathbf{e}}^i, \vec{e}_i)}{\underline{\mu}(q_{\mathbf{e}}^i)} - \frac{\underline{\mu}'(q_{\mathbf{e}}^{i+1}, \vec{e}_{i+1})}{\underline{\mu}(q_{\mathbf{e}}^{i+1})} - \frac{i_{\vec{e}_i} b_{e_i}(q_{\mathbf{e}}^i) - i_{\vec{e}_{i+1}} b_{e_{i+1}}(q_{\mathbf{e}}^{i+1})}{2} \right) \frac{1}{h_n} + o_n \left(\frac{1}{h_n} \right) \end{aligned}$$

Ayant obtenu un développement limité des différents $\mathfrak{L}_{Q, \partial_i^+ Q}^n$ (et des $\mathfrak{L}_{Q, \partial_i^- Q}^n$ non explicités ici), il ne nous reste plus qu'à appliquer la formule de Taylor pour approcher $f(\partial_i^+ Q) - f(Q)$. L'ensemble $\partial_i^+ Q$ étant simplement l'ensemble Q élargi de h_n^e en chaque point $(q_{\mathbf{e}}^j, \vec{e}_j) \in \partial Q$, $j \leq i$ (avec un léger abus, on considérera ici $q_{\mathbf{e}}^j$ comme point de la frontière, au lieu de $p_{\mathbf{e}}^j + \frac{h_n^{e_j}}{2} \vec{e}_j$, sans incidence sur les calculs), on obtient :

$$\begin{aligned} \mathfrak{L}_{Q, \partial_m^+ Q}^n \left(f(\partial_m^+ Q) - f(Q) \right) &= \sum_{y, z \in \partial Q} \frac{\partial_y \partial_z f(Q)}{2} - \sum_{y \in \partial Q} \partial_y f(Q) \left(\frac{1}{h_n} + \frac{\underline{\mu}(\partial Q)}{\mu(Q)} \right. \\ &\quad \left. + \frac{\underline{\mu}'(q_{\mathbf{e}}^m, \vec{e}_m)}{\underline{\mu}(q_{\mathbf{e}}^m)} - \frac{i_{\vec{e}_m} b_{e_m}(q_{\mathbf{e}}^m)}{2} \right) + o_n(1) \end{aligned}$$

et pour tout $i < m$:

$$\begin{aligned} \mathfrak{L}_{Q, \partial_i^+ Q}^n \left(f(\partial_i^+ Q) - f(Q) \right) &= - \sum_{j=1}^i \partial_{(q_{\mathbf{e}}^j, \vec{e}_j)} f(Q) \left(\frac{\underline{\mu}'(q_{\mathbf{e}}^j, \vec{e}_j)}{\underline{\mu}(q_{\mathbf{e}}^j)} - \frac{\underline{\mu}'(q_{\mathbf{e}}^{j+1}, \vec{e}_{j+1})}{\underline{\mu}(q_{\mathbf{e}}^{j+1})} \right. \\ &\quad \left. - \frac{i_{\vec{e}_j} b_{e_j}(q_{\mathbf{e}}^j) - i_{\vec{e}_{j+1}} b_{e_{j+1}}(q_{\mathbf{e}}^{j+1})}{2} \right) + o_n(1) \end{aligned}$$

On remarque que lorsque qu'on somme, une partie du terme en i annule une partie du terme en $i+1$, et il reste :

$$\begin{aligned} \sum_{i=1}^m \mathfrak{L}_{Q, \partial_i^+ Q}^n \left(f(\partial_i^+ Q) - f(Q) \right) &= \sum_{y, z \in \partial Q} \frac{\partial_y \partial_z f(Q)}{2} - \sum_{i=1}^m \partial_{(q_{\mathbf{e}}^i, \vec{e}_i)} f(Q) \left(\frac{1}{h_n} + \frac{\underline{\mu}(\partial Q)}{\mu(Q)} \right. \\ &\quad \left. + \frac{\underline{\mu}'(q_{\mathbf{e}}^i, \vec{e}_i)}{\underline{\mu}(q_{\mathbf{e}}^i)} - \frac{i_{\vec{e}_i} b_{e_i}(q_{\mathbf{e}}^i)}{2} \right) + o_n(1) \end{aligned}$$

Si on reprend les mêmes calculs en remplaçant pour tout $i \leq m$, ∂_i^+ par ∂_i^- , p_e par p_i , et \vec{e}_i par \vec{e}'_i l'arête entrante en p_i^i , on trouve :

$$\begin{aligned} \sum_{i=1}^m \mathfrak{L}_{Q, \partial_i^+ Q}^n \left(f(\partial_i^+ Q) - f(Q) \right) &= \sum_{y, z \in \partial Q} \frac{\partial_y \partial_z f(Q)}{2} + \sum_{i=1}^m \partial_{(p_i^i, \vec{e}'_i)} f(Q) \left(\frac{1}{h_n} - \frac{\mu(\partial Q)}{\mu(Q)} \right. \\ &\quad \left. - \frac{\mu'(p_i^i, \vec{e}'_i)}{\mu(p_i^i)} + \frac{i_{\vec{e}'_i} b_{e'_i}(p_i^i)}{2} \right) + o_n(1) \end{aligned}$$

En ajoutant ces deux sommes on obtient finalement :

$$\begin{aligned} \mathfrak{L}^n f_n(D) &= \mathfrak{L}^n f(Q) \\ &= \sum_{y, z \in \partial Q} \partial_y \partial_z f(Q) + \sum_{y=(p, \vec{e}) \in \partial Q} \partial_y f(Q) \left(-2 \frac{\mu(\partial Q)}{\mu(Q)} - 2 \frac{\mu'(p, \vec{e})}{\mu(p)} + i_{\vec{e}} b_e(p) \right) + o_n(1) \\ &= \mathfrak{L}f(Q) + o_n(1) \\ &= \mathfrak{L}f(D) + o_n(1) \end{aligned}$$

par continuité uniforme de $\mathfrak{L}f$ sur \mathcal{G}^N .

Nous avons donc montré que la suite de fonction $\mathfrak{L}^n f_n$ converge uniformément vers $\mathfrak{L}f$ sur $g_i^{-1}(D_i)$, et il nous reste à montrer qu'elle est uniformément bornée sur \mathcal{G}^N . Puisque nous ne sommes ici plus intéressé par la convergence, nous utiliserons la notation $O_n(\cdot)$ avec une signification évidente, le caractère uniforme étant cette fois pris sur \mathcal{G}^N .

Soit donc maintenant $D \in \mathcal{G}^N$, $n \in \mathbb{N}$ et $Q := g_n(D)$. On note encore $(p_e^i)_{1 \leq i \leq m_e}$ et $(p_i^i)_{1 \leq i \leq m_i}$ l'indexation des frontières extérieure et intérieure de Q telle que décrite dans l'exemple 3.0.6, et $(\partial_i^+ Q)_{1 \leq i \leq m_e}$ et $(\partial_i^- Q)_{1 \leq i \leq m_i}$ les ensembles correspondants. Pour tout $i \leq m_e$, on a $f(\partial_i^+ Q) - f(Q) = O_n(h_n)$, $\frac{\mu^n(\partial_i^+ Q)}{\mu^n(Q)} = 1 + O_n(h_n)$ et $\alpha_{\partial_i^+ Q} = O_n\left(\frac{1}{h_n^2}\right)$. On peut alors écrire, avec la convention $\partial_0^+ Q = Q$:

$$\begin{aligned} \sum_{i=1}^{m_e} \mathfrak{L}_{Q, \partial_i^+ Q}^n \left(f(\partial_i^+ Q) - f(Q) \right) &= \sum_{i=1}^{m_e} \left(f(\partial_i^+ Q) - f(Q) \right) \frac{\mu^n(\partial_i^+ Q)}{\mu^n(Q)} \alpha_{\partial_i^+ Q} \\ &= \sum_{i=1}^{m_e} \left(f(\partial_i^+ Q) - f(Q) \right) \alpha_{\partial_i^+ Q} + O_n(1) \\ &= \sum_{i=1}^{m_e} \sum_{q \in Q} \frac{\mu^n(q)}{\mu^n(p_e^i)} \mathcal{L}_{q, p_e^i}^n \left(f(\partial_i^+ Q) - f(\partial_{i-1}^+ Q) \right) + O_n(1) \end{aligned}$$

et similairement pour l'autre somme :

$$\sum_{i=1}^{m_i} \mathfrak{L}_{Q, \partial_i^- Q}^n \left(f(\partial_i^- Q) - f(Q) \right) = \sum_{i=1}^{m_i} \sum_{q \notin Q} \frac{\mu^n(q)}{\mu^n(p_i^i)} \mathcal{L}_{q, p_i^i}^n \left(f(\partial_i^- Q) - f(\partial_{i-1}^- Q) \right) + O_n(1)$$

Pour bien comprendre pourquoi ces deux sommes s'annulent mutuellement, nous allons fixer $i \in \llbracket 1, m_e \rrbracket$ et étudier le terme correspondant dans la première somme. En reprenant les calculs précédents, on remarque que pour tout $q \in R_e$, $e \in E_{p_e^i}$, tel que $\{q, p_e^i\} \in E_n$, on a :

$$\frac{\mu^n(q)}{\mu^n(p_e^i)} \mathcal{L}_{q, p_e^i}^n = \frac{2p_n(p_e^i, q)}{h_n^2} + O_n\left(\frac{1}{h_n}\right)$$

On définit l'ensemble :

$$F := \{(p, q, \vec{e}) \in \text{Fr}_i \times \text{Fr}_e \times \vec{E} : p + h_n^e \vec{e} = q\}$$

Soient $(q_1, \vec{e}_1), \dots, (q_{k_0}, \vec{e}_{k_0})$, $k_0 \in \mathbb{N}$ les éléments de $\partial Q \times \vec{E}_{p_e^i}$ tels que $p_e^i + h_n^{e_k}/2 \vec{e}_k = q_k$, $k = 1, \dots, k_0$, et $(q_{k_0+1}, \vec{e}_{k_0+1}), \dots, (q_{k_0+k_1}, \vec{e}_{k_0+k_1})$ ceux de $Q^c \times \vec{E}_{p_e^i}$ tels que $p_e^i + h_n^{e_k}/2 \vec{e}_k = q_k$, $k = k_0 + 1, \dots, k_0 + k_1$. Le passage de $\partial_{i-1}^+ Q$ à $\partial_i^+ Q$ consiste en l'ajout de l'ensemble $I_n(p_e^i)$, lui-même constitué des segments $[p_e^i, q_k]$, $k = 1, \dots, k_0 + k_1$. Pour approcher la quantité $f(\partial_i^+ Q) - f(\partial_{i-1}^+ Q)$, nous allons donc utiliser la formule de Taylor, en nous servant de ces différents segments. Dans le but de clarifier les notations, nous écrirons simplement $\partial_{(p_e^i, \vec{e}_k)} f(Q)$ pour $\partial_{(q_k, \vec{e}_k)} f(Q \cup \bigcup_{j=1}^{k_0} [p_e^i, q_j])$, c'est-à-dire, on considère Q élargi « jusqu'au point p_e^i ». On a alors :

$$\begin{aligned} f(\partial_i^+ Q) - f(\partial_{i-1}^+ Q) &= \left(- \sum_{k=1}^{k_0} \partial_{(q_k, \vec{e}_k)} f(\partial_{i-1}^+ Q) + \sum_{k=k_0+1}^{k_1} \partial_{p_e^i, e_k} f(Q) \right) \frac{h_n}{2} + O_n(h_n^2) \\ &= \left(- \sum_{k=1}^{k_0} \partial_{(p_e^i, \vec{e}_k)} f(Q) + \sum_{k=k_0+1}^{k_1} \partial_{p_e^i, e_k} f(Q) + O_n(h_n) \right) \frac{h_n}{2} + O_n(h_n^2) \\ &= \left(- \sum_{k=1}^{k_0} \partial_{(p_e^i, \vec{e}_k)} f(Q) + \sum_{k=k_0+1}^{k_1} \partial_{p_e^i, e_k} f(Q) \right) \frac{h_n}{2} + O_n(h_n^2) \end{aligned}$$

donc :

$$\begin{aligned} &\sum_{q \notin Q} \frac{\mu^n(q)}{\mu^n(p_e^i)} \mathcal{L}_{q, p_e^i}^n \left(f(\partial_i^- Q) - f(\partial_{i-1}^- Q) \right) \\ &= \sum_{k=1}^{k_0} \frac{p_n(p_e^i, e_k)}{h_n} \left(\sum_{k=k_0+1}^{k_1} \partial_{p_e^i, e_k} f(Q) - \sum_{k=1}^{k_0} \partial_{(p_e^i, \vec{e}_k)} f(Q) \right) + O_n(1) \\ &= \frac{1}{h_n} \left(\sum_{k=1}^{k_0} p_n(p_e^i, e_k) \sum_{k=k_0+1}^{k_1} \partial_{p_e^i, e_k} f(Q) - \left(1 - \sum_{k=k_0+1}^{k_1} p_n(p_e^i, e_k) \right) \sum_{k=1}^{k_0} \partial_{(p_e^i, \vec{e}_k)} f(Q) \right) + O_n(1) \\ &= \frac{1}{h_n} \sum_{k=1}^{k_0} \partial_{(p_e^i, \vec{e}_k)} f(Q) + O_n(1) \end{aligned}$$

d'après la condition (6.11) sur f . Par continuité, cette dernière quantité peut se réécrire :

$$\frac{1}{h_n} \sum_{k=1}^{k_0} \partial_{(q_k, \vec{e}_k)} f(Q) + O_n(1)$$

En faisant la même chose pour chaque p_e^i , $i = 1, \dots, m_e$ et en sommant, on obtient :

$$\sum_{i=1}^{m_e} \mathfrak{L}_{Q, \partial_i^+ Q}^n \left(f(\partial_i^+ Q) - f(Q) \right) = \frac{1}{h_n} \sum_{y \in \partial Q} \partial_y f(Q) + O_n(1)$$

et symétriquement :

$$\begin{aligned} \sum_{i=1}^{m_i} \mathfrak{L}_{Q, \partial_i^- Q}^n \left(f(\partial_i^- Q) - f(Q) \right) &= \frac{1}{h_n} \sum_{y \in \partial^c Q} \partial_y f(Q) + O_n(1) \\ &= -\frac{1}{h_n} \sum_{y \in \partial Q} \partial_y f(Q) + O_n(1) \end{aligned}$$

et donc finalement :

$$\mathfrak{L}^n[f_n](D) = \mathfrak{L}^n f(Q) = O_n(1)$$

□

Le problème qui se pose maintenant en vue de montrer la convergence des processus discrets vers un processus continu, est qu'il n'existe pas en général de suite de fonction $(f_n)_{n \in \mathbb{N}}$ convergeant vers une fonction f , et telle que $\mathfrak{L}^n f_n$ converge uniformément vers $\mathfrak{L}f$. Nous ne pouvons donc pas appliquer exactement la même démarche que dans la section précédente. A l'heure actuelle, nous n'avons pas déterminé avec certitude de méthode pour contourner ce problème, mais nous allons tâcher de présenter une piste qui semble raisonnable.

Pour tout $n \in \mathbb{N}$, on note $D^n := (D_t^n)_{t \geq 0}$ la solution du problème de martingale associé à \mathfrak{L}^n , et :

$$\tau_N^n := \inf\{t \geq 0 : C(D_t^n) > N\}$$

où on a défini la fonction $C : \mathcal{G} \rightarrow \mathbb{N} \cup \{\infty\}$ qui à un élément de \mathcal{G} associe le nombre de ses composantes connexes.

Conjecture 6.2.4. *La suite $(D_{t \wedge \tau_N^n}^n)_{t \geq 0}$ admet une sous-suite qui converge vers un processus $(D_{t \wedge \tau_N})_{t \geq 0}$, solution du problème de martingale associé à \mathfrak{L} jusqu'au temps :*

$$\tau_N := \lim_{n \rightarrow \infty} \tau_N^n = \inf\{t \geq 0 : C(D_t) > N\}$$

Idee de démonstration. Pour un $n \in \mathbb{N}$ fixé, on définit deux processus D^+ et D^- , qui d'un état Q ne peuvent passer qu'à $V(Q)$ ou \mathring{Q} . On pose :

$$\begin{aligned} M^+ &:= \max_{Q \in \mathcal{D}^n} \left\{ \sum_{Q' \in \mathcal{V}^+(Q)} \mathfrak{L}_{Q, Q'}^n \right\} \sim \frac{1}{h_n^2} \\ m^+ &:= \min_{Q \in \mathcal{D}^n} \mathfrak{L}_{Q, \mathring{Q}}^n \sim \frac{1}{h_n^2} \\ M^- &:= \min_{Q \in \mathcal{D}^n} \mathfrak{L}_{Q, V(Q)}^n \sim \frac{1}{h_n^2} \\ m^- &:= \max_{Q \in \mathcal{D}^n} \left\{ \sum_{Q' \in \mathcal{V}^-(Q)} \mathfrak{L}_{Q, Q'}^n \right\} \sim \frac{1}{h_n^2} \end{aligned}$$

et on définit ces coefficients comme étant les taux de saut respectivement de Q à $V(Q)$ et de Q à \mathring{Q} pour D^+ (pour les deux premiers), puis pour D^- (pour les deux derniers). On couple ensuite ces processus de telle sorte que $D^- \subset D^n \subset D^+$ sur un intervalle de temps $[0, T]$ p.s. (ou plus exactement on couple D^n avec une famille de processus de même loi que D^+ et D^- , sinon le couplage ne tient que jusqu'au moment où D^n atteint $\mathcal{G} \setminus \mathcal{D}^n$). Enfin, on fait les mêmes calculs que pour le lemme 6.1.4 pour conclure à la compacité relative de la famille $(D^n)_n$.

Dans un deuxième temps, grâce à la convergence de $\mathfrak{L}^n f_n$ vers $\mathfrak{L}f$ montrée dans la proposition 6.2.3, et en justifiant que le temps passé en $g_i^{-1}(\mathcal{D}_i)^c$ sur tout intervalle $[0, T \wedge \tau_N]$ tend vers 0 lorsque i tend vers l'infini, on espère pouvoir montrer que :

$$\forall t > 0, \quad \int_0^{t \wedge \tau_N^n} \mathfrak{L}^n f_n(D_s^n) dD_s^n - \int_0^{t \wedge \tau_N} \mathfrak{L}f(D_s^n) dD_s^n \xrightarrow{\mathbb{P}} 0$$

ce qui, par le corollaire 3.3, chapitre 3 de Ethier et Kurtz [29], implique :

$$\int_0^{t \wedge \tau_N^n} \mathfrak{L}^n f_n(D_s^n) dD_s^n \xrightarrow{\mathcal{L}} \int_0^{t \wedge \tau_N} \mathfrak{L}f(D_s) dX_s$$

Par la convergence des espérances, on peut alors en déduire que le processus :

$$f(D_t \wedge \tau_N) - f(D_0) - \int_0^{t \wedge \tau_N} \mathfrak{L}f(D_s) dD_s$$

est une martingale. Ceci ne garantit pas en revanche que le processus D est markovien, et il faudrait pour ça, par exemple, montrer l'unicité de la solution du problème de martingale associé à \mathfrak{L} (cf théorème 4.2, chapitre 4 de [29]). \square

Remarque 6.2.5. *Observons que τ_N est le premier temps de sortie de \mathcal{G}^N , qui est fermé, donc $C(D_{\tau_N}) \leq N$.*

Ce résultat n'est toujours pas suffisant, car nous n'avons pas montré que le nombre de composantes connexes n'explose pas en temps fini. En étudiant les trajectoires des processus D^n par un couplage avec des processus de vie et de mort, on espère en fait pouvoir prouver la conjecture suivante :

Conjecture 6.2.6. *La famille de processus $(D^n, C(D^n))_n$ à valeurs dans $\mathcal{G} \times \mathbb{N}$ est relativement compacte.*

Ceci implique notamment que $\tau_N \rightarrow \infty$ presque sûrement quand $N \rightarrow \infty$, et résout le problème du nombre de composantes connexes.

une question qui reste à se poser est celle de l'existence d'un nombre suffisant de fonctions $f \in \mathcal{D}(\mathfrak{L})$ telles que $\mathfrak{L}f$ soit continue. En particulier, on sait que les fonctionnelles de la forme F_f , $f \in \mathcal{F}$ ne rentrent pas dans cet ensemble.

Conjecture 6.2.7. *Soient $(D_t)_{t \geq 0}$ un processus à valeurs dans \mathcal{G} , et pour tout $f \in \mathcal{C}^2(\mathcal{G})$ le processus M^f défini par :*

$$\forall t \geq 0, \quad M_t^f := f(D_t) - f(D_0) - \int_0^t \mathfrak{L}[f](D_s) ds$$

Si M^g est une martingale (relativement à la filtration engendrée par D) pour toute fonction $g \in \mathcal{D}(\mathfrak{L})$ telle que $\mathfrak{L}g$ est continue sur \mathcal{G} , alors M^f l'est aussi, pour tout $f \in \mathcal{D}(\mathfrak{L})$.

Démonstration. Il s'agit de trouver, pour tout $f \in \mathcal{D}(\mathfrak{L})$, une suite $(g_n)_{n \in \mathbb{N}}$ de $\mathcal{D}(\mathfrak{L})$, convergeant uniformément vers f , telle que la suite $(\mathfrak{L}g_n)_{n \in \mathbb{N}}$ est uniformément bornée et converge simplement vers $\mathfrak{L}f$, et que pour tout $n \in \mathbb{N}$, $\mathfrak{L}g_n$ est continue. \square

Bibliographie

- [1] David Aldous and Persi Diaconis. Shuffling cards and stopping times. *American Mathematical Monthly*, 93(5) :333–348, 1986.
- [2] David Aldous and Persi Diaconis. Strong uniform times and finite random walks. *Advances in Applied Mathematics*, 8(1) :69–97, 1987.
- [3] William J. Anderson. *Continuous-time Markov chains*. Springer Series in Statistics : Probability and its Applications. Springer-Verlag, New York, 1991. An applications-oriented approach.
- [4] Svetlana V. Anulova. Diffusion processes with singular characteristics. In *Stochastic differential systems (Proc. IFIP-WG 7/1 Working Conf., Vilnius, 1978)*, volume 25 of *Lecture Notes in Control and Information Sci.*, pages 264–269. Springer, Berlin-New York, 1980.
- [5] Dominique Bakry, Ivan Gentil, and Michel Ledoux. *Analysis and geometry of Markov diffusion operators*, volume 348 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Cham, 2014.
- [6] Philippe Barbe and Michel Ledoux. *Probabilité*. Collection Enseignement sup. Mathématiques. EDP Sciences, 2007.
- [7] Alain Berberian. La cité de la peur, une comédie familiale. *Les Nuls*, 1994.
- [8] Patrick Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics : Probability and Statistics. John Wiley & Sons, Inc., New York, second edition, 1999. A Wiley-Interscience Publication.
- [9] Hubert E. Bray. Elementary properties of the Stieltjes integral. *Ann. of Math. (2)*, 20(3) :177–186, 1919.
- [10] Pierre Brémaud. *Markov chains*, volume 31 of *Texts in Applied Mathematics*. Springer-Verlag, New York, 1999. Gibbs fields, Monte Carlo simulation, and queues.
- [11] Alain Bretto, Alain Faisant, and François Hennecart. *Éléments de théorie des graphes*. Collection IRIS. [IRIS Collection]. Springer, Paris, 2012.
- [12] Mark Brown and Yi-Shi Shao. Identifying coefficients in the spectral representation for first passage time distributions. *Probability in the Engineering and Informational Sciences*, 1 :69 – 74, January 1987.
- [13] Philippe Carmona, Frédérique Petit, and Marc Yor. Beta-gamma random variables and intertwining relations between certain Markov processes. *Revista Matemática Iberoamericana*, 14(2) :311–367, 1998.
- [14] Djalil Chafaï and Aldéric Joulin. Intertwining and commutation relations for birth–death processes. *Bernoulli*, 19(5A) :1855–1879, November 2013.
- [15] Bertrand Cloez and Claire Delplancke. Intertwinings and Stein’s magic factors for birth-death processes. *ArXiv e-prints*, September 2016.
- [16] Donald L. Cohn. *Measure theory*. Birkhäuser, Boston, Mass., 1980.

- [17] Guillaume Copros. Existence condition of strong stationary times for continuous time markov chains on discrete graphs. *Journal of Theoretical Probability*, February 2017.
- [18] Konele Coulibaly-Pasquier and Laurent Miclo. On the evolution by duality of domains on manifolds. Work in progress.
- [19] Claude Dellacherie and Paul-André Meyer. *Probabilités et potentiel*. Hermann, Paris, 1975. Chapitres I à IV, Édition entièrement refondue, Publications de l'Institut de Mathématique de l'Université de Strasbourg, No. XV, Actualités Scientifiques et Industrielles, No. 1372.
- [20] Persi Diaconis. *Group representations in probability and statistics*, volume 11 of *Institute of Mathematical Statistics Lecture Notes—Monograph Series*. Institute of Mathematical Statistics, Hayward, CA, 1988.
- [21] Persi Diaconis. The cutoff phenomenon in finite Markov chains. *Proceedings of the National Academy of Sciences of the United States of America*, 93(4) :1659–1664, 1996.
- [22] Persi Diaconis and James A. Fill. Examples for the theory of strong stationary duality with countable state spaces. *Probability in the Engineering and Informational Sciences*, 4(2) :157–180, 1990.
- [23] Persi Diaconis and James A. Fill. Strong stationary times via a new form of duality. *The Annals of Probability*, 18(4) :1483–1522, 1990.
- [24] Persi Diaconis and Laurent Miclo. On times to quasi-stationarity for birth and death processes. *Journal of Theoretical Probability*, 22(3) :558–586, 2009.
- [25] Persi Diaconis and Laurent Saloff-Coste. Separation cut-offs for birth and death chains. *The Annals of Applied Probability*, 16(4) :2098–2122, 2006.
- [26] Persi Diaconis and Mehrdad Shahshahani. Generating a random permutation with random transpositions. *Probability Theory and Related Fields*, 57 :159–179, 01 1981.
- [27] Joseph L. Doob. *Stochastic processes*. John Wiley & Sons, Inc., New York ; Chapman & Hall, Limited, London, 1953.
- [28] Eugene B. Dynkin. *Markov processes*. Springer Berlin Heidelberg, Berlin, Heidelberg, 1965.
- [29] Stewart N. Ethier and Thomas G. Kurtz. *Markov processes*. Wiley Series in Probability and Mathematical Statistics : Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1986. Characterization and convergence.
- [30] William Feller. The parabolic differential equations and the associated semi-groups of transformations. *Annals of Mathematics. Second Series*, 55 :468–519, 1952.
- [31] James A. Fill. Time to stationarity for a continuous-time Markov chain. *Probability in the Engineering and Informational Sciences*, 5(1) :61–76, 1991.
- [32] James A. Fill. Strong stationary duality for continuous-time Markov chains. I. Theory. *Journal of Theoretical Probability*, 5(1) :45–70, 1992.
- [33] James A. Fill. An interruptible algorithm for perfect sampling via Markov chains. *The Annals of Applied Probability*, 8(1) :131–162, 1998.
- [34] James A. Fill. On hitting times and fastest strong stationary times for skip-free and more general chains. *Journal of Theoretical Probability*, 22(3) :587–600, 2009.
- [35] James A. Fill. The passage time distribution for a birth-and-death chain : strong stationary duality gives a first stochastic proof. *Journal of Theoretical Probability*, 22(3) :543–557, 2009.

- [36] James A. Fill and Jonas Kahn. Comparison inequalities and fastest-mixing Markov chains. *The Annals of Applied Probability*, 23(5) :1778–1816, 2013.
- [37] James A. Fill and Vince Lyzinski. Strong stationary duality for diffusion processes. *Journal of Theoretical Probability*, 29(4) :1298–1338, December 2016.
- [38] James A. Fill, Motoya Machida, Duncan J. Murdoch, and Jeffrey S. Rosenthal. Extension of Fill’s perfect rejection sampling algorithm to general chains (extended abstract). In *Monte Carlo methods (Toronto, ON, 1998)*, volume 26 of *Fields Inst. Commun.*, pages 37–52. Amer. Math. Soc., Providence, RI, 2000.
- [39] Mark Freidlin and Shuenn-Jyi Sheu. Diffusion processes on graphs : stochastic differential equations, large deviation principle. *Probability Theory and Related Fields*, 116(2) :181–220, 2000.
- [40] Mark I. Freidlin and Alexander D. Wentzell. Diffusion processes on graphs and the averaging principle. *The Annals of Probability*, 21(4) :2215–2245, 1993.
- [41] Yu Gong, Yong-Hua Mao, and Chi Zhang. Hitting time distributions for denumerable birth and death processes. *Journal of Theoretical Probability*, 25(4) :950–980, 2012.
- [42] David Griffeath. A maximal coupling for Markov chains. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 31 :95–106, 1974/75.
- [43] Einar Hille. *Functional Analysis and Semi-Groups*. American Mathematical Society Colloquium Publications, vol. 31. American Mathematical Society, New York, 1948.
- [44] Nobuyuki Ikeda and Shinzo Watanabe. *Stochastic differential equations and diffusion processes*, volume 24 of *North-Holland Mathematical Library*. North-Holland Publishing Co., Amsterdam ; Kodansha, Ltd., Tokyo, second edition, 1989.
- [45] Samuel Karlin and James McGregor. Coincidence properties of birth and death processes. *Pacific Journal of Mathematics*, 9(4) :1109–1140, 1959.
- [46] Samuel Karlin and Howard M. Taylor. *A first course in stochastic processes*. Academic Press, 2nd edition, 1975.
- [47] Samuel Karlin and Howard M. Taylor. *A second course in stochastic processes*. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York-London, 1981.
- [48] David G. Kendall. Stochastic processes occurring in the theory of queues and their analysis by the method of the imbedded markov chain. *Ann. Math. Statist.*, 24(3) :338–354, 09 1953.
- [49] Harold J. Kushner and Paul Dupuis. *Numerical Methods for Stochastic Control Problems in Continuous Time*, volume 24. Springer, 01 2001.
- [50] Michel Ledoux. The geometry of Markov diffusion generators. *Annales de la Faculté des Sciences de Toulouse. Mathématiques. Série 6*, 9(2) :305–366, 2000. Probability theory.
- [51] Antoine Lejay. On the constructions of the skew Brownian motion. *Probab. Surv.*, 3 :413–466, 2006.
- [52] David A. Levin, Yuval Peres, and Elizabeth L. Wilmer. *Markov chains and mixing times*. American Mathematical Society, Providence, RI, 2009. With a chapter by James G. Propp and David B. Wilson.
- [53] Paul Lévy. *Théorie de l’addition des variables aléatoires*. Collection des monographies des probabilités. Gauthier-Villars, 1937.
- [54] Thomas M. Liggett. *Continuous time Markov processes*, volume 113 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2010. An introduction.

- [55] Pawel Lorek and Ryszard Szekli. Strong stationary duality for Möbius monotone Markov chains. *Queueing Systems. Theory and Applications*, 71(1-2) :79–95, 2012.
- [56] Pawel Lorek and Ryszard Szekli. Strong stationary duality for Möbius monotone Markov chains : examples. *Probability and Mathematical Statistics*, 36(1) :75–97, 2016.
- [57] Pawel Lorek, Filip Zagórski, and Michal Kulis. Strong stationary times and its use in cryptography. *ArXiv e-prints*, September 2017.
- [58] Vince Lyzinski. *Intertwinings, interlacing eigenvalues, and strong stationary duality for diffusions*. ProQuest LLC, Ann Arbor, MI, 2013. Thesis (Ph.D.)—The Johns Hopkins University.
- [59] Yong-Hua Mao and Pan Zhao. Strong stationary duality for discrete time Möbius monotone Markov chains on \mathbb{Z}_+^d . *Statistics & Probability Letters*, 123 :183–192, 2017.
- [60] Peter Matthews. Mixing rates for a random walk on the cube. *Society for Industrial and Applied Mathematics. Journal on Algebraic and Discrete Methods*, 8(4) :746–752, 1987.
- [61] Peter Matthews. A strong uniform time for random transpositions. *Journal of Theoretical Probability*, 1(4) :411–423, October 1988.
- [62] Peter Matthews. Some sample path properties of a random walk on the cube. *Journal of Theoretical Probability*, 2(1) :129–146, 1989.
- [63] Laurent Miclo. Processus de markov inhomogènes finis.
- [64] Laurent Miclo. On ergodic diffusions on continuous graphs whose centered resolvent admits a trace. *Journal of Mathematical Analysis and Applications*, 437(2) :737–753, 2016.
- [65] Laurent Miclo. Strong stationary times for one-dimensional diffusions. *Annales de l’Institut Henri Poincaré Probabilités et Statistiques*, 53(2) :957–996, 2017.
- [66] Jesper Møller and Katja Schladitz. Extensions of Fill’s algorithm for perfect simulation. *Journal of the Royal Statistical Society. Series B. Statistical Methodology*, 61(4) :955–969, 1999.
- [67] Ben Morris and Yuval Peres. Evolving sets, mixing and heat kernel bounds. *Probability Theory and Related Fields*, 133(2) :245–266, 2005.
- [68] Evita Nestoridis. Optimal strong stationary times for random walks on the chambers of a hyperplane arrangement. *ArXiv e-prints*, July 2017.
- [69] James R. Norris. *Markov chains*, volume 2 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, 1998. Reprint of 1997 original.
- [70] Pierre Patie, Mladen Savov, and Yixuan Zhao. Intertwining, Excursion Theory and Krein Theory of Strings for Non-self-adjoint Markov Semigroups. *ArXiv e-prints*, June 2017.
- [71] James W. Pitman. One-dimensional Brownian motion and the three-dimensional Bessel process. *Advances in Applied Probability*, 7(3) :511–526, 1975.
- [72] Nicolai I. Portenko. Stochastic differential equations with generalized drift vector. *Theory of Probability & Its Applications*, 24(2) :338–353, 1980.
- [73] Yuri V. Prokhorov. Convergence of random processes and limit theorems in probability theory. *Theory of Probability & Its Applications*, 1(2) :157–214, 1956.
- [74] James G. Propp and David B. Wilson. Exact sampling with coupled Markov chains and applications to statistical mechanics. In *Proceedings of the Seventh International Conference on Random Structures and Algorithms (Atlanta, GA, 1995)*, volume 9, pages 223–252, 1996.

- [75] Daniel Revuz and Marc Yor. *Continuous martingales and Brownian motion*, volume 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, third edition, 1999.
- [76] Leonard C. G. Rogers and James W. Pitman. Markov functions. *The Annals of Probability*, 9(4) :573–582, 1981.
- [77] Gian-Carlo Rota. On the foundations of combinatorial theory. I. Theory of Möbius functions. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 2 :340–368 (1964), 1964.
- [78] Walter Rudin. *Real and complex analysis*. McGraw-Hill Book Co., New York-Toronto, Ont.-London, 1966.
- [79] Wolfgang M. Schmidt. *Diophantine approximations and Diophantine equations*, volume 1467 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1991.
- [80] Lynn A. Steen and J. Arthur Seebach, Jr. *Counterexamples in topology*. Holt, Rinehart and Winston, Inc., New York-Montreal, Que.-London, 1970.
- [81] Daniel W. Stroock and S. R. S. Varadhan. Diffusion processes with boundary conditions. *Comm. Pure Appl. Math.*, 24 :147–225, 1971.
- [82] Daniel W. Stroock and S. R. Srinivasa Varadhan. Diffusion processes with continuous coefficients. II. *Communications on Pure and Applied Mathematics*, 22 :479–530, 1969.
- [83] Daniel W. Stroock and S. R. Srinivasa Varadhan. *Multidimensional diffusion processes*. Classics in Mathematics. Springer-Verlag, Berlin, 2006. Reprint of the 1997 edition.
- [84] Satoshi Takanobu. On the uniqueness of solutions of stochastic differential equations with singular drifts. *Publ. Res. Inst. Math. Sci.*, 22(5) :813–848, 1986.
- [85] Satoshi Takanobu. On the existence of solutions of stochastic differential equations with singular drifts. *Probab. Theory Related Fields*, 74(2) :295–315, 1987.
- [86] Satoshi Takanobu and Shinzo Watanabe. On the existence and uniqueness of diffusion processes with Wentzell’s boundary conditions. *J. Math. Kyoto Univ.*, 28(1) :71–80, 1988.
- [87] Elke Thönnies. Perfect simulation of some point processes for the impatient user. *Advances in Applied Probability*, 31(1) :69–87, 1999.
- [88] Hermann Thorisson. Future independent times and Markov chains. *Probability Theory and Related Fields*, 78(1) :143–148, 1988.
- [89] Erik A. Van Doorn. *Stochastic monotonicity and queueing applications of birth-death processes*, volume 4 of *Lecture Notes in Statistics*. Springer-Verlag, New York-Berlin, 1981.
- [90] Alexander D. Venttsel’. On boundary conditions for multidimensional diffusion processes. *Theory of Probability & Its Applications*, 4(2) :164–177, 1959.
- [91] Stephen Willard. *General Topology*. Addison-Wesley, 1st edition, 1970.
- [92] Marc Yor. *Intertwinings of Bessel processes*. Department of Statistics, University of California, 1988.
- [93] Kosaku Yosida. On the differentiability and the representation of one-parameter semi-group of linear operators. *Journal of the Mathematical Society of Japan*, 1(1) :15–21, September 1948.