

Sur l'utilisation des relations d'entrelacement dans l'étude des générateurs de Markov auto-adjoints. Application aux inégalités spectrales et fonctionnelles et à l'analyse de sensibilité

Clément Steiner

▶ **To cite this version:**

Clément Steiner. Sur l'utilisation des relations d'entrelacement dans l'étude des générateurs de Markov auto-adjoints. Application aux inégalités spectrales et fonctionnelles et à l'analyse de sensibilité. Mathématiques [math]. Université Toulouse 3 (Paul Sabatier), 2022. Français. NNT : . tel-03727812

HAL Id: tel-03727812

https://theses.hal.science/tel-03727812

Submitted on 19 Jul 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université Fédérale
Toulouse Midi-Pyrénées

THÈSE

En vue de l'obtention du
DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE
Délivré par l'Université Toulouse 3 - Paul Sabatier

Présentée et soutenue par

Clément STEINER

Le 8 juillet 2022

Sur l'utilisation des relations d'entrelacement dans l'étude des générateurs de Markov auto-adjoints. Application aux inégalités spectrales et fonctionnelles et à l'analyse de sensibilité

Ecole doctorale : **EDMITT - Ecole Doctorale Mathématiques, Informatique et Télécommunications de Toulouse**

Spécialité : **Mathématiques et Applications**

Unité de recherche :
IMT : Institut de Mathématiques de Toulouse

Thèse dirigée par
Patrick CATTIAUX et Aldéric JOULIN

Jury

M. Djalil CHAFAï, Rapporteur
M. Pierre-André ZITT, Rapporteur
Mme Clémentine PRIEUR, Examinatrice
M. Nathael GOZLAN, Examinateur
M. Patrick CATTIAUX, Directeur de thèse
M. Aldéric JOULIN, Co-directeur de thèse

Biais du survivant : biais de sélection consistant à surévaluer les chances de succès d'une initiative en concentrant l'attention sur les sujets ayant réussi mais qui sont des exceptions statistiques plutôt que des cas représentatifs.

Remerciements

Bonjour lecteur. Comme tu peux le constater, cette section "remerciements" est particulièrement touffue. Afin que tu puisses t'y retrouver, elle a été découpée en sections, qui te permettront de localiser le passage qui t'intéresse (et de te raccrocher au plus vite au fil de la soutenance). Attention, certaines personnes sont mentionnées plusieurs fois dans des passages différents ! Et ne fais pas trop cas de l'ordre d'apparition des uns et des autres, celui-ci est plus le résultat des fluctuations de ma mémoire que d'une quelconque volonté.

1 Comment j'ai aimé les maths

Cette première section fait la part belle à tous ceux que j'ai pu croiser dans le monde mathématique ces dernières années, que ce soit comme camarades étudiants, collègues enseignants ou enseignants-chercheurs, et bien-sûr les amis de galère.

1.1 Aux membres du jury

Comme il se doit, je ne peux commencer ces remerciements sans m'adresser d'abord aux membres du jury présents aujourd'hui (ici ou à distance). D'abord un très grand merci à Djalil et Pierre-André d'avoir accepté de rapporter ma thèse, et de l'avoir fait avec tant de bienveillance. Merci pour vos remarques nombreuses et constructives, et bravo d'être venu à bout de ces 150 pages parfois un peu verbeuses, je l'avoue. C'est en partie grâce à vous que j'appréhende cette soutenance avec confiance et sérénité. Un grand merci également à Clémentine et Nathaël, qui m'ont fait l'honneur d'accepter de faire partie de mon jury. Vous apportez à celui-ci une diversité que j'apprécie beaucoup. Je remercie également Olivier d'avoir accepté l'invitation, ainsi que pour l'année de travail écoulée. Tu m'as aidé à sortir de ma zone de confort mathématique ("confort", c'est peut-être un bien grand mot), ce qui n'était pas gagné ! Chapeau, et merci. Enfin, je ne peux pas oublier mes directeurs. Merci Patrick d'avoir accepté d'encadrer ma thèse. Les événements ont fait que nous n'avons pas pu travailler autant ensemble que prévu, mais ta présence et ton écoute m'ont toujours accompagné durant ces quatre années. Je

te remercie également pour ta participation à ma formation. Tu fais partie de ces enseignants qui m'ont fait entrevoir la richesse et la diversité des mathématiques, en toute humilité et en toute simplicité. Le mot de la fin (de cette section !) va naturellement pour Aldéric. Merci d'avoir tenu pendant ces quatre années, parfois difficiles, toujours intenses. La thèse est une aventure humaine complexe, avec ses hauts très hauts et ses bas très bas, et bien malin celui qui pourrait en prédire les péripéties. Je suis heureux et fier d'être ici aujourd'hui pour pouvoir dire qu'on y est arrivé. Je pense que nous avons tous deux beaucoup appris, mathématiquement et humainement, et je garderai toujours ce bagage avec moi, quels que soient les chemins que je serai amené à emprunter. Pax melior est quam iustissimum bellum.

1.2 À la jeunesse

Que ceux qui ne figurent pas dans cette section n'en prennent pas ombrage, mais il faut bien mettre la barre de la jeunesse quelque-part. En l'occurrence, ce sera une jeunesse symbolique, la jeunesse qui rame, mais aussi celle qui espère.

À mes collègues et amis doctorants

C'est probablement une platitude terrible que de dire que la thèse n'est pas une aventure en solitaire, et que rien n'aurait été possible sans mes compagnons de galère. Une platitude peut-être, mais cela reste vrai néanmoins, et il me faut leur dire à quel point leur soutien m'a été précieux pendant ces quatre années.

Ma première pensée va naturellement à mes cobureaux de l'INSA, anciens et présents. D'abord Jessica, on s'est suivis de près, du début idyllique du stage à la fin tourmentée de la thèse. Peu de choses a priori auraient laissé penser que nous parcourrions tout ce chemin ensemble, et je suis très heureux de t'avoir eue à mes côté. Merci d'avoir partagé les joies et les peines, les moments forts comme les petites choses du quotidien. Je te souhaite de trouver ta place dans ton nouveau monde, mais je ne doute pas que tu y brilleras de nouveau ! À Morgane, dont le covid nous aura injustement privée, merci d'avoir su rester l'îlot de sanité dans ce monde de plus en plus flou. Ton sens des réalités m'aura probablement permis plus d'une fois de garder les pieds sur la berge, quand la sombre écume appelait de sa voix sonore. Sans compter sur les connaissances de Python que tu m'as aidé à acquérir, tu m'as mis le pied à l'étrier, et pour ça je te remercie. Un petit coucou à Raphaël aussi, pour ses questions de géométrie toujours bienvenues et jamais idiotes ! J'espère que vous aussi trouvez vos marques dans votre nouvelle vie. Spaciba Ievgueniia ! J'ai été très content de partager cette année et demie avec toi dans le bureau, entre cours de russe, astuces de maths 2 et gâteaux divers et variés, je te souhaite bon courage pour la dernière année, en espérant que nos chemins se croisent à nouveau ! Petit coucou à Thibault, avec qui j'ai pu partager aussi un petit bout d'aventure, et qui... jouit... d'un beau cadre de vie (il eut fallu que j'osasse). Franck, merci pour la fraîcheur que tu as apportée au bureau par ton bref passage ! J'espère que le Maroc est à la hauteur de tes espérances et que la fin de thèse se passera pour le mieux. Enfin Alban, je continue à croire que tu es une de mes plus belles réussites de

ces dernières années. Je te remercie de m'avoir bien aidé avec Python, et pour toutes les conversations qu'on a pu avoir depuis trois ans. Je ne sais pas si je retrouverai un collègue comme toi, je te souhaite de réussir à mener à bien tes projets. Et si jamais, la fonction publique française t'attend les bras ouverts (pour l'instant) ! Je finis par un dernier petit coucou à Cyriel, si tu tombes sur ces lignes, je pense à nos interminables discussions, heureusement que nous ne sommes pas restés trop longtemps au même endroit... Bonne continuation, où que tu sois !

Mon regard se tourne maintenant vers le couloir d'en face. Merci Julie et Iain pour les apéros, raclettes et autres aprems crêpes. Et bon courage pour la fin, je pense que vous tenez le bon bout, et n'oubliez jamais, le pire est toujours devant vous ! Bon courage à Nathanaël aussi, l'aventure ne fait que commencer, mais tu seras bien entouré. Un petit coucou aussi à Josué, tu gères, gros respect ! And a wave to Armando, hope everything goes well. Et il ne faudrait pas oublier nos collègues du bout du couloir. Merci Thibault pour ces quelques années de pauses café et de discussions parfois intenses, toujours intéressantes, et un coucou à Hippolyte, toujours trop rare à la pause ! Merci aussi à Mahmoud pour les discussions sans fin, on est d'accord sur rien mais toujours prêts à en parler, j'aime ça. Je n'oublie bien-sûr pas ceux qui nous ont quittés il y a longtemps déjà, Maria, Léo, Valentin, Willy, Firas, Anouar, vous avez tous laissé votre empreinte d'une façon ou d'une autre, et je suis heureux d'avoir pu marcher dedans à mon tour. J'espère que vous allez bien chacun où que vous soyez.

Et parce qu'il n'y a probablement pas plus galérien que l'ATER, je remercie Julie pour son passage éclair mais mémorable au GMM, pour les bières, les soirées jeux et les restos vegans ! Un petit coucou à Joe aussi, j'espère que ton post-doc se passe bien. Et bien entendu je n'oublie pas Alexandre, compagnon de galère jusque dans les cours de maths 2, tu es lancé vers la soutenance, ne lâche rien !

Mais il n'y a pas que l'INSA à l'IMT, et je ne peux pas oublier mes collègues de la fac. Les premiers auxquels je pense sont naturellement les anciens, ceux de l'âge d'or, du temps de l'équipe ESP, du temps d'avant le Covid et du temps où la fin de la thèse était encore loin... Merci à l'équipe de choc des ESP, Maylis (et ça, c'est important !), Valentin, Kiki (Monsieur le Directeur de l'UT4), William (mon grand-frère de thèse, bien plus balèze que moi), merci de m'avoir intégré et mis le pied à l'étrier. Je n'aurais pas pu rêver de meilleur accueil. Un petit coucou à Hugo en passant, et même à Antoine, nous ne nous sommes jamais rencontrés, mais ta photo trône éternellement sur le petit placard électrique du bureau 201. Le roi est mort, vive le roi ! Et évidemment, à la fois un ancien mais toujours actuel, je n'oublie pas ce cher Fabien. Un espèce d'énergumène comme toi, je ne pense pas en revoir de sitôt, merci pour tes blagues bien pourries et ta gloutonnerie proverbiale, je me dis que mon propre cas n'est peut-être pas désespéré, en fin de compte ;) Et bien-sûr on n'oublie pas notre cher président, le plus ESP des MIP, Baptiste Fedele en personne ! Je suis content que les événements t'aient ramené auprès de nous, on aura passé du sacrément bon temps. À toi et Maylis (et Jade, qui malgré ses nombreuses qualités n'a pas celle d'être ESP), j'espère que je me serai montré digne du séminaire étudiant. Il a survécu au covid, il a survécu à des exposés incompréhensibles, à des goûters bien trop chers pour ce que c'est (toute mon inimitié va à ce sujet au

CNRS, je vous regarde mal), longue vie à lui !

Et tant qu'on parle des vieux, on n'oubliera les collègues MIP et Picard. D'abord Jojo, dont l'engagement et la ténacité n'ont cessé de m'impressionner, et Alexis, j'ai été très heureux de partager l'organisation du séminaire avec toi. Ça me gêne de te compter dans les vieilles, mais merci Michèle, juste d'exister, et d'être venue dynamiter les frontières des équipes pour créer un groupe soudé et ouvert. Si tu arrêtes les maths, tu pourras te reconvertir dans l'animation et l'organisation. J'espère d'ailleurs que ta soutenance sera à la hauteur de tes attentes ;) Idem, merci à l'inénarrable Anthony, pour être ce que tu es, en toute décomplexion. Mes salutations au doyen Florian et à JM, respect pour vous être accrochés vaille que vaille. Respect aussi à Dr Domi, tu as le feu sacré, et ça c'est beau. Je n'oublie pas non plus Sonny, mais j'y reviendrai...

On avance un peu dans le temps maintenant, pour se tourner vers les vieux un peu moins vieux. À commencer par mes cobureaux de l'IMT, Maxime (j'espère que tout va bien pour toi), Gabriel (la première fois que je t'ai entendu parler de copules, j'ai ri, et me voilà en train d'en mettre dans ma thèse...) et Gach (gros respect aussi pour la décision difficile mais courageuse que tu as sue prendre), on était peut-être pas le bureau le plus vivant, mais j'ai pour ma part beaucoup aimé ces moments en votre compagnie ! Toujours dans le même couloir, je n'oublie pas Eva (mais j'y reviendrai aussi !), et en bas Trang et Camille. Vous faites des stats, mais je vous aime bien quand-même !

Toujours dans les vieux moins vieux, un petit coucou à François et à Mathias, merci pour les discussions à bâtons rompus et respect pour avoir tenu jusqu'au bout. Et les Picards aussi ont des vieux moins vieux, comme Dimitri et Elena, je vous souhaite le meilleur !

On peut maintenant terminer notre voyage vers le présent avec les petits jeunes. Vous verrez, pour citer Fabien, quand vous aurez mon âge, tout le monde vous paraîtra jeune. Pour ne pas casser le rythme, on va commencer par les collègues probabilistes et statisticiens. Les plus vieux parmi les jeunes, Sofiane, Jens, Jordan, Clément et Medhi, vous formez une relève aussi hétéroclite qu'intéressante ! Je garde une petite place spéciale pour Laetitia, merci pour tout ce temps passé à discuter, à se remonter le moral et à parler citron et chocolat. Et gros respect pour ton engagement envers les doctorants, et les autres en général. J'ai été honoré de te côtoyer pendant trois ans. Vous tous avez réussi à maintenir le fil de la vie dans un couloir orphelin, et c'est déjà beaucoup.

Les EDP et nos amis fondas ne sont pas en reste concernant la jeunesse. D'abord Corentin, j'ai été très heureux d'organiser à tes côtés le séminaire étudiant. Il est en train de changer, et c'est normal, place au sang neuf ! Merci pour ça à Viviana et Javier, d'avoir accepté de s'occuper de ce vieux et noble bébé. Merci aussi à Paola, Perla, Louis pour les bonnes soirées, vous formez aussi une super équipe ! Bon courage enfin à Irène, Solène, William, que j'ai pu côtoyer avec plaisir, de près ou de loin, pendant ces dernières années.

Enfin, c'est avec une pointe de nostalgie que je m'adresse à tous les "petits nouveaux", maintenant plus si nouveaux que ça. Il n'y a finalement pas si long depuis l'époque où j'étais à votre place, mais le temps fait son œuvre, encore plus durant ce sacerdoce qu'est la thèse, et vous voilà maintenant bien loin de moi. Ou suis-je loin

de vous ? Cela n'a guère plus d'importance. Bon courage à vous tous, quels que soient vos parcours ou vos projets. J'ai apprécié rencontrer certains d'entre vous, souvent regretté de ne pas avoir pu mieux vous connaître. Mais il faut savoir dire au revoir, et j'espère que vous serez la génération qui verra la situation de la recherche publique enfin s'améliorer. Elle n'en prend pas le chemin, et nombreux sont ses adversaires, qui sacrifieraient les statuts des uns pour de bien futiles raisons. Ne soyez pas d'aveugles gardiens du temple, mais n'arrêtez jamais de vous battre pour ce qui en vaut le coup. Que l'intérêt commun soit votre boussole, et la foi en l'humanité votre vaisseau. Perdez l'un des deux, et les abîmes vous tendront les bras. Hasta siempre !

À tous ceux que j'ai croisés sur le chemin de la connaissance

On l'oublie souvent, mais le doctorat n'est jamais que l'aboutissement, et pas toujours la partie la plus agréable, d'un parcours bien plus long. Et celui-ci est, plus encore que la thèse, un chemin collectif. Je voudrais donc remercier ici ceux qui m'ont accompagné depuis 2012, année de notre premier pas dans la cour des grands.

Et que l'on me pardonne de commencer par les plus chers à mon cœur, l'équipe de la première heure, les vrais, les meilleurs. Ce sont bien vous, Charlie (et Morgan !), Maxime, Youlie, Rémy (et Cécile !), Rémi, Hacheuh (et Méli !) et Marie Denat (et Vilius !). Merci infiniment pour ce chemin parcouru tous ensemble depuis les premières heures et toujours aujourd'hui. Il y aurait probablement trop à dire pour être exhaustif, mais je dois mentionner les réveillons, la Corse, les expériences culinaires, les fêtes d'anniversaire et tant d'autres choses... Vous comptez aujourd'hui beaucoup pour moi, et je n'aurais probablement pas pu rêver meilleure équipe pour les dix dernières années. Nos chemins maintenant on pris des tournures différentes, mais pour moi vous êtes toujours là. Merci <3.

Je pense aussi à ceux qui nous ont quittés en route, surtout Quentin et Laurent. J'espère que vos chemins n'ont pas été trop tortueux, et que vous allez bien !

Et comment ne pas évoquer le temps béni de la licence sans penser à Neige. Une scientifique exceptionnelle, retenez bien son nom (il est assez inhabituel, ça ne devrait pas être trop difficile). Un parcours sans faute, une curiosité sans limites, l'envie d'apprendre et de transmettre chevillée au corps, autant de qualités qui forcent le respect et l'admiration. J'ai été très heureux de partager ces premières années avec toi, et je n'ai aucun doute que tu réussiras au-delà de toutes les espérances ! Je repense encore à notre exposé sur les orages, ou à cette nuit des chercheurs mémorable. Et merci pour l'appart ;)

Enfin, la licence n'aurait pas été la même sans tous les camarades des parcours spéciaux. Un deuxième coucou à Eva, j'ai vraiment apprécié toutes ces années à tes côtés, de la licence à la thèse, j'espère que tu trouves ton bonheur à Paris (même si ça semble antinomique). Et bien-sûr tous les autres, mais le temps est passé et je crains ne pas me souvenir de la moitié d'entre vous moitié moins que je le souhaiterais, et me contenterai donc ici d'un salut collectif ! Chacun de vous a contribué à rendre ce parcours vivant, et à nous mener là où nous sommes aujourd'hui.

Un coucou également à l'équipe du M1, et merci pour tous les moments partagés ! Mais je voulais surtout faire un saut dans le futur pour arriver à l'année de l'agreg. Charnière pour beaucoup d'entre nous, c'est probablement celle où le travail d'équipe aura réalisé son plein potentiel. Et je ne pourrais faire l'affront d'oublier ces camarades de galère. Mes première pensées vont aux vrais du M2, Mathilde (et Bruno !), Gazo, Cassagnes, Willetts, Sophie et Sandra, même si on est partis sur des bases un peu compliquées, je suis heureux d'avoir appris à vous connaître, et je vous remercie de continuer à faire vivre ce lien, malgré mon manque d'assiduité. Merci aussi pour les heures de travail intense, et une pensée particulière pour le petit ventilateur de Mathilde, alors que nous nous obstinions à la bibliothèque jusqu'au dernier jour ! Je n'oublie pas Romain, Enzo et un deuxième coucou à Gach. Les autres collègues de l'agreg ne sont pas en reste, Quentin, Simon et Paul (au plaisir de se recroiser autour d'un accordéon !), vous m'avez apporté une façon différente de faire des maths que je n'oublierai pas. Un merci tout particulier à Antoine (et Adé !) également, pour tous tes conseils et les soirées jeux parfois interminables ;) Un coucou à la bande de Mathilde N., Baptiste, Loïc et Benjamin, j'espère que ça va, où que vous soyez ! Coucou également à Nicolas, aux Benoîts, à Manon et Maxence, David, Logan, Spider (cette soirée au Borriquito...), Simon, Nathanaël, Gabriel, Stella, Élodie, Sandrine, Renan, Théo, Ibrahim, Antoine, et j'en oublie peut-être quelques uns. C'est du beau travail qu'on a fait collectivement cette année-là, je ne l'oublierai pas de sitôt.

Le voyage dans le temps est presque fini, et j'adresse maintenant un salut plein de respect à mes camarades du M2RI. Si l'agrégation est, malgré le concours, l'année de l'entraide et du travail de groupe, l'esprit de compétition plane malheureusement bien plus sur le M2RI. En attendant des temps meilleurs, j'ai beaucoup apprécié cette année à vos côtés, et j'espère que vous avez tous trouvé votre voie !

Enfin, je termine par une petite mention à toutes les personnes que j'ai pu rencontrer à droite à gauche, au cours de conférences ou autres écoles d'été. Il serait impossible de tous les nommer, mais je peux mentionner Lucile et ses éléphants, Émeline et ses poissons, Thomas, Armand et la marche du kouign-amann, Baptiste, Michel, Zoé (et pas Agathe), Apolline et ses fantômes, Brenda, Alexandre Legrand (lui-même), et tous les autres. Vous formez une belle communauté, et j'ai passé de très bons moments, tant scientifiquement qu'humainement, avec vous. Merci pour ça, et à tous je vous souhaite le meilleur !

1.3 À la moins-jeunesse

"À la vieillesse", c'était pas terrible, j'avoue. "À l'expérience", c'est mieux, mais je serais bien en peine de distribuer les bons points dans ce domaine. On reste factuel, en n'oubliant pas que ce n'était après tout qu'une jeunesse symbolique.

À nos amis enseignants et enseignants-chercheurs de l'IMT

La tâche est vaste, mais je souhaite remercier les enseignants et enseignants-chercheurs de l'IMT qui m'ont accompagné, à des degrés divers, pendant ces quelques années.

Pour simplifier les choses, je procéderai par regroupements géographiques et temporels.

Au lieu d'avancer dans le temps, nous allons reculer, en commençant par les membres du département GMM de l'INSA. Merci à Laure pour ces quelques années d'enseignement en commun, toujours rapide et efficace (et une autre grand-sœur plus balèze que moi). Merci à Juliette d'apporter un peu de sang neuf au département, et rendez-vous aux prochaines élections ! Merci à Charles pour les conseils et l'expérience dont il nous fait profiter chaque jour, à nous les jeunes pleins de naïveté et d'espoir. Un grand merci à Anthony "Tonton" Réveillac pour avoir pris soin de moi depuis le premier jour (littéralement) et jusqu'à aujourd'hui. Même si tu nous a quittés pour le CSH, je ne t'oublierai pas. Merci Jean-Paul de m'avoir fourni, même si nous ne nous sommes pas vraiment connus, un bon réservoir d'anecdotes intéressantes. Un grand merci à Jean-Yves, d'abord d'avoir assuré la direction du département avec brio, et surtout d'avoir été disponible pendant ces trois ans. Ton oreille attentive m'a été d'une grande aide quand la pandémie rendait nos vies de plus en plus difficiles, et ton humanité mérite d'être, par tous, saluée. Un deuxième coucou à Olivier, merci encore pour cette année de travail constructive et très intéressante, et pour ta participation aux mots croisés. Merci aussi à Mélisande pour sa bonne humeur constante, et pour les chocolats (ça reste entre nous, bien-sûr).

Permettez que je passe à la ligne pour remercier un tandem mythique du GMM : Cathy et Fred. D'abord merci, et chapeau à Cathy pour sa co-gestion du département. Tu es souvent parmi les derniers fidèles une fois l'été arrivé, et parmi les premiers le matin venu, toujours là pour commencer ma journée avec une petite pause, toujours là pour déplorer le manque de respect pour les câbles HDMI, et toujours compréhensive à l'égard des doctorants. C'est aussi grâce à toi si, logistiquement parlant, tout s'est déroulé pour moi comme sur des roulettes. Merci, et bravo. Et ensuite Fred. Fred... Je ne pourrai pas mieux faire ici que te citer toi-même : merci d'être ce que tu es. Même si parfois tu me fatigues, je ne peux pas imaginer le département sans toi. J'ai été très content de travailler avec toi et partager tous ces repas, à discuter de choses et d'autres, surtout d'autres (parfois au grand dam d'Olivier), reste comme ça. Et passe ta HDR.

Merci à Florent, d'avoir maintenu l'ambiance aux afterworks GMM, et d'être toujours là pour discuter avec "les jeunes", comme tu fais la concession de les appeler toi-même. Merci à David pour les discussions toujours intéressantes, et pour l'intérêt que tu portes à la pédagogie. Parlant de pédagogie, un immense merci (oui, immense) à l'équipe des PRAG (dont Laure, même si elle est déjà passée !). Merci à Sandrine de m'avoir fait confiance dès le début et de m'avoir mis le pied à l'étrier. J'ai énormément appris avec toi, tu as grandement contribué à faire de moi l'enseignant que je pense être devenu. Et merci aussi pour ton inspiration en matière de rangement du bureau ! Merci à Loïc pour tous les conseils et la participation à la bonne ambiance des soirées, et surtout merci de jouer le rôle du méchant flic auprès des étudiants, ça m'a bien aidé à l'occasion ;) Merci à Gégé pour ton soutien dans ces dernières année, particulièrement l'ATER, où le cours de maths 2 a pu passer de façon relativement indolore malgré la grosse charge qu'il représente. Et merci pour ton rire qui, même si tu ne le sais pas, m'a

fait par moments beaucoup de bien. Enfin, un grand merci à Adeline, pour avoir cru en moi dès le début et avoir toujours trouvé les mots pour me remotiver. Je suis admiratif de ta ténacité et de ton engagement pour faire de l'enseignement supérieur un lieu destiné à tous, quels que soient les parcours. Ton engagement m'a beaucoup inspiré ces dernières années, et ton dynamisme pédagogique m'a poussé sans cesse à remettre sur le métier mon ouvrage. Je pars la tête pleine d'idées et d'idéaux, et c'est probablement une des plus belles choses que j'ai acquises durant la thèses. Enfin, un grand merci collectif pour l'aide que vous m'avez apportée lors des concours de PRAG. Vos conseils précieux et avisés m'accompagneront longtemps, et j'espère que bientôt nos routes se rencontreront de nouveau.

Je reprends le fil de notre parcours. Merci à Pierre C. pour cette jeunesse que tu apportes aussi, à ta manière, au département. Ça fait du bien d'avoir avec nous quelqu'un qui vit et pense un peu différemment. Un grand merci à Lorick et Sandrine, pour leur générosité et leur attention. C'est aussi grâce à vous que j'ai pu garder les pieds sur terre et la tête hors de l'eau pendant ces quelques années. Merci pour les cours de sabre, évidemment, pour les innombrables conseils, et parfois le réconfort, que vous m'avez si gentiment prodigués. Vous êtes la preuve que même si les choses se passent parfois mal, en faisant ce qu'il faut la chance peut tourner. C'est une leçon dont je me rappelle souvent, et que je ferais bien de ne pas oublier. Encore un très grand merci à vous deux. Merci Robin pour les conseils que tu as pu apporter sur la fin de ma thèse et pour l'intérêt que tu portes à mon parcours. J'apprécie toujours discuter avec toi, tu apportes un point de vue auquel je pense rarement et qui est toujours enrichissant.

Quatre personnes particulièrement importantes arrivent, et méritent amplement leur paragraphe dédié. D'abord un grand merci à Olivier M., pour beaucoup de choses. Pour la confiance que tu m'accordes et le sens de la justice qui t'anime, qui résonne avec mes propres convictions et les enrichit toujours. Merci pour ta connaissance des textes et les conseils que tu es toujours prompt à fournir, dans la plus grande modestie vis-à-vis de tes propres connaissances. Merci enfin pour les inénarrables discussions du midi, toujours variées et enrichissantes, et pour être prêt à me suivre même sur les pentes les plus glissantes. Le GMM est, je crois, très chanceux de t'avoir dans ses murs, et je garderai ton engagement en mémoire au moment de faire mes premières armes dans la vaste et impitoyable arène politique (au sens à la fois noble et large du terme).

Je ne pouvais pas ne pas dédier un passage à Romain. Un immense merci pour avoir toujours été là. Dans les temps calmes comme au cœur des tempêtes, tu es resté à mes côtés pour m'empêcher de couler. Merci pour tout ce temps passé ensemble, à la Maison Blanche, à la Cave Po, en promenade ou en soirée. Merci à toi et à Laurent pour les après-midis et soirée jeux, et pour m'avoir poussé à reprendre du service dans le jeu vidéo. Merci pour toutes les discussions, tous les conseils et le soutien que tu m'as prodigués. Tu es pour moi presque le grand-frère que je n'ai pas, et sans toi la thèse se serait passée bien autrement. Merci du fond du cœur, et bon courage. Je te souhaite que tout aille bien, tu le mérites amplement.

Ta cobureau n'est pas en reste. Un grand merci à Aude pour la gentillesse dont tu as toujours fait preuve à mon égard. À toi comme moi la période récente n'a pas fait

de cadeaux, et ton courage et ton obstination m'ont servi de modèles. Toujours prête à rendre service, toujours encourageante et compréhensive, j'ai été très heureux de te côtoyer pendant ces quelques années. Je te souhaite de trouver la paix après tout ça, tu le mérites aussi.

Enfin, un très grand merci à Violaine. Ton engagement est exemplaire, et je compte bien le reprendre et l'exporter, dans mon humble mesure, où que j'aille. Ta bataille (votre bataille) acharnée pour que les femmes se fassent une place bien méritée en sciences, tes interventions auprès des jeunes pour leur faire aimer les maths, tout ça est si important, j'espère pouvoir le porter au mieux à mon tour. Merci également pour la confiance que tu me témoignes et pour tous les conseils, tant dans mon enseignement au quotidien que pour les concours de PRAG. Toujours disponible et toujours pertinente, je crois pouvoir dire que je suis honoré d'avoir pu travailler à tes côtés. Merci, et bravo.

On arrive à la fin du tunnel. Merci Stéphane pour ta gentillesse et ta dévotion exemplaire à l'égard des étudiants. Et merci aussi, de façon plus étonnante, de m'avoir fait réaliser que l'enseignant n'est pas infaillible, qu'il est aussi sensible et émotionnellement investi. Ton parcours force le respect, et nourrit mon propre engagement, merci. Merci aussi à Philippe pour les conversations toujours intéressantes et pour le point de vue extérieur que tu apportes, qui enrichit toujours la discussion. Un grand merci à Béatrice, pour ta gentillesse et ta disponibilité, et pour avoir assumé pendant 6 ans (!) la direction du département. J'ai beaucoup apprécié pouvoir bénéficier de ton expérience (et disons-le, de ta sagesse) pendant ces quatre années. Et merci pour toutes les cerises ! Tant qu'on parle de la direction, merci à Pascal, qui a dignement repris le fauteuil du directeur. Merci pour ton dynamisme, ta disponibilité également et ton désir de permettre à chacun de s'exprimer, de participer à la discussion et aux décisions du département. Bon courage pour le reste de ton mandat ! Merci à Kévin d'avoir été présent en salle café aux pics de mon addiction, et bon courage pour la suite. Merci aussi à Jérôme, pour ta gentillesse à mon égard, les nombreux conseils que tu m'as donnés et ton assiduité à suivre mon parcours. Merci aussi pour le point de vue que tu apportes lors de nos discussions, souvent différent mais non moins pertinent. Il a contribué à enrichir ma propre perception des choses, ce qui n'allait pas de soi ! Pour ça, merci aussi à Jean-Yves P. et Frédéric C., avec qui les discussions ne manquent jamais d'intérêt.

Enfin, je termine ce tour des enseignants/enseignants-chercheurs par un CR, désolé, mais merci Pierre. Tu es souvent le seul à m'accompagner dans mes errances méridiennes, ça fait plaisir. Merci pour ton dynamisme et la bonne ambiance que tu contribues à créer, j'espère sincèrement que nos routes se croiseront de nouveau !

Permettez-moi de passer maintenant à nos collègues de l'IMT. Il serait illusoire de vouloir tous les citer, mais j'ai pu, en dix an passés dans ces murs, faire nombre de rencontres marquantes.

Je voudrais d'abord remercier ceux qui ont eu l'immense honneur, voir le prestige, de m'avoir en cours. Merci d'abord à l'équipe enseignante du M2RI, Anthony (qui est déjà passé !), Fabrice et Thierry, François, Gersende, Manon et Patrick (qui est déjà passé aussi), vos cours aussi divers qu'intéressants m'ont conforté dans mon choix

de parcours, tant du point de vue recherche qu'enseignement. Une mention spéciale à Clément et Sébastien, pour la confiance que vous m'avez accordée et l'aide que vous m'avez apportée, j'ai beaucoup aimé ces quelques mois. Désolé de vous avoir trahi en retour, sans rancune ! Un grand merci également à Pierre Bousquet, pour le temps qu'il m'a accordé et la qualité de son enseignement. Merci également à Francis de m'avoir fait réaliser que, parfois, on ne peut pas tout faire, et merci à Radu de m'avoir fait voir qu'être un bon chercheur ne fait pas tout.

Un grand merci un peu spécial à l'équipe du M2 agreg, ce n'est pas l'année la plus drôle, et vous avez réussi à nous porter jusqu'au concours. Un grand merci et chapeau bas à Stéphane Lamy, toujours présent et à l'écoute, d'avoir porté cette année. Bon courage à Fanny et Corentin, qui doivent commencer à être rodés ! Merci à l'ensemble de l'équipe enseignante, en vrac Fanny, Thierry, Emily, Pierre, Franck, Franck, Philippe, François, Vincent (Feufeu), et j'en oublie peut-être. Merci du temps que vous accordez à cette formation, lors de laquelle j'ai beaucoup appris. Merci également à tous les doctorants qui y ont donné des colles, certains dont je n'ai malheureusement jamais su le nom.

Enfin, je remercie les enseignants que j'ai pu croiser dans les premières années. C'est vous qui avez fait le mathématicien que je suis aujourd'hui, merci. Une mention particulière pour Jean-Marc Bouclet, le premier enseignants de mathématiques que j'ai rencontré dans le supérieur, qui est pour beaucoup dans mon choix d'orientation, pour Patrick L., qui m'a ouvert aux maths appliquées avec passion, et pour Judith V. et Patrick M. pour leur gentillesse et leur disponibilité.

Les premières années ont été l'occasion, grâce aux parcours spéciaux, de plonger très vite dans le grand bain de la recherche. Pour ça, je remercie particulièrement Jasmin et Xavier, Francesco, Anne-Charline et Marc, qui m'ont mis le pied à l'étrier et m'ont transmis un peu de leur passion pour les belles maths. Un merci spécial à Xavier, grâce à qui j'ai pu partir quatre mois au Canada, probablement une des expériences les plus enrichissantes de ces années-là. Et à ce titre, un grand merci à Anthony Quas de m'avoir accueilli à Victoria, et un coucou à Dan et Joseph, qui se sont occupés de moi au jour le jour. Thanks guys !

J'en profite pour adresser un salut amical à nos collègues des autres disciplines, à qui je doit mon goût pour les mathématiques dans tous leurs états, que j'ai eu le plaisir d'écouter en première année (il y a 10 ans déjà !). Merci Jesse, de m'avoir fait confiance, merci Florence, pour les meilleurs cours de méca de la fac, merci Sophie, pour le pot de fin de semestre, merci Benoît, qui m'a permis de trouver ma voie, merci Jean-Paul B. pour les cours de C hauts en couleur, et tous les autres dont je peine, après tout ce temps, à me rappeler. Vous m'avez donné une ouverture d'esprit scientifique que je m'efforcerai toujours de conserver et de transmettre.

Enfin, il y a les collègues simplement croisés au détour d'un couloir, autour d'une pause café ou le temps d'un séminaire. Merci Pierre P., de t'occuper sans relâche du café du 1R1 et pour les séances de musique, merci à tous les chefs d'équipes, qui font un boulot monstre mais nécessaire, merci à l'équipe de direction, de faire tourner cette grosse machine malgré les tensions et les individualités... particulières, et merci enfin

aux responsables de séminaires, qui maintiennent une vie scientifique de haut niveau à l'IMT. Je suis fier de dire que j'y étais, et d'avoir contribué, à ma modeste échelle, à cette animation.

En parlant d'animation et de coordination, je ne peux pas passer à côté des personnels de soutien, sans qui la grosse machine resterait à l'arrêt.

Aux personnels de soutien, qui portent bien leur nom

C'est naturellement que je me tourne en premier vers les deux meilleures secrétaires (subjectivement) de l'INSA. Un très très très grand merci à Sandrine, toujours dispo, toujours de bon conseil, toujours à l'écoute et prête à aider les jeunes comme les moins jeunes. Que ce soit pour un stylo, pour passer le temps ou faire passer le blues, c'est toujours avec plaisir que je descends te voir. Merci pour tout ça. Merci aussi à Marlène, pour sa disponibilité et son efficacité, pour sa magie d'ADE et sa gestion des listes d'étudiants. À vous deux, vous êtes l'épine dorsale du GMM, sur laquelle tout le monde se repose, consciemment ou pas. Merci beaucoup. Un petit coucou à Brigitte, qui a repris du mieux qu'elle pouvait la place de Marlène quelques temps, et à Monique, qui nous regarde de loin depuis sa retraite bien méritée. Merci aussi aux autres secrétaires, notamment Sylvie R. et Laurence D., dont l'efficacité et la gentillesse m'ont accompagné pendant quatre ans avec plaisir. Un grand merci enfin à toutes ces petites mains de l'ombre, secrétaires pédagogiques ou administratives, agents d'entretien ou de restauration, service technique ou informatique, qui font de l'INSA une maison accueillante et confortable.

La grosse machine n'est pas en reste. Merci à toutes les secrétaires et gestionnaires, qui nous rendent la vie bien plus facile, et qui sont toujours là pour aider les nouveaux arrivants à trouver leurs repères. Merci au service informatique, pour le travail titanesque qu'ils abattent chaque année, et bon courage plus particulièrement à Pierre, qui essaye patiemment et obstinément de faire changer les pratiques. Merci aussi à Dominique et Nathalie (Domi et Nath entre nous, mais je ne devrais pas vous le dire...), pour leur gestion terriblement efficace de la bibliothèque, leur gentillesse et leur disponibilité. On vous doit probablement une partie de notre réussite à l'agreg, et ce n'est pas peu dire ! Merci enfin à toutes les petites mains de l'ombre, toujours présentes et discrètes, sans qui, entre autres, les toilettes seraient bien plus sales (et rien que ça, imaginez...).

Mais la fac, ce sont aussi les secrétaires pédagogiques, et à ce titre je voulais adresser un immense merci à Sylvie Crabos. Pardon pour les autres, mais on vous a élue (officieusement) meilleure secrétaire de la FSI, et haut la main. Merci pour votre disponibilité, vos conseils toujours avisés et votre gentillesse proverbiale. C'est grâce à vous que l'année de l'agreg s'est passée sans le moindre tracas administratif. C'est sans me tromper que je vous remercie pour tout ça de la part de toute la promo, nous ne vous oublierons pas. Merci aussi à Émilie, qui a fait de son mieux pour gérer toujours plus d'étudiants de licence, chapeau.

Et comme l'administration ne s'arrête pas avec la thèse, un grand merci à Agnès Requis, grâce à qui les démarches administratives des doctorants sont comme une prome-

nade de santé sous une pluie battante (je ne peux pas dire que ce soit agréable non plus, mais c'est toujours mieux que le parcours du combattant habituel). Un immense merci à Marie-Hélène également, pour son dévouement auprès des doctorants, son écoute et ses conseils. Si je suis ici aujourd'hui, c'est un peu grâce à toi, merci beaucoup.

2 Parce qu'il n'y a pas que le travail dans la vie

S'il n'est pas peu dire que l'université et les mathématiques m'ont pris un peu de temps dernièrement, il serait heureusement faux de dire que ç'ont été mes seuls centres d'intérêt. Et pour ça je remercie tous ceux qui m'ont permis de changer d'air pendant ces quatre (parfois longues) années.

2.1 À mes amis musiciens, à la culture avec un gros Q

Les premiers auxquels je souhaite m'adresser sont mes amis musiciens de l'Enharmonie. Merci à tous pour le temps passé en répétition, en concert, le temps du travail et le temps du plaisir. Cette bouffée d'air hebdomadaire a été salutaire, malgré le caractère pas toujours facile de notre chef... Merci d'abord aux altos, Anne-Flore, Marielle, Matthieu, et aux anciens Sarah (bientôt de retour ?) et Julien. En toute objectivité, on est le meilleur pupitre de l'orchestre. Un merci tout particulier au comité des fêtes, Capu et Florian, dont l'arrivée récente à fait souffler un nouveau vent de convivialité sur l'orchestre. Et bien-sûr un grand merci à l'ensemble des musiciens qui font vivre cette belle formation : Caro, Caca (toujours là dans nos cœurs !), Leslie, Maryse, Aurélie, Chalu (et Ainhoa !), Agnès, Hélène, Aurélie, Manu, Laurence, Fleur, Romain, Estelle, Pierre, Sébastien, Adrian, Michel, et tous nos amis souffleurs ! Et, évidemment, merci à Serge notre chef de prendre autant de son temps pour nous faire travailler et à Seb pour sa patience et ses conseils toujours avisés.

Bien-sûr, il n'y a pas que l'Enharmonie, et je salue de loin nos amis de l'OUT. Trop nombreux pour être tous cités (et ça fait trop longtemps que je n'y suis plus...), j'ai passé avec vous aussi des années formidables. Un petit coucou aux vrais des vrais, avec qui l'aventure continue toujours, Dada (j'espère que la maman et le bébé vont bien !), Mimi et Manu, j'ai de la chance d'être entouré d'autant de musiciens de qualité et motivés pour avancer tous ensemble. Un coucou également aux musiciens des divers ensembles au sein desquels j'ai pu intervenir ces derniers temps. Je n'oublie évidemment pas mes anciens collègues de l'orchestre de Tournefeuille, toujours au top, ainsi que tous mes enseignants de FM et d'alto, qui m'ont donné le goût du travail musical et le niveau pour y prendre beaucoup de plaisir (en particulier Caroline, merci).

Je voudrais terminer cette partie musicale par deux personnes dont la présence a été très importante ces derniers temps. D'abord Chalu, même si on ne se voit plus beaucoup (d'ailleurs, je ne te vois pas dans la salle !), merci d'être toujours partant pour faire la fête et passer du bon temps. Je me souviens de l'EVG à Budapest, du petit week-end irlandais dans les montagnes, bien-sûr du golden (half-)mile, autant de bons souvenirs qui continuent de me porter dans les temps difficiles. Et ensuite Hu (anonymat

garanti !), merci d'être toujours présente, soit à Toulouse soit au bout du fil depuis Marseille. Merci d'ailleurs de m'y avoir accueilli et de m'avoir fait découvrir ses merveilles, la Bonne Mère (wallah merci à elle), le cours Ju et bien-sûr l'OM et le Vélodrome ! C'est un privilège d'avoir une amie sur laquelle je peux compter à toute heure du jour ou de la nuit, merci d'être là. En bande organisée, personne peut nous canaliser !

Savez-vous ce qui se cache au 71 rue du Taur, au fond d'un couloir sombre et pas bien large, derrière une porte en bois sans charme et au bas d'un escalier humide ? Une cave, bien-sûr, mais pas n'importe laquelle. Une cave où, comme dirait Hugo, se mêlent le grotesque et le sublime : la Cave Poésie. J'ai eu la chance d'y servir bénévolement sans relâche depuis deux ans maintenant, et je tenais à remercier chaleureusement toute l'équipe pour sa bienveillance et la confiance qu'elle m'accorde. Un grand merci à Yann, Jean-Noël, Gisèle, Clémentine, Lou, Benjamin et Mathilde de faire vivre ce lieu et de nous permettre de venir y passer du bon temps. Et un coucou à ceux qui ne sont plus là mais qu'on n'oublie pas, Aoife (tu nous manques), Lola, Paulin, Alexis, Hugo et Pauline, merci à tous pour votre gentillesse et votre confiance. Avec la musique, la cave était (paradoxalement) mon deuxième bol d'air, et j'espère vous revenir très vite !

Et, il ne faut pas les oublier, un petit mot à tous les artistes, intermittents du spectacle, musiciens titulaires, troupes amateures, qui ont connu avec la pandémie une période extrêmement difficile. Merci à eux d'avoir continué, et de continuer, à nous faire vibrer, pleurer, rire, voir le monde sous un nouveau jour et réaliser que, même si on est bien peu de choses, la vie vaut la peine d'être endurée.

Enfin, j'aimerais adresser un petit mot à Toulouse Haidong Gumdo. C'est peut-être pas de la "culture" en soi, mais c'est un art martial, donc quelque part... Merci à notre kwanjang Lorick et notre kyobeom Noémie, ainsi qu'à tous les su-ryeon-saeng du Baeg-Ho Dohjang, avec qui j'ai pu exercer mon corps et mon esprit dans les meilleures conditions. De blessure en rédaction, je n'ai pas pu être présent récemment, mais j'espère pourvoir vous revenir d'ici peu !

2.2 À l'intime

Nous nous acheminons tranquillement vers la fin de ces remerciements. Dans cette pénultième section, je voudrais remercier tous ceux qui, dans la vie de tous les jours, en-dehors du travail ou des loisirs, m'ont apporté leur soutien.

Je vais commencer par toute l'équipe médicale qui m'a suivi pendant ces années. Merci à Nicolas, pour séances de rééducation et de papotages sans fin. Merci au Dr Bousquet, pour son suivi attentif et ses conseils toujours avisés et merci aux Drs Dethieux et Sibaud, pour m'avoir parfois porté à bout de bras jusqu'à la fin de cette aventure. La vie est une chose bien étrange et imprévisible, insatiable et qui sans cesse demande de remettre son ouvrage sur le métier. On ne va alors guère loin avec un mauvais métier, merci à vous de m'avoir aidé à concevoir et renforcer le mien. Pour ça, merci aussi à Romain, qui a fait beaucoup et mérite bien cette deuxième dédicace.

De façon plus anecdotique, merci à toi, l'inconnu, qui sur internet publie ce meme ou cette vidéo de chat. Internet est une culture formidable au foisonnement exception-

nel, n'en déplaise aux tenants élitistes d'une culture embourgeoisée. J'ai pu découvrir, au gré de pérégrinations sans but mais souvent salutaires, un nombre de domaines hallucinant, m'intéresser à l'histoire médiévale, à l'art préhistorique, au Titanic, aux arts plastiques, à la philosophie arabe ou à la chimie culinaire, le cul vissé à ma chaise sans bouger de chez moi. Sans ironie, on vit une époque formidable. Par le truchement d'internet, je remercie (à titre posthume) Arnold Bax et Alfred Schnittke d'avoir composé la BO de ma vie ces dernières années.

Enfin, je voudrais adresser un mot à ma famille. Merci à mes parents de m'avoir fait (déjà, parce que je ne serais clairement pas là sinon), d'avoir cru en moi et de m'avoir toujours poussé et conseillé au cours de ces années. Merci d'avoir subvenu inconditionnellement à mes besoins durant tout ce temps, je mesure la chance que j'ai eue. Je ne le dis pas souvent, mais je vous aime fort tous les deux. Merci aussi à la petite, d'être si différente de moi, et si proche à la fois. Comme dirait Dom dans Fast and Furious, rien n'est plus important que la famille, et je crois qu'on en a tous pris conscience récemment. Merci à Grisou aussi, toujours là et jamais avare de consolation (bien que de plus en plus bavarde avec l'âge). Merci à Dadie et Papoum, qui ont toujours suivi avec le plus grand intérêt mon parcours et ont toujours été fiers de moi. Vous vous estimez peut-être chanceux d'avoir de tels petits-enfants, mais ces derniers sont encore plus fiers de leurs grands-parents. Je mesure là aussi la chance que fut la mienne toutes ces années. Merci à mon parrain et ma marraine, toujours fiers de moi et attentifs, j'espère pouvoir bientôt monter fêter ça avec vous ! Enfin je remercie ma famille au sens plus large, en Bretagne, à Gradignan, et partout en France, vous avez tous contribué à faire la personne que je suis, et nous avons partagé tant de bons souvenirs que je ne saurais ici tous les mentionner. J'espère en vivre encore beaucoup. Merci à tous !

2.3 À l'avenir

Ainsi se terminent ces remerciements. De façon générale, je remercie toutes celles et ceux qui m'ont aidé à avancer pour arriver là aujourd'hui. Le milieu de l'enseignement supérieur est de la recherche est précieux, mais fragile. L'absence de gestion humaine le rend sensible aux individualités problématiques, sa parfois faible politisation l'expose à de nombreuses attaques avec de faibles défenses, et l'élitisme et "l'excellence" gangrènent l'enseignement. Concept mobilisateur s'il en est, qui peut être contre l'excellence ? Quand celle-ci se mue, de la joyeuse émulation, en une concurrence aux financements, un élitisme envers les étudiants, et en indicateurs chiffrés tous moins pertinents les uns que les autres, tout le monde devrait l'être. C'est la même bataille que celle contre la méritocratie, point complètement aveugle aux déterminants sociaux pourtant... déterminants. Je quitte aujourd'hui le supérieur (ou pas, qui sait), mais j'espère sincèrement que, en jouant collectif, le milieu sortira de l'impasse dans lequel il se trouve. Que la recherche publique française ne finisse pas définitivement au rayon des souvenirs.

Table des matières

1 Introduction générale **1**
1.1 Opérateurs de diffusion 2
1.1.1 Aspects historiques 2
1.1.2 Cadre formel 3
1.1.3 La théorie de Bakry-Émery 13
1.1.4 Hypothèses techniques 16
1.2 Inégalités fonctionnelles 17
1.2.1 Inégalité de Poincaré 18
1.2.2 Inégalité de Sobolev logarithmique 23
1.3 Contexte de la thèse et présentation des résultats 27
1.3.1 Relations d'entrelacement 27
1.3.2 Présentation de quelques résultats et plan de la thèse 32

2 Représentation des semi-groupes de Feynman-Kac, application aux inégalités de Sobolev logarithmiques **37**
2.1 Représentation des semi-groupes de Feynman-Kac 38
2.1.1 Éléments de calcul stochastique 38
2.1.2 Heuristique scalaire 39
2.1.3 Régularité du processus et flot tangent 41
2.1.4 Approche perturbative 46
2.2 Inégalité de Sobolev logarithmique 49
2.2.1 Critère général 50
2.2.2 Fonctions monotones 53
2.2.3 Exemples 56
2.3 Perspectives 62

3 Une approche algébrique et géométrique à l'étude spectrale des générateurs **63**
3.1 Éléments d'analyse spectrale 65
3.1.1 Décomposition du spectre 65
3.1.2 Cas du générateur d'Ornstein-Uhlenbeck 67

3.1.3 Trou spectral et caractérisation de Courant-Fischer 73
3.2 Contexte algébrique 76
3.2.1 Algèbre tensorielle 77
3.2.2 Opérateurs de diffusion et Schrödinger sur l'algèbre tensorielle et estimations préliminaires 79
3.2.3 Transformations de Riesz 82
3.3 Estimations spectrales pour $-\mathbf{L}$ 84
3.3.1 Résultat principal 84
3.3.2 Optimalité dans le cas gaussien 89
3.3.3 Perspectives 91

4 Aspects numériques des inégalités de Poincaré et application à l'analyse de sensibilité **93**
4.1 Contexte général 94
4.1.1 Modèle physique 94
4.1.2 Modélisation probabiliste 97
4.2 Analyse de sensibilité 104
4.2.1 Principes de base 104
4.2.2 Indices globaux et liens avec l'inégalité de Poincaré 106
4.3 Aspects numériques de l'inégalité de Poincaré 112
4.3.1 Principe général 112
4.3.2 Généralités sur les éléments finis en dimension deux 115
4.3.3 Domaine et maillage 118
4.3.4 Validation de la méthode 128
4.3.5 Implémentation 134
4.3.6 Conclusion et perspectives 139

Bibliographie **141**

Chapitre 1

Introduction générale

LE BUT de ce chapitre introductif est de présenter le cadre mathématique dans lequel s'inscrivent les résultats présentés dans les chapitres suivants. Ce travail se place à l'interface de l'analyse fonctionnelle, des probabilités et de la statistique. Il traite plus particulièrement des inégalités fonctionnelles (Poincaré, Sobolev logarithmiques, etc.) et des estimations spectrales des opérateurs de diffusion, et de l'application de certaines de ces notions à l'analyse de sensibilité. Le but de cette thèse est d'appliquer, suite à la série de travaux récents de Michel Bonnefont et Aldéric Joulin, des techniques d'entrelacements à ces différents problèmes, et de fournir un exemple applicatif concret autour de certains aspects numériques. Nous rappellerons donc dans un premier temps la définition des opérateurs de diffusion, avant d'introduire les inégalités fonctionnelles d'intérêt pour la suite du travail et enfin les techniques d'entrelacement telles que présentées par Bonnefont et Joulin dans leurs articles sur le sujet.

Table des matières

1.1 Opérateurs de diffusion 2
1.1.1 Aspects historiques 2
1.1.2 Cadre formel 3
1.1.3 La théorie de Bakry-Émery 13
1.1.4 Hypothèses techniques 16
1.2 Inégalités fonctionnelles 17
1.2.1 Inégalité de Poincaré 18
1.2.2 Inégalité de Sobolev logarithmique 23
1.3 Contexte de la thèse et présentation des résultats 27
1.3.1 Relations d'entrelacement 27
1.3.2 Présentation de quelques résultats et plan de la thèse 32

1.1 Opérateurs de diffusion

1.1.1 Aspects historiques

Les considérations autour des opérateurs de diffusion prennent leur source, comme c'est le cas de nombreux opérateurs différentiels, dans la physique. Le point de départ est l'introduction de l'équation de la chaleur par Joseph Fourier en 1822 [40], suivie de la mise en évidence du mouvement brownien par Robert Brown en 1827. Le phénomène de diffusion n'est pas encore compris, mais le formalisme introduit par Fourier incitera Adolf Fick à présenter sa célèbre loi dès 1855 [37] :

$$\frac{\partial C}{\partial t} - \nabla \cdot (\rho \mathcal{D} \nabla C) = 0,$$

pour décrire le mouvement de particules de masse volumique ρ de concentration molaire C avec coefficient de diffusion $\mathcal{D}$. Cette loi sera confirmée et raffinée par James Clark Maxwell [61] et Jožef Stefan [82], afin de prendre en compte l'effet de gradients de charge ou de pression.

Il faudra néanmoins attendre le début du vingtième siècle pour que le lien avec le mouvement brownien soit compris, avec Albert Einstein en 1905 [32], puis Marian Smoluchowski en 1906 [85]. Les outils mathématiques manquent encore néanmoins dans le traitement de ces problèmes, malgré les définitions successives de Louis Bachelier en 1901 [4], Einstein en 1905 [32] ou Paul Langevin en 1908 [57]. Ces premiers tâtonnement théoriques permettront néanmoins des avancées spectaculaires en physique des particules, avec notamment Jean Perrin en 1909 [70], qui obtiendra le prix Nobel en 1923 pour sa confirmation empirique de l'hypothèse atomiste de la matière.

Le formalisme mathématique permettant enfin l'étude de tels processus sera introduit par Norbert Wiener en 1923, qui donnera le premier une description de la mesure du même nom et une définition de l'intégrale brownienne [89]. Son caractère martingale et sa caractérisation en terme de variation quadratique seront établis par Paul Lévy, à partir de 1933 [60], et l'étude précise des équations différentielles stochastiques sera enfin rendue possible en 1944 avec l'introduction par Kiyoshi Itō du calcul stochastique. La théorie a depuis connu de nombreux développements, comme l'intégrale de Stratonovich, ou le calcul de Malliavin et l'intégrale de Skorokhod (voir par exemple [68] à ce sujet).

En parallèle des avancées autour des processus de diffusion, les idées de Max Planck amènent Louis de Broglie à proposer en 1924 les équations qui poseront les fondements de la dualité onde-corpuscule de la lumière [28]. À sa suite, Erwin Schrödinger établira en 1926 l'équation qui porte son nom sur la densité de la probabilité de présence de particules soumises à une force issue d'un potentiel V [78] :

$$i\hbar \frac{\partial \Psi}{\partial t}(t,x) = -\frac{\hbar^2}{2m} \Delta \Psi(t,x) + V(x)\Psi(t,x).$$

Cette équation ouvre par ailleurs la voie aux interprétations probabilistes de la mécanique quantique (popularisées un peu improprement par l'expérience de pensée du chat de Schrödinger). L'étude de telles équations poussera Richard Feynman et Mark Kac à proposer leur célèbre formule en 1949 (d'abord proposée de façon formelle par Feynman dans le cas complexe, puis formulée rigoureusement par Kac dans le cas réel), qui relie la solution d'équations aux dérivées partielles (incluant celle de Schrödinger) à l'espérance de fonctionnelles de certains processus de diffusion [53].

Cette thèse ne s'intéresse pas du tout aux tribulations physiciennes autour de ces théories, mais plutôt à l'étude mathématique des opérateurs de diffusion sous plusieurs aspects, ainsi qu'à celle des semi-groupes de Feynman-Kac. Le but de cette introduction est de donner le contexte mathématique et les outils nécessaires à la compréhension du travail effectué. La présente section se concentre sur les opérateurs, semi-groupes et processus de diffusion, la deuxième est dédiée aux inégalités fonctionnelles d'intérêt avant de terminer par les relations d'entrelacement, le point de départ technique des résultats obtenus.

1.1.2 Cadre formel

Nombres réels et algèbre linéaire Dans toute cette thèse, nous nous focaliserons sur des espaces vectoriels réels. Le corps des nombres réels sera noté $\mathbf{R}$, et $\mathbf{R}^d$ désignera, pour d un entier positif non nul, le produit cartésien usuel de d copies de $\mathbf{R}$. Les entiers naturels seront désignés par $\mathbf{N}$, les entiers relatifs par $\mathbf{Z}$. On notera de même $\mathbf{N}^d$ et $\mathbf{Z}^d$ les multi-indices de taille d à valeurs entières. Si α est un multi-indice, on notera $|\alpha|$ sa norme 1 (la somme des valeurs absolues de ses composantes). Si $\mathbf{K}$ est l'un des ensembles ci-dessus, $\mathbf{K}^*$, $\mathbf{K}_+$ et $\mathbf{K}_-$ désigneront respectivement les éléments non nuls, positifs (ou nuls) ou négatifs (ou nuls) de $\mathbf{K}$.

Si E et F sont deux espaces vectoriels réels, $\mathcal{L}(E,F)$ désignera l'ensemble des applications linéaires de E vers F. Si $E = F = \mathbf{R}^d$, cet espace est représenté par $\mathcal{M}_d(\mathbf{R})$, l'ensemble des matrices $d \times d$ à coefficients réels. On notera $GL_d(\mathbf{R})$ le sous-groupe des matrices inversibles de $\mathcal{M}_d(\mathbf{R})$.

Si u et M sont respectivement un vecteur de $\mathbf{R}^d$ et une matrice de $\mathcal{M}_d(\mathbf{R})$, on notera Mu le produit matrice-vecteur classique, u^T et M^T les transposés de u et M. Pour une norme $\|\cdot\|$ sur $\mathbf{R}^d$, on notera de même la norme subordonnée sur $\mathcal{M}_d(\mathbf{R})$. Si N est une autre matrice de $\mathcal{M}_d(\mathbf{R})$, on notera $\langle M, N\rangle_{HS}$ le produit scalaire de Hilbert-Schmidt entre M et N (la somme des éléments du produit terme à terme de M et N), et $\|\cdot\|_{HS}$ la norme associée.

Si $M \in \mathcal{M}_d(\mathbf{R})$ est une matrice symétrique, on notera respectivement $\rho_+(M)$ et $\rho_-(M)$ sa plus grande et plus petite valeur propre (en valeur absolue). Cette définition peut s'étendre à des matrices non symétriques dont le spectre est complexe, mais on ne rencontrera pas cette situation dans la suite.

Enfin, pour M et N deux matrices symétriques, on dira que $M \geq N$ si $M - N$ est semi-définie positive. Cet ordre n'est pas total, et n'est valable que pour deux matrices symétriques. la définition reste valide si M et N sont deux opérateurs auto-adjoints.

D'autres structures algébriques pourront être introduites au besoin.

Espaces mesurables Sauf mention contraire, on considérera tout au long de cette thèse l'espace euclidien $(\mathbf{R}^d, |\cdot|)$, $d \in \mathbf{N}^*$, où $|\cdot|$ désigne la norme euclidienne standard.

On notera $(e_1, \dots, e_d)$ la base canonique de $\mathbf{R}^d$ et $x \cdot y$ le produit scalaire euclidien classique :

$$x \cdot y = x_1 y_1 + \cdots + x_d y_d.$$

On considérera dans la suite la tribu borélienne de $\mathbf{R}^d$ notée $\mathcal{B}(\mathbf{R}^d)$, et l'on notera dx la mesure de Lebesgue sur cette tribu. Sauf précision contraire, l'expression "presque partout" renverra à cette dernière.

Espaces fonctionnels On notera $\mathcal{C}^0(\mathbf{R}^d, \mathbf{R})$ l'espace des fonctions continues de $\mathbf{R}^d$ dans $\mathbf{R}$. Une fonction $f \in \mathcal{C}^0(\mathbf{R}^d, \mathbf{R})$ est *(L-)lipschitzienne* s'il existe $L \in \mathbf{R}_+$ une constante telle que :

$$\forall x, y \in \mathbf{R}^d,\ |f(x) - f(y)| \leq L|x - y|.$$

On notera $\mathcal{C}^k(\mathbf{R}^d, \mathbf{R})$, $k \in \mathbf{N}^*$, l'espace des fonctions k fois continuement différentiables. Si $f \in \mathcal{C}^1(\mathbf{R}^d, \mathbf{R})$, on notera ∇f son gradient : pour $x \in \mathbf{R}^d$,

$$\nabla f(x) = (\partial_1 f(x), \dots, \partial_d f(x))^T.$$

Si $F : \mathbf{R}^d \to \mathbf{R}^d$ est un champ de vecteurs de classe $\mathcal{C}^1$, on notera $\mathrm{Jac}_x F$ sa matrice jacobienne au point $x \in \mathbf{R}^d$, définie par :

$$\mathrm{Jac}_x F = (\partial_i F_j(x))_{i,j=1,\dots,d}.$$

Cette convention induit, pour F, G deux champs de vecteurs de classe $\mathcal{C}^1$, la formule de composition suivante : pour $x \in \mathbf{R}^d$,

$$\mathrm{Jac}_x(F \circ G) = \mathrm{Jac}_x G \, \mathrm{Jac}_{G(x)} F.$$

Si $f \in \mathcal{C}^2(\mathbf{R}^d, \mathbf{R})$, on notera respectivement $\nabla^2 f$ et Δf sa matrice hessienne et son laplacien, définis pour tout $x \in \mathbf{R}^d$ par :

$$\nabla^2 f(x) = \left(\partial^2_{ij} f(x)\right)_{i,j=1,\dots,d} \text{ et } \Delta f(x) = \mathrm{tr}(\nabla^2 f(x)) = \sum_{i=1}^{d} \partial^2_{ii} f(x).$$

Ces derniers objets pourront être également considérés au sens des distributions le cas échéant.

On notera enfin $\mathcal{C}^\infty(\mathbf{R}^d, \mathbf{R})$ l'espace des fonctions infiniment différentiables dont les dérivées successives sont toutes continues. On considérera également les espaces $\mathcal{C}^k_c(\mathbf{R}^d, \mathbf{R})$ et $\mathcal{C}^k_+(\mathbf{R}^d, \mathbf{R})$ de fonctions de classe $\mathcal{C}^k$ respectivement à support compact et

strictement positives. On considérera sur ces espaces des fonctions bornées au sens de la norme du sup $\|\cdot\|_\infty$, définie pour une fonction f par :

$$\|f\|_\infty = \sup_{x\in\mathbf{R}^d} |f(x)|.$$

Pour terminer, si μ est une mesure sur $\mathcal{B}(\mathbf{R}^d)$, on définit l'espace $L^p(\mu)$, $p \in [1,\infty)$, des fonctions f mesurables (définies presque partout) vérifiant :

$$\int_{\mathbf{R}^d} |f|^p \, d\mu < +\infty.$$

On munit $L^p(\mu)$ de la norme $\|\cdot\|_p$ définie par :

$$\|f\|_p = \left(\int_{\mathbf{R}^d} |f|^p \, d\mu\right)^{1/p}.$$

Dans le cas particulier où $p = 2$, l'espace $L^2(\mu)$ est un espace de Hilbert pour le produit scalaire

$$\langle f, g\rangle = \int_{\mathbf{R}^d} fg \, d\mu.$$

Si f est une fonction de $L^2(\mu)$ telle que sa différentielle (au sens des distributions) est elle-même une fonction de carré intégrable, on dira que f est dans l'*espace de Sobolev* $H^1(\mu)$. C'est également un espace de Hilbert muni du produit scalaire

$$\langle f, g\rangle_{H^1(\mu)} = \langle f, g\rangle + \int_{\mathbf{R}^d} \nabla f \cdot \nabla g \, d\mu,$$

et de norme associée $\|\cdot\|_{H^1(\mu)}$. On notera div_μ la divergence dans $L^2(\mu)$, l'adjoint de ∇ : pour $f \in H^1(\mu)$ et G un champ de vecteurs dont les coordonnées sont dans $H^1(\mu)$, on a :

$$\int_{\mathbf{R}^d} \nabla f \cdot G \, d\mu = \int_{\mathbf{R}^d} f \, \mathrm{div}_\mu G \, d\mu.$$

Enfin, pour $p = \infty$, on définit $L^\infty(\mu)$ l'espace des fonctions essentiellement bornées, muni de la norme $\|\cdot\|_\infty$ (étendue aux fonctions définies presque partout via le supremum essentiel). On parlera volontiers de *bornitude* pour désigner le caractère borné d'une fonction.

Définition des opérateurs

Comme mentionné ci-dessus, les équations de diffusion ont émergé dans le cadre de la physique des particules, pour décrire l'évolution de la densité de présence de ces dernières dans un milieu. Le plus souvent, en sus des collisions inter-particulaires, le système peut être soumis à des forces extérieures (type gravitation, magnétisme, etc.), traduites par un potentiel extérieur. Les opérateurs que nous introduisons seront tous

basés sur ce modèle, et il peut être bon, pour fixer les idées, de garder en tête l'exemple d'un potentiel gravitationnel attractif.

Considérons donc une fonction $V : \mathbf{R}^d \to \mathbf{R}$ de classe $\mathcal{C}^2$. Nous supposerons uniquement dans un premier temps que e^{-V} est intégrable sur $\mathbf{R}^d$ par rapport à la mesure de Lebesgue. Des hypothèses additionnelles plus ou moins techniques pourront se rajouter au fur et à mesure.

Définition 1.1.1 (Mesure de Boltzmann-Gibbs)**.** On définit la *mesure de Boltzmann associée à* V (ou mesure de Gibbs) par :

$$\mu(dx) = \frac{e^{-V(x)}}{Z_V} dx,$$

où $Z_V = \int_{\mathbf{R}^d} e^{-V(x)} dx$.

Le cas particulier de la mesure gaussienne γ_d, définie pour $V(x) = |x|^2/2$, est de première importance, pour laquelle toutes les propriétés exposées ultérieurement sont vérifiées, et avec des constantes explicites. Tout au long de ce chapitre, des apartés "Et la gaussienne dans tout ça ?" signaleront des points d'arrêts où la propriété en cours sera envisagée du point de vue de la mesure gaussienne.

Par hypothèse sur V, μ est une mesure de probabilité, absolument continue par rapport à la mesure de Lebesgue. Par ailleurs, la constante de normalisation Z_V est appelée "fonction de partition", et possède de nombreux liens avec la physique statistique et la thermodynamique. Le potentiel V joue à ce titre le rôle du hamiltonien de la physique statistique (consulter à ce sujet le vieux mais non moins pertinent ouvrage de Landau et Lifschitz [56]).

Les propriétés de cette mesure peuvent être appréhendées selon de nombreux aspects différents (existence d'une inégalité de Poincaré, plongements de Sobolev, structure hilbertienne de $L^2(\mu)$, etc.) et leur étude a donné lieu au développement de nombreux outils. L'un deux, qui nous intéressera plus particulièrement dans cette première section, est le générateur associé à μ. Cet opérateur apparaît très naturellement dans les calculs d'intégrales, et porte donc beaucoup des propriétés de μ précitées. Considérons à titre d'exemple la formule d'intégration par parties. Pour $f, g \in \mathcal{C}_c^\infty(\mathbf{R}^d, \mathbf{R})$, cette formule s'écrit, pour la mesure de Lebesgue :

$$\int_{\mathbf{R}^d} \nabla f(x) \cdot \nabla g(x)\, dx = -\int_{\mathbf{R}^d} f(x) \Delta g(x)\, dx$$

(f et g étant à support compact et l'intégrale sur tout l'espace, les termes de bord n'apparaissent pas). En prenant en compte la densité de μ par rapport à la mesure de Lebesgue, on obtient la nouvelle formule d'intégration par parties :

$$\int_{\mathbf{R}^d} \nabla f(x) \cdot \nabla g(x)\, d\mu(x) = -\int_{\mathbf{R}^d} f(x)(\Delta g(x) - \nabla V(x) \cdot \nabla g(x))\, d\mu(x).$$

Ainsi, l'opérateur $\Delta - \nabla V \cdot \nabla$ joue pour μ le rôle du laplacien pour la mesure de Lebesgue. Il sera systématiquement noté $\mathbf{L}$ (avec de légères variations si plusieurs potentiels différents interviennent). Ces deux opérateurs sont des cas particuliers d'opérateurs *de diffusion*, au sens suivant.

Définition 1.1.2 (Opérateur de diffusion)**.** Un opérateur différentiel d'ordre deux D est dit de *diffusion* s'il existe une application $A = (a_{ij})_{i,j=1,\dots,d}$ sur $\mathbf{R}^d$ à valeurs dans les matrices symétriques semi-définies positives et un champ de vecteurs $b = (b_i)_{i=1,\dots,d}$ sur $\mathbf{R}^d$ tels que pour $f \in \mathcal{C}^\infty(\mathbf{R}^d, \mathbf{R})$, $x \in \mathbf{R}^d$,

$$Df(x) = \sum_{i,j=1}^{d} a_{ij}(x)\partial_{ij}^2 f(x) + \sum_{i=1}^{d} b_i(x)\partial_i f(x).$$

La matrice A est appelée *matrice de diffusion*, le vecteur b la *dérive*. Dans le cas du laplacien, on a $A \equiv I_d$ et $b = 0$, et $A \equiv I_d$, $b = -\nabla V$ pour $\mathbf{L}$. On peut également réécrire un opérateur de diffusion D sous la forme condensée :

$$D = \mathrm{div}_\mu \circ \nabla.$$

Cette écriture sera plus amplement détaillée dans un chapitre ultérieur. Sauf mention contraire, les définitions et propriétés présentées dans cette section peuvent être, ainsi que les preuves afférentes, retrouvées dans la monographie de référence de Dominique Bakry, Ivan Gentil et Michel Ledoux [8] (chapitre 3).

La définition de $\mathbf{L}$ et la formule d'intégration par parties amènent à quelques propriétés immédiates, mais non moins importantes.

Proposition 1.1.3 (Propriétés de $\mathbf{L}$)**.** *Considérons l'opérateur* $\mathbf{L}$ *défini sur les fonctions* $\mathcal{C}_c^\infty(\mathbf{R}^d, \mathbf{R})$. *Alors :*

1. $\mathbf{L}$ *est symétrique ;*
2. $\mathbf{L}$ *est semi-défini négatif, et ne s'annule que pour les fonctions constantes ;*
3. $\mathbf{L}$ *vérifie la propriété de diffusion (ou de changement de variable) : pour* $\varphi : \mathbf{R} \to \mathbf{R}$ *une fonction de classe* $\mathcal{C}^2$ *et* $f \in \mathcal{C}_c^\infty(\mathbf{R}^d, \mathbf{R})$*,on a*

$$\mathbf{L}\varphi(f) = \varphi'(f)\mathbf{L}f + \varphi''(f)|\nabla f|^2.$$

La deuxième assertion est le point de départ de l'étude spectrale de $\mathbf{L}$. En effet, cette propriété garantit que $\sigma(-\mathbf{L})$, le spectre de $-\mathbf{L}$, est contenu dans $\mathbf{R}_+$. Une question importante est de connaître sa structure : admet-il un trou spectral (écart strictement positif entre les deux premières valeurs propres) ? Est-il discret ? continu ? Ces considérations seront plus amplement traitées dans les chapitres suivants, et dans la section dédiée à l'inégalité de Poincaré.

Notons que les fonctions constantes sont (hormis la fonction nulle) exclues a priori de $\mathcal{C}_c^\infty(\mathbf{R}^d, \mathbf{R})$. Nous verrons dans la suite comment étendre le domaine de $\mathbf{L}$ de sorte à inclure ces dernières. On comprend néanmoins schématiquement que, $\mathbf{L}$ étant un opérateur différentiel, il s'annule pour les fonctions constantes.

La forme quadratique $f \mapsto -\int_{\mathbf{R}^d} f\mathbf{L}f \, d\mu$ sera parfois appelée *forme de Dirichlet* associée à $\mathbf{L}$, et notée $\mathcal{E}$. Notons que la formule d'intégration par parties énoncée ci-dessus implique que $\mathcal{E}$ ne dépend de V qu'à travers μ, et donc plusieurs opérateurs $\mathbf{L}$ différents peuvent admettre la même forme de Dirichlet. C'est pourquoi on considérera en général $\mathbf{L}$ comme une donnée du problème et $\mathcal{E}$ comme un sous-produit de cet opérateur, et non l'inverse.

Domaines de définition

Comme mentionné précédemment, les constantes ne font pas a priori partie des fonctions pour lesquelles l'opérateur $\mathbf{L}$ a été défini. Néanmoins, dans la mesure où ces fonctions seraient intégrables contre la mesure μ, on peut leur appliquer l'opérateur comme à la limite L^2 d'une suite de fonctions $\mathcal{C}_c^\infty(\mathbf{R}^d, \mathbf{R})$ (et obtenir ainsi, par continuité, la valeur 0). Les constantes ne sont pas les seules fonctions pour lesquelles un tel prolongement est permis. Bien que, comme souvent, travailler sur l'espace $\mathcal{C}_c^\infty(\mathbf{R}^d, \mathbf{R})$ s'avère fort commode d'un point de vue calculatoire, il est intéressant de déterminer le plus grand domaine possible pour les opérateurs. À ce titre, on définit la norme suivante.

Définition 1.1.4. Soit $f \in \mathcal{C}_c^\infty(\mathbf{R}^d, \mathbf{R})$, on note :

$$\|f\|_{\mathbf{L}} := \left(\|f\|_2^2 + \|\mathbf{L}f\|_2^2\right)^{1/2}.$$

Cette norme est bien définie sur $\mathcal{C}_c^\infty(\mathbf{R}^d, \mathbf{R})$, et c'est bien une vraie norme. On définit alors le domaine de $\mathbf{L}$ comme suit.

Définition 1.1.5. On définit $\mathcal{D}(\mathbf{L})$ comme l'adhérence de $\mathcal{C}_c^\infty(\mathbf{R}^d, \mathbf{R})$ dans $L^2(\mu)$ pour la norme $\|\cdot\|_{\mathbf{L}}$.

En d'autre termes, toute fonction $f \in \mathcal{D}(\mathbf{L})$ peut être vue comme la limite au sens de la norme $\|\cdot\|_{\mathbf{L}}$ d'une suite de fonctions de $\mathcal{C}_c^\infty(\mathbf{R}^d, \mathbf{R})$. Ce choix de norme assure de plus que $\mathcal{D}(\mathbf{L}) \subset L^2(\mu)$, et permet entre autres de prendre en compte les constantes. L'opérateur $\mathbf{L}$ est donc naturellement étendu à ce domaine (et on le notera toujours $\mathbf{L}$). Si V est assez régulier (ce qui sera le cas dans la suite), alors il n'y est plus symétrique, mais *essentiellement auto-adjoint*, la propriété de symétrie étant vérifiée sur $\mathcal{C}_c^\infty(\mathbf{R}^d, \mathbf{R})$, qui est un sous-espace dense de $\mathcal{D}(\mathbf{L})$ (en lien avec la notion d'hypo-ellipticité, voir [8] §1.15.6). La plupart des résultats seront présentés sur $\mathcal{D}(\mathbf{L})$, mais les calculs sur $\mathcal{C}_c^\infty(\mathbf{R}^d, \mathbf{R})$, pour alléger les preuves.

On peut également s'intéresser au domaine $\mathcal{D}(\mathcal{E})$ de la forme de Dirichlet $\mathcal{E}$. Il ne coïncide pas avec celui de $\mathbf{L}$, mais peut être de même obtenu par complétion de $\mathcal{C}_c^\infty(\mathbf{R}^d, \mathbf{R})$ pour une certaine norme. La norme naturelle est ici, vue la définition de $\mathcal{E}$, la norme de $H^1(\mu)$ (qui permet d'inclure les fonctions dont le carré du gradient est intégrable). Finalement, $\mathcal{D}(\mathcal{E})$ correspond à l'adhérence de $\mathcal{C}_c^\infty(\mathbf{R}^d, \mathbf{R})$ dans $H^1(\mu)$, et coïncide donc avec l'espace de Sobolev $H_0^1(\mu)$.

Semi-groupe et processus de diffusion

L'opérateur $\mathbf{L}$ est un opérateur différentiel du second ordre. De plus, la proposition précédente implique que $-\mathbf{L}$ est un opérateur elliptique. Il peut donc être naturel de s'intéresser à l'équation d'évolution associée (surtout dans la perspective d'un lien avec la physique), l'équation parabolique suivante :

$$\begin{cases} \partial_t u(t,x) = \mathbf{L}u(t,x), \quad t \geq 0, x \in \mathbf{R}^d \\ u(0,\cdot) = f. \end{cases} \tag{E}$$

On peut considérer que la condition initiale est suffisamment régulière (au moins de classe $\mathcal{C}^2$) et bornée. On ne traitera pas (ou peu) des considérations liées à ce type d'équations aux dérivées partielles, le lecteur intéressé trouvera ces informations dans la monographie de référence de Lawrence Evans [34] (chapitre 7). De même que l'opérateur $\mathbf{L}$ est l'analogue du laplacien pour la mesure μ, cette équation est analogue à l'équation de la chaleur $\partial_t u = \Delta u$ (on l'appellera parfois elle-même "équation de la chaleur"). En particulier, de nombreuses propriétés de cette équation se transposent à (E). La plupart d'entre elles s'entendent plus simplement en termes de *flot*, ou *semi-groupe*.

Définition 1.1.6. L'application qui à une condition initiale f associe la trajectoire u solution de (E) est appelée *flot de l'équation* (E).

Le flot peut être perçu comme une action de groupe de $(\mathbf{R}, +)$ sur l'espace X des fonctions de $\mathbf{R}^d$ dans $\mathbf{R}$:

$$\begin{array}{rccc} \varphi : & X \times \mathbf{R} & \longrightarrow & X \\ & (f,t) & \longmapsto & u(t,\cdot). \end{array}$$

Plus précisément, on considère l'action du semi-groupe $(\mathbf{R}_+, +)$ sur l'espace X. En particulier, la propriété d'action de groupe implique la relation de composition de flots suivante :

$$\varphi(\varphi(f,t),s) = \varphi(f, t+s),$$

relation parfois nommée directement *propriété de semi-groupe* de φ. Cela suggère la présentation usuelle des semi-groupes comme suit.

Définition 1.1.7 (Semi-groupe)**.** On note $(\mathbf{P}_t)_{t\geq 0}$ la famille d'opérateurs définis, pour f suffisamment régulière, par :

$$\mathbf{P}_t f(x) = u(t,x), \quad t \geq 0, \ x \in \mathbf{R}^d,$$

où u est la solution du problème (E). C'est un *semi-groupe de Markov* au sens suivant :

1. $\mathbf{P}_0 = \mathrm{id}$ (condition initiale),
2. pour tous $s, t \in \mathbf{R}_+$, $\mathbf{P}_t \circ \mathbf{P}_s = \mathbf{P}_{t+s}$ (semi-groupe),
3. si f est positive, $\mathbf{P}_t f$ l'est aussi pour tout t (positivité),
4. pour tout $t \in \mathbf{R}_+$, $\mathbf{P}_t \mathbf{1} = \mathbf{1}$ (conservation de la masse).

Notons que la mention "de Markov" s'applique uniquement au dernier point. Cette propriété n'est pas immédiate, et renvoie à la non-explosion des processus de diffusion. On rencontre parfois la notation $\mathbf{P}_t = e^{t\mathbf{L}}$, héritée de la théorie de Hille-Yosida. Cette dernière fait clairement apparaître la propriété de semi-groupe et peut s'avérer fort utile pour écrire certains résultats, par exemple faisant intervenir la formule de Duhamel. Néanmoins, l'écriture exponentielle pourrait amener certaines confusions, et nous n'en ferons que très peu usage par la suite. Elle justifie toutefois la terminologie de *générateur de Markov*, qui s'applique à l'opérateur $\mathbf{L}$ et que nous utiliserons, elle, beaucoup.

Un tel semi-groupe est parfois appelé *semi-groupe de contraction*. En effet, s'il peut être, au contraire du générateur, pertinent de poser la question de la régularité de l'image par le semi-groupe en fonction de celle de la donnée initiale (plutôt que celle du domaine de définition), cette information est souvent difficilement accessible. En revanche, on peut facilement voir que pour toute fonction f continue bornée, pour tout $t \geq 0$,

$$\|\mathbf{P}_t f\|_\infty \leq \|f\|_\infty.$$

Ainsi, pour tout $t \geq 0$, $\mathbf{P}_t$ est une contraction sur $L^\infty(\mu)$. Une façon simple de visualiser cette propriété (et d'autres) est d'introduire le *processus de diffusion* associé à $(\mathbf{P}_t)_{t\geq 0}$. C'est un type de processus stochastiques basés sur le mouvement brownien qui présente d'étroits liens avec $\mathbf{L}$ et $(\mathbf{P}_t)_{t\geq 0}$. Tous les processus (et variables aléatoires) présents dans ce manuscrit sont définis sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbf{P})$.

Définition 1.1.8 (Processus de diffusion)**.** On considère, pour $x \in \mathbf{R}^d$, $(X_t^x)_{t\geq 0}$ la solution de l'équation différentielle stochastique suivante :

$$dX_t^x = \sqrt{2}dB_t - \nabla V(X_t^x)\,dt, \quad X_0^x = x \ \mathbf{P}\text{-p.s.}, \tag{EDS}$$

où $(B_t)_{t\geq 0}$ désigne le mouvement brownien standard sur $\mathbf{R}^d$. Un tel processus est appelé *processus de diffusion* associé à $\mathbf{L}$.

Ce processus est mesurable pour la filtration brownienne $(\mathcal{F}_t)_{t\geq 0}$ (la plus petite filtration qui rende mesurable le mouvement brownien). Les considérations d'existence et d'unicité seront abordées dans un second temps. La proposition suivante, sous réserve de bonne définition des objets impliqués, donne la représentation probabiliste du semi-groupe.

Proposition 1.1.9. *Soit $f \in \mathcal{C}^0(\mathbf{R}^d, \mathbf{R})$ une fonction bornée. Alors pour tous $t \geq 0$, $x \in \mathbf{R}^d$,*

$$\mathbf{P}_t f(x) = \mathbf{E}[f(X_t^x)],$$

où $\mathbf{E}$ désigne l'espérance par rapport à $\mathbf{P}$.

La propriété de contraction suit immédiatement de cette représentation (sous réserve qu'elle soit bien définie en tout temps). En effet, si f est une fonction continue bornée, alors pour tous $t \geq 0$, $x \in \mathbf{R}^d$,

$$|\mathbf{P}_t f(x)| = |\mathbf{E}[f(X_t^x)]| \leq \mathbf{E}[|f(X_t^x)|] \leq \|f\|_\infty,$$

d'où la contraction.

Outre ses liens avec le semi-groupe, le processus $(X_t)_{t\geq 0}$ (la condition initiale n'apparaîtra que lorsque la situation le nécessitera) possède en soi des propriétés intéressantes. La première est que, à l'instar du mouvement brownien, c'est un processus de Markov homogène : pour tous $t > s \geq 0$, pour toute fonction $f \in \mathcal{C}^0(\mathbf{R}^d, \mathbf{R})$ bornée,

$$\mathbf{E}[f(X_t) \mid \mathcal{F}_s] = \mathbf{E}[f(X_t) \mid X_s],$$

et pour tout $x \in \mathbf{R}^d$,

$$\mathbf{E}[f(X_t) \mid X_s = x] = \mathbf{E}[f(X_{t-s}) \mid X_0 = x].$$

Cette propriété indique que la loi du processus à l'instant t sachant tout le passé jusqu'à l'instant s ne dépend que de la valeur du processus à l'instant s. L'homogénéité assure que cette loi ne dépend en fait que de l'écart temporel $t - s$ et de la valeur initiale du processus. L'ensemble des processus abordés satisferont à ces deux conditions. En outre, il peut être intéressant de s'interroger sur le comportement en temps long de ce type de processus. Cette question donne lieu à la notion de mesure invariante, qui permet d'établir le lien définitif entre μ, $\mathbf{L}$, $(\mathbf{P}_t)_{t\geq 0}$ et $(X_t)_{t\geq 0}$.

Mesure invariante

La notion de mesure invariante est absolument centrale, tant dans l'étude des processus de Markov que celle des opérateurs de diffusion.

Proposition-définition 1.1.10. Soient μ, $\mathbf{L}$, $(\mathbf{P}_t)_{t\geq 0}$ et $(X_t)_{t\geq 0}$ tels que définis ci-dessus. On dit que μ est *invariante* pour $\mathbf{L}$, $(\mathbf{P}_t)_{t\geq 0}$ et $(X_t)_{t\geq 0}$ dans la mesure où :

1. pour toute fonction $f \in \mathcal{D}(\mathbf{L})$, $\displaystyle\int_{\mathbf{R}^d} \mathbf{L}f \, d\mu = 0$;
2. pour toute fonction f continue bornée, $\displaystyle\int_{\mathbf{R}^d} \mathbf{P}_t f \, d\mu = \int_{\mathbf{R}^d} f \, d\mu$;
3. si X_s suit la loi μ pour un certain $s \geq 0$, alors X_t suit la loi μ pour tout $t \geq s$.

La terminologie "invariance" est héritée du dernier point. Cette acception est d'ailleurs facile à comprendre : si la loi du processus atteint μ, alors le passage du temps laissera invariante cette distribution.

Notons que l'existence et l'unicité de la mesure invariante ne vont pas de soi. Dans le présent contexte, l'existence est assurée par la construction des différents objets, et l'unicité est obtenue simplement en raisonnant par l'absurde, grâce à la forme des mesures de Boltzmann-Gibbs.

On rencontre parfois la terminologie de "mesure stationnaire", avec le même sens que "mesure invariante". Nous ne l'emploierons pas ici, car il pourrait prêter à confusion avec un autre concept, celui de *mesure ergodique*, qui elle décrit la convergence du semi-groupe (ou du processus) en temps long.

Définition 1.1.11 (Mesure ergodique)**.** Une mesure π sur $\mathcal{B}(\mathbf{R}^d)$ est dite L^2*-ergodique* pour $(\mathbf{P}_t)_{t\geq 0}$ si pour toute fonction f continue bornée,

$$\left\| \mathbf{P}_t f - \int_{\mathbf{R}^d} f \, d\pi \right\|_{L^2(\pi)} \xrightarrow[t\to+\infty]{} 0.$$

De même, π est ergodique pour $(X_t)_{t\geq 0}$ si pour tout état initial X_0 (aléatoire ou non), toute fonction continue bornée f,

$$\mathbf{E}[f(X_t) \mid X_0] \xrightarrow[t\to+\infty]{} \int_{\mathbf{R}^d} f \, d\pi.$$

La terminologie "ergodique" est héritée de la théorie des systèmes dynamiques. Dans le présent contexte, l'existence d'une mesure ergodique indique que le processus parcourt l'espace de sorte que, au bout d'un temps assez long, chaque région de l'espace a été occupée selon un taux correspondant à sa masse pour la mesure ergodique. Le parcours de l'espace est toujours aléatoire, mais "ordonné" en temps long. Cela peut être rapproché, dans une certaine mesure, du théorème ergodique de Birkhoff [15].

Les cas les plus intéressants correspondent à une coïncidence entre mesure ergodique et mesure invariante. Une telle coïncidence ne va pas de soi, mais des critères simples existent et seront abordés, parmi d'autres hypothèses, ci-après.

Avant d'aborder les différentes hypothèses et critères auxquels nous nous référerons dans la suite, nous proposons un autre point de vue sur tous les objet introduits. Dans une perspective plus descriptive, on peut supposer donné le processus $(X_t)_{t\geq 0}$, par exemple comme réalisation d'un système physique. La connaissance des propriétés du processus est alors de première importance pour l'étude du phénomène afférent. La plupart du temps, les états du système ne sont connus qu'à travers des observables (comme la température, le volume, l'entropie, etc.), soit des fonctions $f(X_t)$ du processus. L'étude du semi-groupe est justifiée dès lors que l'on réalise un grand nombre des ces observations et que, au lieu de leur valeur précise, on s'intéresse à leur moyenne. La connaissance des propriétés du semi-groupe permet alors de comprendre (voire prédire) l'évolution des observables d'intérêt. Finalement, le générateur intervient ici comme l'opérateur régissant la dynamique du semi-groupe, et donc son évolution d'un point de vue infinitésimal. De ce point de vue-là, la mesure invariante n'est plus l'objet d'étude donné a priori, contrairement au contexte de ce travail, mais une propriété afférente du processus étudié.

Et la gaussienne dans tout ça ? Le cas gaussien possède une particularité qui justifie en partie l'intérêt qui lui est porté : le semi-groupe est explicite.

Proposition 1.1.12 (Mehler)**.** *Soit f une fonction continue bornée. Alors pour tout $t \geq 0$, $x \in \mathbf{R}^d$,*

$$\mathbf{P}_t f(x) = \int_{\mathbf{R}^d} f\left(xe^{-t} + y\sqrt{1-e^{-2t}}\right) d\gamma_d(y).$$

On vérifie immédiatement que $\partial_t \mathbf{P}_t = \mathbf{L}\mathbf{P}_t$. Cette formule permet de plus d'établir simplement la plupart des propriété de contraction et de continuité du semi-groupe, sans faire appel au processus de diffusion sous-jacent. La propriété d'ergodicité est également immédiate, et les hypothèses de régularité sur f apparaissent clairement. Notons que ce cas est exceptionnel, la plupart des semi-groupes n'admettent pas de telle représentation explicite. Cela justifie également que ce semi-groupe porte un nom spécifique : le semi-groupe d'*Ornstein-Uhlenbeck*. C'est également le nom donné à l'opérateur (noté souvent $\mathbf{L}_{OU}$) et au processus de diffusion associés. À l'instar du semi-groupe, ce dernier est lui-même explicite. En effet, par variation de la constante, on montre facilement que le processus d'Ornstein-Uhlenbeck admet la représentation suivante, pour tous $t \geq 0$, $x \in \mathbf{R}^d$:

$$X_t^x = e^{-t}x + \sqrt{2}\int_0^t e^{-(t-s)}dB_s.$$

Comme cela sera développé dans le chapitre suivant, l'intégrale ci-dessus est une martingale, et par définition de l'intégrale brownienne, c'est une variable aléatoire de loi gaussienne. Ses moments se calculent aisément :

$$\mathbf{E}[X_t^x] = e^{-t}x + \sqrt{2}\mathbf{E}\left[\int_0^t e^{-(t-s)}dB_s\right] = e^{-t}x,$$

sa matrice de covariance est, comme celle du mouvement brownien, multiple de l'identité. On a, par exemple pour la première composante :

$$\mathrm{Var}((X_t^x)_1) = 2\int_0^t e^{-2(t-s)}ds = 1 - e^{-2t}.$$

Ainsi, pour tout $t \geq 0$ et tout $x \in \mathbf{R}^d$, X_t^x suit une loi gaussienne $\mathcal{N}_d(e^{-t}x, (1-e^{-2t})I_d)$. De cette formule découlent la plupart des propriétés du processus : la bonne définition de l'équation, l'ergodicité, la représentation du semi-groupe, etc. Notons que cette présentation permet la mise en œuvre d'approximations numériques facilitées.

Les points de calcul stochastique nécessaires à la compréhension de ces différents points seront abordés dans le chapitre suivant. Dans la suite, on s'intéressera à la théorie de Bakry-Émery, qui a servi de terreau fertile au présent travail et à de nombreux autres.

1.1.3 La théorie de Bakry-Émery

Cette théorie, d'essence géométrique et énoncée originellement dans un cadre riemannien par Dominique Bakry et Michel Émery en 1985 [10], se base sur l'introduction de deux opérateurs : le *carré du champ* $\mathbf{\Gamma}$ et le *carré du champ itéré* $\mathbf{\Gamma}_2$. Ce dernier est si important dans l'établissement du critère de Bakry-Émery qu'il est parfois appelé "critère $\mathbf{\Gamma}_2$". Nous introduirons donc dans un premier temps ces deux opérateurs, avant d'énoncer le critère dans la plus grande généralité. Le cas euclidien avec mesure de Boltzmann-Gibbs, qui nous intéresse plus particulièrement, sera développé en suivant. Une présentation générale de la théorie pourra être trouvée dans [8], partie I.

Présentation générale

Dans cette première section, nous considérons un opérateur de diffusion général $\mathbf{L} = \mathrm{div} \circ \nabla$. Cet opérateur peut être défini sur un espace (topologique) mesuré assez général, moyennant une structure différentielle. On peut penser à $\mathbf{L}$ de la forme de la Définition 1.1.2 pour fixer les idées, mais cet opérateur peut très bien être défini sur une variété différentielle assez régulière (comme la sphère) ou même, via des opérateurs aux différences, sur des espaces discrets.

On considère donc ici un espace topologique mesuré $(X, \mathcal{B}(X), \mu)$, muni d'une structure différentielle, dont la différentielle sera toujours notée ∇. La forme de l'opérateur $\mathbf{L}$ implique en particulier qu'il est symétrique sur $\mathcal{C}^\infty(X, \mathbf{R})$.

Définition 1.1.13 (Carré du champ)**.** Soient $f, g \in \mathcal{C}^\infty(X, \mathbf{R})$. On définit le *carré du champ* associé à $\mathbf{L}$ par :

$$\mathbf{\Gamma}(f, g) = \frac{1}{2}[\mathbf{L}(fg) - f\mathbf{L}g - g\mathbf{L}f].$$

De même, le carré du champ itéré est donné par :

$$\mathbf{\Gamma}_2(f, g) = \frac{1}{2}[\mathbf{L}\mathbf{\Gamma}(f, g) - \mathbf{\Gamma}(f, \mathbf{L}g) - \mathbf{\Gamma}(\mathbf{L}f, g)].$$

Ces deux opérateurs sont obtenus, en quelque sorte, par des identités de polarisation. On notera souvent dans la suite, pour $f \in \mathcal{C}^\infty(X, \mathbf{R})$,

$$\mathbf{\Gamma}(f) := \mathbf{\Gamma}(f, f) = \frac{1}{2}\mathbf{L}(f^2) - f\mathbf{L}f,$$

et

$$\mathbf{\Gamma}_2(f) := \mathbf{\Gamma}_2(f, f) = \frac{1}{2}\mathbf{L}\mathbf{\Gamma}(f) - \mathbf{\Gamma}(f, \mathbf{L}f).$$

Le domaine de définition de ces opérateurs peut, à l'instar de celui de $\mathbf{L}$, être étendu à un sous-domaine de $\mathcal{D}(\mathbf{L})$. Cela ne sera pas discuté dans le cas général, mais nous y reviendrons dans le cas euclidien. Notons également que cette construction peut s'itérer en définissant une suite d'opérateurs $(\mathbf{\Gamma}_n)_{n\geq 0}$ par récurrence. Certains résultats ont été obtenus dans cette direction, mais elle ne sera pas développée ici.

Ces deux opérateurs permettent d'écrire finalement la *condition de courbure-dimension*, qui donne naissance au critère éponyme.

Définition 1.1.14 (Condition de courbure-dimension)**.** L'opérateur $\mathbf{L}$ satisfait la *condition de courbure-dimension* $CD(\rho, n)$, pour $\rho \in \mathbf{R}$ et $n \in [1, \infty)$, si pour toute fonction $f \in \mathcal{C}^\infty(X, \mathbf{R})$, on a l'inégalité suivante :

$$\mathbf{\Gamma}_2(f) \geq \rho\mathbf{\Gamma}(f) + \frac{1}{n}(\mathbf{L}f)^2.$$

On dira que $\mathbf{L}$ vérifie la *condition de courbure* $CD(\rho, \infty)$, $\rho \in \mathbf{R}$, si pour toute fonction $f \in \mathcal{C}^\infty(X, \mathbf{R})$,

$$\mathbf{\Gamma}_2(f) \geq \rho\mathbf{\Gamma}(f).$$

La terminologie de "courbure dimension" est héritée du contexte riemannien, où $\mathbf{L}$ est l'opérateur de Laplace-Beltrami. Dans ce cadre, le critère est satisfait dès que le tenseur de Ricci est uniformément minoré par ρ et n est supérieur à la dimension de la variété. On parle alors de variété à courbure minorée, et ce terme a notamment servi à généraliser le concept à des espaces métriques mesurés, privés de structure différentielle.

Comme il sera développé ultérieurement, cette condition est reliée à de nombreuses propriétés du générateur $\mathbf{L}$, notamment aux différentes inégalités fonctionnelles qu'il peut satisfaire. La section suivante traite du cas euclidien, et permettra d'établir les premiers critères intéressants dans notre situation, notamment sur l'existence des solutions aux équations différentielles stochastiques.

Le cas de l'espace euclidien

Nous revenons à présent au cas $X = \mathbf{R}^d$, $d\mu(x) \propto e^{-V(x)}dx$ et $\mathbf{L} = \Delta - \nabla V \cdot \nabla$. Dans ce cas, les deux opérateurs carrés du champ se réécrivent assez simplement.

Proposition 1.1.15 (Carré du champ euclidien)**.** *Soient $f, g \in \mathcal{C}^\infty(\mathbf{R}^d, \mathbf{R})$. On a alors*

$$\mathbf{\Gamma}(f,g) = \nabla f \cdot \nabla g, \quad \mathbf{\Gamma}(f) = |\nabla f|^2,$$

et

$$\mathbf{\Gamma}_2(f,g) = \langle \nabla^2 f, \nabla^2 g \rangle_{HS} + \nabla f \cdot \nabla^2 V \nabla g, \quad \mathbf{\Gamma}_2(f) = \|\nabla^2 f\|_{HS}^2 + \nabla f \cdot \nabla^2 V \nabla f.$$

Il apparaît que $\mathbf{\Gamma}$ et $\mathbf{\Gamma}_2$ sont des opérateurs d'ordre respectif 1 et 2. En fait, $\mathbf{\Gamma}$ (resp. $\mathbf{\Gamma}_2$) mesure le défaut de dérivation (le fait de satisfaire la formule de Leibnitz) de $\mathbf{L}$ (resp. de $\mathbf{L\Gamma}$).

La condition de courbure-dimension $CD(\rho, n)$ se réécrit alors :

$$\|\nabla^2 f\|_{HS}^2 + \nabla f \cdot \nabla^2 V \nabla f \geq \rho |\nabla f|^2 + \frac{1}{n}(\mathbf{L}f)^2.$$

Notons que, dans ce contexte, le tenseur de Ricci est constant égal à 0. La condition de courbure ne peut plus alors s'entendre de ce point de vue-là, mais est à comprendre en lien avec le potentiel. En effet, la condition $CD(\rho, \infty)$ s'écrit en particulier :

$$\|\nabla^2 f\|_{HS}^2 + \nabla f \cdot \nabla^2 V \nabla f \geq \rho |\nabla f|^2,$$

et on établit alors facilement la proposition suivante.

Proposition 1.1.16 (Condition $CD(\rho, \infty)$)**.** *L'opérateur $\mathbf{L}$ satisfait $CD(\rho, \infty)$, $\rho \in \mathbf{R}$, si et seulement si $\nabla^2 V \geq \rho I_d$.*

L'inégalité $\nabla^2 V \geq \rho I_d$ est à comprendre uniformément sur $\mathbf{R}^d$, soit $\nabla^2 V(x) \geq \rho I_d$ pour tout $x \in \mathbf{R}^d$.

Dans le cas où ρ est (strictement) positif, cela revient à la convexité (uniforme) du potentiel. Dans le cas général ($\rho \in \mathbf{R}$), le critère indique que le potentiel ne doit pas être "trop concave". Avant de traiter des inégalités fonctionnelles liées à la courbure-dimension, la section suivante établit les principales hypothèses techniques qui seront utilisées dans la suite, en lien avec $\mathbf{L}$, le semi-groupe ou le potentiel.

Et la gaussienne dans tout ça ? Le cas gaussien est souvent considéré comme le cas "de référence" pour la courbure dimension. En effet, on a ici $\nabla^2 V = I_d$, donc γ_d satisfait la condition $CD(1,\infty)$. Comme cela sera développé ultérieurement, la plupart des propriétés de γ_d découleront de ce fait directement et les estimations fournies seront, exceptionnellement, optimales.

1.1.4 Hypothèses techniques

Consciencieusement éludées tout au long de cette introduction, nous établissons ci-dessous les hypothèses techniques nécessaires pour la suite du travail.

La première est que la mesure d'intérêt μ est une mesure de probabilité. Cela implique nécessairement que $e^{-V} \in L^1(dx)$.

Une des propriétés importantes utilisées précédemment est la *complétude* de l'espace pour $\mathbf{\Gamma}$ au sens suivant : il existe une suite de fonctions $(\chi_n)_{n\geq 0} \subset \mathcal{C}_c^\infty(\mathbf{R}^d, \mathbf{R})$ telle que $\chi_n \xrightarrow[n\to+\infty]{} \mathbf{1}$ simplement et $\mathbf{\Gamma}(\chi_n) \leq 1/n$ pour tout $n \in \mathbf{N}^*$. Cette hypothèse est satisfaite pour $\mathbf{R}^d$, mais nécessite d'être mentionnée pour des espaces moins usuels. C'est elle en particulier qui permet de compléter le domaine de $\mathbf{L}$ et d'inclure les constantes.

Les hypothèses sur le semi-groupe peuvent être reliées à celles sur le processus. En effet, la Proposition 1.1.9 assure que, pour $(X_t^x)_{t\geq 0}$ et $(\mathbf{P}_t)_{t\geq 0}$ tels que définis précédemment, on a $\mathbf{P}_t f(x) = \mathbf{E}[f(X_t^x)]$, sous réserve de l'unicité de la solution de (E). On considérera donc cette hypothèse pour la suite.

Cette écriture n'est néanmoins valide que si le processus lui-même existe en tout temps. En effet, les processus de diffusion existent a priori uniquement jusqu'à un temps d'explosion τ_e, qui a priori n'est pas toujours égal à l'infini. Il en est de même pour les équations différentielles ordinaires, mais le temps d'explosion est ici une variable aléatoire (c'est même un temps d'arrêt associé à $(\mathcal{F}_t)_{t\geq 0}$). Cette question est en général très difficile, mais une réponse simple est apportée par le critère de courbure dimension. En effet, elle est équivalente à l'hypothèse de préservation de la masse ($\mathbf{P}_t\mathbf{1} = \mathbf{1}$), et dans ce cadre, Bakry a apporté la réponse suivante en 1986 [5].

Proposition 1.1.17 (Non-explosion)**.** *Soient $(\mathbf{P}_t)_{t\geq 0}$ et $(X_t^x)_{t\geq 0}$ définis comme précédemment. S'il existe $\rho \in \mathbf{R}$ tel que $\mathbf{L}$ satisfait $CD(\rho,\infty)$, alors le processus n'explose pas.*

Cette propriété remarquablement synthétique permet de passer outre les critères nécessitant de la régularité sur les coefficients de l'équation (EDS) ou les critères basés sur des fonctions de Lyapunov.

La dernière propriété à vérifier est l'ergodicité, qui assure la convergence en temps long du semi-groupe (comme du processus). Elle peut être simplement comprise à partir du générateur, en particulier de sa décomposition spectrale. En effet, de même que les matrices symétriques peuvent être décomposées en base orthonormée, les opérateurs de diffusion admettent une décomposition par projecteurs spectraux : pour tout $f \in \mathcal{D}(\mathbf{L})$, on a :

$$-\mathbf{L}f = \int_0^{+\infty} \lambda dP_\lambda f,$$

où l'opposé de $\mathbf{L}$ est considéré pour avoir un spectre dans $\mathbf{R}_+$ (voir [8], chapitre 3). L'opérateur P_λ représente la projection orthogonale dans $L^2(\mu)$ de f sur l'espace caractéristique associé à λ. Le semi-groupe hérite de cette décomposition, via la formule suivante :

$$\mathbf{P}_t f = \int_0^{+\infty} e^{-\lambda t} dP_\lambda f.$$

Il est dès lors plus aisé de comprendre le comportement en temps long de $(\mathbf{P}_t)_{t\geq 0}$. En effet, passer à la limite dans l'égalité ci-dessus donne :

$$\lim_{t\to+\infty} \mathbf{P}_t f = E_0(f),$$

étant donné que $e^{-\lambda t} \to 0$ dès que $\lambda \neq 0$. Toutes ces limites et identités sont à comprendre dans $L^2(\mu)$. Ainsi, $\mathbf{P}_t f$ converge vers la projection de f sur le sous-espace des fonctions qui annulent $\mathbf{L}$. Il a déjà été établi que les fonctions constantes appartiennent à ce sous-espace. Or, l'intégrale de f contre μ correspond exactement à la projection orthogonale de f sur le sous-espace des fonctions constantes. Dès lors on dégage le critère d'ergodicité suivant : le semi-groupe est ergodique dès que $\mathbf{L}f = 0$ implique que f est constante. Dans ce cas-là, par unicité, la mesure invariante est également ergodique.

Résumons les hypothèses obtenues :

- **(H1)** $e^{-V} \in L^1(dx)$ (mesure de probabilité) ;
- **(H2)** complétude pour $\mathbf{\Gamma}$;
- **(H3)** unicité de la solution de (E) ;
- **(H4)** $CD(\rho, \infty)$, $\rho \in \mathbf{R}$ (non-explosion/conservation de la masse) ;
- **(H5)** $\mathbf{L}f = 0$ implique que f est constante (ergodicité).

Sauf mention contraire, toutes ces hypothèses seront considérées implicitement comme valides dans la suite. Elles pourront être affinées ou renforcées au besoin, auquel cas la nouvelle hypothèse sera clairement indiquée.

Dans la partie suivante, nous décrivons les principales inégalités fonctionnelles qui seront étudiées par la suite, et la dernière partie est dédiée à la présentation des relations d'entrelacement et des résultats principaux de la thèse. Encore une fois, sauf mention contraire, les objets μ, $\mathbf{L}$, $(\mathbf{P}_t)_{t\geq 0}$ et $(X_t)_{t\geq 0}$ seront toujours ceux introduits au début de cette partie.

1.2 Inégalités fonctionnelles

Le but premier de cette thèse était l'application des techniques d'entrelacement à l'étude de certaines inégalités fonctionnelles. Cette étude constitue un large pan de l'analyse fonctionnelle, et a depuis longtemps été une thématique très forte à Toulouse, notamment depuis la fin des années 80 autour de Dominique Bakry et Michel Ledoux. La durée de la thèse ne suffit pas à explorer l'intégralité du domaine, c'est pourquoi nous nous concentrerons sur deux grandes familles d'inégalités : les inégalités de Poincaré et de Sobolev logarithmique. Elles sont présentées dans la suite, puis reliées à la condition de courbure-dimension introduite ci-dessus. Les travaux originaux de cette

thèse seront présentés en fin de chapitre, aussi l'ensemble des résultats énoncés ci-après seront clairement référencés dans la littérature.

1.2.1 Inégalité de Poincaré

Probablement la plus connue des inégalités abordées, l'inégalité de Poincaré a été introduite par Henri Poincaré en 1890 dans un article dédié aux équations aux dérivées partielles de la physique mathématique [72]. Abondamment développée depuis lors, elle est devenue un outil proéminent en analyse des équations aux dérivées partielles, mais aussi en analyse spectrale. En effet, elle possède un lien très fort avec le spectre des opérateurs intervenant dans les équations concernées. Généralement formulée dans des espaces de Sobolev sur des domaines de $\mathbf{R}^d$, nous nous concentrerons sur une version légèrement différente, mais plus adaptée aux problématiques probabilistes. Une présentation détaillée de la question, ainsi que les preuves des résultats énoncés ci-dessous peuvent être trouvés dans [8], chapitre 4.

Généralités

L'idée générale de cette inégalité est de contrôler la variance d'une fonction par un terme d'énergie, classiquement la norme L^2 de son gradient, à travers une constante qui dépend uniquement de la mesure. On rappelle donc d'abord la définition de la variance d'une fonction, avant de présenter l'inégalité.

Définition 1.2.1 (Variance)**.** Soit $f \in L^2(\mu)$. On définit sa *variance* comme son moment centré d'ordre deux :

$$\mathrm{Var}_\mu(f) = \int_{\mathbf{R}^d} \left(f - \int_{\mathbf{R}^d} f \, d\mu \right)^2 d\mu = \int_{\mathbf{R}^d} f^2 d\mu - \left(\int_{\mathbf{R}^d} f \, d\mu \right)^2 .$$

L'inégalité de Poincaré s'énonce comme suit.

Définition 1.2.2 (Inégalité de Poincaré)**.** La mesure μ satisfait une *inégalité de Poincaré* de constante $c > 0$, notée $IP(c)$, si pour toute fonction $f \in \mathcal{D}(\mathcal{E})$,

$$\mathrm{Var}_\mu(f) \leq c \int_{\mathbf{R}^d} |\nabla f|^2 d\mu.$$

La plus petite constante pour laquelle cette inégalité est vérifiée est fréquemment nommée "constante de Poincaré", notée $c_P(\mu)$.

Notons que l'espace des fonctions test considérées est important. En effet, la valeur de la constante optimale peut dépendre du choix de cet espace, qu'il est donc bon de fixer. Il est choisi ici pour donner un sens aux deux termes de l'inégalité, le second étant égal à la forme de Dirichlet $\mathcal{E}$ définie en début de chapitre, et pour permettre les intégrations par parties.

Comme indiqué ci-dessus, l'inégalité de Poincaré a un lien très fort avec les inégalités de trou spectral. On dit que l'opérateur $-\mathbf{L}$ admet un *trou spectral* s'il existe $\varepsilon > 0$

tel que pour toute valeur propre $\lambda \in \sigma(-\mathbf{L})$ non nulle, $|\lambda_0 - \lambda| > \varepsilon$. La première valeur propre λ_0 de $-\mathbf{L}$ est connue, et vaut toujours 0, les fonctions constantes étant les fonctions propres associées. Le trou spectral indique donc que 0 n'est pas un point d'accumulation du spectre.

Proposition 1.2.3 (Trou spectral)**.** *L'opérateur $-\mathbf{L}$ admet un trou spectral d'ordre $1/c$, $c > 0$, si et seulement si μ satisfait une $IP(c)$.*

La preuve de ce résultat ne présente pas de difficultés particulières, en s'appuyant sur la décomposition spectrale introduite plus tôt.

Cette équivalence amène à parfois appeler directement l'inégalité de Poincaré "inégalité de trou spectral". C'est en particulier un point très important de l'étude spectrale de $\mathbf{L}$, comme cela sera mis en avant dans les chapitres suivants.

Notons que cette propriété ne dit rien sur le reste du spectre. À ce titre, deux exemples extrêmes illustrent des situations très différentes :

- le cas gaussien : le spectre de $\mathbf{L}_{OU}$ est discret (aucune des valeurs propres n'est un point d'accumulation, et la décomposition spectrale fournit une "vraie" base hilbertienne de $L^2(\mu)$) ;
- le cas exponentiel : si $\mu(dx) \propto e^{-|x|}dx$, alors μ satisfait une $IP(4)$, et $\mathbf{L}$ possède donc bien un trou spectral, mais le reste de son spectre est continu (on montre que $\sigma(-\mathbf{L}) = \{0\} \cup [1/4, +\infty)$), toutes les autres valeurs propres sont des points d'accumulation.

Ces différents comportements seront plus précisément étudiés dans le chapitre dédié (voir également [8], §4.1 et §4.4).

La seconde propriété en lien fort avec l'inégalité de Poincaré est l'ergodicité. En effet, l'inégalité de Poincaré assure la décroissance exponentielle de la norme L^2 du semi-groupe.

Proposition 1.2.4. *La mesure μ satisfait une $IP(c)$, $c > 0$, si et seulement si, pour toute fonction $f \in L^2(\mu)$, pour tout $t \geq 0$,*

$$\left\| \mathbf{P}_t f - \int_{\mathbf{R}^d} f \, d\mu \right\|_2^2 = \mathrm{Var}_\mu(\mathbf{P}_t f) \leq e^{-2t/c} \mathrm{Var}_\mu(f).$$

Cette preuve repose sur un argument très classique lié à l'utilisation du lemme de Grönwall.

Au regard des liens étroits entre l'inégalité de Poincaré et les propriétés de $\mathbf{L}$, il est souvent pertinent de connaître la valeur de la constante optimale $c_P(\mu)$. Malheureusement, cette dernière est très rarement accessible explicitement. Certaines propriétés permettent néanmoins d'en faire des estimations, parfois optimales.

Estimation de la constante de Poincaré

Une des propriétés que l'inégalité de Poincaré a le bon goût de satisfaire, et qui facilite grandement l'estimation de la constante éponyme, est la propriété de tensorisation.

Proposition 1.2.5 (Stabilité par produit)**.** *Soient ν_1 et ν_2 deux mesures de probabilité telles que ν_1 satisfait une $IP(c_1)$ et ν_2 satisfait une $IP(c_2)$. Alors la mesure produit $\nu = \nu_1 \otimes \nu_2$ satisfait une $IP(\max(c_1, c_2))$.*

Cette propriété est rendue possible par la structure de mesure produit, qui permet l'utilisation du théorème de Fubini.

Cette borne n'est pas optimale en général, mais elle permet de déduire la propriété de tensorisation : si μ satisfait une $IP(c)$, alors pour tout entier $n \in \mathbf{N}^*$, $\mu^{\otimes n}$ satisfait une $IP(c)$. Cette propriété n'est peut-être pas souvent applicable, mais a l'intérêt d'exhiber un comportement adimensionnel de la constante de Poincaré. La valeur de $c_P(\mu)$ peut en effet, au vu des propriétés énoncées ci-dessus, poser problème si elle venait à exploser quand la dimension devient trop grande.

La stabilité de l'inégalité est avantageusement utilisée couplée avec des méthodes d'estimation de la constante de Poincaré en petite dimension (un ou deux). Ces méthodes seront abordées dans le dernier chapitre, consacrée aux aspects numériques de cette inégalité.

Des critères existent néanmoins en dimension quelconque pour obtenir des estimations sur $c_P(\mu)$ hors du cadre produit. Nous traiterons ici de deux d'entre eux, parmi les plus répandus : le critère de courbure-dimension et les fonctions de Lyapunov.

Proposition 1.2.6 (Critère de courbure-dimension I)**.** *Si μ satisfait $CD(\rho, \infty)$ pour $\rho > 0$, alors elle satisfait aussi une $IP(1/\rho)$.*

La preuve de cet énoncé fait intervenir une des conséquences les plus importantes de la condition $CD(\rho, \infty)$: la borne exponentielle sur le gradient du semi-groupe.

Lemme 1.2.7. *Supposons que μ satisfait $CD(\rho, \infty)$. Alors pour toute fonction $f \in \mathcal{C}_c^\infty(\mathbf{R}^d, \mathbf{R})$, pour tout $t \geq 0$,*

$$|\nabla \mathbf{P}_t f| \leq e^{-\rho t} \mathbf{P}_t(|\nabla f|).$$

La preuve de ce résultat est assez technique, et nous renvoyons le lecteur à [8], §3.3.6, pour une démonstration précise. On utilisera plus souvent le corollaire suivant, qui résulte de l'application de l'inégalité de Jensen au lemme ci-dessus.

Corollaire 1.2.8. *Sous les mêmes hypothèses que le lemme précédent, on a pour tout $t \geq 0$,*

$$|\nabla \mathbf{P}_t f|^2 \leq e^{-2\rho t} \mathbf{P}_t(|\nabla f|^2).$$

Les estimations sur le gradient du semi-groupe rendues permises par la condition $CD(\rho, \infty)$ permettent d'obtenir facilement une inégalité de Poincaré, mais elles nécessitent l'hypothèse de convexité du potentiel pour être établies (la nécessité de cette hypothèse apparaîtra clairement dans le cadre des relations d'entrelacement). Une des façons de passer outre cette contrainte est de travailler avec des fonctions de Lyapunov.

Définition 1.2.9 (Fonction de Lyapunov)**.** Une fonction $W : \mathbf{R}^d \to [1, +\infty)$ de classe $\mathcal{C}^2$ est une *fonction de Lyapunov* s'il existe deux constantes $\theta, b > 0$ et un rayon $R > 0$ tels que :

$$\mathbf{L}W \leq -\theta W + b\mathbf{1}_{B(0,R)}.$$

Proposition 1.2.10 (Bakry, Barthe, Cattiaux, Guillin '08). *S'il existe une fonction de Lyapunov W telle que définie ci-dessus et telle que $\mu_{|B(0,R)}$ satisfait une $IP(c_R)$, alors μ satisfait une inégalité de Poincaré de constante $\frac{1}{\theta}(1 + bc_R)$.*

Nous renvoyons le lecteur à [7] pour la preuve. Notons que l'inégalité de Poincaré sur la boule est liée, comme nous le verrons plus tard, à un problème aux valeurs propres avec conditions aux bords de Neumann. Ce type de résultats est très puissant pour démontrer qu'une mesure satisfait une inégalité de Poincaré, mais des estimations quantitatives de la constantes sont parfois difficiles à déduire. Néanmoins, il permet d'établir des critères simples et qualitatifs reliés directement au potentiel V.

Corollaire 1.2.11. *Supposons que V satisfait à l'une de ces deux conditions :*

1. *il existe $\alpha > 0$ et $R \geq 0$ tels que, pour $|x| \geq R$,*

$$\langle x, \nabla V(x) \rangle \geq \alpha |x|;$$

2. *il existe $a \in (0,1)$, $c > 0$ et $R \geq 0$ tels que, pour $|x| \geq R$,*

$$a|\nabla V|^2 - \Delta V \geq c.$$

Alors le résultat précédent s'applique.

Ce type de critère est essentiellement qualitatif, et il est difficile de savoir si les constantes obtenues sont optimales. Il permet néanmoins d'établir l'existence d'une inégalité de Poincaré en s'intéressant uniquement au potentiel et en relâchant l'hypothèse de convexité de la condition $CD(\rho, \infty)$. À titre indicatif, les fonctions de Lyapunov qui vérifient les conditions sont (entre autres) de la forme $e^{\gamma|x|}$ pour la première et $e^{\gamma V + \varepsilon}$ pour la seconde.

Et la gaussienne dans tout ça ? Le cas gaussien est particulier à double titre. D'une part le trou spectral est connu explicitement, avec les fonctions propres, et donc la constante de Poincaré aussi, et d'autre part l'estimation donnée par Bakry-Émery est optimale.

En effet, comme il sera développé dans le chapitre dédié à l'analyse spectrale, le spectre de $\mathbf{L}_{OU}$ est discret et connu : $\sigma(-\mathbf{L}_{OU}) = \mathbf{N}$. Le trou spectral est donc égal à 1 et l'espace propre associé est de dimension d. Il est décrit explicitement, engendré par les formes coordonnées $g_i : x \mapsto x_i$, $i \in [\![1, d]\!]$. Le calcul est direct, pour tout $x \in \mathbf{R}^d$:

$$-\mathbf{L}_{OU} g_i(x) = -0 + x \cdot e_i = x_i = g_i(x).$$

Il s'ensuit que $c_P(\gamma_d) = 1$. Comme indiqué dans la section précédente, la borne donnée par le critère de courbure-dimension est égale à 1, ce qui coïncide avec l'optimal. Ce cas de figure est exceptionnel, et contribue à ce que l'on désigne souvent comme le "miracle gaussien". Notons que la propriété de tensorisation s'applique également à cette situation. En effet, la mesure gaussienne γ_d est le produit de d mesures γ_1 pour

lesquelles on montre de même que $c_P(\gamma_1) = 1$. L'estimation donnée par tensorisation est ici optimale.

Les estimations du Lemme 1.2.7 se déduisent aussi facilement dans ce cas, elles sont une conséquence de la formule de Mehler. En effet, pour $f \in \mathcal{C}_c^\infty(\mathbf{R}^d, \mathbf{R})$, on a :

$$\begin{aligned}\nabla \mathbf{P}_t f(x) &= \int_{\mathbf{R}^d} \nabla \left(f(xe^{-t} + y\sqrt{1-e^{-2t}}) \right) d\gamma_d(y) \\ &= e^{-t} \int_{\mathbf{R}^d} \nabla f(xe^{-t} + y\sqrt{1-e^{-2t}})\, d\gamma_d(y),\end{aligned}$$

et le lemme suit en prenant la norme de chaque terme par inégalité triangulaire. L'égalité ci-dessus peut parfois s'écrire $\nabla \mathbf{P}_t f = e^{-t} \mathbf{P}_t(\nabla f)$, ce qui est un évident abus de notation (à comprendre comme une action coordonnée par coordonnée de $\mathbf{P}_t$ sur les $\partial_i f$, voir le chapitre suivant), mais met en évidence une relation de commutation entre le gradient et le semi-groupe. Ce type de relation est appelée "entrelacement" entre ∇ et $\mathbf{P}_t$, et fera l'objet de la dernière partie de cette introduction.

Outre les techniques précitées ou les cas explicites, comme indiqué plus haut, des critères spécifiques existent en dimension 1, dus à la structure particulière de la droite réelle (comme le critère de Muckenhoupt), et seront évoqués dans le dernier chapitre. Nous terminons cette section sur l'inégalité de Poincaré par une amélioration de celle-ci : l'inégalité de Brascamp-Lieb [23].

Inégalité de Brascamp-Lieb

Proposition-définition 1.2.12. Si V est strictement convexe (i.e. $\nabla^2 V > 0$), alors μ satisfait une *inégalité de Brascamp-Lieb* : pour tout fonction $f \in \mathcal{C}_c^\infty(\mathbf{R}^d, \mathbf{R})$,

$$\mathrm{Var}_\mu(f) \leq \int_{\mathbf{R}^d} \nabla f \cdot (\nabla^2 V)^{-1} \nabla f \, d\mu.$$

On peut aisément déduire de cette inégalité l'inégalité de Poincaré sous la condition $CD(\rho, \infty)$.

En effet, si $\nabla V \geq \rho I_d$, $\rho > 0$, alors $(\nabla^2 V)^{-1}$ est toujours symétrique définie positive et vérifie, pour toute fonction f,

$$\nabla f \cdot (\nabla^2 V)^{-1} \nabla f \leq \frac{1}{\rho} |\nabla f|^2.$$

L'inégalité de Brascamp-Lieb devient alors :

$$\mathrm{Var}_\mu(f) \leq \frac{1}{\rho} \int_{\mathbf{R}^d} |\nabla f|^2 d\mu,$$

ce qui est exactement $IP(1/\rho)$. L'établissement de cette inégalité fait appel à des raisonnements que nous ne détaillerons pas ici. Elle donne accès des améliorations des résultats obtenus à partir de l'inégalité de Poincaré, comme il sera discuté plus tard.

Notons néanmoins que, dans le cas gaussien, les inégalités de Poincaré et Brascamp-Lieb coïncident.

De nombreux raffinements de l'inégalité de Poincaré existent (super-Poincaré, Poincaré à poids, etc.) mais ne seront pas abordés ici. La suite de cette partie traite d'une autre propriété, de nature assez différente : l'inégalité de Sobolev logarithmique.

1.2.2 Inégalité de Sobolev logarithmique

D'abord introduites dans le contexte de la théorie de l'information par Adriaan Stam en 1959 [81] puis conceptualisées par Leonard Gross en 1975 [42], les inégalités de Sobolev logarithmiques (ISL) ont depuis été abondamment utilisées en analyse infinie-dimensionnelle. Initialement pensées pour établir des propriétés des semi-groupes (autour notamment des travaux de Nelsen [67], Faris [35] ou Federbush [36]), on a depuis retrouvé leurs traces dans de nombreux domaines (y compris en topologie algébrique, pour la preuve de la conjecture de Poincaré [69]). Comme son nom le laisse penser, elle présente des liens avec les plongements de Sobolev, mais aussi avec l'inégalité de Poincaré ou la condition de courbure-dimension. Comme pour le paragraphe précédent, nous commencerons par une présentation générale de l'inégalité, avant donner des critères d'établissement et de discuter enfin des liens entre inégalités de Sobolev logarithmiques et de Poincaré. Le lecteur avide de science pourra consulter en [8], chapitre 5 pour tout complément d'information et pour les preuves des résultats énoncés ci-après.

Généralités

Les inégalités de Sobolev classiques font apparaître, à l'instar des inégalités de Hölder, des exposants conjugués reliés entre eux par la dimension de l'espace sous-jacent. Le but des inégalités de Sobolev logarithmiques est de s'affranchir de cette dépendance. Elles restent de même nature que les inégalités classiques (contrôle d'une fonctionnelle par une énergie), mais font intervenir une notion d'entropie.

Définition 1.2.13 (Entropie)**.** Soit $f : \mathbf{R}^d \to \mathbf{R}_+^*$ une fonction assez régulière. On définit son *entropie* par le "moment centré logarithmique" suivant :

$$\mathrm{Ent}_\mu(f) = \int_{\mathbf{R}^d} f \log f \, d\mu - \int_{\mathbf{R}^d} f \, d\mu \log\left(\int_{\mathbf{R}^d} f \, d\mu\right).$$

La régularité minimale pour donner un sens à l'entropie est de supposer que $f \in L^1(\mu)$ et $\int_{\mathbf{R}^d} f|\log f| d\mu < +\infty$. L'espace de telles fonctions est parfois noté $L\log L(\mu)$.

L'idée de l'inégalité de Sobolev logarithmique est donc de contrôler l'entropie d'une fonction par la norme L^2 de son gradient, via une constante qui ne dépend que de μ. Elle s'énonce comme suit [42].

Définition 1.2.14 (Inégalité de Sobolev logarithmique)**.** La mesure μ satisfait une *inégalité de Sobolev logarithmique* de constante $c > 0$, notée $ISL(c)$, si pour toute fonction $f \in \mathcal{D}(\mathcal{E})$,

$$\mathrm{Ent}_\mu(f^2) \leq c \int_{\mathbf{R}^d} |\nabla f|^2 d\mu.$$

La plus petite constante pour laquelle cette inégalité est vérifiée est fréquemment nommée "constante de Sobolev logarithmique", notée $c_{SL}(\mu)$.

L'espace de fonctions test est également important. Ici, l'hypothèse $f \in \mathcal{D}(\mathcal{E})$ assure que les deux termes à la fois sont bien définis et permet aussi les intégrations par parties.

Si l'inégalité de Poincaré était fortement liée au trou spectral de $\mathbf{L}$, l'inégalité de Sobolev logarithmique peut se comprendre en termes d'information. En effet, si on suppose que $f : \mathbf{R}^d \to \mathbf{R}_+^*$ est la densité par rapport à μ d'une mesure de probabilité ν, alors l'entropie devient :

$$\mathrm{Ent}_\mu(f) = \int_{\mathbf{R}^d} f \log f \, d\mu = D_{KL}(\nu \| \mu),$$

où $D_{KL}(\nu\|\mu)$ est la divergence de Kullback-Leibler entre μ et ν. De même, si on définit l'information de Fisher entre μ et ν par :

$$I(\nu\|\mu) = \int_{\mathbf{R}^d} \frac{|\nabla f|^2}{f} d\mu,$$

alors l'inégalité $ISL(c)$ devient :

$$D_{KL}(\nu\|\mu) \leq \frac{c}{4} I(\nu\|\mu).$$

C'est sous cette forme qu'elle est le plus utilisée en théorie de l'information, une des premières situations où elle est apparue.

De façon analogue à l'inégalité de Poincaré, l'ISL nous renseigne sur l'ergodicité du semi-groupe.

Proposition 1.2.15. *La mesure μ satisfait une $ISL(c)$, $c > 0$, si et seulement si pour toute fonction $f \in L^1(\mu)$ d'entropie finie,*

$$\mathrm{Ent}_\mu(\mathbf{P}_t f) \leq e^{-4t/c} \mathrm{Ent}_\mu(f).$$

La preuve est analogue à celle de la Proposition 1.2.4, elle repose sur la dérivation en temps du terme de gauche et l'utilisation du lemme de Grönwall. De nature similaire à son pendant pour l'inégalité de Poincaré, l'une ou l'autre de ces propriété peut être utilisée en fonction de la régularité de la condition initiale f. Notons que la régularité L^2 exigée dans la Proposition 1.2.4 est plus restrictive que celle de la proposition ci-dessus qui est donc, de fait, plus puissante.

L'entropie est, à l'instar de la divergence de Kullback-Leibler, une mesure de dissimilarité entre lois de probabilités, et permet donc via la proposition ci-dessus d'établir

une vitesse exponentielle dans des convergences en loi. À ce titre, le théorème d'Otto-Villani décrit une relation entre l'inégalité de Sobolev logarithmique et la distance de Wasserstein 2. Couplé à la proposition précédente, ce résultat permet d'établir un lien très fort entre l'ISL et le transport optimal.

Mentionnons enfin que, dans leur contexte d'introduction par Gross, les inégalités de Sobolev logarithmiques ont été utilisées pour établir l'hypercontractivité du semi-groupe, qui améliore la contraction décrite plus haut. Cette propriété indique non seulement que $(\mathbf{P}_t)_{t\geq 0}$ est en fait une contraction sur $L^p(\mu)$, $p > 1$, mais de plus qu'il améliore la régularité de la fonction de départ : si $f \in L^p(\mu)$, alors dès que $t > 0$, on peut trouver $q(t) > p$ tel que $\mathbf{P}_t f \in L^{q(t)}(\mu)$. Dans le même esprit que la régularisation du semi-groupe de la chaleur, l'hypercontractivité garantit une régularisation dans L^p de la condition initiale.

De nombreux autres résultats font appel à l'inégalité de Sobolev logarithmique, il est impossible de tous les citer ici. Le lecteur intéressé pour consulter avantageusement [2]. Néanmoins, toutes ces propriétés sont reliées à la constante optimale $c_{SL}(\mu)$, qu'il est important, sinon de connaître explicitement, au moins de savoir estimer.

Estimation de la constante de Sobolev logarithmique

Comme cela a été mentionné, la constante optimale de l'inégalité de Sobolev logarithmique a le bon goût de bien se comporter quand la dimension augmente. Le théorème de Fubini permet d'établir, comme pour l'inégalité de Poincaré, le résultat suivant.

Proposition 1.2.16 (Stabilité par produit)**.** *Soient ν_1 et ν_2 deux mesures de probabilité telles que ν_1 satisfait une $ISL(c_1)$ et ν_2 satisfait une $ISL(c_2)$. Alors la mesure produit $\nu = \nu_1 \otimes \nu_2$ satisfait une $ISL(\max(c_1, c_2))$.*

Cette propriété implique donc que si μ satisfait une $ISL(c)$, $c > 0$, alors pour tout $n \in \mathbf{N}^*$, $\mu^{\otimes n}$ satisfait également une $ISL(c)$. Bien qu'inapplicable en-dehors du contexte des mesures produits, ce critère oriente néanmoins les recherches vers une comportement adimensionnel de la constante optimale. Ce point sera plus amplement discuté dans le chapitre suivant.

Pour des mesures plus générales, il existe de nombreuses techniques d'estimation de $c_{SL}(\mu)$. Nous allons ici nous concentrer sur deux d'entre elles : le critère de courbure-dimension et la méthode perturbative de Holley-Stroock.

Proposition 1.2.17 (Critère de courbure-dimension II)**.** *Si μ satisfait $CD(\rho, \infty)$ pour $\rho > 0$, alors elle satisfait aussi une $ISL(2/\rho)$.*

La preuve utilise, à l'instar de celle de la Proposition 1.2.6, les estimations sur le gradient du semi-groupe du Lemme 1.2.7. Sa structure est très similaire à celle que l'on rencontrera dans le chapitre suivant dans le chapitre suivant, dans le cadre duquel on pourra d'ailleurs retrouver ce résultat.

Comme précédemment, le critère de courbure-dimension permet d'établir facilement une estimation de la constante optimale, bien que celle-ci ne soit en générale pas

optimale. Mais encore une fois, l'hypothèse de convexité exigée sur le potentiel est assez restrictive. Elle peut toutefois être relâchée si μ est une perturbation bornée d'une mesure qui elle-même satisfait une ISL : c'est le critère de Holley-Stroock (dont une version assez similaire existe pour l'inégalité de Poincaré).

Proposition 1.2.18 (Holley, Stroock [46]). *S'il existe $\Phi : \mathbf{R}^d \to \mathbf{R}$ une fonction continue bornée et une mesure de probabilité ν telle que $d\mu \propto e^{\Phi} d\nu$ et ν satisfait une $ISL(c)$, $c > 0$, alors μ satisfait une ISL de constante $ce^{2(\sup(\Phi)-\inf(\Phi))}$.*

Ce critère peut naturellement se combiner au précédent et se comprendre uniquement au niveau du potentiel : s'il existe une fonction uniformément convexe U avec $\nabla^2 U \geq \rho I_d$, $\rho > 0$ et une fonction bornée Φ telles que $V = U + \Phi$, alors μ satisfait une ISL de constante $e^{2(\sup(\Phi)-\inf(\Phi))}/\rho$. Notons que le 2 dans l'expression précédente peut être omis sous certaines conditions, voir [8] §5.1.2. La valeur de la constante obtenue dépend en tout cas fortement de Φ, et en fonction du comportement dimensionnel de cette dernière, la borne exponentielle peut être loin de l'optimal. Le chapitre suivant sera dédié à une exploration plus profonde de ces considérations.

Et la gaussienne dans tout ça ? Comme pour l'inégalité de Poincaré, le miracle gaussien se produit pour l'ISL. En effet, la mesure γ_d satisfait une $ISL(2)$, et la borne donnée par le critère de courbure-dimension est donc optimale. Plusieurs approches permettent d'établir ce résultat, la plus simple étant d'identifier les fonctions extrémales en dimension 1 (celles qui saturent l'inégalité) et d'utiliser la tensorisation.

Si les fonctions linéaires saturaient l'inégalité de Poincaré, ce sont ici les exponentielles de fonctions linéaires $x \mapsto e^{ax}$ qui réalisent l'égalité. En effet, pour $f(x) = e^{ax}$, un calcul direct montre que :

$$\mathrm{Ent}_{\gamma_1}(f^2) = 2a^2 e^{2a^2}$$

et

$$\int_{\mathbf{R}^d} (f')^2 d\gamma_1 = a^2 e^{2a^2},$$

ce qui, combiné à l'estimation de Bakry-Émery $c_{SL}(\gamma_1) \leq 2$, établit l'égalité. Le résultat en dimension quelconque suit par tensorisation.

De nombreuses autres techniques d'établissement des ISL existent, et certaines seront développées plus avant dans le chapitre suivant. Un des intérêts de l'inégalité de Sobolev logarithmique est son lien avec l'inégalité de Poincaré introduite plus tôt.

Lien avec l'inégalité de Poincaré

Il a déjà été mentionné que l'inégalité de Sobolev logarithmique est "plus forte" que celle de Poincaré. Plus précisément, elle implique cette dernière selon les modalités suivantes.

Proposition 1.2.19. *Si μ satisfait une $ISL(c)$, $c > 0$, alors μ satisfait également une $IP(c/2)$.*

La démonstration consiste à appliquer l'ISL à la fonction $(1+\varepsilon f)^2$, pour $f \in \mathcal{D}(\mathcal{E})$, couplée à un développement limité du logarithme.

L'inégalité de Sobolev logarithmique est en fait "strictement" plus forte que l'inégalité de Poincaré dans le sens où toutes les mesures qui satisfont l'inégalité de Poincaré ne satisfont pas d'ISL. L'exemple le plus connu est celui de la loi de Laplace (ou double exponentielle). Sa densité par rapport à la mesure de Lebesgue est égale à $e^{-|x|}/2$, et elle satisfait une $IP(4)$. Néanmoins, elle ne satisfait pas d'ISL. Ce point peut être montré élégamment via la concentration de la mesure, voir par exemple [59].

Notons que ces deux inégalités sont des cas particuliers de Φ-entropies (voir par exemple [8], §7.6), qui permettent par exemple d'interpoler entre Poincaré et Sobolev via les inégalités de Latała-Oleszkiewicz [58].

Les inégalités présentées sont également étroitement liées à des propriétés d'intégrabilité exponentielle, et sont donc des outils de choix dans l'étude de la concentration de la mesure. Elles permettent de plus d'établir des plongements des espaces de Sobolev dans certains espaces L^p, ainsi que de préciser la structure des espaces de fonctions lipschitziennes. Ces propriétés, en lien également avec l'inégalité de Sobolev, sont présentées en détail dans [8], sections 4.4, 5.4 et 6.6., et méritent le coup d'œil.

Nous terminons cette introduction par un exposé du contexte dans lequel s'inscrit cette thèse, ainsi qu'une présentation des principaux résultats obtenus.

1.3 Contexte de la thèse et présentation des résultats

Dans cette dernière section, nous introduisons les techniques d'entrelacements dont le développement fut à l'origine de la présente thèse. Après les considérations générales sur ces relations, un bref état de l'art est dressé et les résultats originaux sont énoncés, ainsi que le plan des chapitres à venir.

1.3.1 Relations d'entrelacement

Définition générale

Le nom de "relation d'entrelacement" renvoie à une réalité assez floue et protéiforme. L'idée générale est l'étude des relations de commutation entre opérateurs, qui peuvent être de natures assez différentes.

Définition 1.3.1 (Entrelacement)**.** Deux opérateurs D_1, D_2 vérifient une *relation d'entrelacement* s'il existe un troisième opérateur Λ tel que pour toute fonction f idoine :

$$\Lambda D_1 f = D_2 \Lambda f.$$

La nature des opérateurs peut être d'une grande diversité, aucun de ceux-ci n'est par exemple a priori un opérateur différentiel. En pratique, les opérateurs D_1 et D_2 peuvent être donnés, et Λ permet d'établir une correspondance entre les deux, ou encore D_1 et Λ être prescrits et la relation donne alors des informations sur D_1 à partir de D_2, si celui-ci présente une forme intéressante.

Diverses relations de ce genre ont été établies et étudiées au cours du temps, dans de nombreux domaines, et il n'est ni possible ni pertinent d'en faire la liste, encore moins de s'intéresser à toutes (995 résultats sur arXiv math pour le mot-clé "intertwining"). Un type particulier a déjà été rencontré dans ces pages : le cas où $\Lambda = \nabla$, et $D_1 = \mathbf{P}_t$, le semi-groupe d'Ornstein-Uhlenbeck. On a vu dans ce contexte que $\nabla \mathbf{P}_t f = e^{-t}\mathbf{P}_t(\nabla f)$, ce qui, à un abus de notation près, est exactement une relation d'entrelacement entre $\mathbf{P}_t$ et $e^{-t}\mathbf{P}_t$ via ∇. Cette expression peut également s'entendre au niveau du générateur : on aura alors $\nabla \mathbf{L}_{OU} f = (\mathbf{L}_{OU} - I_d)\nabla f$, moyennant aussi un abus de notations. Ces deux relations sont équivalentes et leur généralisation donne lieu à la famille d'entrelacements qui nous intéressera dans la suite.

Définition 1.3.2 (Entrelacement gradient-générateur)**.** Soit $\mathbf{L}$ un générateur markovien. On dit qu'il existe une *relation d'entrelacement gradient-générateur* pour $\mathbf{L}$ s'il existe un opérateur $\tilde{\mathcal{L}}$ tel que pour toute fonction f assez régulière :

$$\nabla \mathbf{L} f = \tilde{\mathcal{L}}(\nabla f).$$

De même, si $(\mathbf{P}_t)_{t\geq 0}$ est un semi-groupe markovien, on dit qu'il existe une *relation d'entrelacement gradient-semi-groupe* s'il existe un autre semi-groupe $(\tilde{\mathbf{P}}_t)_{t\geq 0}$ tel que pour toute fonction f assez régulière et pour tout $t \geq 0$:

$$\nabla \mathbf{P}_t f = \tilde{\mathbf{P}}_t(\nabla f).$$

C'est à la lumière de cette dernière identité que l'on peut comprendre le mieux la terminologie d'entrelacement. En effet, cette dernière est héritée de la théorie des représentations, où un opérateur d'entrelacement permet de passer d'une représentation d'un groupe donné à une autre (et de les conjuguer si l'entrelacement est un isomorphisme). Ici, on considère des représentations du *semi*-groupe $(\mathbf{R}_+, +)$, d'une part sur l'espace vectoriel des fonctions continues ($t \mapsto \mathbf{P}_t$), d'autre part sur celui des champs de vecteurs continus ($t \mapsto \tilde{\mathbf{P}}_t$). On passe de l'une à l'autre via ∇, qui est bien linéaire. Les relations d'entrelacement sont donc, à ce titre, de vraies relations d'entrelacement de représentation de (semi-)groupes.

En-dehors du cas spécifique de la dimension 1, sur lequel nous reviendrons, les opérateurs $\tilde{\mathbf{P}}$ et $\tilde{\mathcal{L}}$ agissent en général sur des objets différents des semi-groupe et opérateur d'origine. À cette fin, on définit les opérateurs et semi-groupes diagonaux, qui agissent sur les champs de vecteurs, dans un formalisme similaire à celui de [3].

Définition 1.3.3. Soit $F \in \mathcal{C}^\infty(\mathbf{R}^d, \mathbf{R}^d)$ un champ de vecteurs. On définit le *générateur diagonal* $\mathcal{L}$ par :

$$\mathcal{L}F = (\mathbf{L}F_i)_{i\in[\![1,d]\!]}.$$

De même le *semi-groupe diagonal* $(\mathcal{P}_t)_{t\geq 0}$, si F est de surcroît à support compact :

$$\mathcal{P}_t F = (\mathbf{P}_t F_i)_{i\in[\![1,d]\!]}.$$

Les domaines de définition de ces opérateurs sont analogues à leurs pendants agissant sur des champs scalaires, de même les normes induites, construites par produit tensoriel de la norme $\|\cdot\|_{\mathbf{L}}$. Une façon parfois intéressante de voir ces opérateurs est de les représenter sous forme matricielle :

$$\mathcal{L} = \begin{pmatrix} \mathbf{L} & & 0 \\ & \ddots & \\ 0 & & \mathbf{L} \end{pmatrix},$$

auquel cas l'action de $\mathcal{L}$ sur les champs de vecteurs peut se voir comme un produit matriciel (dont la forme implique la terminologie de "diagonal"). Le semi-groupe se comprend alors comme l'exponentielle $\mathcal{P}_t = \exp(t\mathcal{L})$, dans un sens analogue à l'exponentielle matricielle classique. Ces notations nous permettent d'écrire l'entrelacement gaussien sans abus.

Proposition 1.3.4 (Entrelacement gaussien)**.** *Pour toute fonction $f \in \mathcal{D}(\mathbf{L})$,*

$$\nabla \mathbf{L}_{OU} f = (\mathcal{L}_{OU} - I_d)\nabla f,$$

et pour tout $t \geq 0$, avec $\mathbf{P}_t$ le semi-groupe d'Ornstein-Uhlenbeck,

$$\nabla \mathbf{P}_t f = e^{-t}\mathcal{P}_t(\nabla f).$$

On retrouve immédiatement les estimations du Lemme 1.2.7. L'enjeu est alors de parvenir à généraliser cette relation à d'autres mesures, y compris quand le semi-groupe n'est pas explicite. La forme de l'entrelacement n'est alors plus aussi simple, le semi-groupe d'intérêt est de nature assez différente, et nécessite l'introduction des opérateurs de Schrödinger Feynman-Kac.

Définition 1.3.5. Soit $M : \mathbf{R}^d \to \mathcal{M}_d(\mathbf{R})$ une application mesurable. Pour $F \in \mathcal{C}^\infty(\mathbf{R}^d, \mathbf{R}^d)$, on définit l'*opérateur de Schrödinger* associé à $\mathcal{L}$ et M par :

$$\mathcal{L}^M F = \mathcal{L}F - MF,$$

où MF est le produit matrice-vecteur classique. Le semi-groupe associé $(\mathcal{P}_t^M)_{t\geq 0}$ est appelé *semi-groupe de Feynman-Kac* et est défini comme le flot de l'équation $\partial_t u = \mathcal{L}^M u$.

Il est important de remarquer que, a priori, les opérateurs $\mathcal{L}^M$ et $\mathcal{P}_t^M$ ne sont *pas* markoviens. En effet, si $\mathbf{1}$ désigne le champ de vecteurs constant égal à 1, on a $\mathcal{L}^M\mathbf{1} = -M\mathbf{1} \neq 0$ (sauf si $\mathbf{1} \in \ker M(x)$ pour tout $x \in \mathbf{R}^d$, ce qui est faux en général), et donc a priori $\mathcal{P}_t^M\mathbf{1} \neq \mathbf{1}$ pour $t \geq 0$. C'est visible dans le cas gaussien : $\mathcal{L}_{OU}^{I_d}\mathbf{1} = -\mathbf{1} \neq 0$ et $e^{-t}\mathcal{P}_t\mathbf{1} = e^{-t}\mathbf{1} \neq \mathbf{1}$. On dit alors que $(\mathcal{P}_t^M)_{t\geq 0}$ est *sous-markovien*. Les propriété de semi-groupe et de continuité continuent néanmoins d'être vérifiées. En particulier, ce cadre permet d'écrire les relations d'entrelacements pour des générateurs non gaussiens.

Proposition 1.3.6 (Entrelacement général)**.** *Soit $f \in \mathcal{D}(\mathbf{L})$, on a :*

$$\nabla \mathbf{L} f = \mathcal{L}^{\nabla^2 V} \nabla f$$

et pour tout $t \geq 0$:

$$\nabla \mathbf{P}_t f = \mathcal{P}_t^{\nabla^2 V} \nabla f.$$

Notons que la deuxième égalité est en réalité une conséquence directe de la première et de la définition des semi-groupes de Feynman-Kac, et n'apporte en soi aucune information. Ce type de formule est bien étudié depuis longtemps en géométrie différentielle et en physique mathématique, et peut être relié au laplacien de Hodge et à la formule de Bochner-Lichnerowicz-Weitzenböck en géométrie riemannienne (voir par exemple [21]) ou aux laplaciens de Witten en mécanique statistique (voir [44]). Des travaux récents d'Elworthy et Li reprennent également ce principe, voir par exemple [33]. Notons que la "version originale" de l'équation de Schrödinger fait intervenir des fonctions à valeurs scalaires, pour lesquelles une écriture simple est disponible.

Proposition 1.3.7. *Soit $m : \mathbf{R}^d \to \mathbf{R}$ une application mesurable et $f \in \mathcal{C}_c^\infty(\mathbf{R}^d, \mathbf{R})$. Alors le flot $(\mathbf{P}_t^m)_{t\geq 0}$ de l'équation $\partial_u = \mathbf{L}u - mu$ pour $u(0,\cdot) = f$ est donnés pour tous $t \geq 0$ et $x \in \mathbf{R}^d$, par la représentation :*

$$\mathbf{P}_t^m f(x) = \mathbf{E}\left[f(X_t^x) e^{-\int_0^t m(X_s^x) ds} \right],$$

où $(X_t^x)_{t\geq 0}$ est le processus de diffusion introduit au début.

Un des premiers enjeux des chapitres à venir est de discuter une extension de cette représentation à des situations plus générales. Un de ses principaux intérêts, couplée aux relations d'entrelacement, est de fournir une étape intermédiaire dans les estimations du Lemme 1.2.7, pour laisser notamment place à d'éventuelles améliorations. À ce titre, un dernier outil mérite d'être introduit ici : la *h*-transformation, ou transformation de Doob (voir [8], section 1.15).

Définition 1.3.8 (Transformation de Doob)**.** Soit $h : \mathbf{R}^d \to \mathbf{R}$ une fonction assez régulière. On définit la *h-transformée* (ou *transformée de Doob*) de $\mathbf{L}$, pour toute fonction $f \in \mathcal{D}(\mathbf{L})$,

$$\mathbf{L}_h f = \frac{1}{h} \mathbf{L}(hf).$$

De même, le semi-groupe associé $(\mathbf{P}_{t,h})_{t\geq 0}$ s'écrit, pour $t \geq 0$,

$$\mathbf{P}_{t,h} f = \frac{1}{h} \mathbf{P}_t(hf).$$

Comme cela sera abordé dans la suite, l'intérêt de cette transformation est double : d'une part, insérée dans la preuve du Lemme 1.2.7, elle permet d'améliorer la borne obtenue, et d'autre part elle facilite les changements de mesure qui, au niveau des processus, permettent d'obtenir des informations.

D'une certaine façon, la transformation de Doob peut être vue comme une relation d'entrelacement : si on note τ_h l'opérateur de multiplication par h, alors la h-transformation peut se réécrire :

$$\tau_h \mathbf{L}_h = \mathbf{L}(\tau_h f).$$

En ce sens, $\mathbf{L}_h$ peut être considéré comme le conjugué de $\mathbf{L}$ par l'action de groupe multiplicatif de $\mathcal{C}^0_+(\mathbf{R}^d, \mathbf{R}) : h \mapsto \tau_h$. Notons que, en général, le générateur obtenu n'est *pas* markovien, mais un bon choix de fonction h peut néanmoins amener à des résultats intéressants.

La section suivante propose un survol des résultats récents impliquant les outils présentés ci-dessus, avant de clore l'introduction par une présentation des résultats obtenus lors de la thèse.

État de l'art

Cas discret Des relations d'entrelacements ont été établies dans le cadre des processus de naissance et de mort discrets. La plupart des définitions se transposent à ce cas discret, via des opérateurs aux différences.

Dans ce contexte, des relations de commutation ont été démontrées dès 2013 par D. Chafaï et A. Joulin pour les processus de naissance et de mort, et une condition de courbure exprimée en terme de distance de Wasserstein a permis d'obtenir une borne sur la constante optimale de Poincaré et des inégalités de Φ-entropie pour ces processus. De plus, les semi-groupes de Feynman-Kac ont été définis et utilisés pour obtenir des propriétés de contraction sur le semi-groupe de Markov associé. Des résultats ont également été obtenus autour de certaines files d'attente, voir [25].

Dans un cadre analogue, B. Cloez et C. Delplancke ont utilisé ces mêmes outils appliqués à la méthode de Chen-Stein, pour obtenir des vitesses de convergence en temps long pour de nombreux processus discrets (files d'attente, Galton-Watson, etc.). Ces résultats sont basés sur les travaux antérieurs de Chen et Barbour autour de la méthode de Stein, travaux qui laissent percevoir une ouverture analogue dans le cadre des diffusions (voir [26]).

Cas des diffusions C'est néanmoins dans le cadre des diffusions que le plus de travaux ont été menés. Les premiers résultats en dimension 1 par M. Bonnefont et A. Joulin ont pu mettre en évidence le lien entre relation d'entrelacement et transformation de Doob et ont permis à ce titre de retrouver et d'améliorer certains résultats, notamment autour du trou spectral ou des inégalités de Sobolev logarithmiques et Brascamp-Lieb [18]. En particulier, des estimations semblables au Lemme 1.2.7 ont été redémontrées simplement à partir des semi-groupes de Feynman-Kac, comme il sera développé dans le chapitre suivant.

Plus récemment, ces résultats ont pu être utilisés, toujours en dimension 1, pour aller au-delà des estimations de trou spectral. En considérant la caractérisation de Courant-Fisher pour les valeurs propres et des transformations de Doob bien choisies, les auteurs ont obtenu des encadrements sur les valeurs propres des opérateurs de diffusions,

ainsi que des estimations sur les écarts entre deux valeurs propres consécutives [17]. Une généralisation de cette méthode en dimension supérieure est discuté en chapitre deux.

Notons que ces résultats en dimension 1 peuvent également trouver écho en dimension plus grande, par exemple pour des potentiels à symétrie sphériques (tels que $V = \tilde{V}(|\cdot|)$), où l'on peut se ramener à une étude unidimensionnelle [20], ou encore dans le cas des mesures produits.

L'étude en dimension supérieure de ces relations a également apporté des résultats variés, tant autour des inégalités de covariance, Poincaré ou Brascamp-Lieb [3], que de l'étude spectrale des opérateurs de diffusion [19]. Dans ce cas, la transformation de Doob peut prendre deux aspects différents : une transformation scalaire telle que présentée dans la Définition 1.3.8, ou une transformation matricielle, donnant lieu à un gradient à poids. Ces relations de commutation spécifiques ne seront pas traitées ici mais, à titre indicatif, des résultats semblables au suivant peuvent être obtenus.

Proposition 1.3.9 (Arnaudon, Bonnefont, Joulin [3])**.** *Soit $A : \mathbf{R}^d \to GL_d(\mathbf{R})$ une application différentiable telle que $\mathcal{L}A^{-1}$ est symétrique (compris comme* **L** *appliqué à A^{-1} élément par élément). Si $\nabla^2 V - (\mathcal{L}A^{-1})A \geq \rho I_d$, $\rho \in \mathbf{R}$, alors il existe un semi-groupe de Feynman-Kac $(Q_{t,A})_{t\geq 0}$ tel que pour toute fonction $f \in \mathcal{C}_c^\infty(\mathbf{R}^d, \mathbf{R})$,*

$$A\nabla \mathbf{P}_t f = Q_{t,A}(A\nabla f).$$

Nous verrons dans le chapitre suivant que de tels résultats sont de nature assez différente de l'application de la transformation de Doob au semi-groupe de Feynman-Kac introduit précédemment. En particulier, les propriétés du commutant de A peuvent avoir une influence sur les théorèmes de représentation disponibles.

Les travaux menés dans le cadre de cette thèse se situent entièrement dans le cadre des processus de diffusion. Les principaux résultats ainsi qu'un plan d'ensemble de la thèse sont présentés dans la section suivante, qui clôt cette introduction.

1.3.2 Présentation de quelques résultats et plan de la thèse

Le but de cette thèse est d'appliquer les relations d'entrelacement introduites ci-dessus aux différents objets mentionnés en introduction, et plus spécifiquement à l'analyse spectrale et aux inégalités de Sobolev logarithmiques. Les résultats seront présentés en trois chapitres. Le premier traite des inégalités de Sobolev logarithmiques, le second de l'analyse spectrale, et le dernier s'intéresse spécifiquement aux aspects numériques du trou spectral en petite dimension.

Inégalités de Sobolev logarithmiques

Ce premier chapitre se base sur [83], article publié en 2021 dans *Electronic Journal of Probability*. Le but de ce travail était d'établir un nouveau critère et une nouvelle estimation de la constante optimale des inégalités de Sobolev logarithmiques, dans l'idée

de dépasser le critère de courbure-dimension de la Proposition 1.2.17. À cette fin, on introduit dans la formule du semi-groupe un processus perturbé par l'entremise du théorème de Girsanov, qui permet de compenser le manque de convexité du potentiel. Le premier résultat obtenu dans ce sens est un théorème de représentation des semi-groupes de Feynman-Kac qui peut s'énoncer comme suit.

Théorème. *Soit a une fonction réelle assez régulière. Il existe deux processus stochastiques réels $(X_{t,a})_{t\geq 0}$ et $(R_{t,a})_{t\geq 0}$ et un processus stochastique à valeurs matricielles $(J_t^{X_a})_{t\geq 0}$ (qui dépendent également tous de $(X_t)_{t\geq 0}$) tels que pour toute fonction f régulière, on a pour tout $t \geq 0$:*

$$\mathcal{P}_t^{\nabla^2 V}(\nabla f) = \mathbf{E}[R_{t,a} J_t^{X_a} \nabla f(X_{t,a})],$$

où $(\mathcal{P}_t^{\nabla^2 V})_{t\geq 0}$ est le semi-groupe de Feynman-Kac de la Proposition 1.3.6.

Ces processus seront explicités par la suite. Ce résultat (intermédiaire mais déjà intéressant en soi) permet d'introduire un degré de liberté à travers une fonction de perturbation. On en déduit une nouvelle borne "perturbée" κ_a sur $\nabla^2 V$, similaire à celle présente dans la condition de courbure-dimension, qui donne lieu au nouveau critère ci-dessous.

Théorème. *Soit a une fonction réelle assez régulière telle que $\kappa_a > 0$, alors μ satisfait une inégalité de Sobolev logarithmique avec $c_{LSI}(\mu) \leq C_a/\kappa_a$, pour une certaine constante $C_a > 0$ qui ne dépend que de a.*

L'introduction de la fonction de perturbation permet de compenser le manque de convexité du potentiel. Un exemple simple d'application peut être trouvé avec le potentiel quadrique : $V = |\cdot|^4/4$. La hessienne de ce potentiel s'annule en 0, et sort donc du cadre d'application du critère de Bakry-Émery du fait de ce seul point. Le comportement de la constante optimale n'est à ce jour pas connu, en particulier concernant sa dépendance en la dimension. Le critère ci-dessus permet néanmoins d'établir une borne sur cette constante qui ne dépend pas de la dimension.

Un autre exemple est traité, ainsi que la restriction de l'inégalité aux fonctions monotones, cas dans lequel la constante obtenue peut être améliorée.

Analyse spectrale

Le chapitre suivant est basé sur un travail en cours de rédaction, en collaboration avec M. Bonnefont et A. Joulin, qui a pour but d'établir des estimations fines sur les valeurs propres des opérateurs de diffusion. Un résultat récent d'Emanuel Milman [63] établit que la comparaison des hessiennes présente dans le critère de Bakry-Émery induit une comparaison des spectres des opérateurs impliqués. Plus précisément, en utilisant des outils de transport optimal, il montre que les valeurs propres d'un opérateur qui satisfait une condition $CD(\rho, \infty)$, pour $\rho > 0$, se comparent à celle d'un générateur d'Ornstein-Uhlenbeck, qui sont explicites.

L'idée est alors de retrouver ce résultat en travaillant directement sur le générateur. À cette fin, une structure algébrique est déployée pour définir une nouvelle relation

d'entrelacement : l'entrelacement de Riesz. De nature légèrement différente de celle présentée dans cette introduction, elle fait intervenir les transformation de Riesz, qui sont des isométries de L^2. Ces dernières permettent d'estimer les valeurs propre d'un opérateur de diffusion $\mathbf{L}$ tel que défini ci-dessus par dérivations successives. On obtient ainsi le résultat suivant.

Théorème. *Supposons que $\nabla^2 V \geq \rho I_d$ pour un certain $\rho > 0$. Il existe une suite croissante d'entiers $(a_n)_{n\geq 0}$ telle que pour tous $k, n \in \mathbf{N}$ tels que $a_n < k \leq a_{n+1}$,*

$$\lambda_k(-\mathbf{L}) \geq (n+1)\rho.$$

Ce résultat n'est pas en l'état aussi précis que celui de Milman, mais l'approche utilisée est plus constructive, et laisse a priori de nombreux degrés d'améliorations.

La dimension des espaces propres associés est également discutée, en relation notamment avec les travaux récents de Bo'az Klartag [54] sur le sujet. On discute dans une dernière partie de l'optimalité de ce résultat, ainsi que d'éléments sur l'estimation de la suite d'entiers sus-mentionnée.

Aspects numériques de l'inégalité de Poincaré

Le dernier chapitre est basé sur un travail en cours de rédaction en collaboration avec Aldéric Joulin et Olivier Roustant. Commencé sous forme d'un projet étudiant donné à l'INSA Toulouse, il vise à développer des outils d'analyse de sensibilité hors du contexte classique d'indépendance. Ce domaine des statistique s'intéresse à l'influence des variables d'entrées sur la sortie d'un système, représenté le plus souvent par une fonction, ces dernières étant supposées (pour diverses raisons) aléatoires. Cette influence peut être déterminée par plusieurs indices de sensibilité, dont ceux de Sobol' et ceux basés sur la dérivée de la fonction du système (DGSM). L'intérêt de ces deux familles est qu'elles sont liées par la constante de Poincaré des différentes lois en jeu.

Le contexte classique d'étude de ces problèmes est le cas d'entrées mutuellement indépendantes. Il permet entre autres de se ramener à l'étude de lois (et donc d'inégalités de Poincaré) unidimensionnelles. Les calculs numériques des différents indices et des constantes de Poincaré sont donc possibles par différences finies. Ce cadre est néanmoins éloigné (la plupart du temps) de la réalité, et l'objectif de ce travail était de s'en affranchir. Pour permettre les approximations numériques des inégalités de Poincaré, l'hypothèse d'indépendance a été relâchée en supposant que les variables sont regroupées par paires de variables dépendantes, et que ces paires sont elles-mêmes mutuellement indépendantes. On se ramène ainsi au calcul de constantes de Poincaré bidimensionnelles. On discute dans ce contexte de la conjecture suivante, largement validée par l'expérience.

Conjecture. *Soit μ la loi d'un couple de variable aléatoires (X, Y) de covariance $\mathrm{Cov}(X, Y) = \sigma_{(X,Y)} \geq 0$. Alors $c_P(\mu)$ croît avec $\sigma_{(X,Y)}$.*

Notons que cette conjecture est un théorème dans le cas gaussien, comme il sera démontré grâce à la théorie de Barky-Émery.

Ce chapitre présente donc d'abord les bases d'analyse de sensibilité, puis la méthode développée et les outils utilisés (essentiellement la bibliothèque FEniCS pour Python et le mailleur éléments finis GMSH) pour identifier de telles constantes de Poincaré. Cette méthode est ensuite testée sur l'exemple d'un modèle de crue simplifié, déjà utilisé dans le cas de variables indépendantes. Les performances numériques sont également discutées, tant d'un point de vue du temps de calcul que dans la perspective d'un contexte de parcimonie. Les différents aspects de la méthode sont amplement illustrés, et je me ferai un plaisir de transmettre les codes sur simple demande, dans l'attente d'une mise à disposition publique en ligne.

Chapitre 2

Représentation des semi-groupes de Feynman-Kac, application aux inégalités de Sobolev logarithmiques

DANS CE DEUXIÈME CHAPITRE, nous nous intéressons à la version semi-groupe des relations d'entrelacement. Bien qu'équivalente à la version générateur, les deux s'utilisent parfois dans des cadres différents. Le but est ici de les appliquer à des semi-groupes bien choisis afin d'adapter des estimations similaires à celles du Lemme 1.2.7 ou à la preuve de la Proposition 1.2.17 dans des situations plus complexes.

Le point central dans ces preuves sera l'introduction d'une représentation probabiliste des semi-groupes de Feynman-Kac introduits en Définition 1.3.5, couplés avec une transformation de Doob bien choisie, qui fournissent de nouvelles estimations sur les gradients des semi-groupes. Ces résultats de portée assez générale sont exposés en première partie, après quelques rappels de calcul stochastique.

La seconde partie est consacrée à l'application de ces résultats aux inégalités de Sobolev logarithmiques, introduites dans la Définition 1.2.14. Ce chapitre est largement basé sur l'article [83].

Sauf mention contraire, les objets V, μ, $\mathbf{L}$, $(\mathbf{P}_t)_{t\geq 0}$ et $(X_t^x)_{t\geq 0}$ seront ceux introduits dans le chapitre précédent.

Table des matières

2.1 Représentation des semi-groupes de Feynman-Kac 38
2.1.1 Éléments de calcul stochastique 38
2.1.2 Heuristique scalaire 39
2.1.3 Régularité du processus et flot tangent 41
2.1.4 Approche perturbative 46
2.2 Inégalité de Sobolev logarithmique 49
2.2.1 Critère général 50
2.2.2 Fonctions monotones 53
2.2.3 Exemples 56
2.3 Perspectives 62

2.1 Représentation des semi-groupes de Feynman-Kac

Le but de l'article [83] était de fournir un nouveau critère pour garantir l'existence d'une inégalité de Sobolev logarithmique et estimer la constante optimale. L'idée générale était alors, en utilisant relations d'entrelacement et transformations de Doob, de démontrer dans un cadre plus général des estimations sur le gradient des semi-groupes, de même nature que celles du Lemme 1.2.7, permettant de généraliser le critère de courbure-dimension (Proposition 1.2.17). Deux éléments sont alors nécessaires à l'établissement de ces estimations : la représentation probabiliste des semi-groupes de Feynman-Kac, qui permet l'introduction d'une fonction de perturbation via le théorème de Girsanov, et une notion adéquate de courbure, qui généralise celle de Bakry-Émery.

Nous commencerons cette section par des rappels de calcul stochastique, suivis d'une heuristique unidimensionnelle, qui permettra d'établir un lien avec Bakry-Émery à moindre frais techniques, puis énoncerons et prouverons le théorème de représentation, avant d'établir les estimations de gradient attendues.

2.1.1 Éléments de calcul stochastique

Le théorème de représentation au cœur de ce chapitre est basé sur le théorème de Girsanov, qui sera rappelé en temps utile. On notera dans la suite $(B_t)_{t\geq 0}$ le mouvement brownien standard sur $\mathbf{R}^d$, dont on pourra trouver une description par exemple dans [73] (ainsi que de nombreuses informations complémentaires). Il permet de définir les équations différentielles stochastiques mentionnées dans l'introduction :

$$dY_t = f(Y_t)dt + g(Y_t)dB_t,$$

avec f et g des fonctions de classe $\mathcal{C}^1$. La notation différentielle ci-dessus ne traduit pas la dérivation, mais est un raccourci commode pour la formulation intégrale :

$$Y_t = Y_0 + \int_0^t f(Y_s)\, ds + \int_0^t g(Y_s)\, dB_s.$$

Le principal outil de calcul stochastique nécessaire pour la suite est la formule d'Itō [50], qui s'énonce dans ce contexte comme suit.

Proposition 2.1.1 (Itō pour les équations différentielles stochastiques)**.** *Soient $f \in \mathcal{C}^2(\mathbf{R}^d, \mathbf{R})$ et $(Y_t)_{t\geq 0}$ la solution de l'équation différentielle stochastique :*

$$dY_t = \sigma(Y_t)dB_t + b(Y_t)dt,$$

avec $\sigma \in \mathcal{C}^1(\mathbf{R}^d, \mathbf{R})$ et $b \in \mathcal{C}^1(\mathbf{R}^d, \mathbf{R}^d)$. Alors :

$$df(Y_t) = \sigma(Y_t)\nabla f(Y_t) \cdot dB_t + \left[\frac{1}{2}\sigma(Y_t)^2 \Delta f(Y_t) + \nabla f(Y_t) \cdot b(Y_t)\right] dt.$$

Il existe de nombreux variantes et sous-cas de l'équation ci-dessus. Le plus courant est de considérer une fonction σ à valeurs dans les matrices symétriques définies positives. Dans la suite, nous prendrons plutôt, comme décrit au début, $\sigma \equiv \sqrt{2}$ (pour éviter le $1/2$ dans la formule d'Itō) et $b = -\nabla V$. On retrouve alors l'équation (EDS) :

$$dX_t = \sqrt{2}dB_t - \nabla V(X_t),$$

et on obtient la formule synthétique suivante :

$$df(X_t) = \sqrt{2}\nabla f(X_t) \cdot dB_t + \mathbf{L}f(X_t)dt.$$

Cette dernière formule sera la plus utilisée dans la suite. Rappelons également que les hypothèses techniques considérées (plus particulièrement **(H4)**), assurent que le processus solution de (EDS) existe en tout temps.

2.1.2 Heuristique scalaire

Cette heuristique correspond en fait au contexte initial dans lequel Mark Kac a établi sa célèbre formule [53] (inspirée par l'étude des opérateurs de Schrödinger par Richard Feynman). Parfois énoncée dans le cadre des équations différentielles rétrogrades ("backward" en anglais, avec une flèche du temps inversée, cadre qui ne nous intéresse pas ici), elle s'énonce simplement pour des équations antérogrades ("forward" en anglais).

Proposition 2.1.2 (Formule de Feynman-Kac)**.** *On considère l'opérateur de Schrödinger $\tilde{\mathbf{L}} = \mathbf{L} - U$, où U est un potentiel multiplicatif scalaire assez régulier. On définit u comme la solution de $\partial_t u = \tilde{\mathbf{L}}u$, avec condition initiale $u(0, \cdot) = f$ continue bornée. Alors pour tous $t \geq 0$, $x \in \mathbf{R}^d$,*

$$u(t, x) = \mathbf{E}\left[f(X_t^x)e^{-\int_0^t U(X_s^x)ds}\right].$$

Démonstration. La preuve repose sur la formule d'Itō pour les diffusions, couplée à l'intégration par parties pour les processus et aux équations de Kolmogorov pour les semi-groupes. Plus précisément, si on note

$$A_t = \exp\left(-\int_0^t U(X_s^x)\, ds\right),$$

on cherche à montrer que $\tilde{\mathbf{P}}_t f := \mathbf{E}[f(X_t^x)A_t]$ satisfait $\partial_t \tilde{\mathbf{P}}_t f = \tilde{\mathbf{P}}_t \tilde{\mathbf{L}} f$. Cela se fait en utilisant la formule d'Itō.

En effet, on a d'une part $dA_t = -U(X_t^x)A_t dt$, et d'autre part la formule pour les diffusions :

$$df(X_t^x) = \sqrt{2}\nabla f(X_t^x) \cdot dB_t + \mathbf{L}f(X_t^x)dt.$$

Comme U est régulier, A_t est un processus à variations bornées, il s'ensuit :

$$\begin{aligned} d(f(X_t^x)A_t) &= df(X_t^x)A_t + f(X_t^x)dA_t \\ &= \sqrt{2}A_t \nabla f(X_t^x) \cdot dB_t + (\mathbf{L}f(X_t^x) - U(X_t^x)f(X_t^x))A_t dt \\ &= \sqrt{2}A_t \nabla f(X_t^x) \cdot dB_t + \tilde{\mathbf{L}}f(X_t^x)A_t dt, \end{aligned}$$

et en prenant l'espérance :

$$\tilde{\mathbf{P}}_t f = \mathbf{E}\left[\int_0^t \sqrt{2}A_s \nabla f(X_s^x) \cdot dB_s + \int_0^t \tilde{\mathbf{L}}f(X_s^x)A_s ds\right] = \mathbf{E}\left[\int_0^t \tilde{\mathbf{L}}f(X_s^x)A_s ds\right],$$

entendu que l'intégrale brownienne est nulle. Cela donne donc :

$$\partial_t \tilde{\mathbf{P}}_t f = \mathbf{E}[\tilde{\mathbf{L}}f(X_t^x)A_t] = \tilde{\mathbf{P}}_t \tilde{\mathbf{L}} f.$$

Ainsi, la solution de $\partial_t u = \tilde{\mathbf{L}}u$ est bien donnée, pour $t \geq 0$ et $x \in \mathbf{R}^d$, par

$$u(t,x) = \mathbf{E}\left[f(X_t^x)e^{-\int_0^t U(X_s^x)ds}\right].$$

■

Avec la notation des semi-groupes de Feynman-Kac introduite en Définition 1.3.5, cela s'écrit pour tout $t \geq 0$:

$$\mathbf{P}_t^U f = \mathbf{E}\left[f(X_t)e^{-\int_0^t U(X_s)ds}\right],$$

la notation $\mathcal{P}_t$ étant réservée aux semi-groupes à valeurs vectorielles.

Notons que ce semi-groupe admet une interprétation intéressante en termes de processus tués (ou absorbés), dans le cas où U est une fonction positive (présentée par exemple dans [18]). Considérons une variable aléatoire ε de loi exponentielle de paramètre 1 indépendante de $(X_t)_{t\geq 0}$ et définissons :

$$\xi = \inf\left\{t \geq 0 \,:\, \int_0^t U(X_s)\,ds > \varepsilon\right\}.$$

On peut définir $(\tilde{X}_t)_{t\geq 0}$ le processus tué au temps ξ, avec cimetière ∂, comme

$$\tilde{X}_t = X_t \mathbf{1}_{\{t\leq \xi\}} + \partial \mathbf{1}_{\{t>\xi\}}.$$

Alors pour toute fonction f régulière à support compact et $t \geq 0$, on a l'identité

$$\mathbf{P}_t^U f = \mathbf{E}[f(\tilde{X}_t)].$$

En effet, f étant à support compact, on a pour tout $t \geq 0$,

$$\mathbf{E}[f(\tilde{X}_t)] = \mathbf{E}[f(X_t)\mathbf{1}_{\{t \leq \xi\}}],$$

et $\{t \leq \xi\} = \left\{ \int_0^t U(X_s)\, ds \leq \varepsilon \right\}$, soit en conditionnant par $(\mathcal{F}_t)_{t \geq 0}$ la filtration naturelle de $(X_t)_{t \geq 0}$,

$$\begin{aligned}
\mathbf{E}[f(X_t)\mathbf{1}_{\{t \leq \xi\}}] &= \mathbf{E}\left[f(X_t)\mathbf{1}_{\left\{ \int_0^t U(X_s)\, ds \leq \varepsilon \right\}} \right] \\
&= \mathbf{E}\left[f(X_t)\mathbf{P}\left(\varepsilon > \int_0^t U(X_s)\, ds \,\middle|\, \mathcal{F}_t \right) \right] \\
&= \mathbf{E}\left[f(X_t) e^{-\int_0^t U(X_s) ds} \right] = \mathbf{P}_t^U f.
\end{aligned}$$

Une telle interprétation n'est malheureusement pas disponible dans le cadre des relations d'entrelacement introduites en Proposition 1.3.6. Elle subsiste néanmoins dans le cas de la dimension 1.

Dans ce cas particulier, les relations d'entrelacement sont explicites.

Proposition 2.1.3 (Entrelacement en dimension 1)**.** *Pour toute fonction $f \in \mathcal{C}_c^\infty(\mathbf{R}, \mathbf{R})$, pour tout $t \geq 0$,*

$$(\mathbf{P}_t f)' = \mathbf{P}_t^{V''}(f').$$

La preuve est immédiate en calculant la dérivée du générateur. On retrouve dans ce cas immédiatement les estimations du Lemme 1.2.7. En effet, si $V'' \geq \rho > 0$, on a $\int_0^t V''(X_s)\, ds \geq \rho t$, et donc

$$|(\mathbf{P}_t f)'| \leq e^{-\rho t} \mathbf{P}_t(|f'|).$$

L'étape intermédiaire de l'écriture explicite de $(\mathbf{P}_t f)'$ est néanmoins importante. Comme cela a été largement discuté dans [18], elle permet d'apporter des améliorations au critère de courbure-dimension en fournissant des estimations de gradient dans des cas où l'hypothèse de convexité est relâchée. L'enjeu de la représentation dans le cas général est alors double : comprendre la structure des semi-groupes de Feynman-Kac généraux et permettre une amélioration du critère de courbure-dimension, via une approche perturbative.

2.1.3 Régularité du processus et flot tangent

L'utilisation de la formule d'Itō dans la preuve du théorème précédent est un outil fort commode quand la forme de l'intégrande est connue, mais il existe un autre outil, plus spécifiquement adapté aux entrelacements avec le gradient : le flot tangent.

Définition 2.1.4. On appelle *flot tangent* à $\mathbf{L}$ (ou à $(\mathbf{P}_t)_{t \geq 0}$) le flot sur $\mathbf{R}_+$ de l'équation

$$\partial_t u = (\mathcal{L} - \nabla^2 V) u. \tag{2.1}$$

Notons que la solution de cette équation est à valeurs vectorielles. La terminologie de flot tangent vient de la relation d'entrelacement de la Proposition 1.3.6 :

$$\nabla \mathbf{L} = \mathcal{L}^{\nabla^2 V} \nabla.$$

Dans un cadre riemannien, cela correspond non plus à une action de $\mathbf{R}_+$ sur l'ensemble des trajectoires sur la variété, mais bien à une action sur les applications du fibré tangent, d'où l'appellation.

Comme nous l'avons vu ci-dessus, le flot tangent s'exprime facilement en dimension 1, dans la mesure où gradients et fonctions sont essentiellement de la même nature. Ce n'est plus le cas en dimension supérieure, où l'exponentielle de la formule de Feynman-Kac laisse place à un processus matriciel.

Théorème 2.1.5. *Il existe un processus $(J_t^X)_{t\geq 0}$ continu à valeurs matricielles tel que pour tous champ de vecteurs $F \in \mathcal{C}_c^\infty(\mathbf{R}^d, \mathbf{R}^d)$, $t \geq 0$,*

$$\mathcal{P}_t^{\nabla^2 V} F = \mathbf{E}[J_t^X F(X_t)].$$

Démonstration. On considère la condition initiale $x \in \mathbf{R}^d$ pour $(X_t^x)_{t\geq 0}$. L'idée de preuve est semblable à la précédente. Il faut néanmoins identifier ici le processus en jeu. Une approche naïve consiste à dériver l'équation différentielle stochastique satisfaite par $(X_t^x)_{t\geq 0}$, en écrivant

$$X_t^x = x + \sqrt{2} B_t - \int_0^t \nabla V(X_s^x)\, ds.$$

On a donc, pour tout $t \geq 0$,

$$\mathrm{Jac}_x X_t^x = I_d - \int_0^t \mathrm{Jac}_x X_t^x \nabla^2 V(X_t^x)\, ds.$$

On définit alors $(J_t^X)_{t\geq 0}$ comme la solution de l'équation différentielle ordinaire :

$$dJ_t^X = -J_t^X \nabla^2 V(X_t) dt, \tag{2.2}$$

avec la condition initiale $J_0^X = I_d$ (indépendante de $(X_t^x)_{t\geq 0}$). On pose alors, pour $t \geq 0$,

$$\tilde{\mathcal{P}}_t F = \mathbf{E}[J_t^X F(X_t)],$$

et le but est de montrer que $\partial_t \tilde{\mathcal{P}}_t F = \tilde{\mathcal{P}}_t \mathcal{L}^{\nabla^2 V} F$. Pour cela, dans le même esprit que la preuve précédente, on se ramène à l'utilisation de la formule d'Itō. Elle s'écrit pour le champ de vecteurs F :

$$dF(X_t) = \sqrt{2}\mathrm{Jac}_{X_t} F dB_t + \mathcal{L} F(X_t)\, dt,$$

où, pour rappel, $\mathcal{L}$ désigne le générateur diagonal associé à $(X_t)_{t\geq 0}$. De plus, le processus $(J_t^X)_{t\geq 0}$ ainsi défini possède des trajectoires continues à variations bornées (d'après

les hypothèses sur V, voir par exemple [73] V.7). La formule d'Itō donne alors

$$\begin{aligned} d(J_t^X F(X_t)) &= dJ_t^X F(X_t) + J_t^X dF(X_t) \\ &= -J_t^X \nabla^2 V(X_t) F(X_t) dt + \sqrt{2} J_t^X \mathrm{Jac}_{X_t} F dB_t + J_t^X \mathcal{L} F(X_t)\, dt \\ &= \sqrt{2} J_t^X \mathrm{Jac}_{X_t} F dB_t + J_t^X \mathcal{L}^{\nabla^2 V} F(X_t)\, dt. \end{aligned}$$

L'espérance de l'intégrale brownienne ci-dessus est nulle, et on a donc

$$\partial_t \tilde{\mathcal{P}}_t F = \mathbf{E}[J_t^X \mathcal{L}^{\nabla^2 V} F(X_t)] = \tilde{\mathcal{P}}_t(\mathcal{L}^{\nabla^2 V} F),$$

Par unicité, il s'ensuit :

$$\mathcal{P}_t^{\nabla^2 V} F = \tilde{\mathcal{P}}_t F = \mathbf{E}[J_t^X F(X_t)].$$

■

Plusieurs remarques s'imposent à la suite de cette preuve. La première concerne l'unicité de la solution du flot tangent. Celle-ci est relativement facile à établir : considérons deux solutions u_1 et u_2 de (2.1) issues de la même condition initiale. Si on note, $t \geq 0$,

$$g(t) = \|u_1(t,\cdot) - u_2(t,\cdot)\|_2^2,$$

alors

$$\begin{aligned} g'(t) &= 2\langle \mathcal{L}^{\nabla^2 V}(u_1(t,\cdot) - u_2(t,\cdot)), u_1(t,\cdot) - u_2(t,\cdot)\rangle \\ &= 2\langle \mathcal{L}(u_1(t,\cdot) - u_2(t,\cdot)), u_1(t,\cdot) - u_2(t,\cdot)\rangle \\ &\quad - 2\langle \nabla^2 V(u_1(t,\cdot) - u_2(t,\cdot)), u_1(t,\cdot) - u_2(t,\cdot)\rangle. \end{aligned}$$

L'opérateur $\mathcal{L}$ est, à l'instar de $\mathbf{L}$, semi-défini négatif, donc le premier produit scalaire ci-dessus est négatif. De même, par hypothèse il existe $\rho \in \mathbf{R}$ (pas nécessairement positif) tel que pour tout $x \in \mathbf{R}^d$, $\nabla^2 V(x) \geq \rho I_d$, on a donc :

$$g'(t) \leq 0 - 2\rho\langle u_1(t,\cdot) - u_2(t,\cdot), u_1(t,\cdot) - u_2(t,\cdot)\rangle = -2\rho g(t).$$

Le lemme de Grönwall assure alors que pour tout $t \geq 0$,

$$g(t) \leq g(0) e^{-2\rho t} = 0,$$

car $g(0) = \|u_1(0,\cdot) - u_2(0,\cdot)\|_2^2 = 0$ par hypothèse sur la condition initiale. Comme on le voit, l'hypothèse sur V assure non seulement la non-explosion du processus $(X_t)_{t\geq 0}$, mais ici également l'unicité de la solution à l'équation du flot tangent. C'est un point technique vite établi, mais qui a le mérite d'apporter une certaine sérénité méthodologique.

La deuxième remarque est naturellement la comparaison avec le cas unidimensionnel. Ici, le processus $(J_t^X)_{t\geq 0}$ joue le rôle de l'exponentielle de la Proposition 2.1.2, l'équation satisfaite par $(A_t)_{t\geq 0}$ est de même nature que (2.2). Malheureusement, l'écriture

exponentielle n'est que rarement accessible dans le cas général : elle ne l'est que si pour tous $x, y \in \mathbf{R}^d$, les matrices $\nabla^2 V(x)$ et $\nabla^2 V(y)$ commutent.

Deux cas particuliers qui satisfont cette condition méritent d'être mentionnés : le cas où $\nabla^2 V$ est constante (le potentiel est polynomial de degré au plus deux) et le cas des mesures produit. En effet, si $\mu = \mu_1 \otimes \cdots \otimes \mu_d$, il existe $V_1, \ldots, V_d$ d potentiels tels que

$$V(x) = \sum_{i=1}^{d} V_i(x_i), \quad x \in \mathbf{R}^d,$$

et la hessienne de V est alors diagonale : pour $x \in \mathbf{R}^d$,

$$\nabla^2 V(x) = \mathrm{diag}(V_1''(x_1), \ldots, V_d''(x_d)).$$

Dans ce cas, le flot tangent se réécrit comme une collection de formules de Feynman-Kac unidimensionnelles : pour $i \in [\![1, d]\!]$,

$$(\mathcal{P}_t^{\nabla^2 V} F)_i = \left[e^{-\int_0^t V_i''(X_s^i) ds} F_i(X_t) \right].$$

C'est en particulier de ce cas que relève la gaussienne. Cette écriture permet de retrouver l'entrelacement gaussien : pour toute fonction f assez régulière,

$$\mathcal{P}_t^{\nabla^2 V}(\nabla f) = \mathcal{P}_t^{I_d}(\nabla f) = \mathbf{E}\left[e^{-tI_d} \nabla f(X_t) \right] = e^{-t} \mathcal{P}_t(\nabla f).$$

Mais il est possible d'aller plus loin, en retrouvant à partir de la condition $CD(\rho, \infty)$ les estimations du Lemme 1.2.7.

Proposition 2.1.6 (Critère de courbure-dimension III)**.** *Supposons que $\nabla^2 V \geq \rho I_d$, pour $\rho > 0$. Alors pour toute fonction f assez régulière, pour tout $t \geq 0$,*

$$|\nabla \mathbf{P}_t f| \leq e^{-\rho t} \mathbf{P}_t(|\nabla f|).$$

La preuve a le bon goût, une fois les objets proprement introduits, de présenter une technicité relativement faible. Elle introduit néanmoins un concept fondamental pour la suite : les estimations spectrales du processus tangent.

Lemme 2.1.7. *Supposons que $\nabla^2 V \geq \rho I_d$, pour $\rho > 0$. Alors pour tout $t \geq 0$,*

$$\|J_t^X\|_2^2 \leq e^{-2\rho t}.$$

Démonstration. Quitte à transposer, on peut s'intéresser à $(J_t^X)^T$ (cela ne change pas les valeurs propres et facilite les calculs). Il s'agit donc de montrer que, pour tous $y \in \mathbf{R}^d$, $t \geq 0$,

$$y^T J_t^X (J_t^X)^T y \leq e^{-2\rho t} |y|^2.$$

Une fois encore, hors de Grönwall point de salut. On pose donc :

$$g(t) = y^T J_t^X (J_t^X)^T y.$$

Alors g est dérivable et on a :

$$g'(t) = -2y^T J_t^X \nabla^2 V(X_t)(J_t^X)^T y.$$

L'hypothèse $\nabla^2 V \geq \rho I_d$ donne alors :

$$g'(t) \leq -2\rho y^T J_t^X (J_t^X)^T y = -2\rho g(t),$$

et du lemme de Grönwall il suit, pour tout $t \geq 0$,

$$g(t) \leq e^{-2\rho t} g(0) = e^{-2\rho t}|y|^2,$$

ce qu'il fallait démontrer. ■

Ce lemme est la pierre angulaire de la démonstration de la proposition précédente, et l'esprit de sa preuve sera central dans l'établissement des nouvelles estimations du gradient des semi-groupes.

Proposition 2.1.6. On commence par écrire la forme explicite du gradient du semi-groupe : pour $t \geq 0$,

$$\nabla \mathbf{P}_t f = \mathcal{P}_t^{\nabla^2 V}(\nabla f) = \mathbf{E}[J_t^X \nabla f(X_t)].$$

On a, par l'inégalité de Jensen,

$$|\nabla \mathbf{P}_t f| \leq \mathbf{E}[|J_t^X \nabla f(X_t)|] = \mathbf{E}\left[\left(\nabla f(X_t)^T (J_t^X)^T J_t^X \nabla f(X_t)\right)^{1/2}\right].$$

Les estimations du Lemme 2.1.7 donnent alors :

$$|\nabla \mathbf{P}_t f| \leq e^{-\rho t}\mathbf{E}[|\nabla f(X_t)|] = e^{-\rho t}\mathbf{P}_t(|\nabla f|),$$

qui est l'inégalité attendue. ■

Notons que l'hypothèse de positivité sur ρ n'est pas utilisée dans les preuves ci-dessus. On peut effectivement s'en affranchir par ce biais-là, mais ces résultats sont en pratique utilisés pour décrire la décroissance du gradient du semi-groupe. À cette fin, l'hypothèse de positivité sur ρ n'est plus superflue.

Cette proposition n'est malheureusement pas applicable en l'état si l'on relâche l'hypothèse de convexité sur V. Un moyen d'y pallier, celui développé dans [83], consiste à perturber la matrice J_t^X de sorte à compenser localement le manque de convexité du potentiel. Cette méthode est détaillée dans la section suivante, avec quelques points de calcul stochastique nécessaires.

2.1.4 Approche perturbative

L'idée de cette dernière section est d'introduire la fonction de perturbation qui permettra de généraliser les estimations spectrales de $J_t^X(J_t^X)^T$. L'intérêt de la représentation des semi-groupes de Feynman-Kac est de rendre possible, via le théorème de Girsanov et les h-transformations, l'introduction de nombreux degrés de libertés dans l'amélioration des inégalités.

D'un point de vue probabiliste, h-transformation et théorème de Girsanov sont deux faces d'une même pièce. On parlera d'ailleurs volontiers de "perturbation" dans ces deux cas. Nous commençons donc par introduire un processus perturbé et les objets stochastiques afférents.

Définition 2.1.8. Soit $a \in \mathcal{C}_+^\infty(\mathbf{R}^d, \mathbf{R})$. On définit $(X_{t,a})_{t\geq 0}$ comme la solution de l'équation différentielle stochastique :

$$dX_{t,a} = \sqrt{2}dB_t - \nabla V_a(X_{t,a})\, dt,$$

où $V_a = V + \log(a^2)$.

Ce choix de processus et de potentiel n'est pas arbitraire. Si l'on se souvient de la Définition 1.3.8, la mesure invariante du générateur transformé $\mathbf{L}_h$ est explicite : si h est positive et μ-intégrable, alors la mesure $\mu_h = h\mu$ est invariante : pour f suffisamment régulière,

$$\int_{\mathbf{R}^d} \mathbf{L}_h f d\mu_h = \int_{\mathbf{R}^d} \frac{1}{h}\mathbf{L}(hf)h\, d\mu = \int_{\mathbf{R}^d} \mathbf{L}(hf)\, d\mu = 0.$$

Le potentiel associé à cette mesure est $V - \log(h)$, le choix fait dans la définition précédente correspond donc à $h = 1/a^2$, choix qui sera justifié par les calculs ultérieurs.

On notera dans ce contexte μ_a la mesure et $\mathbf{L}_a$ le générateur associés à V_a. Le calcul direct donne donc :

$$\mathbf{L}_a = \mathbf{L} - 2\frac{\nabla a}{a}\cdot\nabla.$$

De plus, la plupart des propriétés sympathiques de $\mathbf{L}$ se transmettent à $\mathbf{L}_a$.

Lemme 2.1.9. *La mesure μ_a est invariante pour $\mathbf{L}_a$. De plus, $\mathbf{L}_a$ est symétrique sur $\mathcal{C}_c^\infty(\mathbf{R}^d, \mathbf{R})$ et semi-défini négatif.*

Démonstration. La preuve est directe. Pour l'invariance, considérons $f \in \mathcal{C}_c^\infty(\mathbf{R}^d, \mathbf{R})$. On a :

$$\int_{\mathbf{R}^d} \mathbf{L}_a f\, d\mu_a = \int_{\mathbf{R}^d} \left(\mathbf{L}f - 2\frac{\nabla a}{a}\cdot\nabla f\right)\frac{1}{a^2}d\mu = \int_{\mathbf{R}^d} \mathbf{L}f\frac{1}{a^2}d\mu - 2\int_{\mathbf{R}^d}\frac{\nabla a}{a}\cdot\nabla f d\mu_a.$$

La formule d'intégration par parties classique donne pour le premier terme :

$$\int_{\mathbf{R}^d} \mathbf{L}f\frac{1}{a^2}d\mu = -\int_{\mathbf{R}^d}\nabla f\cdot\nabla\left(\frac{1}{a^2}\right)d\mu = 2\int_{\mathbf{R}^d}\nabla f\cdot\frac{\nabla a}{a^3}d\mu = 2\int_{\mathbf{R}^d}\nabla f\cdot\frac{\nabla a}{a}d\mu_a,$$

et donc

$$\int_{\mathbf{R}^d} \mathbf{L}_a f \, d\mu_a = 0.$$

De même, si $f, g \in \mathcal{C}_c^\infty(\mathbf{R}^d, \mathbf{R})$,

$$\int_{\mathbf{R}^d} f \mathbf{L}_a g \, d\mu_a = \int_{\mathbf{R}^d} f \mathbf{L} g \, d\mu_a - 2 \int_{\mathbf{R}^d} f \frac{\nabla a}{a} \cdot \nabla g \, d\mu_a.$$

la formule d'intégration par parties donne alors, pour le premier terme :

$$\begin{aligned}\int_{\mathbf{R}^d} f \mathbf{L} g \, d\mu_a &= \int_{\mathbf{R}^d} f \mathbf{L} g \frac{1}{a^2} d\mu = - \int_{\mathbf{R}^d} \nabla \left(f / a^2 \right) \cdot \nabla g \, d\mu \\ &= - \int_{\mathbf{R}^d} \nabla f \cdot \nabla g \, d\mu_a + 2 \int_{\mathbf{R}^d} f \frac{\nabla a}{a} \cdot \nabla g \, d\mu_a.\end{aligned}$$

Soit après simplification :

$$\int_{\mathbf{R}^d} f \mathbf{L}_a g \, d\mu_a = - \int_{\mathbf{R}^d} \nabla f \cdot \nabla g \, d\mu_a.$$

Il suit alors que $\mathbf{L}_a$ est symétrique et, prenant $f = g$, semi-défini négatif. ■

La preuve ci-dessus fait apparaître une forme de Dirichlet $\mathcal{E}_{\mu_a}$ tout à fait analogue à celle du cas classique. On procédera donc, comme précédemment, à la complétion de $\mathcal{C}_c^\infty(\mathbf{R}^d, \mathbf{R})$ en deux domaines $\mathcal{D}(\mathbf{L}_a)$ et $\mathcal{D}(\mathcal{E}_{\mu_a})$. Notons néanmoins que la mesure μ_a n'est pas supposée, a priori, finie. Pour s'assurer néanmoins de l'existence du processus $(X_{t,a})_{t \geq 0}$, on considérera que les hypothèse **(H2)** à **(H5)** mentionnées en introduction sont satisfaites.

Ainsi, on notera $(\mathbf{P}_{t,a})_{t \geq 0}$ le semi-groupe associé à $\mathbf{L}_a$, également ergodique dans $L^2(\mu_a)$ et invariant pour cette mesure. Ce semi-groupe admet également une représentation probabiliste, à l'instar de $(\mathbf{P}_t)_{t \geq 0}$: pour f assez régulière et $t \geq 0$,

$$\mathbf{P}_{t,a} f = \mathbf{E}[f(X_{t,a})].$$

Le lien entre $(\mathbf{P}_t)_{t \geq 0}$ et $(\mathbf{P}_{t,a})_{t \geq 0}$ est au cœur du théorème de représentation énoncé dans [83]. L'introduction des processus permet de travailler directement sur ces derniers, et un résultat important du calcul stochastique porte justement sur les changements de processus : le théorème de Girsanov.

Proposition 2.1.10 (Girsanov pour le mouvement brownien)**.** *Soient $(H_t)_{t \geq 0}$ un processus $(\mathcal{F}_t)_{t \geq 0}$-adapté à valeurs vectorielles et $(Y_t)_{t \geq 0}$ défini par :*

$$Y_t = B_t - \int_0^t H_s ds.$$

Si on définit la mesure de probabilité $\mathbf{Q}$ par :

$$\left. \frac{d\mathbf{Q}}{d\mathbf{P}} \right|_{\mathcal{F}_t} = \exp \left(\int_0^t H_s \cdot dB_s - \frac{1}{2} \int_0^t |H_s|^2 ds \right),$$

pour tout $t \geq 0$, alors l'exponentielle ci-dessus est une $(\mathcal{F}_t)_{t \geq 0}$-martingale et $(Y_t)_{t \geq 0}$ est un mouvement brownien sous la loi $\mathbf{Q}$.

La preuve de ce résultat peut être trouvée dans [73], chapitre IV. L'hypothèse de bornitude sur $(H_t)_{t\geq 0}$ peut être assouplie avec les critères de Novikov ou Kamazaki (voir également [73]), mais sera conservée dans la suite.

Le passage du processus $(X_t)_{t\geq 0}$ à $(X_{t,a})_{t\geq 0}$ peut alors s'effectuer par l'entremise du théorème de Girsanov.

Lemme 2.1.11. *Soit $a \in \mathcal{C}^\infty_+(\mathbf{R}^d, \mathbf{R})$ telle que $|\nabla a|/a$ est bornée. Alors pour toute fonction $f \in \mathcal{C}^\infty_c(\mathbf{R}^d, \mathbf{R})$, tous $t \geq 0$, $x \in \mathbf{R}^d$:*

$$\mathbf{P}_t f(x) = \mathbf{E}[f(X_t^x)] = \mathbf{E}\left[R^x_{t,a} f(X^x_{t,a})\right],$$

où $(R^x_{t,a})_{t\geq 0}$ est une $(\mathcal{F}_t)_{t\geq 0}$-martingale définie par :

$$R^x_{t,a} = \frac{a(X^x_{t,a})}{a(x)} \exp\left(-\int_0^t \frac{\mathbf{L}_a a}{a}(X^x_{s,a})\, ds\right), \quad t \geq 0.$$

Démonstration. La preuve commence par la définition d'une martingale exponentielle ad hoc, qui servira ensuite dans l'application du théorème de Girsanov. Pour alléger les notations, la condition initiale sera omise par la suite.

On commence par appliquer la formule d'Itō à $\log(a(X_{t,a}))$:

$$\log a(X_{t,a}) = \log a + \sqrt{2}\int_0^t \nabla(\log a(X_{s,a})) \cdot dB_s + \int_0^t \mathbf{L}_a(\log a)(X_{s,a})\, ds.$$

Le calcul du terme de droite donne, en passant à l'exponentielle :

$$\begin{aligned}\frac{a(X_{t,a})}{a} &\exp\left(-\int_0^t \frac{\mathbf{L}_a a}{a}(X_{s,a})\, ds\right) \\ &= \exp\left(\sqrt{2}\int_0^t \frac{\nabla a}{a}(X_{s,a}) \cdot dB_s - \int_0^t \left|\frac{\nabla a}{a}\right|^2 (X_{s,a})\, ds\right).\end{aligned}$$

Le terme de gauche est $R_{t,a}$ défini précédemment, et on pose :

$$\left.\frac{d\mathbf{Q}_a}{d\mathbf{P}}\right|_{\mathcal{F}_t} = R_{t,a}.$$

D'après la proposition précédente, $(R_{t,a})_{t\geq 0}$ est donc une $(\mathcal{F}_t)_{t\geq 0}$-martingale et $(\tilde{B}_t)_{t\geq 0}$ défini par :

$$\tilde{B}_t = B_t - \sqrt{2}\int_0^t \frac{\nabla a}{a}(X_{s,a})\, ds,$$

est un mouvement brownien sous la loi $\mathbf{Q}_a$. De plus, le processus $(X_{t,a})_{t\geq 0}$ est solution de l'équation :

$$dX_{t,a} = \sqrt{2} d\tilde{B}_t - \nabla V(X_{t,a})\, dt,$$

donc la loi de $X_{t,a}$ sous $\mathbf{Q}_a$ coïncide avec celle de X_t sous $\mathbf{P}$. En particulier, pour toute fonction $f \in \mathcal{C}^\infty_c(\mathbf{R}^d, \mathbf{R})$,

$$\mathbf{P}_t f = \mathbf{E}[f(X_t)] = \mathbf{E}[R_{t,a} f(X_{t,a})],$$

ce qui termine la preuve. ∎

Notons que, tel le théorème de Girsanov, ce résultat peut s'appliquer à des fonctionnelles de la trajectoire. Plus précisément, pour $t \geq 0$, si $F : \mathcal{C}^0([0,t], \mathbf{R}) \to \mathbf{R}$ est une fonctionnelle assez régulière, alors :

$$\mathbf{E}[F(X_{[0,t]})] = \mathbf{E}[R_{t,a} F(X_{[0,t],a})],$$

où $X_{[0,t]} = (X_s, 0 \leq s \leq t)$ et $X_{[0,t],a} = (X_{s,a}, 0 \leq s \leq t)$. Cette généralisation reste valable si F est à valeurs vectorielles, et sera largement utilisée par la suite.

Ce lemme permet d'exprimer les semi-groupes de Feynman-Kac d'intérêt en fonction du processus perturbé, qui introduira un degré de liberté supplémentaire dans les estimations nécessaires à l'établissement de l'inégalité de Sobolev logarithmique.

Théorème 2.1.12. *Soient $f \in \mathcal{C}_c^\infty(\mathbf{R}^d, \mathbf{R})$ et $a \in \mathcal{C}_+^\infty(\mathbf{R}^d, \mathbf{R})$ telle que $|\nabla a| / a$ est bornée. Alors pour tous $t \geq 0$, $x \in \mathbf{R}^d$,*

$$\mathcal{P}_t^{\nabla^2 V}(\nabla f)(x) = \mathbf{E}[R_{t,a}^x J_t^{X_a^x} \nabla f(X_{t,a}^x)],$$

où $(R_{t,a}^x)_{t\geq 0}$ est la $(\mathcal{F}_t)_{t\geq 0}$-martingale définie précédemment et $(J_t^{X_a^x})_{t\geq 0}$ est solution de l'équation :

$$\begin{cases} dJ_t^{X_a^x} &= -J_t^{X_a^x} \nabla^2 V(X_{t,a}^x) dt, \quad t > 0 \\ J_0^{X_a^x} &= I_d. \end{cases}$$

Démonstration. La preuve est presque immédiate. Dans un souci de lisibilité, on ne fera pas apparaître la condition initiale. On a d'abord, par le Théorème 2.1.5,

$$\mathcal{P}_t^{\nabla^2 V}(\nabla f) = \mathbf{E}[J_t^X \nabla f(X_t)].$$

De plus, quitte à remplacer $(X_t)_{t\geq 0}$ par $(X_{t,a})_{t\geq 0}$, on définit :

$$J_t^{X_a} = I_d - \int_0^t J_s^{X_a} \nabla^2 V(X_{s,a})\, ds.$$

Notons que la modification ne porte que sur le processus, le potentiel en jeu ne change pas. On a alors, par le Lemme 2.1.11,

$$\mathbf{E}[J_t^X \nabla f(X_t)] = \mathbf{E}[R_{t,a} J_t^{X_a} \nabla f(X_{t,a})],$$

ce qui est le résultat annoncé. ■

La section suivante est consacrée à l'application des principes présentés ci-dessus à l'établissement d'un critère pour l'inégalité de Sobolev logarithmique, telle que présentée dans [83].

2.2 Inégalité de Sobolev logarithmique

Cette section reprend le plan de la deuxième partie de [83] : le critère pour l'inégalité de Sobolev logarithmique est d'abord énoncé et prouvé, assorti d'une brève discussion sur les améliorations possibles et la comparaison avec le critère de Holley-Stroock ; suivent une restriction de ce critère aux fonctions monotones (inspirée par [62]) et l'application à deux exemples.

2.2.1 Critère général

Ce critère se base d'une part sur la bornitude de la fonction de perturbation, et d'autre part sur une courbure perturbée, qui généralise la courbure de Bakry-Émery, induite par la hessienne du potentiel.

L'idée de ce résultat est inspiré par un travail récent de Feng-Yu Wang sur les variétés riemanniennes à bord [87].

Théorème 2.2.1. *Soit $a \in \mathcal{C}^\infty_+(\mathbf{R}^d, \mathbf{R})$. On définit la* courbure perturbée *:*

$$\kappa_a = \inf_{x \in \mathbf{R}^d} \left\{ 2\rho_-(\nabla^2 V(x)) - a\mathbf{L}(a^{-1})(x) \right\}.$$

Si a, a^{-1} et $|\nabla a|$ sont bornés et $\kappa_a > 0$, alors pour toute fonction $f \in \mathcal{C}^\infty_c(\mathbf{R}^d, \mathbf{R})$,

$$\mathrm{Ent}_\mu(f^2) \leq \frac{4\|a\|_\infty \|a^{-1}\|_\infty}{\kappa_a} \int_{\mathbf{R}^d} |\nabla f|^2 \, d\mu.$$

La terminologie de "courbure perturbée" est héritée du contexte riemannien dans lequel a été énoncé le critère de Bakry-Émery pour la première fois (en lien avec la courbure de Ricci). Il s'agit ici d'une borne sur la courbure qui peut, dans le cas de mesures sur des variétés différentes de $\mathbf{R}^d$, s'exprimer en fonction du tenseur de Ricci (voir par exemple [10] ou [87]).

Démonstration. Le déroulé de la preuve est très similaire à celui de la Proposition 1.2.17. On considère dans un premier temps $f \in \mathcal{C}^\infty_c(\mathbf{R}^d, \mathbf{R})$ une fonction positive. L'ergodicité et l'invariance pour μ de $(\mathbf{P}_t)_{t \geq 0}$ donnent :

$$\mathrm{Ent}_\mu(f) = -\int_{\mathbf{R}^d} \int_0^{+\infty} \partial_t \left(\mathbf{P}_t f \log \mathbf{P}_t f \right) \, dt \, d\mu = -\int_{\mathbf{R}^d} \int_0^{+\infty} \mathbf{L}[\mathbf{P}_t f] \log \mathbf{P}_t f \, dt \, d\mu.$$

Par intégration par parties suivie de la relation d'entrelacement classique, on a :

$$\mathrm{Ent}_\mu(f) = \int_{\mathbf{R}^d} \int_0^{+\infty} \frac{|\nabla \mathbf{P}_t f|^2}{\mathbf{P}_t f} \, d\mu \, dt = \int_{\mathbf{R}^d} \int_0^{+\infty} \frac{\left| \mathcal{P}_t^{\nabla^2 V}(\nabla f) \right|^2}{\mathbf{P}_t f} \, dt \, d\mu.$$

Le travail d'estimation se concentre sur le numérateur du terme ci-dessus. Plus précisément, afin de simplifier le terme $\mathbf{P}_t f$ au dénominateur, on se repose sur des estimations spectrales rendues possibles par l'utilisation du théorème de Girsanov. En effet, les hypothèses sur a assurent que le Théorème 2.1.12 s'applique, et on a :

$$\mathcal{P}_t^{\nabla^2 V}(\nabla f) = \mathbf{E}[R_{t,a} J_t^{X_a} \nabla f(X_{t,a})],$$

ce qui se réécrit

$$\mathcal{P}_t^{\nabla^2 V}(\nabla f) = 2\mathbf{E}\left[R_{t,a}^{1/2} J_t^{X_a} \nabla \sqrt{f}(X_{t,a}) R_{t,a}^{1/2} \sqrt{f}(X_{t,a}) \right].$$

L'inégalité de Cauchy-Schwarz donne donc, avec le Lemme 2.1.11,

$$\begin{aligned}\left|\mathcal{P}_t^{\nabla^2 V}(\nabla f)\right|^2 &\leq 4\mathbf{E}\left[\left|R_{t,a}^{1/2} J_t^{X_a} \nabla\sqrt{f}(X_{t,a})\right|^2\right] \mathbf{E}\left[R_{t,a} f(X_{t,a})\right] \\ &= 4\mathbf{E}\left[\nabla\sqrt{f}(X_{t,a})^T J_t^{X_a} R_{t,a} (J_t^{X_a})^T \nabla\sqrt{f}(X_{t,a})\right] \mathbf{P}_t f.\end{aligned}$$

Cela implique, dans l'entropie :

$$\mathrm{Ent}_\mu(f) \leq 4\int_{\mathbf{R}^d}\int_0^{+\infty} \mathbf{E}\left[\nabla\sqrt{f}(X_{t,a})^T (J_t^{X_a})^T R_{t,a} J_t^{X_a} \nabla\sqrt{f}(X_{t,a})\right]\, dt\, d\mu.$$

Pour retrouver le terme d'énergie de l'inégalité, on reprend l'idée de la preuve de la Proposition 1.2.17, en s'intéressant à l'analyse spectrale de $J_t^{X_a} R_{t,a} (J_t^{X_a})^T$. On définit donc :

$$J_t^a = J_t^{X_a} \exp\left(-\frac{1}{2}\int_0^t \frac{\mathbf{L}_a a}{a}(X_{s,a})\, ds\right),$$

qui est solution de :

$$dJ_t^a = -J_t^a\left(\nabla^2 V(X_{t,a}) - \frac{1}{2} a\mathbf{L}(a^{-1})(X_{t,a}) I_d\right) dt.$$

En effet, on a d'une part :

$$dJ_t^{X_a} = -J_t^{X_a} \nabla^2 V(X_{t,a})\, dt,$$

et d'autre part :

$$d\left[\exp\left(-\frac{1}{2}\int_0^t \frac{\mathbf{L}_a a}{a}(X_{s,a})\, ds\right)\right] = -\frac{1}{2}\frac{\mathbf{L}_a a}{a}(X_{t,a}) \exp\left(-\frac{1}{2}\int_0^t \frac{\mathbf{L}_a a}{a}(X_{s,a})\, ds\right) dt.$$

De plus, $\mathbf{L}_a(a)/a = -a\mathbf{L}(a^{-1})$, ce qui donne la formule attendue. Comme

$$J_t^{X_a} R_{t,a} (J_t^{X_a})^T = \frac{a(X_{t,a})}{a(x)} J_t^a (J_t^a)^T,$$

on s'intéresse plus particulière aux estimations spectrales pour $J_t^a (J_t^a)^T$.

Ainsi, on pose $\varphi(t) = y^T J_t^a (J_t^a)^T y$, pour $y \in \mathbf{R}^d$. Par symétrie de $\nabla^2 V$, on a :

$$\begin{aligned} d\varphi(t) &= y^T dJ_t^a (J_t^a)^T y + y^T J_t^a (dJ_t^a)^T y \\ &= -y^T J_t^a\left(\nabla^2 V(X_{t,a}) - \frac{1}{2} a\mathbf{L}(a^{-1})(X_{t,a}) I_d\right) (J_t^a)^T y\, dt \\ &\quad - y^T J_t^a\left(\nabla^2 V(X_{t,a}) - \frac{1}{2} a\mathbf{L}(a^{-1})(X_{t,a}) I_d\right)^T (J_t^a)^T y\, dt \\ &= -y^T J_t^a\left(2\nabla^2 V(X_{t,a}) - a\mathbf{L}(a^{-1})(X_{t,a}) I_d\right) (J_t^a)^T y\, dt \\ &\leq -\kappa_a y^T J_t^a (J_t^a)^T y\, dt = -\kappa_a \varphi(t)\, dt, \end{aligned}$$

par définition de κ_a. Ainsi, $\varphi(t) \leq e^{-\kappa_a t}\varphi(0)$ pour tout $t \geq 0$, et donc :

$$y^T J_t^a (J_t^a)^T y \leq e^{-\kappa_a t}|y|^2.$$

On applique alors l'inégalité ci-dessus à $y = \sqrt{\frac{a(X_{t,a})}{a(x)}}\nabla\sqrt{f}(X_{t,a})$ pour obtenir :

$$\mathrm{Ent}_\mu(f) \leq 4\int_{\mathbf{R}^d}\int_0^{+\infty} e^{-\kappa_a t}\mathbf{E}\left[\frac{a(X_{t,a})}{a(x)}|\nabla\sqrt{f}(X_{t,a})|^2\right]\, dt\, d\mu,$$

qui se réécrit :

$$\mathrm{Ent}_\mu(f) \leq 4\int_0^{+\infty} e^{-\kappa_a t}\int_{\mathbf{R}^d}\frac{1}{a}\mathbf{P}_{t,a}\left(a|\nabla\sqrt{f}|^2\right)\, d\mu\, dt.$$

Sachant que $d\mu_a/d\mu = 1/a^2$, la bornitude de a entraîne :

$$\mathrm{Ent}_\mu(f) \leq 4\|a\|_\infty\int_0^{+\infty} e^{-\kappa_a t}\int_{\mathbf{R}^d}\mathbf{P}_{t,a}\left(a|\nabla\sqrt{f}|^2\right)\, d\mu_a\, dt.$$

L'invariance de $\mathbf{P}_{t,a}$ pour μ_a et l'hypothèse sur κ_a donnent :

$$\mathrm{Ent}_\mu(f) \leq \frac{4\|a\|_\infty}{\kappa_a}\int_{\mathbf{R}^d} a|\nabla\sqrt{f}|^2\, d\mu_a.$$

On utilise finalement la bornitude de a^{-1} :

$$\mathrm{Ent}_\mu(f) \leq \frac{4\|a\|_\infty\|a^{-1}\|_\infty}{\kappa_a}\int_{\mathbf{R}^d}|\nabla\sqrt{f}|^2\, d\mu,$$

et la preuve se conclut en remplaçant f par f^2. ■

Remarque. Dans les termes introduits par les auteurs de [3], la matrice de perturbation considérée ici est $A = aI_d$. Pour prendre en compte au mieux la géométrie de $\nabla^2 V$, une extension naturelle de ce résultat serait la généralisation aux perturbations non homothétiques, par exemple de la forme $A = \mathrm{diag}(a_1, \ldots, a_d)$, où $a_1, \ldots, a_d \in \mathcal{C}^\infty_+(\mathbf{R}^d, \mathbf{R})$ sont des fonctions distinctes. Ce résultat n'est malheureusement pas accessible pour l'instant avec les techniques développées ici, en raison de problèmes de commutation entre la matrice résultante $R_{t,A}$ et le processus tangent, malgré de nombreux efforts poussés dans cette direction. En particulier, cela nécessiterait une généralisation du théorème de représentation et des estimations spectrales afférentes.

Remarque (Critère de Holley-Stroock)**.** L'approche perturbative développée ici peut être comparée à celle de Holley-Stroock (présentée pour la première fois dans [46] autour du modèle d'Ising). Pour rappel, si ν est une mesure de probabilité satisfaisant une inégalité de Sobolev logarithmique et $\Phi : \mathbf{R}^d \to \mathbf{R}$ une fonction continue bornée, alors $d\mu \propto e^\Phi d\nu$ satisfait aussi une telle inégalité avec en sus

$$c_{LSI}(\mu) \leq e^{2osc(\Phi)}c_{LSI}(\nu),$$

où $osc(\Phi) = \sup(\Phi) - \inf(\Phi)$. Comme cela a déjà été évoqué, le comportement de $osc(\Phi)$ en la dimension de l'espace peut être problématique. Par exemple, pour $\Phi(x) = \sum_{i=1}^{d} \varphi(x_i)$, on a $osc(\Phi) = d \cdot osc(\varphi)$. Dans ce cas, la borne donnée par ce critère est d'autant moins intéressante que la dimension grandit ce qui, dans le cadre d'applications à la physique des particules ou aux statistiques, peut s'avérer fâcheux. En conformation avec le présent contexte, on choisit ici $\Phi = \log(a^2)$ pour $a \in \mathcal{C}_+^\infty(\mathbf{R}^d, \mathbf{R})$ une fonction bornée. L'inégalité des constantes optimales devient :

$$c_{LSI}(\mu) \leq \|a\|_\infty^4 \|a^{-1}\|_\infty^4 c_{LSI}(\mu_a),$$

de sorte que la méthode de Holley-Stroock se réduit à montrer que μ_a satisfait une inégalité de Sobolev logarithmique. Cela est assuré à peu de frais dès que μ_a vérifie une inégalité de courbure-dimension, soit ici :

$$\inf_{x \in \mathbf{R}^d} \{\rho_-(\nabla^2 V_a(x))\} > 0.$$

Cette condition peut se réécrire directement en fonction de a et V :

$$\inf_{\mathbf{R}^d} \left\{ \rho_- \left(\nabla^2 V + \frac{2}{a} \nabla^2 a - \frac{2}{a^2} \nabla a (\nabla a)^T \right) \right\} > 0.$$

Cela doit être comparé à l'hypothèse sur κ_a, qui peut être reformulée :

$$\inf_{\mathbf{R}^d} \left\{ \rho_-(\nabla^2 V) + \frac{\Delta a}{a} - \nabla V \cdot \nabla a - \frac{2}{a^2} |\nabla a|^2 \right\} > 0.$$

Ces deux expressions ne se comparent pas vraiment, mais la seconde semble plus accessible, comme nous le verrons dans la section dédiée aux exemples.

Comme cela a déjà été abordé dans l'introduction, l'espace des fonctions test considérées est important. Les résultats sont ici présentés et prouvés sur $\mathcal{C}_c^\infty(\mathbf{R}^d, \mathbf{R})$ en utilisant la densité dans $\mathcal{D}(\mathcal{E})$. Mais la considération d'un espace différent peut amener à un raffinement de la constante optimale. C'est l'objet de la section suivante, consacrée aux fonctions monotones.

2.2.2 Fonctions monotones

Outre le fait qu'il n'existe pas de définition canonique des fonctions monotones en dimension supérieure à un, le choix de la restriction à ce domaine en particulier peut sembler incongru. Il est inspiré par l'article de Miclo [62], où il établit en dimension un que la constante de Sobolev logarithmique est atteinte sur l'espace des fonction monotones (en d'autres termes, il existe des fonctions monotones pour lesquelles l'inégalité de Sobolev logarithmique est une égalité). Il n'est au demeurant pas possible d'obtenir un résultat analogue en dimension quelconque (du fait de la multiplicité des notions de monotonie dans ce cadre), mais l'étude de cette restriction apporte une précision intéressante au résultat précédent.

Notons que cette définition est de nature différente de celle de fonctions 2-monotones telle que présentée dans le dernier chapitre.

Définition 2.2.2. Une fonction mesurable $f : \mathbf{R}^d \to \mathbf{R}$ est *monotone* (dans toutes les directions) si pour tout $i = 1, \dots, d$, et pour tous $(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_d) \in \mathbf{R}^{d-1}$ fixés, la fonction partielle $f_i : x_i \mapsto f(x_1, \dots, x_d)$ est monotone.

En particulier, si f est différentiable, alors elle est monotone si et seulement si $\partial_i f$ est de signe constant sur $\mathbf{R}^d$ pour tout i.

On se concentrera dans la suite sur des fonctions f dont toutes les fonctions partielles f_i sont croissantes (resp. décroissantes). Dans de tels cas, on dira que f est elle-même croissante (resp. décroissante).

Les estimations à la base du résultat de cette section viennent (en partie) d'un article de Bakry et Michel [9], en l'honneur desquels la condition suivante a été nommée.

Définition 2.2.3 (Condition *(BM)*)**.** Une fonction $a \in \mathcal{C}^\infty_+(\mathbf{R}^d, \mathbf{R})$ satisfait la *condition de Bakry-Michel* (abrégé *(BM)*) si :

1. pour tout $i, j \in [\![1, d]\!]$, $i \neq j$, $\partial^2_{ij} V_a \leq 0$;
2. pour tout $i \in [\![1, d]\!]$, $\sum_{j=1}^d \partial^2_{ij} V_a$ est majoré.

La nouvelle estimation de $c_{LSI}(\mu)$ pour les fonctions monotones repose en grande partie sur le résultat suivant, qui établit une relation d'ordre entre $(\mathbf{P}_t)_{t\geq 0}$ et $(\mathbf{P}_{t,a})_{t\geq 0}$.

Proposition 2.2.4. *Soient $f \in \mathcal{C}^\infty_+(\mathbf{R}^d, \mathbf{R})$ et $a \in \mathcal{C}^\infty_+(\mathbf{R}^d, \mathbf{R})$ deux fonctions croissantes. Si a satisfait la condition (BM), alors pour tout $t \geq 0$:*

$$\mathbf{P}_{t,a} f \leq \mathbf{P}_t f.$$

Cette proposition est basée sur un lemme dû à Bakry et Michel dans [9], initialement utilisé dans le contexte des inégalités FKG sur $\mathbf{R}^d$ (voir par exemple [39]).

Lemme 2.2.5. *Soit $M : \mathbf{R}^d \to \mathcal{M}_d(\mathbf{R})$ une application mesurable telle que $M_{ij} \leq 0$ pour tous $i \neq j$ et $\sum_{j=1}^d M_{ij}$ est majoré pour tout i. Soit F un champ de vecteurs régulier sur $\mathbf{R}^d$. Alors toutes les composantes de $\mathcal{P}^M_t F$ sont positives dès que le sont toutes celles de F.*

Le lecteur intéressé par la preuve pourra consulter [9]. Ce résultat nous permet de prouver la Proposition 2.2.4.

Démonstration. Les techniques développées ici sont très classiques. Soient $t \geq 0$ fixé et $f \in \mathcal{C}^\infty_+(\mathbf{R}^d, \mathbf{R})$ une fonction croissante. On définit, pour $s \in [0, t]$,

$$\Psi(s) = \mathbf{P}_s(\mathbf{P}_{t-s,a} f).$$

On a $\Psi(0) = \mathbf{P}_{t,a} f$ et $\Psi(t) = \mathbf{P}_t f$, donc la preuve se réduit à démontrer que Ψ est croissante. Pour tout $s \in [0, t]$, on a :

$$\Psi'(s) = \mathbf{P}_s[(\mathbf{L} - \mathbf{L}_a)\mathbf{P}_{t-s,a} f],$$

qui peut se réécrire :

$$\Psi'(s) = \mathbf{P}_s \left[\frac{\nabla a}{a} \cdot \nabla \mathbf{P}_{t-s,a} f \right] = \mathbf{P}_s \left[\frac{\nabla a}{a} \cdot \mathcal{P}_{t-s,a}^{\nabla^2 V_a}(\nabla f) \right].$$

Comme f est croissante, toutes les composantes de ∇f sont positives, et comme a vérifie la condition *(BM)*, le Lemme 2.2.5 assure que toutes les composantes $\mathcal{P}_{t-s,a}^{\nabla^2 V_a}(\nabla f)$ sont également positives. De plus, a est positive et croissante, donc

$$\frac{\nabla a}{a} \cdot \mathcal{P}_{t-s,a}^{\nabla^2 V_a}(\nabla f) \geq 0.$$

Ainsi, par positivité de $\mathbf{P}_s$ (voir la Définition 1.1.7), $\Psi'(s) \geq 0$ et la preuve est achevée. ■

Remarque. En dimension $d = 1$, la forme particulière du semi-groupe de Feynman-Kac $(\mathcal{P}_{t,a}^{\nabla^2 V_a})_{t\geq 0}$ entraîne que la Proposition 2.2.4 reste vraie si l'on suppose uniquement que a est positive et a et f sont croissantes.

La Proposition 2.2.4 nous permet d'adapter la preuve du Théorème 2.2.1 et d'obtenir une nouvelle estimation de $c_{LSI}(\mu)$. De plus, la preuve permet de prendre en compte des fonctions de perturbation non bornées.

Théorème 2.2.6. *Soit $a \in \mathcal{C}_+^\infty(\mathbf{R}^d, \mathbf{R})$ une fonction croissante telle que $|\nabla a|/a$ est borné. On définit une nouvelle courbure perturbée :*

$$\tilde{\kappa}_a = \inf_{x \in \mathbf{R}^d} \left\{ \rho_-(\nabla^2 V(x)) - a\mathbf{L}(a^{-1})(x) \right\}.$$

Si a satisfait la condition (BM) et $\tilde{\kappa}_a > 0$, alors pour toute fonction $f \in \mathcal{C}_+^\infty(\mathbf{R}^d, \mathbf{R})$ croissante,

$$\mathrm{Ent}_\mu(f^2) \leq \frac{2}{\tilde{\kappa}_a} \int_{\mathbf{R}^d} |\nabla f|^2 \, d\mu.$$

Démonstration. Soit $f \in \mathcal{C}_+^\infty(\mathbf{R}^d, \mathbf{R})$ une fonction croissante. La preuve débute comme celle de la Proposition 1.2.17 ou du Théorème 2.2.1. L'entropie s'écrit :

$$\mathrm{Ent}_\mu(f) = \int_{\mathbf{R}^d} \int_0^{+\infty} \frac{\left| \mathcal{P}_t^{\nabla^2 V}(\nabla f) \right|^2}{\mathbf{P}_t f} \, dt \, d\mu,$$

avec la représentation :

$$\mathcal{P}_t^{\nabla^2 V}(\nabla f) = 2\mathbf{E}\left[R_{t,a} J_t^{X_a} \nabla \sqrt{f}(X_{t,a}) \sqrt{f}(X_{t,a}) \right],$$

car $|\nabla a|/a$ est borné. Le Théorème 2.1.12 et l'inégalité de Cauchy-Schwartz donnent ici :

$$\begin{aligned} \left| \mathcal{P}_t^{\nabla^2 V}(\nabla f) \right|^2 &\leq 4\mathbf{E}\left[R_{t,a}^2 |J_t^{X_a} \nabla \sqrt{f}(X_{t,a})|^2 \right] \overbrace{\mathbf{E}[f(X_{t,a})]}^{\mathbf{P}_{t,a} f} \\ &\leq 4\mathbf{E}\left[R_{t,a}^2 |J_t^{X_a} \nabla \sqrt{f}(X_{t,a})|^2 \right] \mathbf{P}_t f, \end{aligned}$$

en utilisant la Proposition 2.2.4. L'entropie vérifie donc :

$$\mathrm{Ent}_\mu(f) \leq 4 \int_{\mathbf{R}^d} \int_0^{+\infty} \mathbf{E}\left[\nabla\sqrt{f}(X_{t,a})^T J_t^{X_a} R_{t,a}^2 (J_t^{X_a})^T \nabla\sqrt{f}(X_{t,a})\right] \, dt \, d\mu.$$

On définit ici

$$J_t^a = J_t^{X_a} \exp\left(-\int_0^t \frac{\mathbf{L}_a a}{a}(X_{s,a})\, ds\right),$$

et le raisonnement spectral développé dans la preuve du Théorème 2.2.1 conduit ici à :

$$\mathrm{Ent}_\mu(f) \leq 4 \int_{\mathbf{R}^d} \int_0^{+\infty} e^{-2\tilde{\kappa}_a t} \mathbf{E}\left[\frac{a(X_{t,a})^2}{a(x)^2} |\nabla\sqrt{f}(X_{t,a})|^2\right] \, dt \, d\mu.$$

Ainsi, par invariance de $(\mathbf{P}_{t,a})_{t\geq 0}$ pour μ_a,

$$\begin{aligned}\mathrm{Ent}_\mu(f) &\leq 4 \int_0^{+\infty} e^{-2\tilde{\kappa}_a t} \int_{\mathbf{R}^d} \mathbf{P}_{t,a}\left(a^2 |\nabla\sqrt{f}|^2\right) \, d\mu_a \, dt \\ &= 4 \int_0^{+\infty} e^{-2\tilde{\kappa}_a t} \int_{\mathbf{R}^d} |\nabla\sqrt{f}|^2 \, d\mu \, dt = \frac{2}{\tilde{\kappa}_a} \int_{\mathbf{R}^d} |\nabla\sqrt{f}|^2 \, d\mu,\end{aligned}$$

et on conclut la preuve en remplaçant f par f^2. ■

On remarque que la nouvelle constante ne fait pas intervenir la norme de la fonction de perturbation, pas plus que la démonstration. À ce titre, ce résultat permet de considérer des perturbations non bornées ; c'est attendu dans la mesure où les fonctions test en jeu sont uniquement supposées croissantes (et positives), mais non nécessairement bornées ou à support compact.

On constate également que la courbure $\tilde{\kappa}_a$ diffère de celle présente dans le Théorème 2.2.1 par le facteur 2 devant la hessienne. Cela est dû à la façon dont l'inégalité de Cauchy-Schwarz est utilisée ici, en tirant parti de la comparaison donnée par la Proposition 2.2.4.

2.2.3 Exemples

Dans cette dernière section, nous illustrons le Théorème 2.2.1 sur deux exemples, l'idée étant principalement d'obtenir des résultats asymptotiques en la dimension. La forme particulière des perturbations introduites (scalaires) rend la méthode particulièrement adaptée aux potentiels dont la hessienne admet beaucoup de symétries, par exemple les potentiels radiaux. Les exemples qui suivent seront donc de ce type, et plus particulièrement des potentiels de Subbotin ou des doubles puits. Notons que des résultats analogues ont été récemment obtenus, avec des techniques différentes et pour des mesures radiales à support compact, par Cattiaux, Guillin et Wu [24].

Potentiels de Subbotin

La première classe d'exemples d'intérêt est celle des potentiels de Subbotin [84]. Cela correspond ici à $V(x) = |x|^\alpha/\alpha$ pour $\alpha > 2$, condition qui assure que V est au moins deux fois différentiable, mais que le critère de Bakry-Émery ne s'applique pas (voir la preuve et remarque subséquentes). La première étape du traitement du problème consiste à calculer la hessienne du potentiel ainsi que sa plus petite valeur propre.

Lemme 2.2.7. *Soit $a \in \mathcal{C}^\infty_+(\mathbf{R}^d, \mathbf{R})$. Alors pour tout $x \in \mathbf{R}^d$,*

$$\rho_-(2\nabla^2 V(x) - a\mathbf{L}(a^{-1})(x)I_d) = 2|x|^{\alpha-2} - a\mathbf{L}(a^{-1})(x).$$

Démonstration. On a d'abord, pour tout $x \in \mathbf{R}^d$ fixé,

$$\nabla^2 V(x) = (\alpha-2)|x|^{\alpha-4}xx^T + |x|^{\alpha-2}I_d.$$

Ainsi, $T_x := 2\nabla^2 V(x) - a\mathbf{L}(a^{-1})(x)I_d$ (vu comme un élément de $\mathcal{L}(\mathbf{R}^d)$), s'écrit comme somme d'un opérateur de rang 1 (projection sur $\mathbf{R}x$) et d'un opérateur de rang plein (multiple de l'identité), dont x est vecteur propre. Comme T_x est symétrique, les vecteurs propres sont tous orthogonaux, et on peut écrire $\mathbf{R}^d = \mathbf{R}x \oplus (\mathbf{R}x)^\perp$. Soient λ une valeur propre non nulle de T_x et y un vecteur propre associé. Alors :

- soit $y \in \mathbf{R}x$, donc $y = \beta x$ pour $\beta \in \mathbf{R}^*$, et on a alors :

  $$\lambda y = T_x y = 2\beta(\alpha-2)|x|^{\alpha-2}x + 2\beta|x|^{\alpha-2}x - \beta a\mathbf{L}(a^{-1})(x)x,$$

  ce qui donne :

  $$\lambda = 2(\alpha-1)|x|^{\alpha-2} - a\mathbf{L}(a^{-1})(x);$$

- soit $y \in (\mathbf{R}x)^\perp$, auquel cas :

  $$\lambda y = T_x y = 2|x|^{\alpha-2}y - a\mathbf{L}(a^{-1})(x)y,$$

  ce qui donne :

  $$\lambda = 2|x|^{\alpha-2} - a\mathbf{L}(a^{-1})(x).$$

Ainsi, pour tout $x \in \mathbf{R}^d$, comme $\alpha > 2$,

$$\rho_-(2\nabla^2 V(x)) - a\mathbf{L}(a^{-1})(x) = \rho_-(T_x) = 2|x|^{\alpha-2} - a\mathbf{L}(a^{-1})(x).$$

■

Dans la suite, le cas particulier $\alpha = 4$ (potentiel quadrique) est traité. En effet, les calculs deviennent rapidement inaccessibles pour des valeurs de α générales, d'autant plus que l'on cherche à déterminer explicitement la dépendance de la constante trouvée en d et en α. Le critère de Bakry-Émery ne s'applique clairement pas à ce cas-là, étant donné que $\rho_-(\nabla^2 V(x))$ s'annule pour $x = 0$.

Théorème 2.2.8. *Soit $\mu(dx) \propto \exp(-|x|^4/4)dx$. Il existe une constante universelle $c > 0$ telle que pour toute fonction $f \in C_c^\infty(\mathbf{R}^d, \mathbf{R})$, on a :*

$$\mathrm{Ent}_\mu(f^2) \leq c \int_{\mathbf{R}^d} |\nabla f|^2 d\mu.$$

En particulier, c ne dépend pas de la dimension d.

On déduit de la preuve que $c = 8e\sqrt{3}$ est une telle constante, dont la valeur dépend néanmoins grandement de la méthode employée.

Démonstration. La première interrogation quant à l'application du Théorème 2.2.1 porte sur le choix de la fonction de perturbation a. En pratique, a doit compenser (localement) le manque de convexité de V (ici aux points où $\nabla^2 V(x) \leq 0$, soit $x = 0$). Un premier choix se porte sur la fonction

$$a(x) = \exp\left(\frac{\varepsilon}{2} \arctan(|x|^2)\right), \quad x \in \mathbf{R}^d.$$

En effet, l'arc tangente est équivalente à l'identité en 0 (point de non-convexité de V) et est constante à l'infini (de sorte que a est bien bornée). De plus, la fonction carré est uniformément convexe sur $\mathbf{R}^d$, ce qui entraîne que la hessienne de a est définie positive au voisinage de l'origine. Finalement, l'exponentielle assure que a est positive et facilite quelque peu les calculs. Notons que ce choix de fonction de perturbation est motivé par des résultats analogues sur le trou spectral, dans les preuves desquels le choix d'une fonction de perturbation proche de la non-intégrabilité fournit des estimations intéressantes sur les constantes optimales (voir par exemple [18, 3]).

L'étape suivante consiste à calculer explicitement κ_a. Avec notre définition de a, il vient :

$$-a\mathbf{L}(a^{-1})(x) = \varepsilon \frac{d + |x|^4(d-4)}{(1+|x|^4)^2} - \varepsilon \frac{|x|^4}{1+|x|^4} - \varepsilon^2 \frac{|x|^2}{(1+|x|^4)^2}, \quad x \in \mathbf{R}^d,$$

et il faut donc minimiser en $x \in \mathbf{R}^d$:

$$2|x|^2 + \varepsilon \frac{d + |x|^4(d-4)}{(1+|x|^4)^2} - \varepsilon \frac{|x|^4}{1+|x|^4} - \varepsilon^2 \frac{|x|^2}{(1+|x|^4)^2},$$

que l'on réécrit, en posant $t = |x|^2$,

$$\kappa_a = \inf_{t \geq 0} \left(2t + \varepsilon \frac{d + t^2(d-4)}{(1+t^2)^2} - \varepsilon \frac{t^2}{1+t^2} - \varepsilon^2 \frac{t}{(1+t^2)^2} \right).$$

L'optimisation des polynômes, et a fortiori des fractions rationnelles, se fait rarement sans douleurs, en particulier quand on cherche à suivre les valeurs des différents paramètres (ici ε et d). Pour simplifier, on se ramène au cas où le minimum est atteint en $t = 0$, ce qui revient à montrer que, pour tout $t \geq 0$,

$$2t^4 - \varepsilon(d+1)t^3 + 4t^2 - \varepsilon(d+5)t + 2 - \varepsilon^2 \geq 0.$$

Notons g la fonction ci-dessus. Il vient immédiatement que $\varepsilon \leq \sqrt{2}$ est une condition nécessaire, mais pas suffisante, pour assurer la positivité de g. Pour simplifier les choses, on peut supposer que g'' est positive. C'est vrai dès que

$$\varepsilon < \frac{8}{\sqrt{3}(d+1)}.$$

Considérons donc $\varepsilon \leq \dfrac{1}{\sqrt{3}(d+1)}$. Avec ce choix, sachant que $d \geq 1$, on a alors pour tout $t \geq 0$:

$$g(t) \geq 2t^4 - \frac{8t^3}{3\sqrt{3}} + 4t^2 - \frac{8t}{\sqrt{3}} + 2 - \frac{1}{6}.$$

On vérifie alors aisément que le terme de droite est positif, et donc g également, sur $\mathbf{R}_+$. On prend alors $\kappa_a = \varepsilon d$, et le Théorème 2.2.1 donne la borne suivante :

$$c_{LSI}(\mu) \leq \frac{4e^{\varepsilon\pi/4}}{\varepsilon d},$$

avec $\varepsilon \leq \dfrac{1}{\sqrt{3}(d+1)}$ (qui implique bien $\varepsilon \leq \sqrt{2}$). Finalement, cette borne est minimisée pour $\varepsilon \in \left(0, \frac{1}{\sqrt{3}(d+1)}\right]$ pour obtenir :

$$c_{LSI}(\mu) \leq \frac{4\sqrt{3}(d+1)}{d} e^{\pi/4\sqrt{3}(d+1)}.$$

La borne ci-dessus est uniformément majorée pour $d \in \mathbf{N}^\star$, et on choisit alors $c = 8e\sqrt{3}$ comme constante universelle du théorème. ■

Remarque. Cette preuve met en lumière les points les plus importants dans l'application du Théorème 2.2.1. En effet, le choix de la fonction de perturbation (ou famille de fonctions) a est central. Mais le point le plus important, et technique, est probablement le calcul explicite de κ_a, dans la mesure où tous les paramètres doivent être suivis avec attention.

À une constante numérique près, la borne sur ε dans la preuve ci-dessus est optimale (avec cette méthode d'optimisation). En effet, le problème revient à montrer que la fonction g définie sur $\mathbf{R}_+$ par

$$g(t) = 2t^4 - \varepsilon(d+1)t^3 + 4t^2 - \varepsilon(d+5)t + 2 - \varepsilon^2, \quad t \geq 0,$$

est positive. Si on suppose que ε est d'ordre $(d+1)^{-r}$ pour un certain $r \in (0,1)$, alors pour d grand, pour tout $t > 0$ fixé,

$$g(t) \sim 2t^4 - d^{1-r}t^3 + 4t^2 - d^{1-r}t + 2 - d^{-2r}.$$

Si on prend $t = 3/d^{1-r}$, il vient alors :

$$g(3/(d+1)^r) \sim \frac{162}{d^{4(1-r)}} + \frac{9}{d^{2(1-r)}} - \frac{1}{2^{2r}} - 1 < 0$$

quand d croît, suite à quoi le minimum de $t \mapsto \varepsilon d + tg(t)$ ne peut être atteint en $t = 0$.

Il est difficile de savoir si la borne trouvée est proche de l'optimal (en ce qui concerne la dépendance en la dimension). On peut néanmoins conjecturer, en se basant sur les résultats de [20], que $c_{LSI}(\mu)$ est d'ordre (au mieux) $1/\sqrt{d}$. En effet, les auteurs de cet article montrent que le trou spectral pour la distribution quadrique est d'ordre $\sqrt{d}$, et donc que la constante de Poincaré pour cette mesure est d'ordre exactement $1/\sqrt{d}$ (et donc que $c_{LSI}(\mu)$ ne peut pas être d'ordre d^{-r} pour $r > 1/2$, voir la Proposition 1.2.19). Dès lors il apparaît incertain que la procédure d'optimisation utilisée permette d'atteindre l'optimal. De meilleures techniques de minimisation constitueraient une amélioration notable dans le cadre de l'utilisation du Théorème 2.2.1 pour l'obtention de constantes explicites.

Remarque. La méthode de Holley-Stroock telle que développée dans la section précédente donne ici, après des calculs guère plus plaisants, à une conclusion comparable. Les constantes en jeu ne sont néanmoins pas explicites et la méthode laisse moins de degrés liberté que celle présentée dans ce chapitre.

Double puits

L'exemple suivant est une perturbation du précédent, bien connu de nos amis physiciens : le potentiel à double puits. On pose $V(x) = |x|^4/4 - \beta|x|^2/2$, où $\beta > 0$ décrit la taille de la région concave. Ce potentiel est convexe à l'infini, mais sa matrice hessienne est définie négative au voisinage de l'origine, le critère de Bakry-Émery ne s'applique donc pas. Néanmoins, on pourrait raisonnablement s'attendre à retrouver un comportement similaire à celui observé dans le Théorème 2.2.8 quand β est petit.

La hessienne de V est, comme dans le cas précédent, calculable explicitement, ce qui conduit au résultat suivant.

Lemme 2.2.9. *Soit $a \in \mathcal{C}^\infty_+(\mathbf{R}^d, \mathbf{R})$. Alors pour tout $x \in \mathbf{R}^d$,*

$$\rho_-(2\nabla^2 V(x) - a\mathbf{L}(a^{-1})(x)I_d) = 2|x|^2 - 2\beta - a\mathbf{L}(a^{-1})(x).$$

Démonstration. La preuve est identique à celle du Lemme 2.2.7. ■

Théorème 2.2.10. *Soit $\mu(dx) \propto \exp(-|x|^4/4 + \beta|x|^2/2)dx$, $\beta \in (0, 1/2)$. Il existe une constante universelle $c_\beta > 0$ telle que, pour toute fonction $f \in \mathcal{C}^\infty_c(\mathbf{R}^d, \mathbf{R})$, on a :*

$$\mathrm{Ent}_\mu(f^2) \leq c_\beta \int_{\mathbf{R}^d} |\nabla f|^2 d\mu.$$

Ici aussi, c_β ne dépend pas de la dimension.

De la preuve on déduit que $c_\beta = \frac{4e}{1-2\beta}$ fait l'affaire, et l'explosion quand $\beta \to 1/2$ est un artefact de calcul qui ne porte a priori pas d'information qualitativement pertinente (lié à la méthode utilisée pour estimer κ_a du Théorème 2.2.1).

Démonstration. Cette preuve est très analogue à la précédente. En particulier, on prend de même pour tout $x \in \mathbf{R}^d$:

$$a(x) = \exp\left(\frac{\varepsilon}{2}\arctan(|x|^2)\right),$$

ce qui donne, pour $t = |x|^2$,

$$\kappa_a = \inf_{t\geq 0}\left(2t - 2\beta + \varepsilon\frac{d + t^2(d-4)}{(1+t^2)^2} - \varepsilon(t-\beta)\frac{t}{1+t^2} - \varepsilon^2\frac{t}{(1+t^2)^2}\right).$$

En suivant la même procédure, on cherche à montrer que le minimum vaut $\varepsilon d - 2\beta$, atteint en $t = 0$. Cela revient à montrer que, pour tout $t \geq 0$,

$$g(t) := 2t^4 - \varepsilon(d+1)t^3 + (4+\beta)t^2 - \varepsilon(d+5)t + 2 - \varepsilon^2 + \beta \geq 0,$$

adjoint de la condition $\varepsilon > 2\beta/d$, pour assurer la positivité de κ_a.

La première condition nécessaire à intervenir est $\varepsilon \leq \sqrt{\beta+2}$. De plus, en regard de la preuve et de la remarque précédentes, on attend de ε qu'il soit d'ordre $\frac{1}{d+1}$. Pour simplifier les calculs, on considère $\varepsilon = \frac{2}{d+1}$. Avec les conditions $\varepsilon > 2\beta/d$ et $\varepsilon \leq \sqrt{\beta+2}$, cela implique que β ne doit pas excéder $d/d+1$ pour tout d, ce qui revient (pour obtenir une borne uniforme en la dimension) à $\beta < 1/2$. Pour résumer, nous avons

$$\varepsilon = \frac{2}{d+1} \quad \text{et} \quad 0 \leq \beta < \frac{1}{2}.$$

Sous ces hypothèses, la fonction g peut être minorée comme suit :

$$g(t) \geq 2t^4 - 2t^3 + 4t^2 - 2t + 1 + \beta, \quad t \geq 0.$$

Le terme de droite est positif sur $\mathbf{R}_+$, ce choix de ε entraîne donc :

$$\kappa_a = \frac{2d}{d+1} - 2\beta.$$

Ainsi, d'après le Théorème 2.2.1,

$$c_{LSI}(\mu) \leq \frac{4(d+1)}{2d(1-\beta) - 2\beta}e^{\frac{\pi}{2(d+1)}}.$$

Cette constante peut être bornée uniformément en $d \in \mathbf{N}^\star$, ce qui donne bien $c_\beta = \frac{4e}{1-2\beta}$ comme constante universelle. ■

Remarque. On rappelle que la restriction sur β est un artefact de calcul, et que de plus $c_\beta \xrightarrow[\beta\to\frac{1}{2}^-]{} +\infty$. Il n'en reste pas moins que le comportement dimensionnel est le même que celui observé pour la distribution quadrique dans le Théorème 2.2.8. Encore une fois, les techniques d'optimisation utilisées sont probablement loin de l'optimal, ce qui illustre le fait qu'une part importante de l'application du Théorème 2.2.1 réside dans l'identification et la mise en œuvre de méthodes de minimisation adaptées.

Remarque. Comme dans le cas précédent, la méthode de Holley-Stroock conduit à des résultats analogues, bien que les calculs soient bien plus complexes, en particulier si l'on souhaite suivre avec précision la dépendance en β.

2.3 Perspectives

Outre l'aspect optimisation, dont on a déjà mentionné qu'il améliorerait la portée du résultat s'il était plus développé, les inégalités de Sobolev logarithmiques sont connues pour leur relation avec les inégalités de concentration. Brièvement, une mesure μ satisfait une inégalité de concentration à vitesse α s'il existe une constante $c > 0$ telle que pour toute fonction $F : \mathbf{R}^d \to \mathbf{R}$ lipschitzienne, pour toute médiane m de F et pour tout $r \geq 0$,

$$\mathbf{P}(F(X) > m + r) \leq e^{-\alpha(r)/c}, \quad X \sim \mu.$$

Cette propriété, de nature en fait très géométrique, décrit la façon dont μ concentre la masse dans l'espace. Elle a trait en dimension 1 à la queue des distributions. L'enjeu des inégalités de concentration est de déterminer la fonction α optimale, celle qui donnera la meilleure décroissance en r à l'infini. Plusieurs cas particuliers ont été étudiés, notamment la concentration exponentielle ($\alpha = |\cdot|$), mais c'est la concentration gaussienne ($\alpha = |\cdot|^2$) qui est la plus recherchée. L'argument de Herbst présenté en introduction indique que toute mesure qui satisfait une inégalité de Sobolev logarithmique satisfait également une inégalité de concentration gaussienne. Explorer les liens entre ce domaine et les résultats présentés ici constituerait alors certainement un travail intéressant. Pour plus d'informations sur les inégalités de concentration, la monographie de Michel Ledoux [59] reste une référence, et les liens évoqués sont aussi présentés dans [8].

Une autre famille d'inégalités présente un fort lien avec les techniques évoquées ici : les inégalités de covariance. Ce nom assez générique désigne le fait de pouvoir contrôler la covariance de deux fonctions par deux termes distincts qui les séparent. Autrement dit, trouver deux fonctionnelles F et G telles que pour toutes fonctions $f, g : \mathbf{R}^d \to \mathbf{R}$ assez régulières et $X \sim \mu$, on a

$$\mathrm{Cov}(f(X), g(X)) \leq F(f)G(g).$$

Ces fonctionnelles peuvent être de nature assez diverses, voir par exemple [3, 16]. Dans l'article [3], les auteurs utilisent des techniques de Feynman-Kac pour obtenir de telles inégalités, techniques qui pourraient être revues à la lumière des présents travaux. De plus, les inégalités de covariance sont liées à de très nombreuses autres inégalités, comme la conjecture de corrélation gaussienne [12] ou les inégalités de concentration [16], et constituent en elles-mêmes un sujet d'étude intéressant.

Enfin, une approche numérique s'inspirant de celle développée en fin de thèse pourrait être instructive, bien qu'applicable essentiellement en petite dimension. Cela reste néanmoins pertinent dans certaines situations, notamment en statistiques, comme le suggère le dernier chapitre pour les inégalités de Poincaré.

Chapitre 3

Une approche algébrique et géométrique à l'étude spectrale des générateurs

DANS CE TROISIÈME CHAPITRE, réalisé en collaboration avec A. Joulin et M. Bonnefont (et objet d'un travail en cours de rédaction), on approfondit la réflexion sur l'analyse spectrale de l'opérateur $\mathbf{L}$ amorcée dans l'introduction (voir par exemple la Proposition 1.2.3). Plus précisément, on forme l'observation suivante.

Étant donné un potentiel V, la condition de courbure-dimension revient à comparer la hessienne de V avec la matrice identité. Or, la matrice identité est elle-même la hessienne du potentiel gaussien ($x \mapsto |x|^2/2$), la condition précitée peut donc être vue comme une comparaison des potentiels via les hessiennes. On sait de plus que le spectre de l'opérateur d'Ornstein-Uhlenbeck est discret, et on peut montrer que, sous la condition de courbure dimension, le spectre de $\mathbf{L}$ l'est aussi. Partant, il est raisonnable de se demander si la comparaison des potentiels peut se transposer aux spectres respectifs.

En d'autres termes, si $\mathbf{L}$ satisfait $CD(\rho, \infty)$ pour $\rho > 0$, peut-on obtenir une inégalité sur les valeurs propres du type :

$$\lambda_k(-\mathbf{L}) \geq \rho \lambda_k(-\mathbf{L}_{OU}),$$

pour tout $k \in \mathbf{N}$?

La pertinence de cette question tient dans le fait que le terme de droite de l'inégalité précédente est toujours explicite. En effet, parmi les nombreuses manifestations du miracle gaussien, il apparaît que le spectre de $\mathbf{L}_{OU}$ est explicite et peut-être complètement décrit, multiplicités comprises. Il est alors naturel de se demander si ces informations peuvent être préservées via la comparaison des potentiels.

Cette question peut être posée dans un contexte plus général, où l'on compare un opérateur de diffusion sur une variété riemannienne quelconque à la loi gaussienne sur l'espace euclidien. C'est en ces termes qu'Emanuel Milman a énoncé la question dans

[63]. A priori fausse en toute généralité (l'auteur donne comme contre-exemple dans ce sens celui de la d-sphère normalisée), il parvient à y répondre par l'affirmative dans le cas euclidien. Il obtient le résultat suivant.

Théorème (Milman '18)**.** *Supposons que $\nabla^2 V \geq \rho I_d$ pour un certain $\rho > 0$. Alors pour tout $k \geq 1$,*

$$\lambda_k(-\mathbf{L}) \geq \rho \lambda_k(-\mathbf{L}_{OU}).$$

Sommairement, sa preuve utilise le théorème de contraction de Caffarelli pour établir que la condition de courbure-dimension induit une contraction de l'espace sous-jacent vers l'espace euclidien muni de la mesure gaussienne. Le coefficient de contraction est donné explicitement en fonction de la courbure et permet de conclure à la contraction du spectre. Notons que ce résultat ne dit rien sur la multiplicité des valeurs propres.

Une question analogue, laissée ouverte par le résultat de Milman, porte sur l'estimation de l'écart entre deux valeurs propres successives. Plus précisément, sous la même hypothèse de convexité uniforme, si λ et λ' sont deux valeurs propres successives *distinctes*, que peut-on dire de la différence $|\lambda - \lambda'|$? Cette question a été étudiée par Michel Bonnefont et Aldéric Joulin dans [19] pour les deux premières valeurs propres (pour lesquelles les espaces propres sont descriptibles explicitement) et dans [17] dans le cas de la dimension un, en se basant sur un article de Jon Johnsen [52] (qui sert de base au présent travail).

Ce chapitre a pour but de donner des éléments de réponse aux deux questions précitées par des moyens plus analytiques en travaillant directement sur l'opérateur $\mathbf{L}$ (et sans passer par des considérations de transport optimal). Les principaux outils mis en place à cette fin sont une structure algébrique en lien avec la dérivée extérieure et la caractérisation de Courant-Fischer des valeurs propres. Ces outils d'analyse spectrale, ainsi que le cas gaussien, sont rappelés en première partie de chapitre. La structure algébrique est introduite en suivant, avant de clore par les estimations spectrales d'intérêt.

Il est important de mentionner que les résultats obtenus dans ce cadre sont partiels, et ne permettent pas, en l'état, de retrouver le résultat de Milman. Il est néanmoins probable que les outils développés permettent d'y arriver par une analyse plus fine. À titre indicatif, le résultat principal de ce chapitre (qui sera développé et prouvé en dernière section) peut se présenter comme suit.

Théorème. *Supposons que $\nabla^2 V \geq \rho I_d$ pour un certain $\rho > 0$. Il existe une suite croissante d'entiers $(a_n)_{n\geq 0}$ telle que pour tous $k, n \in \mathbf{N}$ tels que $a_n < k \leq a_{n+1}$,*

$$\lambda_k(-\mathbf{L}) \geq (n+1)\rho.$$

La suite $(a_n)_{n\geq 0}$ décrit la vitesse de croissance à l'infini des valeurs propres, et est liée à la dimension des espaces propres associés. Plus précisément, les valeurs propres croissent par paliers (liés à la multiplicité), et la suite $(a_n)_{n\geq 0}$ vise à estimer leur taille.

Comme il sera discuté en fin de chapitre, le point d'achoppement de la méthode réside dans une estimation fine de cette suite.

Table des matières

3.1 Éléments d'analyse spectrale 65
3.1.1 Décomposition du spectre 65
3.1.2 Cas du générateur d'Ornstein-Uhlenbeck 67
3.1.3 Trou spectral et caractérisation de Courant-Fischer 73
3.2 Contexte algébrique 76
3.2.1 Algèbre tensorielle 77
3.2.2 Opérateurs de diffusion et Schrödinger sur l'algèbre tensorielle et estimations préliminaires 79
3.2.3 Transformations de Riesz 82
3.3 Estimations spectrales pour $-\mathbf{L}$ 84
3.3.1 Résultat principal 84
3.3.2 Optimalité dans le cas gaussien 89
3.3.3 Perspectives 91

3.1 Éléments d'analyse spectrale

Le but de cette section est double. D'abord rappeler des éléments d'analyse structurelle du spectre (définition de valeur propre, spectres discret et essentiel, caractérisation de Courant-Fischer), ensuite exhiber un cas particulier où tous ces concepts sont illustrés explicitement.

Dans toute la suite, on notera $\langle \cdot, \cdot \rangle$ le produit scalaire dans L^2 sans préciser la mesure de référence si cette dernière est clairement identifiée.

3.1.1 Décomposition du spectre

Si la notion de valeur et espace propres est rarement ambiguë en dimension finie, il en est autrement en dimension infinie. La définition classique du spectre en analyse fonctionnelle fait intervenir la notion de résolvante, et c'est celle que nous utiliserons ici. Les définitions seront données pour l'opérateur $-\mathbf{L}$, mais peuvent s'étendre à des opérateurs sur des espaces de Banach très généraux.

Définition 3.1.1 (Spectre)**.** On appelle *ensemble résolvant de* $-\mathbf{L}$, noté $\rho(-\mathbf{L})$, l'ensemble des nombres complexes λ tels que la résolvante

$$R_\lambda = (\lambda \mathrm{id} + \mathbf{L})^{-1}$$

est continue. On définit alors le *spectre de* $-\mathbf{L}$ comme le complémentaire du résolvant : $\sigma(-\mathbf{L}) = \mathbf{C} \backslash \rho(-\mathbf{L})$.

Il suit immédiatement du caractère défini positif de $-\mathbf{L}$ que $\sigma(-\mathbf{L}) \subset (\mathbf{R}_+)$.

Cette définition coïncide avec les notions classiques de valeurs propres en dimension finie, où la résolvante R_λ existe et est continue dès que $\det(\lambda \mathrm{id} + \mathbf{L}) \neq 0$. Le théorème fondamental de l'algèbre implique alors que le spectre est discret et fini. Une telle alternative n'est pas vraie en dimension infinie, où toutes les applications linéaires ne sont pas continues. En particulier, la propriété fondamentale qui définit les valeurs propres en dimension finie n'est plus satisfaite pour tout le spectre. Il convient donc d'expliciter la partie du spectre qui correspond aux valeurs propres en dimension infinie, qu'on nomme *spectre ponctuel*.

Définition 3.1.2 (Spectre ponctuel)**.** On appelle *spectre ponctuel* de $-\mathbf{L}$, noté $\sigma_p(-\mathbf{L})$, l'ensemble des $\lambda \in \sigma(-\mathbf{L})$ tels que $\lambda \mathrm{id} + \mathbf{L}$ n'est pas injectif.

Si $\lambda \in \sigma_p(-\mathbf{L})$ est une valeur propre, il suit qu'il existe une fonction $f \in L^2(\mu)$ non identiquement nulle telle que, μ-presque sûrement,

$$-\mathbf{L}f = \lambda f.$$

C'est en ce sens que l'on désigne comme *valeurs propres* les valeurs du spectre ponctuel. Une telle fonction f est appelée *fonction propre* associée à λ. Le sous-espace vectoriel de $L^2(\mu)$ engendré par ces fonctions propres est appelé *espace propre* associé à la valeur propre λ, que l'on notera E_λ. On désigne par *multiplicité de la valeur propre* λ la dimension de cet espace, qui peut être infinie (mais toujours bien définie dans le présent cadre hilbertien).

Comme l'espace sous-jacent à $-\mathbf{L}$ est hilbertien, et si ce dernier est auto-adjoint (ce qui est le cas ici), ces espaces propres sont de plus orthogonaux. En effet, considérant $\lambda, \lambda' \in \sigma_p(-\mathbf{L})$ non nulles et $f \in E_\lambda$, $g \in E_{\lambda'}$, il vient

$$\langle f, g \rangle = \frac{1}{\lambda} \langle -\mathbf{L}f, g \rangle = \frac{1}{\lambda} \langle f, -\mathbf{L}g \rangle = \frac{\lambda'}{\lambda} \langle f, g \rangle.$$

Il suit que soit $\lambda = \lambda'$, soit $\langle f, g \rangle = 0$, ce qui établit que deux espaces propres distincts sont orthogonaux.

En dimension finie, l'étude spectrale des opérateurs est souvent motivée par la réduction des endomorphismes. En effet, l'espace peut être présenté comme somme directe de sous-espaces propres, ce qui induit une décomposition dans une base de vecteurs propres adaptée au travail avec cet endomorphisme. Le but en dimension infinie est, dans une certaine mesure, assez similaire. Deux différences fondamentales changent néanmoins radicalement le cadre de cette étude : le nombre a priori infini d'espaces propres et la dimension infinie de ces derniers. D'aucun arguerait qu'on ne serait guère avancé d'avoir décomposé un espace de dimension infinie en somme d'espaces de dimension encore infinie. D'autant plus si ces espaces sont indexés par une famille continue d'indices (au sens de la puissance du continu). Cette remarque pousse à distinguer, parmi les valeurs propres du spectre ponctuel, celles qui sont isolées et de multiplicité finies. Le sous-ensemble du spectre ainsi défini se nomme, à propos, *spectre discret*.

Définition 3.1.3 (Spectre discret)**.** On appelle *spectre discret* de $-\mathbf{L}$, noté $\sigma_{disc}(-\mathbf{L})$, l'ensemble des valeurs propres $\lambda \in \sigma_p(-\mathbf{L})$ qui vérifient les deux propriétés suivantes :

- λ est isolée dans $\sigma(-\mathbf{L})$ (pour la topologie induite par $\mathbf{R}$),
- l'espace propre E_λ est de dimension finie.

On a donc la chaîne d'inclusions (strictes a priori) suivante :

$$\sigma_{disc}(-\mathbf{L}) \subset \sigma_p(-\mathbf{L}) \subset \sigma(-\mathbf{L}).$$

Les valeurs propres isolées de multiplicité finie sont a priori les objets les plus proches des valeurs propres en dimension finie, ce qui justifie en soi leur étude. Étant donné qu'elles ne constituent a priori pas l'intégralité du spectre, on définit le *spectre essentiel* de $-\mathbf{L}$ comme le complémentaire de $\sigma_{disc}(-\mathbf{L})$ dans $\sigma(-\mathbf{L})$. L'objet d'étude principal de ce chapitre sera le spectre discret de $-\mathbf{L}$, on s'intéressera donc à la décomposition suivante :

$$\sigma(-\mathbf{L}) = \sigma_{disc}(-\mathbf{L}) \sqcup \sigma_{ess}(-\mathbf{L}).$$

À la lumière de cette décomposition, on peut affiner la décomposition spectrale donnée en introduction (après la Proposition 1.1.17). Toute fonction $f \in L^2(\mu)$ s'écrit

$$f = \sum_{\lambda \in \sigma_{disc}(-\mathbf{L})} P_\lambda f + \int_{\lambda \in \sigma_{ess}(-\mathbf{L})} dP_\lambda f.$$

De nombreux autres sous-ensembles du spectre existent, donnant lieu à d'autres décompositions, en lien avec d'autres propriétés des valeurs propres (spectre résiduel, spectre continu, etc.). Le choix de cette décomposition n'est néanmoins pas arbitraire, et est motivé par le cas particulier annoncé en introduction : le générateur d'Ornstein-Uhlenbeck. Non content d'être relié à la mesure préférée des probabilistes et statisticiens, cet opérateur a le bon goût d'avoir un spectre essentiel vide : son spectre est uniquement constitué de valeurs propres isolées de multiplicité finie. La question posée en introduction peut alors être complétée : sous l'hypothèse $CD(\rho, \infty)$, le spectre de $-\mathbf{L}$ est-il complètement discret ? Et si oui, les valeurs propres et espaces propres peuvent-ils être reliés à ceux du générateur d'Ornstein-Uhlenbeck ?

Avant de tenter de répondre à cette question, la section suivante s'attelle à une description exhaustive du spectre de l'opérateur $-\mathbf{L}_{OU}$, avant de présenter quelques techniques d'estimation reliées à la caractérisation de Courant-Fischer des valeurs propres.

3.1.2 Cas du générateur d'Ornstein-Uhlenbeck

La diversité pléthorique des opérateurs de diffusion oblige à envisager de nombreuses structures différentes pour le spectre : complètement discret, a contrario complètement essentiel, ou avec une composante discrète et une composante essentielle toutes deux non vides. Cette section s'intéresse au générateur d'Ornstein-Uhlenbeck, qui présente un spectre complètement discret. D'autres exemples seront évoqués en fin de section, pour lesquels le spectre essentiel n'est pas vide.

On rappelle que le générateur d'Ornstein-Uhlenbeck est défini, pour $f \in \mathcal{C}_c^\infty(\mathbf{R}^d, \mathbf{R})$, $x \in \mathbf{R}^d$,

$$\mathbf{L}_{OU} f(x) = \Delta f(x) - x \cdot \nabla f(x),$$

et qu'il admet la mesure gaussienne centrée réduite γ_d comme unique loi invariante. On scinde l'étude du spectre de cet opérateur en deux : d'abord une étude en dimension un, puis la généralisation en dimension quelconque.

Cas de la dimension un

Bien que peu intéressant quant aux propriétés asymptotiques de la grande dimension, le cas des fonctions d'une variable réelle permet d'établir facilement des résultats généralisables en dimension supérieure sous une forme simple à écrire et à comprendre. On rappelle en préambule la définition des polynômes d'Hermite en dimension un.

Définition 3.1.4 (Polynômes d'Hermite)**.** Soit $n \in \mathbf{N}$. Le n-ième polynôme d'Hermite sur $\mathbf{R}$ est défini, pour tout $x \in \mathbf{R}$, par :

$$H_n(x) = (-1)^n e^{x^2/2} \frac{d^n}{dx^n} e^{-x^2/2}.$$

Ces polynômes ont très tôt présenté un exemple de famille de fonctions orthogonales dans $L^2(\gamma_1)$, et on a pour tous $m, n \in \mathbf{N}$:

$$\langle H_n, H_m \rangle = n! \delta_{nm}.$$

On considérera donc préférentiellement dans la suite la famille de polynômes orthonormale $(h_n)_{n \in \mathbf{N}}$ définie, pour tout $n \in \mathbf{N}$, par :

$$h_n(x) = \frac{1}{\sqrt{n!}} H_n(x).$$

Outre l'orthonormalité, cette famille possède d'intéressantes propriétés relatives à la dérivation, qui s'avèrent fort utiles dans l'étude de $-\mathbf{L}_{OU}$. En effet, pour tout $n \in \mathbf{N}$ et $x \in \mathbf{R}$, on a d'une part :

$$h_n'(x) = x h_n(x) - \sqrt{n+1} h_{n+1}(x),$$

et d'autre part, par la formule de Leibnitz,

$$h_{n+1}(x) = \frac{x}{\sqrt{n+1}} h_n(x) - \sqrt{\frac{n}{n+1}} h_{n-1}(x),$$

ce qui donne la seconde formule pour h_n' :

$$h_n'(x) = \sqrt{n} h_{n-1}(x).$$

Ces deux formules permettent un premier pas dans la description de $\sigma(-\mathbf{L}_{OU})$.

Proposition 3.1.5. *Pour tout $n \in \mathbf{N}$, le polynôme h_n est une fonction propre de $-\mathbf{L}_{OU}$ associée à la valeur propre n.*

La preuve de ce résultat est un calcul direct de $-\mathbf{L}_{OU}h_n$ pour tout $n \in \mathbf{N}$, en utilisant les formules ci-dessus.

La première conséquence à la fois immédiate et importante de ce résultat est l'inclusion $\mathbf{N} \subset \sigma_p(-\mathbf{L}_{OU})$. Mais on peut en réalité aller plus loin. En effet, la famille des polynômes d'Hermite $(h_n)_{n\in\mathbf{N}}$ forme une base hilbertienne de $L^2(\gamma_1)$. Cela peut se montrer, par exemple, en remarquant qu'une fonction orthogonale à tous les monômes est nulle presque sûrement. Cela entraîne donc immédiatement que chaque espace propre est de dimension un, ainsi que la décomposition

$$L^2(\gamma_1) = \bigoplus_{n\in\mathbf{N}} \mathrm{Vect}(h_n).$$

Il suit également que $-\mathbf{L}_{OU}$ n'admet pas d'autre valeur propre que les entiers. Dans le cas contraire, toute fonction propre associée serait orthogonale à tous les polynômes d'Hermite (voir la remarque suivant la Définition 3.1.2), et donc serait presque sûrement nulle. On a donc la double inclusion qui assure que $\sigma_p(-\mathbf{L}_{OU}) = \mathbf{N}$.

Enfin, cette décomposition implique que toute fonction $f \in L^2(\gamma_1)$ est complètement décrite par les projecteurs sur les espaces engendrés par les polynômes d'Hermite :

$$f = \sum_{n\in\mathbf{N}} \langle f, h_n\rangle h_n,$$

ce qui induit, en vertu de la décomposition spectrale rappelée ci-dessus, que $\sigma_{ess}(-\mathbf{L}_{OU}) = \emptyset$. Il s'ensuit la description suivante du spectre de $-\mathbf{L}_{OU}$.

Théorème 3.1.6. *Le spectre de $-\mathbf{L}_{OU}$ est discret et égal à $\mathbf{N}$. Toutes les valeurs propres sont de multiplicité 1 et une base hilbertienne de vecteurs propres associés est donnée par les polynômes d'Hermite $(h_n)_{n\in\mathbf{N}}$.*

Une description aussi exhaustive du spectre est en général inaccessible, en-dehors de certains cas très spécifiques. Le superlatif "miracle gaussien" n'est pas usurpé pour qualifier ce résultat, qui justifie les tentatives de comparaison telles que décrites en introduction. De plus, le caractère tensoriel de la mesure gaussienne en dimension quelconque permet une transposition aisée de cette structure, qui fait l'objet de la section suivante.

Dimension quelconque

La plupart des calculs techniques ont été effectués dans la section précédente, celle-ci consistera essentiellement à exploiter la structure tensorielle de la mesure gaussienne pour décrire le spectre de $-\mathbf{L}_{OU}$ en dimension quelconque. À cette fin, on rappelle les notations des fonctions partielles introduites dans le chapitre précédent.

Définition 3.1.7. Soit $f \in L^2(\gamma_d)$. Pour $i \in [\![1, d]\!]$ et $(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_d) \in \mathbf{R}^{d-1}$ fixé, on définit la fonction partielle $f_i : \mathbf{R} \to \mathbf{R}$ par

$$f_i(x_i) = f(x),$$

pour $x_i \in \mathbf{R}$.

On notera $-\mathbf{L}_{OU,1}$ l'opérateur d'Ornstein-Uhlenbeck unidimensionnel de la section précédente. Avec la définition précédente, on a la décomposition suivante de $-\mathbf{L}_{OU}$: pour toute fonction $f \in L^2(\gamma_d)$ et tout $x \in \mathbf{R}^d$,

$$-\mathbf{L}_{OU} f(x) = -\sum_{i=1}^{d} \mathbf{L}_{OU,1} f_i(x_i).$$

Elle découle directement de la définition de $-\mathbf{L}_{OU}$:

$$-\mathbf{L}_{OU} f(x) = -\sum_{i=1}^{d} [\partial_{ii}^2 f(x) + x_i \partial_i f(x)].$$

Une conséquence de cette écriture est que, multiplicité exceptée, on a l'égalité des spectres $\sigma(-\mathbf{L}_{OU}) = \sigma(-\mathbf{L}_{OU,1})$. Plus précisément, on a la description suivante des valeurs propres.

Proposition 3.1.8. *Le spectre de $-\mathbf{L}_{OU}$ est discret, et $\sigma(-\mathbf{L}_{OU}) = \mathbf{N}$. De plus, pour tout $n \in \mathbf{N}$,*

$$\dim(E_n) = \binom{d-1+n}{n},$$

et une base de E_n est donnée par les polynômes d'Hermite d-dimensionnels de degré n.

La dernière assertion de la proposition ci-dessus nécessite préalablement une description des polynômes en question. On rappelle qu'un polynôme homogène de degré $n \in \mathbf{N}$ à d indéterminées s'écrit comme une somme de monômes de la forme

$$M_\alpha = X_1^{\alpha_1} \cdots X_d^{\alpha_d},$$

où $\alpha \in \mathbf{N}^d$, avec $|\alpha| = n$. On peut définir une action de l'ensemble des polynômes homogènes de degré n à d indéterminées sur la famille $h = (h_n)_{n \in \mathbf{N}}$ de polynômes d'Hermite de la façon suivante.

Définition 3.1.9 (Polynômes d'Hermite multidimensionnels)**.** Soient $n \in \mathbf{N}$ et $\alpha \in \mathbf{N}^d$ tel que $|\alpha| = n$. On définit M_α le monôme homogène de degré n à d indéterminées :

$$M_\alpha = X_1^{\alpha_1} \cdots X_d^{\alpha_d}.$$

Le *polynôme d'Hermite de degré n à d variables* associé à M_α est donné par l'action des monômes homogènes sur la familles des polynômes d'Hermite unidimensionnels $h = (h_n)_{n \in \mathbf{N}}$ au sens suivant : pour tout $x \in \mathbf{R}^d$,

$$M_\alpha(h)(x) = \prod_{i=1}^{d} h_{\alpha_i}(x_i).$$

On notera $(\tilde{h}_n)_{n \geq 0}$ la famille de polynômes ainsi définie.

Cette structure en variables séparées fait écho à celle de la loi gaussienne elle-même et du générateur $-\mathbf{L}_{OU}$. En particulier, les polynômes multidimensionnels ainsi définis sont des fonctions propres pour cet opérateur. De plus, le théorème de Fubini permet, via les fonctions partielles, de se ramener à la dimension un pour montrer que la famille $(\tilde{h}_n)_{n\geq 0}$ forme une base hilbertienne de $L^2(\gamma_d)$.

Proposition 3.1.10. *La famille $(\tilde{h}_n)_{n\geq 0}$ forme une base hilbertienne de $L^2(\gamma_d)$ de vecteurs propres pour $-\mathbf{L}_{OU}$.*

Ce résultat permettra, par un raisonnement similaire à celui de la dimension un, de prouver la Proposition 3.1.8.

Démonstration. Cette proposition est en deux parties, on commence par montrer que $(\tilde{h}_n)_{n\geq 0}$ est une famille orthonormale complète de $L^2(\gamma_d)$. L'orthonormalité est essentiellement un jeu d'écriture utilisant la propriété analogue en dimension un. Soient $\alpha, \beta \in \mathbf{N}^d$ deux multi-indices quelconques. On a avec le théorème de Fubini, par définition de γ_d,

$$\begin{aligned}\langle M_\alpha(h), M_\beta(h)\rangle &= \int_{\mathbf{R}^d} \prod_{i=1}^d h_{\alpha_i}(x_i) \prod_{i=1}^d h_{\beta_i}(x_i) d\gamma_d(x) \\ &= \prod_{i=1}^d \int_{\mathbf{R}^d} h_{\alpha_i}(x_i) h_{\beta_i}(x_i) d\gamma_1(x_i) \\ &= \prod_{i=1}^d \delta_{\alpha_i\beta_i},\end{aligned}$$

par orthonormalité de h_{α_i} et h_{β_i} dans $L^2(\gamma_1)$. Il s'ensuit donc :

$$\langle M_\alpha(h), M_\beta(h)\rangle = \delta_{\alpha\beta},$$

et donc l'orthonormalité de $(\tilde{h}_n)_{n\geq 0}$. La complétude vient elle aussi de la propriété analogue en dimension un. Soit $f \in L^2(\gamma_d)$ telle que, pour tout $\alpha \in \mathbf{N}^d$, $\langle M_\alpha(h), f\rangle = 0$. La preuve peut se faire par récurrence descendante. Pour tout $\alpha \in \mathbf{N}^d$, on a par le théorème de Fubini :

$$0 = \int_{\mathbf{R}^d} M_\alpha(h) f d\gamma = \int_{\mathbf{R}} h_{\alpha_1}(x_1) \int_{\mathbf{R}^{d-1}} \prod_{i=2}^d h_{\alpha_i}(x_i) f(x) d\gamma_{d-1}(x_2, \ldots, x_d) d\gamma_1(x_1).$$

Cette égalité est vraie quelle que soit la valeur de $\alpha_1 \in \mathbf{N}$, donc l'intégrale de droite (contre γ_{d-1}) est nulle pour presque tout $x_1 \in \mathbf{R}$. On peut réitérer l'opération avec x_2, x_3, etc., et il s'ensuit par récurrence que $f(x) = 0$ presque partout. La famille $(\tilde{h}_n)_{n\geq 0}$ forme donc bien une base hilbertienne de $L^2(\gamma_d)$.

Reste à montrer que ce sont des fonctions propres pour $-\mathbf{L}_{OU}$. En utilisant la décomposition en opérateurs unidimensionnels rappelée ci-dessus, on a pour tout $\alpha \in \mathbf{N}^d$:

$$\begin{aligned}
-\mathbf{L}_{OU} M_\alpha(h)(x) &= \sum_{i=1}^{d} -\mathbf{L}_{OU,1} \left(\prod_{j=1}^{d} h_{\alpha_j}(x_j) \right)_i (x_i) \\
&= \sum_{i=1}^{d} -\mathbf{L}_{OU,1}(h_{\alpha_i})(x_i) \prod_{\substack{j=1 \\ j\neq i}}^{d} h_{\alpha_j}(x_j) \\
&= \sum_{i=1}^{d} \alpha_i h_{\alpha_i}(x_i) \prod_{\substack{j=1 \\ j\neq i}}^{d} h_{\alpha_j}(x_j) \\
&= \sum_{i=1}^{d} \alpha_i \prod_{j=1}^{d} h_{\alpha_j}(x_j) = \left(\sum_{i=1}^{d} \alpha_i \right) M_\alpha(h)(x).
\end{aligned}$$

On en déduit donc que $M_\alpha(h)$ est une fonction propre de $-\mathbf{L}_{OU}$ associée à la valeur propre $\sum_{i=1}^{d} \alpha_i$.

■

Ce résultat permet de donner la preuve de la Proposition 3.1.8, qui décrit complètement le spectre de $-\mathbf{L}_{OU}$.

Proposition 3.1.8. Comme dans le cas du Théorème 3.1.6, la décomposition spectrale de $L^2(\gamma_d)$ implique que $\sigma_{ess}(-\mathbf{L}_{OU}) = \emptyset$, et donc que $\sigma(-\mathbf{L}_{OU}) = \mathbf{N}$.

La complétude du système $(\tilde{h}_n)_{n\geq 0}$ implique de même que l'espace propre E_n associé à la valeur propre $n \in \mathbf{N}$ est engendré par l'ensemble des polynômes de la forme $M_\alpha(h)$ où $|\alpha| = n$. Ainsi, la dimension de E_n est égale au nombre de polynômes homogènes de degré n à d indéterminées, ou encore au nombre de multi-indices $\alpha \in \mathbf{N}^d$ tels que $\alpha_1 + \cdots + \alpha_d = n$. Un dénombrement élégant de cet ensemble peut être réalisé avec un codage par étoiles et barres. On représente un tel multi-indice par une suite de n étoiles et $d-1$ barres, toutes indiscernables, où les étoiles sont groupées d paquets, chacun séparé du suivant et du précédent par une barre, et dont le i-ème comprend α_i étoiles :

$$\underbrace{******}_{\alpha_1} | \underbrace{****}_{\alpha_2} | \cdots | \underbrace{**}_{\alpha_d}.$$

Il y a bien bijection entre ces codes et les multi-indices d'intérêt. Ces codes sont faciles à dénombrer : il y en a autant que de permutations de ces $n+d-1$ éléments ($(n+d-1)!$), divisé par le nombre de permutations des étoiles ($n!$) et le nombre de permutations des barres ($(d-1)!$), car les étoiles et les barres sont indiscernables. Il en résulte que

$$\dim(E_n) = \frac{(d+n-1)!}{n!(d-1)!} = \binom{d+n-1}{n},$$

ce qui est la formule attendue. ■

Remarque. La technique de dénombrement par les étoiles et les barres valait bien la démonstration !

Les deux propositions précédentes donnent une description exhaustive de $\sigma(-\mathbf{L}_{OU})$. Comme mentionné plus haut, les cas aussi explicites sont très rares, et il est en pratique très difficile d'avoir accès à la description des fonctions propres. Dans ce qui suit, on se concentrera sur deux points principaux : d'une part une estimation des valeurs propres et d'autre part une estimation de la dimension des espaces propres. Certains cas, comme la dimension des premiers espaces propres, sont explicites, bien que les fonctions propres ne le soient pas. Ces cas font l'objet de techniques particulières, notamment dues à Klartag [54] et reprises dans [14].

Avant de développer les outils spécifiques liés à ce problème, on rappelle deux points importants qui constituent le point de départ de la réflexion. Dans tout ce qui suit, on revient à l'opérateur de diffusion $\mathbf{L}$ général.

3.1.3 Trou spectral et caractérisation de Courant-Fischer

Les estimations spectrales auxquelles nous nous intéressons ici portent essentiellement sur les valeurs propres isolées de multiplicité finie de $-\mathbf{L}$, soit son spectre discret. Les questions qui se posent sont, dans l'ordre : l'opérateur $-\mathbf{L}$ admet-il des valeurs propres isolées de multiplicité finie ? Le spectre essentiel est-il vide ? Peut-on estimer les valeurs propres ? Dans cette section sont explorés trois outils qui répondent chacun en partie à ces questions et servent de point d'ancrage pour le développement suivant.

Trou spectral

On a déjà vu dans l'introduction un premier résultat d'analyse spectrale. Dans la Proposition 1.2.3, il a été établi que l'inégalité de Poincaré est équivalente à l'existence d'un trou spectral. Outre une première description de la structure du spectre, cela indique que 0 est une valeur propre isolée dans le spectre. De plus, les deux hypothèses **(H2)** et **(H5)** assurent que $E_0(-\mathbf{L}) \simeq \mathbf{R}$, et donc que 0 est de multiplicité finie.

Il s'ensuit que l'inégalité de Poincaré implique que le spectre discret de $-\mathbf{L}$ est non vide, et que $0 \in \sigma_{disc}(-\mathbf{L})$. Cela ne préfigure rien sur le reste du spectre, comme il sera montré dans les deux exemples suivants.

Exemple (Le cas gaussien)**.** Il n'est nul besoin d'aller bien loin pour trouver un premier exemple simple. Dans le cas de la gaussienne, les résultats vus en introduction permettent d'établir un trou spectral de taille 1 pour la gaussienne. Cela suffit à affirmer que $0 \in \sigma_{disc}(-\mathbf{L}_{OU})$. Mais comme cela a été vu, on peut obtenir une description exhaustive du spectre, qui va bien au-delà de la première valeur propre, information qui n'est pas disponible avec la seule inégalité de Poincaré.

Exemple (Le cas exponentiel unidimensionnel)**.** Bien qu'en-dehors des hypothèses techniques formulées en introduction, le générateur associé à la loi de Laplace μ_L présente un cas très particulier de description du spectre. Pour rappel, cette mesure admet

$x \mapsto e^{-|x|}/2$ comme densité par rapport à la mesure de Lebesgue sur $\mathbf{R}$ (dont le potentiel peut être remplacé par une version régulière de la norme à l'origine avec le même comportement asymptotique), de générateur donné, pour $f \in L^2(\mu_L)$ et $x \in \mathbf{R}$, par :

$$\mathbf{L}f(x) = f''(x) - \text{sign}(x)f'(x).$$

On peut facilement voir que cette mesure satisfait une inégalité de Poincaré $IP(4)$, et par conséquent admet un trou spectral de largeur $1/4$. On a donc bien $0 \in \sigma_{disc}(-\mathbf{L})$. Mais $\lambda = 1/4$ n'est pas une valeur propre, et constitue la base du spectre essentiel. On peut le voir en considérant, pour $\lambda \geq 1/4$, une solution sur $\mathbf{R}_+$ de l'équation différentielle ordinaire :

$$f'' - f' + \lambda f = 0,$$

donnée par la forme générale :

$$f(x) = e^{x/2}\left[A\cos\left(x\sqrt{4\lambda - 1}\right) + B\sin\left(x\sqrt{4\lambda - 1}\right)\right],$$

pour tout $x \in \mathbf{R}_+$, avec $A, B \in \mathbf{R}$ fixés. Sauf dans le cas particulier $A = B = 0$, une fonction de ce type n'est pas de carré intégrable contre la mesure exponentielle. On en déduit que le problème aux valeurs propres n'admet pas de solution non triviale, et donc que $\lambda \geq 1/4$ ne peut être valeur propre. Il s'ensuit donc que $\sigma_{disc}(-\mathbf{L})$ est réduit à 0.

On remarque de plus que ce phénomène n'est pas lié à la régularité du potentiel en 0, mais bien à sa croissance linéaire à l'infini.

On voit, à la lumière de ces exemples, que l'inégalité de Poincaré permet uniquement de montrer que le spectre discret n'est pas vide. Elle ne dit rien sur d'autres éventuelles valeurs propres ou sur le spectre essentiel. Pour décrire ce dernier, de nombreuses techniques existent, nous nous concentrerons sur l'une d'entre elles liée à une condition de Lyapunov.

Spectre essentiel

La caractérisation abordée dans le paragraphe suivant permettra de fournir des estimations assez fines des valeurs propres isolées. En revanche, elle ne permet d'accéder qu'à la base du spectre essentiel. Des caractérisations plus adaptée existent néanmoins, et on peut citer par exemple le critère de Weyl (voir [76], chapitre XIII), qui s'énonce comme suit.

Proposition 3.1.11 (Weyl)**.** *Un réel $\lambda \in \mathbf{R}$ appartient à $\sigma_{ess}(-\mathbf{L})$ si et seulement s'il existe une suite orthonormale non nulle $(f_n)_{n\in\mathbf{N}}$ telle que*

$$\lim_{n\to+\infty} \|\mathbf{L}f_n - \lambda f_n\|_2 = 0.$$

Une telle suite est parfois appelée suite singulière.

Ce résultat peut être utilisé pour montrer que le spectre essentiel de l'opérateur d'Ornstein-Uhlenbeck est vide. Dans un cadre différent, il permet également de montrer que le spectre discret du laplacien sur $L^2(dx)$ est vide.

On ne prouvera pas ce résultat, mentionné à titre indicatif. Dans l'idée de se concentrer d'avantage sur le spectre discret, on s'intéressera d'avantage aux critères suivants, qui permettent d'établir simplement la vacuité du spectre essentiel.

Proposition 3.1.12. *Si*

$$\lim_{|x|\to+\infty} \left[-\Delta W(x) + \frac{1}{2}|\nabla W(x)|^2\right] = +\infty,$$

alors le spectre de $-\mathbf{L}$ *est discret (i.e.* $\sigma_{ess}(-\mathbf{L}) = \emptyset$*).*

La preuve de ce résultat est un peu trop longue pour être mentionnée ici, mais pourra être consultée dans [8] (§4.10). L'idée de la preuve est de montrer, en définissant les opérateurs de Persson, que le bas du spectre essentiel est la limite d'une suite divergeant vers l'infini. Une conséquence intéressante de ce critère est qu'il permet d'établir que, pour $V(x) = |x|^\alpha$, l'opérateur $-\mathbf{L}$ admet un spectre discret dès que $\alpha > 1$ (pour $\alpha = 1$, on a déjà vu que le spectre discret est réduit à $\{0\}$, résultat qui se transpose à l'identique en dimension supérieure).

Le second critère permet de mettre en relation l'inégalité de Sobolev logarithmique vue au chapitre précédent et le spectre essentiel.

Proposition 3.1.13. *Supposons que* μ *satisfait une* $ISL(c)$, $c > 0$. *Alors* $\sigma_{ess}(-\mathbf{L}) = \emptyset$.

Ce résultat est dû à Wang [86], établi dans le cadre plus général des inégalités de F-Sobolev. Notons que le premier critère peut être retrouvé comme un corollaire de ce dernier résultat. Il permet en outre d'établir que, dans notre contexte particulier, la condition $CD(\rho, \infty)$ implique la vacuité du spectre essentiel. Notons enfin qu'en général, cette dernière n'est pas suffisante pour assurer que $\sigma_{ess}(-\mathbf{L}) = \emptyset$, et nécessite en réalité une condition $CD(\rho, n)$, $n < +\infty$, en lien avec une inégalité de Sobolev.

Nous concluons finalement cette section introductive à l'analyse spectrale par la caractérisation de Courant-Fischer pour les valeurs propres.

Caractérisation de Courant-Fischer

Ce résultat bien connu d'analyse spectrale se base sur un principe de min-max (ou max-min) pour les quotients de Rayleigh associés à $-\mathbf{L}$. On rappelle que dans tout ce qui suit, l'opérateur $-\mathbf{L}$ est (essentiellement) auto-adjoint et semi-défini positif.

Définition 3.1.14 (Quotient de Rayleigh)**.** Soit $f \in \mathbf{L}^2(\mu)$ non identiquement nulle. On définit le *quotient de Rayleigh* associé à f par :

$$R(f) = \frac{\int_{\mathbf{R}^d} f(-\mathbf{L}f)\, d\mu}{\int_{\mathbf{R}^d} f^2 d\mu}.$$

Cette notion, héritée de la physique des systèmes vibratoires [74], est intimement liée aux valeurs propres de $-\mathbf{L}$. En effet, si f est une fonction propre associée à la valeur propre λ, il vient immédiatement que $R(f) = \lambda$. Ce quotient mesure, d'une certaine façon, le "défaut de propreté" d'une fonction donnée.

En général, les fonctions propres ne sont pas connues, d'autant plus dans un contexte d'estimation du spectre. Le quotient de Rayleigh seul ne permet donc pas d'estimer efficacement les valeurs propres. À cette fin, on utilise la caractérisation de Courant-Fischer, qui fournit une approche variationnelle de l'estimation de ces dernières (voir [38, 27] pour l'aspect historique, et par exemple [76] pour une présentation moderne).

Proposition 3.1.15 (Courant-Fischer)**.** *Soit $n \in \mathbf{N}$. La n-ième valeur propre de $-\mathbf{L}$, notée $\lambda_n(-\mathbf{L})$, satisfait :*

$$\lambda_n(-\mathbf{L}) = \inf_{\psi_1,\dots,\psi_n \in L^2(\mu)} \sup\{R(f),\ f \in \mathrm{Vect}(\psi_1,\dots,\psi_n)\}.$$

Le sup *et le* inf *peuvent être permutés pour obtenir la formulation duale :*

$$\lambda_n(-\mathbf{L}) = \sup_{\psi_1,\dots,\psi_{n-1} \in L^2(\mu)} \inf\{R(f),\ f \perp \psi_1,\dots,\psi_{n-1}\}.$$

Notons enfin que si $-\mathbf{L}$ n'admet que N valeurs propres, la caractérisation reste vraie pour tout $n \in \mathbf{N}$, avec $\lambda_n(-\mathbf{L}) = \inf \sigma_{ess}(-\mathbf{L})$ pour tout $n > N$.

Dans la suite, nous considérerons uniquement la seconde formulation. En effet, la condition $CD(\rho,\infty)$ implique que le spectre est discret et donc que la caractérisation de Courant-Fischer permet une description exhaustive de ce dernier. De plus, la décomposition spectrale déjà mentionnée rend la seconde écriture d'autant plus pertinente qu'une structure en sous-espaces orthogonaux est donnée par les espaces propres.

On remarque par ailleurs que cette structure permet de raffiner la caractérisation. En effet, on peut montrer que le supremum de la seconde écriture est en fait atteint quand les fonctions $\psi_1,\dots,\psi_{n-1}$ sont les $n-1$ premières fonctions propres. Cette observation ne permet en général pas une estimation plus précise des valeurs propres, mais s'avère très pertinente dans la recherche des multiplicités.

Maintenant les principales base d'analyse spectrales posées, le développement suivant consistera en deux parties : d'abord l'établissement d'une structure algébrique et l'introduction d'une nouvelle relation d'entrelacement, puis les estimations spectrales proprement dites.

3.2 Contexte algébrique

L'idée centrale dans l'estimation des valeurs propres à l'aide de la caractérisation de Courant-Fischer est de faire apparaître une relation d'entrelacement bien choisie en dérivant de façon adéquate à l'intérieur des quotients de Rayleigh. Un première dérivation fait apparaître le gradient des fonctions, une deuxième la hessienne, mais toute dérivation ultérieure nécessite la définition d'un objet de dimension supérieure. Le cadre

de vie algébrique naturel de tels objets est l'algèbre tensorielle, dont les principes de base sont rappelés dans la première partie de la section. Les opérateurs de diffusion sur cette algèbre sont ensuite décrits, avant l'introduction d'une nouvelle famille d'opérateurs qui permet l'écriture de la relation d'entrelacement précitée.

3.2.1 Algèbre tensorielle

Les définitions de cette partie pourront être trouvées dans tout cours de master d'algèbre générale, dont de très bons sont disponibles sur internet, tant en anglais qu'en français. La page Wikipédia sur le sujet pourra être avantageusement consultée sans vergogne.

Les espaces vectoriels seront tous définis sur $\mathbf{R}$, mais la plupart des notions peuvent être étendues à des corps variés (y compris de caractéristique non nulle).

Définition 3.2.1. Soit V un $\mathbf{R}$-espace vectoriel. Pour tout $n \geq 0$, on note $T^n(V) = V^{\otimes n}$ la n-ième *puissance tensorielle* de V. L'*algèbre tensorielle* sur V, notée $T^\bullet(V)$, est l'algèbre graduée définie par :

$$T^\bullet(V) = \bigoplus_{n \geq 0} T^n(V).$$

Par convention, $T^0(V) = \mathbf{R}$.

Dans la suite, on s'intéressera au cas euclidien $V = \mathbf{R}^d$. Pour une fonction $f : \mathbf{R}^d \to T^n(\mathbf{R}^d)$, on peut écrire :

$$f = \sum_{i_1,\dots,i_n=1}^{d} f_{i_1 \cdots i_n} e_{i_1} \otimes \cdots \otimes e_{i_n},$$

où chaque $f_{i_1,\cdots i_n}$ est une fonction à valeurs réelles et $(e_i)_{i=1}^d$ désigne la base canonique de $\mathbf{R}^d$. Cette écriture permet de décrire simplement la régularité de telles fonctions. On dira ainsi qu'une fonction $f : \mathbf{R}^d \to T^n(\mathbf{R}^d)$ définie comme ci-dessus est de classe $\mathcal{C}^r$, pour $r \in \mathbf{N}$, si chacune des fonctions $f_{i_1 \cdots i_n}$ est elle-même de classe $\mathcal{C}^r$. Il s'ensuit naturellement la définition suivante de différentielle, proche dans un certain sens de la dérivée extérieure.

Définition 3.2.2. Pour tout $n \geq 0$ et $f \in \mathcal{C}^1(\mathbf{R}^d, T^n(\mathbf{R}^d))$, on définit la *dérivée tensorielle* de f par :

$$\mathrm{d}f = \sum_{i_1,\dots,i_n,k=1}^{d} \partial_k f_{i_1 \cdots i_n} e_k \otimes e_{i_1} \otimes \cdots \otimes e_{i_n}.$$

On a alors $\mathrm{d}f \in \mathcal{C}^0(\mathbf{R}^d, T^{n+1}(\mathbf{R}^d))$. On définit de même la dérivée r-ième de f, pour $r \in \mathbf{N}^*$, comme $\mathrm{d}^{(r)} f = \underbrace{\mathrm{d} \circ \cdots \circ \mathrm{d}}_{r} f$ (sous l'hypothèse que f est de classe $\mathcal{C}^r$).

Pour $n = 0$, on retrouve le gradient et la hessienne de f respectivement pour $r = 1$ et $r = 2$. Pour $n = 1$ (quand f est un champ de vecteurs), $\mathrm{d}f$ peut être identifié à la jacobienne de f (à transposition près selon la convention choisie pour l'écriture de cette dernière).

L'espace des fonctions à valeurs dans $T^n(\mathbf{R}^d)$ peut être muni d'un produit scalaire compatible avec des estimations spectrales. Pour $f, g : \mathbf{R}^d \to T^n(\mathbf{R}^d)$, on définit le produit scalaire de Hilbert-Schmidt suivant :

$$\langle f, g\rangle_n = \sum_{i_1,\dots,i_n=1}^{d} \int_{\mathbf{R}^d} f_{i_1,\cdots i_n} g_{i_1,\cdots i_n} d\mu,$$

ainsi que $L^2(\mathbf{R}^d, T^n(\mathbf{R}^d))$ l'espace L^2 associé (dans un souci de concision, on ne fera apparaître la mesure μ que dans $L^2(\mu)$, pour les fonctions à valeurs réelles). La norme induite est notée $\|\cdot\|_n$. Pour $n = 0$, le produit ci-dessus correspond au produit scalaire classique des fonctions L^2.

Dans la suite, on considérera la différentielle de fonctions L^2, soit au sens des distributions. De plus, on définit la suite suivante :

$$L^2(\mathbf{R}^d, \mathbf{R}) \xrightarrow{\mathrm{d}_0} L^2(\mathbf{R}^d, \mathbf{R}^d) \to \cdots \to L^2(\mathbf{R}^d, T^n(\mathbf{R}^d)) \xrightarrow{\mathrm{d}_n} L^2(\mathbf{R}^d, T^{n+1}(\mathbf{R}^d)) \to \cdots$$

où d_n désigne la dérivée tensorielle (à l'ordre 1) des fonctions à valeurs dans $T^n(\mathbf{R}^d)$.

Enfin, on note $\mathrm{d}_n^\star$ l'opérateur de divergence, défini comme l'adjoint de d_n dans $L^2(\mathbf{R}^d, T^n(\mathbf{R}^d))$: pour toutes fonctions $f \in L^2(\mathbf{R}^d, T^n(\mathbf{R}^d))$ et $g \in L^2(\mathbf{R}^d, T^{n+1}(\mathbf{R}^d))$, on a

$$\langle f, \mathrm{d}_n^\star g\rangle_n = \langle \mathrm{d}_n f, g\rangle_{n+1}.$$

Cet opérateur est explicite, et un calcul direct donne la formule suivante pour $\mathrm{d}_n^\star$: pour toute fonction $g \in L^2(\mathbf{R}^d, T^{n+1}(\mathbf{R}^d))$,

$$\mathrm{d}_n^\star g = \sum_{i_1,\dots,i_n=1}^{d} \left[\sum_{k=1}^{d} (-\partial_k + \partial_k V) g_{k,i_1,\cdots i_n}\right] e_{i_1} \otimes \cdots \otimes e_{i_n}.$$

Dans le cas $n = 0$ et $V \equiv 0$ (cas de la mesure de Lebesgue), on retrouve la définition classique de la divergence des champs de vecteurs.

Ces définitions ressemblent à celles de l'algèbre extérieure et de la dérivée extérieure. Mais le cadre tensoriel n'admet pas les symétries de cette dernière, et on prendra garde à ne pas confondre les deux différentielles. De même, la suite introduite ci-dessus n'est pas une suite exacte, et a fortiori pas un complexe de cochaînes (comme c'est le cas avec le complexe de De Rham, par exemple). La relative absence de structure permet néanmoins d'itérer la différentielle autant de fois que nécessaire sur des espaces "arbitrairement gros" (alors que toute algèbre extérieure de puissance strictement supérieur à la dimension de l'espace sous-jacent est triviale). Ce point sera particulièrement important dans la recherche des multiplicités des valeurs propres.

3.2.2 Opérateurs de diffusion et Schrödinger sur l'algèbre tensorielle et estimations préliminaires

Le formalisme introduit ci-dessus permet d'étendre la définition de $\mathbf{L}$ aux fonctions à valeurs tensorielles. Cette définition découle de celle de Johnsen dans [52] et d'autres extensions de ces opérateurs aux champs de vecteurs, voir à ce titre le chapitre précédent et les référence afférentes. Les questions de domaines ayant été amplement discutées en introduction, on se restreindra ici aux fonctions infiniment dérivables.

Définition 3.2.3. Soient $n \in \mathbf{N}$ et $f \in \mathcal{C}^\infty(\mathbf{R}^d, T^n(\mathbf{R}^d))$. On définit l'opérateur $\mathcal{L}_n$ par :

$$\mathcal{L}_n f = \sum_{i_1,\dots,i_n=1}^{d} \mathbf{L} f_{i_1,\dots,i_n} e_{i_1} \otimes \cdots \otimes e_{i_n}.$$

En utilisant d_0, on peut réécrire $\mathbf{L}$ comme suit :

$$-\mathbf{L} f = -\Delta f + \mathrm{d}_0 V \cdot \mathrm{d}_0 f.$$

L'opérateur $\mathcal{L}$ du chapitre précédent correspond ici à $\mathcal{L}_1$, et $\mathcal{L}_0 = \mathbf{L}$.

Remarque. Cette définition rend compte de l'action "diagonale" de $\mathcal{L}_n$ telle que décrite en introduction (voir la Définition 1.3.3) et dans le chapitre précédent. En effet, la suite d'opérateurs $(\mathcal{L}_n)_{n\geq 0}$ admet la description récursive suivante : pour tout $n \geq 1$, $\mathcal{L}_n = \mathcal{L}_{n-1} \otimes I_d$.

Plus généralement, et suivant la démarche de [52], chaque opérateur $\mathcal{L}_n$ peut être exprimé en fonction de d_n et $\mathrm{d}_n^\star$. On retrouve de plus une relation d'entrelacement similaire à celle de la Proposition 1.3.6, pour laquelle un nouvel opérateur multiplicatif basé sur la hessienne de V est défini.

Proposition 3.2.4. *Pour tout $n \in \mathbf{N}$,*

$$-\mathcal{L}_n = \mathrm{d}_n^\star \circ \mathrm{d}_n \ \text{ et } \ (-\mathcal{L}_{n+1} + \nabla^2_{n+1} V)_{|\mathrm{Im}(\mathrm{d}_n)} = (\mathrm{d}_n \circ \mathrm{d}_n^\star)_{|\mathrm{Im}(\mathrm{d}_n)},$$

où $\nabla^2_{n+1} V$ est un opérateur multiplicatif défini sur les fonctions à valeurs dans $T^{n+1}(\mathbf{R}^d)$ par

$$\nabla^2_{n+1} V f = \sum_{i_1,\dots,i_{n+1}=1}^{d} \left(\nabla^2 V f_{\star,i_2,\dots,i_{n+1}}\right)_{i_1} e_{i_1} \otimes \cdots \otimes e_{i_{n+1}},$$

et

$$f_{\star,i_2,\dots,i_{n+1}} := \sum_{j=1}^{d} f_{j,i_2,\dots,i_{n+1}} e_j,$$

pour tous $i_2, \cdots, i_{n+1} \in [\![1, d]\!]$ fixés et toute fonction $f : \mathbf{R}^d \to T^{n+1}(\mathbf{R}^d)$.

Cette écriture peut faire penser aux laplaciens de Witten définis sur les formes différentielles, tels que décrits par exemple dans [52]. Les objets sont néanmoins ici de nature différente, étant définis à partir la dérivée tensorielle et non la dérivée extérieure.

Démonstration. La preuve consiste en un calcul direct. Pour le premier point, on considère $f \in \mathcal{C}^\infty(\mathbf{R}^d, T^n(\mathbf{R}^d))$. Par définition de d_n, on a :

$$\begin{aligned}
\mathrm{d}_n^\star \circ \mathrm{d}_n f &= \sum_{i_1,\dots,i_n=1}^{d} \left[\sum_{k=1}^{d} (-\partial_k + \partial_k V)(\mathrm{d}_n f)_{k,i_1,\dots,i_n} \right] e_{i_1} \otimes \cdots \otimes e_{i_n} \\
&= \sum_{i_1,\dots,i_n=1}^{d} \left[\sum_{k=1}^{d} (-\partial_k + \partial_k V)\partial_k f_{i_1,\dots,i_n} \right] e_{i_1} \otimes \cdots \otimes e_{i_n} \\
&= \sum_{i_1,\dots,i_n=1}^{d} \left[-\Delta f_{i_1,\dots,i_n} + \mathrm{d}_0 V \cdot \mathrm{d}_0 f_{i_1,\dots,i_n} \right] e_{i_1} \otimes \cdots \otimes e_{i_n} \\
&= \sum_{i_1,\dots,i_n=1}^{d} \mathbf{L} f_{i_1,\dots,i_n} e_{i_1} \otimes \cdots \otimes e_{i_n} = -\mathcal{L}_n f.
\end{aligned}$$

De même pour le second point, on considère $F \in \mathrm{Im}(\mathrm{d}_n)$. Il existe alors $f \in \mathcal{C}^\infty(\mathbf{R}^d, T^n(\mathbf{R}^d))$ telle que $F = \mathrm{d}_n f$. Par conséquent, il vient :

$$\begin{aligned}
\mathrm{d}_n \circ \mathrm{d}_n^\star F = \mathrm{d}_n \circ \mathrm{d}_n^\star \circ \mathrm{d}_n f &= -\mathrm{d}_n \circ \mathcal{L}_n f \\
&= - \sum_{k,i_1,\dots,i_n=1}^{d} \partial_k (\mathbf{L} f_{i_1,\dots,i_n}) e_k \otimes e_{i_1} \otimes \cdots \otimes e_{i_n} \\
&= - \sum_{k,i_1,\dots,i_n=1}^{d} (\mathcal{L}_1 \mathrm{d}_0 f_{i_1,\dots,i_n} - \nabla^2 V \mathrm{d}_0 f_{i_1,\dots,i_n})_k e_k \otimes e_{i_1} \otimes \cdots \otimes e_{i_n},
\end{aligned}$$

en utilisant l'entrelacement général (voir la Proposition 1.3.6). L'argument de la somme peut être scindé en deux, et on a d'une part, car $\mathcal{L}_n$ est diagonal,

$$\begin{aligned}
- \sum_{k,i_1,\dots,i_n=1}^{d} &(\mathcal{L}_1 \mathrm{d}_0 f_{i_1,\dots,i_n})_k e_k \otimes e_{i_1} \otimes \cdots \otimes e_{i_n} \\
&= - \sum_{k,i_1,\dots,i_n=1}^{d} \mathbf{L} \left(\partial_k f_{i_1,\dots,i_n} \right) e_k \otimes e_{i_1} \otimes \cdots \otimes e_{i_n} \\
&= -\mathcal{L}_{n+1}(\mathrm{d}_n f).
\end{aligned}$$

D'autre part, par définition de d_n,

$$\begin{aligned}
\sum_{k,i_1,\dots,i_n=1}^{d} &(\nabla^2 V \mathrm{d}_0 f_{i_1,\dots,i_n})_k e_k \otimes e_{i_1} \otimes \cdots \otimes e_{i_n} \\
&= \sum_{k,i_1,\dots,i_n=1}^{d} [\nabla^2 V (\mathrm{d}_n f)_{\star,i_1,\dots,i_n}]_k e_k \otimes e_{i_1} \otimes \cdots \otimes e_{i_n} \\
&= \nabla^2_{n+1} V(\mathrm{d}_n f).
\end{aligned}$$

Finalement,

$$\mathrm{d}_n \circ \mathrm{d}_n^\star F = -(\mathcal{L}_{n+1} + \nabla_{n+1}^2 V)F,$$

ce qui termine la preuve. ■

La restriction à $\mathrm{Im}(\mathrm{d}_n)$ dans la deuxième identité met en avant la relation d'entrelacement, que cette proposition permet d'écrire de façon synthétique :

$$\mathrm{d}_n(-\mathcal{L}_n) = \mathrm{d}_n(\mathrm{d}_n^\star \circ \mathrm{d}_n) = (\mathrm{d}_n \circ \mathrm{d}_n^\star)\mathrm{d}_n = (-\mathcal{L}_{n+1} + \nabla_{n+1}^2 V)\mathrm{d}_n.$$

Notons que cette propriété implique directement la symétrie de $\mathcal{L}_n$ pour tout $n \in \mathbf{N}$.

Maintenant que les opérateurs ont été clairement définis, de premières estimations spectrales sur $-\mathcal{L}_n$ et $\nabla_{n+1}^2 V$ peuvent être établies. Celles-ci sont principalement des conséquences des propriétés des opérateurs sous-jacents $-\mathbf{L}$ et $\nabla^2 V$.

Proposition 3.2.5. *Soit $n \in \mathbf{N}$. Supposons que $-\mathbf{L}$ est semi-défini positif et $\nabla^2 V \geq \rho I_d$ pour un certain $\rho > 0$. Alors $-\mathcal{L}_n$ est semi-défini positif et $\nabla_{n+1}^2 V \geq \rho I_d^{\otimes n+1}$ (avec $I_d^{\otimes n+1}$ l'identité de $\mathcal{L}(T^{n+1}(\mathbf{R}^d))$).*

Démonstration. Ces deux points suivent directement de la définition de $\langle \cdot, \cdot \rangle_n$. Soit $f \in \mathcal{C}^\infty(\mathbf{R}^d, T^n(\mathbf{R}^d))$, on a

$$\langle -\mathcal{L}_n f, f \rangle_n = -\sum_{i_1,\dots,i_n=1}^d \int_{\mathbf{R}^d} \mathbf{L} f_{i_1,\dots,i_n} f_{i_1,\dots,i_n} d\mu = \sum_{i_1,\dots,i_n=1}^d \langle -\mathbf{L} f_{i_1,\dots,i_n}, f_{i_1,\dots,i_n} \rangle_0 \geq 0,$$

car $-\mathbf{L}$ est semi-défini positif. L'estimation sur $\nabla_{n+1}^2 V$ revient à montrer que, pour tout $F \in \mathcal{C}^\infty(\mathbf{R}^d, T^{n+1}(\mathbf{R}^d))$,

$$\langle \nabla_{n+1}^2 V F, F \rangle_{n+1} \geq \rho \|F\|_{n+1}^2.$$

Toujours par définition de $\langle \cdot, \cdot \rangle_{n+1}$, on a :

$$\begin{aligned}
\langle \nabla_{n+1}^2 V F, F \rangle_{n+1} &= \sum_{i_1,\dots,i_{n+1}=1}^d \int_{\mathbf{R}^d} \left(\nabla^2 V F_{\star,i_2,\dots,i_{n+1}}\right)_{i_1} F_{i_1,\dots,i_{n+1}} d\mu \\
&= \sum_{i_2,\dots,i_{n+1}=1}^d \int_{\mathbf{R}^d} \sum_{i_1=1}^d \left(\nabla^2 V F_{\star,i_2,\dots,i_{n+1}}\right)_{i_1} F_{i_1,\dots,i_{n+1}} d\mu \\
&= \sum_{i_2,\dots,i_{n+1}=1}^d \int_{\mathbf{R}^d} \langle \nabla^2 V F_{\star,i_2,\dots,i_{n+1}}, F_{\star,i_2,\dots,i_{n+1}} \rangle_1 d\mu \\
&\geq \rho \sum_{i_2,\dots,i_{n+1}=1}^d \int_{\mathbf{R}^d} |F_{\star,i_2,\dots,i_{n+1}}|^2 d\mu,
\end{aligned}$$

par hypothèse sur $\nabla^2 V$. De plus,

$$|F_{\star,i_2,\dots,i_{n+1}}|^2 = \sum_{i_1=1}^d (F_{i_1,i_2,\dots,i_{n+1}})^2,$$

et donc

$$\rho \sum_{i_2,\dots,i_{n+1}=1}^{d} \int_{\mathbf{R}^d} |F_{\star,i_2,\dots,i_{n+1}}|^2 d\mu = \rho \sum_{i_1,\dots,i_{n+1}=1}^{d} \int_{\mathbf{R}^d} (F_{i_1,i_2,\dots,i_{n+1}})^2 d\mu = \rho \|F\|_{n+1}^2,$$

ce qui achève la preuve. ■

Pour aller plus loin dans les estimations spectrales, il est nécessaire de revenir aux quotients de Rayleigh. Les relations d'entrelacement définies plus haut, adjointes de la propriété précédente, le permettent. Elles nécessitent néanmoins l'introduction de la différentielle dans le numérateur, ce qui ne peut pas être fait "gratuitement". Pour pallier à cette défaillance, on introduit une nouvelle famille d'opérateurs différentiels qui ont le bon goût d'être des isométries pour le produit scalaire en jeu : les transformations de Riesz.

3.2.3 Transformations de Riesz

On supposera dans tout ce qui suit que $-\mathbf{L}$ est semi-défini positif et que μ satisfait $CD(\rho,\infty)$. Sous cette dernière hypothèse, le spectre de $-\mathbf{L}$ est discret (voir la Proposition 3.1.13), et sera noté $\{\lambda_i(-\mathbf{L})\}_{i\geq 0}$. De plus, on peut trouver une base hilbertienne de $L^2(\mu)$ de vecteurs propre $(g_i)_{i\geq 0}$ telle que, pour toute fonction $f \in L^2(\mu)$,

$$f = \sum_{i\geq 0} \langle f, g_i\rangle_0 g_i.$$

Ainsi, pour $f \in \mathcal{D}(-\mathbf{L})$,

$$-\mathbf{L}f = \sum_{i\geq 0} \lambda_i(-\mathbf{L})\langle f, g_i\rangle_0 g_i.$$

Notons que les valeurs propres sont comptées avec répétitions (on peut avoir $\lambda_i(-\mathbf{L}) = \lambda_j(-\mathbf{L})$ pour $i \neq j$).

Pour $\alpha \in \mathbf{R}$, on peut définir les puissances $-\mathbf{L}$ comme suit :

$$(-\mathbf{L})^\alpha f = \sum_{i\geq 0} \lambda_i(-\mathbf{L})^\alpha \langle f, h_i\rangle_0 h_i,$$

sous l'hypothèse que $\sum_{i\geq 0} \lambda_i(-\mathbf{L})^{2\alpha}\langle f, h_i\rangle_0^2$ converge. En particulier, comme $-\mathbf{L}$ est semi-défini négatif, cette formule permet de définir $(-\mathbf{L})^{1/2}$ et $(-\mathbf{L})^{-1/2}$ sur les fonctions centrées. Ces définitions se transposent naturellement à l'opérateur tensoriel $-\mathcal{L}_n$ pour tout $n \geq 1$.

Ces puissances d'opérateur permettent de définir les transformations de Riesz mentionnées plus haut. Cette définition est inspirée des travaux de Bakry sur la question (voir [6] et ses précédents articles sur le sujet), bien que le contexte euclidien tensoriel ici présent soit assez différent (en particulier, on ne parlera pas de formes différentielles).

Dans tout ce qui suit, on désignera par "centrée" toute fonction $f : \mathbf{R}^d \to T^n(\mathbf{R}^d)$ telle que

$$\int_{\mathbf{R}^d} f\, d\mu = 0_{T^n(\mathbf{R}^d)},$$

où $0_{T^n(\mathbf{R}^d)}$ est le tenseur nul.

Définition 3.2.6. Soient $n \in \mathbf{N}$ et $f \in \mathcal{C}^\infty(\mathbf{R}^d, T^n(\mathbf{R}^d))$ centrée. On définit la *transformée de Riesz* de f par :

$$\mathcal{U}_n f = \mathrm{d}_n(-\mathcal{L}_n)^{-1/2} f.$$

Comme cela apparaîtra clairement dans la suite, ces transformations apparaissent assez naturellement dans le calcul de certains quotients de Rayleigh. Comme annoncé, ces transformations sont des isométries.

Proposition 3.2.7. *Pour tout $n \in \mathbf{N}$, $\mathcal{U}_n : L^2(\mathbf{R}^d, T^n(\mathbf{R}^d)) \to L^2(\mathbf{R}^d, T^{n+1}(\mathbf{R}^d))$ est une isométrie.*

Démonstration. Soient $f, g \in L^2(\mathbf{R}^d, T^n(\mathbf{R}^d))$ deux fonctions centrées. On prouve que

$$\langle \mathcal{U}_n f, \mathcal{U}_n g \rangle_{n+1} = \langle f, g \rangle_n.$$

Les propriétés de d_n et $\mathrm{d}_n^\star$ donnent :

$$\begin{aligned}
\langle \mathcal{U}_n f, \mathcal{U}_n g \rangle_{n+1} &= \langle \mathrm{d}_n(-\mathcal{L}_n)^{-1/2} f, \mathrm{d}_n(-\mathcal{L}_n)^{-1/2} g \rangle_{n+1} \\
&= \langle (-\mathcal{L}_n)^{-1/2} f, \underbrace{\mathrm{d}_n^\star \mathrm{d}_n}_{-\mathcal{L}_n} (-\mathcal{L}_n)^{-1/2} g \rangle_n \\
&= \langle (-\mathcal{L}_n)^{-1/2} f, (-\mathcal{L}_n)^{1/2} g \rangle_n \\
&= \langle f, g \rangle_n,
\end{aligned}$$

par symétrie $-\mathcal{L}_n$ (et donc de $(-\mathcal{L}_n)^{1/2}$). ∎

La définition des transformations de Riesz et la relation d'entrelacement de la Proposition 3.2.4 laissent attendre l'existence d'une relation similaire pour ces dernières. Elle est établie par la proposition suivante, qui servira de pierre angulaire à la prochaine section.

Proposition 3.2.8. *Pour tout $n \in \mathbf{N}$, on a la relation d'entrelacement suivante :*

$$\mathcal{U}_n(-\mathcal{L}_n) = (-\mathcal{L}_{n+1} + \nabla_{n+1}^2 V)\mathcal{U}_n,$$

attendu que toutes les compositions sont bien définies.

Démonstration. Soit $n \in \mathbf{N}$ fixé. Par définition de la mesure invariante, toute fonction de l'image de $-\mathcal{L}_n$ est centrée, le terme de gauche est donc bien défini sur tout $\mathcal{D}(-\mathcal{L}_n)$ (qui est isomorphe à $\mathcal{D}(-\mathbf{L})^{\otimes n}$). Le terme de droite, en revanche, n'existe que pour les fonctions centrées de ce dernier domaine. Ainsi, sous l'hypothèse que ces restrictions sont respectées, on a :

$$\begin{aligned}
\mathcal{U}_n(-\mathcal{L}_n) &= \mathrm{d}_n(-\mathcal{L}_n)^{-1/2}(-\mathcal{L}_n) \\
&= \mathrm{d}_n(-\mathcal{L}_n)(-\mathcal{L}_n)^{-1/2} \\
&= (-\mathcal{L}_{n+1} + \nabla_{n+1}^2 V)\mathrm{d}_n(-\mathcal{L}_n)^{-1/2} = (-\mathcal{L}_{n+1} + \nabla_{n+1}^2 V)\mathcal{U}_n,
\end{aligned}$$

en utilisant l'entrelacement de la Proposition 3.2.4. ∎

Comme nous le verrons dans la partie suivante, le fait que les transformations de Riesz sont des isométries permet de faire apparaître "gratuitement" les relations d'entrelacement dans les quotients de Rayleigh (ce qui est impossible avec la seule différentielle). Cette propriété justifie l'intérêt de la précédente relation d'entrelacement et rend possible les estimations spectrales annoncées en introduction.

3.3 Estimations spectrales pour $-\mathbf{L}$

Peut-être encore plus que le résultat en lui-même, la technique mise en œuvre présente un fort intérêt. De nature assez générale, elle permet, moyennant de fines estimations sur la dimension de certains noyaux, d'obtenir des bornes optimales sur les valeurs propres de $-\mathbf{L}$. On illustre en particulier cette optimalité dans le cas gaussien, où l'on retrouve les résultats présentés en première partie du chapitre.

Dans toute cette section, on suppose que $-\mathbf{L}$ est semi-défini positif et $\nabla^2 V \geq \rho I_d$ pour un certain $\rho > 0$. Cela implique que le spectre de $-\mathbf{L}$ est discret et positif, et sera toujours noté $\{\lambda_i(-\mathbf{L})\}_{i\geq 0}$ (avec répétitions). On notera enfin, pour tout $n \in \mathbf{N}^*$,

$$k_n := \dim \ker(-\mathcal{L}_n \circ \mathcal{U}_{n-1} \circ \cdots \circ \mathcal{U}_0).$$

Au besoin, on pourra également noter $k_0 = \dim \ker(-\mathbf{L})$ (égal à 1 dans notre contexte).

3.3.1 Résultat principal

Sans plus attendre, le résultat obtenu s'énonce comme suit.

Théorème 3.3.1. *Sous l'hypothèse $\nabla^2 V \geq \rho I_d$ pour $\rho > 0$, on a pour tout $r \in \mathbf{N}^*$,*

$$\lambda_r(-\mathbf{L}) \geq (n+1)\rho,$$

où $n \in \mathbf{N}$ vérifie $\sum_{i=1}^{n} k_i < r \leq \sum_{i=1}^{n+1} k_i$.

Sa preuve nécessite l'utilisation de la quasi-totalité des résultats avancés dans la section précédente. En particulier, la Proposition 3.2.4 y joue un rôle prépondérant. Il faut donc vérifier que les restrictions décrites dans la preuve de cette dernière sont respectées, ce qui est l'objet du lemme suivant.

Lemme 3.3.2. *Soient $n \in \mathbf{N}^*$ et $h \in L^2(\mu)$. Alors $h \in \ker(-\mathcal{L}_n \circ \mathcal{U}_{n-1} \circ \cdots \circ \mathcal{U}_0)^\perp$ si et seulement si $\int_{\mathbf{R}^d} \mathcal{U}_{n-1} \circ \cdots \circ \mathcal{U}_0 h \, d\mu = 0$.*

Démonstration. Soit $h \in L^2(\mu)$.

- Supposons d'abord que $h \in \ker(-\mathcal{L}_n \circ \mathcal{U}_{n-1} \circ \cdots \circ \mathcal{U}_0)^\perp$. Notons d'abord que $\int_{\mathbf{R}^d} \mathcal{U}_{n-1} \circ \cdots \circ \mathcal{U}_0 h \, d\mu = 0$ équivaut à :

$$\langle \mathcal{U}_{n-1} \circ \cdots \circ \mathcal{U}_0 h, C \rangle_n = 0$$

pour tout tenseur constant $C \in T^n(\mathbf{R}^d)$. De plus, si $C \in T^n(\mathbf{R}^d)$, alors $\mathcal{U}_0^\star \circ \cdots \circ \mathcal{U}_{n-1}^\star C \in \ker(-\mathcal{L}_n \circ \mathcal{U}_{n-1} \circ \cdots \circ \mathcal{U}_0)$. En effet, comme chaque $\mathcal{U}_i$ est une isométrie, $\mathcal{U}_i^\star = \mathcal{U}_i^{-1}$ et donc

$$-\mathcal{L}_n \circ \mathcal{U}_{n-1} \circ \cdots \circ \mathcal{U}_0 \circ \mathcal{U}_0^\star \circ \cdots \circ \mathcal{U}_{n-1}^\star C = -\mathcal{L}_n C = 0.$$

Finalement, comme $h \in \ker(-\mathcal{L}_n \circ \mathcal{U}_{n-1} \circ \cdots \circ \mathcal{U}_0)^\perp$,

$$0 = \langle h, \mathcal{U}_0^\star \circ \cdots \circ \mathcal{U}_{n-1}^\star C \rangle_0 = \langle \mathcal{U}_{n-1} \circ \cdots \circ \mathcal{U}_0 h, C \rangle_n,$$

soit $\int_{\mathbf{R}^d} \mathcal{U}_{n-1} \circ \cdots \circ \mathcal{U}_0 h \, d\mu = 0$.

- Supposons à présent que $\int_{\mathbf{R}^d} \mathcal{U}_{n-1} \circ \cdots \circ \mathcal{U}_0 h \, d\mu = 0$. Soit $g \in \ker(-\mathcal{L}_n \circ \mathcal{U}_{n-1} \circ \cdots \circ \mathcal{U}_0)$. L'hypothèse sur h assure que $(-\mathcal{L}_n)^{-1}(\mathcal{U}_{n-1} \circ \cdots \circ \mathcal{U}_0 h)$ est bien défini. Comme chaque $\mathcal{U}_i$ est une isométrie,

$$\begin{aligned}
\langle h, g \rangle_0 &= \langle \mathcal{U}_{n-1} \circ \cdots \circ \mathcal{U}_0 h, \mathcal{U}_{n-1} \circ \cdots \circ \mathcal{U}_0 g \rangle_n \\
&= \langle -\mathcal{L}_n \circ (-\mathcal{L}_n)^{-1} \circ \mathcal{U}_{n-1} \circ \cdots \circ \mathcal{U}_0 h, \mathcal{U}_{n-1} \circ \cdots \circ \mathcal{U}_0 g \rangle_n \\
&= \langle (-\mathcal{L}_n)^{-1} \circ \mathcal{U}_{n-1} \circ \cdots \circ \mathcal{U}_0 h, \underbrace{-\mathcal{L}_n \circ \mathcal{U}_{n-1} \circ \cdots \circ \mathcal{U}_0 g}_{=0} \rangle_n \\
&= 0,
\end{aligned}$$

par symétrie de $\mathcal{L}_n$. La fonction g est quelconque, donc $h \in \ker(-\mathcal{L}_n \circ \mathcal{U}_{n-1} \circ \cdots \circ \mathcal{U}_0)^\perp$.

■

Ce résultat technique nous permet de passer à la preuve du Théorème 3.3.1.

Théorème 3.3.1. On montre en fait le résultat suivant, dont le théorème qui nous intéresse découle immédiatement :

$$\forall n \in \mathbf{N}, \ \lambda_{\sum_{i=1}^n k_i + 1}(-\mathbf{L}) \geq (n+1)\rho. \qquad (\star)$$

Tout d'abord, pour $n = 0$, l'hypothèse $\nabla^2 V \geq \rho I_d$, implique via le critère de Bakry-Émery (voir la Proposition 1.2.6) que $-\mathbf{L}$ satisfait une $IP(1/\rho)$, et donc que $\lambda_1(-\mathbf{L}) \geq \rho$ (voir la Proposition 1.2.3).

Soit donc $n \in \mathbf{N}^*$. Pour rappel, on s'intéresse à des quotients de Rayleigh de la forme

$$R(h) = \frac{\int_{\mathbf{R}^d} h(-\mathbf{L})h \, d\mu}{\int_{\mathbf{R}^d} h^2 d\mu}, \quad h \in \mathcal{D}(-\mathbf{L}).$$

Soit $h \in \mathcal{D}(-\mathbf{L})$ une fonction centrée telle que $h \in \bigcap_{k=1}^{n} \ker(-\mathcal{L}_k \circ \mathcal{U}_{k-1} \circ \cdots \circ \mathcal{U}_0)^{\perp}$. On souhaite montrer dans un premier temps que :

$$\int_{\mathbf{R}^d} h(-\mathbf{L})h \, d\mu = \langle \mathcal{U}_n \circ \cdots \circ \mathcal{U}_0 h, (-\mathcal{L}_{n+1})\mathcal{U}_n \circ \cdots \circ \mathcal{U}_0 h \rangle_{n+1} + \sum_{k=1}^{n+1} R_k, \qquad (\star\star)$$

avec

$$R_k = \langle \mathcal{U}_{k-1} \circ \cdots \circ \mathcal{U}_0 h, \nabla_k^2 V \mathcal{U}_{k-1} \circ \cdots \circ \mathcal{U}_0 h \rangle_k.$$

Le problème d'estimation sera alors séparé en deux, et on pourra utiliser la Proposition 3.2.5 pour éliminer le premier terme et estimer la somme des restes. On procède par récurrence sur $k \leq n$.

Initialisation : Comme h est centrée, $\mathcal{U}_0 h$ existe et, d'après la Proposition 3.2.8,

$$\begin{aligned}\int_{\mathbf{R}^d} h(-\mathbf{L})h \, d\mu &= \langle h, -\mathbf{L}h \rangle_0 = \langle \mathcal{U}_0 h, \mathcal{U}_0(-\mathbf{L}h) \rangle_1 \\ &= \langle \mathcal{U}_0 h, (-\mathcal{L}_1 + \nabla_1^2 V)\mathcal{U}_0 h \rangle_1 = \langle \mathcal{U}_0 h, (-\mathcal{L}_1)\mathcal{U}_0 h \rangle_1 + R_1.\end{aligned}$$

Hérédité : Supposons que, pour un certain $k \in [\![1, n]\!]$,

$$\int_{\mathbf{R}^d} h(-\mathbf{L})h \, d\mu = \langle \mathcal{U}_{k-1} \circ \cdots \circ \mathcal{U}_0 h, (-\mathcal{L}_k)\mathcal{U}_{k-1} \circ \cdots \circ \mathcal{U}_0 h \rangle_k + \sum_{j=1}^{k} R_j.$$

Comme $1 \leq k \leq n$, le Lemme 3.3.2 implique que $\int_{\mathbf{R}^d} \mathcal{U}_{k-1} \circ \cdots \circ \mathcal{U}_0 h \, d\mu = 0$. On en déduit que $\mathcal{U}_k(\mathcal{U}_{k-1} \circ \cdots \circ \mathcal{U}_0 h)$ est bien défini et la relation d'entrelacement de la Proposition 3.2.8 pour $\mathcal{U}_k$ donne, par isométrie :

$$\begin{aligned}&\langle \mathcal{U}_{k-1} \circ \cdots \circ \mathcal{U}_0 h, (-\mathcal{L}_k)\mathcal{U}_{k-1} \circ \cdots \circ \mathcal{U}_0 h \rangle_k + \sum_{j=1}^{k} R_j \\ &\qquad = \langle \mathcal{U}_k \circ \mathcal{U}_{k-1} \circ \cdots \circ \mathcal{U}_0 h, \mathcal{U}_k \circ (-\mathcal{L}_k)\mathcal{U}_{k-1} \circ \cdots \circ \mathcal{U}_0 h \rangle_{k+1} + \sum_{j=1}^{k} R_j \\ &\qquad = \langle \mathcal{U}_k \circ \cdots \circ \mathcal{U}_0 h, (-\mathcal{L}_{k+1} + \nabla_{k+1}^2 V)\mathcal{U}_k \circ \cdots \circ \mathcal{U}_0 h \rangle_{k+1} + \sum_{j=1}^{k} R_j \\ &\qquad = \langle \mathcal{U}_k \circ \cdots \circ \mathcal{U}_0 h, (-\mathcal{L}_{k+1})\mathcal{U}_k \circ \cdots \circ \mathcal{U}_0 h \rangle_{k+1} + \sum_{j=1}^{k+1} R_j,\end{aligned}$$

ce qui établit $(\star\star)$. Les propositions 3.2.5 et 3.2.7 impliquent alors, pour $k \in [\![1, n+1]\!]$,

$$\begin{aligned}R_k &= \langle \mathcal{U}_{k-1} \circ \cdots \circ \mathcal{U}_0 h, \nabla_k^2 V \mathcal{U}_{k-1} \circ \cdots \circ \mathcal{U}_0 h \rangle_k \\ &\geq \rho \|\mathcal{U}_{k-1} \circ \cdots \circ \mathcal{U}_0 h\|_k^2 = \rho \int_{\mathbf{R}^d} h^2 \, d\mu.\end{aligned}$$

La positivité de $-\mathcal{L}_{n+1}$ et $(\star\star)$ induisent donc :

$$\begin{aligned}\int_{\mathbf{R}^d} h(-\mathbf{L})h\,d\mu &= \langle \mathcal{U}_n \circ \cdots \circ \mathcal{U}_0 h, (-\mathcal{L}_{n+1})\mathcal{U}_n \circ \cdots \circ \mathcal{U}_0 h\rangle_{n+1} + \sum_{k=1}^{n+1} R_k \\ &\geq 0 + \sum_{k=1}^{n+1} R_k \geq \sum_{k=1}^{n+1} \rho \int_{\mathbf{R}^d} h^2\,d\mu = (n+1)\rho \int_{\mathbf{R}^d} h^2\,d\mu.\end{aligned}$$

Finalement, on obtient la borne inférieure pour le quotient de Rayleigh $R(h)$:

$$R(h) \geq (n+1)\rho,$$

uniformément en h.

Pour relier cette borne à la caractérisation de Courant-Fischer, il faut identifier un ensemble de fonctions auquel la fonction h est orthogonale. Cela donnera l'indice de la valeur propre associée et renseignera sur sa multiplicité. On rappelle que $h \in \bigcap_{k=1}^{n} \ker(-\mathcal{L}_k \circ \mathcal{U}_{k-1} \circ \cdots \circ \mathcal{U}_0)^{\perp}$ et est centrée (et donc orthogonale aux fonctions constantes). Si on pose $\mathcal{B}_n = \bigcup_{j=1}^{n} \left(\{f_i^j\}_{i=1}^{k_j}\right) \cup \{\mathbf{1}\}$, où $\{f_i^j\}_{i=1}^{k_j}$ est une base orthonormée de $\ker(-\mathcal{L}_j \circ \mathcal{U}_{j-1} \circ \cdots \circ \mathcal{U}_0)$ et $\mathbf{1}$ désigne la fonction constante réelle égale à 1, alors $\mathcal{B}_n$ constitue une telle famille. Le théorème de Courant-Fischer donne alors :

$$\lambda_{|\mathcal{B}_n|}(-\mathbf{L}) \geq \inf_{\substack{h \perp f \\ f \in \mathcal{B}_n}} R(f) \geq (n+1)\rho,$$

et comme $|\mathcal{B}_n| = \sum_{i=1}^{n} k_i + 1$, l'inégalité $(\star)$ est prouvée.

L'assertion du théorème suit aisément. Rappelons que les valeurs propres de $-\mathbf{L}$ sont classées par ordre croissant. Soit $r \in \mathbf{N}^*$, il existe $n \in \mathbf{N}$ tel que :

$$\sum_{i=1}^{n} k_i + 1 \leq r < \sum_{i=1}^{n+1} k_i + 1$$

(avec $\sum_{i=1}^{0} k_i = 0$). Ainsi, $(\star)$ implique que $\lambda_r(-\mathbf{L}) \geq \lambda_{\sum_{i=1}^{n} k_i + 1}(-\mathbf{L}) \geq (n+1)\rho$, et la condition sur r se réécrit $\sum_{i=1}^{n} k_i < r \leq \sum_{i=1}^{n+1} k_i$, ce qui clôt la démonstration. ■

Afin de retrouver précisément le résultat de Milman, il est nécessaire d'estimer la valeur des k_i. La principale difficulté réside dans le fait que les transformations de Riesz ne sont "pas assez explicites" pour permettre une détermination aisée des images réciproques en jeu. Il est néanmoins possible d'effectuer des estimations dans le cas général pour les premières itérations de la méthode.

Proposition 3.3.3. *On a, en toute généralité, $k_0 = 1$ et $k_1 = d$.*

Démonstration. La valeur de k_0 vient des hypothèses techniques formulées sur $-\mathbf{L}$, sous lesquelles le noyau de $-\mathbf{L}$ est réduit aux seules fonctions constantes (hypothèse **(H5)**). Il s'ensuit que $\ker(-\mathbf{L}) \simeq \mathbf{R}$, et donc que $k_0 = 1$.

L'estimation de k_1 est moins directe. Considérons $f \in \ker(-\mathcal{L}_1 \circ \mathcal{U}_0)$. Alors f est centrée (pour être dans le domaine de $\mathcal{U}_0$) et

$$\mathcal{L}_1 \mathcal{U}_0 f = 0.$$

L'hypothèse **(H5)** assure alors que $\mathcal{U}_0 f = C \in \mathbf{R}^d$. Comme les transformation de Riesz sont des isométries, elles sont en particulier inversibles, et on a alors $f = \mathcal{U}_0^\star C$ (l'absence de terme constant étant dû au caractère centré de f). Ainsi, la fonction f est complètement donnée par le vecteur $C \in \mathbf{R}^d$. Plus précisément, on a

$$\mathrm{d}_0(-\mathbf{L})^{-1/2} f = C,$$

soit pour tout $x \in \mathbf{R}^d$,

$$(-\mathbf{L})^{-1/2} f(x) = \langle C, x \rangle_0 + b,$$

avec $b \in \mathbf{R}$. Or, par définition de $(-\mathbf{L})^\alpha$, on a $(-\mathbf{L})^{1/2} b = 0$, et donc, pour tout $x \in \mathbf{R}^d$,

$$f(x) = (-\mathbf{L})^{1/2}(\langle C, x \rangle_0) + c',$$

avec $c' \in \mathbf{R}$ le centrage de f. Cette constante est complètement déterminée également par la donnée de C, de sorte que $\ker(-\mathcal{L}_1 \circ \mathcal{U}_0) \simeq \mathbf{R}^d$. Il vient donc que $k_1 = d$. ■

Dans le Théorème 3.3.1, cela donne :

- $\lambda_n(-\mathbf{L}) = 0$ pour $n = 0$,
- $\lambda_n(-\mathbf{L}) \geq \rho$ pour tout $n \in [\![1, d]\!]$,
- $\lambda_n(-\mathbf{L}) \geq 2\rho$ pour tout $n \geq d + 1$.

Cette conclusion est en accord avec d'autres résultats sur le sujet, par exemple le travail Klartag [54], où l'auteur établit que la dimension de l'espace propre associé à la première valeur propre non nulle (ici E_{λ_1}) est de dimension égale au plus à celle de l'espace sous-jacent (ici $\mathbf{R}^d$). Son cadre de travail est l'opérateur de Laplace avec conditions de Neumann en lien avec la mesure uniforme sur des corps convexes, transposable à l'opérateur $-\mathbf{L}$ traité ici pour une mesure de Boltzmann-Gibbs sur l'espace tout entier (tel que décrit par exemple dans [14]). Sa preuve, très élégante, permet d'estimer par en-dessous cette dimension en étudiant l'injectivité d'applications linéaires bien choisies. Un théorème de structure de cet espace peut en être déduis dans certains cas, que le Théorème 3.3.1 ne permet pas de retrouver. Ce résultat ne s'applique néanmoins qu'à ce dernier espace propre, l'article précité traitant principalement de l'inégalité de Poincaré (et donc de la première valeur propre non nulle).

On conclue ce chapitre sur un retour au cas gaussien. Comme cela a été amplement détaillé plus haut, ce cas est complètement explicite, et l'on connaît à la fois les valeurs propres et la structure (et dimension) des espaces propres. Le miracle gaussien va jusqu'à permettre le calcul explicite des noyaux en jeu et de leur dimension, ce qui montre au passage le caractère optimal du Théorème 3.3.1.

3.3.2 Optimalité dans le cas gaussien

Pour alléger l'écriture, on notera encore $\mathbf{L}$ le générateur du semi-groupe d'Ornstein-Uhlenbeck (en dimension d), invariant par rapport à la mesure gaussienne γ_d. De même ses versions tensorielles seront toujours notées $\mathcal{L}_n$, pour $n \in \mathbf{N}$. On sait de la première partie que $\sigma(-\mathbf{L}) = \mathbf{N}$, et que pour tout $n \in \mathbf{N}$, l'espace propre $E_n(-\mathbf{L})$ associé à la valeur propre n est de dimension $\binom{d-1+n}{n}$, engendré par les polynômes d'Hermite :

$$E_n(-\mathbf{L}) = \mathrm{Vect}\left(M_\alpha(h), \alpha \in \mathbf{N}^d, |\alpha| = n\right).$$

Cette description explicite donne accès à la décomposition spectrale de $-\mathbf{L}$ abordée plus tôt, qui permet de démontrer le résultat suivant.

Théorème 3.3.4. *Pour tout $n \in \mathbf{N}$, $\ker(-\mathcal{L}_n \circ \mathcal{U}_{n-1} \circ \cdots \circ \mathcal{U}_0) = E_n(-\mathbf{L})$. Ainsi*

$$k_n = \dim E_n(-\mathbf{L}) = \binom{d-1+n}{n}.$$

Démonstration. La preuve est en deux étapes, qui visent à l'identification de $\ker(-\mathcal{L}_n \circ \mathcal{U}_{n-1} \circ \cdots \circ \mathcal{U}_0)$. On montre d'abord que les seuls polynômes de degré $k \leq n$ dans ce noyau sont les (combinaisons linéaires des) polynômes d'Hermite de degré n, et ensuite que les seules fonctions de ce noyau sont des polynômes.

Étape 1. Remarquons d'abord que si $F \in E_\lambda(-\mathcal{L}_n)$, alors $\mathrm{d}_n F \in E_{\lambda-1}(-\mathcal{L}_{n+1})$. C'est une conséquence directe de la Proposition 3.2.4 : avec $\nabla^2_{n+1} V = I_d^{\otimes n+1}$, on a

$$\lambda \mathrm{d}_n F = \mathrm{d}_n(-\mathcal{L}_n)F = -\mathcal{L}_{n+1}(\mathrm{d}_n F) + \mathrm{d}_n F,$$

ce qui se réécrit :

$$-\mathcal{L}_{n+1}(\mathrm{d}_n F) = (\lambda - 1)\mathrm{d}_n F.$$

De plus, d'après la Proposition 3.2.8, cette assertion reste vraie en remplaçant d_n par $\mathcal{U}_n$. Rappelons enfin que $\mathcal{U}_n F$ n'est bien défini que pour F centrée. Cette dernière condition implique qu'il n'y a pas de polynôme non trivial de degré strictement inférieur à n dans $\ker(-\mathcal{L}_n \circ \mathcal{U}_{n-1} \circ \cdots \circ \mathcal{U}_0)$.

En effet, soit f une fonction polynomiale de degré k, avec $k < n$. Par définition de $\mathbf{L}$, $\mathbf{L}f$ est une fonction polynomiale de degré k, de même que $(-\mathbf{L})^{-1/2} f$, sous l'hypothèse que f est centrée. Ainsi, $\mathcal{U}_0 f$ est une fonction polynomiale de degré $k-1$ à valeurs vectorielles, dont les coefficients dominants ne dépendent que de ceux de f. Dans l'hypothèse où toutes les conditions de centrage sont vérifiées, $\mathcal{U}_{k-1} \circ \cdots \circ \mathcal{U}_0 f$ est de même un tenseur constant dont les coefficients non triviaux ne dépendent que des coefficients dominants de f. La condition $\int_{\mathbf{R}^d} \mathcal{U}_{k-1} \circ \cdots \circ \mathcal{U}_0 f d\mu = 0$ implique donc que $\mathcal{U}_{k-1} \circ \cdots \circ \mathcal{U}_0 f$ est le tenseur nul. Ainsi, tous les coefficients dominants de f sont nuls, et f est donc en fait une fonction polynomiale de degré $k-1$. Un récurrence descendante directe implique alors que f est en fait constante, et la condition de centrage $\int_{\mathbf{R}^d} f d\mu = 0$ permet de conclure que f est nulle. Ainsi, il n'y a pas de polynôme non

trivial de degré strictement inférieur à n dans le domaine de $\mathcal{U}_{n-1} \circ \cdots \circ \mathcal{U}_0$ (et a fortiori dans le noyau d'intérêt).

Soit maintenant $M_\alpha(h)$ un polynôme d'Hermite de degré n. Étant une fonction propre, toutes les conditions de centrage précédentes sont satisfaites (en vertu de l'observation faite en tout début de preuve). De plus, étant donné que $M_\alpha(h)$ est une fonction polynomiale de degré n, il suit que $\mathcal{U}_{n-1} \circ \cdots \circ \mathcal{U}_0 h$ est une fonction polynomiale constante (non triviale), et donc que $-\mathcal{L}_n \mathcal{U}_{n-1} \circ \cdots \circ \mathcal{U}_0 h$ est le tenseur nul. Ainsi, on en déduit que $M_\alpha(h) \in \ker(-\mathcal{L}_n \circ \mathcal{U}_{n-1} \circ \cdots \circ \mathcal{U}_0)$. De même, si f est une combinaison linéaire de polynômes d'Hermite de degré n, on a encore $f \in \ker(-\mathcal{L}_n \circ \mathcal{U}_{n-1} \circ \cdots \circ \mathcal{U}_0)$. Cela établit le premier point.

Étape 2. Soit maintenant $f \in L^2(\mu)$ une fonction non triviale. La décomposition spectrale donne :

$$f = \sum_{k=0}^{+\infty} \sum_{\substack{\alpha \in \mathbf{N}^d \\ |\alpha|=k}} f_\alpha M_\alpha(h),$$

où la seconde somme porte sur tous les polynômes d'Hermite de degré fixé, et $f_\alpha = \langle f, M_\alpha(h) \rangle_0$. Ainsi,

$$-\mathcal{L}_n \circ \mathcal{U}_{n-1} \circ \cdots \circ \mathcal{U}_0 f = \sum_{k=0}^{+\infty} \sum_{\substack{\alpha \in \mathbf{N}^d \\ |\alpha|=k}} f_\alpha (-\mathcal{L}_n) \circ \mathcal{U}_{n-1} \circ \cdots \circ \mathcal{U}_0 M_\alpha(h).$$

D'après l'étape 1, cette somme n'est bien définie que si $f_\alpha = 0$ pour tout $\alpha \in \mathbf{N}^d$ tel que $|\alpha| < n$ (on retrouve les conditions d'orthogonalité successives pour f). De plus, $f \in \ker(-\mathcal{L}_n \circ \mathcal{U}_{n-1} \circ \cdots \circ \mathcal{U}_0)$ si et seulement si :

$$\sum_{k=n+1}^{+\infty} \sum_{\substack{\alpha \in \mathbf{N}^d \\ |\alpha|=k}} |\alpha|^{-1/2} \dots (|\alpha| - n + 1)^{-1/2} (|\alpha| - n) f_\alpha \mathrm{d}_{n-1} \dots \mathrm{d}_0 M_\alpha(h) = 0.$$

Or, toujours d'après l'observation préliminaire, la famille $(\mathrm{d}_{n-1} \dots \mathrm{d}_0 M_\alpha(h))_{\substack{\alpha \in \mathbf{N}^d \\ |\alpha|=k}}$ est orthogonale. L'identité précédente implique donc que $f_\alpha = 0$ pour tout $\alpha \in \mathbf{N}^d$ tel que $|\alpha| \geq n+1$. Comme f est non triviale, elle s'écrit donc comme une combinaison linéaire de polynômes d'Hermite de degré n.

On en déduit donc que $\ker(-\mathcal{L}_n \circ \mathcal{U}_{n-1} \circ \cdots \circ \mathcal{U}_0) = E_n(-\mathbf{L})$. L'égalité des dimensions suit directement. ■

Il vient alors que l'inégalité dans le Théorème 3.3.1 est en fait une égalité, ce qui permet d'avancer que ce résultat est optimal. Cette affirmation est néanmoins à nuancer. En effet, le caractère complètement explicite des objets dans le cadre gaussien permet de retrouver exactement les bonnes dimensions d'espaces propres et les estimations optimales sur les valeurs propres, mais reste spécifique à ce cadre précis. Dans la plupart des cas, les fonctions propres ne sont pas connues et une observation telle que celle faite en ouverture de la preuve précédente n'est pas accessible.

3.3.3 Perspectives

En pratique, l'utilisation du Théorème 3.3.1 est couplée à l'estimation de la suite $(k_i)_{i\geq 0}$. La difficulté réside dans la compréhension de l'action des transformations de Riesz, en particulier dans leur action sur la structure hilbertienne héritée de la décomposition spectrale. Ces estimations dépassent le cadre de ce chapitre, et les tentatives dans cette direction (dans le cas général) n'ont pour l'instant malheureusement pas donné de résultats concluants.

Une suite enrichissante de ce travail serait donc par exemple d'aller plus loin dans la description des espaces propres, la connaissance de leur structure permettant d'avancer dans l'estimation de la suite $(k_i)_{i\geq 0}$, ainsi que dans celle des multiplicités et des valeurs propres elles-mêmes.

Bien que peu explicites au regard du cas gaussien, d'autres exemples pourraient être traités (dans l'esprit de la fin du chapitre précédent), pour essayer de décrire l'action des transformations de Riesz dans certains cas. La recherche de situations dans lesquelles ces dernières sont plus explicites pourrait également constituer un pas intéressant dans cette direction.

Enfin, Milman propose un grand nombre de conjectures dans sont article initial, dont plusieurs sont encore ouvertes. En particulier quand l'espace sous-jacent n'est plus l'espace euclidien $\mathbf{R}^d$. Une extension du présent travail consisterait à une traduction des résultats dans un cadre plus général, par exemple le cadre riemannien, dans lequel la structure algébrique introduite serait un atout indéniable. Notons que les relations d'entrelacement classiques décrites dans les chapitres précédent ont déjà fait l'objet d'une telle généralisation, voir par exemple le travail de Baptiste Huguet sur la question [47], et que les transformations de Riesz avaient déjà été étudiées par Bakry dans ce contexte.

Chapitre 4

Aspects numériques des inégalités de Poincaré et application à l'analyse de sensibilité

CE DERNIER CHAPITRE revient sur l'inégalité de Poincaré en tant que telle. Son but est d'explorer les liens entre cette dernière et l'analyse de sensibilité, dans le cadre d'une application à l'hydrographie (en petite dimension). Une méthode de calcul numérique de la constante de Poincaré est également développée et implémentée.

Ce travail (aujourd'hui en cours de rédaction) a été effectué en collaboration avec Aldéric Joulin et Olivier Roustant, sur la base d'un projet étudiant donné à l'INSA Toulouse. Il fait suite à l'article de Barthe, Iooss et Roustant [13] qui traitait de ce problème en dimension un. Le présent chapitre aborde la question de la dimension deux, qui avait été explorée dans le cadre de ce projet par deux étudiants de mathématiques appliquées de l'INSA, Jérôme Deveaux et Landry Duguet [31]. Les constantes de Poincaré sont obtenues via la résolution par éléments finis d'un problème aux valeurs propres faisant intervenir le trou spectral. Il n'est pas discuté d'extensions de cette technique aux dimensions supérieures, les éléments finis étant intrinsèquement liés à la petite dimension. On peut néanmoins mentionner un travail récent de Pillaud-Vivien, Bach, Lelièvre, Rudi et Stolz [71], qui estime par Monte-Carlo cette constante en dimension quelconque. Cette approche statistique est cependant un peu moins pertinente en petite dimension, où les éléments finis permettent une meilleure précision.

Le chapitre sera découpé comme suit. Le problème physique est présenté dans une première partie introductive, suivie d'un rappel sur l'analyse de sensibilité et les liens avec l'inégalité de Poincaré. Les aspects numériques liés et les résultats obtenus font l'objet de la dernière section.

Table des matières

4.1 Contexte général 94
4.1.1 Modèle physique 94
4.1.2 Modélisation probabiliste 97
4.2 Analyse de sensibilité 104
4.2.1 Principes de base 104
4.2.2 Indices globaux et liens avec l'inégalité de Poincaré 106
4.3 Aspects numériques de l'inégalité de Poincaré 112
4.3.1 Principe général 112
4.3.2 Généralités sur les éléments finis en dimension deux 115
4.3.3 Domaine et maillage 118
4.3.4 Validation de la méthode 128
4.3.5 Implémentation 134
4.3.6 Conclusion et perspectives 139

4.1 Contexte général

4.1.1 Modèle physique

Le problème d'origine est proposé par Rocquigny [29], repris par Iooss [48], et enfin par Iooss et Lemaître dans [49]. Il consiste à modéliser le coût d'entretien d'une digue au bord d'une rivière, dont le cours peut varier et dont les crues occasionnent un surcoût d'entretien. Le flot de la rivière est modélisé par une simplification des équations de Barré de Saint-Venant 1D [11]. La hauteur h de l'eau est donnée par ces équations, en fonction de la vitesse d'écoulement q, par l'équation :

$$\partial_t(hq) + \partial_x(hq^2) + gh\partial_x h = \frac{gq^2}{K_s^2 h^{1/3}}, \tag{SV}$$

où g représente l'intensité de la pesanteur (supposée constante en temps et en espace) et K_s est le *coefficient de Strickler*, qui permet d'exprimer la quantité de mouvement transmise par le fluide à la paroi de la rivière. Ce dernier dépend de nombreux paramètres, comme le matériau de la paroi, sa rugosité, la végétation, etc. Il est en général évalué empiriquement.

Le domaine de résolution fait intervenir divers paramètres, dont les longueur et largeur du tronçon de rivière étudié, ainsi que les côtes amont et aval du tronçon (qui donnent le dénivelé parcouru et donc l'inclination). Ces paramètres se retrouvent dans la formulation simplifiée ci-dessous.

La principale grandeur d'intérêt dérivée de cette équation est la hauteur maximale de l'eau sur une période de temps donnée (ici une année), notée H, qui s'exprime ainsi

en fonction de la vitesse d'écoulement maximale Q :

$$H = \left(\frac{Q}{BK_s \sqrt{\frac{Z_m - Z_v}{L}}} \right)^{0.6},$$

où B désigne la largeur de la rivière (supposée elle aussi constante), L la longueur du tronçon de rivière (et donc aussi la longueur de la digue), supposé rectiligne, et Z_m et Z_v les cotes (altitudes par rapport à un niveau de référence, souvent la mer) respectivement en amont et aval du tronçon de rivière (la différence $Z_m - Z_v$ est donc le dénivelé parcouru par la rivière). La hauteur de crue maximale S sur une année est donc égale au niveau maximal de la rivière additionné de la côte aval, auquel sont soustraits la hauteur de la berge C_b et la hauteur de la digue H_d, soit :

$$S = H + Z_v - H_d - C_b.$$

Dans le cadre d'applications industrielles, en plus de la hauteur de crue maximale, on s'intéresse au coût annuel de la digue. Celui-ci peut s'exprimer par :

$$C = \mathbf{1}_{S>0} + \left[0.2 + 0.8 \left(1 - e^{-1000/S^4} \right) \right] \mathbf{1}_{S \leq 0} + \frac{1}{20} \max(H_d, 8).$$

Ce coût (exprimé en millions d'euros) comporte trois termes. Le premier correspond au coût de remise en état de la digue suivant une crue, le second à la maintenance annuelle de la digue en l'absence de crue, et le dernier au coût initial de construction de la digue.

Les trois grandeurs d'intérêt (H, S et C) peuvent être obtenues directement à partir des mesures des différents paramètres du modèle. Malheureusement, la plupart de ces observables ne peuvent être connues avec précision, et présentent souvent des fluctuations. Pour cette raison, il est commode de les modéliser par des variables aléatoires, dont les différentes lois sont présentées ci-dessous.

Variable	Grandeur mesurée	Loi
Q	vitesse d'écoulement maximale	$Gum(1013, 558)$ (restreinte à $[500, 3000]$)
K_s	coefficient de Strickler	$\mathcal{N}(30, 64)$ (restreinte à $[15, +\infty[$)
Z_m	cote amont	$\mathcal{T}(49, 50, 51)$
Z_v	cote aval	$\mathcal{T}(49, 50, 51)$
L	longueur du tronçon	$\mathcal{T}(4990, 5000, 5010)$
B	largeur du tronçon	$\mathcal{T}(295, 300, 305)$
H_d	hauteur de la digue	$\mathcal{U}(7, 9)$
C_b	hauteur de la berge	$\mathcal{T}(55, 55.5, 56)$

TABLE 4.1 – Variables du problème

La notation $Gum(a,b)$, pour $a \in \mathbf{R}$, $b \in \mathbf{R}_+^*$, désigne la *loi de Gumbel* de paramètres a et b, dont la densité est donnée, pour tout $x \in \mathbf{R}$, par :

$$f(x) = \frac{1}{\beta} \exp\left[-\left(\frac{x-a}{b} + e^{-\frac{x-a}{b}} \right) \right],$$

et de fonction de répartition :

$$F(x) = \exp\left(-e^{-\frac{x-a}{b}} \right).$$

De même que pour la gaussienne, le paramètre a est un paramètre de localisation, et b un paramètre d'échelle. De même également, on appellera *loi de Gumbel standard* la loi $Gum(0,1)$. Notons néanmoins qu'il n'est pas possible d'obtenir une loi de Gumbel standard par centrage/réduction comme on le ferait pour la gaussienne. Ce choix est classique dans la modélisation des valeurs extrêmes d'une série statistique, voir par exemple l'article initial de Gumbel [43].

La notation $\mathcal{T}(a,c,b)$, avec $a < c < b$, désigne la loi triangulaire sur l'intervalle $[a,b]$ de mode c, dont la densité est donnée, pour tout $x \in \mathbf{R}$, par :

$$f(x) = \frac{2(x-a)}{(b-a)(c-a)} \mathbf{1}_{[a,c]}(x) + \frac{2(b-x)}{(b-a)(b-v)} \mathbf{1}_{]c,b]}(x).$$

Son nom vient du graphe de cette densité, qui représente un triangle dont la base est le segment $[a,b]$ et dont le sommet est situé au-dessus de c. On considère en général la loi $\mathcal{T}(0,1/2,1)$ comme la loi triangulaire standard.

Les paramètres présentés rendent compte des ordres de grandeur rencontrés dans les applications : une altitude de 50m environ par rapport au niveau de la mer (ce qui correspond assez bien à l'altitude à laquelle les cours d'eau présentent une typologie en accord avec les hypothèses), un tronçon de 5km de long pour 300m de large et une digue de 8m de hauteur. À titre indicatif, cela correspond à la typologie de la Garonne un peu avant Bordeaux.

En pratique, la dépendance de S et C en tous les paramètres du modèle est, bien qu'explicite, difficile à appréhender. Les entrées étant aléatoires, on cherche en pratique à identifier dans un premier temps les plus influentes d'entre elles. En d'autres termes, on souhaite estimer comment l'incertitude sur une entrée donnée se traduit sur la sortie. C'est l'objet d'étude de l'analyse de sensibilité, qui permet de définir cette notion d'influence et de la quantifier via des indices de sensibilité. Ces techniques seront plus amplement détaillées dans la section suivante.

Mentionnons enfin que, dans la plupart des travaux cités précédemment, les différentes grandeurs en jeu dans le modèle sont considérées indépendantes. Cela simplifie radicalement l'analyse de sensibilité, et permet de se ramener à des composantes unidimensionnelles, qui ont fait l'objet de l'étude dans [13]. Cette hypothèse peut néanmoins sembler particulièrement forte dans le contexte d'un problème concret. Un des buts de cette partie est d'étudier le cas où les variables ne sont plus supposées indépendantes,

mais corrélées deux à deux (selon des couplages qui seront précisés ultérieurement). La section suivante s'intéresse à une façon de modéliser la dépendance d'un couple de variable aléatoire : la copule.

4.1.2 Modélisation probabiliste

Formalisme fonctionnel

Il existe plusieurs façons de décrire la loi d'un couple de variables aléatoires dépendantes. La copule, introduite par Sklar en 1959 [79], permet de séparer l'étude de la structure de dépendance entre les variables de celle des marges. Les définitions qui suivent pourront être retrouvées dans l'ouvrage de référence de Nelsen [66].

Les copules se définissent au niveau des fonctions de répartition, et nécessitent donc une définition de ces fonctions en dimension deux. Plusieurs approches existent pour étendre le concept (intrinsèquement unidimensionnel) de fonction de répartition à la dimension supérieure, mais un cadre adapté ici pour les décrire est celui des fonctions 2-croissantes.

Définition 4.1.1 (Fonction 2-croissante)**.** Soient I_1 et I_2 deux intervalles de $\mathbf{R}$. Une fonction $H : I_1 \times I_2 \to \mathbf{R}$ est dite *2-croissante* si pour tous $(x_1, x_2), (y_1, y_2) \in I_1 \times I_2$ avec $x_1 \leq x_2$ et $y_1 \leq y_2$, on a :

$$H(x_1, y_1) + H(x_2, y_2) - H(x_1, y_2) - H(x_2, y_1) \geq 0.$$

Notons que cette définition diffère par nature de celle des fonctions croissantes adoptée dans le chapitre 1. Il existe en particulier des fonctions 2-croissantes qui ne sont pas croissantes dans toutes les directions, et réciproquement (voir [66] §2.1). Les fonctions de répartition bivariées sont des cas particuliers de fonctions 2-croissantes. En effet, cette notion de croissance est plus adaptée à la définition probabiliste d'une fonction de répartition, qui calcule la masse de certains ensemble sous une loi de probabilité donnée.

Définition 4.1.2 (Fonction de répartition jointe)**.** Une fonction de répartition jointe H est une fonction 2-croissante définie sur $\mathbf{R}^2$ telle que pour tout $(x_1, y_1) \in \mathbf{R}^2$ fixé :

$$\lim_{y \to -\infty} H(x_1, y) = \lim_{x \to -\infty} H(x, y_1) = 0,$$

et

$$\lim_{x,y \to +\infty} H(x, y) = 1.$$

Les *marges de* H, notées F et G, s'obtiennent par passage à la limite : pour $(x, y) \in \mathbf{R}^2$, on pose :

$$F(x) = \lim_{z \to +\infty} H(x, z) \text{ et } G(y) = \lim_{z \to +\infty} G(z, y).$$

On retrouve comme conséquence immédiate de la définition qu'une fonction de répartition jointe est monotone en chacune des variables, ce qui entraîne la continuité presque partout de telles fonctions, et à valeurs dans $[0,1]$. De plus, les marges ainsi définies sont bien des fonctions de répartition.

Une copule est un cas particulier de fonction de répartition jointe où le domaine considéré est le carré unité $[0,1]^2$.

Définition 4.1.3 (Copule)**.** Une fonction $C : [0,1]^2 \to [0,1]$ est une *copule* si elle vérifie les trois conditions suivantes :

1. C est 2-croissante,
2. C est *enracinée* : pour tous $x, y \in [0,1]$:

$$C(x,0) = C(0,y) = 0,$$

3. pour tous $x, y \in [0,1]$:

$$C(x,1) = x \text{ et } C(1,y) = y.$$

Comme nous le verrons tout à l'heure avec le théorème de Sklar, une copule décrit la structure de dépendance entre deux variables aléatoires de loi uniforme sur $[0,1]$. De plus, la propriété de 2-croissance implique, à l'instar de la monotonie sur $\mathbf{R}$, une certaine régularité sur les copules. En particulier, les copules sont dérivables presque partout parallèlement aux axes (voir [66] Théorème 2.2.7).

L'analogie avec les fonctions de répartition jointes se comprend à la lumière du théorème de Sklar [79], qui permet de faire le lien entre copule et lois marginales.

Théorème 4.1.4 (Sklar)**.** *Soit $H : \mathbf{R}^2 \to [0,1]$ une fonction de répartition jointe avec marges F et G. Alors il existe une copule C telle que pour tout $(x,y) \in \mathbf{R}^2$:*

$$H(x,y) = C(F(x), G(y)).$$

De plus, C est unique si F et G sont continues.

Réciproquement, si F et G sont deux fonctions de répartition et C une copule, alors H définie pour $(x,y) \in \mathbf{R}^2$ par

$$H(x,y) = C(F(x), G(y))$$

est une fonction de répartition jointe.

On renvoie à [66], Théorème 2.3.3, pour la preuve. Ce résultat établit une correspondance bijective entre les structures de dépendance et les copules : à marges prescrites, toute loi jointe peut être exprimée en fonction d'une unique copule. Notons que dans le cas où les marges ne sont pas continues, la copule est déterminée de façon univoque uniquement sur $\mathrm{Im}(F) \times \mathrm{Im}(G)$.

Cette bijectivité permet de considérer le cas inverse : si la loi jointe et les marges sont prescrites, comment peut-on retrouver la copule en jeu ? Cela peut se comprendre

simplement en terme d'inverse généralisé des fonctions de répartition. Si $F : \mathbf{R} \to [0,1]$ est une fonction de répartition, on définit son inverse généralisé F^{-1} sur $[0,1]$ par :

$$F^{-1}(u) = \inf\{x \in \mathbf{R} \,:\, F(x) \geq u\}.$$

Si F est obtenue par intégration d'une densité de probabilité, l'inverse généralisé coïncide avec le "vrai" inverse sur le support de cette densité. Pour la copule, cela se traduit par le résultat suivant.

Proposition 4.1.5. *Soit H une fonction de répartition jointe de marges F et G. On pose, pour $(u,v) \in [0,1]$,*

$$C(u,v) = H(F^{-1}(u), G^{-1}(v)).$$

Alors on a, pour tout $(x,y) \in \mathbf{R}^2$,

$$H(x,y) = C(F(x), G(y)).$$

Cette proposition énonce que les copules sont compatibles avec les changements de variables qui interviennent lors des changements de loi. Notons que C ainsi posée n'est pas toujours une copule : elle peut n'être pas définie sur tout le carré $[0,1]^2$ (certains auteurs parlent alors de *sous-copule*). C'est néanmoins le cas dès que F et G sont continues, ce qui sera le cas dans les applications qui vont suivre.

Cas des variables aléatoires

Les définitions ci-dessus sont posées en-dehors d'un véritable contexte probabiliste. Ces concepts coïncident néanmoins avec les définitions usuelles des fonctions de répartition. En particulier, la fonction de répartition jointe d'un couple de variables aléatoires réelles (X,Y), notée $F_{(X,Y)}$, peut être définie, pour tout $(x,y) \in \mathbf{R}^2$, par :

$$F_{(X,Y)}(x,y) = \mathbf{P}(X \leq x, Y \leq y).$$

Cette fonction est bien 2-croissante, et on a en fait par calcul direct, pour tous $x_1 \leq x_2$ et $y_1 \leq y_2$,

$$\begin{aligned} F_{(X,Y)}(x_1,y_1) + F_{(X,Y)}(x_2,y_2) - F_{(X,Y)}(x_1,y_2) - F_{(X,Y)}(x_2,y_1) \\ = \mathbf{P}(X \in [x_1,x_2], Y \in [y_1,y_2]) \geq 0. \end{aligned}$$

On comprend dès lors pourquoi la notion de 2-croissance est plus adaptée que la croissance directionnelle à la définition des fonctions de répartition bidimensionnelles. On retrouve également la loi des marges par passage à la limite. Par exemple pour X, on a pour tout $x \in \mathbf{R}$:

$$\lim_{y \to +\infty} F_{(X,Y)}(x,y) = \lim_{y \to +\infty} \mathbf{P}(X \leq x, Y \leq y) = \mathbf{P}(X \leq x, Y \in \mathbf{R}) = \mathbf{P}(X \leq x) = F_X(x).$$

Enfin, dans le cas où X et Y sont deux lois à densité par rapport à la mesure de Lebesgue, dériver la fonction de répartition jointe permet de retrouver la densité du couple. En effet, on a pour tout $(x, y) \in \mathbf{R}^2$, par le théorème de Fubini-Tonelli,

$$F_{(X,Y)}(x, y) = \mathbf{P}(X \leq x, Y \leq y) = \int_{-\infty}^{y} \int_{-\infty}^{x} f_{(X,Y)}(s, t) ds dt,$$

et donc d'une part :

$$\partial_x F_{(X,Y)}(x, y) = \int_{-\infty}^{y} \partial_x \left(\int_{-\infty}^{x} f_{(X,Y)}(s, t) ds \right) dt = \int_{-\infty}^{y} f_{(X,Y)}(x, t) dt,$$

et d'autre part :

$$\partial^2_{x,y} F_{(X,Y)}(x, y) = \partial_y \left(\int_{-\infty}^{y} f_{(X,Y)}(x, t) dt \right) = f_{(X,Y)}(x, y).$$

Cette dernière observation permet d'exprimer aisément la densité jointe en fonction de la copule obtenue via le théorème de Sklar.

Proposition 4.1.6. *Soit H une fonction de répartition jointe de marges continues F et G et de copule C. Soit (X, Y) le vecteur aléatoire de fonction de répartition jointe H. Sa densité $f_{(X,Y)}$ est donnée, pour tout $(x, y) \in \mathbf{R}^2$, par :*

$$f_{(X,Y)}(x, y) = \partial^2_{xy} C(F(x), G(y)) f_X(x) f_Y(y).$$

Démonstration. Ce résultat s'obtient directement en calculant la dérivée croisée dans l'expression du Théorème 4.1.4. ■

C'est cette dernière forme qui sera la plus utile pour travailler avec les densités jointes. À ce titre, la copule permet de séparer au sein de la densité la partie liée purement aux marges (produit des densités) de la partie liée à la structure de dépendance. Dans cette optique, il peut être intéressant d'écrire ce résultat dans le cas où les deux marges suivent la loi uniforme $[0, 1]$. On a, pour tout $(x, y) \in [0, 1]^2$,

$$f_{(X,Y)}(x, y) = \partial^2_{xy} C(x, y).$$

Cette expression est parfois elle-même appelée *densité de la copule* C. Ainsi, la structure de dépendance peut être complètement comprise dans le cas où les marges suivent des lois uniformes. Le passage à des marges générales peut se faire simplement par inversion de la fonction de répartition (qui permet le passage de toute mesure à densité par rapport à la mesure de Lebesgue à la mesure uniforme sur $[0, 1]$). De ce point de vue, on a bien séparé les problèmes relevant de la dépendance de ceux liés aux marges.

Avant de pousser plus loin l'étude de ces objets et leur lien avec le problème initial, on présente quelques exemples de copules classiques, dont celle qui sera utilisée dans la suite du développement : la copule gaussienne.

Exemples de copules

Copule produit Le premier exemple dont on pourrait légitimement se préoccuper est celui de la copule produit, qui correspond au cas d'une loi produit. En effet, si H est une fonction de répartition jointe dont les marges F et G sont indépendantes, on a :

$$C(F(x), G(y)) = H(x, y) = \mathbf{P}(X \leq x, Y \leq y) = \mathbf{P}(X \leq x)\mathbf{P}(Y \leq y) = F(x)F(y).$$

La copule est donc définie ici, pour tout $(u, v) \in [0, 1]^2$, par :

$$C(u, v) = uv.$$

Copule de Gumbel Un autre exemple (moins trivial, mais non moins utilisé) de copule classique est la copule de Gumbel (à ne pas confondre avec la loi du même nom). Elle est donnée, pour $(u, v) \in [0, 1]^2$ et $\theta \in [0, 1]$, par :

$$C_\theta(u, v) = u + v - 1 + (1 - u)(1 - v)e^{-\theta \log(1-u) \log(1-v)}.$$

Elle est souvent associée à la loi exponentielle bivariée de Gumbel (toujours à ne pas confondre avec la loi de Gumbel tout court), qui correspond à deux marges exponentielles et dont la copule est celle donnée ci-dessus, et dont la fonction de répartition jointe peut s'écrire comme suit, pour $(x, y)\mathbf{R}^2$:

$$H_\theta(x, y) = \left(1 - e^{-x} - e^{-y} + e^{-(x+y+\theta xy)}\right) \mathbf{1}_{\mathbf{R}_+}(x)\mathbf{1}_{\mathbf{R}_+}(y).$$

Copules archimédiennes Il est important de noter que la copule de Gumbel est une variation d'un cas particulier d'une famille de copules très utilisées : les copules archimédiennes. Une telle copule est définie, pour $(u, v) \in [0, 1]^2$, par :

$$C(u, v) = \varphi^{[-1]}(\varphi(u, \theta) + \varphi(v, \theta), \theta),$$

où $\theta \in \Theta$ est un paramètre, et $\varphi : [0, 1] \times \Theta \to \mathbf{R}_+$ est une fonction continue, strictement décroissante et convexe en sa première variable telle que $\phi(1, \cdot) = 0$, pour laquelle on définit :

$$\varphi^{[-1]}(t, \theta) = \varphi^{-1}(t, \theta)\mathbf{1}_{[0,\varphi(0,\theta)]}(t).$$

On obtient une variante de la copule de Gumbel, dite "copule de survie", basée non sur une fonction de répartition F mais sur une fonction de survie $1 - F$, avec $\varphi(t, \theta) = \log(1 - \theta \log t)$.

L'intérêt de ces copules réside dans le paramètre θ, qui décrit l'intensité de la corrélation entre les variables. C'est également le cas pour la copule classique de Gumbel ci-dessus, où l'on retrouve par exemple la copule produit pour $\theta = 0$.

La dernière copule de ces exemples présente aussi ce type de paramètres : la copule gaussienne.

Copule gaussienne Le principal exemple qui servira dans la suite du travail est la copule gaussienne. Elle n'est néanmoins pas complètement explicite, et nécessite un peu de mise en place préalable. On rappelle qu'un vecteur aléatoire (X, Y) est gaussien si pour tout $(\alpha, \beta) \in \mathbf{R}^2$, la variable réelle $\alpha X + \beta Y$ est gaussienne.

Un vecteur gaussien est, à l'instar de la variable homonyme, caractérisé par ses deux premiers moments. Si (X, Y) est un vecteurs gaussien, son espérance est définie comme le vecteur $\mu = (\mathbf{E}[X], \mathbf{E}[Y])$, et sa covariance comme la matrice :

$$\Gamma = \begin{pmatrix} \mathrm{Var}(X) & \mathrm{Cov}(X,Y) \\ \mathrm{Cov}(X,Y) & \mathrm{Var}(Y) \end{pmatrix}.$$

On note alors $(X, Y) \sim \mathcal{N}_2(\mu, \Gamma)$. L'inégalité de Cauchy-Schwarz implique qu'une matrice de covariance est toujours symétrique semi-définie positive. Quitte à renormaliser chaque marge par son écart-type respectif, on peut toujours se ramener à un vecteur dont la matrice de covariance s'écrit :

$$\Gamma_\rho = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix},$$

où $\rho \in [-1, 1]$ est le coefficient de corrélation entre X et Y. Pour des raisons de symétrie, on considérera dans la suite des vecteurs dont la corrélation est positive. Un vecteur gaussien de loi $\mathcal{N}_2(\mu, \Gamma)$ admet pour densité, pour tout $x \in \mathbf{R}^2$,

$$f(x) = \frac{1}{2\pi\sqrt{\det \Gamma}} \exp\left[-\frac{1}{2}(x-\mu)^T \Gamma^{-1}(x-\mu)\right],$$

si Γ est inversible (soit $\rho \neq 1$). On parlera de vecteur *centré* si $\mu = 0$, et *réduit* si $\Gamma = I_2$. Dans ce cas, la densité est le produit de deux densités gaussiennes centrées réduites unidimensionnelles. On s'intéressera plus spécifiquement au cas où $\mu = 0$ et $\Gamma = \Gamma_\rho$ pour un certain $\rho \in [0, 1[$, car c'est dans ce cadre que sont définies les copules gaussiennes. On a donc, dans ce cas-là, pour tout $(x, y) \in \mathbf{R}^2$,

$$f_\rho(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2(1-\rho^2)}(x^2 + y^2 - 2\rho xy)\right].$$

On notera dans la suite Φ_ρ la fonction de répartition jointe associée à cette dernière densité, et Φ la fonction de répartition de la loi $\mathcal{N}(0, 1)$. Ces deux fonctions nous permettent de définir la copule gaussienne, qui ne s'exprime pas explicitement.

Définition 4.1.7 (Copule gaussienne)**.** La copule gaussienne est définie, pour $\rho \in [0, 1[$ et tout $(u, v) \in [0, 1]^2$, par :

$$C_\rho(u,v) = \Phi_\rho(\Phi^{-1}(u), \Phi^{-1}(v)).$$

La densité d'un vecteur aléatoire (X, Y) de copule C et de marges F et G continues peut s'écrire semi-explicitement : pour tout $(x, y) \in \mathbf{R}$,

$$f_{(X,Y)}(x,y) = \frac{1}{\sqrt{1-\rho^2}} \exp\Big[-\frac{1}{2(1-\rho^2)}\Big(\rho^2[\Phi^{-1}(F(x))^2 + \Phi^{-1}(G(y))^2] - 2\rho\Phi^{-1}(F(x))\Phi^{-1}(G(y))\Big)\Big] f_X(x) f_Y(y).$$

On retrouve bien l'expression de la densité d'un vecteur gaussien centré de covariance Γ_ρ si les marges sont deux gaussiennes centrées réduites.

On comprend de même l'influence du paramètre ρ, qui mesure le "degré de dépendance" des marges. En particulier, on obtient de nouveau la copule produit pour $\rho = 0$. Le cas $\rho = 1$ est intéressant. En effet, la matrice Γ_1 est singulière et aucune des fonctions ci-dessus n'est donc bien définie. On peut néanmoins donner un sens au vecteur aléatoire (X, Y) dans ce cas, si les marges sont toutes deux gaussiennes centrées réduites. En effet, le coefficient de corrélation s'exprime directement comme :

$$\rho = \mathrm{Cov}(X, Y).$$

Le cas $\rho = 1$ correspond donc au cas d'égalité dans Cauchy-Schwarz :

$$\mathrm{Cov}(X, Y) = 1 = \sigma_X \sigma_Y,$$

ce qui implique que X et Y sont colinéaires. Comme ces deux variables sont centrées réduites, on a en fait $X = Y$. Dans le cas où les marges sont gaussiennes, le vecteur "dégénère" en une unique variable aléatoire gaussienne portée par la diagonale $x = y$ du plan. Ce cas limite se comprend également au vu des lignes de niveau de la densité pour différentes valeurs de ρ, comme il sera indiqué dans la dernière section. Notons enfin que cette analyse ne tient plus si une des marges n'est pas gaussienne.

Tous les outils ont été introduits pour permettre la suite du travail. Avant de procéder à des rappels sur l'analyse de sensibilité, revenons à la description du problème physique de départ.

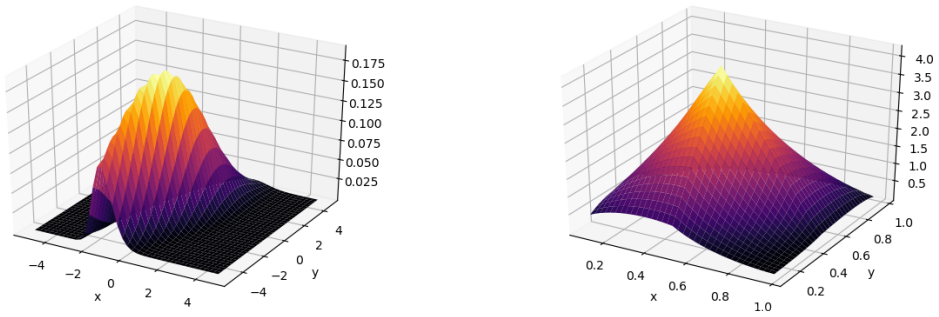
Retour au problème physique

La première direction prise pour s'affranchir de l'hypothèse d'indépendance présente dans les précédents travaux consiste à supposer que les variables sont dépendantes par paires. Si l'on reprend les grandeurs présentées dans le tableau en première partie, il est raisonnable d'apparier certaines variables entre elles. Ainsi, on considère les distributions suivantes :

Couple	Paramètre ρ
(Q, K_s)	0.5
(Z_m, Z_v)	0.3
(L, B)	0.3

TABLE 4.2 – Couples du problème

(tous les couples ci-contre sont liés par des copules gaussiennes C_ρ dont le paramètre est précisé, avec les marges indiquées précédemment). À titre indicatif, on donne le graphe de deux densités proches de celles impliquées dans le modèle.



(a) marges $Gum(0,1)$ et $\mathcal{N}(0,1)$, $\rho = 0.5$

(b) marges $\mathcal{T}(0,1/2,1)$, $\rho = 0.3$

FIGURE 4.1 – Deux copules gaussiennes, marges prescrites et ρ fixé

Le but de la suite de ce travail est de déterminer lesquels de ces trois couples sont les plus influents dans les sorties du modèle (ici S et C). La définition même de "influent" recouvre plusieurs réalités, et nécessite a fortiori des outils précis pour être mesurée. C'est l'objet de l'analyse de sensibilité, dont la section suivante rappelle les principes de base et les outils nécessaires dans la présente situation.

Notons que les variables H_d et C_b n'apparaissent pas ici et sont considérées comme indépendantes des autres (l'incertitude sur la hauteur de la digue étant un paramètre d'ordre industriel indépendant de l'environnement, et la hauteur de la berge étant supposée indépendante du niveau de la rivière). Nous éluderons donc ces grandeurs dans l'analyse, pour lesquelles les techniques unidimensionnelles développées dans [13] s'appliquent directement.

4.2 Analyse de sensibilité

4.2.1 Principes de base

L'idée de départ de l'analyse de sensibilité est la suivante. Étant donnée une entrée d-dimensionnelle $\mathbf{X} = (X_1, \ldots, X_d)$, on cherche à décrire une sortie $\mathbf{Y} = f(\mathbf{X})$ donnée par une fonction $f : \mathbf{R}^d \to \mathbf{R}$ assez régulière (parfois désignée comme *modèle*, ou *code de calcul*). En pratique, plusieurs obstacles s'opposent à cette description, par exemple la complexité de la fonction (qui peut n'être que partiellement, voire pas du tout, explicite, ou dont l'évaluation peut être coûteuse) ou le nombre de variables (qui peut atteindre plusieurs milliers, voire plus, dans des systèmes complexes). Il peut être donc commode de considérer les variables d'entrée comme aléatoires et d'étudier d'un point de vue statistique leur influence sur la sortie.

De nombreux indicateurs existent pour quantifier cette influence, dont Iooss fait par exemple une revue dans [48]. Nous nous concentrerons ici sur deux d'entre eux :

les indices de Sobol' (basés sur la décomposition de la variance), et les indices DGSM (Derivative-based Global Sensitivity Measures). Ce sont toutes deux des familles d'indices globaux, qui font intervenir des grandeurs globales pour f (la variance dans un cas, l'intégrale du gradient sur tout l'espace dans l'autre).

Ces deux indices seront détaillés en suivant, mais on rappelle au préalable la décomposition de Hoeffding-Sobol', qui permet de d'écrire de façon plus simple le problème de sensibilité. Cette formule, au cœur de l'ANOVA (ANalysis Of VAriance), permet de décomposer la variance de la sortie en regroupant les entrées de façon à segmenter le problème. Initialement introduite par Hoeffding dans [45], elle est reformulée par Sobol' en 1990 (en russe) et présentée dans l'article en anglais de Saltelli et Sobol' [77] en 1995.

Dans un souci de simplicité, cette décomposition est énoncée dans le cadre de variables mutuellement indépendantes. Elle se généralise à notre contexte en considérant que chaque variable X_i est à valeurs dans $\mathbf{R}^2$, et représente un couple de variables dépendantes.

Proposition 4.2.1 (Décomposition de Hoeffding-Sobol')**.** *La sortie $f(\mathbf{X})$ admet la décomposition suivante :*

$$f(\mathbf{X}) = f_0 + \sum_{i=1}^{d} f_i(X_i) + \sum_{1\leq i<j\leq d} f_{i,j}(X_i, X_j) + \cdots + f_{1,\dots,d}(X_1, \dots, X_d),$$

où f_0 est une constante $\mathbf{E}[f_I(X_I)|X_J] = 0$ pour tous $I \subset [\![1, d]\!]$ et $J \subsetneq I$.

Pour $I = \{i_1, \dots, i_n\} \subset [\![1, d]\!]$, on a noté $X_I = (X_{i_1}, \dots, X_{i_n})$ et $f_I = f_{i_1,\dots,i_n}$. La condition de centrage de l'espérance conditionnelle correspond à une condition d'orthogonalité dans $L^2(\mu)$ (avec μ la loi de $\mathbf{X}$) sur les termes successifs de la décomposition. Elle permet d'en déduire une décomposition de la variance :

$$\mathrm{Var}(f(\mathbf{X})) = \sum_{I\subset[\![1,d]\!]} \mathrm{Var}(f_I(X_I)),$$

qui sera fondamentale dans l'établissement de la première famille d'indices de sensibilité abordée dans la section suivante. Remarquons néanmoins que la somme ci-dessus comprend 2^d termes distincts, ce que peut s'avérer problématique lorsque le nombre de variables d'entrée augmente.

Notons enfin que cette décomposition peut s'écrire par récurrence en termes d'espérances conditionnelles, suivant un principe d'inclusion-exclusion :

- on pose $f_0 = \mathbf{E}[f(\mathbf{X})]$,
- pour $i \in [\![1, d]\!]$,
$$f_i(X_i) = \mathbf{E}[f(\mathbf{X})|X_i] - f_0,$$
- pour tout $I \subset [\![1, d]\!]$,
$$f_I(X_I) = \mathbf{E}[f(\mathbf{X})|X_I] - \sum_{J\subsetneq I} f_J(X_J).$$

Cette dernière écriture fait clairement l'orthogonalité décrite précédemment. Les fonctions partielles f_I peuvent également être définies de cette façon dans le cas de variables dépendantes, ce que nous ferons pour la suite. Notons que ces fonctions dépendent à la fois du modèle de départ f et de la loi de $\mathbf{X}$.

La décomposition de la variance obtenue est au cœur de la construction des indices de Sobol', qui sont décrits dans la section suivante. Les indices DGSM sont abordés ensuite, et leur lien avec les précédents est discuté en fin de section.

4.2.2 Indices globaux et liens avec l'inégalité de Poincaré

On s'intéresse dans cette section à deux grandes catégories d'indices de sensibilité globaux : les indices de Sobol' et les indices DGSM. Leur présentation est faite de façon claire et succinte dans [55], où de nombreuses références additionnelles pourront être trouvées.

La première famille d'indices a été introduite par Sobol' dans l'article en russe susmentionné puis présentée dans [77]. Ils reposent sur la décomposition de la variance donnée précédemment.

Définition 4.2.2 (Indices de Sobol')**.** On pose $D = \mathrm{Var}(f(X))$ la *variance totale* du modèle. Pour tout $I \subset [\![1,d]\!]$, on définit la *variance partielle* $D_I = \mathrm{Var}(f_I(X_I))$. L'*indice de Sobol'* associé à l'ensemble d'indices I est défini par :

$$S_I = \frac{D_I}{D}.$$

L'indice de Sobol' représente la part de variance expliquée part les variables de I dans la variance totale. Cette famille d'indices est exhaustive, dans le sens où

$$\sum_{I \subset [\![1,d]\!]} S_I = 1.$$

Définition 4.2.3 (Indices de Sobol' totaux)**.** Pour tout $I \subset [\![1,d]\!]$, on définit la *variance totale d'interaction* $D_I^T = \sum_{J \supset I} D_J$. L'*indice de Sobol' total* associé à I est défini par :

$$S_I^T = \frac{D_I^T}{D}.$$

Cette famille représente représente la part de variance expliquée par un ensemble d'indices en prenant en compte toutes les interaction avec les autres indices. C'est cette dernière que nous utiliserons dans notre contexte de variables dépendantes. Plusieurs techniques existent pour calculer empiriquement ces indices, notamment des estimations par Monte-Carlo. Nous reviendrons sur ces dernières en fin de chapitre, en application au problème introduit précédemment.

Remarque. Les définitions données ci-dessus sont évidemment valables dans le cas où $I = i$ est un singleton, auquel cas on notera $S_i := S_{\{i\}}$ et $S_i^T := S_{\{i\}}^T$. Notons que, dans ce

cas-là, on peut définir la fonction d'interaction totale de la variable X_i, pour $i \in [\![1, d]\!]$, par :

$$f_i^T(\mathbf{X}) = \sum_{I \ni i} f_I(X_I),$$

qui permet d'exprimer l'indice S_i^T comme

$$S_i^T = \frac{\mathrm{Var}(f_i^T(\mathbf{X}))}{D}.$$

Ces indices sont néanmoins moins adaptés au contexte de variables dépendantes, et ne seront donc pas utilisés par la suite.

Ces indices sont bien globaux, et sont établis dans l'idée que plus une variable est influente, plus elle contribuera à la variance totale du modèle (dans le sens où de grandes variations sur cette grandeur conduiront à de grandes variations sur la sortie). On peut alors définir un seuil sur les indices de Sobol' au-delà duquel un ensemble de variables est considéré comme influent, où en-deçà duquel on le considérera comme négligeable. Comme il sera détaillé en fin de chapitre, ces indices peuvent s'approcher par Monte-Carlo, via notamment la technique de "pick-and-freeze".

Une autre approche, dans le cas où f est différentiable, consiste à observer que si la dérivée partielle du modèle par rapport à une variable donnée admet des valeurs élevées, on peut attendre que de grandes variations de cette variable conduise à de grandes variations sur la sortie. Proposés par Sobol' et Gershman [80], ils ont été largement étudiés depuis (voir [55]), et présentent des liens intéressants avec l'inégalité de Poincaré.

Définition 4.2.4 (Indices DGSM)**.** Supposons que f est différentiable et que pour tout $i \in [\![1, d]\!]$, $\partial_i f(\mathbf{X}) \in L^2(\Omega)$. On définit alors l'*indice DGSM* associé à $i \in [\![1, d]\!]$ par :

$$\nu_i = \mathbf{E}\left[(\partial_i f(\mathbf{X}))^2\right].$$

À l'instar des indices de Sobol', ces derniers peuvent également être calculés par Monte-Carlo. Dans les cas où la fonction f n'est pas (ou pas assez) explicite, et où le gradient n'est pas directement accessible, il peut être nécessaire de passer d'abord par une approximation de ce dernier. Ces points seront également discutés à la fin du chapitre, ainsi qu'une comparaison qualitative entre les deux familles d'indices.

Dans la plupart des cas, les liens entre ces différents indices sont présentés dans le cas de variables indépendantes. Mais, dans le modèle présenté ci-dessus, les variables sont deux à deux dépendantes (selon la structure de dépendance présentée dans la section précédente). Les définitions données nécessitent donc une adaptation pour être utilisables dans cette situation. Essentiellement, pour $I \subset [\![1, d]\!]$, on définit l'indice DGSM suivant :

$$\nu_I = \mathbf{E}\left[\sum_{i \in I} \partial_i f(\mathbf{X})^2\right].$$

Cette transposition est une extension des indices totaux (comme présentés pour les indices de Sobol) aux indices DGSM. De même on étend la définition de fonction d'interaction totale à une classe d'indices $I \subset [\![1,d]\!]$:

$$f_I^T(\mathbf{X}) = \sum_{J \supset I} f_J(X_J).$$

On a alors $D_I^T = \mathrm{Var}(f_I^T(\mathbf{X}))$, par orthogonalité des fonctions de la décomposition. Notons que, par définition des fonctions partielles, on a également

$$D_I^T = D - \mathrm{Var}\left(\mathbf{E}[f(\mathbf{X})|X_{-I}]\right),$$

où $-I$ désigne le complémentaire de I dans $[\![1,d]\!]$. Dans la suite, on considérera des ensembles d'indices I à deux éléments, qui représentent les paires de variables dépendantes.

Si l'on compare les grandeurs en jeu dans les différentes familles d'indices, on ne manquera pas d'observer qu'ils correspondent aux deux membres d'une inégalité de Poincaré. Cette inégalité permet en effet de les relier entre elles, selon les modalités de la proposition suivante. Dans tout ce qui suit, on notera μ la loi de $\mathbf{X}$ et μ_I la loi de X_I pour $I \subset [\![1,d]\!]$.

Proposition 4.2.5. *Soit $I \subset [\![1,d]\!]$ tel que, pour tout $J \subset [\![1,d]\!]$ avec $I \cap J = \emptyset$, X_I est indépendant de X_J. Supposons que $f \in H^1(\mu)$ et que μ_I satisfait une $IP(c_I)$. On a alors*

$$D_I^T \leq c_I \nu_I.$$

Démonstration. L'hypothèse sur I signifie que les variables de X_I sont indépendantes de toutes les autres. Ainsi, si on note μ_{-I} la loi jointe de toutes ces autres variables, notées X_{-I}, on a $\mu = \mu_I \otimes \mu_{-I}$. On a donc, par orthogonalité des fonctions partielles, et d'après le théorème de Fubini-Tonelli, pour toute fonction $g \in H^1(\mu)$ centrée :

$$\begin{aligned}\mathrm{Var}(g(\mathbf{X})) &= \int_{\mathbf{R}^d} g(x)^2 d\mu(x) \\ &= \int_{\mathbf{R}^{d-|I|}} \left(\int_{\mathbf{R}^{|I|}} g(x)^2 d\mu_I(x_I) \right) d\mu_{-I}(x_{-I}).\end{aligned}$$

Par hypothèse sur μ_I, si on note $\nabla_I = (\partial_i)_{i \in I}$, on a :

$$\int_{\mathbf{R}^{|I|}} g^2 d\mu_I(x_I) \leq c_I \int_{\mathbf{R}^{|I|}} |\nabla_I g(x)|^2 d\mu_I(x_I),$$

et donc

$$\begin{aligned}\mathrm{Var}(g(\mathbf{X})) &\leq \int_{\mathbf{R}^{d-|I|}} \left(c_I \int_{\mathbf{R}^{|I|}} |\nabla_I g(x)|^2 d\mu_I(x_I) \right) d\mu_{-I}(x_{-I}) \\ &= c_I \int_{\mathbf{R}^d} |\nabla_I g(x)|^2 d\mu(x).\end{aligned}$$

Enfin, par construction de f_I^T, on a $\nabla_I f = \nabla_I f_I^T$. En effet, $f(x)$ ne dépend de x_I qu'à travers $f_I^T(x)$, et on a donc, pour $g = f_I^T$,

$$D_I^T = \mathrm{Var}(f_I^T(\mathbf{X})) \leq c_I \int_{\mathbf{R}^d} |\nabla_I f_I^T(x)|^2 d\mu(x) = c_I \int_{\mathbf{R}^d} |\nabla_I f(x)|^2 d\mu(x) = c_I \nu_I.$$

■

Ce dernier résultat permet de proposer une méthode de majoration des indices de Sobol' par les DGSM, via la constante de Poincaré des composantes indépendantes des marges. Il n'est néanmoins pas possible de déterminer l'optimalité de la borne obtenue en général, mais elle permet tout de même de détecter les variables les moins influentes, comme cela sera illustré en fin de chapitre. Malheureusement, comme cela a déjà été mentionné, les constantes de Poincaré ne sont en pratique pas connues explicitement. Dans le cadre du modèle décrit ci-dessus, on ne peut pas non plus utiliser les estimations asymptotiques en la dimension de l'espace. On peut néanmoins réduire la complexité en se ramenant à des lois standard, via un changement d'échelle.

Proposition 4.2.6 (Changement d'échelle)**.** *Soit $I \subset [\![1, d]\!]$ telle que la loi de X_I est paramétrée par deux grandeurs m_I et s_I, notée $\mu_I = \mathcal{L}(m_I, s_I)$. Supposons également que s_I est inversible et que $s_I^{-1}(X_I - m_I)$ suit la loi standard $\tilde{\mu}_I = \mathcal{L}(0, 1)$. Supposons enfin que $\tilde{\mu}_I$ satisfait une $IP(\tilde{c}_I)$. Alors il existe une constante $C = C(m_I, s_I)$ telle que :*

$$D_I^T \leq C \tilde{c}_I \nu_I.$$

Démonstration. Notons $\tilde{X}_I = s_I^{-1}(X_I - m_I)$. Par commodité, on notera $\mathbf{X} = (X_I, X_{-I})$, bien que les indices de I ne soient pas toujours successifs. Les transformations affines ne changent pas structure de dépendance de X_I. En effet, la composition d'une fonction de répartition par ces transformations reste une fonction de répartition, donc quitte à remplacer les fonctions de répartition des marges de départ par les versions changées d'échelle, l'expression de la corrélation ne change pas. On a donc, pour toute fonction $g \in H^1(\mu)$:

$$\mathbf{E}[g(\mathbf{X})] = \mathbf{E}\left[g(m_I + s_I \tilde{X}_I, X_{-I})\right].$$

Ainsi, si on note pour toute fonction $g \in H^1(\mu)$ la translatée $\tilde{g}(x_I, x_{-I}) := g(m_I + s_I x_I, x_{-I})$, on a bien :

$$\mathrm{Var}[g(\mathbf{X})] = \mathrm{Var}[\tilde{g}(\tilde{X}_I, X_{-I})].$$

Pour $g = f_I^T$, on a donc :

$$\mathrm{Var}[f_I^T(\mathbf{X})] = \mathrm{Var}[\tilde{f}_I^T(\tilde{X}_I, X_{-I})].$$

L'inégalité de Poincaré appliquée à la nouvelle mesure $\tilde{\mu}$ donne alors la relation :

$$\begin{aligned} D_I^T &= \mathrm{Var}[\tilde{f}_I^T(\tilde{X}_I, X_{-I})] \\ &\leq \int_{\mathbf{R}^{d-|I|}} \left(\tilde{c}_I \int_{\mathbf{R}^{|I|}} |\nabla_I \tilde{f}_I^T(\tilde{x}_I, x_{-I})|^2 d\tilde{\mu}_I(\tilde{x}_I) \right) d\mu_{-I}(x_{-I}) \\ &= \tilde{c}_I \int_{\mathbf{R}^{d-|I|}} \left(\int_{\mathbf{R}^{|I|}} |\nabla_I \tilde{f}_I(\tilde{x}_I, x_{-I})|^2 d\tilde{\mu}_I(\tilde{x}_I) \right) d\mu_{-I}(x_{-I}). \end{aligned}$$

Finalement, en appliquant le changement de variable inverse, on a :

$$\begin{aligned}\int_{\mathbf{R}^{|I|}} |\nabla_I \tilde{f}_I(\tilde{x}_I, x_{-I})|^2 d\tilde{\mu}_I(\tilde{x}_I) &= \int_{\mathbf{R}^{|I|}} \nabla_I f_I(x_I, x_{-I})^T s_I^T s_I \nabla_I f_I(x_I, x_{-I}) d\mu_I(x_I) \\ &\leq \rho_+(s_I)^2 \int_{\mathbf{R}^{|I|}} |\nabla_I f_I(x_I, x_{-I})|^2 d\mu_I(x_I),\end{aligned}$$

où $\rho_+(s_I)$ désigne la plus grande valeur propre (en module) de s_I. On a donc l'estimation :

$$D_I^T \leq \tilde{c}_I \rho_+(s_I)^2 \int_{\mathbf{R}^{d-|I|}} \left(\int_{\mathbf{R}^{|I|}} |\nabla_I f_I(x_I, x_{-I})|^2 d\mu_I(x_I) \right) d\mu_{-I}(x_{-I}) = \tilde{c}_I \rho_+(s_I)^2 \nu_I,$$

et on peut prendre $C = \rho_+(s_I)$. ■

En pratique, s_I est une matrice diagonale dont les coefficients correspondent aux changements d'échelle nécessaires à appliquer à chaque marge pour obtenir la version "réduite". Notons cependant que cette quantité ne correspond pas toujours à la variance de la loi (c'est le cas par exemple de la loi de Gumbel, dont les paramètres ne correspondent respectivement ni à l'espérance ni à la variance). Comme nous le verrons dans la partie suivante, ce résultat est néanmoins très utile pour réduire la complexité du problème.

Remarque. Dans le cas des lois restreintes à un intervalle, le changement d'échelle s'applique également aux bornes de l'intervalle. Ainsi, à la loi unidimensionnelle $\mathcal{L}(m, s)$ restreinte à l'intervalle $\mathcal{I} = [a, b]$ on fera correspondre la loi $\mathcal{L}(0, 1)$ restreinte à l'intervalle

$$\tilde{\mathcal{I}} = \left[\frac{a-m}{s}, \frac{b-m}{s} \right].$$

Dans le cas d'une loi multidimensionnelle restreinte à un pavé de $\mathbf{R}^d$, on appliquera ce changement à chaque intervalle coordonnée par coordonnée en fonction des paramètres des marges.

Notons que la preuve précédente peut être légèrement modifiée dans le cas où $s_I = \text{diag}(s_i, i \in I)$, pour obtenir le résultat plus fin suivant.

Corollaire 4.2.7. *Sous les hypothèses de la proposition précédente, s'il existe* $(s_i)_{i \in I} \subset \mathbf{R}$ *tels que* $s_I = \text{diag}(s_i, i \in I)$*, alors :*

$$D_I^T \leq \tilde{c}_I \sum_{i \in I} s_i^2 \nu_i.$$

Démonstration. La modification intervient au moment de majorer $s_I^T s_I$ par $\rho_+(s_I)^2 I_{|I|}$, où l'on remarque que pour tout $x_I \in \mathbf{R}^{|I|}$:

$$\nabla_I f_I(x_I, x_{-I})^T s_I^T s_I \nabla_I f_I(x_I, x_{-I}) = \sum_{i \in I} s_i^2 \partial_i f(x_I, x_{-I})^2.$$

La fin de la preuve est inchangée. ■

Notons que, dans le cas où les coefficients de normalisation sont très grands, cette borne peut être très loin de l'optimal. L'utilisation de ce changement d'échelle se justifie néanmoins dans un cadre numérique : le calcul de la constante de Poincaré avant normalisation s'avère très difficile, notamment en raison de la taille des domaines d'étude. Il est donc en pratique plus intéressant de procéder à la renormalisation, qui donne par ailleurs des résultats proches de l'optimal dans certains cas (voir la fin du chapitre).

On termine cette introduction à l'analyse de sensibilité par l'énoncé d'une formule de calcul des variances totales présentée par M. Jansen dans [51], bien utile dans les applications numériques. Elle est connue en anglais sous le nom de "pick-and-freeze", que l'on pourrait malhabilement traduire en français par "fixer et regénérer".

Proposition 4.2.8 (Pick-and-freeze). *Soient $I \subset [\![1, d]\!]$ et Z_I une copie indépendante de X_I. On a alors :*

$$D_I^T = \frac{1}{2}\mathrm{Var}\left(f(X_I, X_{-I}) - f(Z_I, X_{-I})\right).$$

Démonstration. La preuve est un calcul direct. On a, par définition de la covariance,

$$\begin{aligned}\mathrm{Var}\left(f(X_I, X_{-I}) - f(Z_I, X_{-I})\right) = \mathrm{Var}(f(X_I, X_{-I})) + \mathrm{Var}(f(Z_I, X_{-I})) \\ - 2\mathrm{Cov}(f(X_I, X_{-I}), f(Z_I, X_{-I})).\end{aligned}$$

Les deux termes de variance ci-dessus sont égaux (car X_I et Z_I ont la même loi) à la variance totale D. Le terme de covariance s'écrit :

$$\mathrm{Cov}(f(X_I, X_{-I}), f(Z_I, X_{-I})) = \mathbf{E}[f(X_I, X_{-I})f(Z_I, X_{-I})] - \mathbf{E}[f(X_I, X_{-I})]\mathbf{E}[f(Z_I, X_{-I})].$$

De même, le produit des deux termes de droite est égal à $\mathbf{E}[f(\mathbf{X})]^2$. Pour utiliser l'indépendance de X_I et Z_I, on peut conditionner par X_{-I} dans l'espérance du produit :

$$\begin{aligned}\mathbf{E}[f(X_I, X_{-I})f(Z_I, X_{-I})] &= \mathbf{E}\Big[\mathbf{E}[f(X_I, X_{-I})f(Z_I, X_{-I})|X_{-I}]\Big] \\ &= \mathbf{E}\Big[\mathbf{E}[f(X_I, X_{-I})|X_{-I}]\mathbf{E}[f(Z_I, X_{-I})|X_{-I}]\Big].\end{aligned}$$

Le deux espérances conditionnelles sont égales, toujours car X_I et Z_I ont même loi. Ainsi, si on note $Y = \mathbf{E}[f(\mathbf{X})|X_{-I}]$, on a :

$$\mathbf{E}[Y] = \mathbf{E}[\mathbf{E}[f(\mathbf{X})|X_{-I}]] = \mathbf{E}[f(\mathbf{X})],$$

et

$$\mathbf{E}[f(X_I, X_{-I})f(Z_I, X_{-I})] = \mathbf{E}[Y^2].$$

Ainsi,

$$\mathrm{Cov}(f(X_I, X_{-I}), f(Z_I, X_{-I})) = \mathbf{E}[Y^2] - \mathbf{E}[Y]^2 = \mathrm{Var}(Y).$$

Finalement,

$$\mathrm{Var}\left(f(X_I, X_{-I}) - f(Z_I, X_{-I})\right) = 2D - 2\mathrm{Var}(Y) = 2D_I^T,$$

ce qui termine la preuve. ■

On comprend bien alors, à la lumière de la preuve, l'expression "pick-and-freeze" : un sous-ensemble de variables est choisi ("pick") et regénéré indépendamment, les autres sont fixées ("freeze"), et sont utilisées pour conditionner. Le principal atout de cette formule est de permettre d'exprimer les indices de Sobol' non en termes de variances conditionnelles, mais en termes de variances classiques, faciles à estimer par Monte-Carlo.

Cette facilité d'estimation peut sembler invalider la méthode d'estimation via les indices DGSM et l'inégalité de Poincaré. Mais les estimateurs en jeu dans la méthode pick-and-freeze nécessitent deux échantillons indépendants de la loi de $\mathbf{X}$, alors que l'estimation des DGSM par Monte-Carlo n'en nécessite qu'un seul (voir la définition des estimateurs en toute fin de chapitre). Les deux estimateurs demandent le même nombre d'appels à la fonction f, ce qui rend les indices DGSM préférables en situation de parcimonie, par exemple quand les échantillons sont obtenus suite à une expérience physique ou numérique coûteuse.

Dans la partie suivante, on présente une méthode numérique de calcul de ces constantes en petite dimension, basée sur des techniques d'éléments finis. La méthode est dans la droite ligne de l'article de Barthe, Iooss et Roustant en dimension 1 [13], qui sert d'introduction à ces techniques.

4.3 Aspects numériques de l'inégalité de Poincaré

4.3.1 Principe général

On rappelle la définition de l'inégalité de Poincaré (cf Définition 1.2.2). On considère ici μ une mesure de Boltzmann-Gibbs sur $\mathbf{R}^d$ associée à un potentiel V et à un générateur $\mathbf{L}$. L'inégalité s'énonce comme suit.

Définition 4.3.1 (Inégalité de Poincaré)**.** La mesure μ satisfait une *inégalité de Poincaré* de constante $c > 0$, notée $IP(c)$, si pour toute fonction f assez régulière,

$$\mathrm{Var}_\mu(f) \leq c \int_{\mathbf{R}^d} |\nabla f|^2 d\mu.$$

Dans tout ce qui suit, on considérera des fonctions de régularité au moins H^1. Cet espace permet de définir correctement les problèmes variationnels sans restreindre trop l'espace des fonctions test (par rapport à $L^2(\mu)$).

On rappelle que cette inégalité est, via sa constante optimale, reliée à un problème aux valeurs propres. Dans toute la suite, on supposera que l'opérateur $\mathbf{L}$ est semi-défini négatif. Si μ satisfait une inégalité de Poincaré, elle admet un trou spectral (Proposition 1.2.3), on peut donc chercher alors le plus petit $\lambda \in \mathbf{R}_+^*$ tel que :

$$\exists f \in H^1(\mu),\ -\mathbf{L}f = \lambda f.$$

Notons que l'hypothèse de trou spectral n'est pas suffisante pour garantir l'existence de fonctions propres (voir le Chapitre 3). Nous supposerons donc dans la suite que cette

condition est satisfaite (c'est par exemple le cas quand le spectre de $-\mathbf{L}$ est discret), ce qui sera vérifié dans les applications ultérieures. Ce problème peut alors se réécrire sous la forme :

$$\exists u \in H^1(\mu) : \forall v \in H^1(\mu),\ \langle \nabla u, \nabla v \rangle = \lambda \langle u, v \rangle.$$

En effet, supposons que $u \in H^1(\mu)$ est solution du problème précédent. Alors une intégration par parties dans le terme de gauche donne

$$\forall v \in H^1(\mu),\ \langle -\mathbf{L}u, v \rangle = \langle \lambda u, v \rangle,$$

soit

$$\forall v \in H^1(\mu),\ \langle -\mathbf{L}u - \lambda u, v \rangle = 0,$$

et donc

$$-\mathbf{L}u - \lambda u = 0 \quad \mu\text{-}p.s.,$$

ce qui équivaut bien au problème de départ. Il est commode dans ce cadre de définir les opérateurs (bi-)linéaires en jeu.

Définition 4.3.2 (Formulation faible)**.** On définit les formes bilinéaires $a : H^1(\mu) \times H^1(\mu) \to \mathbf{R}$ et $l : H^1(\mu) \times H^1(\mu) \to \mathbf{R}$, pour tous $u, v \in H^1(\mu)$, par

$$a(u,v) = \langle \nabla u, \nabla v \rangle + \langle u, v \rangle, \text{ et } l(u,v) = \langle u, v \rangle.$$

On s'intéresse alors au problème suivant :

$$\exists u \in H^1(\mu) : \forall v \in H^1(\mu),\ a(u,v) = (\lambda + 1) l(u,v).$$

Notons que, dans le cas où μ est supportée sur un domaine borné Ω (assez régulier), on supposera des conditions de Neumann sur le bord du domaine : pour tout $x \in \partial\Omega$,

$$\nabla f(x) \cdot \vec{n}(x) = 0,$$

où $\vec{n}(x)$ représente la normale au bord au point x. Cette hypothèse permet d'éliminer les termes de bord dans l'intégration par parties précédente. Cette équivalence est alors analogue à celle présentée dans [13] en dimension 1. Mentionnons enfin que d'autres conditions aux bords (Dirichlet, Robin, etc.) ne garantissent pas que $\lambda_0(-\mathbf{L}) = 0$, et ne correspondent donc pas en ce sens au présent contexte.

Cette formulation possède alors un avantage majeur : elle permet de résoudre le problème aux valeurs propres par des méthodes d'éléments finis. Cette méthode se base sur le principe suivant :

1. mailler le domaine Ω par un nombre fini d'éléments eux-mêmes finis,
2. construire un espace de fonctions test v et un espace de fonctions candidates u, tous deux de dimension finie,
3. calculer a et l sur ces espaces,
4. résoudre le problème aux valeurs propres.

Le gain réside ici dans le fait de se ramener à un problème en dimension finie, qui peut être traité matriciellement, comme il sera détaillé dans la section suivante. Il est néanmoins à noter que chacune de ces étapes est non-triviale. L'outil utilisé pour cette résolution est FEniCS [1], une interface regroupant des outils de résolution numérique des EDP, du maillage à la résolution. Nous reviendrons en détails dessus dans la suite. Avant cela, présentons rapidement les résultats obtenus précédemment en dimension 1.

A parte I : sur dimension 1

L'étude en dimension 1 fait l'objet du travail de Barthe, Iooss et Roustant [13], dont ce travail est la continuation. Leur article s'intéresse aux inégalités de Poincaré sur des intervalles (et donc sous des conditions au bord de Neumann). Une des premières applications de leurs résultat réside dans l'étude des lois présentées plus tôt dans le problème de sensibilité. En effet, les lois de Gumbel et la loi gaussienne sont tronquées, et il est alors pertinent d'essayer d'estimer la constante de Poincaré de ces nouvelles mesures. Notons que l'étude unidimensionnelle est pertinente dès que les variables sont supposées indépendantes.

Certains résultats exacts sont obtenus, notamment pour la gaussienne, en restriction à certains intervalles. À titre d'exemple, la valeur de la constante de Poincaré est retrouvée explicitement en restriction à des intervalles de la forme $[r_i^{(n)}, r_{i+1}^{(n)}]$, où $r_i^{(n)}$ désigne la i-ième racine du polynôme d'Hermite de degré n sur $\mathbf{R}$ (voir le chapitre précédent). La restriction à un intervalle $[a, b]$ générale est discutée, mais la valeur n'est pas explicite. Notons enfin que les fonctions qui saturent cette inégalité (qui réalisent l'égalité) sont également identifiées.

Afin d'obtenir des résultats explicites sur les constantes du problème (lorsqu'elles ne sont pas connues analytiquement), une méthode par éléments finis est proposée. Le principe en est assez simple : réduire un problème de dimension infinie (sur $H^1(\mu)$) à un problème de dimension finie. Pour cela, l'intervalle d'intérêt I est découpé en N sous-intervalles disjoints de même taille $(I_n)_{n=1}^N$ (avec un pas $h = |I|/N$), sur chacun desquels on définit une fonction g_n. Sur l'intervalle $I = [0, 1]$, ces fonctions sont définies par :

$$g_n(x) = \left(\frac{x}{h} + 1 - n\right) \mathbf{1}_{I_{n-1}}(x) + \left(-\frac{x}{h} + 1 + n\right) \mathbf{1}_{I_n}(x).$$

Chaque g_n est une fonction "chapeau" continue, croissante sur I_{n-1} et décroissante sur I_n, nulle partout ailleurs et égale à 1 en nh. Cette définition s'adapte par translation à tout intervalle compact. On note :

$$G_h = \text{Vect}(g_n, n = 1, \dots, N),$$

qui est un espace de dimension finie. On peut alors résoudre numériquement le problème aux valeurs propres sur cet espace, dont la solution converge quand $h \to 0$ vers la solution du problème initial. Tous les détails de cette démarche sont donnés dans la

1. https ://fenicsproject.org/

section suivante. Cette méthode permet alors aux auteurs d'obtenir de nouvelles estimations de la constante de Poincaré, à savoir :

- loi $\mathcal{T}(-1, 1/2, 1) : c_P(\mu) \simeq 0.173$,
- loi $\mathcal{N}(0,1)$ sur $[-1.87, +\infty[: c_P(\mu) \simeq 0.892$,
- loi $Gum(0,1)$ sur $[-0.92, 3.56] : c_P(\mu) \simeq 1.021$.

Ces grandeurs ont été calculées pour des loi standards sur des intervalles renormalisés, en utilisant l'analogue de la Proposition 4.2.6 en dimension 1.

Le but de la suite de cette section est de présenter la version 2D de cette approche, en rentrant un peu plus dans le détail des méthodes numériques.

4.3.2 Généralités sur les éléments finis en dimension deux

Les définitions et propriétés générales de cette section peuvent être retrouvées dans l'ouvrage de référence d'Allaire [1], §6.3. La première porte sur le domaine d'étude la définition de maillage. On considérera ici des maillages triangulaires, de loin les plus classiques.

La plupart des concepts sont présentés sur des ouverts *polyédriques*. Ce sont des ouverts bornés dont le bord est constitué d'un nombre fini de segments tels que l'intersection deux segments distincts est soit l'ensemble vide soit réduit à une extrémité commune.

Définition 4.3.3 (Maillage)**.** Soit Ω un ouvert connexe polyédrique de $\mathbf{R}^2$. Un *maillage triangulaire* de $\overline{\Omega}$ est un ensemble $\mathcal{T}_h$ de triangles non plats (appelés cellules) du plan $(K_i)_{i=1}^n$ tels que :

1. pour tout $i \in [\![1, n]\!]$, $K_i \subset \Omega$ et $\overline{\Omega} = \bigcup_{i=1}^n K_i$,
2. si $i \neq j$, $K_i \cap K_j$ est soit vide, soit réduit à un sommet commun, soit à une arête commune entière.

On appelle *finesse du maillage* le réel $h = \max_{i=1,\dots,n} \operatorname{diam}(K_i)$ et *taille du maillage* l'entier $|\mathcal{T}_h| := n$.

Pour mailler des domaines non polyédriques du plan, on commence par construire une approximation affine par morceaux du bord telle que le nouveau domaine ainsi délimité est un ouvert polyédrique. Cette approximation du bord est possible dès que ce dernier est assez régulier. On parlera alors par abus de langage de *maillage triangulaire de* Ω.

Une fois le maillage défini, il est nécessaire de définir l'espace des fonctions test en relation à ce dernier. On utilise ici la méthode des éléments finis de Lagrange d'ordre 1. Comme son nom le laisse penser, elle repose sur des polynômes de degré 1. On notera donc dans la suite $\mathbf{R}_1[x_1, x_2]$ l'espace vectoriel des fonctions polynomiales de degré au plus un à deux variables.

Définition 4.3.4. Soit Ω un ouvert borné de $\mathbf{R}^2$ et $\mathcal{T}_h$ un maillage triangulaire de Ω. On définit l'*espace des fonctions test de Lagrange d'ordre 1* comme

$$\mathcal{V}_h = \{v \in \mathcal{C}^0(\overline{\Omega}) \ : \ \forall K \in \mathcal{T}_h, v_{|K} \in \mathbf{R}_1[x_1, x_2]\}.$$

On appelle *nombre de degrés de liberté* de la méthode le nombre de sommets (ou nœuds) de $\mathcal{T}_h$, noté n_{dl}.
On appelle enfin *degrés de liberté* d'une fonction $v \in \mathcal{V}_h$ l'ensemble des valeurs possibles pour v aux nœuds du maillage.

Pour une triangulation donnée $\mathcal{T}_h$, il n'est pas difficile de voir que l'espace $\mathcal{V}_h$ est toujours de dimension finie. En effet, chaque fonction est complètement décrite par au plus $3|\mathcal{T}_h|$ monômes (3 par cellule, car $\dim(\mathbf{R}_1[x_1, x_2]) = 3$). Ainsi, il existe une base $(\varphi_i)_{i=1}^{n_{dl}}$ de $\mathcal{V}$ telle que pour tous $i, j \in [\![1, n_{dl}]\!]$,

$$\varphi_i(t_j) = \delta_{ij},$$

où $(t_i)_{i=1}^{d}$ désigne l'ensemble des sommets de $\mathcal{T}_h$. Toute fonction $v \in \mathcal{V}_h$ se décompose alors comme :

$$v = \sum_{i=1}^{n_{dl}} v(t_i)\varphi_i.$$

Il s'ensuit alors que $\mathcal{V}_h$ est un sous-espace de dimension finie de $H^1(\Omega)$ (voir [1] Lemme 6.3.3 et Proposition 6.3.7 pour plus de détails). En pratique, les fonctions $(\varphi_i)_{i=1}^{n_{dl}}$ sont l'équivalent bidimensionnel des fonctions chapeaux évoquées précédemment. Ici, chaque fonction φ_i est une petite pyramide dont le sommet est t_i, les arêtes les côtés des triangles dont t_i est un sommet et dont la base est constituée de la troisième arête des triangles précités (et qui est nulle partout ailleurs, voir [1] p.181).

On peut alors réécrire le problème variationnel tel que présenté dans la Définition 4.3.2. On cherche alors $u_h \in \mathcal{V}_h$ de la forme

$$u_h = \sum_{i=1}^{n_{dl}} u_h(t_i)\varphi_i,$$

et le test contre toute fonction $v_h \in \mathcal{V}_h$ se réduit à un test contre toutes les fonctions de base. On a alors la formulation suivante :

$$\forall i \in [\![1, n_{dl}]\!], \quad \sum_{j=1}^{n_{dl}} u_h(t_j)a(\varphi_j, \varphi_i) = (\lambda + 1)\sum_{j=1}^{n_{dl}} u_h(t_j)l(\varphi_j, \varphi_i).$$

Pour faire bien voir l'aspect fini-dimensionnel du nouveau problème, on définit la *matrice de rigidité* $\mathcal{K}_h$ comme :

$$\mathcal{K}_h = \big(a(\varphi_i, \varphi_j)\big)_{i,j=1}^{n_{dl}},$$

et la *matrice de masse* M_h par :

$$M_h = \big(l(\varphi_i, \varphi_j)\big)_{i,j=1}^{n_{dl}},$$

toutes deux symétriques (et dont les noms viennent de la mécanique des structures). Si on note enfin $U_h = (u_h(t_j))_{j=1}^{n_{dl}}$ le vecteur des degrés de liberté de u_h, le problème se réécrit :

$$\mathcal{K}_h U_h = (\lambda + 1)M_h U_h.$$

On est donc ramené à un problème aux valeurs propres dit généralisé (de par la présence de M_h) qui porte sur des matrices en dimension n_{dl}. Notons que cette mise en œuvre peut être appliquée à des formes bilinéaires a et l tout à fait générales.

En pratique, un tel problème n'est pas résolu directement. Néanmoins, dans le cas où la matrice de masse est définie positive, on peut calculer sa décomposition de Cholesky $M_h = L_h L_h^T$, et en posant $\widetilde{\mathcal{K}}_h = L_h^{-1}\mathcal{K}_h L_h^{-T}$ et $\widetilde{U}_h = L_h^T U_h$, on se ramène au problème aux valeurs propres classique :

$$\widetilde{\mathcal{K}}_h \widetilde{U}_h = (\lambda + 1)\widetilde{U}_h.$$

La matrice $\widetilde{\mathcal{K}}_h$ est toujours symétrique, on peut donc calculer sa décomposition en base orthonormée, ce qui donne λ et le $\widetilde{U}_h$ correspondant. La solution au problème de départ est alors retrouvée en inversant le changement de variable. Les preuves de convergence lorsque h devient petit font appel à de nombreuses notions annexes qui ne seront pas traitées ici, mais pourront être trouvées en détails dans [1], §6.3.2.

Avant de poursuivre l'étude, plusieurs remarques peuvent être soulevées. D'une part, les fonctions de base ont un support petit devant la taille du domaine dès que le maillage est assez fin. Il en résulte que, en pratique, les matrices $\mathcal{K}_h$ et M_h sont creuses. Des techniques numériques spécifiques existent pour traiter ces cas, ce qui peut parfois accélérer la résolution (voir [1], §13.1). Les erreurs numériques peuvent néanmoins poser problème lors du calcul de la décomposition de Cholesky de M_h (non définie pour une matrice singulière). Cela peut être corrigé au besoin par l'adjonction d'un petite perturbation dans M_h qui l'empêche d'être singulière.

Le maillage lui-même peut poser des problèmes de régularité. En effet, les éléments de base de la décomposition en éléments finis dégénèrent si un triangle du maillage est plat (cas où un sommet se situe au milieu d'une arête). Bien que ce cas soit exclu de fait de la définition de maillage, la déformation d'un maillage déjà existant peut amener à de telles situations. Plus précisément, on peut définir les fonctions de V_h sur des éléments triangulaires de référence, puis les transporter sur le maillage réel via un difféomorphisme. Un triangle plat dans le maillage réel se traduit alors par un changement de variable non inversible, et donc une dégénérescence de la fonction de base associée à cet élément. Il faut donc s'assurer de la régularité du maillage, comme cela sera développé dans la suite.

Enfin, le choix du domaine à mailler est sujet à discussions. La deuxième loi indiquée en Table 4.1 n'est pas à support compact, et il faut alors décider d'un moyen de tronquer son domaine, de sorte à rendre possible la résolution par éléments finis. En lien avec le caractère creux des matrices de rigidité et de masse, une méthode est proposée qui permet également de traiter des problèmes de régularité rencontrés quand les densités s'annulent sur le bord du domaine.

Dans les sections suivantes, nous présentons l'application de tous ces principes au problème de départ. La résolution est séparée en trois parties : d'abord une présentation du choix du domaine et du maillage, ensuite une validation de la méthode sur un cas explicite, enfin l'implémentation.

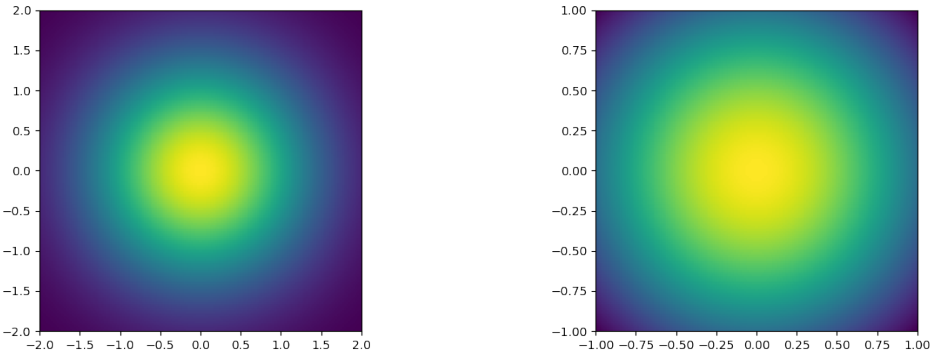
4.3.3 Domaine et maillage

Régions de prédiction

Pour illustrer le raisonnement, les concepts seront d'abord exprimés pour le cas d'un vecteur (X_1, X_2) gaussien centré réduit de matrice de covariance Γ_ρ, pour $\rho \in [0, 1[$. Une première approche pour réduire le domaine d'analyse à un ouvert borné est de tronquer à une valeur fixée arbitrairement, et ce coordonnée par coordonnée.

Définition 4.3.5. On appelle *troncature locale de niveau a* de la loi μ la restriction de μ au carré $[-a, a]^2$.

Pour la loi gaussienne par exemple, dans le cas indépendant, on obtient les histogrammes de densité ci-après.

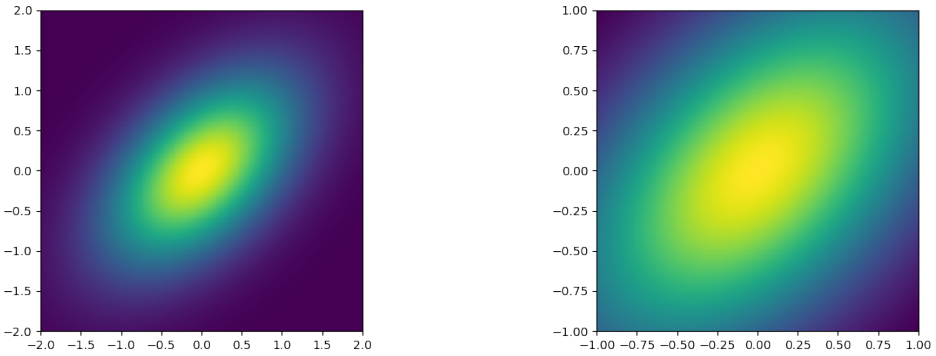


(a) $a = 2$ (b) $a = 1$

FIGURE 4.2 – Densité d'un vecteur gaussien indépendant centré réduit restreint à $[-a, a]^2$

Une couleur plus claire indique une plus haute valeur de la densité.

On observe alors que la densité est nulle (ou presque) sur une partie importante du domaine dans le premier cas. Cela s'explique par la concentration de la loi gaussienne autour de l'origine. La valeur de la densité intervenant dans le calcul de $\mathcal{K}_h$ et M_h, des valeurs trop faibles risqueraient d'aggraver les problèmes liés à la dégénérescence des ces matrices. Une solution serait alors de restreindre à un carré plus petit, comme dans le second cas. Si cette solution peut sembler pertinente dans ce cas précis (vis-à-vis de la valeur de la densité), elle n'est malheureusement pas très robuste. Par exemple pour deux variables corrélées, on obtient pour les mêmes valeurs de a :



(a) $a = 2$ (b) $a = 1$

FIGURE 4.3 – Densité d'un vecteur gaussien de covariance 0.5 centré réduit restreint à $[-a, a]^2$

Le phénomène précédemment décrit s'est aggravé, et la seconde restriction est insuffisante pour y pallier complètement. Ce problème devient de plus en plus important quand ρ se rapproche de 1, quand la masse se concentre sur la diagonale du plan. Il est donc préférable de trouver une méthode de restriction qui s'adapte à la forme de la densité et donne un degré de liberté moins arbitraire qu'une simple troncature par coordonnées. On introduit à cette fin les régions de prédiction.

Définition 4.3.6 (Région de prédiction)**.** Une *région de prédiction* de niveau $\alpha \in]0,1[$ pour le vecteur aléatoire (X_1, X_2) est une région du plan D_α telle que :

$$\mathbf{P}((X_1, X_2) \in D_\alpha) = \alpha.$$

Notons que cette définition diffère de celle d'une région *de confiance*, qui serait basée sur un échantillon du vecteur aléatoire et non sur la densité a priori. Peu d'hypothèses géométriques sont requises dans la précédente définition, bien qu'il existe certaines régions de prédiction particulières : celles données par les lignes de niveau de la densité jointe. S'il existe, pour une région de prédiction D_α donnée pour le vecteur aléatoire (X_1, X_2), une ligne de niveau $\{f_{(X_1,X_2)} = c\}$ telle que $\partial D_\alpha = \{f_{(X_1,X_2)} = c\}$, on appellera cette dernière une *ligne de prédiction* de niveau α.

Ces lignes sont bien connues pour la loi gaussienne, elles sont données par des ellipses.

Proposition 4.3.7. *Soit* $\mathbf{X} = (X_1, X_2)$ *un vecteur gaussien centré de matrice de covariance* $\Gamma_\rho = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$, $\rho \in [0,1[$. *Soit* $\alpha \in]0,1[$ *un niveau de prédiction fixé et* $q(\alpha)$ *le quantile de la loi* $\chi^2(2)$ *associé. On définit la région de prédiction* D_α *dont le bord est donné par l'ellipse :*

$$E_\alpha = \{(x, y) \in \mathbf{R}^2 : x^2 + y^2 - 2\rho xy = (1 - \rho^2)q(\alpha)\}.$$

Alors on a bien $\mathbf{P}(\mathbf{X} \in D_\alpha) = \alpha$.

Démonstration. Comme $\mathbf{X}$ est un vecteur $\mathcal{N}_2(0,\Gamma)$, le vecteur $\Gamma_\rho^{-1/2}\mathbf{X}$ suit la loi $\mathcal{N}_2(0,I_2)$, donc d'après le théorème de Cochran $|\Gamma_\rho^{-1/2}\mathbf{X}|^2 \sim \chi^2(2)$. Ainsi, si on note $q(\alpha)$ le quantile du chi-deux à deux degrés de libertés de niveau α, on a :

$$\mathbf{P}(\mathbf{X}^T\Gamma_\rho^{-1}\mathbf{X} \leq q(\alpha)) = \alpha.$$

L'inverse de Γ_ρ est donnée par :

$$\Gamma_\rho^{-1} = \frac{1}{1-\rho^2}\begin{pmatrix} 1 & -\rho \\ -\rho & 1 \end{pmatrix},$$

et l'équation $\mathbf{X}^T\Gamma_\rho^{-1}\mathbf{X} \leq q(\alpha)$ se réécrit alors :

$$\mathbf{X}^T\begin{pmatrix} 1 & -\rho \\ -\rho & 1 \end{pmatrix}\mathbf{X} - q(\alpha)(1-\rho^2) \leq 0.$$

La matrice Γ_ρ^{-1} est symétrique définie positive, l'équation ci-dessus définit donc un domaine D_α fermé dont le bord est une ellipse d'équation :

$$x^2 + y^2 - 2\rho xy = (1-\rho^2)q(\alpha).$$

Les directions propres de Γ_ρ^{-1} sont les droites d'équation $x = y$ et $x = -y$, qui portent donc les demi-axes de l'ellipse. En prenant $x = y$ dans l'équation ci-dessus, on obtient $x = \pm\sqrt{\frac{1+\rho}{2}q(\alpha)}$, ce qui donne une longueur pour le demi grand axe de $\sqrt{(1+\rho)q(\alpha)}$. De même, pour $x = -y$, on obtient une longueur du demi petit axe égale à $\sqrt{(1-\rho)q(\alpha)}$. L'ensemble E_α est donc bien une ellipse caractérisée ci-dessus délimitant une région D_α telle que

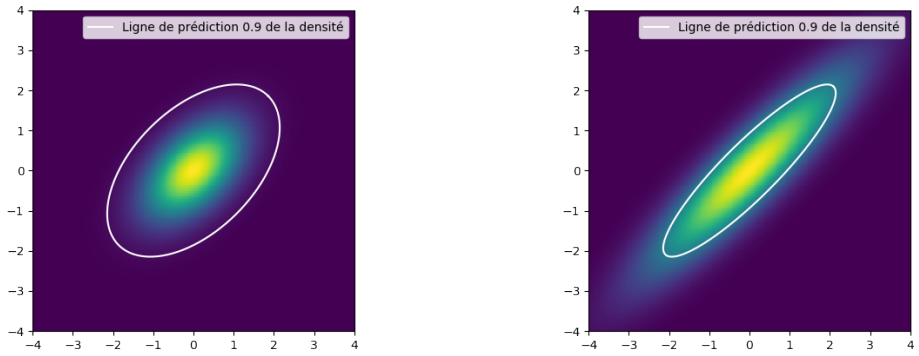
$$\mathbf{P}(X \in D_\alpha) = \alpha.$$

■

Toujours dans le cas gaussien, ces ellipses peuvent être paramétrées assez simplement :

$$\begin{cases} x(t) = \sqrt{\frac{q(\alpha)}{2}}\left(\sqrt{1+\rho}\cos t - \sqrt{1-\rho}\sin t\right) \\ y(t) = \sqrt{\frac{q(\alpha)}{2}}\left(\sqrt{1+\rho}\cos t + \sqrt{1-\rho}\sin t\right) \end{cases}, \; t \in [0, 2\pi[.$$

On représente ci-dessous deux de ces ellipses au niveau $\alpha = 0.9$ pour deux valeurs de ρ différentes.



(a) $\rho = 0.5$ (b) $\rho = 0.9$

FIGURE 4.4 – Lignes de prédiction de niveau 0.9

On observe bien la concentration autour de la diagonale pour de grandes valeurs de ρ, ce qui encourage l'utilisation de telles régions de prédiction en lieu et place des troncatures. Notons que les limites en $\rho \to 0$ et $\rho \to 1$ permettent de retrouver dans un premier cas l'expression d'un disque (la densité d'un vecteur gaussien centré indépendant étant invariante par rotation) et dans un second cas l'équation de la droite $y = x$, correspondant au cas dégénéré évoqué plus tôt.

Afin de généraliser ces régions de prédiction, il est nécessaire de pouvoir passez du vecteur gaussien précédent à un vecteur aux marges prescrites et de copule gaussienne. Considérons le cas d'un vecteur $\mathbf{X} = (X_1, X_2)$ de copule gaussienne de covariance ρ et de marges données par les fonctions de répartitions F_1 et F_2 (de dérivée respective f_1 et f_2). On rappelle que la densité jointe de $\mathbf{X}$ s'écrit :

$$f_{\mathbf{X}}(x_1, x_2) = \frac{1}{\sqrt{1-\rho^2}} \exp\Big[-\frac{1}{2(1-\rho^2)} \Big(\rho^2 [\Phi^{-1}(F_1(x_1))^2 + \Phi^{-1}(F_2(x_2))^2] \\ - 2\rho \Phi^{-1}(F_1(x_1)) \Phi^{-1}(F_2(x_2)) \Big) \Big] f_1(x_1) f_2(x_2),$$

où Φ désigne la fonction de répartition de la loi $\mathcal{N}(0,1)$. Pour une telle loi de $\mathbf{X}$ donnée, on connaît toujours une transformation qui permet de se ramener à un vecteur gaussien (et donc en préservant la copule). C'est l'objet du lemme suivant.

Lemme 4.3.8. *Soit* $\mathbf{X}$ *le vecteur défini précédemment et* $\mathbf{Y} = (Y_1, Y_2)$ *un vecteur gaussien centré de covariance* ρ*. On définit la transformation :*

$$\begin{array}{rccc} \Psi : & \mathbf{R}^2 & \to & \mathbf{R}^2 \\ & (x,y) & \mapsto & (F_1^{-1} \circ \Phi(x), F_2^{-1} \circ \Phi(y)) \end{array}.$$

Alors $\mathbf{X}$ *et* $\Psi(\mathbf{Y})$ *ont la même loi.*

Cette transformation se rencontre parfois dans la littérature sous le nom de *transformation de Nataf* [65], et constitue un outil classique de changement de variable marginal entre deux lois de probabilité.

Démonstration. La preuve est directe. Considérons $g : \mathbf{R}^2 \to \mathbf{R}$ une fonction continue bornée. Pour alléger les calculs, on notera :

$$c(x,y) = \frac{1}{\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2(1-\rho^2)}(\rho^2(x^2+y^2) - 2\rho xy)\right],$$

et f_G la densité de la loi $\mathcal{N}(0,1)$. La densité de $\mathbf{Y}$ est donc donnée, pour tout $(y_1, y_2) \in \mathbf{R}^2$, par :

$$f_Y(y_1,y_2) = c(y_1,y_2) f_G(y_1) f_G(y_2).$$

On calcule alors $\mathbf{E}[g(\Psi(\mathbf{Y}))]$:

$$\mathbf{E}[g(\Psi(\mathbf{Y}))] = \int_{\mathbf{R}^2} g(\Psi(y_1,y_2)) c(y_1,y_2) f_G(y_1) f_G(y_2) dy_1 dy_2$$

On effectue alors le changement de variable $(x_1, x_2) = \Psi(y_1, y_2)$. Les hypothèses sur les lois en jeu assurent que Ψ est un C^1 difféomorphisme de l'intérieur du support de la loi de $\mathbf{X}$ dans $\mathbf{R}^2$ d'inverse donné par :

$$\Psi^{-1}(x_1,x_2) = \left(\Phi^{-1}(F_1(x_1)), \Phi^{-1}(F_2(x_2))\right)$$

et de jacobien :

$$\begin{pmatrix} \frac{F_1'(x_1)}{\Phi' \circ \Phi^{-1}(F_1(x_1))} & 0 \\ 0 & \frac{F_2'(x_2)}{\Phi' \circ \Phi^{-1}(F_2(x_2))} \end{pmatrix} = \begin{pmatrix} \frac{f_1(x_1)}{f_G \circ \Phi^{-1}(F_1(x_1))} & 0 \\ 0 & \frac{f_2(x_2)}{f_G \circ \Phi^{-1}(F_2(x_2))} \end{pmatrix}$$

Le changement de variable donne donc :

$$\begin{aligned} \mathbf{E}[g(\Psi(\mathbf{Y}))] &= \int_{\mathbf{R}^2} g(x_1,x_2) c(\Psi^{-1}(x_1,x_2)) f_G(\Phi^{-1}(F_1(x_1))) f_G(\Phi^{-1}(F_2(x_2))) \times \\ &\qquad \times \frac{f_1(x_1) f_2(x_2)}{f_G(\Phi^{-1}(F_1(x_1))) f_G(\Phi^{-1}(F_2(x_2)))} dx_1 dx_2 \\ &= \int_{\mathbf{R}^2} g(x_1,x_2) c(\Psi^{-1}(x_1,x_2)) f_1(x_1) f_2(x_2) dx_1 dx_2 \\ &= \int_{\mathbf{R}^2} g(x_1,x_2) f_{\mathbf{X}}(x_1,x_2) dx_1 dx_2, \end{aligned}$$

d'après la formule de $f_{\mathbf{X}}$ rappelée ci-dessus. On a donc bien l'égalité en loi entre $\mathbf{X}$ et $\Phi(\mathbf{Y})$. ■

Un corollaire direct de cette proposition est un résultat de transport, qui permet de déformer une région de prédiction gaussienne vers une région de prédiction pour n'importe quel couple de variables aléatoires à copule gaussienne.

Corollaire 4.3.9. *Soit* $\mathbf{X}$ *un vecteur aléatoire de copule gaussienne de covariance* ρ *et* $\mathbf{Y}$ *le vecteur gaussien défini précédemment. Soit* $\alpha \in]0,1[$ *et* D_α *une région de prédiction de niveau* α *pour* $\mathbf{Y}$ *(telle que définie dans la Proposition 4.3.7). Alors* $\Psi(D_\alpha)$ *est une région de prédiction de niveau* α *pour* $\mathbf{X}$.

Démonstration. D'après le lemme précédent, $\mathbf{X}$ et $\Psi(\mathbf{Y})$ ont même loi et Ψ est un difféomorphisme, donc

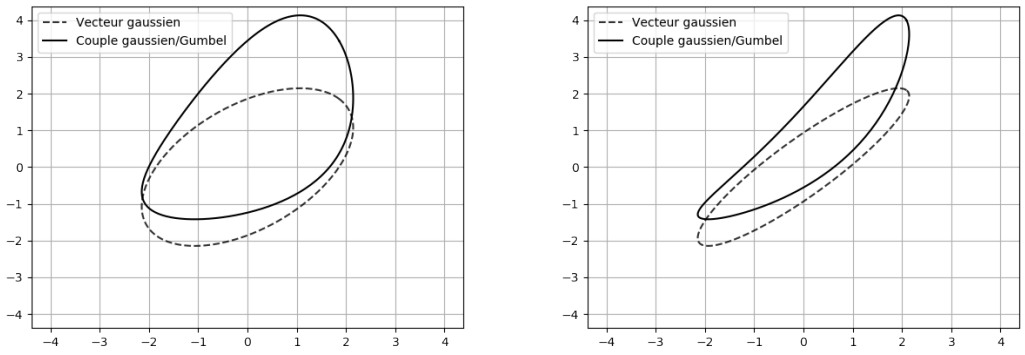
$$\alpha = \mathbf{P}(\mathbf{Y} \in D_\alpha) = \mathbf{P}(\Psi(\mathbf{Y}) \in \Psi(D_\alpha)) = \mathbf{P}(\mathbf{X} \in \Psi(D_\alpha)),$$

ce qui clôt la preuve. ■

À titre d'exemple, on compare les régions de prédiction à niveau fixé obtenues pour la loi $\mathcal{N}_2(0, \Gamma_\rho)$ et le couple $\mathbf{X} = (X_1, X_2)$ de mêmes copule et covariance, où $X_1 \sim \mathcal{N}(0,1)$ et $X_2 \sim Gum(0,1)$. La fonction de répartition de la loi de Gumbel centrée réduite est donnée par $x \mapsto \exp(-e^{-x})$. La fonction Ψ dans ce cas-là est donc donnée par :

$$\Psi(x_1, x_2) = (x_1, -\log(-\log(\Phi(x_2)))).$$

On obtient, pour un niveau 0.9 fixé et deux covariances différentes, les régions suivantes.



(a) $\rho = 0.5$ (b) $\rho = 0.9$

FIGURE 4.5 – Régions de prédiction de niveau $\alpha = 0.9$

Avant de détailler les points relatifs au maillage, notons que cette approche par régions de prédiction est également valable pour les lois à support compact. En effet, le calcul numérique de la densité jointe peut poser problème au voisinage du bord du domaine. Si X_1 est à support dans un intervalle $[a, b]$, alors on a exactement $F_1(a) = 0$ et $F_1(b) = 1$, ce qui entraîne que $\Phi \circ F_1$ n'est pas définie en ces points (sauf à lui donner une valeur infinie). Afin de pallier à ce problème, il peut être nécessaire de laisser une marge entre le bord du domaine réel et le bord du domaine de travail. La

méthode par région de prédiction a l'avantage de proposer un nouveau domaine qui prend en compte la corrélation des marges et dont le bord est assez régulier (car image d'une ellipse par un $\mathcal{C}^1$-difféomorphisme). À titre d'exemple, le régions de prédiction de niveau $\alpha = 0.95$ pour un couple de lois triangulaires standard de copule gaussienne pour différentes valeurs de ρ sont présentées ci-dessous.

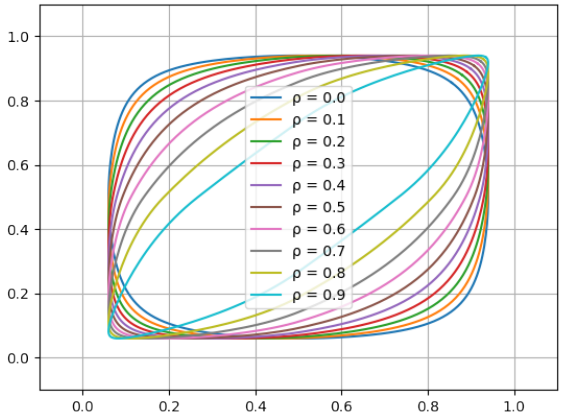


FIGURE 4.6 – Régions de prédiction de niveau 0.95 pour différentes valeurs de ρ (lois triangulaires)

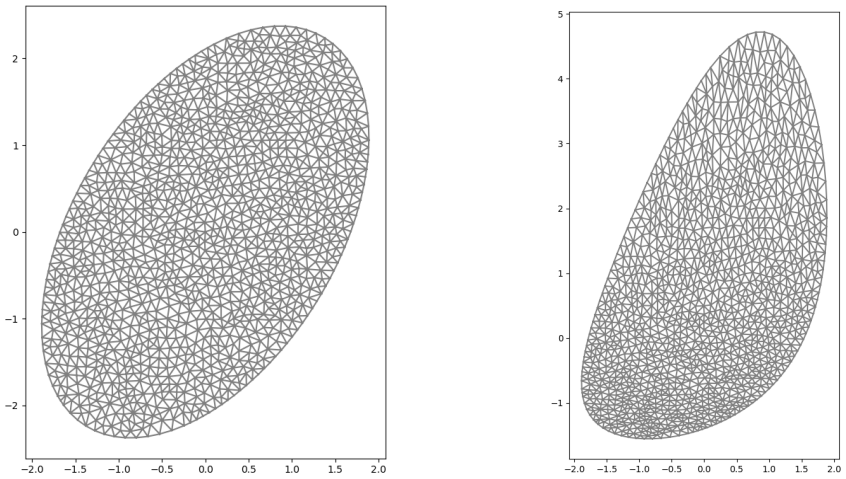
La forme des régions de prédiction ci-dessus diffère beaucoup de celle du support de la loi, qui reste le carré $[0,1]^2$. La méthode proposée permet donc même dans le cas de lois à support compact de réduire drastiquement la taille du domaine étudié, et donc une concentration efficace dans les régions du plan où la densité charge le plus.

Maintenant que les domaines d'étude ont été définis, il est nécessaire de mailler ces derniers pour continuer.

Maillage

Dans toute cette partie, les concepts seront illustrés sur un vecteur de marges gaussienne et Gumbel standards, de copule gaussienne avec covariance $\rho = 0.5$, sans restriction *a priori* de la loi à un domaine compact particulier.

FEniCS comprend la plupart des outils permettant de résoudre un problème par éléments finis du début à la fin. Malheureusement, le mailleur qu'il propose n'est pas très performant lorsque les domaines de travail ont des formes complexes. La première approche développée pour traiter la région de prédiction d'un vecteur non gaussien fut de mailler une ellipse (donnée comme région de prédiction d'un vecteur gaussien), puis de déformer le maillage obtenu via le difféomorphisme Ψ introduit précédemment. Il en résulte alors un maillage le la région de prédiction d'intérêt. À titre d'illustration, voilà le résultat obtenu pour le couple gaussien/Gumbel illustré précédemment.



(a) Vecteur gaussien (b) Vecteur gaussien/Gumbel

FIGURE 4.7 – Maillage des régions de prédiction de niveau $\alpha = 0.9$, $\rho = 0.5$

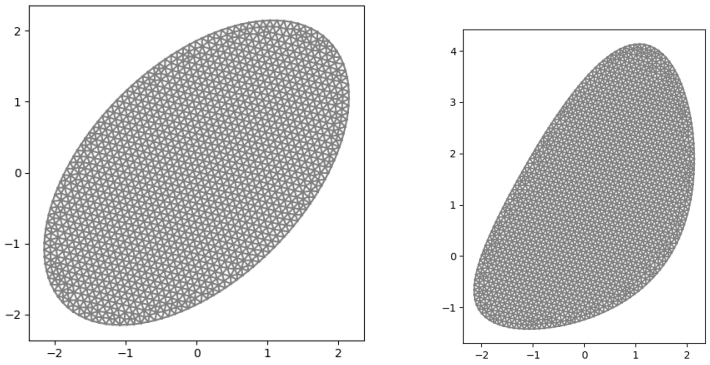
Si le premier maillage semble correct, on observe de fortes déformations dans le second. Comme cela a déjà été abordé, les éléments résultants peuvent poser des problèmes de définition des fonctions de base des éléments finis et donc de résolution numérique sur ce domaine.

Choix a donc été fait de se tourner vers un autre mailleur, bien connu des mécaniciens [2] : GMSH. Logiciel libre de maillage par éléments finis, GMSH s'intègre très bien dans un environnement FEniCS pour mailler des domaines complexes. L'algorithme utilisé est dit "Frontal-Delaunay", basé sur la triangulation de Delaunay [30] et l'algorithme de Bowyer-Watson [22, 88], tel qu'introduit par Rebay dans [75].

Ce mailleur peut prendre en entrée l'équation paramétrée du bord et une grandeur représentant la taille des éléments. Il procède ensuite d'abord à un maillage 1D du bord, avant de mailler l'intérieur. La taille des éléments est choisie en étudiant la stabilité du résultat obtenu en fonction de ce paramètre, comme cela sera abordé plus tard.

Nous ne rentrerons pas plus dans le détail des algorithmes de triangulation, et donnerons seulement, toujours à titre d'illustration, les deux maillages des domaines précédents avec le nouveau mailleur. La taille des éléments est prise à 0.1, pour $\alpha = 0.9$ et $\rho = 0.1$.

2. un grand merci à Morgane et Robin !



(a) Vecteur gaussien (b) Vecteur gaussien/Gumbel

FIGURE 4.8 – Maillage GMSH des régions de prédiction

Dans les deux cas, le maillage obtenu est bien plus régulier. Cette régularité peut s'observer de façon plus précise sur le conditionnement des matrices, et plus particulièrement celui de la matrice de rigidité. À cette fin, on trace ci-dessous le conditionnement de cette dernière, pour $\alpha = 0.9$ et $\rho = 0.5$ fixés, en fonction du nombre de sommets du maillage (en échelle semi-logarithmique).

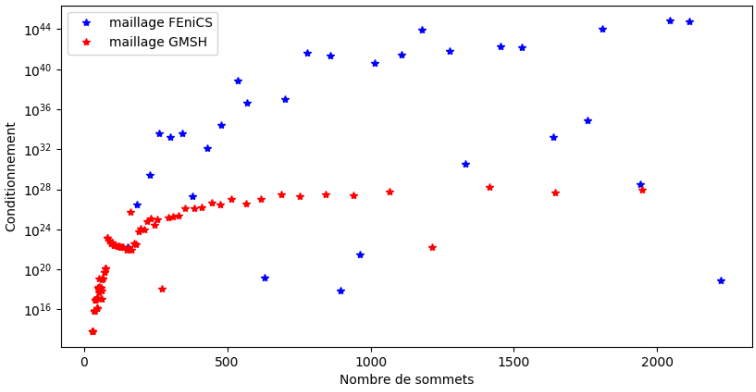
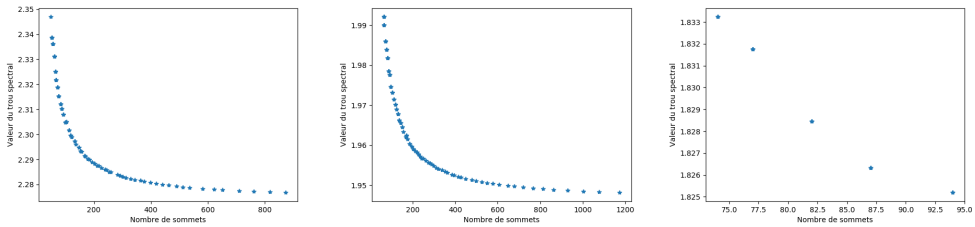


FIGURE 4.9 – Conditionnement de la matrice de rigidité en fonction des éléments

On observe dans les deux cas que plus le nombre d'éléments est grand, plus le conditionnement est élevé. C'est un résultat attendu, car un domaine maillé finement donnera une matrice de rigidité plus grande, et donc potentiellement plus creuse que pour un maillage plus grossier. Néanmoins, le conditionnement est bien plus important pour le maillage FEniCS que pour le maillage GMSH, et de plusieurs ordres de grandeurs. L'intuition visuelle est donc confirmée par l'expérience, ce qui valide l'utilisation de ce dernier.

On ne peut malheureusement pas se contenter d'un maillage grossier, le conditionnement de la matrice de rigidité n'étant pas un objectif en soi. En effet, on attend comme cela a déjà été mentionné que le résultat obtenu par éléments finis converge quand la finesse du maillage tend vers 0. Il est donc a priori nécessaire de trouver un équilibre entre la taille du maillage et la précision attendue. Mais si l'augmentation du nombre de sommets complexifie la résolution, il en va de même avec le paramètre α. En effet, ce dernier détermine la taille de la région d'intérêt, et l'approximation tendra à être d'autant meilleure que α est proche de 1. Les expériences numériques montrent que, dans le présent contexte, la valeur seuil $\alpha = 0.75$ ne semble pas pouvoir être dépassée sans entraîner un défaut de l'algorithme de résolution. On trace ci-dessous, pour différentes valeurs de α et $\rho = 0.5$, la valeur calculée du trou spectral en fonction de la taille du maillage.



(a) $\alpha = 0.65$ (b) $\alpha = 0.75$ (c) $\alpha = 0.8$

FIGURE 4.10 – Valeur du trou spectral en fonction de la taille du maillage

On observe bien dans les deux premiers cas une décroissance qui semble se stabiliser. Ce n'est pas le cas pour $\alpha = 0.8$, où la taille du domaine empêche de mettre en œuvre un maillage trop fin et d'observer une convergence. La valeur $\alpha = 0.75$ semble représenter le meilleur compromis entre la taille du domaine et la finesse permise pour le maillage. Pour le confirmer, on peut tracer une courbe similaire en poussant le maillage plus loin.

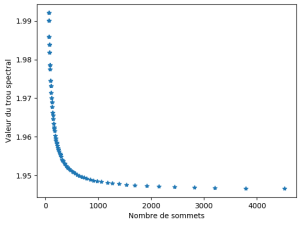
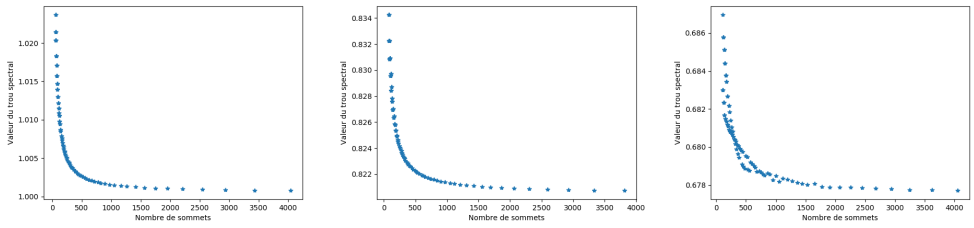


FIGURE 4.11 – Valeur du trou spectral en fonction de la taille du maillage, $\alpha = 0.75$

Ce second graphe conforte la convergence observée précédemment. Ce type de graphe permet d'établir un seuil au-delà duquel la valeur calculée de la constante de

Poincaré n'évolue pratiquement plus, seuil qui pourra être utilisé dans les calculs pratiques.

Cette démarche, opérée ici sur le cas du couple gaussien/Gumbel standard (non restreint a priori) peut être répétée pour tous les cas à traiter selon cette méthode. Notons que des cas différents peuvent donner un niveau α maximal et un seuil de convergence très différents. À titre de comparaison, on donne ci-dessous trois graphes de convergence pour trois valeurs de α différentes dans le cadre du couple gaussien/Gumbel évoqué en 4.3.5.



(a) $\alpha = 0.8$ (b) $\alpha = 0.9$ (c) $\alpha = 0.99$

FIGURE 4.12 – Valeur du trou spectral en fonction de la taille du maillage

On observe dans les trois cas une stabilisation autour de 2500 sommets, pour des valeurs de α proches de 1. La troncature sur la loi borne le domaine de travail et rend possible les calculs pour des valeurs élevées de α. Le dernier graphique montre une certaine instabilité pour les maillages les plus grossiers, qui se corrige quand le nombre de mailles est assez important.

Notons ici une double décroissance du trou spectral calculé : en α à taille du maillage fixée et en la taille du maillage à α fixé. Les deux sont attendues, du fait de la formulation de Courant-Fischer des valeurs propres : augmenter α ou la taille du maillage revient à calculer un infimum sur un espace plus gros, ce qui résulte en une moindre valeur du trou spectral.

Avant de la mettre en œuvre sur les données du problème de sensibilité, il faut néanmoins valider les résultats obtenus par l'algorithme. Pour ce faire, on se base sur un cas explicite : le vecteur gaussien.

4.3.4 Validation de la méthode

La valeur théorique des constantes de Poincaré recherchées est la plupart du temps inconnue. Il est donc difficile de savoir si la valeur vers laquelle converge l'algorithme est proche de la réalité. Il existe néanmoins une famille de lois pour lesquelles la valeur optimale de cette constante est connue, qui vont nous permettre de vérifier les résultats : les vecteurs gaussiens.

Proposition 4.3.10. *Soit $\gamma_\rho = \mathcal{N}_2(0, \Gamma_\rho)$ pour un certain $\rho \in [0, 1[$. Alors $\lambda_1(\gamma_\rho) = \dfrac{1}{1+\rho}$.*

Démonstration. Ce résultat repose sur la théorie de Bakry-Émery telle que présentée dans les chapitres précédents. On rappelle que la densité de γ_ρ est donnée par :

$$f_\rho(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2(1-\rho^2)}(x^2+y^2-2\rho xy)\right],$$

soit $f_\rho(x,y) \propto e^{-V_\rho(x,y)}$, avec $V_\rho(x,y) = \frac{1}{2(1-\rho^2)}(x^2+y^2-2\rho xy)$. La hessienne de V_ρ est donnée par Γ_ρ^{-1}, et ses valeurs propres sont $\frac{1}{1+\rho}$ et $\frac{1}{1-\rho}$. D'après la Proposition 1.2.6, on a donc $\lambda_1(\gamma_\rho) \geq \frac{1}{1+\rho}$. Pour avoir l'égalité, il suffit d'exhiber une fonction propre du générateur $-\mathbf{L}_\rho$ associé. Ce générateur est défini, pour une fonction f assez régulière, par :

$$-\mathbf{L}_\rho f = -\Delta f + \nabla V_\rho \cdot \nabla f,$$

soit, en remplaçant V_ρ par son expression,

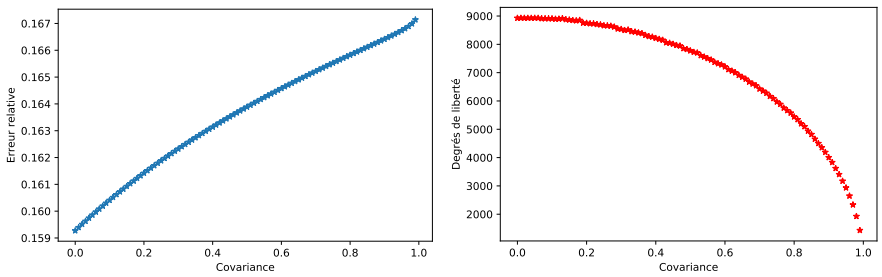
$$-\mathbf{L}_\rho f(x,y) = -\Delta f(x,y) + \frac{1}{1-\rho^2}\begin{pmatrix} x-\rho y \\ y-\rho x\end{pmatrix} \cdot \nabla f(x,y), \quad (x,y) \in \mathbf{R}^2.$$

Dans le cas indépendant, les fonctions propres associées à λ_1 sont les polynômes de degré 1, et vues les directions propres de la matrice de covariance, la fonction $(x,y) \mapsto x+y$ est un bon candidat. En effet, on a

$$\begin{aligned} -\mathbf{L}_\rho(x+y) &= -\Delta(x+y) + \frac{1}{1-\rho^2}\begin{pmatrix} x-\rho y \\ y-\rho x\end{pmatrix} \cdot \nabla(x+y) \\ &= \frac{1}{1-\rho^2}(x-\rho y+y-\rho x) \\ &= (x+y)\frac{1-\rho}{1-\rho^2} = \frac{1}{1+\rho}(x+y). \end{aligned}$$

Ainsi, l'inégalité ci-dessus est saturée pour cette fonction, ce qui implique que le trou spectral est exactement égal à $\frac{1}{1+\rho}$. ■

Le principal intérêt de cet exemple est qu'il fournit une infinité de cas pour valider le résultat sur différentes valeurs de ρ. On représente ci-dessous deux graphes en fonction de ρ, le premier de l'erreur relatif sur la méthode, le second du nombre de degrés de liberté du maillage (à titre indicatif). On a pris $\alpha = 0.95$.



(a) Erreur relative (b) Degrés de liberté

FIGURE 4.13 – Erreur relative de la méthode

La taille des éléments et le niveau de prédiction étant fixés pour tous les calculs, on observe une diminution du nombre de degrés de liberté lorsque ρ augmente. Cela se traduit par une augmentation de l'erreur relative, sans que l'on puisse dire si celle-ci est imputable à la valeur de la corrélation ou uniquement à la diminution du nombre de sommets du maillage. Pour séparer ces deux problèmes, il peu paraître plus judicieux de se concentrer sur une ou deux valeurs de ρ (par exemple 0.3 et 0.5, celles présentes dans le modèle), et d'observer l'erreur obtenue en fonction de α et du nombre de degrés de liberté. Les résultats sont donnés dans les deux tableaux suivants.

α \ n_{dl}	500	1700	6600	18000
0.9	29.8%	29.7%	29.6%	29.6%
0.95	16.3%	16.3%	16.2%	16.2%
0.99	4.36%	4.34%	4.33%	4.33%

(a) $\rho = 0.3$

α \ n_{dl}	500	1700	6600	18000
0.9	30.1%	30.0%	29.9%	29.9%
0.95	16.5%	16.4%	16.4%	16.4%
0.99	4.40%	4.37%	4.37%	4.36%

(b) $\rho = 0.5$

TABLE 4.3 – Erreur relative de la méthode pour le vecteur gaussien en fonction des paramètres du modèle

On observe d'abord que, toutes choses égales par ailleurs, l'erreur relative est très légèrement plus importante dans le cas $\rho = 0.5$. Ces différences se retrouvent pour toutes les valeurs de α et n_{dl}. Elles sont néanmoins relativement faibles en regard de la valeur de l'erreur. On observe également que, dans les deux cas, l'erreur diminue sensiblement lorsque n_{dl} augmente (bien que la différence soit, encore une fois, assez minime). Ainsi, la valeur de la corrélation et le nombre de degrés de liberté du maillage participent conjointement, mais dans une faible mesure, à la variation de l'erreur numérique.

Les différences les plus notables dans la valeur de l'erreur sont néanmoins celles observées pour différentes valeurs de α. Ainsi, une augmentation de l'ordre de 10% sur α entraîne une division de l'erreur par un facteur au mois 6. On veillera ainsi, dans la

mise en œuvre de la méthode, à privilégier les valeurs de α les plus faibles, quitte à opter pour des maillages un peu plus grossiers.

A parte II : approche analytique de la constante de Poincaré

Avant d'appliquer la méthode d'éléments finis développée ci-dessus au problème d'analyse de sensibilité, on présente un résultat intermédiaire qui vise à établir une estimation analytique de la constante de Poincaré d'un couple de variables aléatoires en fonction de leur corrélation. On se place dans le cadre de la section précédente, avec $\mathbf{X}$ un vecteur aléatoire de copule gaussienne de covariance ρ de marges F_1 et F_2, et $\mathbf{Y}$ un vecteur gaussien centré de covariance Γ_ρ. On note $S_{\mathbf{X}}$ le support de la loi de $\mathbf{X}$ et $\mathring{S}_{\mathbf{X}}$ son intérieur.

On définit alors deux fonctions G_1 et G_2 : pour tout $t \in \mathbf{R}$,

$$G_1(t) = \Phi^{-1} \circ F_1(t) \quad \text{et} \quad G_2(t) = \Phi^{-1} \circ F_2(t).$$

Ces deux fonctions interviennent dans le changement de marges tel que décrit dans le Lemme 4.3.8. Plus précisément, la fonction de transport inverse Ψ^{-1} s'exprime, pour tout $(x_1, x_2) \in \mathring{S}_{\mathbf{X}}$, par :

$$\Psi^{-1}(x_1, x_2) = (G_1(x_1), G_2(x_2)).$$

La précision de l'intérieur du support assure que la fonction Ψ définie plus haut est bien un difféomorphisme. Sous ces hypothèses, entendu que toutes les fonctions de répartition en jeu sont continues, Ψ induit en fait un $\mathcal{C}^1$-difféomorphisme *global* de $\mathbf{R}^2$ dans $\mathring{S}_{\mathbf{X}}$. On obtient alors le résultat suivant.

Théorème 4.3.11. *Sous les hypothèses ci-dessus, on a pour tout $\rho \in [0, 1[$:*

$$c_P(\mathbf{X}) \leq (1+\rho) \sup_{\mathbf{R}^2} \{\max(I, II)\},$$

où I et II sont deux fonctions réelles données, pour tout $(x, y) \in \mathbf{R}^2$, par :

$$I(x) = \frac{f_G\left(G_1(x)\right)^2}{f_1(x)^2} \mathbf{1}_{\mathring{S}_{\mathbf{X}}}(x, y), \quad II(y) = \frac{f_G\left(G_2(y)\right)^2}{f_2(y)^2} \mathbf{1}_{\mathring{S}_{\mathbf{X}}}(x, y).$$

Notons que I et II ne dépendent pas de ρ. De plus, la définition de I et II est donnée à l'intérieur du support de la loi de $\mathbf{X}$, et arbitrairement fixée à 0 en-dehors (où le numérateur et le dénominateur sont tous deux nuls).

Démonstration. Ce résultat repose essentiellement sur l'application du Lemme 4.3.8 : les couples $\mathbf{X}$ et $\Psi(\mathbf{Y})$ ont même loi. Il s'ensuit que pour toute fonction $f : \mathbf{R}^2 \to \mathbf{R}$ assez régulière, on a :

$$\mathrm{Var}(f(\mathbf{X})) = \mathrm{Var}(f \circ \Psi(\mathbf{Y})).$$

Ainsi, l'inégalité de Poincaré pour la gaussienne (qui découle de la Proposition 4.3.10) donne :

$$\mathrm{Var}(f(\mathbf{X})) \leq (1+\rho)\mathbf{E}[|\nabla(f \circ \Psi)(\mathbf{Y})|^2].$$

Le gradient de la fonction ci-dessus est donné par :

$$\nabla(f \circ \Psi) = \mathrm{Jac}(\Psi)\nabla f \circ \Psi,$$

dont il suit :

$$\mathbf{E}[|\nabla(f \circ \Psi)(\mathbf{Y})|^2] = \mathbf{E}[|\mathrm{Jac}(\Psi)(\mathbf{Y})\nabla f(\Psi(\mathbf{Y}))|^2].$$

Pour se ramener à une inégalité de Poincaré pour $\mathbf{X}$, on doit le faire apparaître dans le membre de droite ci-dessus. En utilisant le Lemme 4.3.8 de nouveau, on a :

$$\begin{aligned}\mathbf{E}[|\mathrm{Jac}(\Psi)(\mathbf{Y})\nabla f(\Psi(\mathbf{Y}))|^2] &= \mathbf{E}[|\mathrm{Jac}(\Psi)(\Psi^{-1}(\mathbf{X}))\nabla f(\mathbf{X})|^2] \\ &= \mathbf{E}[\nabla f(\mathbf{X})^T \mathrm{Jac}(\Psi)(\Psi^{-1}(\mathbf{X}))^T \mathrm{Jac}(\Psi)(\Psi^{-1}(\mathbf{X}))\nabla f(\mathbf{X})].\end{aligned}$$

Comme vu dans la preuve du lemme sus-mentionné, le jacobien de Ψ est diagonal, et on a alors :

$$\mathrm{Jac}(\Psi)^T \mathrm{Jac}(\Psi) = \mathrm{Jac}(\Psi)^2 = \mathrm{diag}\left(\frac{f_G^2}{(f_1 \circ G_1^{-1})^2}, \frac{f_G^2}{(f_2 \circ G_2^{-1})^2}\right),$$

et donc par définition de Ψ^{-1} :

$$\mathrm{Jac}(\Psi)^2 \circ \Psi^{-1} = \mathrm{diag}\left(\frac{(f_G \circ G_1)^2}{f_1^2}, \frac{(f_G \circ G_2)^2}{f_2^2}\right).$$

On a ainsi pour presque tout $\mathbf{X}$:

$$\mathrm{Jac}(\Psi)^2(\Psi^{-1}(\mathbf{X})) = \begin{pmatrix} I(X_1) & 0 \\ 0 & II(X_2) \end{pmatrix},$$

et conséquemment :

$$\begin{aligned}\nabla f(\mathbf{X})^T \mathrm{Jac}(\Psi)^2(\Psi^{-1}(\mathbf{X}))\nabla f(\mathbf{X}) &\leq \max(I(X_1), II(X_2))|\nabla f(\mathbf{X})|^2 \\ &\leq \sup_{\mathbf{R}^2} \max(I, II)|\nabla f(\mathbf{X})|^2.\end{aligned}$$

Cela donne pour la variance :

$$\mathrm{Var}(f(\mathbf{X})) \leq (1+\rho) \sup_{\mathbf{R}^2} \max(I, II)\mathbf{E}[|\nabla f(\mathbf{X})|^2],$$

et on a ainsi l'estimation annoncée :

$$c_P(\mathbf{X}) \leq (1+\rho) \sup_{\mathbf{R}^2} \max(I, II).$$

■

Le comportement en ρ de la borne obtenue dépend uniquement du choix de la copule. Notons néanmoins que les fonctions I et II dépendent également de ce choix, mais pas de la valeur de la corrélation. On a alors pour le trou spectral :

$$\lambda_1(\mathbf{X}) \geq \frac{C(\mathbf{X})}{1+\rho},$$

où $\lambda_1(\mathbf{X})$ désigne par abus de notation la première valeur propre non nulle de l'opérateur de diffusion associé à la loi de $\mathbf{X}$ et $C(\mathbf{X}) = \sup_{\mathbf{R}^2} \max(I, II)^{-1}$. La dépendance en ρ semble alors complètement encodée par la copule, indépendamment du choix des marges, et la décroissance en ce paramètre est cohérente avec les expériences numériques. Notons enfin que cette borne est optimale : si les deux marges sont gaussiennes, les deux fonctions I et II sont constantes égales à 1, et on retrouve la valeur du trou spectral indiquée dans la Proposition 4.3.10.

On peut également revenir sur la conjecture énoncée à la fin de l'introduction. Si on pose $g(y_1, y_2) = y_1 + y_2$ pour $(y_1, y_2) \in \mathbf{R}^2$, on a d'une part :

$$\mathrm{Var}(g(\mathbf{Y})) = \mathrm{Var}(Y_1) + \mathrm{Var}(Y_2) + 2\mathrm{Cov}(Y_1, Y_2) = 2(1+\rho),$$

et d'autre part :

$$\mathbf{E}[|\nabla g(\mathbf{Y})|^2] = \mathbf{E}[2] = 2.$$

Le trou spectral gaussien implique que $c_P(\mathbf{Y}) \leq (1+\rho)$, et le calcul précédent donne :

$$\mathrm{Var}(g(\mathbf{Y})) = 2(1+\rho) = (1+\rho)\mathbf{E}[|\nabla g(\mathbf{Y})|^2].$$

L'inégalité $IP(1+\rho)$ est donc saturée par g, et on en déduit donc que $c_P(\mathbf{Y}) = 1+\rho$. Notons que l'on peut retrouver ce résultat par tensorisation dans le cas indépendant (voir la Proposition 1.2.5).

On ne peut malheureusement pas être aussi affirmatif dans le cas général, bien que la décroissance de la borne du Théorème 4.3.11 et le comportement en ρ observé numériquement dans la partie précédente soient en faveur de la conjecture.

Le principal écueil de ce résultat est la valeur de la constante $C(\mathbf{X})$. Elle n'est bien-sûr pas explicite, et son estimation est difficile à obtenir. Dans le cas d'un couple gaussien/Gumbel, les expériences numériques sont en accord avec la borne résultante, bien que cette dernière (obtenue également numériquement et de façon assez sommaire) soit loin de l'optimal. C'est d'autant plus vrai dans le cas particulier $\rho = 0$ (marges indépendantes), où l'on sait déjà par tensorisation (voir la Proposition 1.2.5) que $\lambda_1(\mathbf{X}) \geq 1$. La borne donnée dans ce cas-là par le théorème précédent est malheureusement bien plus faible (de l'ordre de 10^{-2}), de sorte que l'estimation n'apporte ici pas d'information supplémentaire.

Dans le cas de marges à support compact (par exemple pour les couples de lois triangulaires), il n'est pas évident que les fonctions I et II soient bornées. Si elles ne le sont pas, la constante $C(\mathbf{X})$ est nulle, et la borne obtenue sur le trou spectral est alors triviale. Une façon d'assurer la bornitude de I et II est de demander que la densité

de chaque marge soit uniformément minorée par une constante strictement positive sur $S_{\mathbf{X}}$ (ce qui est par exemple le cas pour les lois uniformes, mais pas pour les lois triangulaires).

On peut discuter à ce titre du cas de deux marges uniformes sur $[0,1]$. Les deux fonctions I et II coïncident, et on a pour tout $x \in [0,1]$:

$$I(x) = f_G(\Phi^{-1}(x))^2.$$

et donc

$$\sup_{\mathbf{R}^2} \max(I, II) = \sup_{x\in[0,1]} f_G(\Phi^{-1}(x))^2.$$

Ce supremum est atteint pour $\Phi^{-1}(x) = 0$ (soit $x = 1/2$) et vaut $1/2\pi$. On a ainsi $C(\mathbf{X}) = 2\pi$, et la borne du théorème devient, pour tout $\rho \in [0,1[$:

$$\lambda_1(\mathbf{X}) \geq \frac{2\pi}{1+\rho}.$$

Pour $\rho = 0$, on sait par tensorisation que $\lambda_1(\mathbf{X}) \geq \pi^2$ (voir par exemple [31] ou [8] §4.5.2). La borne obtenue par le Théorème 4.3.11 est donc plus proche de l'optimal que dans le cas de la loi de Gumbel, mais n'apporte malheureusement pas d'information nouvelle. Ce cas est jusqu'à maintenant le seul, avec celui des marges gaussiennes, pour lequel toutes les grandeurs en jeu sont connues et peuvent être comparées avec précision.

Comprendre le comportement de I et II de façon fine serait d'un grand intérêt dans l'utilisation du Théorème 4.3.11, notamment pour fournir des estimations utiles dans la validation de la méthode présentée dans ce chapitre.

4.3.5 Implémentation

La méthode développée ci-dessus permet d'effectuer le réglage de l'algorithme pour les couples présents dans le problème de sensibilité. Les résultats de trou spectral obtenus sont présentés dans le tableau ci-dessous, avec la valeur de chaque paramètre. Les marges sont ramenées à leurs paramètres standards.

Marges	Covariance ρ	Niveau α	Degrés de liberté	Trou spectral
Gumbel/Gaussienne	0.5	0.995	4829	0.67
Triangulaires	0.3	0.995	10178	18.1

TABLE 4.4 – Valeur approchée du trou spectral pour les couples du modèle

Le premier couple est restreint a priori au domaine $\left[-\frac{15}{8}, +\infty\right[\times \left[-\frac{513}{558}, \frac{1987}{558}\right]$, et le second au carré $[0,1]^2$. Le temps d'exécution est de 139.5s pour le premier couple et de 181.0s pour le second.

Il reste maintenant à calculer la constante de normalisation présente dans la Proposition 4.2.6. Pour cela, il faut identifier les transformations en jeu qui permettent de ramener chacun des cas au cas standard.

Dans la suite, on notera $\mathbf{X} = (X_1, X_2, X_3, X_4, X_5, X_6) = (Q, K_s, Z_m, Z_v, L, B)$. Les hypothèses formulée précédemment indiquent que, pour $I = \{1,2\}$, $I' = \{3,4\}$ et $I'' = \{5,6\}$, les couples X_I, $X_{I'}$ et $X_{I''}$ sont mutuellement indépendants. On est donc bien dans le cadre d'application des Propositions 4.2.5 et 4.2.6.

Couple (Q, K_s)

Les marges initiales de $X_I = (Q, K_s)$ sont respectivement $Gum(1013, 558)$ et $\mathcal{N}(30, 64)$. La transformation de changement d'échelle se fait donc via les grandeurs :

$$m_I = \begin{pmatrix} 1013 \\ 30 \end{pmatrix} \text{ et } s_I = \begin{pmatrix} 558 & 0 \\ 0 & 8 \end{pmatrix},$$

avec le changement de domaine évoqué plus haut. Ainsi, dans ce cas comme dans les suivants, la matrice s_I entre dans le cadre d'application du Corollaire 4.2.7. De plus, la Proposition 1.2.3 permet d'estimer la constante de Poincaré $c_{\bar{I}}$ du couple standard en fonction du trou spectral approché. On en déduit donc :

$$D_I^T \leq \frac{1}{0.67}(558^2 \nu_1 + 8^2 \nu_2).$$

Couple (Z_m, Z_v)

Les marges initiales de $X_{I'} = (Z_m, Z_v)$ sont toutes deux $\mathcal{T}(49, 50, 51)$. La fonction de répartition $F_{a,c,b}$ de la loi triangulaire $\mathcal{T}(a, c, b)$ est de plus donnée par l'expression :

$$F_{a,c,b}(x) = \begin{cases} 0 & \text{si } x \leq a \\ \dfrac{(x-a)^2}{(b-a)(c-a)} & \text{si } a < x \leq c \\ 1 - \dfrac{(b-x)^2}{(b-a)(b-c)} & \text{si } c < x \leq b \\ 1 & \text{si } x > b. \end{cases}$$

Si $c = (a+b)/2$, on en déduit que $\dfrac{X-a}{b-a} \sim \mathcal{T}(0, 1/2, 1)$. La transformation de changement d'échelle se fait donc via les grandeurs :

$$m_{I'} = \begin{pmatrix} 49 \\ 49 \end{pmatrix} \text{ et } s_{I'} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

Ainsi, on a $\rho_+(s_{I'}) = 2$ (ici, il n'y a pas de différence entre la Proposition 4.2.6 et le Corollaire 4.2.7). On en déduit donc d'après le trou spectral ci-dessus :

$$D_{I'}^T \leq \frac{2^2}{18.1} \nu_{I'}.$$

Couple (L, B)

Ce cas est très similaire au précédent, bien que les deux marges de $X_{I''} = (L, B)$ ne présentent pas les mêmes paramètres. La première suit la loi $\mathcal{T}(4990, 5000, 5010)$, et la seconde la loi $\mathcal{T}(295, 300, 305)$. La transformation de changement d'échelle se fait donc via les grandeurs :

$$m_{I'} = \begin{pmatrix} 4990 \\ 295 \end{pmatrix} \text{ et } s_{I'} = \begin{pmatrix} 20 & 0 \\ 0 & 5 \end{pmatrix}.$$

On en déduit donc, d'après le Corollaire 4.2.6 et le trou spectral ci-dessus :

$$D^T_{I''} \leq \frac{20^2}{18.1}\nu_5 + \frac{5^2}{18.1}\nu_6.$$

Calcul des indices

La dernière étape de l'implémentation consiste à calculer les indices DGSM et la variance totale afin d'obtenir des bornes explicites. Dans le cadre de ce modèle simplifié, il est également possible de calculer les indices de Sobol' totaux. Les deux cas s'obtiennent par Monte-Carlo, avec des techniques néanmoins légèrement différentes.

Les indices DGSM nécessitent le calcul du carré des dérivées partielles de la fonction d'intérêt, puis une estimation de l'intégrale de ces derniers. Cette estimation peut se faire par Monte-Carlo, et le calcul des dérivées partielles par différences finies selon la procédure ci-dessous. Par exemple pour l'indice ν_1 associé à X_1 pour une fonction f :

1. on génère un échantillon $\mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(n)}$ de copies indépendantes de $\mathbf{X}$;
2. en chaque point $\mathbf{X}^{(i)}$, on calcule le quotient :

$$\Delta^{\varepsilon}_{1,i} f = \frac{f(X_1^{(i)} + \varepsilon, X_2^{(i)}, \ldots, X_d^{(i)}) - f(\mathbf{X})}{\varepsilon},$$

   où ε est choisi assez petit. Si f est dérivable, on a $\Delta^{\varepsilon}_{1,i} f \xrightarrow[\varepsilon \to 0]{} \partial_1 f(\mathbf{X})$;
3. on considère alors l'estimateur de Monte-Carlo :

$$\hat{\nu}_1 = \frac{1}{n} \sum_{i=1}^{n} (\Delta^{\varepsilon}_{1,i} f)^2.$$

   Pour ε assez petit, c'est un estimateur consistant sans biais de ν_1.

Cette procédure doit être répétée pour chacun des indices à calculer. Chaque réalisation de $\hat{\nu}_1$ nécessite $2n$ appels de la fonction f (pour le calcul de chaque $\Delta^{\varepsilon}_{1,i} f$). Dans un cadre où la fonction est coûteuse, ou résulte d'une expérience physique, ces appels peuvent être limités en nombre. Pour cela, on considérera des valeurs relativement faibles de n, pour observer la performance de la borne obtenue en contexte de parcimonie.

Le calcul des indices de Sobol' fait appel à des variance partielles. Ces dernières sont obtenues empiriquement à partir de l'estimateur de Sobol'-Jansen [51], tel que présenté par exemple dans [41], qui se base sur la formule "pick-and-freeze" (voir la Proposition 4.2.8). La procédure est alors la suivante, pour une fonction f et un sous-ensemble d'indices $I \subset [\![1, d]\!]$:

1. on génère un échantillon $\mathbf{X}^{(1)}, \ldots, \mathbf{X}^{(n)}$ de copies indépendantes du vecteur aléatoire d'intérêt ;

2. on génère indépendamment $Z_I^{(1)}, \ldots, Z_I^{(n)}$ des copies indépendantes de X_I ;

3. on considère l'estimateur de Monte-Carlo :

$$\hat{D}_I^T = \frac{1}{n} \sum_{i=1}^{n} \left(f\left(X_I^{(i)}, X_{-I}^{(i)}\right) - f\left(Z_I^{(i)}, X_{-I}^{(i)}\right) \right)^2.$$

C'est un estimateur consistant sans biais de D_I^T : on a, par définition de Z_I, $\mathbf{E}[f(X_I, X_{-I}) - f(Z_I, X_{-I})] = 0$. Cet estimateur est de plus asymptotiquement efficace (au sens des estimateurs sans biais : sa variance est asymptotiquement minimale parmi les estimateurs sans biais de la même quantité). On observe enfin la propriété remarquable que $\hat{D}_I^T = 0$ presque sûrement dès que $D_I^T = 0$ (autrement dit, l'estimateur est exact dès que les variables ne sont pas influentes). Ces propriétés peuvent être retrouvées, entre autres, dans [41].

Le calcul de cet estimateur demande également $2n$ appels à la fonction f, ce qui peut poser les mêmes questions de parcimonie que pour l'estimateur des indices DGSM. Notons néanmoins que le calcul de $\hat{v}_i$ n'utilise qu'un seul échantillon de $\mathbf{X}$, quand celui de $\hat{D}_I$ nécessite de regénérer une copie indépendante d'une partie de l'échantillon. Cela peut poser problème dans certaines situations (par exemple quand les variables sont issues d'un système physique), où l'on pourra préférer l'estimation via les indices DGSM (bien que ces derniers soient en eux-mêmes plus difficiles à interpréter que les indices de Sobol'). En effet, dans les situations pratiques, l'échantillon est obtenu avant le calcul des estimateur et indépendamment de ceux-ci. Avec les méthodes décrites ci-dessus, on disposera donc de moitié moins de variables pour l'estimation des indices de Sobol' que pour celle des indices DGSM ce qui, si l'échantillon de départ n'est pas assez grand, peut poser des problèmes de stabilité des estimateurs (une taille trop faible d'échantillon ne garantit par exemple pas la normalité asymptotique du TCL ou la convergence presque sûre de la loi des grands nombres).

Afin d'évaluer uniquement la performance de ces derniers, on estimera les indices de Sobol' avec de grands échantillons, de sorte à pouvoir considérer la valeur obtenue comme une valeur exacte. De même pour le calcul de la variance totale D, qui se fait via l'estimateur sans biais classique.

Le graphe ci-dessous représente les bornes obtenues via les indices DGSM, en regard des indices de Sobol' pour la fonction C du modèle. Les bornes pour H_d et C_b sont obtenues par le même procédé, mais via des constantes de Poincaré unidimensionnelles. Les tailles d'échantillons sont précisées dans chaque cas.

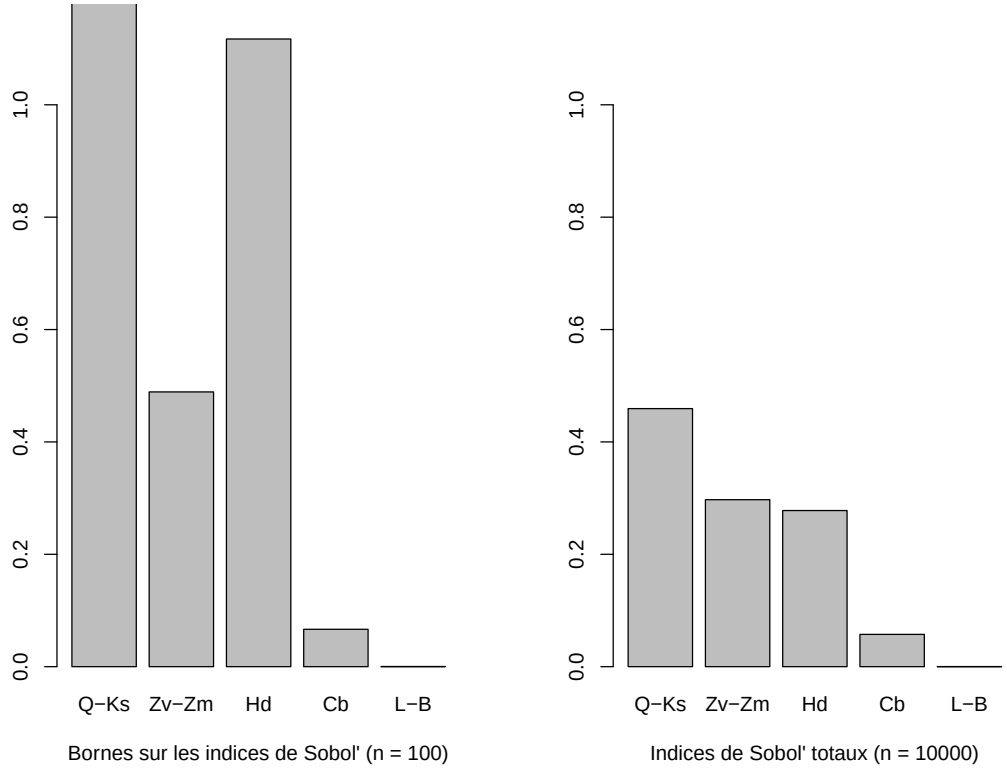


FIGURE 4.14 – Bornes sur les indices de Sobol' pour la fonction C

Les résultats sont bons pour les couples de lois triangulaires ((Z_m, Z_v) et (L, B)) ainsi que pour C_b : la borne obtenue est proche de la valeur réelle des indices, surtout pour C_b et (L, B). En particulier, l'estimation via DGSM permet de retrouver les variables peu influentes C_b et (L, B) mises en évidence par le calcul des indices de Sobol'. Les résultats pour (Q, K_s) et H_d donnent des bornes supérieures à 1, qui sont donc inexploitables pour des indices de sensibilité (qui sont par construction inférieurs à 1). Ce défaut est lié à la régularité de la fonction C, et en particulier au caractère hautement non linéaire en les différentes variables, particulièrement marqué sur la variable H_d.

Dans le cas de la fonction S, les résultats sont toujours bons pour les lois triangulaires, et meilleurs pour la variable H_b. On observe en réalité que, sur les plages de valeurs concernées, la fonction S est, du fait de certaines des valeurs extrêmes en jeu, presque linéaire en chacune des variables (sauf K_s et, dans une moindre mesure, Q). Par ailleurs, les résultats observés en lien avec Q peuvent également être expliqués en partie par le caractère "extrême" de la loi de Gumbel. Cette dernière est en effet utilisée ici pour décrire les valeurs extrêmes d'un phénomène (le débit de la rivière), et la plage de valeurs considérée peut expliquer la valeur de la borne obtenue après normalisation.

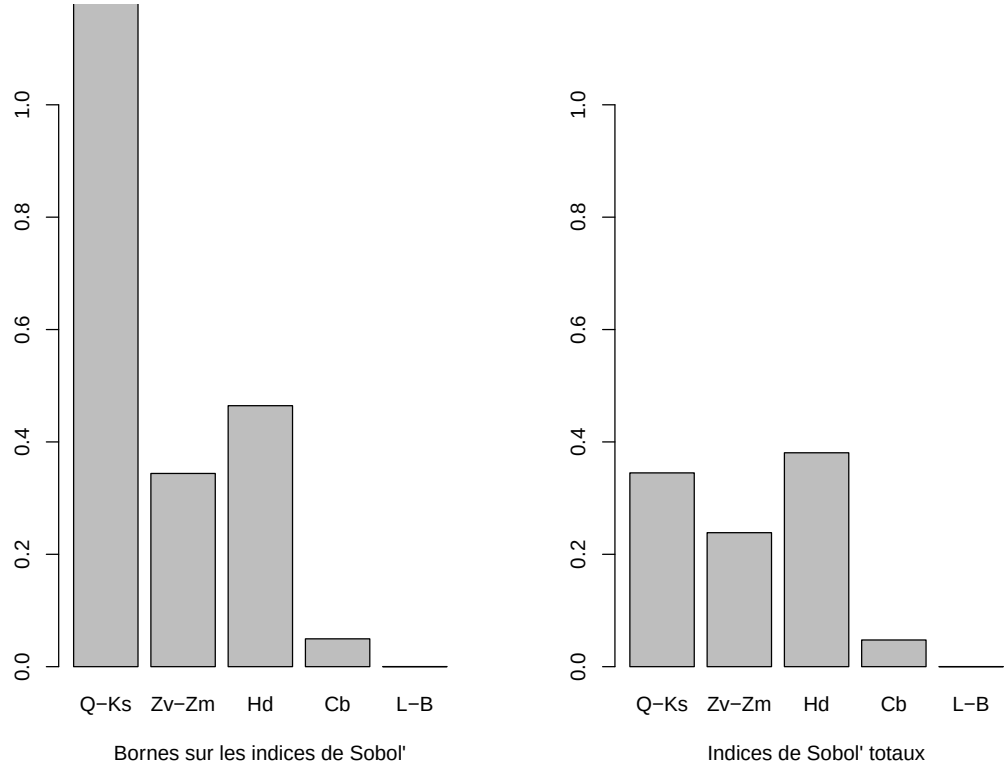


FIGURE 4.15 – Bornes sur les indices de Sobol' pour la fonction S

On peut également souligner que, couple (Q, K_s) excepté, la borne obtenue rend compte du même ordre d'importance des couples de variables que les indices de Sobol' (H_d, puis (Z_v, Z_m), C_b et (L, B)).

4.3.6 Conclusion et perspectives

La méthode proposée a permis, moyennant certains réglages, d'obtenir de nouvelles bornes sur les indices de Sobol' via les indices DGSM. Des estimations avec de petites tailles d'échantillons ont montré des conclusions pertinentes et exploitables, dans les limites permises par le modèle. Si les techniques utilisées peuvent difficilement s'étendre en dimension supérieure (même avec des moyens matériels plus importants), elles permettent de dépasser le cadre des variables indépendantes en considérant des couples de variables liées, tout en proposant une implémentation assez simple et facile à paramétrer. À cette fin, les codes R et Python peuvent être obtenus sur demande, en attendant la publication sur une plateforme idoine. Les temps de calcul sont également très raisonnables, y compris pour un ordinateur peu puissant, à l'instar de celui sur lequel ces lignes sont écrites.

Le comportement de l'algorithme dans le cas du couple (Q, K_s) pourrait néanmoins être questionnés dans de futurs développements. Le lien entre le caractère "extrême" de

la loi de Gumbel et les valeurs obtenues semble lié ici aux constantes de normalisation et ne renseigne pas beaucoup sur la stabilité de l'algorithme dans ces cas précis. Notons cependant que le calcul direct (sans normalisation préalable) a été essayé pour (Q, K_s), sans résultats probants (les valeurs extrêmes des marges semblant en effet affecter fortement la performance de l'algorithme).

D'autres pistes de maillage pour la résolution par éléments finis pourraient également être explorées, notamment en prenant en compte les valeurs mêmes de la densité. En effet, la valeur de la densité est rarement homogène, même sur une région de prédiction fixée, et il pourrait être alors pertinent de s'intéresser à des maillages qui rendent compte de cette hétérogénéité. De tels maillages pourraient par exemple reposer sur la génération d'un échantillon de la loi d'intérêt, dans la mesure où cette opération n'est pas elle-même trop coûteuse. Un réflexion pourrait aussi être conduite autour des valeurs de conditionnement obtenues pour les matrices de rigidité.

Un approfondissement des résultats théoriques pourrait renforcer la méthode utilisée ici, par exemple une compréhension plus fine des grandeurs en jeu dans le Théorème 4.3.11. S'il semble que les constantes obtenues sont loin de l'optimal dans le cas de variables indépendantes (le seul où des valeurs suffisamment précises, voire exactes, sont connues), la question reste ouverte pour des couples non indépendants. On pourrait également s'intéresser à la méthode de preuve en elle-même, au sein de laquelle la borne uniforme utilisée sur la matrice $\begin{pmatrix} I(X_1) & 0 \\ 0 & II(X_2) \end{pmatrix}$ pourrait être discutée. Enfin, il serait intéressant de réussir à élargir la palette d'exemples d'applications, par exemple en incluant un algorithme d'approximation de la constante $C(\mathbf{X})$.

Pour conclure, la méthode présentée ici peut être appliquée à d'autres choix de fonctions (un peu dans l'esprit de [55]), par exemple les fonctions de Morris [64], qui pourraient mettre en lumière l'influence de la fonction sur les bornes obtenues de façon peut-être plus claire que le modèle de crue.

Bibliographie

[1] Grégoire ALLAIRE : *Analyse numérique et optimisation*. Éditions de l'École Polytechnique, Paris, 2012.

[2] Cécile ANÉ, Sébastien BLACHÈRE, Djalil CHAFAÏ, Pierre FOUGÈRES, Ivan GENTIL, Florent MALRIEU, Cyril ROBERTO et Grégory SCHEFFER : *Sur les inégalités de Sobolev logarithmiques*. Numéro 10 de Panoramas et synthèses. Société Mathématique de France, Paris, 2000.

[3] Marc ARNAUDON, Michel BONNEFONT et Aldéric JOULIN : Intertwinings and generalized Brascamp–Lieb inequalities. *Revista Matemática Iberoamericana*, 34(3): 1021–1054, 2018.

[4] Louis BACHELIER : *Théorie de la spéculation*. Gauthier-Villars, Paris, 1900.

[5] Dominique BAKRY : Un critère de non-explosion pour certaines diffusions sur une variété riemannienne complète. *In* Paul André MEYER, éditeur : *Comptes-rendus de l'Académie des Sciences de Paris*, volume 303, pages 23–26. Académie des Sciences, Paris, 1986.

[6] Dominique BAKRY : Étude des transformations de Riesz dans les variétés riemanniennes à courbure de Ricci minorée. *In* Jacques AZÉMA, Marc YOR et Paul André MEYER, éditeurs : *Séminaire de Probabilités XXI*, volume 1247 de *Lecture Notes in Mathematics*, pages 137–172. Springer, Berlin, Heidelberg, 1987.

[7] Dominique BAKRY, Franck BARTHE, Patrick CATTIAUX et Arnaud GUILLIN : A simple proof of the Poincaré inequality for a large class of probability measures. *Electronic Communications in Probability*, 13:60–66, 2008.

[8] Dominique BAKRY, Ivan GENTIL et Michel LEDOUX : *Analysis and Geometry of Markov Diffusion Operators*. Numéro 348 de Grundlehren der mathematischen Wissenschaften. Springer, Berlin, Heidelberg, 2014.

[9] Dominique BAKRY et Dominique MICHEL : Sur les inégalités FKG. *In* Jacques AZÉMA, Marc YOR et Paul André MEYER, éditeurs : *Séminaire de Probabilités XXVI*, volume 1526 de *Lecture Notes in Mathematics*, pages 170–188. Springer, Berlin, Heidelberg, 1992.

[10] Dominique BAKRY et Michel ÉMERY : Diffusions hypercontractives. *In* Jacques AZÉMA et Marc YOR, éditeurs : *Séminaire de Probabilités XIX 1983/84*, volume 1123 de *Lecture Notes in Mathematics*, pages 177–206. Springer, Berlin, Heidelberg, 1985.

[11] Adhémar BARRÉ DE SAINT-VENANT : Théorie du mouvement non permanent des eaux, avec application aux crues des rivières et a l'introduction de marées dans leurs lits. *Comptes rendus hebdomadaires des séances de l'Académie des sciences*, 73:147–154, 1871.

[12] Franck BARTHE : L'inégalité de corrélation gaussienne (d'après Thomas Royen). *Séminaire Bourbaki*, 1124:1–15, 2017.

[13] Franck BARTHE, Bertrand IOOSS et Olivier ROUSTANT : Poincaré inequalities on intervals – application to sensitivity analysis. *Electronic Journal of Statistics*, 11(2): 3081–3119, 2017.

[14] Franck BARTHE et Bo'az KLARTAG : Spectral gaps, symmetries and log-concave perturbations. *Bulletin of the Hellenic Mathematical Society*, 64:1–31, 2020.

[15] George David BIRKHOFF : Proof of the ergodic theorem. *Proceedings of the National Academy of Sciences USA*, 17(12):656–660, 1931.

[16] Sergueï BOBKOV, Friedrich GÖTZE et Christian HOUDRÉ : On Gaussian and Bernoulli Covariance Representations. *Bernoulli*, 7(3):439–451, 2001.

[17] Michel BONNEFONT et Alderic JOULIN : A note on eigenvalues estimates for one-dimensional diffusion operators. *Bernoulli*, 28(1):64–86, 2022.

[18] Michel BONNEFONT et Aldéric JOULIN : Intertwining Relations for One-Dimensional Diffusions and Application to Functional Inequalities. *Potential Analysis*, 41(4):1005–1031, 2014.

[19] Michel BONNEFONT et Aldéric JOULIN : Intertwinings, second-order Brascamp–Lieb inequalities and spectral estimates. *Studia Mathematica*, 260(3):285–316, 2021.

[20] Michel BONNEFONT, Aldéric JOULIN et Yutao MA : Spectral gap for spherically symmetric log-concave probability measures, and beyond. *Journal of Functional Analysis*, 270(7):2456–2482, 2016.

[21] Jean-Pierre BOURGUIGNON : The "magic" of Weitzenböck formulas. *In Variational Methods*, volume 4 de *Progress in Nonlinear Differential Equations and Their Applications*. Birkhäuser, Boston, 1990.

[22] Adrian BOWYER : Computing Dirichlet tessellations. *The Computer Journal*, 24(2): 162–166, 1981.

[23] Herm BRASCAMP et Elliott LIEB : On Extensions of the Brunn-Minkowski and Prékopa-Leindler theorems, including inequalities for log concave functions, and with an application to the diffusion equation. *Journal of Functional Analysis*, 22(4): 366–389, 1976.

[24] Patrick CATTIAUX, Arnaud GUILLIN et Liming WU : Poincaré and logarithmic Sobolev inequalities for nearly radial measures. *arXiv :1912.10825*, page 24, 2019.

[25] Djalil CHAFAÏ et Aldéric JOULIN : Intertwining and commutation relations for birth–death processes. *Bernoulli*, 19(5A):1855–1879, 2013.

[26] Bertrand CLOEZ et Claire DELPLANCKE : Intertwinings and Stein's magic factors for birth-death processes. *Annales de l'Institut Henri Poincaré Probabilités et Statistiques*, 55(1):341–377, 2019.

[27] Richard COURANT : Über die Eigenwerte bei den Differentialgleichungen der mathematischen Physik. *Mathematische Zeitschrift*, 7(1–4):1–57, 1920.

[28] Louis de BROGLIE : *Recherches sur la théorie des Quanta*. Masson, Paris, 1924.

[29] Étienne de ROCQUIGNY : La maîtrise des incertitudes dans un contexte industriel. Première partie : une approche méthodologique globale basée sur des exemples. *Journal de la société française de statistique*, 147(3):33–71, 2006.

[30] Boris DELAUNAY : Sur la sphère vide. À la memoire de Georges Voronoï. *Bulletin de l'Academie des Sciences de l'URSS. Classe des sciences mathematiques et naturelles*, 6:793–800, 1934.

[31] Jérôme DEVEAUX et Landry DUGUET : Discovering influential variables with 2D Poincaré inequalities. Mémoire de D.E.A., INSA Toulouse, 2021.

[32] Alber EINSTEIN : Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Annalen der Physik*, 322(8):549–560, 1905.

[33] David ELWORTHY et Xue-Mei LI : Formulae for the derivatives of heat semigroups. *Journal of Functional Analysis*, 125:252–286, 2019.

[34] Lawrence EVANS : *Partial differential equations*. Numéro 19 de Graduate studies in mathematics. American mathematical society, Providence (R.I.), 2ème édition, 2010.

[35] William FARIS : Product spaces and Nelson's inequality. *Helvetica Physica Acta*, 48(5):721–730, 1975.

[36] Paul FEDERBUSH : Partially Alternate Derivation of a Result of Nelson. *Journal of Mathematical Physics*, 10(1):50–52, 1968.

[37] Adolf FICK : Über Diffusion. *Annalen der Physik und Chemie*, 170(1):59–86, 1855.

[38] Ernst FISCHER : Über quadratische Formen mit reellen Koeffizienten. *Monatshefte für Mathematik und Physik*, 16:234–249, 1905.

[39] Cees FORTUIN, Pieter KASTELEYN et Jean GINIBRE : Correlation Inequalities on Some Partially Ordered Sets. *Communications in Mathematical Physics*, 22(2):89–103, 1971.

[40] Joseph FOURIER : *Théorie analytique de la chaleur*. Firmin Didiot, Père et Fils, Paris, 1822.

[41] Jana FRUTH, Olivier ROUSTANT et Sonja KUHNT : Total interaction index : A variance-based sensitivity index for second-order interaction screening. *Journal of Statistical Planning and Inference*, 147:212–223, 2014.

[42] Leonard GROSS : Logarithmic Sobolev Inequalities. *American Journal of Mathematics*, 97(4):1061–1083, 1975.

[43] Emil GUMBEL : Les valeurs extrêmes des distributions statistiques. *Annales de l'Institut Henri Poincaré*, 5(2):115–158, 1935.

[44] Bernard HELFFER : *Semiclassical analysis, Witten laplacians, and statistical mechanics*. Numéro 1 de Series on partial differential equations and applications. World Scientific, 2002.

[45] Wassily HOEFFDING : A Class of Statistics with Asymptotically Normal Distribution. *Annals of Mathematical Statistics*, 19(3):293–325, 1948.

[46] Richard HOLLEY et Daniel STROOCK : Logarithmic Sobolev inequalities and stochastic Ising models. *Journal of Statistical Physics*, 46(5):1159–1194, 1987.

[47] Baptiste HUGUET : Intertwining relations for diffusions in manifolds and applications to functional inequalities. *Stochastic Processes and their Applications*, 145:1–28, 2022.

[48] Bertrand IOOSS : Revue sur l'analyse de sensibilité globale de modèles numériques. *Journal de la société française de statistique*, 152(1):3–25, 2011.

[49] Bertrand IOOSS et Paul LEMAÎTRE : *A Review on Global Sensitivity Analysis Methods*, volume 56 de *Operations Research/Computer Science Interfaces Series*, chapitre 5, pages 101–122. Springer US, 2015.

[50] Kiyoshi ITŌ : Stochastic integral. *Proceedings of the Imperial Academy*, 20(8):519–524, 1944.

[51] Michiel JANSEN : Analysis of variance designs for model output. *Computer Physics Communications*, 117(1–2):35–43, 1999.

[52] Jon JOHNSEN : On the spectral properties of Witten-Laplacians, their range projections and Brascamp-Lieb's inequality. *Integral Equations and Operator Theory*, 36(3):288–324, 2000.

[53] Mark KAC : On distributions of certain Wiener functionals. *Transactions of the American Mathematical Society*, 65(1):1–13, 1949.

[54] Bo'az KLARTAG : A Berry-Esseen type inequality for convex bodies with an unconditional basis. *Probability Theory and Related Fields*, 145:1–33, 2009.

[55] Matieyendou LAMBONI, Bertrand IOOSS, Anne-Laure POPELIN et Fabrice GAMBOA : Derivative-based global sensitivity measures : general links with Sobol' indices and numerical tests. *Mathematics and Computers in Simulation*, 87:45–5451, 2013.

[56] Lev LANDAU et Evgenii LIFCHITZ : *Physique théorique. Tome V : Physique statistique*. Mir, Moscou, 2e édition, 1967.

[57] Paul LANGEVIN : Sur la théorie du mouvement brownien. *Comptes-rendus de l'Académie des Sciences*, 146:530–532, 1908.

[58] Rafal LATAŁA et Krzysztof OLESZKIEWICZ : Between Sobolev and Poincaré. *In* Vitaly MILMAN et Gideon SCHECHTMAN, éditeurs : *Geometric Aspects of Functional Analysis*, volume 1745 de *Lecture Notes in Mathematics*, pages 147–168. Springer, Berlin, Heidelberg, 2000.

[59] Michel LEDOUX : *The Concentration of Measure Phenomenon*. Numéro 89 de Mathematical Surveys and Monographs. American mathematical society, Providence (R.I.), 2001.

[60] Paul LÉVY : *Processus stochastiques et mouvement brownien*. Gauthier-VIllars, Paris, 1948.

[61] James MAXWELL : On the Dynamical Theory of Gases. *Philosophical Transactions of the Royal Society of London*, 157:49–88, 1867.

[62] Laurent MICLO : Monotonicity of the extremal functions for one-dimensional inequalities of logarithmic Sobolev type. *In* Catherine DONATI-MARTIN, Michel ÉMERY, Alain ROUAULT et Christophe STRICKER, éditeurs : *Séminaire de Probabilités XLII*, volume 1979 de *Lecture Notes in Mathematics*, pages 103–130. Springer, Berlin, Heidelberg, 2009.

[63] Emanuel MILMAN : Spectral Estimates, Contractions and Hypercontractivity. *Journal of Spectral Theory*, 8(2):669–714, 2018.

[64] Max MORRIS : Factorial Sampling Plans for Preliminary Computational Experiments. *Technometrics*, 33(2):61–174, 1991.

[65] André NATAF : Détermination des distributions dont les marges sont données. *Comptes-rendus de l'Académie des Sciences*, 245:42–43, 1962.

[66] Roger NELSEN : *An Introduction to Copulas*. Springer, 1999.

[67] Edward NELSON : A quartic interaction in two dimensions. *In Proceedings of the Conference on the Mathematical Theory of Elementary Particles*, pages 69–73, 1966.

[68] Ivan NOURDIN et Giovanni PECCATI : *Normal Approximations with Malliavin Calculus*. Numéro 192 de Cambridge Tracts in Mathematics. Cambridge University Press, Cambridge, 2012.

[69] Grigori PERELMAN : The entropy formula for the Ricci flow and its geometric applications. *arXiv :math/0211159*, 2002.

[70] Jean PERRIN : Mouvement brownien et réalité moléculaire. *Annales de chimie et de physique*, 18(8):5–114, 1909.

[71] Loucas PILLAUD-VIVIEN, Francis BACH, Tony LELIÈVRE, Alessandro RUDI et Gabriel STOLTZ : Statistical Estimation of the Poincaré constant and Application to Sampling Multimodal Distributions. *In Proceedings of the International Conference on Artificial Intelligence and Statistics (AISTATS)*, volume 108 de *Proceedings in Machine Learning Research*, pages 2753–2763, 2020.

[72] Henri POINCARÉ : Sur les Equations aux Derivees Partielles de la Physique Mathematique. *American Journal of Mathematics*, 12(3):211–291, 1890.

[73] Philip PROTTER : *Stochastic integration and differential equations*. Numéro 21 de Stochastic modelling and applied probability. Springer, Berlin, Heidelberg, 2. ed. édition, 2005.

[74] John RAYLEIGH : *The Theory of Sound*, volume 1. McMillan Co., 1877.

[75] Stefano REBAY : Efficient unstructured Mesh Generation by Means of Delaunay Triangulation and Bowyer-Watson Algorithm. *Journal of Computational Physics*, 106(1):125–138, 1993.

[76] Michael REED et Barry SIMON : *Methods of modern mathematical physics IV : Analysis of Operators*. Academic Press inc., San Diego, California, 1978.

[77] Andrea SALTELLI et Ilya SOBOL' : Sensitivity analysis for nonlinear mathematical models : numerical experience. *Mathematical Modeling & Computational Experiment*, 7(11):16–28, 1995.

[78] Erwin SCHRÖDINGER : An Undulatory Theory of the Mechanics of Atoms and Molecules. *Physical Review*, 28(6):1049–1070, 1926.

[79] Abe SKLAR : Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut de Statistiques de l'Université de Paris*, 8:229–231, 1959.

[80] Ilya SOBOL' et A. GERSHMAN : On an alternative global sensitivity analysis estimator. *In Proceedings of SAMO, Belgirate, 1995*, pages 33–35, 1995.

[81] Adriaan STAM : Some inequalities satisfied by the quantities of information of Fisher and Shannon. *Information and Control*, 2(2):101–112, 1959.

[82] Josef STEFAN : Über das Gleichgewicht und die Bewegung, insbesondere die Diffusion von Gasgemengen. *Sitzungsberichte der Mathematisch-Naturwissenschaftlichen Classe der Kaiserlichen Akademie der Wissenschaften Wien*, 2:63–124, 1871.

[83] Clément STEINER : A Feynman-Kac approach for Logarithmic Sobolev Inequalities. *Electronic Journal of Probability*, 26:1–19, 2021.

[84] Mikhail SUBBOTIN : On the Law of Frequency of Error. *Matematicheskii Sbornik*, 32(2):296–301, 1923.

[85] Marian von SMOLUCHOWSKI : Zur kinetischen Theorie der Brownschen Molekularbewegung und der Suspensionen. *Annalen der Physik*, 326(14):756–780, 1906.

[86] Feng-Yu WANG : Functional Inequalities for Empty Essential Spectrum. *Journal of Functional Analysis*, 170(1):219–245, 2000.

[87] Feng-Yu WANG : Modified Curvatures on Manifolds with Boundary and Applications. *Potential Analysis*, 41(3):699–714, 2014.

[88] David WATSON : Computing the n-dimensional Delaunay tessellation with application to Voronoi polytopes. *The Computer Journal*, 24(2):167–172, 1981.

[89] Norbert WIENER : Differential-Space. *Journal of Mathematical Physics*, 2(2):131–174, 1923.

RÉSUMÉ

Cette thèse s'inscrit dans une série de travaux menés par Aldéric Joulin, Michel Bonnefont et leurs collaborateurs, dont le but est d'appliquer des relations d'entrelacement pour obtenir des informations sur certains générateurs de Markov. Le présent travail traite de trois propriétés plus spécifiquement : les inégalités de Poincaré, de Sobolev logarithmiques et les estimations spectrales. Les deux inégalités ci-dessus sont des outils classiques d'analyse infini-dimensionnelle, qui présentent des liens intéressants avec les générateurs sus-mentionnés et les propriétés de la mesure de Boltzmann-Gibbs associée.

Le premier chapitre propose une approche basée sur les semi-groupes de Feynman-Kac pour déduire de nouvelles estimations en relation avec l'inégalité de Sobolev logarithmique. Les résultats sont obtenus en utilisant le lien entre générateurs et processus stochastiques, via les semi-groupes de Feynman-Kac.

Le deuxième chapitre pousse la réflexion amorcée par l'inégalité de Poincaré sur l'estimation des valeurs propres desdits générateurs en proposant une vision algébrique pour l'estimation de ces dernières. Ce travail fait écho à un article récent d'Emanuel Milman sur le sujet, où une approche par transport optimal était privilégiée. Les multiplicités associées sont également discutées, en lien notamment avec de récents travaux de Franck Barthe et Boaz Klartag.

Enfin, le dernier chapitre traite d'une application des inégalités de Poincaré à l'analyse de sensibilité, où ces dernières sont utilisées pour faire un lien entre deux familles d'indices de sensibilité. Une méthode d'estimation en lien avec cette inégalité en dimension deux est proposée, basée sur des techniques d'éléments finis.

ABSTRACT

This thesis is part of a series of works carried out by Aldéric Joulin, Michel Bonnefont et alius, which aims at using intertwining relations to infer properties of some Markov generators. The present work deals specifically with three properties : Poincaré inequalities, logarithmic Sobolev inequalities and spectral estimates. Both above inequalities are widely use tools in infinite-dimensional analysis, that relate to the latter generators and underlying Boltzmann-Gibbs invariant distribution. In the first chapter, a method based on Feynman-Kac semigroups is proposed to infer new estimates, in relation to the logarithmic Sobolev inequality. The connexion between generators and stochastic processes is explored via a representation theorem for Feynman-Kac semigroups.

In the second chapter, an algebraic approach to the estimation of eigenvalues of the aforementioned generators is discussed. This work echoes a related recent article by Emanuel Milman, in which he used optimal transport results in this very purpose. Multiplicities are addressed as well, in relation to the recent work of Franck Barthe and Boaz Klartag.

In the last chapter, the relation between Poincaré inequalities and sensitivity analysis is investigated, particularly in order to compare two types of sensitivity indices. An estimation method related to this inequality is developed in dimension two, using finite elements methods.