



HAL
open science

Chemins rugueux issus de processus discrets

Olga Lopusanschi

► **To cite this version:**

Olga Lopusanschi. Chemins rugueux issus de processus discrets. Probabilités [math.PR]. Sorbonne Université, 2018. Français. NNT : 2018SORUS074 . tel-02383238

HAL Id: tel-02383238

<https://theses.hal.science/tel-02383238v1>

Submitted on 27 Nov 2019

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.

THÈSE DE DOCTORAT DE L'UNIVERSITÉ PIERRE ET MARIE CURIE

Spécialité : Probabilités

École doctorale : « Sciences Mathématiques de Paris-Centre » (ED 386)

réalisée

au Laboratoire de Probabilités, Statistique et Modélisation

présentée par

Olga LOPUSANSCHI

pour obtenir le grade de :

DOCTEUR DE L'UNIVERSITÉ PIERRE ET MARIE CURIE

Sujet de la thèse :

CHEMINS RUGUEUX ISSUS DE PROCESSUS DISCRETS

soutenue le 18 janvier 2018

devant le jury composé de :

M ^{me}	Laure COUTIN	Rapporteur
M.	Lorenzo ZAMBOTTI	Directeur de thèse
M.	Damien SIMON	Co-encadrant
M.	Ismael BAILLEUL	Examineur
M.	Paul GASSIAT	Examineur
M.	Thierry LÉVY	Examineur

Après avis de Laure Coutin (rapporteur) et Samy Tindel (rapporteur).

Remerciements

Lorsqu'on commence une thèse, comme, par ailleurs, lorsqu'on commence tout autre vaste projet dans la vie, sa fin se perd pour nous dans les brouillards d'un horizon très lointain. On a du mal à croire les esprits avertis qui nous disent qu'on l'atteindra rapidement avant d'en faire l'expérience : on pense avoir tellement de temps, assez pour tout ce qu'on a prévu de faire, mais aussi pour exploiter les imprévus et les occasions qui se présenteront. Et puis, avant même de s'en apercevoir, c'est à nous de dire à ceux qui entame l'aventure : « Le temps passe tellement vite. »

Cependant, même si on n'a pas vu les années s'écouler, tout ce qui a fait notre vie pendant la thèse est là pour donner une mesure au temps : les articles et les livres qu'on a étudiés, les dizaines de pages de notre manuscrit, les voyages, les conférences et les séminaires, tous ce qu'on a appris sur la vie et le travail de chercheur et surtout, surtout, toutes les personnes qu'on a connues. Toutes ces gens qui faisaient partie de notre vie au sein du laboratoire ou qu'on a rencontrés lors d'une école d'été, qu'on a simplement croisés ou qui sont devenues de bons amis, tous ceux avec qui on a travaillé ensemble, bu des cafés et/ou des bières et passé parfois des nuits blanches, qui nous ont soutenus, aidés et accompagnés tout le long de ces années parfois difficiles, se trouvent au coeur de nombreux souvenirs qui nous permettent de compter les semaines, les mois et les années.

Je garde à l'esprit ce que je dois à chacun d'entre eux, mais comme ce n'est pas le moment d'écrire des mémoires, je vais surtout parler de ceux dont la présence pendant cette période a été pour moi la plus marquante, à commencer par mes directeur et co-encadrant de thèse, Lorenzo et Damien. La chance est parfois pour quelque chose dans le travail de thèse et lorsqu'on a de bons directeurs, comme dans mon cas, elle y est pour beaucoup. Leur disponibilité, leur gentillesse, leur tolérance

à mes éternels retards n'ont eu d'égal que leur grande patience et l'aide immense qu'ils m'ont apportée pour mes recherches. Je voudrais également remercier Damien en tant que mon co-auteur : c'était un plaisir de travailler avec toi, et les choses que j'ai apprises sur la rédaction d'articles, je te le dois en grande partie.

Je dois toutes ma reconnaissance également à ceux sans qui cette soutenance n'aurait pas eu lieu, à savoir mes rapporteurs et les autres membres du jury : Laure Coutin, Samy Tindel, Thierry Lévy, Paul Gassiat et Ismael Bailleul. Merci pour le temps que vous avez pris de me lire, pour vos remarques et vos conseils.

Pour m'avoir soutenue durant toute ma vie de doctorant, je dis un grand merci à tous les membres permanents du laboratoire, que j'ai croisés plus ou moins souvent, avec qui j'ai discutés plus ou moins longuement et dont certains avaient été mes professeurs à l'époque où j'étais encore étudiante : Thomas, Gilles, François, M. Zhan Shi, Mme Irina Kourkova, Amaury, Omer, Laurent, Romain, Cédric, Quentin, Camille et tant d'autres, mais également les membres du secrétariat : Florence, Josette, Valérie, sans qui les formalités administratives m'auraient donné beaucoup de fil à retordre, et M. Comets, notre directeur. Je remercie également deux autres enseignants-chercheurs de Paris VI, Adrien et Sylvain qui m'ont beaucoup aidée et soutenue dans mon parcours universitaire.

Enfin, ces quatre ans auraient été bien moins mémorables sans ceux qui ont partagé avec moi l'expérience de la thèse : Bastien, Nelo, Sandro, Wangru, Liping, Gilles, avec qui j'ai partagé un même bureau, Sarah, Anna et Vu-lan, avec qui j'ai eu le plaisir d'organiser le groupe de travail des doctorants, Yvain, Xan, Florian, Guillaume, Carlo, Henri, Guillermo, les deux Nicolas, Paul, Adrien, un autre Florian, Blandine, Richard, avec qui on a partagé les sorties et les conférences, et tous les autres qui ont toujours été de si bonne compagnie ! Je voudrais remercier particulièrement Sanjay, Pierre-Antoine, Eric, Cyril et Joon pour leur amitié et leur soutien.

Ce dernier paragraphe s'adresse à ceux qui ont été et sont ma vie en dehors du monde universitaire. Ma famille, depuis la Moldavie, n'a cessé de me soutenir depuis le jour où elle m'a donné la possibilité de venir en France et jusqu'à aujourd'hui, et ma thèse leur doit beaucoup. Avec ma soeur, qui est en train de finir sa thèse, on s'est beaucoup encouragé mutuellement. Mes amis m'ont aidée à ne pas m'enfermer dans ce travail et d'y revenir constamment avec un nouveau regard. Enfin, mon

mari Laurent a dû supporter mes doutes, mes déceptions, mais aussi l'absence d'horaires de travail et son imprévisibilité, et ma reconnaissance pour son soutien n'a d'égal que mon admiration pour sa patience infinie.

Je voudrais dédier ce travail à mes deux grand-pères, dont un n'est malheureusement plus de ce monde. Ils avaient tous deux le potentiel de faire une thèse, mais la vie en a décidé autrement. Je suis très honorée d'avoir pu réaliser en partie leur rêve.

Résumé

Le présent travail se veut une contribution à l'extension du domaine des applications de la théorie des chemins rugueux à travers l'étude de la convergence des processus discrets, qui permet un nouveau regard sur plusieurs problèmes qui se posent dans le cadre du calcul stochastique classique.

Nous étudions la convergence en topologie rugueuse, d'abord des chaînes de Markov sur graphes périodiques, ensuite des marches de Markov cachées, et ce changement de cadre permet d'apporter des informations supplémentaires sur la limite grâce à *l'anomalie d'aire*, invisible en topologie uniforme. Nous voulons montrer que l'utilité de cet objet dépasse le cadre des équations différentielles.

Nous montrons également comment le cadre des chemins rugueux permet d'encoder la manière dont on plonge un modèle discret dans l'espace des fonctions continues, et que les limites des différents plongements peuvent être différenciées précisément grâce à l'anomalie d'aire.

Nous définissons ensuite les temps d'occupation itérés pour une chaîne de Markov et montrons, en utilisant les sommes itérées, qu'ils donnent une structure combinatoire aux marches de Markov cachées. Nous proposons une construction des chemins rugueux en passant par les sommes itérées et la comparons à la construction classique, faite par les intégrales itérées, pour trouver à la limite deux types de chemins rugueux différents, non-géométrique et géométrique respectivement.

Pour finir, nous illustrons le calcul et la construction de l'anomalie d'aire et nous donnons quelques résultats supplémentaires sur la convergence des sommes et temps d'occupation itérés.

Abstract

Through the present work, we hope to contribute to extending the domain of applications of rough paths theory by studying the convergence of discrete processes and thus allowing for a new point of view on several issues appearing in the setting of classical stochastic calculus.

We study the convergence, first of Markov chains on periodic graphs, then of hidden Markov walks, in rough path topology, and we show that this change of setting allows to bring forward extra information on the limit using *the area anomaly*, which is invisible in the uniform topology. We want to show that the utility of this object goes beyond the setting of differential equations.

We also show how rough paths can be used to encode the way we embed a discrete process in the space of continuous functions, and that the limits of these embeddings differ precisely by the area anomaly term.

We then define the iterated occupation times for a Markov chain and show using iterated sums that they form an underlying combinatorial structure for hidden Markov walks. We then construct rough paths using iterated sums and compare them to the classical construction, which uses iterated integrals, to get two different types of rough paths at the limit: the non-geometric and the geometric one respectively.

Finally, we illustrate the computation and construction of the area anomaly and we give some extra results on the convergence of iterated sums and occupation times.

Table des matières

Remerciements	i
Résumé	v
Abstract	vii
1 Introduction	1
1.1 Le mot d'introduction	1
1.2 Cadre et motivations du présent travail	2
1.2.1 Un peu d'histoire	2
1.2.2 Journal de bord	5
1.2.3 Structure générale de la thèse	11
1.2.4 Notations	12
1.3 Le drift d'aire dans la convergence de modèles markoviens en topologie rugueuse	13
1.3.1 Éléments de la théorie des chemins rugueux	13
1.3.2 Aire de Lévy et drift d'aire	24
1.3.3 Marches de Markov cachées et chaînes de Markov sur graphes périodiques.	34
1.3.4 Les plongements	40
1.3.5 Les équations aux différences	46
1.3.6 Autres résultats préliminaires et sources d'inspiration	46
1.3.7 Présentation des principaux résultats	48
1.3.8 Applications	54

1.4	Convergence de sommes itérées pour des modèles markoviens et étude de la renormalisation	56
1.4.1	Sommes et intégrales itérées	56
1.4.2	Temps d'occupation itérés et théorie ergodique	57
1.4.3	Produits <i>shuffle</i> et <i>quasi-shuffle</i>	60
1.4.4	Présentation des principaux résultats	62
1.5	Le mot de la fin	66
2	Lévy area with a drift as a renormalization limit of Markov chains on periodic graphs	69
2.1	Introduction	69
2.1.1	General motivations	69
2.1.2	Structure of the present article	71
2.1.3	An easy discrete example: rotating sums of Bernoulli r.v.	72
2.1.4	Main result	74
2.2	Tools and additional results	80
2.2.1	Additional results on Λ -invariant Markov chains	80
2.2.2	Decomposition into pseudo-excursions	81
2.2.3	Area process and rough paths	84
2.2.4	The group $G^2(E)$	85
2.3	Proof of main theorem and comments	87
2.3.1	Proof of theorem 2.1.1	87
2.3.2	Properties of the area anomaly	96
2.4	Applications and open questions	97
2.4.1	Application to an SDE	97
2.4.2	A three dimensional model with a non-trivial area anomaly	98
2.4.3	Application to a stochastic differential equation on $SU(2)$	100
2.4.4	Open questions	100
3	Area anomaly and generalized drift of iterated sums for hidden Markov walks	103
3.1	Introduction	103

3.1.1	On the importance of discrete models in the rough paths theory.	103
3.1.2	Structure of the article and notations.	107
3.2	Main results and map of the paper	109
3.2.1	Rough paths models from microscopic models.	109
3.2.2	Combinatorial structure behind the iterated sums of hidden Markov walks.	119
3.3	From Hidden Markov Walks to rough paths	126
3.3.1	Theory of pseudo-excursions for hidden Markov walks.	126
3.3.2	Elements of rough paths theory	129
3.3.3	Embeddings	132
3.3.4	Getting back to theorem 3.2.3	138
3.4	Iterated structures behind discrete time and discrete space Markov chains.	147
3.4.1	Shuffle and quasi-shuffle products.	147
3.4.2	From geometric to non-geometric rough paths through hidden Markov walks	149
3.4.3	Getting back to proposition 3.2.1.	152
3.5	Open questions	157
4	Additional results on area anomaly and iterated occupation times	159
4.1	A case study for the area anomaly	159
4.1.1	The two models as hidden Markov walks	161
4.1.2	The two models as Markov chains on periodic graphs	162
4.1.3	Computing the area anomaly for the two models	163
4.2	Methods for computing the area anomaly	172
4.2.1	A construction using the loop-erased random walk	172
4.2.2	A formula using the stochastic matrix of $(R_n)_n$	178
4.3	Computations related to combinatorial structures of hidden Markov walks	183
4.3.1	A connection between the iterated occupation times of hidden Markov walks and the resolvent	183

4.3.2	Asymptotics of iterated sums/occupation times of hidden Markov walks	185
	Bibliographie	195

Chapitre 1

Introduction

1.1 Le mot d'introduction

Dans les mathématiques, comme dans l'histoire, il y a ceux qui découvrent de nouvelles terres, il y a les pionniers qui y vont pour explorer les contrées inconnues et ceux qui s'y établissent ensuite, il y a ceux qui cherchent à trouver toutes les richesses de ce nouveau monde et ceux qui tentent de les mettre à profit en les exportant vers d'autres pays, il y a enfin ceux qui contribuent à rendre connu leur jeune pays et ceux qui s'appliquent à le rendre accueillant pour tous les étrangers qui seraient de passage.

Le domaine des chemins rugueux, ajouté sur la carte des mathématiques il y a à peine vingt ans, s'est avéré une terre riche, qui a attiré de grands chercheurs et a fourni un nombre impressionnant de résultats importants. Ceux dont les problématiques font apparaître des équations différentielles y puisent volontiers des idées et des outils. Cependant d'autres hésitent à s'y aventurer : ils pensent que les ressources qu'ils y trouveraient ne leur seraient pas indispensables, et, comme le domaine est technique, y regarder de près ne leur paraît pas bien utile.

En exportant les outils des chemins rugueux vers un nouveau domaine (celui de la convergence de processus discrets) et en montrant, à travers, par exemple, des constructions « à la main », qu'on n'a pas besoin de s'enfoncer dans cette théorie pour apprécier la beauté des paysages et y trouver des ressources, notre but est de donner de bons arguments de tenter l'aventure. Autrement dit, nous espérons

que le présent travail, en élargissant le champ d'application des chemins rugueux, contribuera à ébranler un peu plus la réputation d'obscurité que cette théorie garde encore auprès de certains chercheurs.

1.2 Cadre et motivations du présent travail

1.2.1 Un peu d'histoire

Les bases de la théorie des chemins rugueux furent posées en 1998 dans [65] par T. Lyons, dans le but de permettre la définition des intégrales et des équations différentielles pour des processus qui ont, jusque là, échappé à la théorie classique de l'intégration.

Le problème qui se pose est le suivant. Soit V un espace vectoriel de dimension finie $d \geq 2$ et $x : [0, 1] \rightarrow V$ un chemin continu dans V . Pour pouvoir résoudre une équation différentielle (ED) comme, par exemple,

$$dy_t = \sum_{k=1}^d f^{(i)}(y_t) dx_t^{(i)} \quad (1.1)$$

il faut qu'on puisse donner un sens à l'intégrale par rapport aux incréments de x . Or, si la régularité α de x est strictement inférieure à $1/2$, la théorie d'intégration de Young [82] ne s'applique plus. La même question se pose pour les équations aux dérivées partielles (EDP).

Même si, au premier abord, la topologie rugueuse peut paraître une montagne infranchissable et donner davantage l'envie de la contourner que de l'escalader, l'idée sous-jacente est aussi simple qu'importante : si le chemin x de (1.1) est de régularité inférieure à $1/2$, une suite $(x_t^n)_t$ qui converge vers x en topologie uniforme ne permet pas forcément d'approcher y , solution de (1.1), par la suite $(y_t^n)_t$ donnée par

$$dy_t^n = \sum_{k=1}^d f^{(i)}(y_t^n) (dx_t^n)^{(i)} \quad (1.2)$$

Ceci s'explique par le fait que toute l'information sur la convergence de la suite x^n nécessaire pour cela n'est pas prise en compte à la limite. Il faut donc trouver un

moyen d'« enrichir » les objets pour enregistrer cette information manquante.

L'idée de ce moyen vient du développement de Taylor. On crée, à partir de x , un *chemin rugueux* $(\mathbf{x}_t)_t$, objet avec $\lfloor 1/\alpha \rfloor$ niveaux (où $\lfloor \cdot \rfloor$ désigne la partie entière d'un réel) et dont chaque niveau k est un élément de $V^{\otimes k}$ qui satisfait aux propriétés algébriques d'une intégrale itérée et dépend en partie des niveaux précédents (on définit ainsi les intégrales itérées abstraites de x). On cherche ensuite à approcher $(\mathbf{x}_t)_t$ par une suite de chemins rugueux bien choisis et on donne ainsi un sens à l'équation (1.1) à travers sa version rugueuse.

Avant la naissance de la théorie des chemins rugueux, en plus des solutions données au cas par cas, de nombreuses méthodes de généralisation, qui continuent à être explorées aujourd'hui encore, avaient déjà été développées. Citons-en quelques unes.

En 1942, dans [49], K. Itô avait donné une définition rigoureuse de l'intégrale par rapport au mouvement brownien par analogie avec les intégrales de Riemann. À la suite des travaux de P. Lévy et de K. Itô, P. Malliavin proposa de nouvelles approches pour le calcul différentiel aléatoire, et les applications du calcul de Malliavin à la résolution des équations différentielles stochastiques (EDS) furent très fructueuses (par exemple, [79], [72], ou encore [73]) et très prisées surtout dans le domaine de la finance. D'autres méthodes numériques ([54], [37]) deviennent d'autant plus intéressantes que le perfectionnement des technologies rend les algorithmes de simulation qui en sont issus (la méthode Monte-Carlo, les schémas d'Euler-Maruyama et de Milstein, etc.) de plus en plus efficaces. L'analyse par ondelettes, issue de la physique et développée par Y. Meyer [71], S. Jaffard [50] et d'autres, en permettant une étude plus fine de la régularité d'une fonction, a permis de mieux approcher les ED(P) ([19], [22]) et d'en tirer des méthodes numériques de résolution ([74]).

La théorie des chemins rugueux semble avoir fait la synthèse de ces méthodes : elle utilise les intégrales itérées tout comme le calcul de Malliavin lorsqu'il se tourne vers le chaos gaussien ; elle permet un contrôle de régularité en « enrichissant » l'objet de niveaux supplémentaires, et les ondelettes permettent de décomposer une fonction en niveaux de régularité ; elle s'appuie sur les techniques du développement limité, tout comme beaucoup de schémas numériques ; enfin, la plupart des exemples concrets d'application sont fournis par les intégrales stochastiques. C'est en grande

partie pour cette raison que cette théorie a fourni des outils à la fois plus généraux et plus difficilement maniables que les autres méthodes.

Même si les chemins rugueux sont à valeurs dans des algèbres et des groupes de Lie, l'approche de T. Lyons était surtout analytique. M. Gubinelli a développé, quelques années plus tard, une théorie algébrique des chemins rugueux, en mettant en évidence les structures algébriques, en particulier les algèbres de Hopf des arbres enracinés [46], cachées derrière les constructions rugueuses ([38], [39]). Depuis, ce point de vue est privilégié dans le domaine des chemins rugueux, car il permet un plus haut degré d'abstraction.

Des travaux d'unification et de synthèse de la théorie des chemins rugueux ont été faits par T. Lyons et Z. Qian [67], par P. Friz et N. Victoir [34], P. Friz et M. Hairer [32], A. Lejay [59], et parmi les nombreux apports à l'application de cette théorie, on pourrait citer I. Bailleul [4], L. Coutin [21], P. Gassiat [27], R. Catellier [17], etc.

Malgré le fait que ses outils ne prennent en charge que des objets à régularité positive, la théorie des chemins rugueux a inspiré plusieurs contributions à la résolution de certaines équations aux dérivées partielles stochastiques (EDPS) comme l'équation KPZ [41], l'équation de la chaleur avec un bruit blanc dans le temps [24] ou encore l'équation de type Burgers avec un bruit blanc multiplicatif [44].

Une première généralisation de ces méthodes permettant de traiter la présence de termes à régularité négative a été réalisée en 2012, lorsque M. Gubinelli, P. Imkeller et N. Perkowski ont proposé dans [40] une théorie de multiplication des distributions en utilisant les paraproduits de Bony [10].

En 2013, M. Hairer opère une deuxième généralisation, plus vaste, à travers sa théorie des structures de régularité [42] qui permet en particulier de construire, à l'instar des chemins rugueux, des objets abstraits à plusieurs niveaux. L'une des questions centrales de cette théorie est celle de la renormalisation, dont une théorie algébrique détaillée est présentée dans le récent papier de M. Hairer, L. Zambotti et Y. Bruned [15], les implications pour la théorie des chemins rugueux dans [14] et quelques applications dans [13].

Malgré le fait qu'aussi bien la théorie des chemins rugueux que celle des structures de régularité présentent beaucoup de résultats très abstraits, les deux tendent

à établir un lien étroit avec le calcul stochastique, d'une part en transposant les principaux objets de celui-ci dans de nouvelles structures, d'autre part, en complétant celui-ci par des résultats que les nouveaux outils permettent d'atteindre. Nous espérons y contribuer en employant la théorie des chemins rugueux pour étudier des modèles bien connus sous un nouvel angle.

1.2.2 Journal de bord

Recherche de nouvelles applications pour les chemins rugueux. Initialement, nous cherchions des objets probabilistes dont nous pouvions proposer une étude fondée sur des outils simples de la théorie des chemins rugueux qui, tout en apportant des éléments nouveaux par rapport aux études classiques, seraient accessibles à des non-spécialistes; nous nous intéressions avant tout à ceux qui permettraient de sortir les chemins rugueux du cadre des EDS. Après avoir étudié plusieurs possibilités à travers le mouvement brownien de Liouville ([35]), le mouvement brownien sur le groupe des difféomorphismes du cercle ([69], [30]) ou encore les espaces de Fock ([7], [70]), nous nous sommes concentrés sur l'étude de *modèles discrets*.

Modèles discrets. Notre intérêt pour les modèles discrets (ou les processus discrets) s'est éveillé en découvrant divers avantages que ceux-ci peuvent présenter. D'abord, ils permettent de s'adresser à un public très large : on les utilise aussi bien en modélisation qu'en méthodes numériques, les marches aléatoires branchantes et les arbres aléatoires font toujours couler beaucoup d'encre d'imprimante, certains exemples de chaînes de Markov sont accessibles même aux élèves de master, et on peut souvent en donner une représentation intuitive en faisant un dessin.

Ensuite, le cadre discret a été relativement peu étudié en lien avec les chemins rugueux : le principal but de la théorie étant de résoudre efficacement des EDS, on opère surtout avec des processus continus et on procède rarement par discrétisation (cependant, une méthode a été proposée dans [23]).

Enfin, et surtout, la structure simple de certains modèles (dans \mathbb{R}^2 ou \mathbb{R}^3 , par exemple), doublée d'une représentation géométrique, permet de calculer explicitement la limite du processus discret à la fois en topologie uniforme et en topologie

rugueuse et de montrer ce que cette dernière permet d'apporter en plus par rapport à la première.

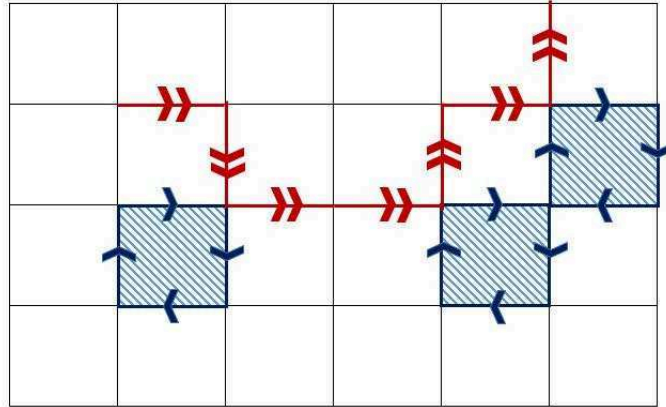


FIGURE 1.1 – MAS sur \mathbb{R}^2 avec une boucle tous les deux pas de temps.

Question 1. *On considère une marche aléatoire simple (MAS) sur \mathbb{Z}^2 et un deuxième processus comme dans la figure 1.1 : une MAS sur \mathbb{Z}^2 à laquelle on ajoute des boucles (hachurées sur le dessin) tous les deux pas de temps. En topologie uniforme les deux processus convergent, après renormalisation, vers le mouvement brownien. Mais est-ce vraiment la même limite ?*

La réponse est non : on peut distinguer les limites de ces processus en topologie rugueuse. Les chemins rugueux permettent d'enregistrer des informations qui disparaissent à la limite en topologie uniforme sous la forme d'objets qui ne dépendent pas de la trajectoire du processus limite. Nous avons choisi de nous intéresser au premier de ces objets non-triviaux, qui est précisément celui qui différencie les limites de la question 1, *l'anomalie d'aire*.

Anomalie d'aire. Illustrons l'anomalie d'aire en analysant les convergences évoquées dans la question 1, inspirée par l'exemple des *bulles d'aire* (micro-perturbations d'une trajectoire qui ne changent pas sa limite, mais qui influencent l'aire accumulée à la limite) utilisé souvent par A. Lejay pour introduire les chemins rugueux (comme dans [59]).

Lorsqu'un processus discret converge vers le mouvement brownien, l'information qui manque au niveau trajectorien, et que les chemins rugueux contiennent, est l'*accumulation d'aire signée*, i.e. l'aire générée par les fluctuations, et en particulier les boucles, du processus (cette aire est définie dans la section 1.3.2 de la présente introduction). Dans le cas de la MAS, cette accumulation d'aire ne donne pas de nouvelles informations par rapport à la trajectoire, le chemin rugueux limite étant composé du mouvement brownien et de son aire signée $(\mathcal{A}_s)_s$ (appelée *aire de Lévy*) :

$$\left(B_{t/6}, \mathcal{A}_{t/6} \right)_t$$

La seule trace des boucles qui persiste à la limite est un ralentissement dans le temps traduit par le facteur $t/6$.

Par contre, dans la limite du modèle représenté dans la figure 1.1, on voit apparaître à la limite un terme déterministe supplémentaire, qui rend compte de l'aire générée par toutes ces boucles disparues dans la trajectoire limite :

$$\left(B_{t/6}, \mathcal{A}_{t/6} + \frac{t}{3} \right)_t$$

Le terme $t/3$ est une anomalie d'aire, car il ne peut pas être déduit de la trajectoire limite (le mouvement brownien). Ceci montre, en particulier, que cette anomalie d'aire n'est pas reliée au drift du niveau trajectorien, la MAS étant centrée (tous les six pas de temps).

L'anomalie d'aire issue de la convergence de modèles discrets sera au cœur de notre démarche et aura une double vocation :

- elle fournira de nouveaux critères de convergence de processus discrets, qui dépassent le cadre du calcul stochastique classique, et un nouveau regard sur les domaines qui opèrent avec des modèles discrets (par exemple, l'étude des *équations aux différences*, présentées dans la section 1.3.5 de l'introduction, et abordées également dans le chapitre 2, section 2.4.1) ;
- pour certains modèles, elle permet d'expliquer certaines difficultés liées au passage à la limite et dont on pense intuitivement qu'ils sont liés à une accumulation de « boucles » dans le processus (par exemple, dans le cas des processus dans un milieu aléatoire, comme supposé dans [5]) ;

- elle permettra d'introduire dans le domaine des chemins rugueux, qui reste très abstrait, avec relativement peu d'illustrations en dehors des EDS, des constructions et des calculs explicites avec l'utilisation d'outils probabilistes et géométriques classiques.

Le modèle présenté dans la figure 1.1 étant trivial, nous avons cherché des processus avec des propriétés markoviennes, avec des corrélations temporelles à faible portée et en nous inspirant des phénomènes produits par un champ magnétique (comme dans [31]). Après avoir construit un premier exemple sur la base de variables de Bernoulli corrélées (présenté dans le chapitre 2, section 2.1.3), nous l'avons généralisé à la classe des *chaînes de Markov sur graphes périodiques*.

Chaînes de Markov sur graphes périodiques. Une fois construites les chaînes de Markov sur graphes périodiques (présentées dans la section 1.3.3 de l'introduction et dans le chapitre 2, section 2.1.4), nous nous sommes rendu compte que notre définition était très proche de celle des modèles issus de divers domaines comme la mécanique statistique, la cristallographie, la biologie (plusieurs exemples sont présentés dans la section 2.1.4 du chapitre 2) et dont la convergence en topologie uniforme (souvent vers le mouvement brownien) est beaucoup étudiée.

L'étude de la convergence en topologie rugueuse a été facilitée par la possibilité d'un découpage en *v.a.i.i.d. plus des fluctuations résiduelles* de ce type de processus (expliqué dans la théorie des pseudo-excursions, chapitre 2, section 2.2.2). Nous avons montré que ces fluctuations résiduelles généraient une anomalie d'aire à la limite.

Question 2. *Comment donner une expression explicite de l'anomalie d'aire ?*

Nous avons répondu à cette question une première fois dans le cadre des chaînes de Markov sur graphes périodiques au chapitre 2, section 2.1.4 (en particulier dans le théorème 2.1.1) dont un résumé est donné dans la section 1.3.7 de l'introduction. Précisons simplement que les chaînes de Markov sur graphes périodiques nous ont permis non seulement de fournir une formule probabiliste avec interprétation géométrique de l'anomalie d'aire, mais aussi de la calculer explicitement pour certains modèles (présentés dans le chapitre 4).

Dans les méthodes que nous avons développées pour cette classe de processus, un rôle-clé est joué par *les chaînes de Markov sur espaces d'états finis* que nous avons extraites de ces modèles, ce qui nous a donné l'idée d'étendre nos résultats à une classe de processus plus grande : *les marches de Markov cachées*.

Marches de Markov cachées. Alors que, dans le cas des chaînes de Markov sur graphes périodiques, c'est le caractère périodique des processus qui nous a permis de déduire l'existence d'une chaîne de Markov sous-jacente pour chacun d'entre eux, les marches de Markov cachées sont définies directement comme des processus dépendant d'une chaîne de Markov (voir la section 1.3.3 de l'introduction ou la section 3.2.1 du chapitre 3). Une chaîne de Markov sur graphe périodique peut également avoir une interprétation en tant que marche de Markov cachée, mais avec une chaîne de Markov sous-jacente souvent différente et plus complexe, ce qui rend les calculs et l'exploitation du côté géométrique du modèle plus difficiles. Nous allons comparer les deux représentations dans la section 1.3.3 de l'introduction et dans le cas de deux modèles concrets dans la section 4.1 du chapitre 4.

Les principaux outils que nous avons développés s'appliquant également aux marches de Markov cachées, nous avons pu généraliser le théorème 2.1.1 à cette classe de processus (théorème 3.2.1 du chapitre 3).

Nous avons ensuite exploré une autre direction de généralisation, inspirée toujours par les "bulles d'aire" d'A. Lejay. Pour étudier la convergence d'un processus discret à valeurs dans \mathbb{R}^d dans la topologie uniforme, nous le plongeons d'abord dans l'espace des fonctions continues en adoptant tacitement l'interpolation linéaire par morceaux. Mais si l'on s'intéressait de plus près à la question des plongements, que se passerait-il ?

Classes d'équivalence pour les plongements. Puisque ce qui nous intéresse est de regarder le processus "de loin", choisir un autre plongement semble inutile tant qu'on reste en topologie uniforme (les plongements sont présentés dans la partie 1.3.4 de l'introduction, et plus en détails dans la section 3.3.3 du chapitre 3). Et pourtant...

Convergence des structures combinatoires des marches de Markov cachées. Ces structures combinatoires nous ont conduit à donner une définition généralisée de la notion de temps d'occupation, classique dans la théorie ergodique, et à étudier *les sommes itérées* des marches de Markov.

Question 4. *Peut-on étudier la convergence des ces structures combinatoires en topologie rugueuse ?*

La réponse est oui, et cela permet également de construire, à partir d'un même modèle discret, des chemins rugueux limites géométriques et non-géométriques.

1.2.3 Structure générale de la thèse

Après avoir exposé la chronologie de nos résultats, nous allons en présenter l'organisation. Le présent travail contient quatre chapitres :

- *l'introduction* (le chapitre 1), où nous présentons de manière synthétique le cadre de nos travaux, nos principaux résultats et leur intérêt. Elle se trouve divisée en deux parties : la première, 1.3, contient les principaux résultats concernant la convergence des chaînes de Markov cachées et l'anomalie d'aire (une partie du chapitre 3, chapitre 2, une partie du chapitre 4) ; la deuxième (1.4) regroupe les résultats concernant les structures combinatoires sous-jacentes aux marches de Markov cachées (les parties restantes des chapitres 3 et 4).
- *le chapitre 2*, où nous construisons les chaînes de Markov sur graphes périodiques, montrons leur dépendance d'une chaîne de Markov sur un espace d'états fini et développons une théorie des pseudo-excursions, avant d'étudier leur convergence en topologie rugueuse et de présenter l'anomalie d'aire.

Nous étudions aussi plusieurs modèles concrets inspirés de la mécanique statistique et présentons un exemple simple d'équation aux différences où apparaît l'anomalie d'aire.

Ce chapitre est notre premier article [64], qui a récemment été accepté pour publication par la revue Stochastic Processes and their Applications.

- *le chapitre 3*, qui est dédié aux résultats généraux de convergence des chemins rugueux associés à des marches de Markov cachées. Nous étudions le terme de l'anomalie d'aire qui apparaît dans ces limites en fonction des plongements choisis et présentons une manière de définir des classes d'équivalence pour ces derniers.

Nous montrons aussi que les marches de Markov cachées sont liées à des objets combinatoires sous-jacents (les temps d'occupation itérés), qui convergent (sous certaines conditions) vers des chemins rugueux non-géométriques.

Nous présentons aussi un lien entre les produits *shuffle* et *quasi-shuffle* et les chemins rugueux géométriques et non-géométriques respectivement. Ensuite, à partir d'un processus discret (toujours une marche de Markov cachée), nous construisons un chemin rugueux canonique (géométrique) par intégrales itérées à partir d'un plongement du processus et un chemin rugueux non-géométrique à partir de sommes itérées, et nous comparons leurs limites, dont la première sera un chemin rugueux au sens de Stratonovich, avec anomalie d'aire, et la deuxième un chemin rugueux au sens d'Itô, avec un drift de niveau deux.

Ce chapitre est notre deuxième article [63] tel qu'il a été soumis sur HAL.

- *le chapitre 4*, où nous avons rassemblé les résultats calculatoires : nous fournissons plusieurs méthodes pour calculer l'anomalie d'aire et la calculons explicitement pour deux modèles, donnons des résultats concernant les espérances de sommes itérées de marches de Markov et leur convergence, etc. Ces calculs, qui ne sont pas encore publiés, illustrent et complètent certains aspects des chapitres précédents.

1.2.4 Notations

- V désigne un espace vectoriel de dimension finie $d \geq 2$;
- $[r]$ désigne la partie entière d'un réel r ;
- \otimes désigne tantôt le produit tensoriel entre deux éléments d'un espace vectoriel V , tantôt une loi de multiplication pour les éléments de $\mathbb{R} \otimes V \oplus \dots \oplus V^{\otimes k}$,

où $k \in \mathbb{N}^*$ (on choisira la bonne définition en fonction des éléments auxquels on applique \otimes);

- $\mathcal{C}^{p-var}([0, T], B)$ est l'espace de fonctions à p -variation finie, à valeurs dans un espace de Banach B ;
- $\mathcal{C}^{\alpha-Höl}([0, T], B)$, avec $\alpha \in]0, 1[$, est l'espace des fonctions α -Hölder, à valeurs dans un espace de Banach B ;
- L'opérateur de dilatation sur $\mathbb{R} \otimes V \oplus \dots \oplus V^{\otimes n}$, où $n \in \mathbb{N}^*$, est l'opérateur de scaling δ_α , $\alpha \in \mathbb{R}$ défini par

$$\delta_\alpha(g^{(1)}, \dots, g^{(n)}) = (\alpha g^{(1)}, \alpha^2 g^{(2)}, \dots, \alpha^n g^{(n)})$$

1.3 Le drift d'aire dans la convergence de modèles markoviens en topologie rugueuse

1.3.1 Éléments de la théorie des chemins rugueux

Comme nous l'avons déjà dit, la théorie des chemins rugueux s'appuie à la fois sur l'analyse et l'algèbre. Nous présenterons les principaux points de repère pour comprendre la nature des objets qu'elle manipule, en nous inspirant de plusieurs sources et en mettant l'accent avant tout sur les outils et les notions dont nous aurons besoin.

Cette section n'est pas indispensable à lire pour ceux qui connaissent déjà la théorie. Le lecteur qui, au contraire, voudra approfondir certains points trouvera une présentation concise de la théorie dans [66] ou dans l'introduction à un ouvrage de référence (par exemple, [34] ou [32]), et une présentation plus détaillée dans [58]. Pour ceux qui veulent étudier la théorie à la fois dans ses aspects algébriques et analytiques, il y a l'ouvrage de référence [34]. Enfin, pour ceux qui s'intéressent également aux structures de régularité, [32] fournit une présentation très claire du domaine, de nombreux exercices (corrigés ou non) et un lien avec la théorie de M. Hairer.

Idée générale.

Dans l'introduction, on avait dit qu'un chemin rugueux était un objet à plusieurs niveaux dont chacun a les propriétés algébriques d'une intégrale itérée.

Lorsqu'un chemin $x : [0, 1] \rightarrow V$ entre dans le cadre de la théorie de Young, ses intégrales itérées sont bien définies et on peut s'en servir directement pour construire le chemin rugueux associé. Ainsi, si x est de régularité $\alpha > 1/2$, pour $n \geq 1$, le *chemin rugueux canonique d'ordre n* , appelé *signature tronquée de x à l'ordre n* , est défini, pour tous les s, t tels que $0 \leq s < t \leq 1$, par :

$$S_n(x)_{s,t} \stackrel{\text{déf}}{=} \left(\int_s^t dx_s, \int_{s < s_1 < s_2 < t} dx_{s_1} \otimes dx_{s_2}, \dots, \int_{s < s_1 < \dots < s_n < t} dx_{s_1} \otimes \dots \otimes dx_{s_n} \right) \quad (1.3)$$

Pour tout $n \geq 2$, l'objet $S_n(x)_{s,t}$ ainsi défini vit dans l'espace $V \oplus V^{\otimes 2} \oplus \dots \oplus V^{\otimes n}$, où \otimes est le produit tensoriel sur V .

Le problème se pose si la régularité α de x est inférieure à $1/2$. Dans ce cas, pour $N \geq \lceil 1/\alpha \rceil$ et pour tout $t \in [0, 1]$, on définit un objet

$$\mathbf{x}_t = (\mathbf{x}_t^{(1)}, \dots, \mathbf{x}_t^{(N)})$$

comme un élément abstrait de $V \oplus V^{\otimes 2} \oplus \dots \oplus V^{\otimes N}$, dont le premier niveau est $x_t - x_0$ et dont les niveaux suivants en dépendent en partie et satisfont certaines règles propres aux intégrales multiples. On peut ainsi construire un chemin continu

$$\begin{aligned} \mathbf{x} : [0, 1] &\rightarrow \bigoplus_{k=1}^N V^{\otimes k} \\ t &\mapsto \mathbf{x}_t \end{aligned}$$

qui sera *un chemin rugueux* correspondant à x dans $V \oplus V^{\otimes 2} \oplus \dots \oplus V^{\otimes N}$.

Remarque. Pour emmagasiner toute l'information nécessaire au passage à la limite, on n'a pas besoin d'aller au-delà du niveau $\lceil 1/\alpha \rceil$ (par exemple, pour approcher le mouvement brownien dont la régularité est $1/2^-$, on va construire des chemins rugueux à deux niveaux), car tous les niveaux supérieurs peuvent être entièrement déduits des précédents. Ce résultat classique de la théorie des chemins

<i>Stratonovich</i>	<i>Itô</i>
- surtout pour des processus de régularité $\alpha \geq 1/3$	- surtout pour des processus de régularité $\alpha < 1/3$
- a besoin de « regarder dans le futur »	- ne s'appuie pas sur des prévisions
- satisfait le théorème de dérivation des fonctions composées et l'intégration par partie	- a besoin de termes additionnels dans les niveaux des chemins rugueux, et ne satisfait donc pas à ces règles
- interpolation géodésique	- interpolation constante par morceaux

TABLE 1.1 – Intégrale d'Itô vs. intégrale de Stratonovich dans le cadre des chemins rugueux

rugueux sera énoncé de manière formelle dans la proposition 1.3.1. Maintenant

ce n'est plus le chemin rugueux qu'on définit à partir des intégrales itérées (ou multiples), ce sont les intégrales itérées de x qu'on peut définir grâce au chemin rugueux comme suit :

$$\int_{0 < s_1 < \dots < s_i < t} x_{s_1} \otimes \dots \otimes dx_{s_i} \stackrel{\text{déf}}{=} \mathbf{x}_t^{(i)} \quad (1.4)$$

Dans le cadre probabiliste, les règles auxquelles les niveaux des chemins rugueux devront satisfaire dépendent aussi du type d'intégration qu'on choisit : Itô ou Stratonovich. L'intégration par parties s'applique au second, mais pas au premier. Ce critère permet aussi de séparer les chemins rugueux en deux grands groupes : *les chemins rugueux géométriques*, correspondant à l'intégration de Stratonovich, et *les chemins rugueux non-géométriques*, correspondant, en particulier, à l'intégration d'Itô. Le tableau 1.1 donne quelques points de comparaison entre les deux types de construction d'intégrales.

Ainsi, les chemins rugueux géométriques ont ceci de particulier que *les parties symétriques de leurs niveaux peuvent être déduites des niveaux antérieurs*, et donc on peut adopter une écriture simplifiée en les enlevant. Par exemple, le chemin

rugueux (géométrique) du mouvement brownien $(B_t)_t$ dans \mathbb{R}^d

$$\mathbb{B}_t = (B_t, \int_{0 < s_1 < s_2 < t} \circ dB_{s_1} \otimes \circ dB_{s_2})$$

peut se réécrire en ne gardant que l'information antisymétrique de la deuxième composante

$$\mathbb{B}_t = (B_t, \mathcal{A}_t) \tag{1.5}$$

où $\mathcal{A}_t = 1/2 \int_{0 < s_1 < s_2 < t} dB_{s_1} \otimes dB_{s_2} - dB_{s_2} \otimes dB_{s_1}$ est précisément l'aire de Lévy qu'on avait rencontrée dans la section 1.2.2, en étudiant la convergence de la question 1.

Algèbres tensorielles et chemins rugueux.

Tout d'abord, on va décrire de manière plus exacte l'espace où un chemin rugueux prend ses valeurs. Soit, comme avant, V un espace vectoriel de dimension finie. On a vu qu'un chemin rugueux défini à partir d'un chemin dans V est donné par (1.4) ou (1.3), et son k -ième niveau est donc un élément de $V^{\otimes k}$. Pour $a_i \in V^{\otimes i}$, on a

$$a = \bigoplus_{k=1}^n a_k \simeq (a_1, \dots, a_n)$$

et, par conséquent, un chemin rugueux peut être vu comme une somme (directe) finie de produits tensoriels, ce qui est analogue à un développement de Taylor d'une fonction à plusieurs variables.

Pour définir proprement les opérations sur les chemins rugueux, on va passer par l'espace

$$T^{(n)}(V) = \bigoplus_{k=0}^n V^{\otimes k} \tag{1.6}$$

avec $V^{\otimes 0} = \mathbb{R}$. On munit $T^{(n)}(V)$ d'une structure d'algèbre en définissant les opérations internes suivantes :

- *Addition* : $\forall a, b \in T^{(n)}(V), a + b = (a_0 + b_0, a_1 + b_1, \dots, a_n + b_n)$
- *Multiplication* : $\forall a, b \in T^{(n)}(V), a \otimes b = (c_0, c_1, \dots, c_n),$ où $c_k = \sum_{i=0}^k a_i \otimes b_{k-i}$

et de la loi externe

$$\forall \lambda \in \mathbb{R}, \lambda a = (\lambda a_0, \lambda a_1, \dots).$$

On a alors la propriété suivante, qui met en relief l'aspect algébrique des chemins rugueux.

Propriété 1.3.1. *L'ensemble $(T^{(n)}(V), \otimes)$ est une algèbre non-commutative, dont l'unité (l'élément neutre) est $\mathbf{1} = (1, 0, 0, \dots)$.*

Comme l'élément du niveau 0 (appartenant à $V^{\otimes 0}$) est redondant dans le cas des chemins rugueux, il nous suffit de considérer l'ensemble

$$T_1^{(n)}(V) = \{a \in T^{(n)}(V) : a_0 = 1\} \quad (1.7)$$

qui satisfait à la propriété ci-dessous.

Proposition 1.3.1. *$(T_1^{(n)}(V), \otimes)$ est un groupe.*

Puisqu'on a l'isomorphisme

$$T_1^{(n)}(V) \simeq \bigoplus_{k=1}^n V^{\otimes k}$$

par abus de langage, on omettra dans l'écriture l'élément du niveau 0 (i.e. on écrira (a_1, \dots, a_n) au lieu de $(1, a_1, \dots, a_n)$).

Comme les formules (1.4) et (1.3) donnent des cas particuliers de chemins continus dans $T_1^{(n)}(V)$, on peut donner une première définition générale et intuitive d'un chemin rugueux, inspirée de [59].

Définition 1.3.1. *Un chemin rugueux est un chemin continu dans $T_1^{(n)}(V)$ pour un certain $n \geq 1$, i.e. un élément de $\mathcal{C}([0, 1], T_1^{(n)}(V))$.*

Ceci implique, entre autres, qu'on peut parler de chemins rugueux sans nécessairement partir d'un chemin dans V .

À partir de la proposition 1.3.1, on peut déduire quelques caractéristiques des chemins rugueux.

Propriété 1.3.2. *Soit $n \geq 2$ et soit $(\mathbf{x}_t)_{t \in [0, 1]}$ un chemin rugueux dans $T_1^{(n)}(V)$.*

- Les incréments de \mathbf{x} sont donnés par :

$$\forall 0 \leq s < t \leq 1, \mathbf{x}_{s,t} = \mathbf{x}_s^{-1} \otimes \mathbf{x}_t$$

où on note par g^{-1} l'inverse de $g \in T_1^{(n)}(V)$.

- Les incréments de \mathbf{x} satisfont la relation de Chen ([18]) :

$$\forall 0 \leq s < u < t \leq 1, \mathbf{x}_{s,t} = \mathbf{x}_{s,u} \otimes \mathbf{x}_{u,t}.$$

Remarque. La relation de Chen dans la propriété 1.3.2 se montre à partir de la définition des incréments donnée au même endroit.

En ce qui concerne les chemins rugueux géométriques, comme les parties symétriques de leurs niveaux dépendent des niveaux précédents, ils sont à valeurs dans $G^n(V)$, un sous-groupe de $T_1^{(n)}(V)$ formé de tous les éléments auxquels cette condition s'applique. Par exemple, dans le cas $n = 2$, $G^2(V)$ peut être défini à partir de

$$\forall (g^1, g^2) \in G^2(V), \text{Sym}(g^2) = \frac{1}{2}g^1 \otimes g^1.$$

On peut alors définir un chemin rugueux géométrique de la manière suivante.

Définition 1.3.2 ([59]). *Un chemin rugueux géométrique est un chemin continu sur $G^n(V)$ pour un certain $n \geq 1$, i.e. un élément de $\mathcal{C}([0, 1], G^n(V))$.*

Nous avons déjà mentionné qu'un chemin géométrique peut être représenté aussi bien avec que sans sa partie symétrique. Dans ce cas, choisir une représentation revient à choisir la loi qu'on met sur $G^2(V)$.

Ainsi, on peut munir $G^n(V)$ de la loi « \otimes » héritée de $T^{(n)}(V)$, ou alors de la loi antisymétrique \wedge , qui permet d'enlever la partie symétrique des éléments. Ainsi, sur $G^2(V)$, ceci se traduit, dans le premier cas, par

$$\forall (a, b), (a', b') \in G^2(V), (a, b) \wedge (a', b') = (a + a', b + b' + a \otimes a')$$

et, dans le deuxième cas, par

$$\forall (a, b), (a', b') \in G^2(V), (a, b) \wedge (a', b') = (a + a', b + b' + \frac{1}{2}(a \otimes a' - a' \otimes a)). \quad (1.8)$$

Remarque. En utilisant 1.8, on peut fournir une interprétation géométrique pour le deuxième niveau de $G^2(V)$: ce sera *l'aire signée* du chemin enregistré au premier niveau. Par ailleurs, la manière dont le deuxième niveau est formé dans (1.8) rappelle la formule de concaténation de deux aires signées donnée dans (1.22).

En particulier, sur $(G^2(V), \wedge)$, l'expression du mouvement brownien rugueux est donnée par la formule (1.5).

Avant de pouvoir étudier la convergence des chemins rugueux, il nous faut définir les topologies adéquates, ainsi qu'un moyen de mesurer leur régularité.

Topologies sur $T_1^{(N)}(V)$ et $G^N(V)$ et chemins rugueux α -Hölder.

Vu sa structure d'espace vectoriel, on peut munir $T_1^{(N)}(V)$ de la norme homogène (i.e. pour tous $g \in T^{(N)}(V)$, $\lambda > 0$, $|||\delta_\lambda g||| = \lambda |||g|||$) définie par :

$$\forall g \in T_1^{(N)}(V), |||g||| = \max_{i=1, \dots, N} \{ |g^{(i)}|_{V^{\otimes i}}^{1/i} \}. \quad (1.9)$$

Cette norme nous permet de définir l'ensemble des fonctions α -Hölder sur $T_1^{(N)}(V)$, ainsi qu'une norme sur cet espace.

Définition 1.3.3. Pour $\alpha \in]0, 1[$ et $N \geq 1$, on définit $\mathcal{C}^{\alpha-H\ddot{o}l}([0, 1], T_1^{(N)}(V))$, l'espace des chemins α -Hölder à valeurs dans $T_1^{(N)}(V)$, par

$$\forall \mathbf{x} \in \mathcal{C}^{\alpha-H\ddot{o}l}([0, 1], T_1^{(N)}(V)), \sup_{0 \leq u < v \leq 1} \frac{|||\mathbf{x}_{u,v}|||}{|u - v|^\alpha} < \infty. \quad (1.10)$$

De plus,

$$|||\mathbf{x}|||_\alpha = \max_{i=1, \dots, N} \sup_{0 \leq u < v \leq 1} \frac{|x_{u,v}^{(i)}|_{V^{\otimes i}}^{1/i}}{|u - v|^\alpha} \quad (1.11)$$

définit une norme (homogène) et, en supposant $x_0 = y_0$,

$$d_\alpha(\mathbf{x}, \mathbf{y}) = \max_{i=1, \dots, N} \sup_{0 \leq u < v \leq 1} \frac{|(\mathbf{y}_{u,v}^{-1} \otimes \mathbf{x}_{u,v})^{(i)}|_{V^{\otimes i}}}{|u - v|^{\alpha i}} \quad (1.12)$$

définit une distance (inhomogène) sur cet espace.

Le groupe $G^N(V)$ peut également être muni de la norme (1.9). Néanmoins, on lui préfère une autre (en particulier pour des raisons de topologie) qui s'appuie sur une autre définition de $G^N(V)$ faite à partir des signatures tronquées à l'ordre N des chemins à variation finie.

Définition 1 (définition 3.3.3 et propriété). *On pose*

$$G^N(V) = \{S_N(\gamma)_{0,1} : \gamma \in \mathcal{C}^{1-var}([0, 1], V)\}$$

où $S_N(\gamma)_{0,1}$ est comme dans (1.3).

Elle nous dit, en particulier, que chaque élément de $T_1^{(n)}(V)$ dont les niveaux respectent la règle de l'intégration par parties est la représentation de la signature d'un chemin dans V à variation finie.

La topologie de $G^N(V)$ est déterminée alors par *la norme de Carnot-Caratheodory* :

$$\forall g \in G^N(V), \|g\| := \inf \left\{ \int_0^1 |dx| : x \in \mathcal{C}^{1-var}([0, 1], V) \text{ and } S_N(x)_{0,1} = g \right\}$$

où $|\cdot|_V$ est la norme euclidienne induite sur V .

La norme de Carnot-Caratheodory est :

- homogène : $\|\delta_\lambda g\| = |\lambda| \|g\|$ for $\lambda \in \mathbb{R}$
- symétrique : $\|g\| = \|g^{-1}\|$
- sous-additive : $\|g \otimes h\| \leq \|g\| + \|h\|$.

De plus, elle définit une distance \mathbf{d} sur $G^N(V)$ par

$$\begin{aligned} \mathbf{d} : G^N(V) \times G^N(V) &\rightarrow \mathbb{R}_+ \\ (g, h) &\mapsto \|g^{-1} \otimes h\| \end{aligned}$$

Muni de cette distance, $G^N(V)$ acquiert une structure bien particulière.

Propriété 1.3.3. $(G^N(V), \mathbf{d})$ est un espace géodésique.

Rappelons brièvement ce que ceci veut dire. Lorsqu'on a un espace métrique, on appelle *géodésique* une courbe (pas nécessairement unique) qui minimise la distance entre deux points de cet espace. On parle d'*espace géodésique* pour désigner un espace métrique où à chaque couple de points correspond (au moins) une géodésique.

L'exemple le plus immédiat d'un espace géodésique est tout espace vectoriel de dimension finie muni de la norme euclidienne. Dans ce cas, les géodésiques sont les lignes droites (la notion de géodésique est en fait une généralisation de la ligne droite à des espaces métriques plus généraux). Le fait que, pour tout $N \geq 2$, $(G^N(V), \mathbf{d})$ est un espace géodésique est montré dans le chapitre 7 de [34]. Pour $g, h \in G^N(V)$, la géodésique qui les relie correspond à l'incrément $g^{-1} \otimes h$.

On a le lien suivant, très utile, entre les normes qu'on peut mettre sur $G^N(V)$, donc, dans notre cas, entre la norme Carnot-Carathéodory et la norme héritée de $T^{(n)}(V)$.

Propriété 1.3.4. Toutes les normes homogènes sont équivalentes sur $G^N(V)$.

La norme de Carnot-Carathéodory présente plus de propriétés utiles, mais est plus difficile à calculer. Grâce à la propriété 1.3.4, on peut fournir une estimation assez facile pour elle.

Propriété 1.3.5. Il existe deux constantes $c, C > 0$ telles qu'on ait l'encadrement suivant pour la norme Carnot-Carathéodory $\|\cdot\|$:

$$\forall g \in G^N(V), c \sum_{i=1}^N |g^{(i)}|_{V^{\otimes i}}^{1/i} \leq \|g\| \leq C \sum_{i=1}^N |g^{(i)}|_{V^{\otimes i}}^{1/i} \quad (1.13)$$

ou, alternativement, il existe deux constantes $c', C' > 0$ telles qu'on ait l'encadrement suivant pour la norme Carnot-Carathéodory $\|\cdot\|$:

$$\forall g \in G^N(V), c' \|\|g\|\| \leq \|g\| \leq C' \|\|g\|\| \quad (1.14)$$

où $\|\| \cdot \|\|$ est comme dans (1.9).

Nous avons maintenant les outils pour énoncer la définition des chemins rugueux géométriques α -Hölder.

Définition 1.3.4. *Pour $\alpha \in]0, 1[$ et $N \geq 1$, on définit $\mathcal{C}^{\alpha-H\ddot{o}l}([0, 1], G^N(V))$, l'espace des chemins α -Hölder à valeurs dans $G^N(V)$, par*

$$\forall \mathbf{x} \in \mathcal{C}^{\alpha-H\ddot{o}l}([0, 1], G^N(V)), \quad \sup_{0 \leq u < v \leq 1} \frac{\|\mathbf{x}_u^{-1} \otimes \mathbf{x}_v\|}{|u - v|^\alpha} < \infty \quad (1.15)$$

De plus,

$$\|\mathbf{x}\|_\alpha = \sup_{0 \leq u < v \leq 1} \frac{\|\mathbf{x}_{u,v}\|}{|u - v|^\alpha} \quad (1.16)$$

définit une norme (homogène) sur $\mathcal{C}^{\alpha-H\ddot{o}l}([0, 1], G^N(V))$ dont on peut déduire une métrique (homogène)

$$\mathbf{d}_\alpha(\mathbf{x}, \mathbf{y}) = \|\mathbf{y}^{-1} \otimes \mathbf{x}\|_\alpha = \sup_{0 \leq u < v \leq 1} \frac{\mathbf{d}(\mathbf{x}_{u,v}, \mathbf{y}_{u,v})}{|u - v|^\alpha} \quad (1.17)$$

où \mathbf{d} est la distance induite sur $G^N(V)$ par la norme Carnot-Carathéodory.

Remarque. Les chemins rugueux géométriques α -Hölder font bien évidemment partie des chemins rugueux α -Hölder de la définition 1.3.3.

Pour finir, nous allons donner le résultat annoncé antérieurement, qui justifie le lien entre la régularité et le nombre de niveaux d'un chemin rugueux et dit en particulier que, pour tout élément de $\mathcal{C}^{\alpha-H\ddot{o}l}([0, 1], G^N(V))$ avec $N \geq \lfloor 1/\alpha \rfloor$, il suffit de considérer seulement ses $\lfloor 1/\alpha \rfloor$ premiers niveaux.

Proposition 1.3.1. *Soit $\alpha > 0$. Pour $N \geq \lfloor 1/\alpha \rfloor$, l'application*

$$\begin{aligned} \pi_\alpha : \mathcal{C}^{\alpha-H\ddot{o}l}([0, 1], G^N(V)) &\rightarrow \mathcal{C}^{\alpha-H\ddot{o}l}([0, 1], G^{\lfloor 1/\alpha \rfloor}(V)) \\ \mathbf{x} = (\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(N)}) &\mapsto (\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(\lfloor 1/\alpha \rfloor)}) \end{aligned}$$

est une bijection. En particulier, un chemin dans V qui est α -Hölder a besoin de $\lfloor 1/\alpha \rfloor$ niveaux pour enregistrer toute l'information pertinente.

Suppléments.

Dans ce qui précède, nous avons présenté des notions et des outils concernant les chemins rugueux hölderiens qui nous intéresseront par la suite. Pourtant, initialement la théorie des chemins rugueux a été développée pour des chemins à p -variation finie ($p \geq 1$) et seulement ensuite étendue par M. Gubinelli aux chemins hölderiens dans [38]. Revenons aux sources de la théorie et présentons brièvement le cadre choisi par T. Lyons.

Pour définir des chemins rugueux à p -variation finie, on peut utiliser la définition, plus générale, de chemins à p -variation finie dans un groupe.

Définition 1.3.5 ([33]). *Soit $(G, \otimes, \mathbf{1})$ un groupe muni d'une norme homogène $\|\cdot\|$, i.e. telle que*

$$\forall g \in G, \forall \alpha \in \mathbb{R}_+, \|T_\alpha g\| = \alpha \|g\|$$

où T_α est un opérateur de scaling.

Soit $x : [0, 1] \rightarrow G$ un chemin continu. Pour $p \in [1, \infty[$, on dit que x est à p -variation finie si

$$\sup_{(t_i) \subset \mathcal{D}_{[0,1]}} \left(\sum_i \|x_{t_{i+1}} - x_{t_i}\|^p \right)^{\frac{1}{p}} < \infty \quad (1.18)$$

où $\mathcal{D}_{[0,1]}$ est l'ensemble des subdivisions de $[0, 1]$.

On note $\mathcal{C}^{p\text{-var}}([0, 1], G)$ l'ensemble des chemins continus $x : [0, T] \rightarrow G$ qui satisfont (1.18).

La définition 1.3.5 nous permet de parler de chemins rugueux (resp. de chemins rugueux géométriques) de degré n à p -variation finie grâce à la norme (homogène) (1.9) définie sur $T_1^{(n)}(V)$ (resp. la norme Carnot-Carathéodory sur $G^n(V)$).

Définition 1.3.6. *Pour $p \geq 1$, un p -chemin rugueux sur V est un élément $\mathbf{x} \in \mathcal{C}^{p\text{-var}}([0, 1], T^{(\lfloor p \rfloor)}(V))$. Un p -chemin rugueux géométrique sur V est un élément $\mathbf{x} \in \mathcal{C}^{p\text{-var}}([0, 1], G^{\lfloor p \rfloor}(V))$.*

Pour un chemin à p -variation finie, tout comme pour les chemins hölderiens dans la proposition 1.3.1, on peut établir un lien entre p et le nombre de niveaux suffisant à enregistrer l'information nécessaire.

Propriété 1.3.6. *Pour tout $n \geq p$, il existe une bijection*

$$\pi_p : \mathcal{C}^{p-var}([0, 1], T_1^{(n)}(V)) \rightarrow \mathcal{C}^{p-var}([0, 1], T_1^{([p])}(V)).$$

Remarque. Même si les espaces $\mathcal{C}^{1/p-Höl}([0, 1], G^N(V))$ et $\mathcal{C}^{p-var}([0, 1], G^N(V))$ ne sont pas les mêmes, les principaux théorèmes sur les chemins rugueux s'appliquent aux deux de manière équivalente. Pour les détails de cette identification, le lecteur pourra voir [34].

1.3.2 Aire de Lévy et drift d'aire

Lorsque nous avons défini la loi antisymétrique (1.8) sur $G^2(V)$, nous avons mentionné que le niveau deux, dans ce cas, pouvait être interprété comme une aire signée correspondant au premier niveau (en particulier, l'aire de Lévy est l'aire signée correspondant au mouvement brownien multidimensionnel).

Surtout, nous allons voir - et c'est là l'un des points importants du présent travail - que les limites de certains modèles discrets qui convergent tous vers le mouvement brownien en topologie uniforme ne sont pas les mêmes en topologie rugueuse : elles se différencient par un processus déterministe linéaire, dont la pente est précisément *l'anomalie* ou *le drift d'aire*. L'étude de l'anomalie d'aire dans le cadre des chemins rugueux est au cœur d'une grande partie du présent travail (par exemple, théorèmes 3.2.1 et 3.2.3 et applications dans le 4).

Cadre déterministe : aire signée.

La manière la plus simple d'expliquer l'aire signée est de l'illustrer à partir d'un exemple facile dans le plan.

Dans la figure 1.3, pour les courbes $\gamma_1 : [0, 1] \rightarrow \mathbb{R}^2$ et $\gamma_2 : [0, 1] \rightarrow \mathbb{R}^2$, l'aire signée est simplement l'aire du secteur délimité par γ_1 (resp. γ_2) et le segment $[\gamma_1(s); \gamma_1(t)]$ (resp. $[\gamma_2(s); \gamma_2(t)]$) prise avec le signe qui dépend de la position de la courbe par rapport au segment. Plus généralement, pour une courbe comme $\gamma_3 : [0, 1] \rightarrow \mathbb{R}^2$, l'aire signée correspondante est la somme des aires signées représentées par les secteurs hachurés.

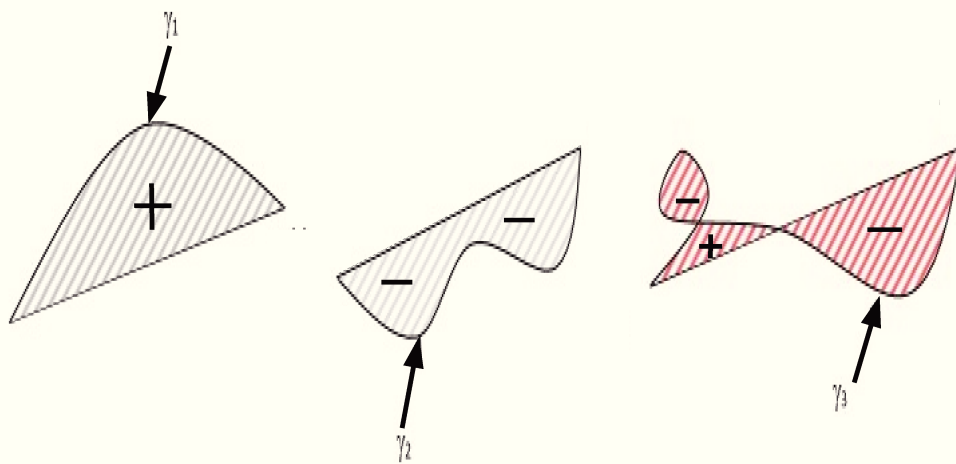


FIGURE 1.3 – Aire signée de trois courbes $\gamma_1, \gamma_2, \gamma_3 : [s, t] \rightarrow \mathbb{R}^2$ (avec les signes correspondants).

Même si la notion d'aire est avant tout géométrique, elle est étroitement liée à la théorie de l'intégration, et sa définition analytique concerne avant tout les processus qui répondent aux critères de la théorie classique de l'intégration.

Définition 1.3.7 (aire signée). *Soit $x : [0, 1] \rightarrow \mathbb{R}^d$, $(x_t)_t = (x_t^{(1)}, \dots, x_t^{(d)})_t$ un chemin dans \mathbb{R}^d pour lequel $\int_s^t x_u^{(i)} dx_u^{(j)}$ est bien définie pour tous $1 \leq i, j \leq d$ et $0 \leq s < t \leq 1$. Alors l'aire signée entre les temps s et t avec $0 \leq s < t \leq 1$ est la matrice antisymétrique $a \in \mathcal{M}(d, d)$ donnée par*

$$a_{s,t}^{(ij)}(x) = \frac{1}{2} \int_{s < s_1 < s_2 < t} dx_{s_1}^{(i)} dx_{s_2}^{(j)} - dx_{s_1}^{(j)} dx_{s_2}^{(i)} \quad (1.19)$$

Un cas particulier dans la construction de l'aire signée sont les chemins linéaires par morceaux, comme le chemin $\ell : [0, 1] \rightarrow \mathbb{R}^2$ dans la figure 1.4. Dans ce cas, la

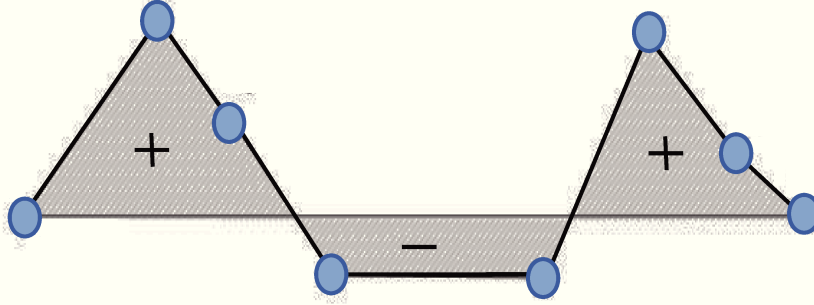


FIGURE 1.4 – Aire signée d'un chemin linéaire par morceaux.

formule (1.19) devient, de fait, une somme double sur les accroissements linéaires de ℓ :

$$a_{s,t}^{(ij)}(\ell) = \frac{1}{2} \sum_{1 \leq k < m \leq n} (\Delta \ell_{s_k})^{(i)} (\Delta \ell_{s_m})^{(j)} - (\Delta \ell_{s_k})^{(j)} (\Delta \ell_{s_m})^{(i)} \quad (1.20)$$

où $\Delta \ell_{s_m} = \ell_{s_m} - \ell_{s_{m-1}}$, et avec $n = 8$ pour le chemin ℓ représenté dans la figure 1.4.

Propriété 1.3.7. *En tant qu'objet géométrique, l'aire signée d'un chemin $\gamma : [0, 1] \rightarrow \mathbb{R}^d$ a plusieurs propriétés simples.*

- *Invariance par translation par un vecteur $v \in \mathbb{R}^d$: $a_{s,t}(\gamma + v) = a_{s,t}(\gamma)$.*
- *Invariance par rotation par un angle $\theta \in]0, 2\pi[$: $a_{s,t}(\theta(\gamma)) = a_{s,t}(\gamma)$.*
- *Action de la dilatation : $\forall c > 0$, $a_{s,t}(c\gamma) = c^2 a_{s,t}(\gamma)$.*
- *Action de la symétrie : étant donné γ' symétrique de γ par rapport à un hyperplan, $a_{s,t}(\gamma') = -a_{s,t}(\gamma)$.*

Une propriété très importante, du point de vue à la fois géométrique et analytique, est donnée par *la loi de concaténation des aires*. Prenons encore une fois un exemple dans le plan. À partir de deux chemins $\gamma_1 : [s, t] \rightarrow \mathbb{R}^2$ et $\gamma_2 : [u, v] \rightarrow \mathbb{R}^2$, on peut construire un chemin $\gamma : [s, t + v - u] \rightarrow \mathbb{R}^2$ comme $\gamma = \gamma_1 \cdot \gamma_2$, où « \cdot » est l'opérateur de concaténation des chemins : on va translater γ_2 de $\gamma_1(t) - \gamma_2(u)$. Alors, grâce à la propriété de l'invariance de l'aire par translation, l'aire signée de γ sera la somme des aires signées de γ_1 , de γ_2 et du triangle déterminé par $\gamma(s)$, $\gamma(t)$ et $\gamma(t + v - u)$.

Dans la figure 1.5, l'aire signée correspondant à γ est la somme des deux régions rayées vers la droite et de la région rayée vers la gauche (chacune avec le signe correspondant). Du point de vue analytique, cela se traduit par la formule :

$$\begin{aligned} a_{s,t+v-u}^{(ij)}(\gamma) &= a_{s,t}^{(ij)}(\gamma_1) + a_{u,v}^{(ij)}(\gamma_2) + \frac{1}{2} \left(\gamma_{s,t}^{(i)} \gamma_{t,t+v-u}^{(j)} - \gamma_{s,t}^{(j)} \gamma_{t,t+v-u}^{(i)} \right) \\ &= a_{s,t}^{(ij)}(\gamma) + a_{t,t+v-u}^{(ij)}(\gamma) + \frac{1}{2} \left(\gamma_{s,t}^{(i)} \gamma_{t,t+v-u}^{(j)} - \gamma_{s,t}^{(j)} \gamma_{t,t+v-u}^{(i)} \right) \end{aligned} \quad (1.21)$$

où $\gamma_{s,t} = \gamma(t) - \gamma(s)$. On peut en déduire la formule générale de décomposition de l'aire signée d'un chemin $\gamma : [0, 1] \rightarrow \mathbb{R}^d$:

$$a_{s,t}^{(ij)}(\gamma) = a_{s,u}^{(ij)}(\gamma) + a_{u,t}^{(ij)}(\gamma) + \frac{1}{2} \left(\gamma_{s,u}^{(i)} \gamma_{u,t}^{(j)} - \gamma_{s,u}^{(j)} \gamma_{u,t}^{(i)} \right) \quad (1.22)$$

avec $s < u < t$.

Cadre déterministe : drift d'aire.

Contrairement à l'aire signée, le drift d'aire est une notion liée à la convergence de chemin, et plus précisément à la question suivante : si une suite de chemins converge, leurs aires signées convergent-elles vers l'aire signée du chemin limite ?

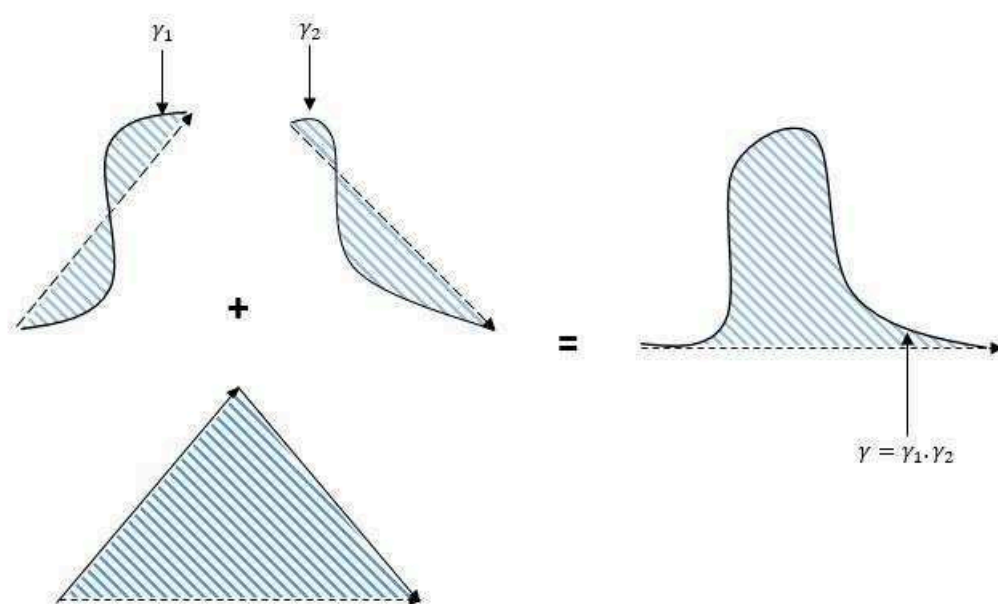


FIGURE 1.5 – La concaténation de deux chemins et de leurs aires.

Prenons un exemple classique, celui des bulles d'aire, utilisé souvent par A. Lejay. On considère la suite de fonctions $\gamma_n : [0, 1] \rightarrow \mathbb{R}^2$, $\gamma_n(t) = (\sin(n^2t)/n, \cos(n^2t)/n)$. Il est évident qu'en topologie uniforme, γ_n converge vers 0. Regardons maintenant les composantes de l'aire signée, par exemple $a_{s,t}^{(12)}(\gamma_n)$:

$$a_{s,t}^{(12)}(\gamma_n) = \frac{1}{2} \int_s^t \frac{\sin(n^2u)}{n} d\left(\frac{\cos(n^2u)}{n}\right) - \frac{\cos(n^2u)}{n} d\left(\frac{\sin(n^2u)}{n}\right) \rightarrow \frac{1}{2}(t-s)$$

On voit qu'elle ne converge surtout pas vers $a_{s,t}^{(12)}(0)$: en même temps que γ_n converge vers 0, le mouvement en spirale génère une accumulation d'aire qui n'est pas perceptible en topologie uniforme, mais sera visible lorsqu'on étudie la limite de l'aire signée. Cette aire qu'on accumule est exactement *le drift d'aire*.

Définition 1.3.8 (drift d'aire). *Pour une suite de chemins $\gamma_n : [0, 1] \rightarrow \mathbb{R}^d$ qui convergent en topologie uniforme vers $\gamma : [0, 1] \rightarrow \mathbb{R}^d$, on appelle drift ou anomalie d'aire la différence entre l'aire signée de γ , $a_{s,t}(\gamma)$, et la limite de la suite constituée des aires signées $a_{s,t}(\gamma_n)$.*

La définition 1.3.8 indique à la fois l'utilité et l'interprétation du drift d'aire : il enregistre l'information sur la suite de chemins qui se perd à la limite au niveau trajectorien. Ce constat nous fournit aussi le lien avec les chemins rugueux, dont le but est précisément d'enregistrer l'information manquante.

Le drift d'aire est un des objets-clé du présent travail, car il permet de donner une application intéressante des chemins rugueux à la convergence de modèles discrets. À travers des calculs explicites de la limite rugueuse et en particulier du drift d'aire, on pourra :

- distinguer des limites identiques au niveau trajectorien (pour la convergence en topologie uniforme) en fonction du drift d'aire ;
- étudier « à la main » la construction des chemins rugueux limite pour une famille particulière de processus ;
- montrer explicitement, même pour les non-spécialistes du domaine, ce que les chemins rugueux peuvent apporter en plus au calcul stochastique grâce à ces objets « enrichis ».

Cadre probabiliste : aire de Lévy.

Définitions, historique, loi. On va maintenant transposer les objets de la section précédente dans un cadre probabiliste. Dans un cadre probabiliste, l'équivalent de l'aire signée est l'aire stochastique :

Définition 1.3.9 (aire stochastique). *Soit $(X_t)_t$ un processus aléatoire à valeurs dans \mathbb{R}^d tel que $\int_s^t X_u^{(i)} dX_v^{(j)}$ ait un sens pour tous $1 \leq i, j \leq d$. On appelle aire stochastique l'aire signée (aléatoire) de $(X_t)_t$.*

On arrive maintenant à un objet qui a un rôle central dans le présent travail : l'aire de Lévy.

Définition 1.3.10. *Soit $(B_t)_t$ un mouvement brownien standard sur \mathbb{R}^d . On appelle aire de Lévy l'aire stochastique du $(B_t)_t$. L'aire de Lévy \mathcal{A} est donnée par la formule :*

$$\mathcal{A}_{s,t}^{(ij)} = A_{s,t}^{(ij)}(B) = \frac{1}{2} \int_{s < u_1 < u_2 < t} dB_{u_1}^{(i)} dB_{u_2}^{(j)} - dB_{u_1}^{(j)} dB_{u_2}^{(i)}. \quad (1.23)$$

La première construction de l'aire de Lévy est due à Paul Lévy et a été faite comme suit. Soient $(X_n^{(1)})_n$ and $(X_n^{(2)})_n$ deux suites de chemins linéaires par morceaux entre les points dyadiques $\frac{k}{2^n}$, et telles que, pour $i = 1, 2$, $X^{(i)}(n)(\frac{k}{2^n}) = B_{\frac{k}{2^n}}^{(i)}$, avec $B^{(1)}$ et $B^{(2)}$ deux mouvements browniens standards. Paul Lévy a montré que le processus

$$A_{s,t}(X_n) = \frac{1}{2} \iint_{s < u_1 < u_2 < t} dX_n^{(1)}(u_1) dX_n^{(2)}(u_2) - dX_n^{(2)}(u_1) dX_n^{(1)}(u_2)$$

converge vers l'aire de Lévy $\mathcal{A}_{s,t}^{(12)} = \frac{1}{2} \iint_{s < u_1 < u_2 < t} dB_{u_1}^1 dB_{u_2}^2 - dB_{u_1}^2 dB_{u_2}^1$ (l'interpolation linéaire peut également être utilisée pour calculer l'aire stochastique d'autres processus gaussiens, dont les incréments satisfont certaines conditions, comme expliqué dans [34] et comme montré dans [21]).

Avec $\theta \in \mathbb{R}$ et $z \in \mathbb{C}$, la loi de l'aire de Lévy est donnée par

$$\mathbb{E}[e^{\theta \mathcal{A}_{0,1}^{(12)}} | B_1 = z] = \frac{\theta}{\sin \theta} e^{-\frac{|z|^2}{2}(\theta \coth \theta - 1)}$$

et a été calculée d'abord par P. Lévy dans [68], ensuite par M. Yor dans [81] en utilisant les processus de Bessel, ou encore par D. Levin et M. Wildon dans [61] à l'aide des produits *shuffle* que nous présenterons dans la section 1.4.3 dans l'introduction, et aussi dans le chapitre 3.

L'aire de Lévy a également été définie pour un mouvement brownien sur un espace de Banach par M. Ledoux, T. Lyons et Z. Qian dans [57], pour un mouvement brownien libre par M. Capitaine et C. Donati-Martin dans [16] (résultat complété par N. Victoir dans [78]), et récemment pour le q -mouvement brownien par A. Deya et R. Schott dans [26].

Suppléments.

Même s'il y a quelques travaux sur les applications de l'aire de Lévy dans le cadre du calcul stochastique classique (par exemple, dans [8] elle a été mise en rapport avec les processus de Bessel à travers les polynômes de Legendre), son importance apparaît de manière évidente surtout dans les domaines des probabilités non-commutatives (en particulier, à travers une construction du mouvement brownien

non-commutatif, qui permet aussi le passage vers le chemin rugueux du mouvement brownien standard) et celui des approximations numériques.

Les approximations numériques des EDS : Euler-Maruyama vs. Milstein.

Nous présentons ici deux schémas d'approximations numériques pour les EDS dont la comparaison permet de montrer, d'une part, l'utilité de l'aire de Lévy dans ce cadre et, d'autre part, la présence latente, dans le domaine des approximations, de la confrontation entre la représentation simple d'un chemin et sa version rugueuse. Ce dernier point offre une manière d'analyser les avantages et désavantages de chaque représentation au cas par cas. Pour une introduction plus détaillée aux approximations numériques des EDS, on peut se référer, par exemple, au livre [54].

Supposons qu'on veuille approcher l'EDS non-linéaire

$$Y_t = Y_0 + \int_0^t V_0(Y_s) ds + \sum_{i=1}^d \int_0^t V_i(Y_s) dB_s^{(i)}$$

où $V_j : \mathbb{R}^N \rightarrow \mathbb{R}^N$ sont des champs vectoriels lisses, indépendants et non-commutatifs (pour le crochet de Lie), et $(B_t)_t$ est un mouvement brownien à valeurs dans \mathbb{R}^d . L'idée qui est à la base des approximations numériques est la même que pour les chemins rugueux : l'approximation par le développement de Taylor. Le développement de Taylor à l'ordre 1 en version Itô donne *la méthode d'Euler-Maruyama* :

$$\tilde{Y}_{n+1} = \tilde{Y}_n + hV_0(\tilde{Y}_n) + \sum_{i=1}^d V_i(\tilde{Y}_n)(B_{t_{n+1}}^{(i)} - B_{t_n}^{(i)})$$

Si on fait un développement à l'ordre 2, on obtient *la méthode de Milstein* :

$$\tilde{Y}_{n+1} = \tilde{Y}_n + hV_0(\tilde{Y}_n) + \sum_{i=1}^d V_i(\tilde{Y}_n)(B_{t_{n+1}}^{(i)} - B_{t_n}^{(i)}) + \sum_{i,j=1}^d V_{ij}(\tilde{Y}_n) J_{ij}(t_n, t_{n+1})$$

où $J_{ij}(t_n, t_{n+1}) = \int_{t_n}^{t_{n+1}} \int_{t_n}^s dB_s^{(i)} dB_t^{(j)}$, et on voit ainsi apparaître l'aire de Lévy sous la forme $\mathcal{A}_{0,t}^{(ij)} = J_{ij} - \frac{1}{2} J_i J_j$.

Dans quelles situations faut-il utiliser la méthode de Milstein, plus lourde et plus complexe ? Les deux méthodes assurent une convergence faible à l'ordre 1, et, par conséquent, si l'on a juste besoin de calculer les moments, l'information

supplémentaire fournie par l'aire de Lévy serait superflue et la méthode d'Euler-Maruyama doit être privilégiée. D'autre part, la convergence forte pour la première méthode est de l'ordre de $h^{1/2}$, alors que pour la seconde méthode elle est de l'ordre de h , rendant cette dernière plus appropriée lorsqu'on doit considérer la convergence de chemins. Ceci est intimement lié aux questions qui ont inspiré la théorie des chemins rugueux. Dans [25], les auteurs proposent un schéma de type Milstein sans le terme d'aire pour certaines EDS dépendant du mouvement brownien fractionnaire, mais ils utilisent des outils de la théorie des chemins rugueux pour en étudier la convergence.

Groupe de Heisenberg et extension centrale. Nous allons maintenant voir le lien entre l'aire de Lévy et l'algèbre à travers la notion *d'extension centrale* en général et *le groupe de Heisenberg* en particulier.

Si G est un groupe de Lie et H un groupe de Lie abélien, alors **l'extension centrale de G par H** est un groupe de Lie \bar{G} tel que, dans la séquence exacte (i.e. telle que l'image de chaque homomorphisme est le noyau de l'homomorphisme suivant),

$$1 \rightarrow H \rightarrow \bar{G} \rightarrow G \rightarrow 1 \quad (1.24)$$

l'image de H soit dans le centre de \bar{G} .

Dans le cas où $\bar{G} = G \times H$, on peut définir la loi interne de \bar{G} par

$$(g_1, h_1) \star (g_2, h_2) = (g_1 g_2, h_1 h_2 \gamma(g_1, g_2))$$

où la fonction $\gamma : G \times G \rightarrow H$ est lisse et satisfait

$$\gamma(g_1 g_2, g_3) \gamma(g_1, g_2) = \gamma(g_1, g_2 g_3) \gamma(g_2, g_3)$$

la dernière condition étant nécessaire, en particulier, pour assurer l'associativité de la loi. On appelle γ un *2-cocycle du groupe G à valeurs dans H* .

Le groupe d'Heisenberg est un exemple particulier (et peut-être le plus usuel) d'extension centrale.

Définition 1.3.11 (le groupe d'Heisenberg). *On pose $G = \mathbb{R}^2$ et $H = \mathbb{R}$. Alors $\gamma : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$, $\gamma((x, y)(x', y')) = xy'$ définit une extension centrale sur \mathbb{R}^2 . On appelle $G \times H$ le groupe d'Heisenberg de dimension 3.*

On peut aussi donner une définition plus générale : en partant d'un espace vectoriel V et d'une forme symplectique (forme bilinéaire, symétrique, non dégénérée) ω sur V , on définit le groupe d'Heisenberg $H(V) = V \times \mathbb{R}$ comme l'extension centrale de V par \mathbb{R} muni de l'opération

$$(h_1, r_1) \star (h_2, r_2) = (h_1 + h_2, r_1 + r_2 + \frac{1}{2}\omega(h_1, h_2)).$$

Regardons maintenant comment cette construction s'applique à l'étude d'une limite rugueuse avec anomalie d'aire. Pour $V = \mathbb{R}^2$, on peut prendre

$$\omega((x, y), (x', y')) = xy' - yx'.$$

On remarque que, dans ce cas, l'opération $r_1 + r_2 + \frac{1}{2}(xy' - yx')$ peut être identifiée à la concaténation des aires signées de deux courbes γ et γ' , d'accroissements (x, y) et (x', y') respectivement, et d'aires signées r_1 et r_2 respectivement. Ainsi, l'extension centrale offre les moyens d'opérer avec *des aires pures*, i.e. des objets de type $(0, a)$, où $0 \in V$ et $a \in \mathbb{R}$. Si on se réfère au schéma 1.24, ces objets sont contenus dans l'image de $H = \mathbb{R}$, et sont donc dans le centre de $V \times \mathbb{R}$. C'est ce schéma qui est à la base de la décomposition des processus discrets qui permet de dégager les composantes fournissant l'anomalie d'aire à la limite.

Passons maintenant à la partie probabiliste, qui fera apparaître l'aire de Lévy. La définition suivante est donnée pour le groupe de Heisenberg de dimension 3, mais elle peut être facilement généralisée à un groupe de Heisenberg de dimension $2d + 1$.

Définition 1.3.12 (mouvement brownien sur le groupe d'Heisenberg). *Soit $(B_t)_t$ un mouvement brownien bidimensionnel et $(\mathcal{A}_t)_t = (\int_0^t B_s^1 dB_s^2 - B_s^2 dB_s^1)_t$ le processus correspondant à son aire de Lévy. Alors $(B_t, \mathcal{A}_t)_t$ est un mouvement brownien sur le groupe de Heisenberg de dimension 3, dont la fonction caractéristique est donnée*

par

$$\mathbb{E}[\exp(i(\alpha B_t^1 + \beta B_t^2 + i\xi \mathcal{A}_t))] = \frac{1}{\cosh \xi t} \exp\left(\frac{\alpha^2 + \beta^2}{2\xi \coth \xi t}\right)$$

où α, β, ξ sont des réels.

Dans le présent travail, nous obtenons et analysons des processus du type $(B_t, \mathcal{A}_t + at)_t$, où a est un nombre déterministe. Au vu de la définition 1.3.12, cet objet est un mouvement brownien sur le groupe d'Heisenberg (de dimension 3) auquel on ajoute un processus déterministe $(0, 0, at)_t$ à valeur dans le centre du groupe, i.e. dans $\text{Ker}(\gamma) \times H$.

1.3.3 Marches de Markov cachées et chaînes de Markov sur graphes périodiques.

Présentation générale des marches de Markov cachées.

Les modèles de Markov cachés sont des processus discrets utilisés pour la modélisation dans des domaines comme les statistiques ([36], contenant une présentation de ce type de modèles et leur lien avec les réseaux bayésiens), le traitement du signal ([75], contenant également une introduction aux modèles), la biologie ([80]) etc.

Nous avons choisi d'appeler les processus de cette classe *marches de Markov cachées*, car ils peuvent être vus comme une généralisation des marches aléatoires simples. Le processus donné par les incréments d'une marche de Markov cachée sera appelé *chaîne de Markov cachée*.

Définition 1.3.13 (marches et chaînes de Markov cachées sur $T^{(n)}(V)$). *Soit $(R_k)_k$ une chaîne de Markov sur E et, pour $n \geq 1$, soit $(\mathbb{X}_m)_m$ un processus sur $T_1^{(n)}(V)$ tel que, pour tout $k \geq 1$ et pour $u_1, \dots, u_k \in E$,*

- $\mathbb{P}\left(\mathbb{X}_k^{-1} \otimes \mathbb{X}_{k+1} \mid R_k = u_k, \dots, R_0 = u_0\right) = \mathbb{P}\left(\mathbb{X}_k^{-1} \otimes \mathbb{X}_{k+1} \mid R_k = u_k\right)$ et $\mathbb{P}\left(\mathbb{X}_k^{-1} \otimes \mathbb{X}_{k+1} \mid R_k = u_k\right) = \mathbb{P}\left(\mathbb{X}_1^{-1} \otimes \mathbb{X}_2 \mid R_1 = u_k\right)$
- Sous $\mathbb{P}(\cdot \mid \sigma(R))$, les incréments $\mathbb{X}_k^{-1} \otimes \mathbb{X}_{k+1}$ sont indépendants.

On dit alors que $(R_m, \mathbb{X}_m)_m$ est une marche de Markov cachée sur $E \times T^{(n)}(V)$ (ou, s'il n'y a pas de risque de confusion, que $(X_m)_m$ est une marche de Markov cachée sur $T^{(n)}(V)$).

On appelle le processus $(R_m, F_m)_m$, où $F_l = \mathbb{X}_{l-1}^{-1} \otimes \mathbb{X}_l$, chaîne de Markov cachée.

Le schéma de dépendance entre $(R_n)_n$ et $(\mathbb{X}_n)_n$ est donné dans la figure 1.6.

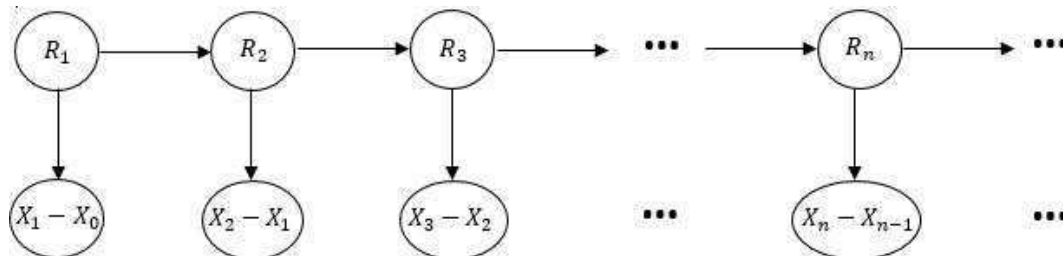


FIGURE 1.6 – Les relations de dépendance pour une marche de Markov cachée $(R_n, X_n)_n$.

Remarque. Dans le cas particulier où $n = 1$ dans la définition 1.3.13, on obtient une marche de Markov cachée $(R_k, X_k)_k$ sur $E \times V$ et les incréments sont donnés par $X_{l+1} - X_l$. Les définitions correspondant aux marches de Markov cachées et chaînes de Markov cachées dans ce cas particulier seront énoncées dans le chapitre 3 (définitions 3.2.2 et 3.2.1 respectivement).

La théorie des pseudo-excursions pour les chaînes de Markov cachées.

Soit $(R_n)_n$ une chaîne de Markov irréductible sur un espace d'états fini E . La théorie des excursions pour les chaînes de Markov s'applique alors à $(R_n)_n$. En particulier, on peut définir la séquence des temps d'arrêt :

$$T_0 = 0$$

$$\forall n \geq 1, T_n = \inf\{k > T_{n-1} : R_k = R_0\}$$

(comme définie dans (3.6) du chapitre 3). Alors $(R_{T_k}, R_{T_k+1}, \dots, R_{T_{k+1}-1})$ est la k -ième excursion de $(R_n)_n$ en R_0 . Mais qu'en est-il alors de la trajectoire de $(X_n)_n$ pendant cette excursion ? La trajectoire $(X_{T_k}, X_{T_k+1}, \dots, X_{T_{k+1}})$ n'est pas une excursion, car on ne revient pas forcément au point de départ. Dans ce cas, il devient intéressant d'étudier les incréments non-triviaux $X_{T_{k+1}} - X_{T_k}$.

Définition 1.3.14 (définition 3.3.1 du 3). Soit $(R_n, X_n)_n$ une chaîne de Markov cachée sur $E \times V$ et soit $(T_k)_k$ une suite de temps d'arrêt comme dans (3.6). On appelle la séquence $(X_{T_{k-1}}, \dots, X_{T_k})$ la k -ième pseudo-excursion de $(X_n)_n$ et la variable $X_{T_k} - X_{T_{k-1}}$ l'incrément de la k -ième pseudo-excursion.

Par analogie avec les excursions, on donne maintenant quelques propriétés élémentaires liées aux pseudo-excursions :

- Les variables $X_{T_{k+1}} - X_{T_k}$ sont i.i.d.
- Pour $i \neq j$, les trajectoires $(X_{T_i}, \dots, X_{T_{i+1}})$ et $(X_{T_j}, \dots, X_{T_{j+1}})$ sachant R_0 sont indépendantes et de même loi.

La proposition suivante vient illustrer l'importance des pseudo-excursions pour la convergence des chaînes de Markov cachées.

Proposition 1 (proposition 3.3.1). Soit $(R_n, X_n)_n$ une chaîne de Markov comme dans la définition 1.3.13 et telle que, pour un certain $M > 0$, $|X_{n+1} - X_n| \leq M$ p.s. et pour tout n . Alors, pour n assez grand, X_n se rapproche en probabilité d'une somme S_n de $\lfloor n\mathbb{E}[X_{T_1}]^{-1} \rfloor$ v.a.i.i.d., au sens où

$$\forall \epsilon > 0, \mathbb{P}(|X_n - S_n| > \epsilon) \rightarrow 0.$$

Cette théorie est présentée dans la section 3.3.1 du chapitre 3.

Chaînes de Markov sur graphes périodiques.

Présentation générale. Les chaînes de Markov sur graphes périodiques ont ceci de particulier qu'elles sont à la fois chaînes de Markov sur un espace d'états infini et chaînes de Markov cachées. Le chapitre 2 sera consacré à l'étude de cette classe de processus indépendamment de la théorie développée pour les chaînes de Markov cachées, car elle présente un intérêt particulier : de nombreux modèles de réseaux ([77], [55] etc.) en sont issus, et l'étude de leur convergence pourrait être approfondie à travers les outils proposés (i.e. en passant en topologie rugueuse pour prendre en compte le drift d'aire). Nous donnerons plus d'exemples de ces modèles dans la section 2.1.4.

La définition de cette classe de processus repose sur la définition du graphe périodique donnée ci-dessous.

Définition 1.3.15 (graphe périodique, propriété 2.1.1). *Soit V un espace vectoriel fini-dimensionnel et G un sous-ensemble infini de V tel que*

$$G = \bigsqcup_{\lambda \in \Lambda} \lambda.G_0$$

où G_0 est un sous-ensemble fini de G , $\Lambda \subset V$ un réseau, et où $\lambda.G_0$ désigne la translation de G_0 par $\lambda \in \Lambda$. On dit alors que $G \subset V$ est un graphe périodique.

Remarque. Le graphe G_0 de la décomposition du graphe G n'est pas forcément unique.

Définition 2 (définition 2.1.2, chapitre 2). *Soit $G \subset V$ comme dans 1.3.15. Soit $(X_n)_{n \in \mathbb{N}}$ une chaîne de Markov à valeurs dans G et de matrice de transition Q . On dit que $(X_n)_{n \in \mathbb{N}}$ est Λ -invariant si et seulement si, pour tous $x, y \in G$ et tous $\lambda \in \Lambda$, $Q(x + \lambda, y + \lambda) = Q(x, y)$.*

La figure 1.7 montre que deux graphes périodiques très différents peuvent avoir le même graphe sous-jacent G_0 . Puisque $G = \bigsqcup_{\lambda \in \Lambda} \lambda.G_0$, chaque point de $x \in G$ peut se décomposer comme $x = (\pi_0(x), \pi_1(x))$, où $\pi_0 : G \rightarrow G_0$ associe à $x \in G$ sa composante dans G_0 et $\pi_1 : G \rightarrow \Lambda$ associe à $x \in G$ sa composante dans Λ . C'est justement cette décomposition qui permet de mettre en lumière une chaîne de Markov sous-jacente à $(X_n)_n$.

Propriété 1 (propriété 2.2.1 du chapitre 2). *Soit $(X_n)_n$ une chaîne de Markov sur un graphe périodique G comme dans la définition 1.3.15. Alors $(\pi_0(X_n))_n$ est une chaîne de Markov sur G_0 .*

Remarque. Contrairement à ce qu'on pourrait croire, $(\pi_0(X_n), X_n)_n$ n'est pas forcément une marche de Markov cachée. Nous allons en parler dans la section 1.3.3.

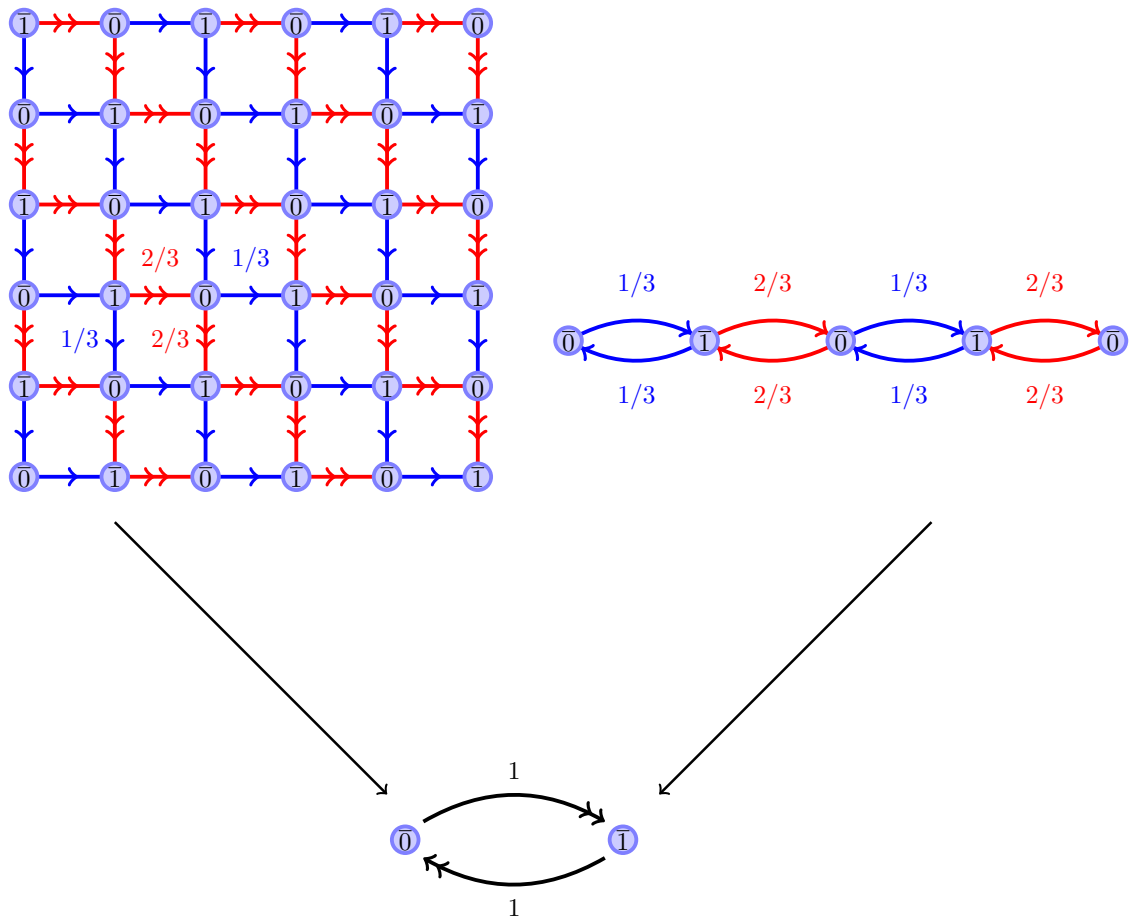


FIGURE 1.7 – Deux chaînes de Markov sur graphes périodiques avec un même sous-graphe donné par une chaîne de Markov déterministe à deux états : un mouvement déterministe sur une droite (à droite) et une chaîne de Markov bidimensionnelle (à gauche) qui passent tous les deux p.s. d'un sommet décoré 1 (resp. 2) à un sommet décoré 2 (resp. 1).

Les pseudo-excursions pour les chaînes de Markov sur graphes périodiques. L'idée est la même que celle que nous avons présentée pour les marches de Markov cachées dans la section 1.3.3. Nous utilisons, cette fois, la chaîne $(\pi_0(X_n))_n$ dont les excursions permettront de définir les pseudo-excursions de la chaîne $(X_n)_n$.

Nous en présenterons une construction détaillée dans la section 2.2.2 du chapitre 2.

Chaîne de Markov sur graphe périodique ou marche de Markov cachée ?

Même si la propriété 2.2.1 (réénoncée plus haut) peut nous porter à croire que les chaînes de Markov sur graphes périodiques sont simplement des cas particuliers des marches de Markov cachées, ce n'est pas tout à fait exact.

Certes, il y a des exemples où $(\pi_0(X_n), X_n)_n$ en est une, comme dans le cas du modèle des sommes tournantes des v.a.i.i.d. de Bernoulli, présenté à la fois comme chaîne de Markov sur graphe périodique et comme marche de Markov cachée dans la section 4.1, mais il y en a aussi où les conditions de la définition 1.3.13 ne sont pas satisfaites, comme c'est le cas du modèle "tourbillon" ("*swirl*" model), présenté au même endroit. Car, si pour tout $n \geq 1$, la loi de $X_{n+1} - X_n$ dépend bien de $\pi_0(X_n)$, on n'a pas de condition relativement à son indépendance des autres $\pi_0(X_k)$.

Ceci dit, on peut construire une marche de Markov cachée à partir d'une chaîne de Markov sur graphe périodique : il suffit de prendre pour la chaîne de Markov sous-jacente $(R_n)_n$ l'ensemble de toutes les valeurs possibles des incréments de $(X_n)_n$ avec les probabilités associées (une telle décomposition a été faite sur un exemple de la section 4.1 du chapitre 4). Dans ce cas, chaque incrément ne dépendra trivialement que de lui-même.

Il y a cependant un intérêt pratique à ne pas regarder une chaîne de Markov sur graphe périodique comme une marche de Markov cachée : vu les contraintes plus faibles sur la dépendance des incréments de $(X_n)_n$ de la chaîne sous-jacente dans le cas de la première, l'espace d'états correspondant à $(\pi_0(X_n))_n$ est souvent plus petit que celui de la chaîne $(R_n)_n$, ce qui facilite les calculs, en particulier ceux de l'anomalie d'aire (comme nous allons le voir sur l'exemple du modèle "tourbillon" dans la section 4.1 du chapitre 4).

<i>CMGP</i>	<i>MMC</i>
<ul style="list-style-type: none"> - est une chaîne de Markov - est définie indépendamment de sa chaîne de Markov sous-jacente - peut donner naissance à une marche de Markov cachée - est liée à une représentation géométrique - a des contraintes plus faibles sur la dépendance de ses incréments de la chaîne sous-jacente - permet de simplifier les calculs, en particulier ceux liés à l'anomalie d'aire 	<ul style="list-style-type: none"> - a des propriétés markoviennes induites par la chaîne sous-jacente - est définie à partir d'une chaîne de Markov sous-jacente - l'inverse n'est pas vrai - permet de représenter des processus plus abstraits - demande l'indépendance des incréments sachant les états correspondants de la chaîne sous-jacente - rend les calculs plus complexes

TABLE 1.2 – Comparaison de chaînes de Markov sur graphes périodiques (CMGP) et marches de Markov cachées (MMC).

Une synthèse de ce qui vient d'être dit est présentée dans le tableau 1.2.

1.3.4 Les plongements

Lorsqu'on veut étudier la limite d'un processus discret $(X_n)_n$, on doit commencer par le plonger dans l'espace des processus continus, ce qui veut dire : choisir un plongement, i.e. la manière de relier entre eux les points du processus.

Dans \mathbb{R}^d , il y a un choix immédiat et naturel : *le plongement géodésique* (illustré dans 1.8). Par plongement géodésique, nous entendons ici un plongement *d'interpolation géodésique* : pour une suite de points dans un espace géodésique, ce plongement consiste à joindre les points consécutifs par des courbes (continues) qui minimisent les distances entre elles. En particulier, pour une suite de points dans \mathbb{R}^d , le plongement géodésique est l'interpolation linéaire. En effet, si on se borne à étudier la limite de $(X_n)_n$ en topologie uniforme, tout autre plongement, pourvu qu'il soit à variation finie, *ne change pas la nature de la limite*. Plus précisément, en topologie uniforme, de deux choses l'une : soit le plongement ne change pas la

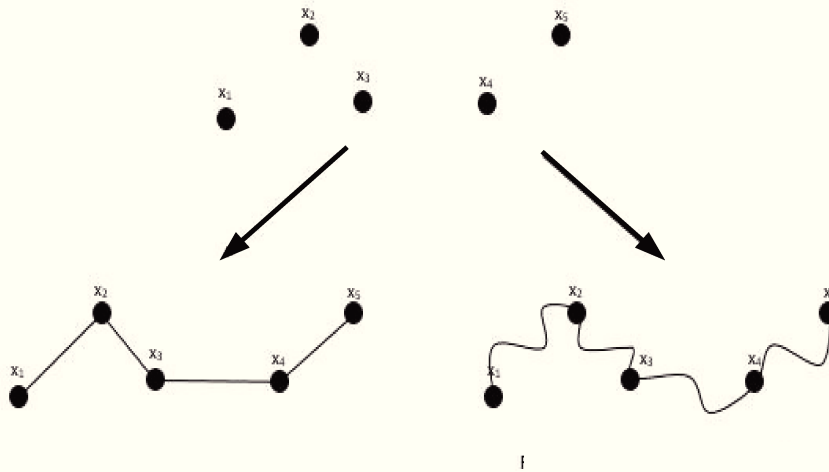


FIGURE 1.8 – Deux plongements pour un nuage de points : plongement linéaire par morceaux (à gauche) et plongement construit à partir de différentes courbes dans \mathbb{R}^2 (à droite).

nature de la limite par rapport à celle de l'interpolation linéaire, soit il ne converge pas.

Définition et classes d'équivalence.

Nous donnons une première définition des plongements directement à partir du processus discret.

Définition 3 (définition 3.2.3 du chapitre 3). *Un plongement d'une suite $(x_n)_n \in \mathcal{S}$, où \mathcal{S} est un espace métrique, est une suite de chemins continus dans \mathcal{S} $\rho = (\rho_N)_N$ tels que*

- $\forall N \geq 1, \rho_N : [0, N] \rightarrow \mathcal{S}$;
- $\forall N \geq 1, \forall i \in \{1, \dots, N\}, \rho_N(i) = x_i$;
- $\forall k < N, \forall t \in [0, k], \rho_k(t) = \rho_N(t)$ (le chemin ρ_N est identique au chemin ρ_{N-1} jusqu'au moment $N - 1$).

Une deuxième définition, équivalente à la première, définit une manière de relier deux points consécutifs de la suite par un chemin continu (plongement élémentaire) et construit ensuite le plongement en concaténant ces courbes.

Définition 4 (définition 3.3.4 du chapitre 3). *Soit $(x_n)_n$ une suite à valeurs dans un espace métrique \mathcal{S} et $(\beta_n)_n$ une suite de chemins continus $\beta_n : [0, 1] \rightarrow \mathcal{S}$ tels que $\beta_n(0) = x_n$ et $\beta_n(1) = x_{n+1}$. On définit la suite $(\rho_n)_n$ par*

$$\rho_N = \beta_1 \cdot \dots \cdot \beta_N$$

où \cdot est l'opérateur de concaténation de chemins. Alors on dit que $(\rho_n)_n$ est un plongement pour $(x_n)_n$. En particulier, le plongement géodésique est le plongement qui minimise les distances entre les points $(x_n)_n$.

Une fois dans le cadre continu, pour pouvoir étudier la limite du plongement, il faut trouver la bonne renormalisation de l'espace-temps.

Définition 1.3.16. *Soit \mathcal{S} un espace métrique qui admet une famille d'opérateur de renormalisation $(T_\alpha)_{\alpha>0}$. Pour $\alpha, \beta > 0$ et $F : \mathbb{R}_+ \rightarrow W$, on pose*

$$T_{\alpha,\beta}(F)(t) = T_\alpha F(\beta t)$$

Soit $(\rho_N)_N$ un plongement de $(X_n)_n$ à valeurs dans \mathcal{S} . Pour $f : \mathbb{N} \rightarrow \mathbb{R}_+$, la suite

$$\iota_N(t) = T_{f(N),N} \rho_N = T_{f(N)} \rho_N(Nt)$$

définit un plongement renormalisé de $(X_n)_n$.

On revient maintenant au cas où l'espace métrique est un espace vectoriel V de dimension finie. Comme on l'a déjà dit, en topologie uniforme, tous les plongements renormalisés d'un processus discret $(X_n)_n$ qui convergent ont, modulo une constante, la même limite. *Mais ce n'est plus le cas en topologie rugueuse.* Ainsi, on peut distinguer de manière plus fine les limites de différents plongements, en les groupant par classes d'équivalence au sens rugueux :

Définition 5 (définition 3.3.6 du chapitre 3). *Soit $(x_n)_n$ une suite à valeurs dans l'espace vectoriel V . On dit que deux plongements de $(x_n)_n$ au sens de la définition*

3.3.4, ρ_N et ρ'_N , sont équivalents au sens rugueux ssi leurs plongements renormalisés respectifs ι_N et ι'_N (au sens de la définition 1.3.16) ont la même limite dans la topologie rugueuse.

On va maintenant fournir un représentant aux classes d'équivalence ainsi formées. Vu la définition 3.3.3 de $G^2(V)$, lorsqu'il est muni de la loi antisymétrique \wedge , un élément $(y, a) \in G^2(V)$ (où $y \in V$) peut être interprété comme une information sur un chemin $x \in \mathcal{C}^{1-var}([s, t], V)$ tel que :

$$x_t - x_s = y, \quad A_{s,t}(x) = a \quad (1.25)$$

À partir de là, on peut définir une équivalence : un chemin $x' : [s, t] \rightarrow V$ est équivalent à x si et seulement si il satisfait (1.25) :

$$x' \sim x \iff x' \text{ satisfait (1.25).}$$

Ce qu'identifie cette équivalence, ce sont l'accroissement entre s et t et l'accumulation d'aire, *et non pas* la trajectoire spécifique de chaque chemin. L'élément (y, a) devient alors le représentant de tous ces chemins, ce qui est illustré dans la figure 1.9. Passons maintenant aux plongements. Dans le chapitre 3 (section 3.2.1), nous verrons que la limite d'un plongement dans la topologie rugueuse est déterminée à la fois par ses incréments et par l'aire signée propre à ce plongement dans l'interprétation antisymétrique de $G^2(V)$ (ou par la double intégrale dans son interprétation symétrique). On peut donc décrire une classe d'équivalence au sens rugueux des plongements de $(x_n)_n$ par une suite de point dans $G^2(V)$.

Propriété 1.3.8. *Soit $(x_n)_n$ une suite à valeurs dans V . On se place dans le cadre de la définition 3.3.6. Une classe d'équivalence au sens rugueux peut être représentée par une suite $(y_n, a_n)_n \in G^2(V)$.*

Démonstration. On choisit un plongement de $\tilde{\rho}_N = \beta_1 \cdot \dots \cdot \beta_N$ de la suite $(x_n)_n$ comme représentant d'une classe d'équivalence aux sens rugueux. On pose :

$$\begin{aligned} y_n &= x_n - x_{n-1} \\ a_n &= A_{0,1}(\beta_n) \end{aligned}$$

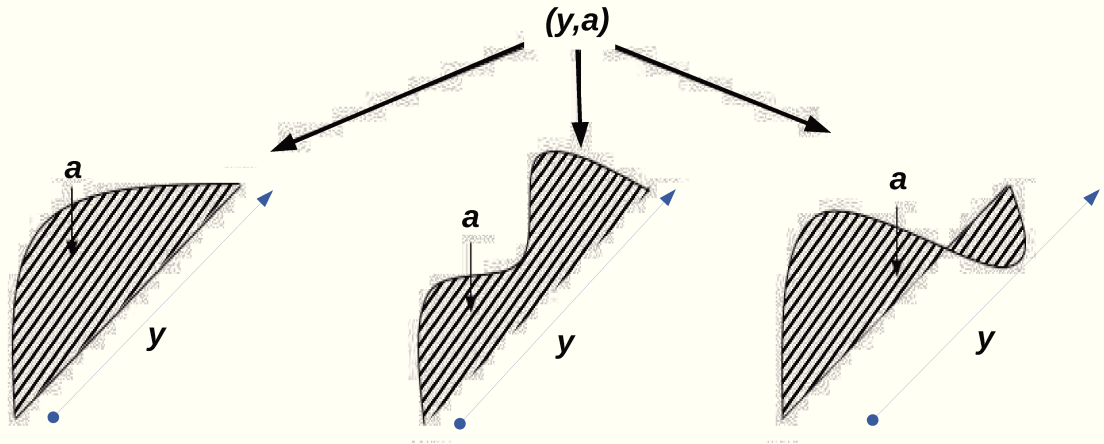


FIGURE 1.9 – Trois chemins qui peuvent être représentés par le même $(y, a) \in G^2(\mathbb{R}^2)$.

où $A_{s,t}(\gamma)$ est l'aire signée de γ entre s et t . On pose ensuite

$$\tilde{\rho}_N = \bigotimes_{k=1}^N (y_k, a_k)$$

et on reconstruit ainsi le chemin rugueux correspondant à $\tilde{\rho}_N$. Donc $\tilde{\rho}_N$ détermine aussi bien $\tilde{\rho}_N$ que sa convergence dans la topologie rugueuse. Par conséquent, il détermine entièrement la classe d'équivalence de $\tilde{\rho}_N$. \square

L'universalité du plongement géodésique dans $G^2(V)$.

Le résultat 1.3.8 a comme conséquence le fait que la convergence de tout plongement particulier d'une séquence $(x_n)_n \in V^{\mathbb{N}}$ peut se ramener à la convergence du plongement géodésique d'une suite dans $G^2(V)$.

Proposition 1.3.2. *Soit $(x_n)_n$ une suite à valeurs dans V . On se place dans le*

cadre de la définition 3.3.6. Soit $(y_n, a_n)_n \in G^2(V)^{\mathbb{N}}$ correspondant à une classe d'équivalence au sens rugueux comme expliqué dans 1.3.8. Alors il suffit de calculer la limite (dans $G^2(V)$) du plongement géodésique renormalisé de $(y_n, a_n)_n$ pour connaître la limite rugueuse de tous les plongements (assez réguliers, par exemple $(\rho_N)_N$ tel que les ρ_N soient à variation bornée et uniformément bornés sur N dans la topologie correspondante) contenus dans la classe d'équivalence respective.

Remarque. La proposition du chapitre 3 (3.3.3) est une conséquence qui découle de cette propriété : la convergence d'un plongement d'une chaîne de Markov cachée vers le mouvement brownien rugueux avec une anomalie d'aire découle de la convergence vers cette limite du plongement géodésique d'une suite bien choisie de $G^2(V)$. C'est en ce sens que les théorèmes 3.2.1 et 3.2.2 du chapitre 3 sont des corollaires du théorème 3.2.3 du même chapitre.

Plongement pour une chaîne de Markov cachée.

Dans les sections précédentes, les résultats ont été présentés dans un cadre déterministe. Nous allons maintenant faire la transition vers le cadre probabiliste pour les marches de Markov cachées.

- Pour une chaîne de Markov $(R_n, F_n)_n$ sur $E \times V$, pour tout $u \in E$, soit $V_u \subset V$ l'ensemble de toutes les réalisations possibles de F_1 sous la loi $\mathbb{P}(\bullet | R_1 = u)$.
- Pour $u \in E$, à chaque $y \in V_u$ on associe une courbe $f^{y;u} : [0, 1] \rightarrow V$ à variation finie et telle que $f^{y;u}(0) = 0$ et $f^{y;u}(1) = y$ (correspondant à l'incrément y), et on note \mathcal{B}_u l'ensemble de ces courbes.
- Pour tout $k \geq 1$, on associe à F_k une v.a. β_k dans le sens où, pour un $u \in E$ et sous la loi $\mathbb{P}(\bullet | R_k = u)$, β_k donne, pour chaque réalisation $y \in V_u$ de F_k l'élément correspondant $f^{y;u} \in \mathcal{B}_u$.

Pour tout $k \geq 1$, la loi de β_k est décrite par la loi de F_k :

$$\forall u \in E, \forall y \in V_u, \mathbb{P}(X_1 = y | R_1 = u) = \mathbb{P}(\beta_1 = f^{y;u} | R_1 = u).$$

1.3.5 Les équations aux différences

Les équations aux différences sont l'analogie discret des équations différentielles (ED). Mais alors que, pour un processus réel d -dimensionnel $(x_t)_t$ l'équation

$$dy_t = \sum_{i=1}^d f^{(i)}(y_t) dx_t^{(i)} \quad (1.26)$$

n'est pas toujours bien définie (et c'est pour pallier ce problème en particulier qu'on passe en topologie rugueuse), on peut toujours définir, pour une suite déterministe $(x_n)_n$ à valeurs dans \mathbb{R}^d , une suite $(y_n)_n$ par la récurrence :

$$y_{n+1} - y_n = \sum_{i=1}^d f^{(i)}(y_n) (x_{n+1} - x_n)^{(i)}. \quad (1.27)$$

Les équations aux différences (comme (1.27)) peuvent être étudiées dans deux directions : soit on cherche à les résoudre directement, i.e. à trouver une expression explicite pour $(Y_n)_n$ à partir de la récurrence ; soit on considère l'équation renormalisée

$$y_{n+1} - y_n = \epsilon \sum_{i=1}^d f^{(i)}(y_n) (x_{n+1} - x_n)^{(i)} \quad (1.28)$$

et on étudie sa convergence lorsque $\epsilon \rightarrow 0$. Cette convergence s'étudie, le plus souvent, au cas par cas, par les méthodes d'approximations numériques.

1.3.6 Autres résultats préliminaires et sources d'inspiration

Nous introduisons ici trois théorèmes qui reviennent régulièrement dans nos preuves, tout en expliquant leur importance pour nos preuves et résultats.

Le premier théorème est issu de l'article [12] et montre la convergence d'une somme de v.a.i.i.d. centrées, en topologie rugueuse, vers le mouvement brownien rugueux.

Théorème 1.3.1 (théorème du Donsker en version chemins rugueux, [12]). *Soit*

$(\xi_n)_n$ une suite de v.a.i.i.d. centrées dans \mathbb{R}^d . On pose, pour $t \in [0, 1]$,

$$W_t^n = \sum_{k=1}^{\lfloor nt \rfloor} \xi_k + (nt - \lfloor nt \rfloor) \xi_{\lfloor nt \rfloor + 1}.$$

Soit $(\mathbb{W}_t^n)_{t \in [0,1]}$ le chemin rugueux correspondant à $(W_t^n)_{t \in [0,1]}$ (alternativement, on peut dire que c'est la signature de $(W_t^n)_{t \in [0,1]}$ tronquée au niveau 2, i.e. $S_2(W_t^n) \in T^{(2)}(V)$).

Alors, pour tout $\alpha < 1/2$, on a la convergence suivante en loi dans la topologie $\mathcal{C}^{0,\alpha\text{-Hölder}}([0, 1], G^2(\mathbb{R}^d))$:

$$(\mathbb{W}_t^n)_{t \in [0,1]} \xrightarrow[n \rightarrow \infty]{} (\mathbb{B}_t)_{t \in [0,1]}$$

où \mathbb{B} est le mouvement brownien rugueux dans \mathbb{R}^d .

Anomalie d'aire pour les modèles issus des marches de Markov cachées.

Le théorème 1.3.1 nous a à la fois inspiré le choix de la classe des marches de Markov cachées pour une construction explicite de l'anomalie d'aire et aidé à montrer nos résultats.

La théorie des pseudo-excursions permet de mettre en relief un découpage en variables i.i.d. d'une marche de Markov cachée, et permet donc d'appliquer le théorème 1.3.1. Mais, dans ce cas, on ne prend pas en compte les fluctuations trajectorielles lors d'une pseudo-excursion. Ce sont ces fluctuations ou, plus précisément, l'aire stochastique générée par elles qui se perd à la limite en topologie uniforme et qui génère l'anomalie d'aire en topologie rugueuse.

Pour intégrer l'anomalie d'aire à la limite de la somme des pseudo-excursions (pour la convergence des marginales fini-dimensionnelles, et même celle des processus), nous utilisons la version générale du théorème de Slutsky (par exemple, théorème 3.1 dans [9]).

Théorème 1.3.2 (théorème de Slutsky, [9]). *Soit (\mathcal{S}, d) un espace métrique et $(X_n)_n, (Y_n)_n$ deux suites de v.a. à valeurs dans \mathcal{S} telles que*

$$d(X_n, Y_n) \xrightarrow[n \rightarrow \infty]{} 0.$$

S'il existe une v.a. X à valeurs dans S telle que

$$d(X_n, X) \xrightarrow[n \rightarrow \infty]{} 0$$

alors $d(Y_n, X) \xrightarrow[n \rightarrow \infty]{} 0$.

Enfin, pour montrer la tension de la suite une fois l'aire stochastique intégrée aux chemins rugueux des pseudo-excursions (comme dans la preuve du théorème 2.1.1 du chapitre 2), nous utilisons une version du critère de Kolmogorov (adapté de [9] ou de [12]).

Théorème 1.3.3 (critère de Kolmogorov). *Soit V un espace vectoriel de dimension finie. On considère une suite de chemins rugueux $(\mathbb{X}_t^n)_{t \in [0,1]}$ dans $G^2(V)$ telle que*

$$\forall p > 1, \exists c_p > 0, \forall s, t \in [0, 1], \sup_n \mathbb{E} \left[\mathbf{d}(\mathbb{X}_t^n, \mathbb{X}_s^n)^{4p} \right] \leq c_p |t - s|^{2p-1}.$$

Alors la suite $(\mathbb{X}_t^n)_{t \in [0,1]}$ est tendue dans $\mathcal{C}^{\alpha-Höl}([0, 1], G^2(V))$ pour tout $\alpha < 1/2$.

1.3.7 Présentation des principaux résultats

Convergence des chaînes de Markov sur graphes périodiques.

Tout d'abord, commençons par le théorème sur la convergence des chaînes de Markov sur graphes périodiques que nous avons démontré dans [64] et dont la preuve peut être reliée à une intuition géométrique.

Théorème 1.3.4 (théorème 2.1.1 du chapitre 2). *Soit G un graphe Λ -périodique sur un espace vectoriel de dimension finie V . Soit $(X_n)_n$ une chaîne de Markov Λ -invariante, à valeurs dans G , telle que*

$$\exists R > 0, \forall n \in \mathbb{N}, |X_{n+1} - X_n|_V \leq R \text{ p.s.}$$

et telle que $(\pi_0(X_n))_{n \in \mathbb{N}}$ soit irréductible. On définit la suite des temps d'arrêt

$$\begin{aligned} T_0 &= 0, \\ T_{k+1} &= \inf \{n > T_k : \pi_0(X_n) = \pi_0(X_0)\}, \quad k \geq 0. \end{aligned}$$

Soit $v = \mathbb{E}[T_1]^{-1} \mathbb{E}[X_{T_1}] \in V$ et $\beta = \mathbb{E}[T_1] \in \mathbb{R}_+^*$. À une réduction de dimension et une transformation linéaire de G près, on peut supposer que la matrice de covariances de $X_{T_1} - T_1 v$ est CI_n , avec $C > 0$.

Soit $(\tilde{X}_n)_{n \in \mathbb{N}}$ le processus défini par $\tilde{X}_n = X_n - nv$. Pour tout $\tau > 0$, on a la convergence en loi

$$\left(\delta_{\sqrt{C^{-1}\beta}} \iota^{(N)}(\tilde{X}_\bullet, A_\bullet(\tilde{X}))_t \right)_{0 \leq t \leq \tau} \xrightarrow{N \rightarrow \infty} (B_t, \mathcal{A}_t + t\Gamma)_{0 \leq t \leq \tau} \quad (1.29)$$

dans la topologie $\mathcal{C}^{\alpha\text{-Hölder}}([0, \tau], G^2(V))$ pour tout $\alpha < 1/2$, avec B un mouvement Brownien sur V , \mathcal{A} son aire de Lévy et Γ une matrice déterministe asymétrique, l'anomalie d'aire, dont une formule explicite est donnée en (2.19).

Éléments de preuve. L'idée de base est la suivante : il faut séparer le processus discret rugueux en deux parties, dont l'une nous donnera la convergence vers le mouvement brownien rugueux et l'autre va générer l'anomalie d'aire.

Pour ce faire, on s'appuie sur la théorie des pseudo-excursions. Les variables $\tilde{X}_{T_{k+1}} - \tilde{X}_{T_k}$ sont i.i.d. et centrées (car $\tilde{X}_{T_{k+1}} - \tilde{X}_{T_k} = X_{T_{k+1}} - X_{T_k} - \mathbb{E}[T_1]$ par construction).

Par conséquent, si on pose $\tilde{\lambda}(X)_n = \sum_{k=1}^n \tilde{X}_{T_k} - \tilde{X}_{T_{k-1}}$, le chemin rugueux correspondant au plongement géodésique de $(\tilde{\lambda}(X)_n)_n$ converge vers le mouvement brownien rugueux $(\delta_{\sqrt{C}}(B_t, \mathcal{A}_t))_t$ par le théorème de type Donsker pour les chemins rugueux rappelé dans le théorème 1.3.1.

L'autre partie du processus $(\tilde{X}_k)_k$ est composée des trajectoires

$$(\tilde{X}_{T_k}, \tilde{X}_{T_{k+1}}, \dots, \tilde{X}_{T_{k+1}-1}), \quad k \geq 0.$$

Justement, en topologie uniforme, ces fluctuations ne se traduisent, à la limite, que par l'apparition d'une constante multiplicative, $\mathbb{E}[T_1]$, qui correspond à la longueur moyenne d'une excursion. En particulier, pour tout $t > 0$, $(n^{-1/2} \tilde{X}_{[tT_n]})_t$ converge, en topologie uniforme, vers $(\sqrt{C\beta^{-1}}B_t)_t$.

En topologie rugueuse, ces fluctuations permettent une accumulation d'aire au niveau 2 du chemin rugueux, ce qui donne l'anomalie d'aire Γ . Plus précisément, jusqu'au temps T_n , la contribution à l'aire stochastique par les fluctuations de la trajectoire est donnée par la somme des aires stochastiques $A_{0,T_1}(\tilde{X}) + A_{T_1,T_2}(\tilde{X}) +$

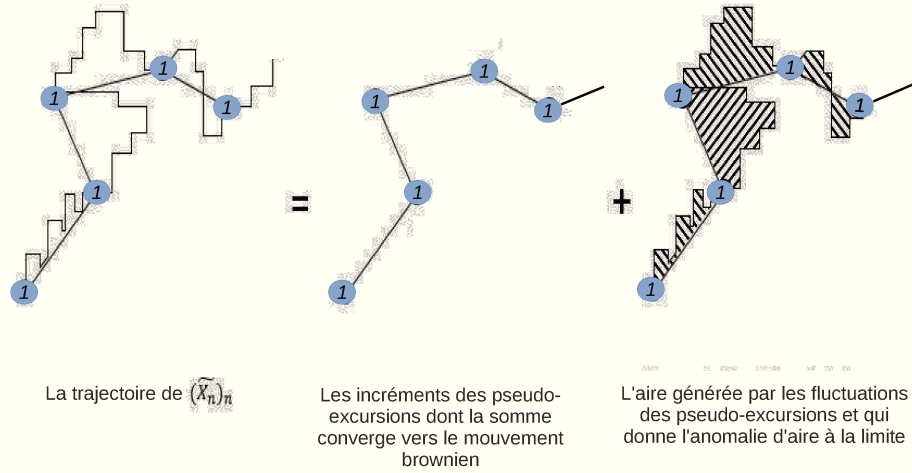


FIGURE 1.10 – La décomposition des pseudo-excursions d’une chaîne de Markov sur graphe périodique en incréments (v.a.i.i.d. dont la somme converge vers le mouvement brownien) et fluctuations (qui génèrent l’anomalie d’aire).

$\dots + A_{T_{n-1}, T_n}(\tilde{X})$. Ceci est schématiquement illustré dans la figure 1.10.

Par la loi des grands nombres, on récupère la convergence p.s.

$$\frac{A_{0, T_1}(\tilde{X}) + A_{T_1, T_2}(\tilde{X}) + \dots + A_{T_{n-1}, T_n}(\tilde{X})}{n} \xrightarrow[n \rightarrow \infty]{} \mathbb{E} [A_{T_1}(\tilde{X})].$$

Ainsi, par le théorème de Slutsky 1.3.2, pour tout $t > 0$, $\delta_{n^{-1/2}}(\tilde{X}_{\lfloor tT_n \rfloor}, \mathcal{A}_{\lfloor tT_n \rfloor}(\tilde{X}))$ converge vers $\delta_{\sqrt{C\beta^{-1}}}(B_t, \mathcal{A}_t + t\Gamma)$, où $\Gamma = C^{-1}\mathbb{E} [A_{T_1}(\tilde{X})]$.

Pour passer de $\tilde{X}_{\lfloor tT_k \rfloor}$ à $\tilde{X}_{\lfloor tn \rfloor}$, on décompose $\tilde{X}_{\lfloor tn \rfloor} = \tilde{X}_{\lfloor tT_{\kappa(n)} \rfloor} + \tilde{X}_{\lfloor tn \rfloor} - \tilde{X}_{\lfloor tT_{\kappa(n)} \rfloor}$, avec $\kappa(n)$ tel que $T_{\kappa(n)} \leq n < T_{\kappa(n)+1}$, on montre que $n^{-1/2}(\tilde{X}_{\lfloor tn \rfloor} - \tilde{X}_{\lfloor tT_{\kappa(n)} \rfloor})$ converge vers 0 en topologie rugueuse, puis on applique une fois de plus le théorème de Slutsky.

Enfin, une fois qu’on a récupéré la convergence des marginales fini-dimensionnelles, on montre la tension de la suite en topologie rugueuse, en utilisant une version du critère de Kolmogorov pour les géodésiques qu’on a rappelé dans le théorème 1.3.3. \square

Convergence des chaînes de Markov cachées sur $G^2(V)$ et deux corollaires.

Le théorème suivant généralise le théorème 2.1.1 dans deux directions : d'une part, on considère une classe bien plus large de processus, les chaînes de Markov cachées ; d'autre part, on généralise l'espace dans lequel vivent ces processus, en passant d'un espace vectoriel de dimension finie V au groupe de Lie nilpotent $G^2(V)$.

Théorème 1.3.5 (version du théorème 3.2.3 du chapitre 3). *Soit $(R_n)_n$ une chaîne de Markov irréductible sur E (espace d'états fini) et $(R_n, \mathbb{X}_n)_n$ une marche de Markov cachée sur $E \times G^2(V)$ dont les incréments sont p.s. bornés, i.e.*

$$\exists K > 0, \forall k \geq 1, \left\| \mathbb{X}_k^{-1} \otimes \mathbb{X}_{k+1} \right\| \leq K.$$

On suppose que $\mathbb{E} \left[\mathbb{X}_{T_1}^{(1)} \right] = 0$, où T_1 est comme dans (3.6). On pose $\beta = \mathbb{E}[T_1]$. On choisit un plongement de $(\mathbb{X}_k^{(1)})_k$ $\rho = (\rho_N)_N$ qui peut être encodé par la suite $(\mathbb{X}_n)_n$ au sens de la proposition 1.3.2.

Alors, modulo une réduction à un sous-espace, il existe $C > 0$ tel que la dilation $\delta_{(n, \beta^{-1}C)^{-1/2}}$ du plongement linéaire de $(\mathbb{X}_k)_k$ converge dans $\mathcal{C}^\alpha([0, 1], G^2(V))$ pour $\alpha < 1/2$ vers le mouvement brownien rugueux avec un drift limite au niveau 2, i.e. vers

$$\left(\mathbb{B}_t^{Strat} \otimes (0, \Gamma_\rho t) \right)_{t \in [0, 1]}$$

où \mathbb{B}^{Strat} est le mouvement brownien standard augmenté de ses intégrales doubles au sens de Stratonovich et Γ_ρ une matrice déterministe antisymétrique dépendant du plongement ρ (qui est dans la classe d'équivalence représentée par $(\mathbb{X}_n)_n$ au sens de la propriété 1.3.8) et donnée par la formule (3.12) :

$$\Gamma_\rho = C^{-1} \mathbb{E} \left[\int_{0 < s < t < T_1} \rho_{T_1}(s) \otimes \rho_{T_1}(t) - \rho_{T_1}(t) \otimes \rho_{T_1}(s) \right]$$

Alternativement, on peut décomposer Γ_ρ comme $\Gamma_\rho = \Gamma + \Gamma_0$, où Γ est l'aire générée

par les incréments du chemin $(\mathbb{X}_0^{(1)}, \dots, \mathbb{X}_{T_1}^{(1)})$, i.e. l'anomalie d'aire

$$\Gamma = C^{-1} \mathbb{E} \left[\text{Antisym} \left(\sum_{1 \leq k < m \leq T_1} \Delta \mathbb{X}_k^{(1)} \otimes \Delta \mathbb{X}_m^{(1)} \right) \right] \quad (1.30)$$

et Γ_0 est l'aire totale générée par les fluctuations de ρ entre X_{T_k} et $X_{T_{k+1}}$ pour $k = 0, \dots, T_1 - 1$, i.e.

$$\Gamma_0 = C^{-1} \mathbb{E} \left[\text{Antisym} \left(\sum_{1 \leq k \leq T_1 - 1} \mathbb{X}_k^{(2)} \right) \right] \quad (1.31)$$

Éléments de preuve. On décompose le plongement géodésique de X_1, \dots, X_n en trois parties :

- une partie contenant les pseudo-excursions jusqu'au moment $T_{\kappa(n)}$, qu'on appellera \mathcal{P}_t^n ,
- une partie contenant l'« aire pure », c'est-à-dire toute l'aire du niveau 2 qui n'est pas générée par les pseudo-excursions (y compris l'aire générée par les fluctuations comme dans la figure 1.10), qu'on appellera \mathcal{A}_t^n ,
- une partie résiduelle qui contient le chemin rugueux entre $T_{\kappa(n)}$ et n , qu'on appellera \mathcal{R}_t^n .

Après quoi, en utilisant le théorème 1.3.1 et quelques lemmes techniques, nous montrons que

$$\begin{aligned} (\delta_{(n\beta^{-1}C)^{-1/2} \mathcal{P}_t^n})_{t \in [0,1]} &\xrightarrow[n \rightarrow \infty]{} (\mathbb{B}_t)_{t \in [0,1]} \text{ en loi} \\ (\delta_{(n\beta^{-1}C)^{-1/2} \mathcal{A}_t^n})_{t \in [0,1]} &\xrightarrow[n \rightarrow \infty]{} ((0, \Gamma + \Gamma_0)t)_{t \in [0,1]} \text{ en probabilité} \\ (\delta_{(n\beta^{-1}C)^{-1/2} \mathcal{R}_t^n})_{t \in [0,1]} &\xrightarrow[n \rightarrow \infty]{} 0 \text{ en probabilité} \end{aligned}$$

et nous pouvons conclure ensuite en utilisant le théorème de Slutsky 1.3.2. \square

On peut déduire de ce résultat les théorèmes 3.2.1 et 3.2.2 présentés au chapitre 3.

Corollaire 1 (théorème 3.2.1). *Soit $(R_n, X_n)_n$ une chaîne de Markov cachée sur $E \times V$ telle que*

$$\exists K > 0, \forall k \geq 1, |X_{k+1} - X_k| \leq K$$

(comme définie dans (3.8) du chapitre 3). On suppose que $\mathbb{E}[X_{T_1}] = 0$, où T_1 est comme dans (3.6). On pose $\beta = \mathbb{E}[T_1]$.

Alors, modulo une réduction de dimension, il existe $C > 0$ tel que la dilatation $\delta_{(n\beta-1C)^{-1/2}}$ du chemin rugueux correspondant au plongement géodésique de $(X_k)_k$ converge dans $\mathcal{C}^\alpha([0, 1], G^2(V))$ pour $\alpha < 1/2$ vers le mouvement brownien rugueux avec un drift limite au niveau 2, i.e. vers

$$\left(\mathbb{B}_t^{Strat} \otimes (0, \Gamma t)\right)_{t \in [0, 1]}$$

où \mathbb{B}^{Strat} est le mouvement brownien standard augmenté de ses intégrales doubles au sens de Stratonovich, Γ est l'anomalie d'aire de la formule (1.30), i.e.

$$\Gamma = C^{-1} \mathbb{E} \left[\sum_{1 \leq k < m \leq T_1} \Delta X_k \otimes \Delta X_m - \Delta X_m \otimes \Delta X_k \right] \quad (1.32)$$

Ce résultat est la généralisation du théorème 2.1.1, au sens où les chaînes de Markov sur graphes périodiques peuvent être interprétées comme des marches de Markov cachées.

Enfin, le deuxième corollaire montre comment le théorème 3.2.3 permet de traiter les cas où on veut choisir un plongement autre que Donsker pour la chaîne de Markov cachée, par la propriété 1.3.2 sur la représentation d'un plongement dans V par le plongement géodésique d'une suite dans $G^2(V)$.

Corollaire 2 (théorème 3.2.2 du chapitre 3). *Soit $(R_n, X_n)_n$ une chaîne de Markov cachée sur $E \times V$ telle que la condition (3.8) soit satisfaite. On suppose que $\mathbb{E}[X_{T_1}] = 0$, où T_1 est comme dans (3.6). On pose $\beta = \mathbb{E}[T_1]$.*

On choisit un plongement $\rho = (\rho_N)_N$ pour $(X_k)_k$ tel que décrit dans 3.3.4 et tel que ρ soit à variation finie, i.e. les ρ_N soient uniformément bornées en N dans $\mathcal{C}^{1-var}([0, 1], V)$.

Alors, modulo une réduction de dimension, il existe $C > 0$ telle que la dilatation $\delta_{(N\beta-1C)^{-1/2}}$ du chemin rugueux correspondant au plongement ρ de $(X_k)_k$ converge dans $\mathcal{C}^\alpha([0, 1], G^2(V))$ pour $\alpha < 1/2$ vers le mouvement brownien rugueux avec un

drift limite au niveau 2, i.e. vers

$$\left(\mathbb{B}_t^{Strat} \otimes (0, \Gamma_\rho t) \right)_{t \in [0,1]}$$

où \mathbb{B}^{Strat} est le mouvement brownien standard augmenté de ses intégrales doubles au sens de Stratonovich et Γ_ρ une matrice déterministe antisymétrique dépendant du plongement ρ par la formule (3.12) :

$$\Gamma_\rho = C^{-1} \mathbb{E} \left[\int_{0 < s < t < T_1} \rho_{T_1}(s) \otimes \rho_{T_1}(t) - \rho_{T_1}(t) \otimes \rho_{T_1}(s) \right]$$

1.3.8 Applications

Le calcul d'anomalie d'aire.

Les méthodes de calcul. Soit $(R_n, X_n)_n$ une chaîne de Markov cachée sur $E \times V$ comme dans la définition 1.3.13. On suppose que X a des excursions centrées (i.e. $\mathbb{E}[X_{T_1}] = 0$, où T_1 est défini dans (3.6)). Nous donnons ici une présentation générale de deux méthodes de calcul : l'une plutôt géométrique, l'autre plutôt algébrique.

La méthode géométrique : l'algorithme inspiré des « loop-erased random walks ». Pour calculer l'aire stochastique $A_{T_1}(X)$, l'idée de base est de décomposer le cycle $(R_0, R_1, \dots, R_{T_1})$ en cycles plus petits et qui ne contiennent pas R_0 , en faisant en parallèle la décomposition de $(X_0, X_1, \dots, X_{T_1})$; on recommence ensuite avec chacun des nouveaux cycles et ainsi de suite, jusqu'au moment où on arrive à une décomposition qui permet des calculs faciles. Pour beaucoup de chaînes de Markov sur graphes périodiques issues de mécanique statistique, cette méthode est assez facile à appliquer grâce aux symétries des modèles.

Cet algorithme, ainsi que les formules de calculs de l'anomalie d'aire qui en sont issues, seront présentés au chapitre 4, section 4.2.1.

La méthode algébrique, à partir de la matrice stochastique de $(R_n)_n$. La loi de $(R_n, X_n)_n$ dépend, d'une part, de la matrice stochastique Q de $(R_n)_n$ et, d'autre part, de la valeur $\mathbb{E}[X_{n+1} - X_n | R_n = u]$, où $u \in E$. Cette méthode donne l'anomalie d'aire en fonction de ces deux paramètres. Elle sera présentée dans la

section 4.2.2 du chapitre 4.

Étude de modèles. Dans le chapitre 4, section 4.1, nous proposons aussi une étude comparative de deux modèles : *les sommes tournantes de v.a.i.i.d. de Bernoulli* et le *"swirl" model*, le modèle "tourbillon". Nous présentons d'abord les deux modèles en tant que chaînes de Markov sur graphes périodiques et en tant que chaînes de Markov cachées (représentations qui sont identiques pour le premier, différentes pour le second) et nous calculons ensuite explicitement l'anomalie d'aire pour chacun des deux, en illustrant le découpage en excursions.

Application aux équations aux différences.

L'application aux équations aux différences permet de voir comment le drift limite du niveau 2 des chemins rugueux peut apparaître dans le drift limite d'une suite d'équations aux différences de la forme (1.27). En particulier, on a le résultat suivant, qui est aussi une généralisation de la formule de la section 2.4.1 du chapitre 2.

Proposition 1.3.3. *Soit $(R_n, X_n)_n$ une chaîne de Markov sur $E \times V$, avec un plongement continu qui satisfait aux conditions du théorème 3.2.2. On considère la suite d'équations aux différences :*

$$\Delta Y_n = N^{-1/2} \sum_{i=1}^d f^{(i)}(Y_{n-1})(\Delta X_n)^{(i)}$$

où $\Delta X_n = X_n - X_{n-1}$ et f est une fonction dérivable à valeurs dans V . Alors cette suite d'équations converge dans la topologie rugueuse vers l'EDS limite

$$\begin{aligned} dY_t &= \sum_{i=1}^d f^{(i)}(Y_t) dB_t^{(i)} + \frac{1}{2} K \sum_{i=1}^d f'^{(i)}(Y_t) f^{(i)}(Y_t) dt \\ &+ \frac{1}{2} \sum_{1 \leq i < j \leq d} \Gamma^{(ij)} (f'^{(i)}(Y_t) f^{(j)}(Y_t) - f^{(i)}(Y_t) f'^{(j)}(Y_t)) dt \end{aligned}$$

où Γ est l'anomalie d'aire et K dépend de la partie symétrique du drift du second niveau.

1.4 Convergence de sommes itérées pour des modèles markoviens et étude de la renormalisation

1.4.1 Sommes et intégrales itérées

Définitions. Les intégrales itérées sont des objets bien mieux connus que les sommes itérées, car elles ont des applications notoires, par exemple dans la théorie du chaos gaussien ou dans le domaine des probabilités non-commutatives ([7]), ou encore comme moyen d'encoder des variétés ([18]). De plus, et ce n'est pas sans rapport avec les applications déjà citées, elles sont les briques de base (abstraites ou pas) pour la construction des chemins rugueux.

Donnons-en une brève définition tout de même.

Définition 1.4.1. Soit $(x_t)_t$ une courbe à variation finie dans \mathbb{R}^d . L'intégrale itérée d'ordre k par rapport à $(x_t)_t$ est définie par la formule (3.2)

$$S_{x,k}(t) = \int_{0 < s_1 < \dots < s_k < t} dx_{s_1} \otimes \dots \otimes dx_{s_k}$$

Les composantes de $S_{x,k}(t)$ seront, pour $i_1, \dots, i_k \in \{1, \dots, d\}$, données par

$$\begin{aligned} S_{x,k;i_1,\dots,i_k}(t) &= \left\langle e_{i_1} \otimes \dots \otimes e_{i_k}, \int_{0 < s_1 < \dots < s_k < t} dx_{s_1} \otimes \dots \otimes dx_{s_k} \right\rangle \\ &= \int_{0 < s_1 < \dots < s_k < t} dx_{s_1}^{(i_1)} \dots dx_{s_k}^{(i_k)} \end{aligned} \quad (1.33)$$

où les e_i sont des vecteurs de la base canonique.

Les sommes itérées peuvent être vues comme des intégrales itérées sur des chemins linéaires par morceaux, ou alors comme la version discrète des intégrales itérées.

Définition 1.4.2. Soit $(y_n)_n$ une suite à valeurs dans \mathbb{R}^d . Pour $(y_n)_n$, la somme itérée d'ordre k est définie comme

$$\tilde{S}_{y,k}(n) = \sum_{1 \leq l_1 < \dots < l_k \leq n} (\Delta y)_{l_1} \otimes \dots \otimes (\Delta y)_{l_k} \quad (1.34)$$

où $(\Delta y)_k = y_k - y_{k-1}$.

Les composantes de $\tilde{S}_{y,k;i_1,\dots,i_k}(n)$ seront, pour $i_1, \dots, i_k \in \{1, \dots, d\}$,

$$\begin{aligned} \tilde{S}_{y,k;i_1,\dots,i_k}(n) &= \left\langle e_{i_1} \otimes \dots \otimes e_{i_k}, \sum_{1 \leq l_1 < \dots < l_k \leq n} (\Delta y)_{l_1} \otimes \dots \otimes (\Delta y)_{l_k} \right\rangle \\ &= \sum_{1 \leq l_1 < \dots < l_k \leq n} (\Delta y)_{l_1}^{(i_1)} \dots (\Delta y)_{l_k}^{(i_k)} \end{aligned} \quad (1.35)$$

où les e_i sont des vecteurs de la base canonique.

Lien avec les chemins rugueux. Nous avons déjà vu que les intégrales itérées permettent de remplir les niveaux des chemins rugueux pour les courbes intégrables (intégrale d'ordre k pour le niveau k). L'intérêt des sommes itérées se révèle lorsqu'on veut étudier la convergence d'un processus discret en topologie rugueuse.

On a alors la propriété suivante, qui est à la base de nos constructions dans le théorème 2.1.1 du chapitre 2 et le théorème 3.2.1 du chapitre 3.

Propriété 1.4.1. *Soit $(x_n)_n$ une suite dans un espace vectoriel fini-dimensionnel V . Modulo un facteur de continuité, les sommes itérées de $(x_n)_n$ fournissent le plongement géodésique pour le chemin rugueux correspondant à $(x_n)_n$.*

1.4.2 Temps d'occupation itérés et théorie ergodique

Les temps d'occupation itérés sont aux sommes itérées d'un processus ce que le temps d'occupation classique est au processus lui-même (ou presque). Mais notre but n'est pas seulement de généraliser la notion du temps d'occupation, mais, à travers cette généralisation, de mettre en évidence le lien entre chemins rugueux et combinatoire pour les marches de Markov cachées. On verra également de quelle manière les chemins rugueux permettent de généraliser la théorie ergodique.

Théorie ergodique pour les chaînes de Markov. On rappelle ici brièvement des résultats et des notions qu'on va généraliser par la suite.

Définition 1.4.3 (temps d'occupation). *Soit $R = (R_n)_n$ une chaîne de Markov sur un espace d'états E . On appelle temps d'occupation par R d'un état $u \in E$*

jusqu'au temps n la variable

$$L_{u;n}(R) = \sum_{k=1}^n \mathbf{1}_{R_k=u} \quad (1.36)$$

Théorème 1.4.1 (théorème ergodique). *Soit $(R_n)_n$ une chaîne de Markov irréductible récurrente positive sur un espace d'états E . Alors, pour tout $u \in E$, on a la convergence p.s.*

$$\frac{L_{u;n}(R)}{n} = \frac{\sum_{k=1}^n \mathbf{1}_{R_k=u}}{n} \xrightarrow[n \rightarrow \infty]{} \pi(u) \quad (1.37)$$

où π est la mesure de probabilité invariante de $(R_n)_n$.

Plus généralement, pour toute fonction $f : E \rightarrow \mathbb{R}_+$ intégrable par rapport à π , on a la convergence p.s.

$$\frac{\sum_{k=1}^n f(R_k)}{n} \xrightarrow[n \rightarrow \infty]{} \sum_{u \in E} f(u)\pi(u) = \langle f | \pi \rangle \quad (1.38)$$

Dans le cas d'une chaîne de Markov irréductible récurrente positive, on a également un théorème central limite (TCL) (inspiré de [52]).

Théorème 1.4.2 (TCL pour les chaînes de Markov). *Soit $(R_n)_n$ une chaîne de Markov irréductible récurrente positive sur un espace d'états E et soit π sa mesure de probabilité invariante. On pose*

$$T_1 = \inf\{n \geq 1 : R_n = R_0\}$$

Alors, pour tout pour toute fonction $f : E \rightarrow \mathbb{R}_+$ intégrable par rapport à π , on a la convergence en loi

$$\frac{\sum_{k=1}^n f(R_k) - n \langle f | \pi \rangle}{\sqrt{nC_f}} \xrightarrow[n \rightarrow \infty]{} \mathcal{N}(0, 1) \quad (1.39)$$

où $C_f > 0$ est une constante qui dépend de f .

En particulier, on a, pour tout $u \in E$, la convergence en loi

$$\frac{\sum_{k=1}^n \mathbf{1}_{R_k=u} - n\pi(u)}{\sigma\sqrt{n}} \xrightarrow[n \rightarrow \infty]{} \mathcal{N}(0, 1) \quad (1.40)$$

où $\sigma > 0$.

Définition des temps d'occupation itérés. Avant d'en donner une version probabiliste, on va formuler une définition déterministe et combinatoire.

Définition 1.4.4. Soit E un ensemble et $r = (r_n)_n \in E^{\mathbb{N}}$. Pour tout k -uplet $(u_1, \dots, u_k) \in E^k$, le temps d'occupation itéré de (u_1, \dots, u_k) jusqu'au temps n est donné par :

$$L_{u_1, \dots, u_k; n}(r) = \sum_{1 \leq n_1 < \dots < n_k \leq n} \mathbf{1}_{r_{n_1} = u_1} \dots \mathbf{1}_{r_{n_k} = u_k} \quad (1.41)$$

Alternativement, on peut utiliser la formule (3.21) du chapitre 3, qui met davantage l'accent sur la combinatoire :

$$L_{u_1, \dots, u_k; n}(R) = \text{card} \{(n_1, \dots, n_k) \in \Delta_k(n); R_{n_1} = u_1, \dots, R_{n_k} = u_k\}$$

où $\Delta_k(n)$ est donné par

$$\Delta_k(n) = \{(n_1, \dots, n_k) \in \mathbb{N}^k; 1 \leq n_1 < n_2 < \dots < n_k \leq n\}$$

Un théorème ergodique généralisé peut être formulé pour les temps d'occupation itérés, comme nous le verrons dans le chapitre 3 (proposition 3.2.2). Les temps locaux itérés apparaissent également dans une forme de TCL en version rugueuse (dédié de la 3.2.3 du chapitre 3).

Temps d'occupation et sommes itérés. Les temps d'occupation itérés se révèlent utiles pour étudier les sommes itérées d'une marche de Markov cachée $(R_n, X_n)_n$ sur $E \times V$. Dans ce cas, on a la décomposition de la formule (3.24) du chapitre 3

$$\mathbb{E}[S_{X,k}(N) | \sigma(R)] = \sum_{u_1, \dots, u_k \in E} f(u_1) \otimes \dots \otimes f(u_k) L_{u_1, \dots, u_k; N}(R)$$

où $f(u) = \mathbb{E}_\nu[\Delta X_1 | u]$.

On voit ainsi que, dans $S_{X,k}(N)$, tout ce qui dépend de N et de $(R_n)_n$ est concentré dans $L_{u_1, \dots, u_k; N}(R)$. Ceci est particulièrement intéressant pour l'analyse de la convergence en N des marches de Markov cachées.

Remarque. Lorsque E est fini et $V = \mathbb{R}$, il est facile de déduire que

$$C^{-1/2} \sum_{u \in E} f(u) \left(\frac{\sum_{k=1}^n \mathbf{1}_{R_n=u} - n\pi(u)}{\sqrt{n}} \right) \xrightarrow{n \rightarrow \infty} \mathcal{N}(0, 1) \quad (1.42)$$

et cette convergence est fonction de celle du vecteur $(L_{u;n} - n\pi(u))_{u \in E}$. Il serait donc intéressant d'étudier la convergence de ce vecteur en topologie rugueuse, y compris pour son lien avec les marches de Markov cachées donné par la formule (3.24) du chapitre 3, ce dont nous allons parler dans la section 1.4.4 de l'introduction.

1.4.3 Produits *shuffle* et *quasi-shuffle*

Les produits *quasi-shuffle* ont été introduits par M. Hoffman dans son article [45]. Ils représentent une généralisation des produits *shuffle* et sont liés aux algèbres de Hopf.

Définitions. Les définitions qu'on donnera ici sont proches de celles du [47], même si on laisse de côté certains détails algébriques.

Définition 1.4.5 (produit *quasi-shuffle*). Soit k un corps contenu dans \mathbb{C} , A un ensemble dénombrable de « lettres » et kA l'espace vectoriel dont A est la base. On suppose que kA est muni d'un produit « \circ » qui est associatif et commutatif. Soit

$$\langle A \rangle = \text{Vect}\{a_1 \dots a_n : a_i \in A\} \quad (1.43)$$

l'algèbre (non-commutative) de polynômes sur A ($\langle A \rangle$ est l'espace vectoriel (sur k) engendré par tous les « mots » qu'on peut former à partir des « lettres » de A). On définit alors le produit *quasi-shuffle* « $*$ » sur $\langle A \rangle$ de manière récursive par

$$\forall a, b \in A, \forall u, v \in \langle A \rangle, au * bv = a(u * bv) + b(au * v) + (a \circ b)(u * v) \quad (1.44)$$

et en demandant que le mot vide qu'on note 1 soit l'élément neutre pour ce produit.

Exemple : Pour $a, b, c \in A$, $ab * c = abc + cab + acb + (a \circ c)b + a(b \circ c)$.

On a le résultat suivant (prouvé dans [47]).

Propriété 1.4.2. $(\langle A \rangle, *)$ est une algèbre commutative.

Le produit *shuffle*, qui a été défini avant le produit *quasi-shuffle*, ne contient plus le terme dépendant de « \circ ».

Définition 1.4.6 (produit *shuffle*). Soit k un corps contenu dans \mathbb{C} , A un ensemble dénombrable de "lettres". Soit

$$\langle A \rangle = \text{Vect}\{a_1 \dots a_n : a_i \in A\} \quad (1.45)$$

l'algèbre (non-commutative) de polynômes sur A ($\langle A \rangle$ est l'espace vectoriel (sur k) engendré par tous les « mots » qu'on peut former à partir des « lettres » de A). On définit alors le produit *shuffle* « \sqcup » sur $\langle A \rangle$ de manière récursive par

$$\forall a, b \in A, \forall u, v \in \langle A \rangle, au \sqcup bv = a(u \sqcup bv) + b(au \sqcup v) \quad (1.46)$$

et en demandant que le mot vide qu'on note 1 soit l'élément neutre pour ce produit.

Exemple : Pour $a, b, c \in A$, $ab \sqcup c = abc + cab + acb$.

La propriété 1.4.2 reste vraie pour « \sqcup » : $(\langle A \rangle, \sqcup)$ est une algèbre commutative.

De plus, en définissant sur $\langle A \rangle$ le coproduit

$$\forall w \in \langle A \rangle, \Delta(w) = \sum_{uv=w} u \otimes v \quad (1.47)$$

on a le résultat suivant ([47], avec des légères modifications) qui nous permet de faire le lien avec les chemins rugueux.

Théorème 1.4.3. $(\langle A \rangle, *, \Delta)$ et $(\langle A \rangle, \sqcup, \Delta)$ sont des algèbres de Hopf.

Remarque. Les produits \sqcup et $*$ sont des opérateurs de concaténation, alors que Δ peut être interprété comme opérateur de déconcaténation. Intuitivement, ils fournissent donc l'essentiel pour un cadre d'algèbre de Hopf : non seulement on

sait composer deux éléments de l'algèbre, mais on peut aussi retrouver toutes les manières dont un élément de l'algèbre peut être obtenu.

Application aux sommes et aux intégrales itérées. Nous allons maintenant brièvement expliquer comment ces deux produits s'appliquent aux sommes et intégrales itérées. Plus précisément, on va montrer que le produit « \sqcup » donne la loi du produit des intégrales itérées par rapport à une courbe $(x_t)_t$ à valeurs dans \mathbb{R}^d , alors que le produit « $*$ » décrit le produit des sommes itérées pour une suite $(y_n)_n \in (\mathbb{R}^d)^\mathbb{N}$.

Dans le cas des intégrales itérées, d'ordres k et l respectivement, d'un chemin x dans V , le produit de leur projections sur $e^{i_1} \otimes \dots \otimes e^{i_k}$ et $e^{j_1} \otimes \dots \otimes e^{j_l}$ respectivement (formule (1.33)) donne la projection de l'intégrale itérées d'ordre $k+l$ sur $e^{p_1} \otimes \dots \otimes e^{p_{k+l}}$, où $p_1 \dots p_{k+l} = i_1 \dots i_k \sqcup j_1 \dots j_l$.

Lorsqu'on multiplie les projections sur $e^{i_1} \otimes \dots \otimes e^{i_k}$ et $e^{j_1} \otimes \dots \otimes e^{j_l}$ de deux sommes itérées correspondant à une suite dans V (formule (1.35)), on retrouve la projection sur $e^{p_1} \otimes \dots \otimes e^{p_{k+l}}$ de la somme itérée d'ordre $k+l$, mais aussi des termes supplémentaires qui viennent du fait que les diagonales, de mesure de Lebesgue nulle et donc négligeables dans le cas des intégrales, sont prises en compte sous la mesure de comptage. Ces termes supplémentaires déterminent le caractère *quasi-shuffle* du produit.

Tout ceci sera expliqué plus longuement au chapitre 3, dans la section 3.4.1.

1.4.4 Présentation des principaux résultats

Le passage entre chemins rugueux géométriques et chemins rugueux non-géométriques.

Soit $(R_k, F_k)_k$ une chaîne de Markov cachée sur $E \times V$ (comme définie dans 1.3.13) et $(R_n, X_n)_n$ la marche de Markov cachée correspondante (i.e. pour $k \geq 1$, $F_k = X_k - X_{k-1}$). On sait qu'à partir d'un plongement suffisamment régulier de $(X_n)_n$, on peut construire un chemin rugueux géométrique dans $G^2(V)$ de manière canonique.

Soit $(X_t^{(N)})_t$ un plongement à variation finie de $(X_n)_n$. Le chemin rugueux

canonique dans $G^2(V)$ qui lui correspond est la signature de niveau 2

$$\forall t \in [0, 1], S_2(X^{(N)})_{0,t} = \left(X_t^{(N)}, \int_{0 < s_1 < s_2 < t} dX_{s_1}^{(N)} \otimes dX_{s_2}^{(N)} \right)$$

qui est un élément de $\mathcal{C}^{1-var}([0, 1], G^2(V))$.

On peut ensuite étudier la convergence en loi de $(\delta_{N^{-1/2}} S_2(X^{(N)})_{0,t})_t$ dans la topologie $\mathcal{C}^{\alpha-H\ddot{o}l}([0, 1], G^2(V))$ pour $\alpha < 1/2$. La limite (si elle existe) est *un chemin rugueux géométrique α -Hölder* pour $\alpha < 1/2$.

Nous allons montrer comment on aboutit à un chemin rugueux non-géométrique à partir d'une marche de Markov cachée $(R_n, X_n)_n$ en utilisant les sommes itérées. Nous commençons par construire la suite $(\mathbb{Y}_n)_n$ dans $T^{(2)}(V)$

$$\mathbb{Y}_n = \left(X_n, \sum_{1 \leq k_1 < k_2 \leq n} F_{k_1} \otimes F_{k_2} \right) \in T^{(2)}(V)$$

et c'est ensuite qu'on lui associe un plongement $(\mathbb{Y}_t^{(N)})_{t \in [0,1]}$ dont on étudiera la convergence.

Il est facile de remarquer que $(\mathbb{Y}_n)_n$ n'est pas dans $G^2(V)$, car le niveau 2 n'est pas égal à $\frac{1}{2} X_n \otimes X_n$:

$$\begin{aligned} \text{Sym} \left(\sum_{1 \leq k_1 < k_2 \leq n} F_{k_1} \otimes F_{k_2} \right) &= \frac{1}{2} \left(\sum_{1 \leq k_1 < k_2 \leq n} F_{k_1} \otimes F_{k_2} + \sum_{1 \leq k_1 < k_2 \leq n} F_{k_2} \otimes F_{k_1} \right) \\ &= \frac{1}{2} X_n \otimes X_n - \frac{1}{2} \sum_{k=1}^n F_k^{\otimes 2} \end{aligned}$$

Le résultat suivant montre qu'on peut construire un plongement de $(\mathbb{Y}_n)_n$ qui converge vers un chemin rugueux non-géométrique :

Proposition 1.4.1 (proposition 3.2.1 du chapitre 3). *Soit $(X_t^{(N)})_t$ le plongement par interpolation linéaire de $(X_n)_n$. On construit aussi le plongement par interpolation linéaire*

$$\forall N \in \mathbb{N}^*, \forall t \in [0, 1], Z_t^{(N)} = \sum_{k=1}^{\lfloor Nt \rfloor} F_k^{\otimes 2} + (Nt - \lfloor Nt \rfloor) F_{\lfloor Nt \rfloor + 1}^{\otimes 2}$$

de $Z_k = \sum_{j=1}^k F_j^{\otimes 2}$ et on pose

$$\forall N \in \mathbb{N}^*, \forall t \in [0, 1], \mathbb{Y}_t^{(N)} = (0, Z_t^{(N)})^{-1} \otimes S_2(X^{(N)})_{0,t}$$

On suppose de plus que $\mathbb{E}[X_{T_1}] = 0$, où T_1 est comme dans (3.6), et que la matrice de covariance $\mathbb{V}(X_{T_1})$ vaut CI_d pour un certain $C > 0$. On pose $\beta = \mathbb{E}[T_1]$.

Alors on a la convergence suivante dans la topologie $\mathcal{C}^{\alpha-H\ddot{o}l}([0, 1], T^{(2)}(V))$ avec $\alpha < 1/2$:

$$\left(\delta_{(nC\beta^{-1})^{-1/2}} \mathbb{Y}_t^{(N)} \right)_{t \in [0,1]} \xrightarrow{N \rightarrow \infty} \left(\mathbb{B}_t^{It\hat{o}} \otimes (0, Mt) \right)_{t \in [0,1]}$$

où $(\mathbb{B}_t^{It\hat{o}})_t$ est le mouvement brownien rugueux au sens d'Itô et M est une matrice déterministe donnée par

$$M = C^{-1} \mathbb{E} \left[\sum_{1 \leq k_1 < k_2 \leq T_1} F_{k_1} \otimes F_{k_2} \right]$$

(formule (3.18) du chapitre 3). En particulier, $(\mathbb{B}_t^{It\hat{o}} \otimes (0, Mt))_t$ est un chemin rugueux non-géométrique.

Les constructions des limites géométrique et non-géométrique peuvent être résumées à travers le schéma 1.11. La relation qui relie la limite géométrique et la limite non-géométrique est la suivante :

$$\mathbb{B}_t^{It\hat{o}} \otimes (0, Mt) = \mathbb{B}_t^{Strat} \otimes (0, \Gamma t) \otimes \left(0, tC^{-1} \mathbb{E} \left[\sum_{k=1}^{T_1} F_k^{\otimes 2} \right] \right)^{-1} \quad (1.48)$$

Convergence de la structure combinatoire des marches de Markov cachées.

On pose $L_n(r) = (L_{u;n}(r))_{u \in E}$, qui sera un vecteur à valeurs entières. On a alors le résultat suivant, qui met en relief une classe de convergence « universelle » pour les chaînes de Markov cachées :

Proposition 2 (proposition 3.2.3 du chapitre 3). *Soit $(R_n)_n$ une chaîne de Markov*

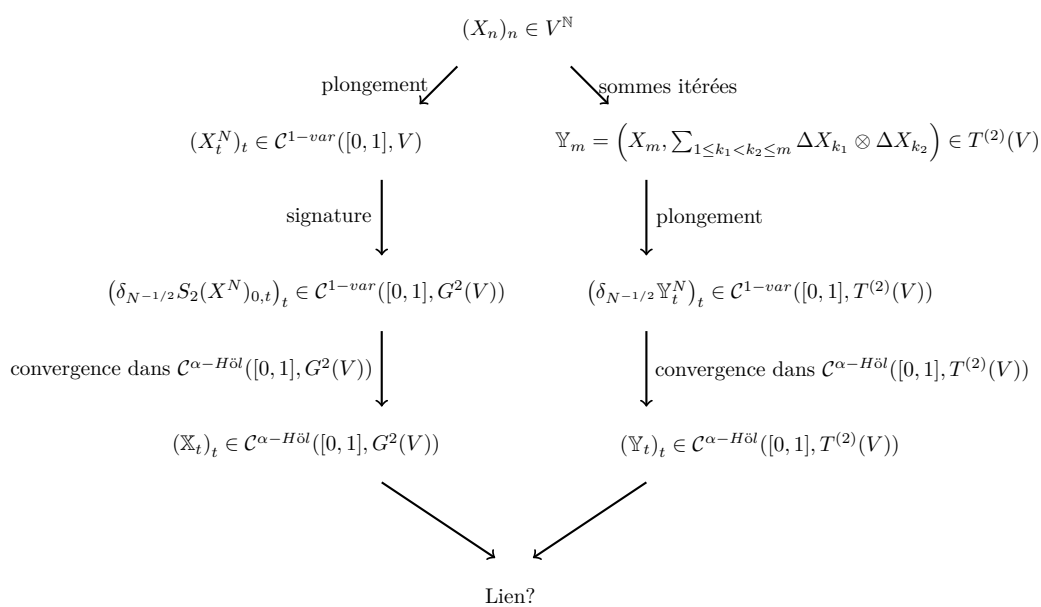


FIGURE 1.11 – Chemins rugueux géométrique (en bas à gauche) et non-géométrique (en bas à droite) comme limites de processus continus issus du même processus discret $(X_n)_n$.

irréductible sur un espace d'états fini E . Pour $u \in E$ et pour $n \geq 1$, on pose

$$\tilde{L}_n(R) = (L_{u;n}(R) - n\pi(u))_{u \in E} \in \mathbb{R}^{\text{Card}(E)}$$

où π est la mesure invariante de $(R_n)_n$.

Alors, modulo une réduction de dimension, le plongement linéaire de la proposition 3.4.3 de $(\tilde{L}_n(R))_n$ et les intégrales itérées associées $S_{\tilde{L}_n(R),k}(t)$ pour $k \geq 2$ converge dans la topologie $C^{\alpha-Hö}([0, 1], T_1^{(2)}(V))$ pour $\alpha < 1/2$ vers

$$\begin{aligned} S_{B,1}(t) &= B_t, \\ S_{B,2}(t) &= \int_{0 < s_1 < s_2 < t} dB_{s_1} \otimes dB_{s_2} + Mt \end{aligned}$$

où $(B_t)_t$ est un mouvement brownien $\text{Card}(E)$ -dimensionnel, M est une matrice déterministe donnée par (3.18).

1.5 Le mot de la fin

Faisons une synthèse de tout ce que nous avons présenté au fil des sections, paragraphes, définitions et théorèmes : qu'est-ce que nous avons finalement obtenu ? Si nous devons le résumer en quelques mots, nous le ferions presque sous la forme d'un slogan de campagne : "du discret dans le rugueux, du rugueux dans le discret".

Nous avons voulu montrer que l'introduction des structures discrètes dans le cadre des chemins rugueux, d'une part, et l'application des outils liés aux chemins rugueux, en particulier à travers l'anomalie d'aire, aux modèles discrets étudiés dans le cadre du calcul stochastique, d'autre part, permettaient un nouveau regard sur les deux domaines et ouvraient des perspectives intéressantes, le tout d'une manière suffisamment accessible pour le plus grand nombre.

Mais nous avons probablement déjà assez parlé de ce que nous avons fait dans notre journal de bord et, pour ne pas ennuyer le lecteur, parlons plutôt de ce que nous n'avons pas (encore) fait.

Pour illustrer notre propos, nous n'avons étudié que des modèles, certes inspirés de la mécanique statistique, mais plus simples que ceux qu'on rencontre en réalité. Il faudrait maintenant passer à ces derniers, et voir toutes les implications que le

changement de regard sur leur convergence et l'éventuelle présence d'une anomalie d'aire pourraient avoir en pratique. La même question se pose pour des modèles plus complexes issus du calcul stochastique classique comme, par exemple, les marches aléatoires en milieux aléatoires (on citera ici une fois de plus [5]).

Nous avons choisi une classe de modèles assez large, mais qui permet, lors de l'étude de la convergence, l'utilisation d'une méthode bien particulière, et qui nous a bien facilité la tâche : le découpage en (pseudo-)excursions. La prochaine étape serait d'abandonner cette facilité et de chercher de nouveaux moyens pour étudier la convergence en topologie rugueuse de processus moins sympathiques (comme les processus faiblement mélangeants ou les modèles mentionnés au paragraphe précédent).

Pour les exemples que nous avons donnés, nous avons également privilégié la simplicité et la possibilité de donner une interprétation géométrique du modèle pour faciliter la compréhension des objets (en particulier, en illustrant la formation d'anomalie d'aire). Sans nécessairement sortir du cadre des marches de Markov cachées, nous pourrions maintenant construire des exemples en prenant pour chaînes de Markov sous-jacentes des processus plus complexes (ASEP, urnes de Polya, etc.). Les conditions sur les moments des variables, bien qu'elles correspondent à celles qu'on a souvent en pratique, pourraient être encore affaiblies.

Du côté combinatoire, il serait intéressant de pousser plus loin l'étude des temps d'occupation itérés. En dehors de la généralisation des théorèmes liés au temps d'occupation classique, on peut se demander si l'on peut rigoureusement attribuer à ces objets combinatoires, dans le cadre de la convergence de processus discrets, un rôle similaire à celui qu'ont les niveaux d'un chemin rugueux dans le continu. Accessoirement, nous étudions la question de savoir s'il est possible de déduire la convergence d'une marche de Markov cachée de celle de ses temps d'occupation itérés.

Mais il faut s'arrêter avant que la liste « à faire » ne dépasse la liste « déjà fait ». Nous espérons que le présent travail aura contribué à ouvrir quelques nouvelles directions de recherche et à montrer que les chemins rugueux étaient une terre accueillante et pleine de ressources même pour ceux qui n'y sont que de passage.

Chapter 2

Lévy area with a drift as a renormalization limit of Markov chains on periodic graphs

Ce chapitre est notre article [64] écrit en collaboration avec D. Simon tel qu'il a été accepté pour publication dans la revue Stochastic Processes and their Applications (SPA) et mis en ligne sur arXiv (<https://arxiv.org/abs/1604.08947>), excepté des changements mineurs comme l'absence d'abstract ou le changement de numérotation pour les références.

2.1 Introduction

2.1.1 General motivations

Many papers deal with the convergence of discrete vector-valued processes to Brownian motion in the spirit of Donsker's theorem, i.e. using uniform convergence. This is useful for finite dimensional marginals but not for the study of differential equations: uniform topology is too weak to ensure proper approximation of integrals driven by a process of low regularity. Terry Lyons solved this problem by creating the *rough path theory*, which Martin Hairer generalized to *regularity structures*. In

both cases, when working in dimensions higher than 2, the processes are lifted to a more complex structure whose topology ensures the continuity of the solution map of SDEs (the *Itô map*).

Let us have a quick look at the theory of rough paths. The main idea is to build out of the initial processes involved more elaborate structures that allow to register all the relevant information. Loosely speaking, a (continuous) process in \mathbb{R}^d (with $d \geq 2$) is considered as a first-level information, and we build the corresponding rough path by adding a few more levels. The number of necessary levels is determined by the regularity of the process (if $(X_t)_{t \geq 0}$ is of regularity $\alpha \in (0, 1)$, we need $\lfloor 1/\alpha \rfloor$ levels) and each level is given by an iterated integral of a tensorial product (a double integral for level two, a triple one for level three and so on). In particular, for a process $X = (X^1, \dots, X^d)$, the second level is determined by the increments of the process (the first level) and the *stochastic area*: a process on the space of $d \times d$ antisymmetric matrices given by

$$A_{s,t}(X) = \left(\iint_{s < u < v < t} dX_u^i dX_v^j - dX_u^j dX_v^i \right)_{1 \leq i, j \leq d} \quad (2.1)$$

with the convention $A_{0,t}(X) = A_t(X)$. A detailed introduction to rough paths can be found in [58] and [66], and more exhaustive treatments in [34] or [32].

Up to now, rough path theory has been either applied to processes of lower regularity than the Brownian motion (for example, fractional Brownian motion) or to the Brownian motion itself, which resulted in a kind of *rewriting of classical stochastic calculus*. Surprisingly – and it is one of the main motivations for this paper – it is actually possible to go beyond the stochastic calculus as we know it. The second level of the Brownian rough path is made of a symmetric and an antisymmetric part. Whereas we have the choice between Itô and Stratonovich integration for the symmetric part, the antisymmetric one is given by the stochastic area of the Brownian motion, the *Lévy area*, and is not affected by the choice of the integration scheme for the symmetric part. However, for some sequences approximating the Brownian motion, there is room on the second level for an extra term, the *area drift* or *area anomaly*.

When approximating an SDE driven by a Brownian motion in a classical way, we are looking for a sequence which does not have an area anomaly at the limit.

On the contrary, theorem 2.1.1 concentrates on a class of models which may exhibit a non-zero area anomaly at the limit. More precisely, we prove that Markov chains on periodic graphs (roughly speaking, graphs that are constructed by translation of a given finite graph) to which we add the area component, converge in the suitable rough path topology to $(B_t, \mathcal{A}_t + t\Gamma)_{t \in [0, \tau]}$, where $(B_t)_{t \geq 0}$ is the d -dimensional Brownian motion, $(\mathcal{A}_t)_t = (A_t(B))_t$ its associated Lévy area and Γ is a $d \times d$ antisymmetric matrix which represents the *area anomaly*.

One of our main goals is to show how getting a non-trivial area anomaly can be used to build new rough paths above Brownian motion and thus to go beyond classical calculus. Consequently, this is *not* a question of classical stochastic integration but a kind of *completion* of classical integration. Such deformations of classical stochastic calculus have been foreseen but never illustrated by explicit, simple, discrete Markovian processes with natural geometric interpretation. In [31], the authors exhibit the area anomaly of a magnetic field but this only concerns the continuous case. The article [60] studies how the area anomaly influences the behaviour of SDEs. The present paper is an attempt to fill in some of the blanks mentioned above and to show that area anomaly from theorem 2.1.1 is in fact a generic property of some renormalized discrete Markov chain and should be taken into account in the study of SDEs originating from many discrete processes.

Acknowledgements. D. S. thanks Q. Berger for valuable discussions on random walks in random environments at the origin of some constructions from the present paper. D. S. and O. L. want to thank the referee of the present article for valuable remarks and suggestions on bibliography on periodic graphs and their generalizations. D. S. is partially funded by the Grant ANR-14CE25-0014 (ANR GRAAL).

2.1.2 Structure of the present article

The present article is organized into five sections, the introduction being the first one. We present a very simple example of a Markov chain which exhibits a non-zero area anomaly in 2.1.3.

After introducing some useful definitions like those of Markov chains $(X_n)_{n \in \mathbb{N}}$ on periodic graphs and of stochastic signed area $(A_n(X))_{n \in \mathbb{N}}$ associated to them, we state our main result in 2.1.4. We then present a historical overview of some results that have a connection to ours. The section ends by a discussion on the consequences of theorem 2.1.1 on the universality class of the multidimensional Brownian motion, and expresses some caveats about the continuous description of the large size limit of discrete models.

In section 2.2, we describe the general settings of our theorem and state some useful results: in particular, we present a decomposition of the process $(X_n, A_n(X))_{n \in \mathbb{N}}$ which is based on excursion theory and inspired by renormalization ideas.

Section 2.3 is dedicated to the proof of our main result (theorem 2.1.1) which is a generalization to our class of Markov chains of the Donsker-type theorem for rough paths from [12].

Finally, in section 2.4, we present some applications of our result: we introduce a model in 3D for which we compute Γ by numerical simulations and we give an example of application to an SDE. We end the section with a list of open questions which arise in connection with the area anomaly Γ .

2.1.3 An easy discrete example: rotating sums of Bernoulli

r.v.

Let $(U_n)_{n \in \mathbb{N}^*}$ be a sequence of independent Bernoulli random variables such that $\mathbb{P}(U_1 = 1) = 1 - \mathbb{P}(U_1 = -1) = p$. We define two complex-valued processes $(Z_n)_{n \in \mathbb{N}}$ and $(Z'_n)_{n \in \mathbb{N}}$ in the following way: $(Z_n)_n$ is the random walk with increments chosen uniformly in $\{1, i, -1, -i\}$ and $(Z'_n)_n$ satisfies $Z'_0 = 0$ a.s. and, for $n \geq 1$, $Z'_n = \sum_{k=1}^n i^{k-1} U_k$. We set $X_n = \mathcal{R}(Z_n)$, respectively $X'_n = \mathcal{R}(Z'_n)$, and $Y_n = \mathcal{I}(Z_n)$, respectively $Y'_n = \mathcal{I}(Z'_n)$. A classical exercise in probability consists in checking that the laws of Z_n/\sqrt{n} and of $Z'_n/\sqrt{2np(1-p)}$ both converge to a normal law $\mathcal{N}(0, 1)$. Moreover, the processes $(X_n, Y_n)_n$ and $(X'_n, Y'_n)_n$ embedded in continuous time by linear interpolation converge both in law to a standard Brownian motion in the uniform topology. The discrete stochastic area of the process $(X_n, Y_n)_n$ is

defined as

$$A_n(X, Y) = \frac{1}{2} \sum_{1 \leq k < l \leq n} (X_k - X_{k-1})(Y_l - Y_{l-1}) - (Y_k - Y_{k-1})(X_l - X_{l-1}) \quad (2.2)$$

and $A(X', Y')_n$ is defined in the same way for the second process. The process $(A_n(X, Y)/n)_n$ embedded in continuous time is known to converge to the Lévy area of the Brownian motion; the present paper deals with the rescaled Lévy area $(A_n(X', Y')/(2np(1-p)))_n$ of the second process $(X'_n, Y'_n)_n$ and shows that, in the correct topology, it converges to the Lévy area of the Brownian motion with an *additional* drift γ . This drift is easily evaluated as:

$$\gamma = \lim_{n \rightarrow \infty} \frac{\mathbb{E}[A(X', Y')_n]}{2np(1-p)} = \frac{(2p-1)^2}{8p(1-p)} \quad (2.3)$$

and some additional computations show that the limit of the first higher cumulants of $(A_n(X', Y')/(2np(1-p)))_n$ coincide with the ones of the classical Lévy area.

Figure 2.1 describes the process (X'_n, Y'_n) as a Markov process in \mathbb{Z}^2 . In particular, the graph G_0 is induced on $\{0, 1, 2, 3\}$ by the Markov chain (X'_n, Y'_n) (we glue together the edges that connect a vertex to points of the same type). Figure 2.2 presents histograms of the marginal laws of $(X_n, Y_n, A_n(X, Y))$ and $(X'_n, Y'_n, A(X', Y')_n)$ obtained by numerical simulations. This figure shows that, in the large n limit and with the classical rescalings, the two processes are very similar, except for the additional drift γ in the Lévy area (the *area anomaly*). Up to our knowledge, such a limit process in continuous time has never been described.

Intuition about the similarities and differences between the two processes can be quickly explained by the following renormalization argument. The increments of $(Z_n - Z_{n-1})_n$ are independent, whereas only the increments $(Z'_{4n+4} - Z'_{4n})_n$ are independent. In a time interval $\{4n, 4n+1, 4n+2, 4n+3, 4n+4\}$, the increments of $(Z'_n)_n$ are bounded and thus do not contribute to the Brownian limit in the uniform topology; however, during the same time interval, correlations among these increments produce non-centered random areas. From a renormalization point of view, the local time correlations are irrelevant for the uniform topology but relevant for the rough path topology.

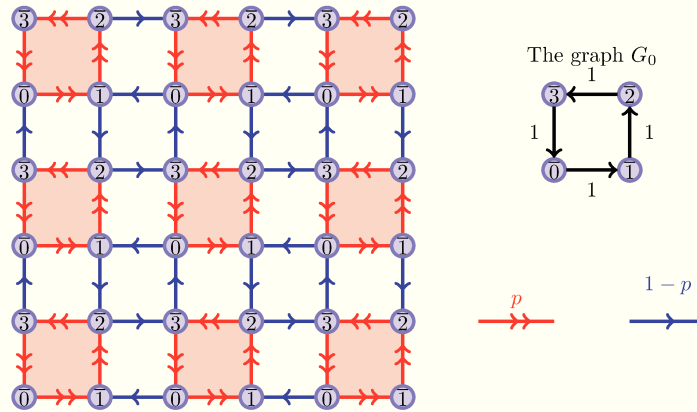


Figure 2.1: Markov chain $(Z'_n)_n$ of section 2.1.3. On the left hand side, its authorized transitions on the periodic graph $G = \mathbb{Z}^2$; on the top right hand side, its deterministic projection on the graph G_0 identified to $\mathbb{Z}/4\mathbb{Z}$. If $p \uparrow 1$, the chain tends to make more and more loops around the red shaded orbits.

2.1.4 Main result

Main theorem.

We start by properly defining the framework for our theorem.

Definition 2.1.1 (periodic subgraph). *Let E be a finite-dimensional vector space. A periodic subgraph of E is a infinite subset G of E such that:*

- (i) *all the points are separated,*
- (ii) *G is invariant under the translation action of a lattice $\Lambda \subset E$ on G .*

Property 2.1.1. *The graph can be decomposed as $G = \bigsqcup_{\lambda \in \Lambda} \lambda.G_0$ where G_0 is a finite subset of G and $\lambda.G_0$ is the translation of G_0 by $\lambda \in \Lambda$.*

This property means that any point x of G can be parametrized in a unique way as (λ, x_0) where $\lambda \in \Lambda$ and $x_0 \in G_0$. We write $\lambda = \pi_\Lambda(x)$ and $x_0 = \pi_0(x)$ for the two projections. We use alternatively the notation x or (λ, x_0) for a point in G .

An illustration of this property is the decomposition of the model from example 2.1.3 detailed in figure 2.1.

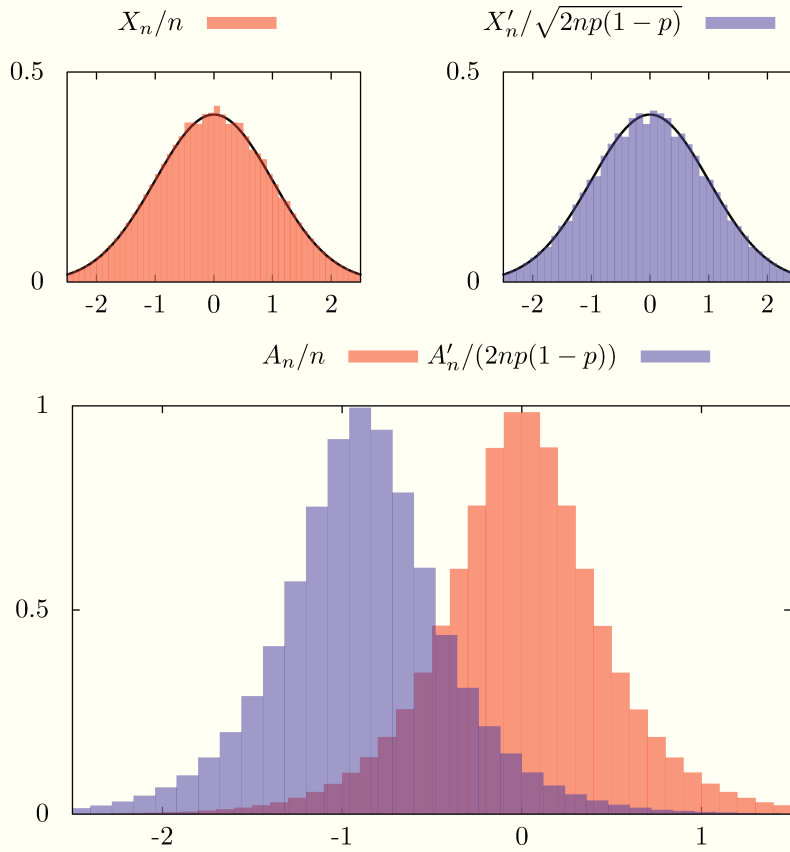


Figure 2.2: Empirical distributions of X_n/\sqrt{n} (top left), $X'_n/\sqrt{2np(1-p)}$ (top right), with the expected normal laws, and empirical distributions of the Lévy areas (below) A_n/n (red, right) and $A_n(X', Y)/(2np(1-p))$ (blue, left). The parameters are chosen as $n = 250000$ and $p = 0.9$. Data are accumulated over 64000000 independent realizations. The empirical means of $A_n(X, Y)$ and A'_n are $2.89 \cdot 10^{-5}$ and -0.88874 and their empirical standard deviations are 0.500031 and 0.499989 . The theoretical values are $\mathbb{E}[A_n(X, Y)/n] = 0$, $\mathbb{E}[A_n(X', Y)/(2np(1-p))] = 8/9 = 0.888\dots$ and $\sigma = 1/2$.

Definition 2.1.2 (invariant Markov chain on G). *Let G be a periodic subgraph of E . A G -valued Markov chain $(X_n)_{n \in \mathbb{N}}$ with transition law Q on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is Λ -invariant if and only if, for all $x, y \in G$ and for all $\lambda \in \Lambda$, $Q(x + \lambda, y + \lambda) = Q(x, y)$.*

As it will be explained section 2.2.1, such a Markov chain $(X_n)_n$ induces a Markov chain $(\pi_0(X_n))_n$ on G_0 . If the latter is irreducible, we can define a sequence of stopping times for it:

$$T_0 = 0, \\ T_{k+1} = \inf \{n > T_k : \pi_0(X_n) = \pi_0(X_0)\}, \quad k \geq 0.$$

Since G_0 is finite, for any initial law μ on G_0 , T_1 has finite moments of all orders.

For an E -valued sequence $(x_n)_n$, we introduce its continuous rescaled version given by

$$x_t^N = \frac{x_{\lfloor Nt \rfloor} + (Nt - \lfloor Nt \rfloor)(x_{\lfloor Nt \rfloor + 1} - x_{\lfloor Nt \rfloor})}{\sqrt{N}}$$

as in the classical Donsker theorem. Then the rough path corresponding to $x^{(N)}$ is defined as

$$\iota^{(N)}(x_\bullet, A_\bullet(x))_t := \left(x_t^N, A_t(x^N) \right) \quad (2.4)$$

where $A_t(x^N)$ is given by the formula (2.1). These variables belong to the space $G^2(E)$, which is described in section 2.2.4. We denote by δ_ϵ the standard homogeneous dilatation on $G^2(E)$: $\delta_\epsilon(x, a) = (\epsilon x, \epsilon^2 a)$.

We can now state our main theorem.

Theorem 2.1.1. *Let G be a Λ -periodic graph on a finite dimensional vector space E . Let $(X_n)_n$ be a G -valued Λ -invariant Markov chain on G with bounded increments (i.e. there exists $R > 0$ such that $|X_{n+1} - X_n|_E \leq R$ a.s.) and such that $(\pi_0(X_n))_{n \in \mathbb{N}}$ is irreducible.*

Let $v = \mathbb{E}[T_1]^{-1} \mathbb{E}[X_{T_1}] \in E$ and $\beta = \mathbb{E}[T_1] \in \mathbb{R}_+^$. Let $(\tilde{X}_n)_{n \in \mathbb{N}}$ be the E -valued process defined by $\tilde{X}_n = X_n - nv$. Up to a dimensional reduction and a linear transformation of the graph G , the covariance matrix of $X_{T_1} - T_1 v$ may always be assumed to be CI_d with $C > 0$ and d the dimension of V .*

For any $\tau > 0$, we have the following convergence in distribution:

$$\left(\delta_{\sqrt{C^{-1}\beta}} \iota^{(N)}(\tilde{X}_\bullet, A_\bullet(\tilde{X}))_t \right)_{0 \leq t \leq \tau} \xrightarrow{N \rightarrow \infty} (B_t, \mathcal{A}_t + t\Gamma)_{0 \leq t \leq \tau} \quad (2.5)$$

in the topology of $\mathcal{C}^{0,\alpha\text{-Hölder}}([0, \tau], G^2(E))$ for $\alpha < 1/2$, where B is a Brownian motion on E , \mathcal{A} its Lévy area as defined by classical stochastic calculus and Γ a constant antisymmetric matrix, the area anomaly, given by the antisymmetric part of $C^{-1}\mathbb{E} \left[\sum_{1 \leq k < l \leq T_1} \Delta \tilde{X}_k \otimes \Delta \tilde{X}_l \right]$ (detailed formula in (2.19)).

Some historical background.

Until now, in the rough path setting, most of the convergence theorems have dealt with processes with i.i.d. centered increments. Based on the Stroock-Varadhan's result from [76], a Donsker theorem for rough paths in the $\mathcal{C}^\alpha([0, 1], \mathbb{R}^d)$ topology for $\alpha < 1/2$ has been proved in theorem 3 from [12], which our theorem 1.1 generalizes. In [6], the discrete sequence converging to the Brownian rough path is constructed by the concatenation (in the rough path sense) of renormalized i.i.d. copies of the cubature formula on Wiener space. The main result of [20] gives sufficient conditions for convergence in distribution of a random walk on $G^N(\mathbb{R}^d)$ to a Lévy process on $G^N(\mathbb{R}^d)$ in a suitable rough path topology. While [6] and [20] are also generalizations of theorem 3 from [12], they do not apply to the class of processes described in our theorem 1.1, as the increments of these processes are not necessarily i.i.d. Moreover, none of these results is concerned by the study of the area anomaly, as we will see that it is trivial when the discrete process has i.i.d. increments.

A more general setting, which encompasses that of rough paths, is given by random walks on different types of groups. A result due to Wehn (see, for example, [11], theorem 1.3, for details) states that, when μ is a centered probability measure on a connected Lie group, μ^{*n} (the n th convolution of μ) converges to the Wiener measure (under certain conditions on μ). In [1], the main result states that, when μ is a probability measure with finite support on a discrete group of polynomial volume growth (nilpotent Lie groups, and in particular $G^N(\mathbb{R}^d)$, are of polynomial volume growth), μ^{*n} converges to the heat kernel of a centered left-invariant sub-Laplacian on a certain simply connected nilpotent Lie group. In both cases, we deal with i.i.d. increments and no area anomaly is exhibited at the limit. However,

in [1] the possibility of a non-centered measure μ^{*n} is taken into account, just to show that the asymptotic behavior is similar to the non-centered case modulo a transformation by a multiplicative function (which is equivalent to re-centring μ). What this shows in particular is that *our area anomaly is not a question of the process drift* but a new phenomenon.

In the uniform topology, the convergence of processes similar to ours is widely studied. In [29] and [53], authors have already considered the convergence of random walks on periodic graphs and Markov chains on graphs respectively, and in [5] an invariance principle has been proved for a certain class of random walks in random environment. In [55] and [48], the authors study the convergence of a random walk, symmetric and non-symmetric respectively, on a lattice graph through the convergence of the corresponding discrete heat kernel. Their lattice graph is a generalization to a Riemannian manifold of the notion of periodic graph from our article (a detailed theory on lattice graphs and the finite quotient graphs, as well as their properties, can be found in chapters 3 and 4 from [77]). While there is no room for an area anomaly for the reversible random walks from [55] (the reversibility of the process implies a zero area anomaly), the loops that can be present in the processes from [5] and [48] might generate a non-zero limit stochastic area drift.

Discussion on the renormalization constant.

Let us now stop briefly to explain the choice of our renormalization constant, namely v . It is immediate to ask why we didn't simply set $\tilde{X}_n = X_n - \mathbb{E}[X_n]$. In this case, we argue that, first, we can not get an explicit infinite constant (of the type nv), second, it is a *sufficient* but not a *necessary* drift and finally, this would not allow us to get an explicit expression of Γ of the type (2.19).

As a consequence of the ergodic theorem applied to the Markov chain $(\pi_0(X_n))_n$, we could have also defined v as

$$v = \sum_{x_0 \in G_0} \nu(x_0) \sum_{y \in G} Q(x^*, y)(y - x^*)$$

where ν the invariant probability of $(\pi_0(X_n))_n$ and x^* a representative of the

equivalence class of x_0 in G ($z \sim z'$ if $\pi_0(z) = \pi_0(z')$). This is similar to the approach the authors adopt in [48]. Our choice of the expression of the renormalization constant is motivated by the desire to highlight the centering of the excursions, which is indispensable for applying the Donsker type theorem from [12].

Symmetry of the process and area anomaly.

In [48], $\rho_{\mathbb{R}}(\gamma_p)$, the analogue of the constant v called the asymptotic direction, is zero if and only if the random walk on the quotient graph X_0 is symmetric.

In our case, the symmetry of the random walk on G_0 is not necessary anymore for v to be zero. In the example from 2.1.3, the random walk on G_0 is not symmetric whatever value of p we choose, and nevertheless $v = 0$ for $p = 1/2$.

On the other hand, the symmetry of $(X_n)_n$ (or its reversibility) will be a sufficient condition for $v = 0$ and also for $\Gamma = 0$. The most immediate example is provided by the framework of the classical Donsker theorem, i.e. the case when we are dealing with sums of i.i.d. centred r.v.

However, we can have $(X_n)_n$ non-symmetric and simultaneously $v = 0$ and $\Gamma \neq 0$ (see example 2.1.3), which shows that the area anomaly *is not a product of the drift of the process*. We can also have $v \neq 0$ and $\Gamma = 0$ in the case of i.i.d. non-centered r.v. (we need, of course, to re-center the variables before we pass to the limit).

Consequences of theorem 2.1.1.

The hypotheses of theorem 2.1.1 are satisfied in many models coming from statistical mechanics where jumps in space are often local. Up to our knowledge, the area anomaly is a new feature never described in any model, even if the examples that we present look natural. One may wonder whether this area anomaly is relevant. We now explain why it is the case.

The general philosophy beyond renormalization and large scale limits of discrete models is to build continuous models such that they are large scale limits of various discrete models and such that it is possible to compute directly with them.

Phrased in a provocative way, our theorem implies in particular that *a two-dimensional standard Brownian motion may not be the same as two independent one-dimensional Brownian motions* as soon as one wishes to use it to drive a

stochastic differential equation. The difference lies in the area anomaly Γ which is irrelevant at the level of the positions $(B_{t_k})_k$ but is relevant in non-linear SDEs.

Thus, when several Brownian motions emerge in the description of the limit of discrete processes, the consequence of the previous theorem is that one *needs in general to wonder* about the presence of area anomalies between components before writing down any stochastic integration.

Hopefully in many cases, it is easy to prove without any calculation that the area anomaly is zero: as it has already been mentioned, this is the case for reversible processes. However, for irreversible Markov chains, especially useful in non-equilibrium statistical mechanics, one should expect in general a non-zero anomaly.

A detailed study of the area anomaly Γ and its generalization to a larger class of processes will be present in a next paper in preparation.

2.2 Tools and additional results

2.2.1 Additional results on Λ -invariant Markov chains

Property 2.2.1. *Let $(X_n)_{n \in \mathbb{N}}$ be a Λ -invariant Markov chain on a periodic subgraph G of E as in definition 2.1.1. Then the process $(\pi_0(X_n))_{n \in \mathbb{N}}$ is a G_0 -valued Markov chain.*

Proof. Let f be any bounded Borel function $G_0 \rightarrow E$.

$$\begin{aligned} \mathbb{E}[f(\pi_0(X_{n+1})) | \mathcal{F}_n] &= \mathbb{E} \left[\sum_{\lambda \in \Lambda} \mathbf{1}_{\pi_\Lambda(X_{n+1})=\lambda} f(\pi_0(X_{n+1})) \middle| \mathcal{F}_n \right] \\ &= \sum_{\lambda \in \Lambda} \mathbb{E} \left[\mathbf{1}_{\pi_\Lambda(X_{n+1})=\lambda} f(\pi_0(X_{n+1})) \middle| \mathcal{F}_n \right] = \sum_{\lambda \in \Lambda} (Qg_\lambda)(X_n) \end{aligned}$$

where $g_\lambda(x) = \mathbf{1}_{\pi_\Lambda(x)=\lambda} f(\pi_0(x))$ by the Markov property for $(X_n)_n$. The invariance of Q gives now:

$$\begin{aligned} (Qg_\lambda)(x) &= \sum_{y \in G} Q(x, y) g_\lambda(y) = \sum_{y_0 \in G} Q(x, y) \mathbf{1}_{\pi_\Lambda(y)=\lambda} f(\pi_0(y)) \\ &= \sum_{y_0 \in G_0} Q((\pi_\Lambda(x), \pi_0(x)), (\lambda, y_0)) f(y_0) = \sum_{y_0 \in G_0} Q((0, \pi_0(x)), (\lambda - \pi_\Lambda(x), y_0)) f(y_0) \end{aligned}$$

by Λ -invariance for Q . Summation over Λ eliminates the dependence on $\pi_\Lambda(x)$ and we thus obtain:

$$\mathbb{E}[f(\pi_0(X_{n+1}))|\mathcal{F}_n] = (Q_0 f)(\pi_0(X_n))$$

with $Q_0(x_0, y_0) = \sum_{\lambda \in \Lambda} Q((0, x_0), (\lambda, y_0))$. \square

Moreover, similar calculations show that the process $(\pi_\Lambda(X_n))_n$ knowing the process $(\pi_0(X_n))_n$ is a heterogeneous Markov chain whose rates¹ depend on the $(\pi_0(X_n))_n$.

2.2.2 Decomposition into pseudo-excursions

We start with a general definition of pseudo-excursions for an E -valued sequence:

Definition 2.2.1. *Let $(x_n)_{n \in \mathbb{N}}$ be an E -valued sequence and $(T_k)_{k \in \mathbb{N}}$ be a strictly increasing sequence in \mathbb{N} such that $T_0 = 0$ and $T_{k+1} - T_k = L_k > 0$. We introduce the sequence $\tilde{\lambda}_p(x) := x_{T_p}$. Let o be an additional cemetery point added to E . The pseudo-excursions $\widetilde{\text{Exc}}^{(p)}(x)$ of the sequence $(x_n)_{n \in \mathbb{N}}$ are then defined as $E \cup \{o\}$ -valued processes through:*

$$\widetilde{\text{Exc}}^{(k)}(x)_n = \begin{cases} x_{T_k+n} - \tilde{\lambda}_k(x) & \text{if } 0 \leq n \leq L_k \\ o & \text{if } n > L_k \end{cases}$$

The global trajectory $(x_n)_n$ can be recovered from the excursions by:

$$x_n = \tilde{\lambda}_{\kappa(n)}(x) + \widetilde{\text{Exc}}^{(\kappa(n))}(x)_{n-T_{\kappa(n)}} \quad (2.6)$$

where $\kappa(n)$ is the unique integer such that $T_{\kappa(n)} \leq n < T_{\kappa(n)+1}$.

We will now give a definition of pseudo-excursions which applies to a specific class of G -valued sequences we are interested in (with G as in definition 2.1.1). For this purpose, we will slightly change definition 2.2.1.

For $(x_n)_{n \in \mathbb{N}} \in G^{\mathbb{N}}$, we introduce the sequence of excursion times of $(\pi_0(x_n))_{n \in \mathbb{N}}$

¹the probability of a jump between λ and λ' between times k and $k+1$ is $Q_k(\lambda, \lambda' | \pi_0(X)) = Q((\lambda, X_k), (\lambda', X_{k+1})) / Q_0(X_k, X_{k+1})$.

from its original point:

$$\begin{aligned} T_0 &= 0, \\ T_{k+1} &= \inf \{n > T_k : \pi_0(x_n) = \pi_0(x_0)\}, \quad k \geq 0. \end{aligned}$$

Definition 2.2.2. Let $(x_n)_{n \in \mathbb{N}}$ be a G -valued sequence such that $(\pi_0(x_n))_{n \in \mathbb{N}}$ is recurrent (i.e. each point of G_0 appears an infinity of times in the sequence). Set $\lambda_k(x) = \pi_\Lambda(x_{T_k})$ and $L_k = T_{k+1} - T_k$ (L_k is the duration of an excursion). Let o be an additional cemetery point added to G . The pseudo-excursions $\text{Exc}^{(k)}(x)$ of the sequence $(x_n)_{n \in \mathbb{N}}$ are defined as $G \cup \{o\}$ -valued processes through:

$$\text{Exc}^{(k)}(x)_n = \begin{cases} x_{T_k+n} - \lambda_k(x) & \text{if } 0 \leq n \leq L_k \\ o & \text{if } n > L_k \end{cases} \quad (2.7)$$

Although the above definition can be viewed as a particular case of definition 2.2.1, its interest consists in exploiting the decomposition of elements of G in $\Lambda \times G_0$ -valued couples. This enables us to translate only the Λ -valued component and thus to start each pseudo-excursion from a point $y \in G$ such that $\pi_0(y) = \pi_0(x_0)$. Moreover, as we keep here close to the classical definition of excursions, we can deal with the Markov chain $(X_n)_{n \in \mathbb{N}}$ (and not only $(\pi_0(X_n))_{n \in \mathbb{N}}$) and make computations of the Lévy area easier. In the rest of the article, we will prefer definition 2.2.2 when we talk of a (recurrent) G -valued sequence, and definition 2.2.1 will apply whenever we make a statement concerning any E -valued sequence.

One immediately checks that $\text{Exc}^{(k)}(x)_0 = \pi_0(x_0)$ and $\text{Exc}^{(k)}(x)_{L_k} = \pi_0(x_0) + \lambda_{k+1}(x) - \lambda_k(x)$. Our construction of pseudo-excursions makes them invariant under translation by an element of Λ : $\text{Exc}^{(k)}(\mu + x) = \text{Exc}^{(k)}(x)$ where μ is any element of Λ .

Property 2.2.2. Let $(X_n)_n$ be a Λ -invariant Markov chain on the periodic graph G such that the projection $(\pi_0(X_n))_n$ is irreducible. The $(G \cup \{o\})^{\mathbb{N}}$ -valued random variables $(\text{Exc}^{(k)}(X))_{k \in \mathbb{N}}$ are independent and identically distributed.

Proof. The proof relies on the repetitive use of the strong Markov property. Let $n \in \mathbb{N}^*$ and let $f_0, f_1, \dots, f_n : (G \cup \{o\})^{\mathbb{N}} \rightarrow \mathbb{R}$ be bounded measurable functions.

The random times T_k are stopping times, which are finite almost surely. For $k \leq n-1$, the random variables $\mathcal{E}_k = \text{Exc}^{(k)}(X)$ are \mathcal{F}_{T_k} -measurable, and we obtain:

$$\mathbb{E}_x [f_0(\mathcal{E}_0) \dots f_{n-1}(\mathcal{E}_{n-1}) f_n(\mathcal{E}_n) | \mathcal{F}_{T_n}] = f_0(\mathcal{E}_0) \dots f_{n-1}(\mathcal{E}_{n-1}) \mathbb{E}_x [f_n(\mathcal{E}_n) | \mathcal{F}_{T_n}]$$

The strong Markov property thus yields:

$$\begin{aligned} \mathbb{E}_x [f_n(\mathcal{E}_n) | \mathcal{F}_{T_n}] &= \mathbb{E}_{X_{T_n}} [f_n(\mathcal{E}_0)] = \mathbb{E}_{(\lambda_n(x), \pi_0(x))} [f_n(\mathcal{E}_0)] \\ &= \mathbb{E}_{(\lambda_0(x), \pi_0(x))} [f_n(\text{Exc}^{(0)}(X + \lambda_n(x) - \lambda_0(x)))] = \mathbb{E}_{(\lambda_0(x), \pi_0(x))} [f_n(\mathcal{E}_0)] \end{aligned}$$

a.s., where the last equality is deduced from the Λ -invariance of the process and of the pseudo-excursions. One now remarks that the last term does not depend anymore on X_{T_n} . By recursion, we obtain the final result:

$$\mathbb{E}_x [f_0(\mathcal{E}_0) \dots f_{n-1}(\mathcal{E}_{n-1}) f_n(\mathcal{E}_n)] = \mathbb{E}_x [f_0(\mathcal{E}_0)] \mathbb{E}_x [f_1(\mathcal{E}_0)] \mathbb{E}_x [f_n(\mathcal{E}_0)]$$

□

Corollary 2.2.1. *The random variables $(\lambda_{k+1}(X) - \lambda_k(X))_{k \in \mathbb{N}}$ are also i.i.d.*

Proof. This follows directly from $\lambda_{k+1}(X) - \lambda_k(X) = \text{Exc}^{(k)}(X)_{L_k} - \text{Exc}^{(k)}(X)_0$ and the independence of the pseudo-excursions. □

Remark: If G is a periodic graph and $M \in GL_n(\mathbb{R})$, then $MG = \{Mx; x \in G\}$ is again a periodic graph (with possibly degenerate vertices). If $(X_n)_n$ is a Λ -invariant Markov chain on G , then $(MX_n)_n$ is again a $M\Lambda$ -invariant Markov chain on MG . We assume all through the paper that $\mathbb{R}^n = \text{span } \Lambda$; if this is not the case, we embed the graph G in the smaller space $\text{span } \Lambda$ isomorphic to some \mathbb{R}^n . Let \mathbf{C} be the covariance matrix of the increment $\lambda_1(X) - \lambda_0(X)$, we always assume that $\mathbf{C} = M^* I_n M$. If $\text{rank } \mathbf{C} < n$, we again embed our Markov chain in a smaller graph in a smaller space such that $\text{rank } \mathbf{C} = n$. Then, up to reduction to a smaller space and up to an invertible linear transformation of the graph, we may always assume that $\mathbf{C} = I_n$. In particular, under $\mathbf{C} = I_n$, the Donsker embedding of the random walk $(\lambda_k(x))_k$ converges to a standard Brownian motion on \mathbb{R}^n . Moreover,

the covariance matrix \mathbf{C} and the drift v are the analogues of the Albanese metric and the asymptotic direction respectively from the article [48].

Example of section 2.1.3. The Markov chain (Z'_n) fits into this framework with $G = \mathbb{Z}^2$. The non-zero elements of the transition matrix Q are represented in 2.1. The matrix Q is Λ -invariant with $\Lambda = (2\mathbb{Z})^2$. The set $G_0 = \{(0,0), (1,0), (0,1), (1,1)\}$ may be identified to $\mathbb{Z}/4\mathbb{Z}$, and so the Markov process $(\pi_0(Z'_n))_n$ is actually deterministic and corresponds to the shift $x \mapsto x + 1$ as in figure 2.1.

2.2.3 Area process and rough paths

Definition 2.2.3 (area sequence). *Let E be a finite-dimensional vector space. Let $(e_i)_{1 \leq i \leq d}$ be a basis of E . We write $x^{(i)}$ for the i -th coordinate of a vector $x \in E$ w.r.t. the basis $(e_i)_{1 \leq i \leq d}$. For any E -valued sequence $(x_n)_{n \in \mathbb{N}}$, we introduce the sequence of antisymmetric $d \times d$ matrices $(A_n(x))_{n \in \mathbb{N}}$ defined by $A_0(x) = A_1(x) = 0$ and, for any $n \geq 2$,*

$$A_n^{ij}(x) = \sum_{1 \leq k < l \leq n} \left((\Delta x^{(i)})_k (\Delta x^{(j)})_l - (\Delta x^{(j)})_k (\Delta x^{(i)})_l \right) \quad (2.8)$$

with $(\Delta u)_k = u_k - u_{k-1}$ for any sequence $(u_n)_n$.

This definition can be tied easily to the stochastic area $A_t(x^{(N)})$ of $x^{(N)}$ from formula 2.4: it is easy to check that

$$A_t(x^{(N)}) = \frac{A_{\lfloor Nt \rfloor}(x) + (Nt - \lfloor Nt \rfloor)(A_{\lfloor Nt \rfloor + 1}(x) - A_{\lfloor Nt \rfloor}(x))}{N} \quad (2.9)$$

Property 2.2.3 (decomposition of an area sequence along excursions). *Let $(x_n)_{n \in \mathbb{N}}$ and $(T_k)_{k \in \mathbb{N}}$ be as in definition 2.2.1. Then the following decomposition holds:*

$$A_{T_n}^{ij}(x) = \sum_{p=0}^{n-1} A_{L_p}^{ij}(\widetilde{\text{Exc}}^{(p)}(x)) + A_n^{ij}(\tilde{\lambda}(x)) \quad (2.10)$$

In the particular case when $(x_n)_{n \in \mathbb{N}}$ and $(T_k)_{k \in \mathbb{N}}$ are as in definition 2.2.2, we

have:

$$A_{T_n}^{ij}(x) = \sum_{p=0}^{n-1} A_{L_p}^{ij}(\text{Exc}^{(p)}(x)) + A_n^{ij}(\lambda(x)) \quad (2.11)$$

Proof. By definition, the l.h.s. uses a double sum over $1 \leq k < l \leq T_n$. We split the interval $\{1, 2, \dots, T_n\}$ into $J_p = \{T_p + 1, \dots, T_{p+1}\}$ for $p = 0, \dots, n-1$ and we classify the indices k and l : either they are in the same subset J_p or they belong respectively to J_{p_1} and J_{p_2} with $p_1 < p_2$.

In the first case, the sum over $T_r + 1 \leq k < l \leq T_{r+1}$ gives the area of the r -th excursion $A_{L_r}^{ij}(\text{Exc}^{(r)}(x))$.

In the second case, the sum over $k \in J_{p_1}$ and $l \in J_{p_2}$ factorizes into two sums, evaluated as telescopic sums respectively to $\lambda_{p_i+1}(x) - \lambda_{p_i}(x)$ for $i = 1, 2$. The remaining sum over $0 \leq p_1 < p_2 \leq n-1$ gives the (signed) area of $(\lambda_k(x))_{k \in \mathbb{N}}$ between 0 and n . \square

We need a last lemma, easy to prove, from linear algebra, about the transformation of the area under a linear transformation M of the space E .

Lemma 2.2.1 (covariance of the area). *Let $(x_n)_n$ be an E -valued sequence and $M \in GL_n(E)$. The area process $(A_n^{ij}(x))_n$ of the sequence $(Mx_n)_n$ in E is given by:*

$$A_n^{ij}(Mx) = \sum_{1 \leq k, l \leq n} M_{ik} M_{jl} A_n^{kl}(x)$$

2.2.4 The group $G^2(E)$

The general construction.

In this section, we rewrite some results from the rough path theory from [34] (in particular from chapter 7) in order for them to correspond to the case of a finite-dimensional vector space E on \mathbb{R} . We concentrate on the case that is of interest to this article, namely $N = 2$. For more details and the general case $N \geq 2$ see [34] or [32].

We introduce the *tensorial truncated algebra* $T^{(2)}(E) = \bigoplus_{k=0}^2 E^{\otimes k}$, where \otimes is the tensorial product on E ($E^{\otimes 0} = \mathbb{R}$) and \bigoplus denotes a direct sum. Endowed with

the multiplication law

$$(a_0, a_1, a_2) \otimes_2 (b_0, b_1, b_2) = (a_0 b_0, a_0 b_1 + b_0 a_1, a_0 b_2 + b_0 a_2 + a_1 \otimes a_2), \quad (2.12)$$

it is a non-commutative algebra with unit element $(1, 0_E, 0_{E \otimes E})$.

To $x \in \mathcal{C}^{1-var}([s, t], E)$ (the set of all continuous paths of finite 1-variation), we associate the element of $T^{(2)}(E)$ given by

$$S_2(x)_{s,t} = \left(1, \int_{s < u < t} dx_u, \int_{s < u_1 < u_2 < t} dx_{u_1} \otimes dx_{u_2} \right) \in T^{(2)}(E)$$

This object satisfies *Chen's relation*, i.e., for $0 \leq s < r < t \leq 1$,

$$S_2(x)_{s,t} = S_2(x)_{s,r} \otimes_2 S_2(x)_{r,t} \quad (2.13)$$

and, in this particular case, \otimes_2 can be viewed as a path concatenation operator.

As in section 7.5.1 in [34], we define the set $G^2(E)$ by

$$G^2(E) = \{S_2(x)_{0,1} : x \in \mathcal{C}^{1-var}([0, 1], E)\}$$

We now denote by $\mathbb{X} = (1, \mathbb{X}^{(1)}, \mathbb{X}^{(2)})$ an element of $G^2(E)$, where $\mathbb{X}^{(1)}$ stands for the first-order and $\mathbb{X}^{(2)}$ for the second-order increments. Implicitly, $\mathbb{X}_{s,t} = S_2(x)_{s,t}$ for some $x \in \mathcal{C}^{1-var}([0, 1], E)$, and $\mathbb{X}_t = \mathbb{X}_{0,t}$. Since the symmetrical part of $\mathbb{X}^{(2)}$ depends on $\mathbb{X}^{(1)}$ (as $\int y dy = \frac{1}{2}y^2$), we can cut off redundant information by transforming $\mathbb{X}^{(2)}$ into $\mathbf{X}_t^{(2);i,j} = \int_0^t (X_s^i - X_0^i) dX_s^j - \int_0^t (X_s^j - X_0^j) dX_s^i$ for $1 \leq i, j \leq \dim(E)$. Under this new form, the element \mathbb{X} belongs to the space $\bigoplus_{k=0}^2 E^{\wedge k}$, where \wedge is the antisymmetric tensor product on E : for $u, v \in E$, $u \wedge v = u \otimes v - v \otimes u$.

For commodity reasons, we can use a more informal notation by neglecting the first component (the identity element) of \mathbb{X} .

Of course, the elements of the type $(x_t, A_t(x))$, which are the ones we are interested in, belong to $G^2(E)$. In particular, we can isolate from the operation \otimes_2 a very important property of the stochastic area, namely, for $0 \leq s < t$

$$A_t^{ij}(x) = A_s^{ij}(x) + A_{s,t}^{ij}(x) + \frac{1}{2}(x_s^{(i)}(x_t - x_s)^{(j)} - x_s^{(j)}(x_t - x_s)^{(i)}) \quad (2.14)$$

The Carnot-Caratheodory distance on $G^2(E)$.

It is natural to ask what is the shortest path in E for a given signature. The answer to this question allows to define the *Carnot-Caratheodory norm on $G^2(E)$* by

$$\|g\| := \inf \left\{ \int_0^1 |dx| : x \in \mathcal{C}^{1-var}([0, 1], E) \text{ and } S_2(x)_{0,1} = g \right\} \quad (2.15)$$

where $|\cdot|_E$ is a restriction to E of the Euclidean norm.

Since the norm thus defined is homogeneous ($\|\delta_\lambda g\| = |\lambda| \|g\|$ for $\lambda \in \mathbb{R}$), symmetric ($\|g\| = \|g^{-1}\|$) and sub-additive ($\|g \otimes h\| \leq \|g\| + \|h\|$), it induces a left-invariant, continuous metric \mathbf{d} on $G^2(E)$ through the application

$$\begin{aligned} \mathbf{d} : G^2(E) \times G^2(E) &\rightarrow \mathbb{R}_+ \\ (g, h) &\mapsto \|g^{-1} \otimes_2 h\| \end{aligned} \quad (2.16)$$

In this case, $(G^2(E), \mathbf{d})$ is a geodesic space (in the sense of definition 5.19 from [34]). It is also a Polish space (corollary 7.50 from [34]).

The Carnot-Caratheodory norm is difficult to use for practical estimations but we can give it a good upper bound:

Proposition 2.2.1. *There exists $\nu > 0$ such that, for \mathbf{d} defined as above, for any $\mathbb{X} \in G^2(E)$ and $0 \leq s < t \leq 1$, we have:*

$$\mathbf{d}(\mathbb{X}_s, \mathbb{X}_t) = \|\mathbb{X}_{s,t}\| \leq \nu \left(\left| \mathbf{X}_{s,t}^{(1)} \right|_E + \left| \mathbf{X}_{s,t}^{(2)} \right|_{E \otimes E}^{\frac{1}{2}} \right) \quad (2.17)$$

2.3 Proof of main theorem and comments

2.3.1 Proof of theorem 2.1.1

We denote by $|\cdot|$ the absolute value on \mathbb{R} , by $|\cdot|_E$ the euclidean norm on the finite-dimensional vector space E and by $|\cdot|_{E \otimes E}$ the induced matrix norm on $E \otimes E$: $|A|_{E \otimes E} = \sup_{|x|_E=1} |Ax|_E$. We also use the norm $\|\cdot\|$ on $G^2(E)$ and the associated distance $\mathbf{d}(\cdot, \cdot)$ (formula (2.16)). We set, for $u, v \in G^2(E)^l$, $\mathbf{d}_l((u^1, \dots, u^l), (v^1, \dots, v^l)) = \sum_{i=1}^l \mathbf{d}(u^i, v^i)$: \mathbf{d}_l is a distance on $G^2(E)^l$.

For any $n \in \mathbb{N}$, we define $\kappa(n)$ as the unique integer such that $T_{\kappa(n)} \leq n < T_{\kappa(n)+1}$, where the T_n s are as in definitions 2.2.1 or 2.2.2 (as has already been done in section 2.2.2).

Proof. Since $\tilde{\lambda}_k(\tilde{X}) = \lambda_k(X) - T_k v$ (with $\lambda_k(X)$ and $\tilde{\lambda}_k(\tilde{X})$ as in definitions 2.2.2 and 2.2.1 respectively), the process $(\tilde{\lambda}_k(\tilde{X}))_k$ is an E -valued centered random walk (not Λ -valued because of the correction). As it has been stated in the theorem, up to a dimensional reduction and a linear transformation of the graph G , the covariance matrix of $\lambda_1(X) - T_1 v$ may always be assumed to be CI_n with $C > 0$, so each (E -valued) increment has a covariance matrix equal to CI_n .

The main idea of the proof of theorem 2.1.1 is to use the theory of pseudo-excursions from section 2.2.2 and the decomposition from property 2.2.3 in order to extract convergence to the standard Brownian rough path through the process $(\tilde{\lambda}_n(\tilde{X}))_{n \in \mathbb{N}}$ using theorem 3 [12], convergence to the area anomaly through the independence of pseudo-excursions, and tightness from additional results on pseudo-excursions. Consequently, the proof of theorem 2.1.1 is divided into 4 steps:

- lemma 2.3.1: convergence of the centered discrete process $\left(\iota^{(N)}(\tilde{\lambda}_\bullet(\tilde{X}), A_\bullet(\tilde{\lambda}(\tilde{X})))_t \right)_{0 \leq t \leq \tau}$
- lemma 2.3.2: convergence of the extracted process $\left(\iota^{(N)}(\tilde{X}_{T_\bullet}, A_{T_\bullet}(\tilde{X}))_t \right)_{0 \leq t \leq \tau}$ and emergence of the area anomaly (drift) Γ
- lemma 2.3.3: convergence of finite-dimensional marginals of the full process $\left(\iota^{(N)}(\tilde{X}_\bullet, A_\bullet(\tilde{X}))_t \right)_{0 \leq t \leq \tau}$
- lemma 2.3.4: tightness of the sequence $\left(\iota^{(N)}(\tilde{X}_\bullet, A_\bullet(\tilde{X})) \right)_{N \in \mathbb{N}}$

Lemma 2.3.1. *The process $\left(\iota^{(N)}(\tilde{\lambda}_\bullet(\tilde{X}), A_\bullet(\tilde{\lambda}(\tilde{X})))_t \right)_{0 \leq t \leq \tau}$ converges in distribution to the Lévy lift on $G^2(E)$ of a Brownian motion $(B_t)_{t \geq 0}$:*

$$\left(\delta_{\sqrt{C^{-1}}} \iota^{(N)}(\tilde{\lambda}_\bullet(\tilde{X}), A_\bullet(\tilde{\lambda}(\tilde{X})))_t \right)_{0 \leq t \leq \tau} \xrightarrow[N \rightarrow \infty]{(d)} (B_t, \mathcal{A}_t)_{0 \leq t \leq \tau}$$

in the topology of $\mathcal{C}^{0, \alpha\text{-Holder}}([0, \tau], G^2(E))$ for $\alpha < 1/2$.

Proof. This is a direct consequence of the Donsker-type theorem for a sequence of i.i.d. centered $G^2(\mathbb{R}^2)$ -valued random variables from [12]. In this article, the

authors use a central limit theorem for centered i.i.d. variables on a nilpotent Lie group in order to prove the convergence of finite-dimensional distributions, and Kolmogorov's criterion to prove the tightness of the sequence. \square

Lemma 2.3.2. *The sequence of processes $(\delta_{\sqrt{C-1}} \iota^{(N)}(\tilde{X}_{T_\bullet}, A_{T_\bullet}(\tilde{X}))_t)_{0 \leq t \leq \tau}$ converges in distribution to $(B_t, \mathcal{A}_t + t\Gamma)_{0 \leq t \leq \tau}$ in the topology of $\mathcal{C}^{0,\alpha\text{-Holder}}([0, \tau], G^2(E))$ for $\alpha < 1/2$, with Γ given by (2.19).*

This is the part of the proof where the area anomaly Γ first appears. We will see that, between 2.3.1 and 2.3.2, nothing changes on the first level of the new sequence, since the embedding is obtained by linear interpolation and therefore does not keep track of the trajectory between T_n and T_{n+1} . Simultaneously, a complementary term appears on the second level, in the expression of the stochastic area. This is due to the fact that, whereas the specific trajectory of an excursion is not memorized, its area is registered in the continuous embedding.

Proof. We have trivially by definition 2.2.1:

$$\tilde{X}_{T_n} = \tilde{\lambda}_n(\tilde{X})$$

Moreover, property 2.2.3 applied to \tilde{X} gives:

$$A_{T_n}^{ij}(\tilde{X}) = A_n^{ij}(\tilde{\lambda}(\tilde{X})) + \sum_{p=0}^{n-1} A_{L_p}^{ij}(\widetilde{\text{Exc}}^{(p)}(\tilde{X})) \quad (2.18)$$

Each term $A_{L_p}^{ij}(\widetilde{\text{Exc}}^{(p)}(\tilde{X}))$ represents exactly the area of the $(p+1)$ -th excursion and the total sum is the complementary second-level term mentioned above.

Let us decompose using (2.8):

$$\begin{aligned}
A_{L_p}^{ij}(\widetilde{\text{Exc}}^{(p)}(\tilde{X})) &= \sum_{1 \leq k < l \leq L_p} \left((\Delta \widetilde{\text{Exc}}^{(p)}(\tilde{X})^{(i)})_k (\Delta \widetilde{\text{Exc}}^{(p)}(\tilde{X})^{(j)})_l \right. \\
&\quad \left. - (\Delta \widetilde{\text{Exc}}^{(p)}(\tilde{X})^{(j)})_k (\Delta \widetilde{\text{Exc}}^{(p)}(\tilde{X})^{(i)})_l \right) \\
&= A_{L_p}^{ij}(\text{Exc}^{(p)}(X)) \\
&\quad + \left(\sum_{1 \leq k < l \leq L_p} \left((X_{T_p+l} - X_{T_p+l-1}) - (X_{T_p+k} - X_{T_p+k-1}) \right) \right)^{(i)} v^{(j)} \\
&\quad - v^{(i)} \left(\sum_{1 \leq k < l \leq L_p} \left((X_{T_p+l} - X_{T_p+l-1}) - (X_{T_p+k} - X_{T_p+k-1}) \right) \right)^{(j)}
\end{aligned}$$

We set

$$\begin{aligned}
\mathbf{Corr}_p^{ij}(X) &= \left(\sum_{1 \leq k < l \leq L_p} \left((X_{T_p+l} - X_{T_p+l-1}) - (X_{T_p+k} - X_{T_p+k-1}) \right) \right)^{(i)} v^{(j)} \\
&\quad - v^{(i)} \left(\sum_{1 \leq k < l \leq L_p} \left((X_{T_p+l} - X_{T_p+l-1}) - (X_{T_p+k} - X_{T_p+k-1}) \right) \right)^{(j)}
\end{aligned}$$

and we call this term the *area drift correction*. Since the increments of X are bounded by a certain $R > 0$, we deduce that

$$|\mathbf{Corr}_p^{ij}(X)| \leq K_{v,R} L_p^2$$

where $K_{v,R}$ is a constant depending on v and R . Likewise, we obtain

$$|A_{L_p}^{ij}(\text{Exc}^{(p)}(X))| \leq K'_R L_p^2$$

where K'_R is a constant depending on R . We can thus conclude that all the $A_{L_p}^{ij}(\widetilde{\text{Exc}}^{(p)}(\tilde{X}))$ are integrable. Moreover, these variables are i.i.d., since $A_{L_p}^{ij}(\text{Exc}^{(p)}(X))$ and $\mathbf{Corr}_p^{ij}(X)$ depend only on $\text{Exc}^{(p)}(X)$. Thus, by the law of large numbers the

following convergence holds:

$$\frac{1}{n} \sum_{p=0}^{n-1} A_{L_p}^{ij}(\widetilde{\text{Exc}}^{(p)}(\tilde{X})) \xrightarrow[n \rightarrow \infty]{\text{a.s.}} \mathbb{E} \left[A_{L_0}^{ij}(\widetilde{\text{Exc}}^{(0)}(\tilde{X})) \right] = \mathbb{E} \left[A_{L_0}^{ij}(\text{Exc}^{(0)}(X)) \right] + \mathbb{E} \left[\mathbf{Corr}_0^{ij}(X) \right]$$

Slutsky's theorem for metric spaces states that, for two sequences $(X_n)_n$ and $(Y_n)_n$ on a metric space (S, ρ) and such that $\rho(X_n, X) \rightarrow 0$ and $\rho(X_n, Y_n) \rightarrow 0$, then $\rho(Y_n, X) \rightarrow 0$ (see, for example, [9], theorem 3.1). Applying it to the sequences

$$\begin{aligned} X_N &= \left(\iota^{(N)} \left(\tilde{\lambda}_\bullet(\tilde{X}), A_\bullet(\tilde{X}) \right) \right)_t \otimes (0, 0, t\Gamma) \Big|_{0 \leq t \leq \tau} \\ Y_N &= \left(\iota^{(N)}(\tilde{X}_{T_\bullet}, A_{T_\bullet}(\tilde{X})) \right)_{0 \leq t \leq \tau} \end{aligned}$$

we can conclude by using the result from lemma 2.3.1 that

$$\left(\delta_{\sqrt{C^{-1}}} \iota^{(N)}(\tilde{X}_{T_\bullet}, A_{T_\bullet}(\tilde{X})) \right)_{0 \leq t \leq \tau} \xrightarrow[N \rightarrow \infty]{(d)} (B_t, \mathcal{A}_t + t\Gamma)_{0 \leq t \leq \tau}$$

where the coefficients of the $d \times d$ (with $d = \dim(E)$) matrix Γ are given by

$$\Gamma^{ij} = C^{-1} \left(\mathbb{E} \left[A_{L_0}^{ij}(\text{Exc}^{(0)}(X)) \right] + \mathbb{E} \left[\mathbf{Corr}_0^{ij}(X) \right] \right) \quad (2.19)$$

The matrix Γ is the announced area anomaly. It is immediate from definition 2.8 that $A_{L_0}^{ij}(\widetilde{\text{Exc}}^{(0)}(\tilde{X})) = -A_{L_0}^{ji}(\widetilde{\text{Exc}}^{(0)}(\tilde{X}))$, which implies that Γ is antisymmetric. \square

Lemma 2.3.3. *For any $t_1 < t_2 < \dots < t_k \in \mathbb{R}_+$, we have*

$$\begin{aligned} & \left(\delta_{\sqrt{C^{-1}\beta}} \iota^{(n)} \left(\tilde{X}_\bullet, A_\bullet(\tilde{X}) \right) \Big|_{t_1}, \dots, \delta_{\sqrt{C^{-1}\beta}} \iota^{(n)} \left(\tilde{X}_\bullet, A_\bullet(\tilde{X}) \right) \Big|_{t_k} \right) \\ & \xrightarrow[n \rightarrow \infty]{(d)} ((B_{t_1}, \mathcal{A}_{t_1} + t_1\Gamma), \dots, (B_{t_k}, \mathcal{A}_{t_k} + t_k\Gamma)) \end{aligned}$$

In this lemma, we pass from the embeddings of an extracted sequence $\left(\iota^{(N)}(\tilde{X}_{T_\bullet}, A_{T_\bullet}(\tilde{X})) \right)_{0 \leq t \leq \tau}$ to the embeddings of the full sequence $\left(\iota^{(N)}(\tilde{X}_\bullet, A_\bullet(\tilde{X})) \right)_{0 \leq t \leq \tau}$. We show that in the term $\delta_{\sqrt{n^{-1}}} \left(\tilde{X}_{[nt]}, A_{[nt]}(\tilde{X}) \right)$ the only part that counts at the limit is the one given by the excursions up to time n , i.e. $\delta_{\sqrt{n^{-1}}} \left(\tilde{X}_{\kappa([nt])}, A_{\kappa([nt])}(\tilde{X}) \right)$. At the same time, the constant β appears in the renormalization, since we have to take into consideration the approximate length of an excursion up to time $n \frac{n}{\kappa(n)}$, and β is

precisely the a.s. limit of this sequence.

Proof. Set $\tilde{\mathbb{X}}_n = (\tilde{X}_n, A_n(\tilde{X}))$. With the upper bound from proposition 2.2.1 (since it doesn't play an important role here, we suppose that $\nu = 1$), for $t \in [0, \tau]$, we get the inequality:

$$\mathbf{d} \left(\delta_{\sqrt{n-1}} \tilde{\mathbb{X}}_{T_{\kappa}(\lfloor nt \rfloor)}, \delta_{\sqrt{n-1}} \tilde{\mathbb{X}}_{\lfloor nt \rfloor} \right) \leq \frac{1}{\sqrt{n}} \left| \tilde{X}_{\lfloor nt \rfloor} - \tilde{X}_{T_{\kappa}(\lfloor nt \rfloor)} \right|_E + \frac{1}{\sqrt{n}} \left| A_{T_{\kappa}(\lfloor nt \rfloor), \lfloor nt \rfloor}(\tilde{X}) \right|_{E \otimes E}^{\frac{1}{2}}$$

We are going to use this decomposition in order to prove that $\mathbf{d} \left(\delta_{\sqrt{n-1}} \tilde{\mathbb{X}}_{T_{\kappa}(\lfloor nt \rfloor)}, \delta_{\sqrt{n-1}} \tilde{\mathbb{X}}_{\lfloor nt \rfloor} \right)$ converges in probability to 0.

We set $\tilde{R} = R + |v|_E$. First, it is easy to see that if $k' \leq k < k''$, by triangular inequality we have a.s.

$$\left| \tilde{X}_{k-k'} \right|_E \leq \sum_{l=1}^{k-k'} \left| \tilde{X}_l - \tilde{X}_{l-1} \right|_E \leq \tilde{R}(k - k') \leq \tilde{R}(k'' - k') \quad (2.20)$$

Next, since $A_n(\tilde{X})$ is a $d \times d$ matrix (with $d = \dim(E)$), we have

$$\left| A_{k-k'}(\tilde{X}) \right|_{E \otimes E} \leq \sum_{i=1}^d \max_{j=1, \dots, d} \left| A_{k-k'}^{ij}(\tilde{X}) \right| \leq d \tilde{R}^2 (k - k')^2 \leq d \tilde{R}^2 (k'' - k')^2 \quad (2.21)$$

Further on, by strong Markov property, for $\epsilon > 0$, the Chebyshev's inequality, together with (2.20), implies, for the first term,

$$\begin{aligned} \mathbb{P} \left(\frac{1}{\sqrt{n}} \left| \tilde{X}_{\lfloor nt \rfloor} - \tilde{X}_{T_{\kappa}(\lfloor nt \rfloor)} \right|_E > \epsilon \right) &\leq \frac{\mathbb{E} \left[\left| \tilde{X}_{\lfloor nt \rfloor - T_{\kappa}(\lfloor nt \rfloor)} \right|_E^2 \right]}{n \epsilon^2} \leq \frac{\tilde{R}^2 \mathbb{E} \left[(T_{\kappa(\lfloor nt \rfloor)+1} - T_{\kappa(\lfloor nt \rfloor)})^2 \right]}{n \epsilon^2} \\ &= \frac{\tilde{R}^2 \mathbb{E} [T_1^2]}{n \epsilon^2} \end{aligned}$$

and, for the second term, together with (2.21),

$$\mathbb{P} \left(\frac{1}{\sqrt{n}} \left| A_{T_{\kappa}(\lfloor nt \rfloor), \lfloor nt \rfloor}(\tilde{X}) \right|_{E \otimes E}^{\frac{1}{2}} > \epsilon \right) \leq \frac{\mathbb{E} \left[\left| A_{T_{\kappa}(\lfloor nt \rfloor), \lfloor nt \rfloor}(\tilde{X}) \right|_{E \otimes E} \right]}{n \epsilon^2} \leq \frac{d \tilde{R}^2 \mathbb{E} [T_1^2]}{n \epsilon^2}$$

Hence, taking into consideration the fact that $\mathbb{E}[T_1^2] < \infty$, we obtain

$$\begin{aligned} \mathbb{P}\left(\mathbf{d}\left(\delta_{\sqrt{n-1}}\tilde{\mathbb{X}}_{T_{\kappa(\lfloor nt \rfloor)}}, \delta_{\sqrt{n-1}}\tilde{\mathbb{X}}_{\lfloor nt \rfloor}\right) > \epsilon\right) &\leq \mathbb{P}\left(\frac{1}{\sqrt{n}}\left|\tilde{X}_{\lfloor nt \rfloor} - \tilde{X}_{T_{\kappa(\lfloor nt \rfloor)}}\right|_E > \frac{\epsilon}{2}\right) \\ &+ \mathbb{P}\left(\frac{1}{\sqrt{n}}\left|A_{T_{\kappa(\lfloor nt \rfloor)}, \lfloor nt \rfloor}(\tilde{X})\right|_{E \otimes E}^{\frac{1}{2}} > \frac{\epsilon}{2}\right) \xrightarrow{n \rightarrow \infty} 0 \end{aligned}$$

As $\kappa(n)$ is the number of excursions up to a time n of $(\pi_0(X_n))_n$, the ergodic theory tells us that $\kappa(n)/n \rightarrow 1/\beta$ a.s. as $n \rightarrow \infty$. Consequently, the above convergence in probability combined with the result from lemma 2.3.2 implies

$$\delta_{\sqrt{n-1}C^{-1}\beta}\tilde{\mathbb{X}}_{\lfloor nt \rfloor} \xrightarrow[n \rightarrow \infty]{(d)} (B_t, \mathcal{A}_t + t\Gamma)$$

What is now left to do is pass from $\tilde{\mathbb{X}}_{\lfloor nt \rfloor}$ to $\iota^{(n)}(\tilde{X}_\bullet, A_\bullet(\tilde{X}))_t$, and in order to do that we have to study the convergence of $\left|\tilde{X}_{\lfloor nt \rfloor+1} - \tilde{X}_{\lfloor nt \rfloor}\right|_E / \sqrt{n}$ and $\left|A_{\lfloor nt \rfloor+1}(\tilde{X}) - A_{\lfloor nt \rfloor}(\tilde{X})\right|_{E \otimes E} / n$.

We start with $\left|\tilde{X}_{\lfloor nt \rfloor+1} - \tilde{X}_{\lfloor nt \rfloor}\right|_E / \sqrt{n} \leq \tilde{R} / \sqrt{n} \rightarrow 0$ a.s. as $n \rightarrow \infty$. Similarly, by formula 2.14 and using the fact that $|ab - cd|/2 \leq a^2 + b^2 + c^2 + d^2$, we conclude to the following convergence in probability:

$$\begin{aligned} \frac{\left|A_{\lfloor nt \rfloor+1}(\tilde{X}) - A_{\lfloor nt \rfloor}(\tilde{X})\right|_{E \otimes E}}{n} &\leq \frac{\tilde{R}^2 + \left|\tilde{X}_{\lfloor nt \rfloor} - \tilde{X}_{T_{\kappa(\lfloor nt \rfloor)}}\right|_E^2}{n} \\ &\leq \frac{\tilde{R}^2((T_{\kappa(\lfloor nt \rfloor)+1} - T_{\kappa(\lfloor nt \rfloor)})^2 + 1)}{n} \xrightarrow[n \rightarrow \infty]{\mathcal{P}} 0 \end{aligned}$$

We conclude by Slutsky's theorem that

$$\delta_{\sqrt{C^{-1}\beta}}\iota^{(n)}(\tilde{X}_\bullet, A_\bullet(\tilde{X}))_t \xrightarrow[n \rightarrow \infty]{(d)} (B_t, \mathcal{A}_t + t\Gamma)$$

It is now easy to pass to the multivariate case. Choose $t_1 < t_2 < \dots < t_l \in \mathbb{R}_+$. Then we have immediately

$$\begin{aligned} \mathbb{P}\left(\mathbf{d}_l\left((\delta_{\sqrt{n-1}}\tilde{\mathbb{X}}_{T_{\kappa(\lfloor nt_1 \rfloor)}}, \dots, \delta_{\sqrt{n-1}}\tilde{\mathbb{X}}_{T_{\kappa(\lfloor nt_l \rfloor)}}), (\delta_{\sqrt{n-1}}\tilde{\mathbb{X}}_{\lfloor nt_1 \rfloor}, \dots, \delta_{\sqrt{n-1}}\tilde{\mathbb{X}}_{\lfloor nt_l \rfloor})\right) > \epsilon\right) \\ \leq \sum_{i=1}^l \mathbb{P}\left(\mathbf{d}\left(\delta_{\sqrt{n-1}}\tilde{\mathbb{X}}_{\lfloor nt_i \rfloor}, \delta_{\sqrt{n-1}}\tilde{\mathbb{X}}_{T_{\kappa(\lfloor nt_i \rfloor)}}\right) > \frac{\epsilon}{l}\right) \xrightarrow{n \rightarrow \infty} 0 \end{aligned}$$

Applying once again the result from lemma 2.3.2, we obtain

$$\left(\delta_{\sqrt{n^{-1}C^{-1}\beta}} \tilde{\mathbb{X}}_{t_1}, \dots, \delta_{\sqrt{n^{-1}C^{-1}\beta}} \tilde{\mathbb{X}}_{t_k} \right) \xrightarrow[n \rightarrow \infty]{(d)} ((B_{t_1}, \mathcal{A}_{t_1} + t_1 \Gamma), \dots, (B_{t_k}, \mathcal{A}_{t_k} + t_k \Gamma))$$

and we conclude by applying Slutsky's theorem as in the univariate case. \square

Lemma 2.3.4. *The sequence $(\iota^{(n)}(\tilde{X}_\bullet, A_\bullet(\tilde{X})))_{n \geq 0}$ is tight in α -Hölder topology for $\alpha < 1/2$.*

Proof. As in [12], we apply here the Kolmogorov's criterion. In order to do so, it will be enough to prove that, for $\tau > 0$ fixed, for any $p > 1$ there exists a positive constant c_p such that, for all $0 \leq s < t \leq \tau$,

$$\sup_n \mathbb{E} \left[\mathbf{d} \left(\iota^{(n)}(\tilde{X}_\bullet, A_\bullet(\tilde{X}))_t, \iota^{(n)}(\tilde{X}_\bullet, A_\bullet(\tilde{X}))_s \right)^{4p} \right] \leq c_p |t - s|^{2p-1}$$

since $(2p - 1)/(4p) \rightarrow 1/2^-$ as $p \rightarrow \infty$.

Choose $a > 0$. By proposition 2.2.1 and applying to $(X_n)_{n \in \mathbb{N}}$ the Markov property, we get:

$$\begin{aligned} \mathbb{E} \left[\mathbf{d} \left(\iota^{(n)}(\tilde{X}_\bullet, A_\bullet(\tilde{X}))_t, \iota^{(n)}(\tilde{X}_\bullet, A_\bullet(\tilde{X}))_s \right)^a \right] &= \mathbb{E} \left[\left\| \iota^{(n)}(\tilde{X}_\bullet, A_\bullet(\tilde{X}))_{s,t} \right\|^a \right] \\ &= \mathbb{E} \left[\mathbb{E} \left[\left\| \iota^{(n)}(\tilde{X}_\bullet, A_\bullet(\tilde{X}))_{s,t} \right\|^a \middle| \tilde{X}_{\lfloor ns \rfloor} \right] \right] = \mathbb{E} \left[\mathbb{E}_{\tilde{X}_{\lfloor ns \rfloor}} \left[\left\| \iota^{(n)}(\tilde{X}_\bullet, A_\bullet(\tilde{X}))_{t-s} \right\|^a \right] \right] \\ &= \mathbb{E} \left[\left\| \iota^{(n)}(\tilde{X}_\bullet, A_\bullet(\tilde{X}))_{t-s} \right\|^a \right] \end{aligned}$$

Since $(\iota^{(n)}(\tilde{X}_\bullet, A_\bullet(\tilde{X}))_t)_{0 \leq t \leq \tau}$ is constructed by linear connections between the points $\iota^{(n)}(\tilde{X}_\bullet, A_\bullet(\tilde{X}))_{k/n}$ for $k = 0, \dots, \lfloor n\tau \rfloor$, the properties of geodesic interpolation imply that it is sufficient to prove that

$$\frac{1}{n^{2p}} \mathbb{E} \left[\left\| \tilde{\mathbb{X}}_k \right\|^{4p} \right] \leq c_p \left(\frac{k}{n} \right)^{2p}$$

for $k = 0, \dots, \lfloor n\tau \rfloor$, uniformly over $n \geq 1$. As in [12], this follows immediately if we prove, for all $p > 1$,

$$\mathbb{E} \left[\left\| \tilde{\mathbb{X}}_n \right\|^{4p} \right] = O(n^{2p})$$

Here, Chen's relation (formula (2.13)) gives

$$\tilde{\mathbb{X}}_n = \tilde{\mathbb{X}}_{T_{\kappa(n)}} \otimes_2 \tilde{\mathbb{X}}_{T_{\kappa(n)},n}$$

where \otimes_2 is the product on $G^2(E)$ from section 2.2.4 (it can also be interpreted as a path concatenation operator). As mentioned in section 2.2.4, the norm $\|\cdot\|$ is sub-additive. Using strong Markov property and the inequality:

$$\forall a, b \geq 0, \quad (a + b)^p \leq 2^p(a^p + b^p), \quad (2.22)$$

we arrive to an initial upper bound:

$$\mathbb{E} \left[\|\tilde{\mathbb{X}}_n\|^{4p} \right] \leq 2^{4p} \left(\mathbb{E} \left[\|\tilde{\mathbb{X}}_{T_{\kappa(n)}}\|^{4p} \right] + \mathbb{E} \left[\|\tilde{\mathbb{X}}_{T_{\kappa(n)},n}\|^{4p} \right] \right)$$

On one hand, as $\kappa(n) \leq n$ a.s., we have

$$\mathbb{E} \left[\|\tilde{\mathbb{X}}_{T_{\kappa(n)}}\|^{4p} \right] \leq \max_{l=1,\dots,n} \mathbb{E} \left[\|\tilde{\mathbb{X}}_{T_l}\|^{4p} \right] = O(n^{2p})$$

since $\tilde{\mathbb{X}}_{T_l}$ is a product of l centred i.i.d. variables ($\tilde{\mathbb{X}}_{T_l} = \tilde{\mathbb{X}}_{T_1} \otimes_2 \tilde{\mathbb{X}}_{T_1,T_2} \otimes_2 \dots \otimes_2 \tilde{\mathbb{X}}_{T_{l-1},T_l}$), and therefore $\mathbb{E} \left[\|\tilde{\mathbb{X}}_{T_l}\|^{4p} \right] = O(l^{2p})$ as it was proved in [12].

On the other hand, proposition 2.2.1 (with the convention $\nu = 1$), and the inequality (2.22), together with the upper bounds from (2.20) and (2.21), give

$$\mathbb{E} \left[\|\tilde{\mathbb{X}}_{T_{\kappa(n)},n}\|^{4p} \right] \leq 2^{4p} \tilde{R}^{4p} (d^{2p} + 1) \mathbb{E} \left[(T_{\kappa(n)+1} - T_{\kappa(n)})^{4p} \right] = 2^{4p} \tilde{R}^{4p} (d^{2p} + 1) \mathbb{E} \left[T_1^{4p} \right]$$

We therefore obtain

$$\mathbb{E} \left[\|\tilde{\mathbb{X}}_n\|^{4p} \right] \leq 2^{4p} \left(O(n^{2p}) + 2^{4p} \tilde{R}^{4p} (d^{2p} + 1) \mathbb{E} \left[T_1^{4p} \right] \right) = O(n^{2p})$$

which achieves the proof. □

□

2.3.2 Properties of the area anomaly

Let us briefly discuss the formula of Γ given by (2.19). The main term, the one that we concentrate on, is given by $\mathbb{E} \left[A_{L_0}^{ij}(\text{Exc}^{(0)}(X)) \right]$, which is the expectation of the stochastic area of an excursion. This is what we were intuitively expecting: the oscillations of the process along an excursion are not visible at the limit in the uniform topology but they generate stochastic area that influences the second level of a rough path through a drift.

The second term, $\mathbb{E} \left[\mathbf{Corr}_0^{ij}(X) \right]$, comes from the fact that the excursions are not necessarily centered. It can be seen as a trace on the second level of the rough path of the fact that the excursions have been re-centered. Of course, if the excursions have zero mean from the beginning as in example 2.1.3, this correction term is zero.

These remarks imply that the area anomaly of a Markov chain on periodic graphs *depends on the drift and on the stochastic area of a pseudo-excursion*. Let us see how this is different from the area drift generated exclusively by the drift of the process, as in [6]. Notice that a deterministic drift appears at level 2 in corollary 3.4 but it is not an area anomaly as it depends entirely on the drift we assign to the Brownian motion $(B_t)_{t \geq 0}$ (and consequently to W). However, in this case, we recover supplementary terms on the second level, and we do not have the area+drift scheme anymore. It is important that this corollary allows us to consider the convergence of drifted processes and their stochastic areas without centering them, which is indispensable in the Donsker-type theorems.

The presence of area anomaly in the limit stochastic area is the reason why the Itô map sometimes fails to be continuous in the uniform topology, and thus, in order to correctly approach an SDE driven by the Brownian motion, we need to be sure to get a zero area anomaly, as it has already been mentioned in the introduction. A discussion on this topic, as well as a method of approaching the Brownian motion and its Lévy area through the rough path of a cubature formula on Wiener space can be found in [6].

As it has already been mentioned, the area anomaly is zero when the process is symmetric/reversible, in particular when we consider the sum of centered i.i.d.r.v. as in the classical Donsker setting. We can use this fact to construct a sequence of processes which generate area anomaly "artificially": we start by a process with

piecewise linear i.i.d. centered increments and we replace every increment by a path of bounded variation and a stochastic area with a non-zero mean. For example, in the case of the sequence constructed by concatenating i.i.d. copies of the cubature formula on Wiener space from [6], we concatenate to every copy of the cubature a centered random variable C_n from $\mathcal{C}^{1-var}([0, 1], \mathbb{R}^d)$ such that the C_n s are i.i.d. and their stochastic areas are of non-zero mean Γ . In this case, by the law of large numbers, the area anomaly is equal to Γ (modulo some renormalization constant).

2.4 Applications and open questions

2.4.1 Application to an SDE

Stochastic differential equations may arise as limits of discrete difference equations. We consider here the simple case of a two-dimensional process $(X_n, Y_n)_{n \in \mathbb{N}}$ and the difference equation:

$$U_{n+1} - U_n = \epsilon[f(U_n)(X_{n+1} - X_n) + g(U_n)(Y_{n+1} - Y_n)] \quad (2.23)$$

where (U_n) is an \mathbb{R} -valued process. If $(X_n, Y_n)_n$ is a random walk converging towards a standard Brownian motion in \mathbb{R}^2 under Donsker's embedding and if ϵ varies as $\epsilon = 1/\sqrt{N}$ where N is the parameter of the Donsker embedding, then U_n converges to the solution of the SDE

$$dU_t = f(U_t)dB_t^1 + g(U_t)dB_t^2 \quad (2.24)$$

with the Itô prescription.

We may now substitute a Markov chain on a periodic graph, like the ones described previously, to the random walk. Now, the suitable framework is rough path theory with an area anomaly. If we choose the process from example 2.1.3, an easy computation inspired from the proof of theorem 2.1.1, which consists in dividing the process into excursions of length 4, shows that the limit process solves

the SDE:

$$dU_t = f(U_t)dB_t^1 + g(U_t)dB_t^2 \quad (2.25a)$$

$$+ \frac{1}{2}[f'(U_t)f(U_t) + g'(U_t)g(U_t)]Kdt \quad (2.25b)$$

$$+ \frac{1}{2}[f'(U_t)g(U_t) - f(U_t)g'(U_t)]\gamma dt \quad (2.25c)$$

where γ is the area anomaly (2.3) and K is the variance of the variables $\lambda_k = X_{4k} - X_{4(k-1)}$. The term (2.25b) is a well-known term in classical stochastic calculus similar to the Itô/Stratonovitch correction. The term (2.25c) is new and requires the area anomaly.

As explained in [34], this behavior is indeed general and generalizes to any Markov chain on a periodic graph satisfying theorem 2.1.1, *mutatis mutandis*. Here again, one notices that both terms (2.25b) and (2.25c) are produced by the coarse-graining procedure based on excursions which leads to non-trivial renormalization terms.

2.4.2 A three dimensional model with a non-trivial area anomaly

We extend the model presented in the introduction to dimension three. This extension is interesting for two main reasons: no particular role is played by the roots of unity as in the introduction and we may choose arbitrary jump rates; moreover, the area anomaly is now an antisymmetric three-by-three matrix which can be arbitrary. Such a process can then be used to obtain a Brownian motion on $SU(2)$ with an area anomaly with the classical identification between $\mathfrak{su}(2)$ and \mathbb{R}^3 and solving the rough differential equation $dU_t = U_t dB_t$.

The graph G is \mathbb{Z}^3 , the lattice Λ is $(2\mathbb{Z})^3$ and the fundamental domain is thus $G_0 = (\mathbb{Z}/2\mathbb{Z})^3$. The only jumps allowed are those between x and $x \pm e_k$ where e_k is one of the three vectors of the canonical basis. The coefficients $Q(x, x')$ depend only on the classes modulo 2 of each coordinate of x and x' . A jump $\pm e_k$ changes the modulo class by 1 on the coordinate k . Once projected onto G_0 , the two jumps $x \pm e_k$ give the same transition on the cube. Q is then parametrized by 8×6

Origin in G	Proj. on G_0	$+e_1$	$-e_1$	$+e_2$	$-e_2$	$+e_3$	$-e_3$
$(2k, 2l, 2m)$	$(0, 0, 0)$	$u/2$	$v/2$	$u/2$	$v/2$	0	0
$(2k + 1, 2l, 2m)$	$(1, 0, 0)$	0	0	$u/2$	$v/2$	$u/2$	$v/2$
$(2k, 2l + 1, 2m)$	$(0, 1, 0)$	$u/3$	$v/3$	$v/3$	$u/3$	$u/3$	$v/3$
$(2k + 1, 2l + 1, 2m)$	$(1, 1, 0)$	$v/2$	$u/2$	0	0	$u/2$	$v/2$
$(2k, 2l, 2m + 1)$	$(0, 0, 1)$	$u/2$	$v/2$	0	0	$v/2$	$u/2$
$(2k + 1, 2l, 2m + 1)$	$(1, 0, 1)$	$v/3$	$v/3$	$u/3$	$v/3$	$v/3$	$u/3$
$(2k, 2l + 1, 2m + 1)$	$(0, 1, 1)$	0	0	$v/2$	$u/2$	$v/2$	$u/2$
$(2k + 1, 2l + 1, 2m + 1)$	$(1, 1, 1)$	$v/2$	$u/2$	$v/2$	$u/2$	0	0

Table 2.1: Parameter of the dynamics of the cubic model used for the numerical results. The numerical simulations are made for $u = 9/10$ and $v = 1 - u = 1/10$ in order to bias the Markov chain to stay in a cube, so that it can develop a non-zero area anomaly Γ .

parameters (the cardinal of Q_0 times the number of directions).

In order to kill in a natural way the asymptotic drift v of the process, we assume a central symmetry such that $Q(x, x \pm e_k) = Q(x + (1, 1, 1), x + (1, 1, 1) \mp e_k)$. The model is thus parametrized by $24 = 8 \cdot 6/2$ parameters. In the generic case, the area anomaly Γ is non-zero.

Simulations are made for parameters chosen as in figure 2.1 with $u = 9/10$ and $v = 1/10$ for over $64 \cdot 10^6$ simulations and the process is observed at time $n = 40000$. We obtain the following values for both coordinates of $X_n^{(i)}/\sqrt{n}$:

- the empirical means are -0.0025 , -0.0020 and -0.0025 .
- the empirical covariance matrix has three coefficients 0.03001 on the diagonal and the other coefficients are all below 10^{-8} .
- the empirical third cumulants are all three below 10^{-6}
- the kurtosis are all three 3.0007 , 3.0006 and 3.0004 .

The empirical values for the area $A_n^{ij}/(\sigma_i \sigma_j n)$ normalized by the empirical standard deviations of the coordinates are:

- empirical mean $\Gamma^{12} = 1.500$, $\Gamma^{23} = 1.500$ and $\Gamma^{31} = -1.500$ (the area anomalies),

- empirical standard deviations are 0.5011, 0.5011 and 0.5011.
- empirical third cumulants $-1.46 \cdot 10^{-4}$, $-6.9 \cdot 10^{-5}$ and $1.1 \cdot 10^{-4}$
- empirical fourth cumulants 0.12533, 0.12535 and 0.12532.

All the cumulants correspond to a normal law for the coordinates and a Lévy drifted area for the area process.

2.4.3 Application to a stochastic differential equation on $SU(2)$

The previous example can be used to study the effect of the area anomaly on an SDE. We identify \mathbb{R}^3 with $\mathfrak{su}(2)$ through the correspondence $(x, y, z) \rightarrow x\sigma_1 + y\sigma_2 + z\sigma_3$ where the $\sigma_1, \sigma_2, \sigma_3$ are the three Pauli matrices. We may now build the process with values in $SU(2)$ defined by:

$$U_{n+1}^\epsilon = U_n^\epsilon \frac{I + \epsilon i(X_{n+1} - X_n)}{I - \epsilon i(X_{n+1} - X_n)} \quad (2.26)$$

If $(X_n)_n$ is a standard random walk on $\mathfrak{su}(2)$, the process $(X_n)_n$ converges in law after rescaling to the Brownian motion on $\mathfrak{su}(2)$ by Donsker's theorem and U_n^ϵ with the correct scaling of ϵ converges in law to the canonical Brownian motion on $SU(2)$.

If we replace the random walk by the process $(X_n)_n$ from the previous section, X_n converges in law after rescaling to the Brownian motion on $\mathfrak{su}(2)$ with an area anomaly $\Gamma \in \mathfrak{su}(2) \wedge \mathfrak{su}(2)$. Thus U_n^ϵ converges now to the solution of the SDE $dU_t = U_t \circ dB_t$ which *must* be interpreted in the rough path sense since traditional stochastic calculus cannot take Γ into account.

2.4.4 Open questions

The present result leads to some open questions both about the limit process with the area anomaly and about the discrete models which may converge to such a limit.

It would be interesting to understand how the area anomaly fits in the Fock space description of Brownian Motion: the question is non-trivial because the Lévy area belongs to the second chaos and the presence of an area anomaly adds a zero-chaos component to the Lévy area.

Next, two-dimensional Brownian motion is known to exhibit conformal symmetry and it is natural to ask how our limit process behaves under conformal transformations.

Since we focus on the area drift, an important question that arises is: can Girsanov's theorem be extended to cancel the area anomaly by a change of measure on the rough path space?

In the introduction, it has been mentioned that other models than those we study, like the ones in a random environment from [5] or the ones on a lattice graph from [48], might generate a non-zero limit drift on the second level of the rough path. Moreover, there is a detail indicating that this drift might be the analogue of our area anomaly: the drift we remove from each of them before studying the convergence of the process is the analogue of the drift nv from the present article. It would thus be interesting to see if we can get a convergence result for the models of these two articles in rough path topology and what the stochastic area limit drift looks like in this case.

Our proof exhibits striking similarities with [3]. Our "internal" G_0 space seems to play the same role as the compact sphere in their paper and their proof also uses theory of rough paths to control convergences. We would like to know if one can build models on Riemannian manifolds which may exhibit area anomalies. A good hint might be given by the construction from [48].

In [2], the main result from [1] is generalized to sub-Laplacians with drift (i.e. we are here in a continuous setting). Just like in [1], the author gets a Berry-Esseen-like estimate of the heat kernel. The drift of the Laplacian is not automatically linked with the area anomaly and is considered here to be more like an additional problem than a central object that should be studied. We may use these results for studying our limit motion, i.e. the Brownian motion rough path with area anomaly.

We can also ask ourselves whether the present paper can lead to generalizing or improving the result from [6].

Finally, Brownian motion belongs to the larger family of Lévy processes and,

consequently, we may wonder whether Lévy processes may be also enhanced with area components and approximated by suitable discrete processes.

Chapter 3

Area anomaly and generalized drift of iterated sums for hidden Markov walks

Ce chapitre est notre article [63] écrit en collaboration avec D. Simon tel qu'il a été mis en ligne sur HAL (<https://hal.archives-ouvertes.fr/hal-01586794>), excepté des changements mineurs comme l'absence d'abstract ou le changement de numérotation pour les références.

3.1 Introduction

3.1.1 On the importance of discrete models in the rough paths theory.

From continuous to discrete setting. Rough paths theory was introduced by T. Lyons in 1998 (see, for example, [65]) in order to provide a deterministic setting to stochastic differential equations (SDEs) of the type

$$dy_s = f(y_s)[dx_s] \tag{3.1}$$

where $(y_s)_s$ is a path in $\mathbb{R}^{d'}$, $(x_s)_s$ is a path in \mathbb{R}^d of Hölder regularity $\alpha < 1$ (which is often the case for stochastic processes) and $f : \mathbb{R}^{d'} \rightarrow \text{End}(\mathbb{R}^d, \mathbb{R}^{d'})$. Whenever the classical Young integration [82] fails (which is the case for $\alpha < 1/2$), paths may be lifted (in a non-unique way) to a larger, more abstract, space, *the space of rough paths*, for which existence and uniqueness of a solution map become easier to prove.

The idea behind this theory relies on the Taylor expansion: if we want to suitably approach (3.1), the trajectory level is not enough to register all the relevant information when passing to the limit. Thus, we have first to make an "expansion" of the path $(x_t)_t$ by constructing, on the grounds of certain rules, the associated *rough path* $(\mathbf{x}_t)_t$, which is a continuous path in $V \oplus \dots \oplus V^{[1/\alpha]}$ and then to suitably approach it.

A particularity of the rough paths theory is that it has been developed in a continuous setting (we approximate continuous processes by other continuous processes that are smoother). Since the main goal of the rough paths theory is giving a general method for solving SDEs, the exploration of the discrete setting may seem irrelevant other than for SDE approximation (see [23]). And yet, new directions of research can be found in this domain as it appears, for example, in [62], where the authors study discrete rough integrals. We believe that developing a theory of discrete models in the rough path setting can enrich the classical stochastic calculus with new tools and results. The present paper concentrates on two issues related to this field: the ways of constructing a rough path out of a discrete model and the explicit construction of new objects arising at the limit in rough path topology, as well as the role they may play in classical stochastic calculus.

Analysing rough paths issues through discrete models. A finite-variation path $\gamma : [0, T] \rightarrow V$ can be canonically lifted to a rough path using the sequence of iterated integrals:

$$S_{\gamma,n}(t) = \int_{0 < s_1 < s_2 < \dots < s_n < t} d\gamma(s_1) \otimes d\gamma(s_2) \otimes \dots \otimes d\gamma(s_n) \in V^{\otimes n}, \quad t \in [0, T] \quad (3.2)$$

For any $N \geq 2$, the corresponding canonical lift coincides with *the step- N signature* of γ given by

$$S_N(\gamma)_{0,t} = (1, S_{\gamma,1}(t), \dots, S_{\gamma,N}(t)) \in T_1^{(N)}(V), \quad t \in [0, T] \quad (3.3)$$

where $T_1^{(N)}(V) = \{(a_0, \dots, a_N \in \bigoplus_{k=0}^N V^{\otimes k} | a_0 = 1\}$ with the convention $V^{\otimes 0} = \mathbb{R}$.

For α -Hölder paths with $\alpha < 1$, such integrals may not be well-defined and some of the values of the vectors $S_{\gamma,n}(t)$ may be *postulated* provided that they satisfy certain algebraic relations and some topological properties. The absence of a canonical lift of Hölder paths to a larger space leads to some natural questions at the basis of the present paper.

The first one, answered since the very beginning of rough paths theory, deals with how classical stochastic integrals (Itô and Stratonovitch) fit within this theory. The set of rough paths is separated in two categories, one for each type of integration:

- *geometric rough paths*, whose components satisfy the chain rule, for Stratonovich integration;
- *non-geometric rough paths*, whose components do not satisfy it, for Itô integration.

Based on discrete sequences we give explicit constructions for both types of integrals as limit processes and discuss the link with geometric and non-geometric rough paths. In particular, we will see how discrete constructions can give rise to non-geometric rough paths as limit processes.

The second issue consists in understanding the nature of the non-trivial objects in the space of rough paths; for example, two-dimensional area bubbles [60] obtained as the limit of the paths:

$$\gamma^{(n)}(t) = \left(\frac{1}{\sqrt{n}} \cos(nt), \frac{1}{\sqrt{n}} \sin(nt) \right)$$

have a well-defined rough path limit different from the constant path $\gamma(t) = (0, 0)$ (which would be the limit in the uniform topology). In this case, $S_{\gamma,1}(t) = 0$ but $S_{\gamma,2}(t) = \frac{1}{2}te_1 \wedge e_2$, with e_1 and e_2 vectors of the canonical basis of \mathbb{R}^2 .

The third (and central) question is to ask whether such non-trivial limits can coexist with the classical structure of stochastic integrals. We can give a positive answer to this question by studying well-chosen discrete models with suitable time correlations. In [64], we have constructed a class of processes on a finite-dimensional vector space V (Markov chains on periodic graphs) that converge, under certain conditions, in rough path topology (see section 3.3.2 for details) to

$$(S_{B,1}(t), S_{B,2}(t) + \Gamma t)_{t \in [0,1]} \quad (3.4)$$

where $(B_t)_t$ is a standard Brownian motion on the vector space V and Γ a deterministic antisymmetric matrix called *the area anomaly*. While $S_{B,1}(t)$ and $S_{B,2}(t)$ are obtained with classical stochastic calculus, Γ is a new object.

The present paper provides a much more general construction of such limits. The main idea relies on the following: if one wants to be able to control Γ in a limit of the type (3.4), one has to modify the underlying discrete models in the following way:

- one may try to create area-bubble-like corrections to the trajectories;
- these corrections have to remain small so that $S_{B,1}$ does not change;
- the corrected trajectories have to satisfy the Markov property.

Our solution consists in

- introducing time-correlations in the discrete processes, tuned so that they affect only $S_{B,2}$ after renormalization, through the framework of *the hidden Markov walks*, of which the Markov chains on periodic graphs are an example, and
- showing how the area anomaly depends on the way we *embed* the discrete paths in the set of continuous paths.

This last point brings into light the difference played by embeddings in the uniform and rough paths topologies. In the first case, it is merely a way to make sense of the convergence of a discrete path and one uses usually linear interpolation by

default. In the second case, it is a mean to influence and even change the nature of the limit through the non-trivial object which is the area anomaly.

We also show how the convergence of any embedding can be studied through the convergence of a discrete sequence in $T_1^{(2)}(V)$ (more precisely, in one of its sub-groups: $G^2(V)$). This allows, in particular, to make a classification of the embeddings of a given discrete process based on the Γ they provide at the limit.

Another important question concerns the combinatorial structures that can be derived from rough paths built out of hidden Markov walks. While the algebraic structure of rough paths, based on *shuffle products*, provides nice combinatorial interpretations of iterated integrals, the iterated sums have more intricate multiplication rules, which can be interpreted using *the quasi-shuffle products*. This subject will be discussed in section 3.4.

3.1.2 Structure of the article and notations.

General structure of the article. The present article is divided into five main sections. The first two sections give a general presentation of our work as follows:

- in the present section 3.1, we describe the motivations and the main goals of our work;
- in section 3.2, we state our main results and give some additional comments and illustrations to them.

The next two sections are more technical:

- section 3.3 is dedicated to the convergence of rough paths constructed out of hidden Markov walks: we give a detailed description of the framework and a proof of our main theorem (3.3.4);
- in section 3.4, we explore the combinatorial properties of hidden Markov chains and the associated rough paths through iterated occupation times and their link with shuffle and quasi-shuffle products.

The last section 3.5 is dedicated to further research perspectives and open questions.

Acknowledgements. D. S. is partially funded by the Grant ANR-14CE25-0014 (ANR GRAAL). D.S. and O.L. thank Lorenzo Zambotti for the stimulating discussions and useful suggestions.

Notations. Throughout this paper, the following objects will often appear.

- $\mathcal{C}^{1-var}([0, T], B)$ is the space of functions valued in a Banach space B with bounded variations.
- V denotes a finite-dimensional vector space. We endow V with the euclidean norm $|\cdot|_V$ and $V \otimes V$ with the induced matrix norm: $|A|_{V \otimes V} = \sup_{|x|_V=1} |Ax|_E$.
- E denotes a finite state space (for Markov chains).
- $T_1^{(k)}(V)$ is the graded vector space:

$$T_1^{(k)}(V) = V \oplus \dots \oplus V^{\otimes k}$$

and $G^2(V)$ a subgroup of $T_1^{(2)}(V)$ defined by:

$$\forall (a, b) \in G^2(V), \text{Sym}(b) = \frac{1}{2}(a \otimes a) \quad (3.5)$$

(the symmetrical part of the level $V \otimes V$ depends entirely on the level V of the object).

- Instead of $(1, a_1, \dots, a_k) \in T_1^{(k)}(V)$, we will write $(a_1, \dots, a_k) \in \bigoplus_{l=1}^k V^{\otimes l}$.
- The operator " \otimes " can denote:
 - a tensor product of two elements in V (in the formula $V \otimes V$, for example);
 - an operation on $T_1^{(2)}(V)$: $(a, b) \otimes (a', b') = (a + a', b + b' + a \otimes a')$, where the operation $a \otimes a'$ is in the sense of a tensor product.
- The operator " \wedge " is an antisymmetric law on $G^2(V)$: $(a, b) \wedge (a', b') = (a + a', b + b' + \frac{1}{2}(a \otimes a' - a' \otimes a))$;
- δ_ϵ : the standard dilatation operator on $G^2(V)$: $\delta_\epsilon(a, b) = (\epsilon a, \epsilon^2 b)$.

3.2 Main results and map of the paper

3.2.1 Rough paths models from microscopic models.

Main theorems: rough paths from hidden Markov walks.

Definition of hidden Markov walks. The Donsker theorem, generalized to rough paths theory in [12], shows that random walks do not leave any room for area anomaly: area bubbles are killed by the independence between random variables at each time step. Following [64], we introduce the framework of *hidden Markov walks* (HMW) whose time-correlations, as it has already been mentioned, allow additional area accumulation on the second level of the corresponding rough paths.

We first give a definition of *hidden Markov chains* (HMC), which are the "bricks" from which hidden Markov walks are constructed:

Definition 3.2.1 (the hidden Markov chain). *Let E be a finite state space and V a finite-dimensional vector space. A hidden Markov chain is a discrete process $(R_n, F_n)_n$ on $E \times V$ such that $R = (R_n)_n$ is a Markov chain on E and under the conditional law $\mathbb{P}(\cdot | \sigma(R))$ the variables F_n are independent and*

$$\forall k \geq 1, \forall u \in E, \mathbb{P}(F_k | R_k = u) = \mathbb{P}(F_1 | R_1 = u)$$

Hidden Markov walks are constructed out of hidden Markov chains just as simple random walks can be constructed as sums of i.i.d. variables:

Definition 3.2.2 (the hidden Markov walk). *Let E be a finite state space and V a finite-dimensional vector space and $(R_n, F_n)_n$ a hidden Markov chain on $E \times V$. For all $n \geq 1$, set $X_n = \sum_{k=1}^n F_k$. The discrete process $(R_n, X_n)_n$ on $E \times V$ is called a hidden Markov walk.*

Remarks.

- Throughout the paper, the transition matrix of the process (R_n) will be written Q and the conditional law of $F_k = X_k - X_{k-1}$ knowing $R_k = u \in E$ will be written $\nu(\bullet | u)$ and expectations under this law $\mathbb{E}_\nu[\bullet | u]$.

- The process $(X_n)_n$ may take values in a more general space than V (and we will encounter such examples throughout the paper). The only thing that changes in this case is the definition of the increments of the process (see property 3.3.1).

A Donsker-type result. The following theorem is a result on convergence of rough paths constructed by piecewise linear interpolation (like in the classical Donsker theorem) out of HMW. It is a generalization of our result from [64] to a larger class of processes.

We define the excursion times of the process $(R_k)_k$ as:

$$T_0 = 0 \tag{3.6}$$

$$\forall n \geq 1, T_n = \inf\{k > T_{n-1} : R_k = R_0\} \tag{3.7}$$

Theorem 3.2.1. *Let $(R_n, X_n)_n$ be a hidden Markov walk on $E \times V$ such that $(R_n)_n$ is irreducible and the increments of $(X_n)_n$ are a.s. uniformly bounded, i.e.*

$$\exists K > 0, \forall j \in \mathbb{N}, |F_j|_V \leq K \text{ a.s.} \tag{3.8}$$

Set $\beta = \mathbb{E}[T_1]$. Under the conditions $\mathbb{E}[X_{T_1}] = 0$ and $\mathbb{E}[X_{T_1}^{\otimes 2}] = CI_d$ for a certain $C > 0$, the rough path of the piecewise linear interpolation of $(X_k)_{k \leq N}$, renormalized by $\delta_{(\beta^{-1}NC)^{-1/2}}$, converges in the rough path topology $\mathcal{C}^\alpha([0, 1], G^2(V))$ for $\alpha < 1/2$ to a rough path whose levels are given by:

$$\begin{aligned} S_{B,1}(t) &= B_t \\ S_{B,2}(t) &= \int_{0 < s_1 < s_2 < t} \circ dB_{s_1} \otimes \circ dB_{s_2} + \Gamma t \end{aligned}$$

where $(B_t)_t$ is a standard d -dimensional Brownian motion and Γ is a deterministic antisymmetric matrix, the area anomaly, given by

$$\Gamma = C^{-1} \frac{1}{2} \mathbb{E} \left[\sum_{1 \leq k_1 < k_2 \leq T_1} F_{k_1} \otimes F_{k_2} - F_{k_2} \otimes F_{k_1} \right] \tag{3.9}$$

Remarks.

- The condition (3.8) requires a.s. uniform bound for F_j s, whereas the classical Donsker theorem (and the its rough path variant from [12]) requires a uniform bound for the moments. On one hand, most of the concrete models we can apply theorem 3.2.1 to satisfy (3.8); on the other hand, as we will see further on, the variables whose role is equivalent to the ones of i.i.d.r.v. from the classical Donsker theorem are, in our case, the pseudo-excursions.
- The definition of β implies that it may depend on the initial law of $(R_n)_n$, whereas the limit does not. The reason for this is simple: β registered the slowing down in time of the limit process which, as we will see, depends on the length of an excursion of $(R_n)_n$.
- We can actually always suppose the covariance matrix of $\mathbb{X}_{T_1}^{(1)}$ (the first component of $\mathbb{X}_{T_1} = (\mathbb{X}_{T_1}^{(1)}, \mathbb{X}_{T_1}^{(2)})$) to be equal to CI_d for a certain $C > 0$ in theorem 3.2.3 (idem for the covariance matrix of X_{T_1} in theorems 3.2.1 and 3.2.2). The reasons for this are similar to those explained in the remark that follows immediately corollary 2.1 from [64]: we can embed the process $\mathbb{X}_{T_1}^{(1)}$ in a smaller space than V , and afterwards use a linear transformation to get the identity matrix modulo a constant.
- We can still get a result similar to theorem 3.2.1 if the covariance matrix of X_{T_1} is not diagonal. The difference is that $(B_t)_t$ will have a covariance matrix depending on X_{T_1} and will not be a standard Brownian motion anymore (see proposition 3.2.1).
- The second-level drift is actually the stochastic area of an excursion modulo a multiplicative constant, i.e. it is built as the area anomaly from [64].
- $S_{B,2}(t)$ and can be decomposed into a symmetric part, given by $1/2(B_t \otimes B_t)$, and an antisymmetric part, given by $\mathcal{A}_t + \Gamma t$, where $(\mathcal{A}_t)_t$ is the Lévy area of $(B_t)_t$, i.e.

$$\mathcal{A}_t = \frac{1}{2} \int_{0 < s_1 < s_2 < t} dB_{s_1} \otimes dB_{s_2} - dB_{s_2} \otimes dB_{s_1}$$

This decomposition highlights the fact that $(S_{B,1}(t), S_{B,2}(t))_t$ is a geometric rough path. Note that the Lévy area does not depend on the choice of integration (Itô or Stratonovich). If we endow the space of rough paths

with the antisymmetric tensor product \wedge instead of the ordinary one \otimes , the $S_{B,2}(t)$ will only keep track of the antisymmetric part of the limit: this is the approach we have adopted in our previous article [64].

Examples to which theorem 3.2.1 applies.

- In case of a simple random walk ($E = \{0\}$) with a piecewise linear embedding, we get the Donsker theorem for rough paths from [12] (in particular, $\Gamma = 0$), with slightly different moment conditions.
- A less trivial example is *the sum of rotating Bernoulli i.i.d.r.v.* which has been presented in our previous article [64]: $X_n = \sum_{k=1}^n i^k U_k$, where the U_k s are i.i.d. Bernoulli variables. In this case, $E = \{1, 2, 3, 4\}$ and $\Gamma \neq 0$.
- More generally, theorem 3.2.1 applies to any Markov chain on periodic graphs as the ones described in [64].

The question of embeddings. If we want to build a rough path associated to a sequence $(x_n)_n \in V^{\mathbb{N}}$, we first need to associate to $(x_n)_n \in V^{\mathbb{N}}$ a continuous path in V for which building a rough path makes sense. In other words, we need to *embed* the discrete sequence in the space of continuous, sufficiently regular paths in V .

Definition 3.2.3. *We define an embedding of $(x_n)_n \in V^{\mathbb{N}}$ as a sequence $\rho = (\rho_N)_N$ of continuous paths in V such that*

- $\forall N \geq 1, \rho_N : [0, N] \rightarrow V$;
- $\forall N \geq 1, \forall i \in \{1, \dots, N\}, \rho_N(i) = x_i$;
- $\forall k < N, \forall t \in [0, k], \rho_k(t) = \rho_N(t)$.

We can construct an embedding by curve concatenation in the following way: we start by connecting, for all $n \in \mathbb{N}$, x_n and x_{n+1} by a continuous curve β_n such that:

$$\beta_n : [0, 1] \rightarrow V, \beta_n(0) = x_n, \beta_n(1) = x_{n+1} \quad (3.10)$$

The embedding is then given by

$$\rho_N = \beta_1 \cdot \dots \cdot \beta_N \quad (3.11)$$

where \cdot is the operation of path concatenation in V . In order to pass to the limit, we define $\iota_N : [0, 1] \rightarrow V$ as $\iota_N(t) = f(N)\rho_N(Nt)$, where $f : \mathbb{N} \rightarrow \mathbb{R}_+$ is the renormalization function (for details on embeddings see section 3.3.3).

In theorem 3.2.1 we have used the piecewise linear embedding, i.e. we have connected two consecutive points by linear interpolation. While this is the most common embedding, it is not unique. In order to define a more general embedding for a hidden Markov walk, we proceed as follows.

Lemma 3.2.1 (construction of an embedding for HMW). *Consider a hidden Markov chain $(R_n, F_n)_n$ on $E \times V$. We construct a sequence of processes $(\rho_N)_N$ as follows:*

- for any $u \in E$, denote by $V_u \subset V$ the set of all possible realizations of F_1 under the law $\mathbb{P}(\bullet | R_1 = u)$;
- for $u \in E$, to every $y \in V_u$ we associate in a measurable way, for the usual Borel σ -algebras, a curve $f^{y;u} : [0, 1] \rightarrow V$ with bounded variation such that $f^{y;u}(0) = 0$ and $f^{y;u}(1) = y$, and we denote the set of these curves by \mathcal{B}_u ;
- for all $k \geq 1$, we associate to F_k a random variable β_k in the sense that, for a $u \in E$ and under the law $\mathbb{P}(\bullet | R_k = u)$, β_k gives for every realization y of F_k from V_u the corresponding element $f^{y;u}$ from \mathcal{B}_u ;
- for all $N \geq 1$, we set $\rho_N = \beta_1 \cdot \dots \cdot \beta_N$.

Then $(\rho_N)_N$ is an embedding of $(X_n)_n$ in the sense of definition 3.2.3.

Remark. Lemma 3.2.1 relies on the following idea: we can then "enlarge" a hidden Markov chain $(R_n, F_n)_n$ in $E \times V$ to the sequence of triplets $(R_n, F_n, \beta_n)_n$ in $E \times V \times \mathcal{C}^{1-var}([0, 1], V)$, and construct an embedding $(\rho_N)_N$ out of $(\beta_k)_k$ as in (3.11).

This leads to the following generalization of theorem 3.2.1.

Theorem 3.2.2. *Let $(R_n, F_n)_n$ be a hidden Markov chain on $E \times V$ such that $(R_n)_n$ is irreducible and condition (3.8) is satisfied, and denote by $(R_n, X_n)_n$ the corresponding Markov walk. Denote by $\rho = (\rho_n)_n$ an embedding for $(X_n)_n$ constructed as in lemma 3.2.1 with bounded variation, i.e. such that the ρ_N s are uniformly bounded over N in $\mathcal{C}^{1-var}([0, 1], V)$.*

Set $\beta = \mathbb{E}[T_1]$. Under the conditions $\mathbb{E}[X_{T_1}] = 0$ and $\mathbb{E}[X_{T_1}^{\otimes 2}] = CI_d$ for a certain $C > 0$, the rough path canonically constructed out of ρ_N , renormalized through the dilation operator $\delta_{(\beta^{-1}NC)^{-1/2}}$, converges in the rough path topology $\mathcal{C}^{\alpha-H\ddot{o}l}([0, 1], G^2(V))$ with $\alpha < 1/2$ to a rough path given by:

$$\begin{aligned} S_{B,1}(t) &= B_t \\ S_{B,2}(t) &= \int_{0 < s_1 < s_2 < t} \circ dB_{s_1} \otimes \circ dB_{s_2} + \Gamma_\rho t \end{aligned}$$

where $(B_t)_t$ is a standard d -dimensional Brownian motion and Γ_ρ is an antisymmetric deterministic matrix depending on ρ as follows:

$$\Gamma_\rho = C^{-1} \mathbb{E} \left[\frac{1}{2} \int_{0 < s_1 < s_2 < T_1} d\rho_{T_1}(s_1) \otimes d\rho_{T_1}(s_2) - d\rho_{T_1}(s_2) \otimes d\rho_{T_1}(s_1) \right]. \quad (3.12)$$

Remarks.

- Γ_ρ is the stochastic area of the curve ρ_{T_1} between the times 0 and T_1 .
- For the ρ -embedding, we have chosen curves of finite variation. This is a sufficient but not a necessary condition for the result of theorem 3.2.1 to be true. We maintain it for commodity reason, a generalisation to a weaker condition on the embedding (for example, curves of finite p -variation with a suitable p) being a question of some additional computations.

Examples to which theorem 3.2.2 applies: All the models to which theorem 3.2.1 applies and for which we connect consecutive points by curves of finite variation: the round-point model (see figure 3.2), models for traffic with road intersections, random walks on deformed networks and deformed periodic graphs etc.

The general theorem. In both theorems 3.2.1 and 3.2.2, we have lifted paths in V to geometric paths in $G^2(V)$. As we will see in section 3.3.3 (proposition 3.3.3), any sufficiently regular embedding ρ of a sequence in V can be encoded in a rough path in $G^2(V)$, so that we can start directly with a sequence in $G^2(V)$ and associate to it a rough path, i.e. a sufficiently regular embedding in $G^2(V)$.

The embeddings for sequences in $G^2(V)$ can be defined by analogy with the ones on V . In particular, we can define for sequences valued in $G^2(V)$ the *geodesic embedding*, i.e. the embedding that connects two consecutive points by a curve whose length is equal to the distance between them. When considering two points $g, h \in G^2(V)$, their geodesic embedding is the curve in $G^2(V)$ determined by the increment $g^{-1} \otimes h$. In figure 3.1, examples of geodesic curves corresponding to different elements of $G^2(\mathbb{R}^2)$ are presented.

Remark. The geodesic embedding of a sequence in $G^2(V)$ can be represented using the increments of the sequence, just like the linear interpolation of a sequence in \mathbb{R}^d .

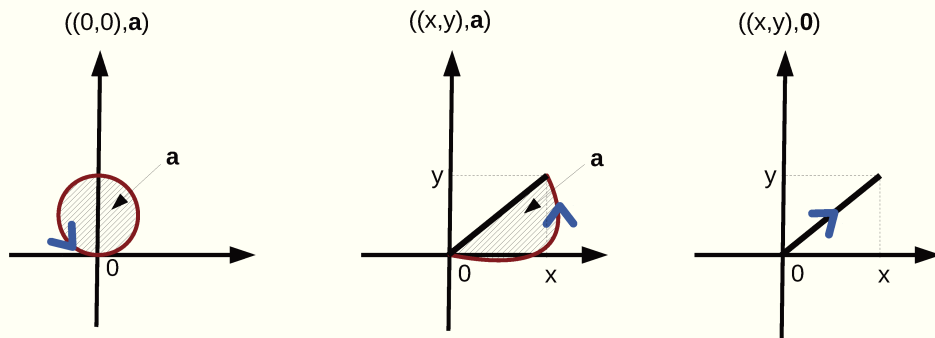


Figure 3.1: Examples of geodesic curves corresponding to elements of $G^2(V)$.

Thus, theorem 3.2.2 becomes a corollary of the following, more abstract result.

Theorem 3.2.3. Let $(R_n)_n$ be an irreducible Markov chain on E and $(R_n, \mathbb{X}_n)_n$ a

hidden Markov walk on $E \times G^2(V)$ such that

$$\exists K > 0, \forall j \in \mathbb{N}, \left\| \mathbb{X}_j^{-1} \otimes \mathbb{X}_{j+1} \right\| < K \text{ a.s.} \quad (3.13)$$

We denote by $\mathbb{X}^{(1)}$ the first component of \mathbb{X} and set $\beta = \mathbb{E}[T_1]$.

Under the conditions $\mathbb{E}[\mathbb{X}_{T_1}^{(1)}] = 0$ and $\mathbb{E}[(\mathbb{X}_{T_1}^{(1)})^{\otimes 2}] = CI_d$ (where d is the dimension of V), the geodesic embedding of $(\mathbb{X}_k)_{k \leq N}$, renormalized through the dilation $\delta_{(\beta^{-1}NC)^{-1/2}}$, converges in the rough path topology $C^{\alpha-H\ddot{o}l}([0, 1], G^2(V))$ with $\alpha < 1/2$ to the rough path given by:

$$\begin{aligned} S_{B,1}(t) &= B_t \\ S_{B,2}(t) &= \int_{0 < s_1 < s_2 < t} \circ dB_{s_1} \otimes \circ dB_{s_2} + \tilde{\Gamma}t \end{aligned}$$

where $(B_t)_t$ is a d -dimensional Brownian motion and $\tilde{\Gamma}$ is a deterministic matrix given by the antisymmetric part of $C^{-1}\mathbb{E}[\mathbb{X}_{T_1}^{(2)}]$.

Remarks.

- Theorems 3.2.1 and 3.2.2 can be seen as particular cases of theorem 3.2.3: if $\rho = (\rho_N)_N$ is a sequence of curves that can be encoded by the sequence $(\mathbb{X}_n)_n$ (as in proposition 3.3.3), we can compute Γ_ρ as in (3.12) and $\Gamma_\rho = \tilde{\Gamma}$.
- Setting

$$\mathbb{F}_j = \left(F_j, \iint_{0 < s_1 < s_2 < 1} d\beta_j(s_1) \otimes d\beta_j(s_2) \right)$$

where the β_i s are the curves that define ρ as in (3.11), we have $\mathbb{X}_n = \otimes_{j=1}^n \mathbb{F}_j$, and condition (3.13) becomes similar to (3.8):

$$\exists K > 0, \forall j \in \mathbb{N}, \|\mathbb{F}_j\| < K \text{ a.s.}$$

- Γ_ρ can be decomposed in two parts: the first one is the Γ that emerges from applying theorem 3.2.1 to the HMW $(R_n, \mathbb{X}_n^{(1)})_n$, the second one comes from applying the law of large numbers to $(\mathbb{X}_n^{(2)})_n$ (see section 3.3.4).
- If we endow the space of rough paths with the antisymmetric tensor product

\wedge , we can state an antisymmetric version of theorem 3.2.3.

Proof. See section 3.3.4. Since theorems 3.2.1 and 3.2.2 are particular cases of theorem 3.2.3, we shall not prove them separately. \square

An easy example: the diamond and round-point models.

Let us now see an example that shows what kind of processes correspond to each of the three theorems.

Consider the irreducible Markov chain $(R_n)_n$ on $E = \{1, 2, \dots, 8\}$ as represented in figure 3.2. The stochastic matrix Q corresponding to it is given by

$$Q = \begin{pmatrix} 0 & p & 0 & 1-p & 0 & 0 & 0 & 0 \\ p & 0 & 0 & 0 & 0 & 1-p & 0 & 0 \\ 0 & p & 0 & 1-p & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1-p & 0 & p & 0 \\ p & 0 & 0 & 0 & 0 & 1-p & 0 & 0 \\ 0 & 0 & 1-p & 0 & 0 & 0 & 0 & p \\ 0 & 0 & 1-p & 0 & 0 & 0 & 0 & p \\ 0 & 0 & 0 & 0 & 1-p & 0 & p & 0 \end{pmatrix} \quad (3.14)$$

By associating to each state of E a vector of \mathbb{R}^2 as in the same figure, we define a HMC $(R_n, F_n)_n$, out of which we can eventually construct a HMW $(R_n, X_n)_n$.

If we apply the piecewise linear embedding to $(X_n)_n$, we get the diamond model which is in the left lower corner of figure 3.2. It is obvious that theorem 3.2.1 applies to this model.

We can obtain a round-point model by applying to $(X_n)_n$ an embedding of round-arched openings (right lower corner of figure 3.2). This model can thus be derived from the diamond model by *a change of embedding*.

We will now rewrite both models in the rough path setting, to which theorem 3.2.3 applies. Since in both cases we construct piecewise smooth embeddings, the rough path resulting from the canonical lift is a geometric rough path, and thus taking values in $G^2(V)$. For commodity reasons, we choose to endow the set $G^2(V)$ with the antisymmetric law \wedge , so that we keep only the antisymmetric part (the signed area) of the second-level component. In the case of the diamond model, the

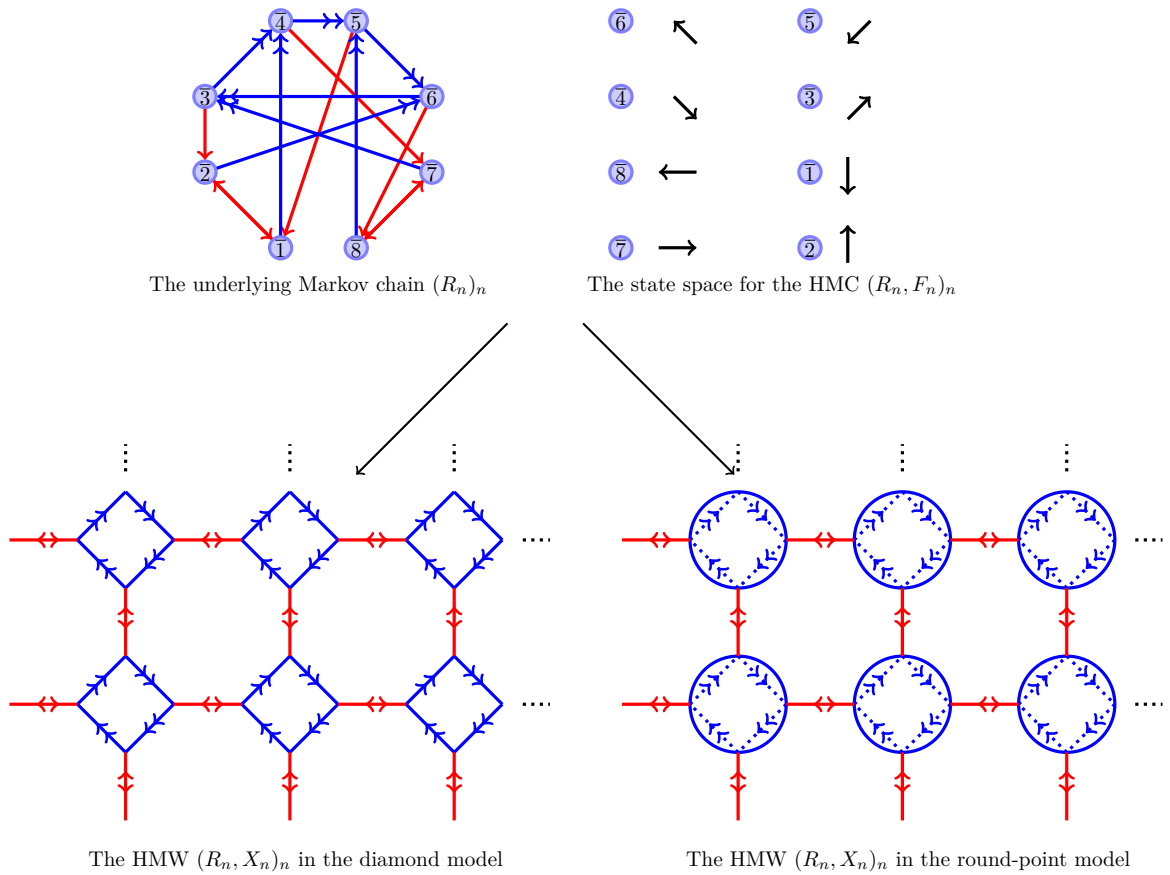


Figure 3.2: Construction of diamond model and round-point model out of the same HMC $(R_n, F_n)_n$.

use of the piecewise linear (geodesic) embedding together with the condition (3.5) for $G^2(V)$ elements imply that $(\mathbb{X}_n)_n$ is composed as

$$\mathbb{X}_N = \bigwedge_{k=1}^N (F_k, 0)$$

In the case of the round-point model, for $k \geq 1$, each round-arched opening can be encoded in an element F_k and a certain $a_k \in \mathbb{R}$, where the a_k s are the signed areas added by the round-arched openings when passing from the piecewise linear embedding of the diamond model to the new embedding. In particular, the a_k s are

the antisymmetric part of the second-level component. The sequence $(\mathbb{X}_n)_n$ is then defined as

$$\mathbb{X}_N = \bigwedge_{k=1}^N (F_k, a_k)$$

$(R_n, (F_n, a_n))_n$ is a hidden Markov chain on $E \times G^2(V)$, with the a_n s given by the signed areas of the corresponding arches of circle.

In order to take the limit, we then construct a rough path by applying the geodesic embedding to $(\mathbb{X}_n)_n$.

3.2.2 Combinatorial structure behind the iterated sums of hidden Markov walks.

Iterated sums and non-geometric rough paths.

From iterated integrals to iterated sums. We have already see that iterated integrals of type $S_{\gamma,k}$ from (3.2) (or algebraic objects satisfying their properties) are what rough paths are composed of.

Let us now pass from continuous objects to discrete ones by analogy. For a sequence $x = (x_n)_n$ in V , we define *the iterated sums* as

$$\tilde{S}_{x,l}(N) = \sum_{1 \leq j_1 < \dots < j_l \leq N} \Delta x_{j_1} \otimes \dots \otimes \Delta x_{j_l} \quad (3.15)$$

where $\Delta x_i = x_i - x_{i-1}$, and these objects can be interpreted as a discrete analogue of the iterated integrals. We will compare some algebraic properties of iterated sums $\tilde{S}_{x,l}$ with those of iterated integrals $S_{\rho,l}$, where $\rho = (\rho_N)_N$ is an embedding of x , and see how the discrete setting of iterated sums allows to isolate a particular combinatorial structure, *the iterated occupation times*.

Moreover, using the iterated sums $\tilde{S}_{x,l}(N)$, we can construct for the sequence x a discrete analogue of the step- l signature $S_l(\rho_N)$ of ρ_N constructed as in (3.3) using iterated integrals.

Definition 3.2.4. For a sequence $(x_k)_k$ in V , we define the step- l discrete signature as

$$\tilde{S}_l(x)_{1,N} = \left(\tilde{S}_{x,1}(N), \dots, \tilde{S}_{x,l}(N) \right) \in T_1^{(l)}(V) \quad (3.16)$$

One may wonder what the relation between $(S_l(\rho_N))_N$ and $(\tilde{S}_l(x)_{1,N})_N$ is and what a study on its convergence may reveal. We will partially try to answer these questions hereunder in the setting of HMW.

Non-geometric paths as limits of iterated sums of HMW. Consider a HMW $(R_n, X_n)_n$ on $E \times V$ and set

$$\mathbb{Y}_n = \tilde{S}_2(X)_{1,N} = \left(X_n, \sum_{1 \leq k_1 < k_2 \leq n} F_{k_1} \otimes F_{k_2} \right) \in T_1^{(2)}(V) \quad (3.17)$$

The following proposition shows that $(\mathbb{Y}_n)_n$ converges in rough path topology to a non-geometric rough path with a second-level deterministic drift.

Proposition 3.2.1. *Let $(R_n, X_n)_n$ be a HMW on $E \times V$ such that the corresponding F_k s satisfy (3.8). Choose $(X_t^{(N)})_t$ to be the piecewise linear embedding of $(X_n)_n$, i.e.*

$$\forall N \in \mathbb{N}^*, \forall t \in [0, 1], X_t^{(N)} = \sum_{k=1}^{\lfloor Nt \rfloor} X_k + (Nt - \lfloor Nt \rfloor) F_{\lfloor Nt \rfloor + 1}$$

and suppose $X_0 = 0$ a.s. Consider the geodesic embedding of $Z_k = \sum_{j=1}^k F_j^{\otimes 2}$ given by

$$\forall N \in \mathbb{N}^*, \forall t \in [0, 1], Z_t^{(N)} = \sum_{k=1}^{\lfloor Nt \rfloor} F_k^{\otimes 2} + (Nt - \lfloor Nt \rfloor)^2 F_{\lfloor Nt \rfloor + 1}^{\otimes 2}$$

and set

$$\forall N \in \mathbb{N}^*, \forall t \in [0, 1], \mathbb{Y}_t^{(N)} = (0, Z_t^{(N)})^{-1} \otimes S_2(X^{(N)})_{0,t}$$

Suppose in addition that $\mathbb{E}[X_{T_1}] = 0$, where T_1 is as in (3.6), and that the covariance matrix $\mathbb{E}[X_{T_1}^{\otimes 2}]$ is equal to CI_d for some $C > 0$. We also set $\beta = \mathbb{E}[T_1]$.

Then $(\mathbb{Y}_t^{(N)})_t$ is a well-defined embedding for $(\mathbb{Y}_n)_n$ and we have the following convergence in the $\mathcal{C}^{\alpha-H\delta l}([0, 1], T_1^{(2)}(V))$ topology for $\alpha < 1/2$:

$$\left(\delta_{(NC\beta^{-1})^{-1/2}} \mathbb{Y}_t^{(N)} \right)_{t \in [0,1]} \xrightarrow{N \rightarrow \infty} \left(\mathbb{B}_t^{It\delta} \otimes (0, Mt) \right)_{t \in [0,1]}$$

where $(\mathbb{B}_t^{It\delta})_t$ is the standard Brownian motion on V enhanced with iterated integrals

in the Itô sense and M is a deterministic matrix given by

$$M = C^{-1} \mathbb{E} \left[\sum_{1 \leq k_1 < k_2 \leq T_1} F_{k_1} \otimes F_{k_2} \right] \quad (3.18)$$

In particular, $(\mathbb{B}_t^{\text{Itô}} \otimes (0, Mt))_t$ is a non-geometric rough path.

Proof. The proof can be found in section 3.4.2. □

Iterated occupation times of HMW as underlying combinatorial structures of rough paths.

Definition and combinatorial nature of iterated occupation times. In the case of a HMW $(R_n, X_n)_n$ on $E \times V$ ($(R_n)_n$ is irreducible as before), the conditional measure $\nu(\bullet|u)$ is important for determining the trajectory on V . However, when it comes to algebraic properties, it actually plays only a small role; most of the algebraic properties of rough paths are already present at the level of the discrete Markov chain $(R_n)_n$. They are concentrated in combinatorial objects contained in the HMW, *the iterated occupation times*.

The definition of a HMW implies that

$$\mathbb{E}[X_n | \sigma(R)] = \sum_{u \in E} f(u) \sum_{k=1}^n \mathbf{1}_{R_k=u} \quad (3.19)$$

where $R = (R_n)_n$ and $f(u) = \mathbb{E}_\nu[X_1 - X_0 | u]$. The occupation time of u by $(R_n)_n$ up to time n

$$L_{u;n}(R) = \sum_{k=1}^n \mathbf{1}_{R_k=u} \quad (3.20)$$

is a natural r.v. at the heart of ergodic-type theorems. We can generalize the notion of occupation time (3.20) as follows.

Definition 3.2.5 (iterated occupation times). *Let $(R_n)_n$ be an irreducible Markov chain on E . For any $k \in \mathbb{N}^*$ and any elements $(u_1, \dots, u_k) \in E^k$, the iterated occupation time of (u_1, \dots, u_k) by $(R_n)_n$ at time $n \in \mathbb{N}$ is defined as:*

$$L_{u_1, \dots, u_k; n}(R) = \text{card} \{ (n_1, \dots, n_k) \in \Delta_k(n); R_{n_1} = u_1, \dots, R_{n_k} = u_k \} \quad (3.21)$$

where $\Delta_k(n)$ is the set

$$\Delta_k(n) = \left\{ (n_1, \dots, n_k) \in \mathbb{N}^k; 1 \leq n_1 < n_2 < \dots < n_k \leq n \right\} \quad (3.22)$$

This cardinal can be written as an iterated sum of products of indicator functions

$$L_{u_1, \dots, u_k; N}(R) = \sum_{1 \leq l_1 < \dots < l_k \leq N} \mathbf{1}_{R_{l_1} = u_1} \dots \mathbf{1}_{R_{l_k} = u_k} \quad (3.23)$$

Further on, we can easily generalize relation (3.19) through the following decomposition.

Property 3.2.1.

$$\begin{aligned} \mathbb{E} \left[\tilde{S}_{X,l}(N) \middle| \sigma(R) \right] &= \mathbb{E} \left[\sum_{1 \leq j_1 < \dots < j_l \leq N} \Delta X_{j_1} \otimes \dots \otimes \Delta X_{j_l} \middle| \sigma(R) \right] \\ &= \sum_{(u_1, \dots, u_l) \in E^l} (f(u_1) \otimes \dots \otimes f(u_l)) L_{u_1, \dots, u_l; N}(R) \end{aligned} \quad (3.24)$$

where $R = (R_n)_n$, $X = (X_n)_n$ and $f(u) = \mathbb{E}_\nu [X_1 - X_0 | u]$ as before.

Remark. Formula 3.24 tells us that, considering under the law $\mathbb{P}(\bullet | \sigma(R))$, the expectations of iterated sums of a HMW are linear combinations of occupation times of the underlying Markov chain. Conversely, an iterated sum of the type

$$\sum_{(u_1, \dots, u_l) \in E^l} v L_{u_1, \dots, u_l; N}(R)$$

with $v \in V^{\otimes l}$ corresponds (under the same probability law) to the expectation of an iterated sum of order l of a certain HMW depending on $(R_n)_n$.

From shuffle to quasi-shuffle products. A particularity of the Stratonovich iterated integrals is that they satisfy *the chain rule*; for example, the product of two components of the first levels satisfies:

$$S_{\gamma,1}(t)_i \otimes S_{\gamma,1}(t)_j = S_{\gamma,2}(t)_{ij} + S_{\gamma,2}(t)_{ji} \quad (3.25)$$

where x_i is the i -th coordinate of a vector $x \in V$. Formula (3.25) is actually an illustration of the fact that the multiplication of Stratonovich iterated integrals is a *shuffle product* introduced in [28]. This product can be identified with a particular set of permutations giving all the ways of interlacing two ordered sets while preserving the original order of components of each of them.

In the case of Itô integrals and iterated sums. We have seen that the combinatorial properties of the latter concentrate in the iterated occupation times, which may themselves be understood, through formula (3.23), as basis vectors of \mathbb{R}^E . This formal identification establishes a relation between iterated integrals/sums of rough paths with the present iterated occupation time. In particular, one checks easily for any elements $u \neq v$

$$L_u(n)L_v(n) = L_{u,v}(n) + L_{v,u}(n) \quad (3.26)$$

which is very reminiscent of (3.25). However, for $u = v$, the formula is modified to:

$$L_u(n)L_u(n) = 2L_{u,u}(n) + L_u(n) \quad (3.27)$$

The emergence of the last term is due to the discrete nature of the sums (in the context of stochastic iterated integrals, it is due to the extra-drift of the Itô integral). This extra-term implies that we need a more general product than the shuffle product to characterize the multiplication of combinatorial structures behind HMW.

Property 3.2.2. *The product of two iterated sums/occupation times of a HMW can be identified with a quasi-shuffle product, where the quasi-shuffle product is as in definition 3.4.1.*

The quasi-shuffle products, introduced in [45] and extended, for example, in [47], are a generalization of the shuffle product: while the latter supposes that components from two different sets can never coincide when interlacing the sets, the former allows for sets to overlap (following certain rules).

Remark: Given a HMW on $E \times V$ and an embedding for it, the structure given by the shuffle product to the space of the corresponding (Stratonovich) iterated

integrals, on one hand, and the one given to the space of the corresponding iterated sums/occupation times by the quasi-shuffle product on the other gives rise to an interesting illustration of the formal algebraic definitions given in [43] for the geometric and non-geometric rough paths.

Asymptotics of the combinatorial structure of HMW. Whereas the convergence of iterated sums of a HMW $(R_n, X_n)_n$ on $E \times V$ can not be directly expressed through the convergence of their combinatorial structure, there is an interesting link between both of them.

Following theorem 3.2.1 and the classical ergodic theorem, we want to study the large time values and dynamics of $L_{u_1, \dots, u_k; n}(R)$ using the results presented above. We provide two types of estimates: almost sure limits of renormalized iterated occupation times and convergence in law to anomalous Brownian motion in the rough path topology for corrections to the almost sure limits.

First, the iterated occupation times satisfy the following scaling limit.

Proposition 3.2.2 (almost sure convergence). *Let $(R_n)_{n \in \mathbb{N}}$ be an irreducible Markov chain on a finite set E and $L_{u_1, \dots, u_k; n}(R)$ as in definition 3.2.5. Then, for any $k \in \mathbb{N}$, the following convergence holds almost surely:*

$$\frac{L_{u_1, \dots, u_k; \lfloor Nt \rfloor}(R)}{N^k} \xrightarrow{N \rightarrow \infty} (1/k!) t^k \pi(u_1) \dots \pi(u_k) \quad (3.28)$$

where π is the invariant measure for $(R_n)_n$.

Proof. We give only the main idea of the proof: using the theory of pseudo-excursions, we prove that $L_{u_1, \dots, u_k; \lfloor Nt \rfloor}(R)$ can be approximated by a polynomial of degree N^k , and show that the coefficient corresponding to this degree is given by the following a.s. convergence obtained by applying the classical ergodic theorem:

$$N^{-k} \binom{N}{k} L_{u_1}(R) \dots L_{u_k}(R) \xrightarrow{N \rightarrow \infty} (1/k!) \pi(u_1) \dots \pi(u_k)$$

□

Remarks.

- Proposition 3.2.2 can be seen as a generalized version of the classical ergodic theorem which is also proved using a decomposition into excursions (see, for example, [51]).
- The stochastic matrix Q of $(R_n)_n$ appears here through the invariant measure π .

We now present the results on the convergence in law. We set $L_n(r) = (L_{u;n}(r))_{u \in E}$. In this case, we get the integer-valued vector $L_n(r) \in \mathbb{N}^{\text{Card}(E)}$. We then have the following result which brings into light a universality class of convergence for the HMW:

Proposition 3.2.3. *Let $(R_n)_n$ be an irreducible Markov chain on the finite state space E . For $n \geq 1$, set $\tilde{L}_n(R) = (L_{u;n}(R) - n\pi(u))_{u \in E} \in \mathbb{R}^{\text{Card}(E)}$, where π is the invariant probability measure of $(R_n)_n$. Then, modulo the appropriate renormalization and a multiplicative constant, the piecewise linear embedding from proposition 3.4.3 of $(\tilde{L}_n(R))_n$ and the associated iterated integrals $S_{\tilde{L}_n(R),k}(t)$ with $k \geq 2$ converge in rough path topology of $\mathcal{C}^{\alpha-H\ddot{o}l}([0, 1], T_1^{(2)}(V))$ with $\alpha < 1/2$ towards*

$$S_{B,1}(t) = B_t,$$

$$S_{B,2}(t) = \int_{0 < s_1 < s_2 < t} dB_{s_1} \otimes dB_{s_2} + Mt$$

where $(B_t)_t$ is a $\text{Card}(E)$ -dimensional Brownian motion and M is a deterministic matrix defined as in (3.18).

Proof. Follows directly from proposition 3.2.1. □

Remark. Notice that the Brownian motion is not standard here. This comes from the fact that the covariance matrix of an excursion here is not diagonal.

We have the following immediate corollary, which can be viewed as a kind of central limit theorem for iterated sums.

Corollary 3.2.1. *Let $(R_n)_n$ be an irreducible Markov chain on the finite state space E . For any $u, v \in E$, set*

$$\tilde{L}_{u,v;N}(R) = \sum_{1 \leq l_1 < l_2 \leq N} (\mathbf{1}_{R_{l_1}=u} - \pi(u))(\mathbf{1}_{R_{l_2}=v} - \pi(v))$$

Then for any $u, v \in E$, we have the following convergence in law in the uniform topology:

$$N^{-1} \tilde{L}_{u,v;N}(R) \xrightarrow{N \rightarrow \infty} \int_{0 < s_1 < s_2 < 1} dB_{s_1}^{(u)} dB_{s_2}^{(v)} + M_{u,v}$$

where B is a $\text{Card}(E)$ -dimensional Brownian motion and the matrix $M_{u,v}$ is the (u, v) entrance of the $\text{Card}(E) \times \text{Card}(E)$ matrix M from proposition 3.2.3.

Remark. We can generalize the result of corollary 3.2.1 for iteration times of higher order by setting, for any $u_1, \dots, u_k \in E$,

$$\tilde{L}_{u_1, \dots, u_k; N}(R) = \sum_{1 \leq l_1 < \dots < l_k \leq N} (\mathbf{1}_{R_{l_1} = u_1} - \pi(u_1)) \dots (\mathbf{1}_{R_{l_k} = u_k} - \pi(u_k))$$

However, in this case the additional terms at the limit (all the terms aside the iterated integral of Brownian motion) are more difficult to express explicitly, even if they depend on M and $(B_t)_t$.

3.3 From Hidden Markov Walks to rough paths

3.3.1 Theory of pseudo-excursions for hidden Markov walks.

Another definition of HMW. Definitions 3.2.1 and 3.2.2 can be summed up in the following definition-property, which includes also the HMW taking values in $E \times G^2(V)$, where $G^2(V)$ is as defined in notations of section 3.1.2 or in definition 3.3.3.

Property 3.3.1. A process $(R_n, X_n)_n$ is a hidden Markov walk on $E \times V$ (resp. on $E \times G^2(V)$) in the sense of definition 3.2.2 if and only if there exists a sequence of r.v. $(F_n)_n$ with values in V (resp. $G^2(V)$) such that

- under $\mathbb{P}(\bullet | \sigma(R))$, with $R = (R_n)_n$, $(F_n)_n$ is a sequence of independent r.v.;
- the distribution of F_n knowing $R_n = u$ is $\nu(\bullet | u)$ for all $n \in \mathbb{N}^*$;
- $\forall n \in \mathbb{N}^*$, $X_n = \sum_{k=1}^n F_k$ (resp. $X_n = \otimes_{k=1}^n F_k$ in $G^2(V)$).

In particular, we have $F_k = X_k - X_{k-1}$ (resp. $F_k = X_{k-1}^{-1} \otimes X_k$ in $G^2(V)$).

Theory of pseudo-excursions. As it has already been mentioned in the introduction, HMW are a way of generalizing simple random walks: instead of being i.i.d. variables, the increments are variables depending on a Markov chain on a finite state space E : to each state $u \in E$ we associate a supplementary object (a vector, a curve etc.) which gives the increment. However, we can derive a sum of i.i.d.r.v. from a HMW using *the theory of pseudo-excursions*. For commodity reasons, we state it for HMW on $E \times V$, but it is equally valid for those on $E \times G^2(V)$.

If $(R_n)_n$ is an irreducible Markov chain on a finite state space E , we can apply to it all the excursion theory we have for Markov chains. Thus, even if the hidden Markov walk $(R_n, X_n)_n$ may not even be a Markov chain, we can construct a theory of pseudo-excursions for it based on the results we have for $(R_n)_n$.

We start by defining a sequence of stopping times for the Markov chain $(R_n)_n$ as in (3.6). We then have the following definition, which is a property at the same time.

Definition 3.3.1 (pseudo-excursions). *Let $(R_n, X_n)_n$ be a HMW on $E \times V$ and the sequence $(T_k)_k$ as defined in (3.6). We call $(X_{T_{k-1}}, X_{T_{k-1}+1}, \dots, X_{T_k})$ the k -th pseudo-excursion of $(X_n)_n$ and $X_{T_k} - X_{T_{k-1}}$ the k -th pseudo-excursion increment.*

Property 3.3.2. *We have the following basic properties for pseudo-excursions:*

- *The variables $X_{T_{k+1}} - X_{T_k}$ are i.i.d. random variables.*
- *For $i \neq j$, the trajectories $(X_{T_i+1}, \dots, X_{T_{i+1}})$ and $(X_{T_j+1}, \dots, X_{T_{j+1}})$ are independent and have the same law.*

Proof. Let $k \neq m$. The pseudo-excursions $X_{T_{k+1}} - X_{T_k}$ and $X_{T_{m+1}} - X_{T_m}$ are independent knowing $(R_{T_k+1}, \dots, R_{T_{k+1}})$ and $(R_{T_m+1}, \dots, R_{T_{m+1}})$ respectively (this follows from property 3.3.1). Since $k \neq m$, the excursions $(R_{T_k+1}, \dots, R_{T_{k+1}})$ and $(R_{T_m+1}, \dots, R_{T_{m+1}})$ are i.i.d., and thus $X_{T_{k+1}} - X_{T_k}$ and $X_{T_{m+1}} - X_{T_m}$ are independent and have the same law. \square

Remark. Let us briefly justify the choice of the term "pseudo-excursions". The excursion part comes from the fact that $X_{T_{k+1}} - X_{T_k}$ depends on the excursion $(R_{T_k+1}, \dots, R_{T_{k+1}})$. Why "pseudo"? Because an ordinary excursion is the trajectory of a process before returning to its starting point, and the increments $X_{T_{k+1}} - X_{T_k}$ are non-trivial, i.e. we do not return to our starting point.

The pseudo-excursions are particularly interesting when it comes to convergence of a hidden Markov walk, as they tell us that it can be studied almost as the convergence of a sum of well-chosen i.i.d.r.v.

Proposition 3.3.1. *Let $(R_n, X_n)_n$ be a hidden Markov walk as in definition 3.2.2. Suppose that there exists $K > 0$ such that $|X_{n+1} - X_n| \leq K$ a.s. There exists a sequence $(V_n)_n$ of i.i.d.r.v. such that, setting, for all $n \geq 1$,*

$$S_n = \sum_{k=1}^{\lfloor n\mathbb{E}[T_1]^{-1} \rfloor} V_k$$

we have

$$\forall \epsilon > 0, \mathbb{P}(|X_n - S_n| > \epsilon) \rightarrow 0$$

Proof. Let $\kappa(n)$ be an integer such that $T_{\kappa(n)} \leq n < T_{\kappa(n)+1}$. We decompose

$$X_n = \sum_{i=1}^{\kappa(n)} (X_{T_i} - X_{T_{i-1}}) + \sum_{i=T_{\kappa(n)+1}}^n (X_i - X_{i-1})$$

The first sum contains $\kappa(n)$ i.i.d. variables, and moreover $\kappa(n)/n \xrightarrow[n \rightarrow \infty]{} \mathbb{E}[T_1]^{-1}$ a.s. by the ergodic theorem.

We will now prove that the second sum converges to zero in probability. For $\epsilon > 0$, we have:

$$\begin{aligned} \mathbb{P}\left(\left|\sum_{i=T_{\kappa(n)+1}}^n X_i - X_{i-1}\right| > \epsilon\right) &\leq \mathbb{P}\left(\sum_{i=T_{\kappa(n)+1}}^{T_{\kappa(n)+1}} |X_i - X_{i-1}| > \epsilon\right) \\ &\leq \mathbb{P}\left(K(T_{\kappa(n)+1} - T_{\kappa(n)}) > \epsilon\right) \\ &\leq K\epsilon^{-1}\mathbb{E}[T_1] \rightarrow 0 \end{aligned}$$

We then obtain the desired result by applying Slutsky's theorem (for reference, see

[9], theorem 3.1). □

Proposition 3.3.1 is very useful in the sense that it allows to adopt the convergence theorems we have for sums of i.i.d. random variables to the class of hidden Markov walks.

3.3.2 Elements of rough paths theory

The reader who is well familiar with rough paths may skip this part, and those who would like a more detailed study may refer, for example, to [34] or to [32].

Some definitions. In general, a rough path is defined as a continuous path on $T_1^{(N)}(V)$ for $N \geq 1$. When we want to associate a rough path to a V -valued path γ of regularity $\alpha < 1$, we choose $N = \lfloor 1/\alpha \rfloor$ and fill each level following some rules.

Remark. The way of lifting a path in V to a rough path is not unique.

In particular, the regularity of the rough path has to be coherent with the regularity of the initial path (for details on the regularity issues, see, for example, [66] or [34]). Throughout this article, we are interested in rough paths that convergence towards the Brownian motion whose regularity is $1/2^-$, so the rough paths we operate correspond to the following definition.

Definition 3.3.2. *Let $1/3 < \alpha < 1/2$. An α -Hölder rough path $(\mathbf{x}_t)_t$ is an element of $\mathcal{C}^{\alpha-H\ddot{o}l}([0, 1], T_1^{(2)}(V))$. Moreover, if $(\mathbf{x}_t)_t$ is an element of $\mathcal{C}^{\alpha-H\ddot{o}l}([0, 1], G^2(V))$, it is an α -Hölder geometric rough path.*

A more informal way to state definition 3.3.2 is the following:

- a geometric rough path on $G^2(V)$ endowed with the law \otimes inherited from $T_1^{(2)}(V)$: $(a, b) \otimes (a', b') = (a + a', b + b' + a \otimes a')$;
- a geometric rough path on $G^2(V)$ endowed with the antisymmetric law \wedge , which eliminates the symmetric part of the second level (it is seen as redundant since it depends entirely on the first level): $(a, b) \wedge (a', b') = (a + a', b + b' + \frac{1}{2}a \otimes a' - a' \otimes a)$;

- a non-geometric rough path on $T_1^{(2)}(V)$ endowed with the law \otimes (we do not consider the antisymmetric law \wedge since the symmetric part of the second level contains elements that do not depend on the first level).

An important property of rough paths is that they satisfy *Chen's relation*, i.e.

$$\forall 0 \leq s < u < t \leq 1, \mathbf{x}_{s,u} \otimes \mathbf{x}_{u,t} = \mathbf{x}_{s,t} \quad (3.29)$$

or, alternatively, in the antisymmetric setting (for $G^2(V)$),

$$\forall 0 \leq s < u < t \leq 1, \mathbf{x}_{s,u} \wedge \mathbf{x}_{u,t} = \mathbf{x}_{s,t} \quad (3.30)$$

where $\forall 0 \leq s < u$, $\mathbf{x}_{s,u} = \mathbf{x}_u^{-1} \otimes \mathbf{x}_s$. This reflects, in particular, the fact that \mathbf{x}_t can be decomposed into a "sum" of increments of \mathbf{x} .

The *Brownian motion rough path* is the most important practical example of rough path in the probabilistic setting. Since stochastic calculus allows us to define integrals with respect to the Brownian motion (Itô or Stratonovich), we can directly construct the rough path as:

$$\mathbb{B}_{s,t} = (B_t - B_s, \int_{s < u < v < t} \circ dB_u \otimes \circ dB_v) \quad (3.31)$$

In the antisymmetric setting of $G^2(V)$, we only keep the antisymmetric part of (3.31) and we obtain *the enhanced Brownian motion*:

$$\mathbb{B}_{s,t} = (B_t - B_s, \mathcal{A}_{s,t}) \quad (3.32)$$

where \mathcal{A} is the stochastic signed area of B , called *the Lévy area*. The enhanced Brownian motion is a Brownian motion on $G^2(V)$.

The group $G^2(V)$: alternative definition and topology. Since our main theorems of convergence (3.2.1, 3.2.2, 3.2.3) are dealing with geometric rough paths, we give more details about the group $G^2(V)$.

In the introduction, we have stated that $G^2(V)$ is a subgroup of $T_1^{(2)}(V)$ whose elements satisfy condition (3.5) (i.e. the symmetrical part of the second level depends entirely on the first level). We will now give an alternative, more analytical

definition.

Definition 3.3.3. *The free nilpotent group $G^2(V)$ is defined as:*

$$G^2(V) = \{S_2(\gamma)_{0,1} : \gamma \in \mathcal{C}^{1-var}([0, 1], V)\}$$

where $S_2(\gamma)_{0,1}$ is as in (3.3).

The topology of $G^2(V)$ is induced by the *Carnot-Caratheodory norm*, which gives the length of the shortest path corresponding to a given signature:

$$\forall g \in G^2(V), \|g\| := \inf \left\{ \int_0^1 |dx| : x \in \mathcal{C}^{1-var}([0, 1], V) \text{ and } S_2(\gamma)_{0,1} = g \right\} \quad (3.33)$$

where $|\cdot|_V$ is a restriction to V of the Euclidean norm. The norm thus defined is homogeneous ($\|\delta_\lambda g\| = |\lambda| \|g\|$ for $\lambda \in \mathbb{R}$), symmetric ($\|g\| = \|g^{-1}\|$) and sub-additive ($\|g \otimes h\| \leq \|g\| + \|h\|$), it induces a continuous metric \mathbf{d} on $G^2(V)$ through the application

$$\begin{aligned} \mathbf{d} : G^2(V) \times G^2(V) &\rightarrow \mathbb{R}_+ \\ (g, h) &\mapsto \|g^{-1} \otimes h\| \end{aligned} \quad (3.34)$$

In this case, $(G^2(V), \mathbf{d})$ is a geodesic space (in the sense of definition 5.19 from [34]) and a Polish space.

Another useful norm on $G^2(V)$, homogeneous but neither symmetric nor sub-additive, is given by:

$$\forall (a, b) \in G^2(V), |||(a, b)||| = \max\{|a|_V, |b|_{V \otimes V}^{1/2}\}$$

Since all homogeneous norms are equivalent on $G^2(V)$, this new norm gives us a rather easy way to get an estimate of the Carnot-Caratheodory norm:

Property 3.3.3. *There exist two positive constants c_1 and c_2 such that*

$$\forall (a, b) \in G^2(V), c_1 |||(a, b)||| \leq \|(a, b)\| \leq c_2 |||(a, b)||| \quad (3.35)$$

A convergence criterion for rough paths. We usually prove the convergence of a sequence of rough paths using pointwise convergence plus a tightness criterion.

However, for some cases, we can also use the following result, deduced from exercise 2.9 in [32]:

Proposition 3.3.2. *Consider a sequence of rough paths $(\mathbb{X}^{(N)})_N$ which takes values in $\mathcal{C}^{\alpha-H\ddot{o}l}([0, 1], G^2(V))$ with $\alpha < 1/2$ such that we have the uniform bound*

$$\sup_N \mathbb{E} \left[\sup_{s \neq t} \frac{\|\mathbb{X}_{s,t}^{(N)}\|}{|t-s|^{1/2}} \right] < \infty$$

and the pointwise convergence (in probability)

$$\forall t \in [0, 1], \mathbb{X}_{0,t}^{(N)} \xrightarrow[N \rightarrow \infty]{} \mathbb{X}_{0,t}$$

for a certain rough path \mathbb{X} in $\mathcal{C}^{\alpha-H\ddot{o}l}([0, 1], G^2(V))$.

Then $\mathbb{X}^{(N)}$ converges in probability to \mathbb{X} in the $\mathcal{C}^{\alpha-H\ddot{o}l}([0, 1], G^2(V))$ topology with $\alpha < 1/2$.

3.3.3 Embeddings

Definition and equivalence classes in the case of a finite-dimensional vector space. In order to study the limit of a discrete process, we need to properly define this convergence in the continuous space, where the limit process lives. This is when embeddings come on the scene.

In the introduction, we have given a general definition 3.2.3 of an embedding and a way of constructing an embedding for a given sequence of points (3.11). The following property highlights the fact that the method is consistent with the definition:

Definition 3.3.4. *Let $(x_n)_n$ be a sequence with values in a finite-dimensional vector space V . Consider a sequence $(\beta_n)_n$ of continuous paths $\beta_n : [0, 1] \rightarrow V$ and such that $\beta_n(0) = x_n$ and $\beta_n(1) = x_{n+1}$. Let \cdot be the operator of path concatenation*

given by

$$\begin{aligned} & \forall \gamma : [s, t] \rightarrow V, \forall \gamma' : [s', t'] \rightarrow V, \forall u \in [s, t + t' - s'], \\ (\gamma \cdot \gamma')(u) &= \begin{cases} \gamma(u) & \text{if } u \in [s, t] \\ \gamma'(u - t + s') + \gamma(t) - \gamma'(s') & \text{if } u \in [t, t + t' - s'] \end{cases} \end{aligned} \quad (3.36)$$

Then the sequence $\rho_N = \beta_1 \cdot \dots \cdot \beta_N$ is an embedding of $(x_n)_n$.

In some cases, we can study the convergence of the embedding thus obtained by shrinking the space scale of the paths, i.e. by *rescaling the embedding*. We then have the following definition:

Definition 3.3.5. *Let V be a finite-dimensional vector space. Consider a V -valued embedding $\rho_N = \beta_1 \cdot \dots \cdot \beta_N$ as defined in 3.3.4. Let $f : \mathbb{N} \rightarrow \mathbb{R}_+$ be an increasing function. The sequence $\iota_N = f(N)\rho_N$ defines a rescaled embedding.*

Even if, for a given discrete process, we have a large choice of embeddings, we should be aware of the fact that the limit we get depends of the embedding we adopt (some embeddings may not even admit a limit). For a sequence $(x_n)_n$ with values in a vector space V , we can define a way of dividing the embeddings into equivalence classes by *building rough paths out of our embeddings* (when possible) and considering the convergence in the rough path topology:

Definition 3.3.6. *Let $(x_n)_n$ be a sequence with values in a (finite-dimensional) vector space V . We say that two embeddings $\rho = (\rho_N)_N$ and $\rho' = (\rho'_N)_N$ of $(x_n)_n$ are equivalent (in the rough path sense) if there exists a function $f : \mathbb{N} \rightarrow \mathbb{R}_+$ such that the rough paths $\mathbf{\iota}_N$ and $\mathbf{\iota}'_N$ corresponding to $\iota_N = f(N)\rho_N$ and $\iota'_N = f(N)\rho'_N$ respectively converge in distribution to the same limit \mathbf{l} in the rough path topology.*

We shall now see how this definition is important for generalizing the convergence of HMW in rough path topology.

From sequences in $G^2(V)$ to rough paths. For $n \in \mathbb{N}$, we consider a sequence $(g_k)_k = (g_k^{(1)}, g_k^{(2)})_k$ in $G^2(V)$. As in the case of a sequence in V , we can pass to the continuous framework by associating a set of embeddings to $(g_k)_k$ as follows.

Definition 3.3.7. Consider a sequence $(g_k)_k \in G^2(V)^{\mathbb{N}}$. Consider a sequence of rough paths $(\rho_N)_N$ such that $\rho_N : [0, N] \rightarrow G^2(V)$ which satisfy

$$\forall 1 \leq k \leq N, \rho_N(k-1)^{-1} \otimes \rho_N(k) = g_k.$$

We say that $(\rho_N)_N$ is an embedding for $(g_k)_k$.

Just like in the case of a vector space V , a particularly important embedding is *the geodesic embedding*. In general, a *geodesic* (or a geodesic curve) on a metric space is a curve (or one of the curves if there is no unicity) minimizing the distance between two points; in other words, it is a notion generalizing the straight line from euclidean geometry. The geodesic embedding consists in connecting two consecutive elements of a sequence by geodesic curves.

Definition 3.3.8 (the geodesic embedding in $G^2(V)$). Consider a sequence $(g_k)_k \in G^2(V)^{\mathbb{N}}$. We call *geodesic embedding of $(g_k)_k$* the sequence of rough paths

$$\begin{aligned} \rho_N : [0, N] &\rightarrow G^2(V) \\ t &\mapsto \bigotimes_{k=1}^{\lfloor t \rfloor} g_k \otimes \delta_{t-\lfloor t \rfloor}(g_{\lfloor t \rfloor+1}) \end{aligned} \quad (3.37)$$

The universality of the geodesic embedding in $G^2(V)$ for HMW. We can now explain what we mean by the universality of the geodesic embedding. The geodesic embedding on $G^2(V)$ allows us to describe further beyond the equivalence classes the convergence of different embeddings of a hidden Markov walk. More precisely, the limit of all the embeddings that are equivalent in the rough path sense is the same as the limit of the geodesic embedding of a well-chosen sequence in $G^2(V)$.

Proposition 3.3.3. Consider a process $(X_n)_n$ on V and suppose that there exists $(\rho_N)_N$ an embedding of $(X_n)_n$ with bounded variation (i.e. a uniform bound over N in 1-variation topology) constructed, for example, as in lemma 3.2.1, and such that $(N^{-1/2}\rho_N(Nt))_{t \in [0,1]}$ converges in law in $\mathcal{C}^{\alpha-H\ddot{o}l}([0,1], G^2(V))$ topology for $\alpha < 1/2$. Then there exists a sequence $g_n = (g_n^{(1)}, g_n^{(2)})_n \in G^2(V)$ such that its geodesic embedding rescaled by the operator $\delta_{N^{-1/2}}$ converges to the same limit in $\mathcal{C}^{\alpha-H\ddot{o}l}([0,1], G^2(V))$ topology with $\alpha < 1/2$ as any embedding equivalent to ρ_N in the sense of definition 3.3.6.

Proof. Denote by $\boldsymbol{\rho}_N$ the geometric rough path (the 2-step signature) corresponding to ρ_N . As before, we denote by $\|\cdot\|$ the Carnot-Caratheodory norm defined in (3.33). We will prove the result only in the particular case

$$\exists c > 0, \forall \alpha' \in]0, 1/2[, \forall N \geq 1, \forall 0 \leq u < v \leq N, \left\| \boldsymbol{\rho}_N(u)^{-1} \otimes \boldsymbol{\rho}_N(v) \right\| \leq c|v - u|^{\alpha'}$$

in order for it not to be too technical, and also because most of the times the embeddings are smooth. Set

$$\forall k \geq 1, g_k = \boldsymbol{\rho}_k(k-1)^{-1} \otimes \boldsymbol{\rho}_k(k)$$

By definition of the embedding 3.3.4, this means that

$$\forall k \geq 1, \forall N \geq k, g_k = \boldsymbol{\rho}_N(k-1)^{-1} \otimes \boldsymbol{\rho}_N(k)$$

The geodesic embedding for the sequence $(g_k)_k$ is given by

$$\begin{aligned} \forall t \in [0, 1], \boldsymbol{\rho}_N(Nt) &= \bigotimes_{k=1}^{\lfloor Nt \rfloor} g_k \otimes \delta_{Nt - \lfloor Nt \rfloor}(g_{\lfloor Nt \rfloor + 1}) \\ &= \boldsymbol{\rho}_N(\lfloor Nt \rfloor) \otimes \delta_{Nt - \lfloor Nt \rfloor}(\boldsymbol{\rho}_N(\lfloor Nt \rfloor)^{-1} \otimes \boldsymbol{\rho}_N(\lfloor Nt \rfloor + 1)) \end{aligned}$$

It is now left to prove that $(\delta_{N^{-1/2}} \boldsymbol{\rho}_N(Nt))_{t \in [0,1]}$ and $(\delta_{N^{-1/2}} \boldsymbol{\rho}_N(Nt))_{t \in [0,1]}$ converge towards the same limit in rough path topology.

For all $\alpha' \in]0, 1/2[, N \geq 1$ and $0 \leq s < t \leq 1$, we have that

$$N^{-1/2} \frac{\|\boldsymbol{\rho}_N(Ns)^{-1} \otimes \boldsymbol{\rho}_N(Nt)\|}{|s - t|^{\alpha'}} \leq N^{-\alpha'} \frac{\|\boldsymbol{\rho}_N(Ns)^{-1} \otimes \boldsymbol{\rho}_N(Nt)\|}{|s - t|^{\alpha'}} \leq c$$

which implies that $\|\delta_{N^{-1/2}} \boldsymbol{\rho}_N(N\cdot)\|_{\alpha' - H\delta l; [0,1]}$ is uniformly bounded over N . Since $\boldsymbol{\rho}_N$ is a geodesic approximation of $\boldsymbol{\rho}_N$, proposition 5.2.2 from [34] tells us that $\|\boldsymbol{\rho}_N(N\cdot)\|_{\alpha' - H\delta l; [0,1]}$ can be bounded from above by $\|\boldsymbol{\rho}_N(N\cdot)\|_{\alpha' - H\delta l; [0,1]}$ (modulo a multiplicative constant), consequently

$$\sup_N \mathbb{E} \left[\|\delta_{N^{-1/2}} \boldsymbol{\rho}_N(N\cdot)\|_{\alpha' - H\delta l; [0,1]} \right] < \infty$$

and we deduce that the sequence $\delta_{N^{-1/2}}(\boldsymbol{\rho}_N(Nt))_t$ is tight in $\mathcal{C}^{\alpha-H\ddot{o}l}([0, 1], G^2(V))$ for $\alpha < 1/2$ (see corollary A.12 from [34]).

We then compute

$$\begin{aligned} & \sup_{t \in [0,1]} \left\| \delta_{N^{-1/2}} \boldsymbol{\rho}_N(Nt)^{-1} \otimes \delta_{N^{-1/2}} \boldsymbol{\rho}_N(Nt) \right\| \\ & \leq N^{-1/2} \left(\sup_{t \in [0,1]} \left\| \boldsymbol{\rho}_N(Nt)^{-1} \otimes \boldsymbol{\rho}_N(\lfloor Nt \rfloor) \right\| + \sup_{t \in [0,1]} \left\| \delta_{Nt - \lfloor Nt \rfloor} (\boldsymbol{\rho}_N(\lfloor Nt \rfloor)^{-1} \otimes \boldsymbol{\rho}_N(\lfloor Nt \rfloor + 1)) \right\| \right) \\ & \leq cN^{-1/2} \left(\sup_{t \in [0,1]} (Nt - \lfloor Nt \rfloor)^\alpha + \sup_{t \in [0,1]} (Nt - \lfloor Nt \rfloor) \right) \xrightarrow{N \rightarrow \infty} 0 \end{aligned}$$

for some $\alpha < 1/2$, where the last inequality comes from the fact that

$$\left\| (\boldsymbol{\rho}_N(\lfloor Nt \rfloor))^{-1} \otimes \boldsymbol{\rho}_N(Nt) \right\| \leq c(Nt - \lfloor Nt \rfloor)^\alpha, \quad \left\| \boldsymbol{\rho}_N(\lfloor Nt \rfloor)^{-1} \otimes \boldsymbol{\rho}_N(\lfloor Nt \rfloor + 1) \right\| \leq c.$$

By Slutsky's theorem, this gives us pointwise convergence of $\delta_{N^{-1/2}}(\boldsymbol{\rho}_N(Nt))_t$ which achieves the proof. \square

Proposition 3.3.3 shows the interest of theorem 3.2.3: rather than operating with concatenation of continuous paths in each particular case, we can manipulate a more general object in $G^2(V)$. The advantage is a general result for an equivalence class and a more convenient setting for computations.

Embeddings for differential and difference equations. As it has already been mentioned, given a differential equation $dY_t = f(Y_t)dX_t$, its solution map $\xi(Y_0, X) = Y$ is not always continuous in uniform topology, which means that we have to choose an appropriate sequence of continuous paths $(X^{(n)})_n$ approaching X in order for $\xi(Y_0, X^{(n)}) = Y^{(n)}$ to suitably approach Y . The rough paths theory gives us a universal recipe for making this choice: a suitable sequence $(X^{(n)})_n$ is such that the corresponding sequence of rough paths $(\mathbb{X}^{(n)})_n$ converges to the rough path \mathbb{Y} of Y in the rough path topology.

Even if the rough paths theory deals only with continuous objects, the above statement has surprising and important consequences in the discrete setting. A crucial difference is that, in the continuous case, the limit is given, and we have to find a suitable sequence that convergence towards it. On the contrary, if we have a

sequence of difference equations, the limit is not known a priori and thus will be function of the embedding. Otherwise said, *in the first case we have a given limit and several possible approximations, whereas in the second one we have a given discretization and several possible limits.*

Let us now see how this fact applies to the analysis of the discrete analogue of differential equations, *the difference equations*. While solving these equations directly can have useful applications, it is also of great interest to consider their convergence after renormalization, i.e. the convergence of

$$\Delta Y_n = \epsilon \sum_{i=1}^d f^{(i)}(Y_{n-1})(\Delta X_n)^{(i)} \quad (3.38)$$

such that $\epsilon \rightarrow 0$. This is where that the embedding comes on the scene.

Property 3.3.4. *Consider a difference equation as in (3.38) and suppose that $(X_n)_n$ allows two different embeddings of finite variation ρ_N and ρ'_N . Then, if ρ_N and ρ'_N are equivalent in the rough path sense, they define the same limit equation in the rough path topology (if it exists).*

In particular, it means that, in the case of rough paths in $G^2(V)$, the difference between the corresponding limit equations will depend on the second-level component of the rough path, which contains *the area anomaly and the square bracket of the process*. Let us now see how this translates in the setting of the classical stochastic calculus in the case where $(X_n)_n$ is a hidden Markov walk.

In our article [64], we have discussed the convergence of a difference equation driven by the sum of the Bernoulli "turning" variables. In that example, we have implicitly supposed a piecewise linear, geodesic embedding. What happens if we change it? The additional term we obtain at the limit by choosing a hidden Markov walk as driving process is a drift consisting of two components. The first one, featuring the constant K , depends on the square bracket of the process, i.e. is a common term in stochastic calculus. The second one, containing the area anomaly Γ , is a new term that can only be brought up by means of rough path analysis. However, *both of them depend on the embedding*: the first one due to the construction of the square bracket as a limit, the second one due to the fact that different embeddings may generate different stochastic area anomalies at the limit,

as it has been mentioned earlier. Thus, the choice of an embedding is a problem that has consequences on the classical stochastic calculus even independently of the rough paths theory.

3.3.4 Getting back to theorem 3.2.3

Notations: For $g \in G^2(V)$, we denote by $g^{(1)}$ its first level component and by $g^{(2)}$ the second-level one. We denote by δ is the standard dilatation operator on $G^2(V)$ (i.e. $\delta_\epsilon(g^{(1)}, g^{(2)}) = (\epsilon g^{(1)}, \epsilon^2 g^{(2)})$).

We start with two preliminary lemmas.

Lemma 3.3.1. *Let $(\xi_n)_n$ be a sequence of i.i.d. $G^2(V)$ -valued centred random variables with bounded moments of all orders, i.e.*

$$\forall p \geq 1, \mathbb{E}[|\xi_n|^p] < \infty$$

Furthermore, let $(k_n)_n$ be a sequence of \mathbb{N} -valued r.v. such that

- $k_0 = 0$ a.s.
- $\mathbb{P}(\forall n \geq 0, k_{n+1} \in \{k_n, k_n + 1\}) = 1$
- $\frac{k_n}{n} \xrightarrow[n \rightarrow \infty]{} a \in \mathbb{R}_+^*$ a.s.

Set $\Xi_n = \bigotimes_{k=1}^{k_n} \xi_k$ and for $t \in [0, 1]$, $\Xi_t^n = \Xi_{[nt]} \otimes \delta_{nt-[nt]}(\xi_{k_{[nt]+1}})$. We then have the following convergence:

$$\left(\delta_{(n\sigma^2 a)^{-1/2}}(\Xi_t^n) \right)_{t \in [0,1]} \xrightarrow[n \rightarrow \infty]{(d)} (\mathbb{B}_t)_{t \in [0,1]}$$

where $(\mathbb{B}_t)_{t \in [0,1]}$ is the enhanced Brownian motion, and $\sigma^2 = \mathbb{E}[|\xi_0^{(1)}|^2]$.

Remark. The difference with the Donsker-type theorem from [12] is that the sums of i.i.d. r.v. are a function of k_n and not simply of n .

Proof. We will proceed by the classical method, proving first the convergence of the finite-dimensional marginals of the process and then its tightness.

For the first part, we use the result from in [12], which states that a (renormalized) sum of i.i.d.r.v. in $G^2(V)$ converges in law to the Brownian rough path. Since k_n/n converges a.s. and that, moreover, $(k_n)_n$ is increasing a.s., a sum of k_n i.i.d.r.v. yields the same type of convergence, but with a change in time depending on a . Otherwise said, for $t \in [0, 1]$, we have

$$\delta_{(\sigma^2 n)^{-1/2}} \Xi_t^n = \delta_{(k_n/n)^{1/2}} \delta_{(\sigma^2 k_n)^{-1/2}} \Xi_t^n \xrightarrow[n \rightarrow \infty]{(d)} \sqrt{a} \mathbb{B}_t$$

and we can then generalize this convergence to any finite-dimensional marginals using the independence of the variables and Slutsky's theorem (theorem 3.1 in [9]) as in the classical Donsker theorem.

In order to prove the tightness of the process in $\mathcal{C}^\alpha([0, 1], G^2(V))$ for $\alpha < 1/2$, we can use the Kolmogorov's criterion (as already exposed in the proof of theorem 1.1 from [64]), i.e. we have to prove that, for any $p > 1$, there exists $c > 0$ such that

$$\mathbb{E} \left[n^{-2p} \mathbf{d}(\Xi_t^n, \Xi_s^n)^{4p} \right] \leq c |t - s|^{2p}$$

for any $s, t \in [0, 1]$. The definition of $\mathbf{d}(\cdot, \cdot)$ implies that

$$\mathbf{d}(\Xi_t^n, \Xi_s^n) = \left\| \delta_{n s - \lfloor n s \rfloor} \xi_{k_{\lfloor n s \rfloor + 1}}^{-1} \otimes \Xi_{\lfloor n s \rfloor}^{-1} \otimes \Xi_{\lfloor n t \rfloor} \otimes \delta_{n t - \lfloor n t \rfloor} \xi_{k_{\lfloor n t \rfloor + 1}} \right\|. \quad (3.39)$$

If $s, t \in [k/n, (k+1)/n[$, then this expression becomes

$$\mathbf{d}(\Xi_t^n, \Xi_s^n) = n(t - s) \left\| \xi_{k_{\lfloor n s \rfloor + 1}} \right\|$$

which further on gives us

$$\mathbb{E} \left[n^{-2p} \mathbf{d}(\Xi_t^n, \Xi_s^n)^{4p} \right] \leq M_p n^{2p} |t - s|^{4p} \leq M_p |t - s|^{2p}$$

where M_p is such that $\mathbb{E} \left[\|\xi_0\|^{4p} \right] \leq M_p$.

If $s \in [i/n, (i+1)/n[$ and $t \in [(i+1)/n, (i+2)/n[$, the expression (3.39) becomes:

$$\begin{aligned} \mathbf{d}(\Xi_t^n, \Xi_s^n) &= \left\| \delta_{ns - \lfloor ns \rfloor} \xi_{k_{\lfloor ns \rfloor + 1}}^{-1} \otimes \Xi_{\lfloor ns \rfloor}^{-1} \otimes \Xi_{\lfloor ns \rfloor + 1} \otimes \delta_{nt - \lfloor nt \rfloor} \xi_{k_{\lfloor nt \rfloor + 1}} \right\| \\ &= \left\| \delta_{1 - (ns - \lfloor ns \rfloor)} \xi_{k_{\lfloor ns \rfloor + 1}} \otimes \delta_{nt - \lfloor nt \rfloor} \xi_{k_{\lfloor nt \rfloor + 1}} \right\| \\ &= (\lfloor nt \rfloor - ns) \left\| \xi_{k_{\lfloor ns \rfloor + 1}} \right\| + (nt - \lfloor nt \rfloor) \left\| \xi_{k_{\lfloor nt \rfloor + 1}} \right\| \\ &\leq 2n(t - s) \left(\left\| \xi_{k_{\lfloor ns \rfloor + 1}} \right\| \right) \end{aligned}$$

and we can conclude as in the first case.

Finally, for $s \in [i/n, (i+1)/n[$ and $t \in [(i+l)/n, (i+l+1)/n[$ with $l \geq 2$, we use the properties of $\mathbf{d}(\cdot, \cdot)$ to get

$$\mathbf{d}(\Xi_t^n, \Xi_s^n)^{4p} \leq 2^{4p} \left(\left\| \delta_{1 - (ns - \lfloor ns \rfloor)} \xi_{k_{\lfloor ns \rfloor + 1}} \right\|^{4p} + \left\| \bigotimes_{j=k_{\lfloor ns \rfloor + 2}}^{k_{\lfloor nt \rfloor}} \xi_j \right\|^{4p} + \left\| \delta_{nt - \lfloor nt \rfloor} \xi_{k_{\lfloor nt \rfloor + 1}} \right\|^{4p} \right)$$

and consequently

$$\mathbb{E} \left[n^{-2p} \mathbf{d}(\Xi_t^n, \Xi_s^n)^{4p} \right] \leq 2^{4p} \left(2\mathbb{E} \left[n^{-2p} \|\xi_0\|^{4p} \right] + \mathbb{E} \left[\left\| n^{-2p} \bigotimes_{j=k_{\lfloor ns \rfloor + 2}}^{k_{\lfloor nt \rfloor}} \xi_j \right\|^{4p} \right] \right).$$

The first part of the right-hand side inequality can be bounded using the fact that there exists $M_p > 0$ such that $\mathbb{E} \left[n^{-2p} \|\xi_0\|^{4p} \right] \leq M_p n^{-2p} \leq M_p |t - s|^{2p}$, as $|t - s| > 1/n$. We transform the second one using the independence of variables, and then what is left to prove is that there exists $c' > 0$ such that

$$\begin{aligned} \mathbb{E} \left[n^{-2p} \left\| \bigotimes_{j=k_{\lfloor ns \rfloor + 2}}^{k_{\lfloor nt \rfloor}} \xi_j \right\|^{4p} \right] &= \mathbb{E} \left[n^{-2p} \left\| \bigotimes_{j=1}^{k_{\lfloor nt \rfloor} - k_{\lfloor ns \rfloor} - 1} \xi_j \right\|^{4p} \right] \\ &\leq c' |t - s|^{2p}. \end{aligned}$$

We use the result $\mathbb{E} \left[\left\| \bigotimes_{k=1}^m \xi_k \right\|^{4p} \right] = O(m^{2p})$ proven in [12] and the inequality

$k_{j+l} - k_j \leq l$ to get:

$$\begin{aligned} \mathbb{E} \left[\left\| \left(\bigotimes_{j=1}^{k_{\lfloor nt \rfloor} - k_{\lfloor ns \rfloor} - 1} \xi_j \right) \right\|^{4p} \right] &= \mathbb{E} \left[\left\| \left(\bigotimes_{j=1}^{k_{i+l} - k_i - 1} \xi_j \right) \right\|^{4p} \right] \\ &= O((k_{i+l} - k_i)^{2p}) = O((l-1)^{2p}) = O((n|t-s|)^{2p}). \end{aligned}$$

We thus get a uniform bound, which achieves the proof. \square

Lemma 3.3.2. *Consider $(k_n)_n$ a sequence of r.v. as the one in lemma 3.3.1. Let $((0, C_n))_n$ be a sequence of uniformly bounded r.v. taking values in the centre of the group $G^2(V)$ and such that we have the following a.s. convergence for any $t \in [0, 1]$:*

$$\delta_{n-1/2} \bigotimes_{i=1}^{k_{\lfloor nt \rfloor}} (0, C_i) \xrightarrow[n \rightarrow \infty]{} (0, atM) \quad (3.40)$$

where M is a deterministic matrix and $a \in \mathbb{R}_+^*$ is as in lemma 3.3.1. Then we have the following convergence in probability

$$\left(\delta_{n-1/2} \bigotimes_{i=1}^{k_{\lfloor nt \rfloor}} (0, C_i) \right)_{t \in [0,1]} \xrightarrow[n \rightarrow \infty]{} (0, atM)_{t \in [0,1]}$$

in the rough path topology $\mathcal{C}^{\alpha-H\ddot{o}l}([0, 1], G^2(V))$ for $\alpha < 1/2$.

Proof. We will use here proposition 3.3.2. The pointwise convergence is here a hypothesis of the lemma, so what is left to prove is the uniform bound for the Carnot-Caratheodory norm. On one hand, the upper bound for the Carnot-Caratheodory norm deduced from the norm equivalence (3.35) implies that

$$\forall (0, y) \in G^2(V), \quad \|(0, y)\| \leq |y|_{V \otimes V}^{1/2}$$

Using the fact that the sequence $(k_n)_n$ is such that, for $m \leq n$, $k_n - k_m \leq n - m$, we obtain

$$\left\| \bigotimes_{i=k_{\lfloor ns \rfloor} + 1}^{k_{\lfloor nt \rfloor}} (0, C_i) \right\|^2 = \left\| \left(0, \sum_{i=k_{\lfloor ns \rfloor} + 1}^{k_{\lfloor nt \rfloor}} C_i \right) \right\|^2 \leq \left| \sum_{i=k_{\lfloor ns \rfloor} + 1}^{k_{\lfloor nt \rfloor}} C_i \right|_{V \otimes V} \leq (\lfloor nt \rfloor - \lfloor ns \rfloor) \xi$$

where ξ is such that for all $i \in \mathbb{N}$, $|C_i|_{V \otimes V} \leq \xi$. We then have:

$$\forall n \geq 1, \mathbb{E} \left[\sup_{s \neq t} \frac{\left\| \delta_{n^{-1/2}} \otimes_{i=k_{[ns]}+1}^{k_{[nt]}} (0, C_i) \right\| \right] \leq \mathbb{E} \left[\sup_{s \neq t} \frac{([nt] - [ns])^{1/2} \xi^{1/2}}{(n|t-s|)^{1/2}} \right]$$

which is uniformly bounded as the $(0, C_i)$ s are supposed to be uniformly bounded. \square

We will now give the proof of theorem 3.2.3. Before doing so, we will give a slightly different formulation of it and explain why we do so. As before, E denotes a finite state space and V a finite-dimensional vector space.

Theorem 1 (theorem 3.2.3 for non-centred variables). *Let $(R_n, \mathbb{X}_n)_n$ be a hidden Markov walk on $E \times G^2(V)$ that satisfies condition (3.13). As in property 3.3.1, we can decompose*

$$\forall n \geq 1, \mathbb{X}_n = \bigotimes_{k=1}^n F_k$$

where $F_k = \mathbb{X}_{k-1}^{-1} \otimes \mathbb{X}_k$.

Furthermore, set $\beta = \mathbb{E}[T_1]$, $v = \beta^{-1} \mathbb{E}[\mathbb{X}_{T_1}^{(1)}]$. Since the F_k s are in $G^2(V)$, $F_k = (F_k^{(1)}, 1/2(F_k^{(1)} \otimes F_k^{(1)} + a_k))$, where a_k is the antisymmetric part of the second component. We then set $\tilde{F}_k = (F_k^{(1)} - v, 1/2(F_k^{(1)} - v) \otimes (F_k^{(1)} - v) + a_k)$ and $\tilde{\mathbb{X}}_n = \bigotimes_{k=1}^n \tilde{F}_k$.

In this case, the geodesic embedding of $(\tilde{\mathbb{X}}_n)_n$ in $G^2(V)$ will be

$$\forall t \in [0, 1], \tilde{\mathbb{X}}_t^n = \bigotimes_{k=1}^{[nt]} \tilde{F}_k \otimes \delta_{nt-[nt]} \tilde{F}_{[nt]+1}.$$

Let $\rho = \rho_{NN}$ be an embedding encoded by $(\tilde{\mathbb{X}}_n)_n$ as shown in proposition 3.3.3. If $\mathbb{X}_{T_1}^{(1)}$ is non-degenerate, we can suppose the covariance matrix of $\mathbb{X}_{T_1}^{(1)}$ is CI_d without loss of generality and we have the following convergence in $\mathcal{C}^\alpha([0, 1], G^2(V))$ for $\alpha < 1/2$:

$$\left(\delta_{(nC\beta^{-1})^{-1/2} \tilde{\mathbb{X}}_t^n} \right)_{t \in [0, 1]} \xrightarrow[n \rightarrow \infty]{(d)} \left(\mathbb{B}_t^{\text{Strat}} \otimes (0, t\Gamma_\rho) \right)_{t \in [0, 1]}$$

where $\mathbb{B}^{\text{Strat}}$ is the Brownian motion enhanced with second-level Stratonovich inte-

grals, i.e. as the second-level limit from theorem 3.2.3, and Γ_ρ is the deterministic antisymmetric matrix $\tilde{\Gamma}$ from theorem 3.2.3, i.e. Γ_ρ can be represented both as in formula (3.12) and as

$$\Gamma_\rho = C^{-1} \left(\mathbb{E} \left[\frac{1}{2} \sum_{1 \leq p < m \leq T_1} \tilde{F}_p^{(1)} \otimes \tilde{F}_m^{(1)} - \tilde{F}_m^{(1)} \otimes \tilde{F}_p^{(1)} \right] + \mathbb{E} \left[\sum_{p=1}^{T_1} a_p \right] \right). \quad (3.41)$$

Remarks.

- $(R_n, \tilde{X}_n)_n$ is a HMW in the sense of definition 3.2.2, and it satisfies the conditions from theorem 3.2.3: in particular, $\mathbb{E} [\tilde{X}_{T_1}^{(1)}] = 0$. Thus, the interest of the present version of theorem 3.2.3 resides in the fact that it allows to treat the case of more general HMW on $G^2(V)$ by recentering their excursions.
- If we compare the statement of the theorem 3.2.2 given in the introduction to the present one, a fundamental difference is that the first one was more analytic, whereas the present one is more algebraic. We have avoided the "heavy" rough path formulation in the introduction as the rough path setting appears only further in the paper (section 3.3.2).
- In particular, instead of presenting the first and second level as random processes in the uniform topology $(S_{B,i}(t))$, the limit is presented here as the rough path process $(\mathbb{B}_t^{Strat} \otimes (0, t\Gamma_\rho))_{t \in [0,1]}$. By doing so, we stress the area anomaly obtained at the limit and given by $(0, t\Gamma_\rho)$.

Proof. As already mentioned, since the F_k s are in $G^2(V)$, the particular form of the elements from this group stated in (3.5) gives the decomposition:

$$\forall k \geq 1, F_k^{(2)} = \frac{1}{2} F_k^{(1)} \otimes F_k^{(1)} + a_k \quad (3.42)$$

where a_k is the antisymmetric part of $F_k^{(2)}$, $\text{Antisym}(F_k^{(2)})$.

Thus, using (3.42), we have the following decomposition for the \tilde{F}_k s:

$$\forall k \geq 1, \tilde{F}_k = (\tilde{F}_k^{(1)}, \frac{1}{2} \tilde{F}_k^{(1)} \otimes \tilde{F}_k^{(1)}) \otimes (0, a_k) \quad (3.43)$$

where $(\tilde{F}_k^{(1)}, \frac{1}{2}\tilde{F}_k^{(1)} \otimes \tilde{F}_k^{(1)}) \in G^2(V)$ is an element with a symmetric second component and $(0, a_k) \in G^2(V)$ is a "pure area" element (i.e. such that the first component equal to zero and the second component antisymmetric) and is in the centre of the group.

We denote by $\kappa(n)$ the rank of the excursion to which n belongs, i.e. the unique integer such that $T_{\kappa(n)} \leq n < T_{\kappa(n)+1}$, where the T_i s are as defined in (3.6). Based on the fact that the elements of type $(0, a) \in G^2(V)$ are in the centre of the group and thus commute with all the others, we have the following decomposition of $\tilde{\mathbb{X}}_t^n$:

$$\begin{aligned}
\tilde{\mathbb{X}}_t^n &= \bigotimes_{k=1}^{\lfloor nt \rfloor} \left((\tilde{F}_k^{(1)}, \frac{1}{2}\tilde{F}_k^{(1)} \otimes \tilde{F}_k^{(1)}) \otimes (0, a_k) \right) \otimes \delta_{nt-\lfloor nt \rfloor} \tilde{F}_{\lfloor nt \rfloor+1} \\
&= \bigotimes_{k=1}^{\kappa(\lfloor nt \rfloor)} \left(\sum_{p=T_{k-1}+1}^{T_k} \tilde{F}_p^{(1)}, \frac{1}{2} \sum_{p=T_{k-1}+1}^{T_k} \tilde{F}_p^{(1)} \otimes \tilde{F}_p^{(1)} + \sum_{T_{k-1}+1 \leq p < m \leq T_k} \tilde{F}_p^{(1)} \otimes \tilde{F}_m^{(1)} \right) \\
&\quad \otimes \left(\bigotimes_{k=1}^{\kappa(\lfloor nt \rfloor)} (0, \sum_{p=T_{k-1}+1}^{T_k} a_p) \right) \otimes \left(\bigotimes_{k=T_{\kappa(\lfloor nt \rfloor)}+1}^{\lfloor nt \rfloor} \tilde{F}_k \right) \otimes \delta_{nt-\lfloor nt \rfloor} \tilde{F}_{\lfloor nt \rfloor+1} \\
&= \bigotimes_{k=1}^{\kappa(\lfloor nt \rfloor)} \left(\sum_{p=T_{k-1}+1}^{T_k} \tilde{F}_p^{(1)}, \frac{1}{2} \left(\sum_{p=T_{k-1}+1}^{T_k} \tilde{F}_p^{(1)} \right)^{\otimes 2} \right) \\
&\quad \otimes \left(\bigotimes_{k=1}^{\kappa(\lfloor nt \rfloor)} (0, \frac{1}{2} \sum_{T_{k-1}+1 \leq p < m \leq T_k} \tilde{F}_p^{(1)} \otimes \tilde{F}_m^{(1)} - \tilde{F}_m^{(1)} \otimes \tilde{F}_p^{(1)} + \sum_{p=T_{k-1}+1}^{T_k} a_p) \right) \\
&\quad \otimes \left(\bigotimes_{k=T_{\kappa(\lfloor nt \rfloor)}+1}^{\lfloor nt \rfloor} \tilde{F}_k \right) \otimes \delta_{nt-\lfloor nt \rfloor} \tilde{F}_{\lfloor nt \rfloor+1} \\
&= \mathcal{P}_t^n \otimes \mathcal{A}_t^n \otimes \mathcal{R}_t^n
\end{aligned}$$

where

$$\mathcal{P}_t^n = \bigotimes_{k=1}^{\kappa(\lfloor nt \rfloor)} \left(\sum_{p=T_{k-1}+1}^{T_k} \tilde{F}_p^{(1)}, \frac{1}{2} \left(\sum_{p=T_{k-1}+1}^{T_k} \tilde{F}_p^{(1)} \right)^{\otimes 2} \right)$$

is the term that concatenates the excursions,

$$\mathcal{A}_t^n = \bigotimes_{k=1}^{\kappa(\lfloor nt \rfloor)} \left(\bigotimes_{k=1}^{\kappa(\lfloor nt \rfloor)} (0, \frac{1}{2} \sum_{T_{k-1}+1 \leq p < m \leq T_k} \tilde{F}_p^{(1)} \otimes \tilde{F}_m^{(1)} - \tilde{F}_m^{(1)} \otimes \tilde{F}_p^{(1)} + \sum_{p=T_{k-1}+1}^{T_k} a_p) \right)$$

is a "pure area" process that takes into consideration the antisymmetric part of the

$F_k^{(2)}$ s, as well as the stochastic areas of the pseudo-excursions, and

$$\mathcal{R}_t^n = \left(\bigotimes_{k=T_{\kappa(\lfloor nt \rfloor)+1}}^{\lfloor nt \rfloor} \tilde{F}_k \right) \otimes \delta_{nt-\lfloor nt \rfloor} \tilde{F}_{\lfloor nt \rfloor+1}$$

is the rest left from the geodesic embedding.

We will now compute separately the limit of each of the three terms. Let us first consider the term $(\delta_{n^{-1/2}} \mathcal{P}_t^n)_{t \in [0,1]}$. By construction, the variables

$$\left(\sum_{p=T_{k-1}}^{T_k} \tilde{F}_p^{(1)}, \frac{1}{2} \left(\sum_{p=T_{k-1}+1}^{T_k} \tilde{F}_p^{(1)} \right)^{\otimes 2} \right)$$

are i.i.d. and centred. Moreover, $\kappa(\lfloor nt \rfloor)/n \xrightarrow[n \rightarrow \infty]{} t\beta^{-1}$ a.s., since $\kappa(n)$ is the number of full excursions accomplished until time n . Thus, since the function κ satisfies the conditions of lemma 3.3.1, we can deduce that:

$$\left(\delta_{(nC\beta^{-1})^{-1/2}} \bigotimes_{k=1}^{\kappa(\lfloor nt \rfloor)} \left(\sum_{p=T_{k-1}}^{T_k} \tilde{F}_p^{(1)}, \frac{1}{2} \left(\sum_{p=T_{k-1}+1}^{T_k} \tilde{F}_p^{(1)} \right)^{\otimes 2} \right) \right)_{t \in [0,1]} \xrightarrow[n \rightarrow \infty]{(d)} \left(\mathbb{B}_t^{Strat} \right)_{t \in [0,1]}$$

in the rough path topology.

We now have to study the convergence of $(\delta_{n^{-1/2}} \mathcal{A}_t^n)_{t \in [0,1]}$. We notice that the two tensor products containing $\kappa(\lfloor nt \rfloor)$ terms are "sums" of i.i.d.r.v., since each term depends entirely on a different excursion. The law of large numbers thus applies to both of them, and we get that, a.s., for $t \in [0, 1]$ fixed:

$$\delta_{(n\beta^{-1})^{-1/2}} \bigotimes_{k=1}^{\kappa(\lfloor nt \rfloor)} \left(0, \frac{1}{2} \sum_{T_{k-1}+1 \leq p < m \leq T_k} \tilde{F}_p^{(1)} \otimes \tilde{F}_m^{(1)} - \tilde{F}_m^{(1)} \otimes \tilde{F}_p^{(1)} \right) \xrightarrow[n \rightarrow \infty]{} (0, t\Gamma)$$

where

$$\Gamma = \mathbb{E} \left[\frac{1}{2} \sum_{1 \leq p < m \leq T_1} \tilde{F}_p^{(1)} \otimes \tilde{F}_m^{(1)} - \tilde{F}_m^{(1)} \otimes \tilde{F}_p^{(1)} \right]$$

is the area anomaly we recover.

For the second part of the sum, we have, for $t \in [0, 1]$ fixed:

$$\delta_{(n\beta^{-1})^{-1/2}} \bigotimes_{k=1}^{\kappa(\lfloor nt \rfloor)} (0, \sum_{p=T_{k-1}+1}^{T_k} a_p) \xrightarrow{n \rightarrow \infty} (0, t\Gamma_0)$$

where

$$\Gamma_0 = \mathbb{E} \left[\sum_{p=1}^{T_1} a_p \right]$$

is generated by the antisymmetric part of the second-level components of the F_k s.

Since κ is as the function from lemma 3.3.2, the functions \tilde{F}_p are a.s. uniformly bounded under condition (3.13) and T_1 has finite moments of all order, the conditions of lemma 3.3.2 are satisfied, and we deduce the following convergence in probability in rough path topology:

$$(\delta_{n^{-1/2}} \mathcal{A}_t^n)_{t \in [0,1]} \xrightarrow{n \rightarrow \infty} (0, t\beta^{-1}\Gamma_\rho)_{t \in [0,1]} \quad (3.44)$$

where $\Gamma_\rho = \Gamma + \Gamma_0$.

We have the residue $(\delta_{n^{-1/2}} \mathcal{R}_t^n)_{t \in [0,1]}$ left to deal with. Its first part contains $\lfloor nt \rfloor - T_{\kappa(\lfloor nt \rfloor)} \leq (T_{\kappa(\lfloor nt \rfloor)+1} - 1) - T_{\kappa(\lfloor nt \rfloor)}$ terms, and $\|\delta_{nt - \lfloor nt \rfloor} \tilde{F}_{\lfloor nt \rfloor + 1}\| \leq \|\tilde{F}_{\lfloor nt \rfloor + 1}\|$ adds one more term. Moreover, since by (3.13) the F_k s are uniformly bounded, there exists $K > 0$ such that $\sup_k \|\tilde{F}_k\| < K$. Therefore, we get:

$$\mathbb{E} \left[\sup_{s \neq t} \frac{\|\delta_{n^{-1/2}} ((\mathcal{R}_s^n)^{-1} \otimes \mathcal{R}_t^n)\|}{|t - s|^{1/2}} \right] \leq K \mathbb{E} \left[\sup_{s \neq t} \frac{T_{\kappa(\lfloor nt \rfloor)} - T_{\kappa(\lfloor ns \rfloor)} + \lfloor nt \rfloor - \lfloor ns \rfloor}{(n|t - s|)^{1/2}} \right].$$

This quantity is bounded since $\lfloor nt \rfloor = \lfloor ns \rfloor$ for $s, t \in [i/n, (i+1)/n]$ for any $i = 1, \dots, n-1$. We also have the convergence in probability

$$\forall t \in [0, 1], \delta_{n^{-1/2}}(\mathcal{R}_t^n) \xrightarrow{n \rightarrow \infty} 0.$$

We can thus use once again proposition 3.3.2 to conclude to the following convergence in probability in rough path topology:

$$(\delta_{n^{-1/2}}(\mathcal{R}_t^n))_{t \in [0,1]} \xrightarrow{n \rightarrow \infty} 0.$$

Finally, putting altogether the convergences of $(\delta_{n^{-1/2}}\mathcal{P}_t^n)_{t \in [0,1]}$, of $(\delta_{n^{-1/2}}\mathcal{A}_t^n)_{t \in [0,1]}$ and of $(\delta_{n^{-1/2}}\mathcal{R}_t^n)_{t \in [0,1]}$, our result follows from Slutsky's theorem:

$$\left(\delta_{(nC\beta^{-1})^{-1/2}}\mathcal{P}_t^n \otimes \mathcal{A}_t^n \otimes \mathcal{R}_t^n\right)_{t \in [0,1]} \xrightarrow[n \rightarrow \infty]{(d)} \left(\mathbb{B}_t^{Strat} \otimes (0, t\Gamma_\rho)\right)_{t \in [0,1]}.$$

□

3.4 Iterated structures behind discrete time and discrete space Markov chains.

In this section we study more thoroughly the algebraic properties of HMW through the corresponding iterated occupation times.

3.4.1 Shuffle and quasi-shuffle products.

The definitions and properties from this section are mainly inspired from [47].

Definitions.

Definition 3.4.1 (the quasi-shuffle product). *Let (A, \cdot) be an algebra, $a, b \in A$, and x, y obtained by concatenation of a finite number of elements from A . Then the quasi-shuffle product $*$ is defined recursively by:*

$$ax * by = [a, b](x * y) + a(x * by) + b(ax * y)$$

where $[\cdot, \cdot]$ is a commutative and associative operation on A .

The shuffle product can be viewed as a particular case of the quasi-shuffle product. More specifically, it is the case when $[a, b] = 0$ for any $a, b \in A$. Contrary to the shuffle product, the quasi-shuffle product is not always commutative.

Definition 3.4.2 (shuffle product). *Let (A, \cdot) be an algebra, $a, b \in A$, and x, y obtained by concatenation of a finite number of elements from A . Then the shuffle*

product \sqcup is defined recursively by:

$$ax \sqcup by = a(x \sqcup by) + b(ax \sqcup y).$$

A more informal way of putting it is as follows.

Definition 3.4.3 (and corollary of 3.4.2). *Consider the set of permutation $\Sigma_{m,n} = \{\sigma \in \Sigma_{m+n} : \sigma(1) < \dots < \sigma(m), \sigma(m+1) < \dots < \sigma(m+n)\}$. For $x = x_1 \dots x_m$ and $y = y_1 \dots y_n$ with $x_i, y_j \in A$, the shuffle product of x and y is given by the sum of images of the concatenated word ab under all the permutations of $S_{m,n}$: $x \sqcup y = \sum_{\sigma \in \Sigma_{m,n}} \sigma(xy)$ (for example, $ab \sqcup c = abc + acb + cab$).*

Quasi-shuffle structure of iterated sums. One of the main example of sets we can endow with a shuffle product is the set of iterated integrals on a vector space V . Denote by $\mathbf{e}_n^{i_1, \dots, i_n} = e_{i_1} \otimes \dots \otimes e_{i_n}$ and $\mathbf{e}_m^{j_1, \dots, j_m} = e_{j_1} \otimes \dots \otimes e_{j_m}$ where the e_k s are vectors from the canonical basis of V . We then define, for an integrable path y in V ,

$$\begin{aligned} S_{y,n;i_1, \dots, i_n}(t) &= \langle \mathbf{e}_n^{i_1, \dots, i_n}, S_{y,n}(t) \rangle \\ &= \int_{0 < t_1 < \dots < t_n < t} dy_{t_1}^{i_1} \dots dy_{t_n}^{i_n} \end{aligned}$$

and idem for $\mathbf{e}_m^{j_1, \dots, j_m}$, with $\langle \cdot, \cdot \rangle$ the induced scalar product on $V^{\otimes l}$. The shuffle product of two iterated integrals depending both on the same path y is then given by:

$$S_{y,n;i_1, \dots, i_n}(t) S_{y,m;j_1, \dots, j_m}(t) = S_{y,n+m;(i_1, \dots, i_n) \sqcup (j_1, \dots, j_m)}(t) \quad (3.45)$$

showing that the shuffle product of two iterated integrals is a linear combination of iterated integrals.

If we now consider a sequence $x = (x_n)_n$ in V , the components of the iterated

sum $\tilde{S}_{x,n}(N)$ from (3.24) given by

$$\begin{aligned} \tilde{S}_{x,n;i_1,\dots,i_n}(N) &= \langle \mathbf{e}_{\mathbf{n}}^{i_1,\dots,i_n}, \tilde{S}_{x,n}(N) \rangle \\ &= \sum_{1 \leq k_1 < \dots < k_n \leq N} x_{k_1}^{i_1} \dots x_{k_n}^{i_n} \end{aligned} \quad (3.46)$$

are discrete versions of (3.45).

However, the product of two elements of this kind is not a shuffle product anymore, as shown below.

$$\begin{aligned} \tilde{S}_{x,n;i_1,\dots,i_n}(N) \tilde{S}_{x,m;j_1,\dots,j_m}(N) &= \tilde{S}_{x,n+m;(i_1,\dots,i_n)\sqcup(j_1,\dots,j_m)}(N) \\ &\quad + r_{x,n,m;i_1,\dots,i_n;j_1,\dots,j_m}(N) \end{aligned} \quad (3.47)$$

where $r_{x,n,m;i_1,\dots,i_n;j_1,\dots,j_m}(N)$ is a rest that comes from the fact that, when multiplying two iterated sums, we get sets of indices that are not necessarily strictly ordered. Moreover, the result of the multiplication can not be in the form of a function \tilde{S} so we can not apply the new product exclusively to $\mathbf{e}_{\mathbf{n}}^{i_1,\dots,i_n}$ and $\mathbf{e}_{\mathbf{m}}^{j_1,\dots,j_m}$. We thus need a more general product, that would better keep track of this rest. The product of two iterated sums of the type $S_{x,n;i_1,\dots,i_n}(N)$ is a quasi-shuffle product, as already stated in property 3.2.2.

3.4.2 From geometric to non-geometric rough paths through hidden Markov walks

Geometric rough paths and shuffle products.

We know that a piecewise linear embedding of $(X_n)_n$ is canonically represented by a (geometric) rough path as follows.

Property 3.4.1. *Denote by $(X_t^{(N)})_t$ a smooth embedding of $(X_n)_n$. The canonical rough path in $G^2(V)$ corresponding to this embedding is given by the 2-step signature*

$$\forall t \in [0, 1], S_2(X^{(N)})_{0,t} = (X_t^{(N)}, \int_{0 < s_1 < s_2 < t} dX_{s_1}^{(N)} \otimes dX_{s_2}^{(N)}).$$

The path $t \mapsto S_2(X^{(N)})_{0,t}$ is in $\mathcal{C}^{1-var}([0, 1], G^2(V))$.

We thus obtain a sequence of geometric rough paths and can study the convergence in law of $(\delta_{N^{-1/2}} S_2(X^{(N)})_{0,t})_t$ in the topology of $\mathcal{C}^{\alpha-H\ddot{o}l}([0, 1], G^2(V))$ for $\alpha < 1/2$. The limit (if it exists) will be a *geometric α -Hölder rough path*, i.e. an element of $\mathcal{C}^{\alpha-H\ddot{o}l}([0, 1], G^2(V))$, for $\alpha < 1/2$.

Since $G^2(V)$ is the space in which geometric rough paths take values, an element $(g^{(1)}, g^{(2)}) \in G^2(V)$ needs to satisfy the shuffle product relation in the sense of the following lemma.

Property 3.4.2. *An element $(g^{(1)}, g^{(2)}) \in T_1^{(2)}(V)$ is in $G^2(V)$ if and only if it satisfies the relation*

$$\langle g^{(1)}, e_i \rangle \langle g^{(1)}, e_j \rangle = \langle g^{(2)}, e_i \sqcup e_j \rangle \quad (3.48)$$

where e_i, e_j are in the canonical basis of V and $e_i \sqcup e_j = e_i \otimes e_j + e_j \otimes e_i$.

Proof. We can decompose the second component into a symmetric and an antisymmetric part :

$$g^{(2)} = \text{Sym}(g^{(2)}) + \text{Antisym}(g^{(2)}).$$

On one hand, we have

$$\begin{aligned} \langle \text{Sym}(g^{(2)}), e_i \sqcup e_j \rangle &= \frac{1}{2} \langle g^{(1)} \otimes g^{(1)}, e_i \otimes e_j + e_j \otimes e_i \rangle \\ &= \langle g^{(1)}, e_i \rangle \langle g^{(1)}, e_j \rangle . \end{aligned}$$

On the other hand, for the antisymmetric term we get

$$\langle \text{Antisym}(g^{(2)}), e_i \sqcup e_j \rangle = \langle \text{Antisym}(g^{(2)}), e_i \otimes e_j \rangle - \langle \text{Antisym}(g^{(2)}), e_j \otimes e_i \rangle = 0$$

which achieves the proof. □

Remark. The condition (3.48) is a version of the relation which is part of the abstract definition of geometric rough paths given in [43].

A discrete construction for non-geometric rough paths.

We now present a way of constructing non-geometric rough paths out of a hidden Markov walk $(R_n, X_n)_n$. Instead of first constructing a continuous path $(X_t^{(N)})_t$ in V and then associate a rough path to it, we will first construct a sequence $(\mathbb{Y}_n)_n$ in $T_1^{(2)}(V)$ and then associate an embedding to it.

Let us consider the sequence $(\mathbb{Y})_n$ as in (3.17), i.e.

$$\mathbb{Y}_n = \left(X_n, \sum_{1 \leq k_1 < k_2 \leq n} F_{k_1} \otimes F_{k_2} \right) \in T_1^{(2)}(V)$$

and explain how connect the non-geometric nature of the corresponding rough path to the quasi-shuffle product.

In the case of the sequence $(\mathbb{Y}_k)_k$, we have:

$$\begin{aligned} \langle X_n, e_i \rangle \langle X_n, e_j \rangle &= \sum_{k_1, k_2=1}^n F_{k_1}^{(i)} F_{k_2}^{(j)} \\ &= \sum_{1 \leq k_1 < k_2 \leq n} F_{k_1}^{(i)} F_{k_2}^{(j)} + \sum_{1 \leq k_2 < k_1 \leq n} F_{k_1}^{(i)} F_{k_2}^{(j)} + \sum_{k=1}^n F_k^{(i)} F_k^{(j)} \end{aligned} \quad (3.49)$$

$$= \langle \sum_{1 \leq k_1 < k_2 \leq n} F_{k_1} \otimes F_{k_2}, e_i \sqcup e_j \rangle + \langle \sum_{k=1}^n F_k^{\otimes 2}, e_i \otimes e_j \rangle \quad (3.50)$$

$$= \langle \sum_{1 \leq k_1 < k_2 \leq n} F_{k_1} \otimes F_{k_2} + \frac{1}{2} \sum_{k=1}^n F_k^{\otimes 2}, e_i \sqcup e_j \rangle .$$

Once again, we see that (3.50) does not coincide with (3.48). Moreover, if we identify e_i as corresponding to the index k_1 and e_j as corresponding to the index k_2 , the line (3.49) in the computation above can be identified with *the quasi-shuffle product*: the sums over $\{k_1 < k_2\} \cup \{k_2 < k_1\}$ give the shuffle product part, whereas the last sum over $\{k_1 = k_2\}$ gives the quasi-shuffle rest.

From all the above, we can draw the following conclusion:

Property 3.4.3. *The sequence $(\mathbb{Y}_n)_n$ takes values in $T_1^{(2)}(V) \setminus G^2(V)$. In particular, any embedding $(\mathbb{Y}_t^{(N)})_t$ of $(\mathbb{Y}_n)_n$, no matter how smooth, will be a non-geometric*

rough path.

3.4.3 Getting back to proposition 3.2.1.

We want to study the convergence of (3.17) and the convergence of the rough paths $(\mathbb{Y}_t^{(N)})_{t \in [0,1]}$ which are obtained from the sequence $(\mathbb{Y}_n)_n$ by a continuous embedding in the sense of definition 3.3.4.

The first thing we need to notice is that the sequence $(\mathbb{Y}_n)_n$ is not in $G^2(V)$. In the present case, we have

$$\begin{aligned} \text{Sym} \left(\sum_{1 \leq k_1 < k_2 \leq n} F_{k_1} \otimes F_{k_2} \right) &= \frac{1}{2} \left(\sum_{1 \leq k_1 < k_2 \leq n} F_{k_1} \otimes F_{k_2} + \sum_{1 \leq k_1 < k_2 \leq n} F_{k_2} \otimes F_{k_1} \right) \\ &= \frac{1}{2} X_n \otimes X_n - \frac{1}{2} \sum_{k=1}^n F_k^{\otimes 2} \end{aligned}$$

and thus we have an extra term that does not depend on the first level X_n .

In order to study the convergence of the non-geometric rough paths $(\mathbb{Y}_t^{(N)})_t$ associated to $(\mathbb{Y}_n)_n$ in property 3.4.3, we need to answer the following questions:

- What embedding do we choose for $(\mathbb{Y}_n)_n$?
- What is the topology to consider when studying the convergence?
- What will be the nature of the limit rough path (if it exists)?

We start with a few preliminary results. First, we want to show that property 3.4.1 can be used to construct a suitable embedding for $(\mathbb{X}_n)_n$. This is a consequence of the following lemma:

Lemma 3.4.1. *Choose $(X_t^{(N)})_t$ to be the geodesic embedding of $(X_n)_n$ and suppose $X_0 = 0$ a.s. For all $N \geq 1$ and $1 \leq i \leq n$, we have:*

$$S_2(X^{(N)})_{0,i/N} = \mathbb{Y}_i \otimes \left(0, \sum_{k=1}^i F_k^{\otimes 2} \right).$$

Proof. For any $N \geq 1$ and $t \in [0, 1]$, we can decompose the second component of $S_2(X^{(N)})_{0,t}$ into a symmetric and an antisymmetric part as follows:

$$\begin{aligned} \int_{0 < s_1 < s_2 < t} dX_{s_1}^{(N)} \otimes dX_{s_2}^{(N)} &= \frac{1}{2} \int_{0 < s_1 < s_2 < t} dX_{s_1}^{(N)} \otimes dX_{s_2}^{(N)} + dX_{s_2}^{(N)} \otimes dX_{s_1}^{(N)} \\ &\quad + \frac{1}{2} \int_{0 < s_1 < s_2 < t} dX_{s_1}^{(N)} \otimes dX_{s_2}^{(N)} - dX_{s_2}^{(N)} \otimes dX_{s_1}^{(N)} \end{aligned}$$

Since $(X_t^{(N)})_t$ is piecewise linear, the symmetric part becomes

$$\frac{1}{2} \int_{0 < s_1 < s_2 < t} dX_{s_1}^{(N)} \otimes dX_{s_2}^{(N)} + dX_{s_2}^{(N)} \otimes dX_{s_1}^{(N)} = (X_t^{(N)})^{\otimes 2}.$$

The antisymmetric part corresponds to the stochastic area of the process $(X_t^{(N)})_t$, $A_{i/N}(X^{(N)})$, which, for a piecewise linear process, can be expressed by second iterated sums, i.e.

$$\begin{aligned} \forall i \in \{1, \dots, N\}, A_i(X) &:= A_{i/N}(X^{(N)}) \\ &= \frac{1}{2} \int_{0 < s_1 < s_2 < i/N} dX_{s_1}^{(N)} \otimes dX_{s_2}^{(N)} - dX_{s_2}^{(N)} \otimes dX_{s_1}^{(N)} \\ &= \frac{1}{2} \sum_{1 \leq k_1 < k_2 \leq i} F_{k_1} \otimes F_{k_2} - F_{k_2} \otimes F_{k_1}. \end{aligned}$$

Since we have

$$\forall i \in \{1, \dots, N\}, \sum_{1 \leq k_1 < k_2 \leq i} F_{k_1} \otimes F_{k_2} + F_{k_2} \otimes F_{k_1} = (X_{i/N}^{(N)})^{\otimes 2} - \sum_{k=1}^i F_k^{\otimes 2}$$

it is straightforward to show

$$\sum_{1 \leq k_1 < k_2 \leq i} F_{k_1} \otimes F_{k_2} = \int_{0 < s_1 < s_2 < i/N} dX_{s_1}^{(N)} \otimes dX_{s_2}^{(N)} - \sum_{k=1}^i F_k^{\otimes 2}. \quad (3.51)$$

We thus conclude that:

$$\forall N \geq 1, \forall 1 \leq i \leq N, S_2(X^{(N)})_{0,i/N} = \left(X_i, \sum_{1 \leq k_1 < k_2 \leq i} F_{k_1} \otimes F_{k_2} \right) \otimes \left(0, \sum_{k=1}^i F_k^{\otimes 2} \right).$$

□

We can now give an answer to the two questions concerning the choice of the embedding and of the topology:

Lemma 3.4.2. *Choose $(X_t^{(N)})_t$ to be the geodesic embedding of $(X_n)_n$ and suppose $X_0 = 0$ a.s. Consider the geodesic embedding of $Z_k = \sum_{j=1}^k F_j^{\otimes 2}$ given by*

$$\forall N \in \mathbb{N}^*, \forall t \in [0, 1], Z_t^{(N)} = \sum_{k=1}^{\lfloor Nt \rfloor} F_k^{\otimes 2} + (Nt - \lfloor Nt \rfloor)^2 F_{\lfloor Nt \rfloor + 1}^{\otimes 2}$$

and set

$$\forall N \in \mathbb{N}^*, \forall t \in [0, 1], \mathbb{Y}_t^{(N)} = (0, Z_t^{(N)})^{-1} \otimes S_2(X^{(N)})_{0,t}.$$

Then $t \mapsto \mathbb{Y}_t^{(N)}$ is an embedding for $(\mathbb{Y}_n)_n$. Furthermore, $t \mapsto \mathbb{Y}_t^{(N)}$ is in $\mathcal{C}^{1-var}([0, 1], T_1^{(2)}(V))$ and is a non-geometric rough path.

Proof. The path $t \mapsto \mathbb{Y}_t^{(N)}$ is constructed by concatenation of two rough paths, $(S_2(X^{(N)})_{0,t})_t$ being a smooth rough path and $(0, Z_t^{(N)})_t$ a smooth path in $T_1^{(2)}(V)$, which implies that $t \mapsto \mathbb{Y}_t^{(N)}$ is an element of $\mathcal{C}^{1-var}([0, 1], T_1^{(2)}(V))$. Moreover, property 3.4.3 implies that $t \mapsto \mathbb{Y}_t^{(N)}$ is a non-geometric rough path.

The fact that $t \mapsto \mathbb{Y}_t^{(N)}$ is an embedding for $(\mathbb{X}_n)_n$ follows from lemma 3.4.1, which tells us that:

$$\forall N \geq 1, \forall 1 \leq i \leq N, \mathbb{Y}_{i/N}^{(N)} = (0, Y_{i/N}^{(N)})^{-1} \otimes S_2(X^{(N)})_{0,i/N}.$$

□

Remark. Another way of expressing $t \mapsto \mathbb{Y}_t^{(N)}$ is through *the rough path bracket* which concentrates the symmetric part of the second-level component (see definition 5.5 from [32]), and implies:

$$\forall t \in [0, 1], \mathbb{Y}_t^{(N)} = (X_t^{(N)}, A_t(X^{(N)})) \otimes (0, \frac{1}{2}[\mathbb{Y}^{(N)}]_t)$$

where

$$\forall t \in [0, 1], [\mathbb{Y}^{(N)}]_t = X_t^{(N)} \otimes X_t^{(N)} - Z_t^{(N)}$$

and $t \mapsto A_t(X^{(N)})$ is, as before, the stochastic area of $t \mapsto X_t^{(N)}$.

We can now proceed to the main proof of proposition 3.2.1.

Proof. Theorem 3.2.1 implies the following convergence in law in the space $\mathcal{C}^{\alpha-H\ddot{o}l}([0, 1], G^2(V))$ for $\alpha < 1/2$:

$$\left(\delta_{(NC\beta^{-1})^{-1/2}}S_2(X^{(N)})_{0,t}\right)_{t \in [0,1]} \xrightarrow{N \rightarrow \infty} \left(\mathbb{B}_t^{Strat} \otimes (0, \Gamma t)\right)_{t \in [0,1]} \quad (3.52)$$

where Γ is the area anomaly from 3.2.1.

Next, we apply the decomposition in pseudo-excursions to Z_n i.e.

$$\forall n \geq 1, Z_n = \sum_{j=0}^{\kappa(n)-1} \sum_{i=T_j+1}^{T_{j+1}} F_i^{\otimes 2} + \sum_{i=T_{\kappa(n)}+1}^n F_i^{\otimes 2}$$

and, by the law of large numbers, we deduce the convergence of the finite-dimensional marginals

$$\forall t \in [0, 1], \delta_{(NC\beta^{-1})^{-1/2}}(0, Z_t^{(N)}) \xrightarrow{N \rightarrow \infty} \left(0, tC^{-1}\mathbb{E}\left[\sum_{i=1}^{T_1} F_i^{\otimes 2}\right]\right).$$

We can now prove the tightness of the sequence of processes $(0, Z_t^{(N)})_t$ in $\mathcal{C}^{\alpha-H\ddot{o}l}([0, 1], T_1^{(2)}(V))$ for $\alpha < 1/2$, by using a version of the Kolmogorov criterion. Theorem 3.10 from [32] implies that it is enough to prove

$$\exists K > 0, \forall s, t \in [0, 1], \forall N \geq 1, N^{-1}\mathbb{E}\left[|Z_{s,t}^{(N)}|_{V \otimes V}\right] \leq K|t - s|.$$

We use the fact that the F_k s are uniformly bounded:

$$\begin{aligned} N^{-1}\mathbb{E}\left[|Z_{s,t}^{(N)}|_{V \otimes V}\right] &= N^{-1}\mathbb{E}\left[\left|\sum_{i=[Ns]+1}^{[Nt]} F_i^{\otimes 2}\right|_{V \otimes V}\right] \\ &\leq K|t - s| \end{aligned}$$

where K comes from the condition (3.8).

We thus conclude to the following convergence in probability in the $\mathcal{C}^{\alpha-H\ddot{o}l}([0, 1], T_1^{(2)}(V))$

topology for $\alpha < 1/2$:

$$\left(\delta_{(NC\beta^{-1})^{-1/2}}(0, Z_t^{(N)})\right)_{t \in [0,1]} \xrightarrow{N \rightarrow \infty} \left(0, tC^{-1}\mathbb{E} \left[\sum_{i=1}^{T_1} F_i^{\otimes 2} \right] \right)_{t \in [0,1]}. \quad (3.53)$$

Using the Slutsky's theorem (theorem 3.1 from [9]), and combining (3.52) and (3.53), we get

$$\left(\delta_{(NC\beta^{-1})^{-1/2}}\mathbb{Y}_t^{(N)}\right)_{t \in [0,1]} \xrightarrow{N \rightarrow \infty} \left(\mathbb{B}_t^{Strat} \otimes (0, t(\Gamma - C^{-1}\mathbb{E} \left[\sum_{i=1}^{T_1} F_i^{\otimes 2} \right])\right)_{t \in [0,1]} \quad (3.54)$$

where \mathbb{B}^{Strat} is the standard Brownian motion $(B_t)_t$ on V enhanced with its iterated integrals in the Stratonovich sense.

At the same time, we can compute

$$\Gamma - C^{-1}\mathbb{E} \left[\sum_{i=1}^{T_1} F_i^{\otimes 2} \right] = C^{-1}\mathbb{E} \left[\frac{1}{2} \sum_{1 \leq k_1 < k_2 \leq T_1} F_{k_1} \otimes F_{k_2} - F_{k_2} \otimes F_{k_1} \right] - C^{-1}\mathbb{E} \left[\frac{1}{2} \sum_{i=1}^{T_1} F_i^{\otimes 2} \right] \quad (3.55)$$

$$\begin{aligned} &= C^{-1}\mathbb{E} \left[\sum_{1 \leq k_1 < k_2 \leq T_1} F_{k_1} \otimes F_{k_2} \right] - \frac{1}{2}C^{-1}\mathbb{E} [X_{T_1}^{\otimes 2}] \\ &= C^{-1}\mathbb{E} \left[\sum_{1 \leq k_1 < k_2 \leq T_1} F_{k_1} \otimes F_{k_2} \right] - \frac{1}{2}I_d = M - \frac{1}{2}I_d \end{aligned} \quad (3.56)$$

where $M = C^{-1}\mathbb{E} \left[\sum_{1 \leq k_1 < k_2 \leq T_1} F_{k_1} \otimes F_{k_2} \right]$ (as stated in (3.18)) and $\mathbb{E} [X_{T_1}^{\otimes 2}]$ is non other than the covariance matrix of X_{T_1} since the excursions are centred.

The expression (3.55) shows that $(0, Mt)$ is a non-geometric rough path: while the first term of the sum is antisymmetric, the second one is symmetric and can not be expressed through the first component of the rough path (which is zero). This shows that the limit of the $(\mathbb{Y}_t^{(N)})_t$ is a non-geometric rough path.

The expression (3.56) shows that we can rewrite the limit using Itô integration, as

$$\forall 0 \leq s < t \leq 1, \mathbb{B}_{t,s}^{Strat} = \mathbb{B}_{t,s}^{Itô} \otimes (0, \frac{1}{2}(t-s)I_d).$$

We thus obtain the convergence:

$$\left(\delta_{(NC\beta^{-1})^{-1/2}}\mathbb{Y}_t^{(N)}\right)_{t\in[0,1]} \xrightarrow{N\rightarrow\infty} \left(\mathbb{B}_t^{It\delta} \otimes (0, tM)\right)_{t\in[0,1]}$$

which achieves the proof. \square

Remark. We find in the expression of the limit (3.54) the decomposition of a non geometric rough path in a geometric rough path plus an element in the centre of $T_1^{(2)}(V)$, highlighted in particular in [32] and expressed by

$$\forall 1/3 < \alpha < 1/2, \mathcal{C}^\alpha([0, 1], T_1^{(2)}(V)) \simeq \mathcal{C}^\alpha([0, 1], G^2(V)) \oplus \mathcal{C}^{2\alpha}([0, 1], V \otimes V).$$

3.5 Open questions

We have seen that hidden Markov walks are a natural generalization of the class of Markov chains on periodic graphs when one wants to study the area anomaly using techniques related to the theory of excursions, in particular allowing to analyse the convergence of a HMW as that of a sum of i.i.d. variables. It would be therefore interesting to know whether we can obtain results on area anomaly on other processes which present only weak time-correlations that could be ignored when passing to the limit in uniform convergence topology (for example, the α -mixing processes described in [9]).

A further study of the iterated occupation times $L_{u_1, \dots, u_k; N}(R)$ 3.2.5 of a Markov chain $(R_n)_n$ could also prove fruitful. Several directions can be considered: a generalization of some results of the ergodic theory (in particular, the connection with the invariant measure), their study as random combinatorial objects (using their representation as cardinals), the construction and analysis of an abstract vector space generated by this objects (which would in particular allow an abstract representation of a HMW depending on $(R_n)_n$), etc.

Furthermore, using the example of HMW, we would like to see if there is a connection between the iterated occupation times as combinatorial structures and the Hopf algebras which describe the combinatorics of (abstract) rough paths (as in [43]).

We have seen that, in the case of HMW, iterated sums allow, for example, to get a decomposition of rough paths that isolates the area anomaly or to propose an interesting construction of non-geometric rough paths. It can prove useful to continue studying these objects (possibly for more general processes) as they may contribute on one hand to developing the discrete setting in the rough paths theory and on the other hand to construct some concrete examples which allow explicit computations, which is often difficult when it comes to rough paths.

It could be interesting to use the framework we present in this paper to study in a new way finite difference equations and in particular their asymptotics and continuous limits; or, in the reverse way, if new discrete models could be imagined to discretize efficiently classical (stochastic) differential equations.

Overall, we hope that the present paper will provide the reader with good arguments to be interested in the role of the area anomaly in the classical stochastic calculus and that of discrete processes in the classical rough paths setting.

Chapter 4

Additional results on area anomaly and iterated occupation times

4.1 A case study for the area anomaly

In order to illustrate our main results, as well as the construction and computation of area anomaly for processes presented in chapters 3 and 2, we will study two models which are hidden Markov walks *and* Markov chains on periodic graphs.

- The first one is *the rotating sum of Bernoulli i.i.d.r.v.*, presented in chapter 2, section 2.1.3, as a first example of Markov chains on periodic graphs. As we have already seen, it is a discrete 2-dimensional process $(Z_n)_n$ where

$$Z_n = \sum_{k=1}^n i^{k-1} U_k \quad (4.1)$$

where the U_k s are i.i.d.r.v. of distribution $\mathbb{P}(U_k = 1) = 1 - \mathbb{P}(U_k = -1) = p$. This model is represented on the left side of the figure 4.1.

- The second one is *the "swirl" model*, a Markov chain $(\hat{Z}_n)_n$ on the \mathbb{Z}^2 lattice obtained by a slight change in the first one, as shown in figure 4.1.

We can see that both models can be constructed by assigning an orientation and a probability to every edge of the \mathbb{Z}^2 network following a particular periodicity.

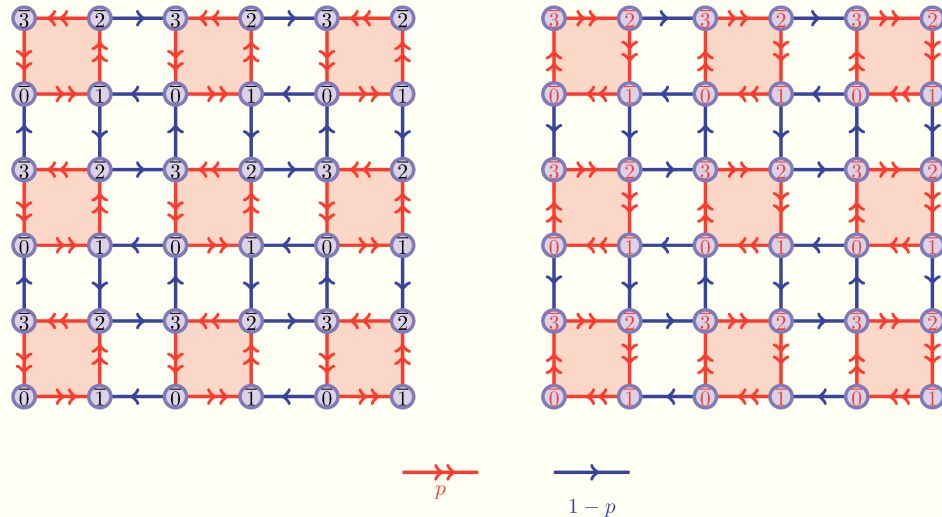


Figure 4.1: The two models: rotating sum of Bernoulli i.i.d.r.v. (on the left) and "swirl" model (on the right). The latter is obtained from the first one by changing the orientation of the red edges. For both models the labels of the vertices indicate one of the four configurations of edges leaving the vertex, i.e. the law of $Z_{n+1} - Z_n$ ($\hat{Z}_{n+1} - \hat{Z}_n$ resp.) knowing Z_n (\hat{Z}_n resp.).

One may wonder why we chose two models that are almost identical at first sight but it is this similarity that will help us illustrate some interesting phenomena. In particular, we will see that the slight modification necessary to pass from one model to another

- does not change the limit in the uniform topology (*modulo* a multiplicative constant), *but*
- may change the area anomaly,
- may give rise to a representation of the model as Markov chain on periodic graph which is different from its representation as a hidden Markov walk (which is the case of the "swirl" model), and

- may complicate the computations of the area anomaly by introducing extra randomness (in this particular case, it will randomize the length of pseudo-excursions).

4.1.1 The two models as hidden Markov walks

We will use here the notions and notations from chapter 3.

Rotating sum of Bernoulli i.i.d.r.v. We naturally set $V = \mathbb{R}^2$ (so it can be identified canonically to \mathbb{C}) and, for all $k \in \mathbb{N}^*$,

$$F_k = i^{k-1}U_k$$

so that, for all $n \in \mathbb{N}$, $Z_n = \sum_{k=1}^n F_k$. We notice that all the F_l s are independent and we can deduce from figure 4.1 that, for all $k \in \mathbb{N}^*$, the law of F_k depends entirely on $r_k \equiv k[4]$. It follows that setting

$$E = \mathbb{Z}/4\mathbb{Z}$$

and choosing $(R_n)_n$ to be a deterministic walk on E with

$$R_n = n[4] \text{ a.s.}$$

we have a representation of the rotating sum of Bernoulli i.i.d.r.v. as a HMW $(R_n, Z_n)_n$ on $E \times V$.

We notice that the pseudo-excursions are of length $T_1 = 4$ a.s. in this case, and thus the variables $\sum_{l=1}^4 i^{k+l-1}U_{k+l}$ are i.i.d. (moreover, they are centred).

"Swirl" model. Let, once again, $V = \mathbb{R}^2$. We consider the process $(\hat{Z}_n)_n$ obtained from $(Z_n)_n$ as explained in figure 4.1. We then set

$$\hat{F}_k = \hat{Z}_{n+1} - \hat{Z}_n.$$

Although the \hat{F}_k s are entirely determined by the label of the vertex they correspond to (see figure 4.1), the choice of $\mathbb{Z}/4\mathbb{Z}$ as \hat{E} and the sequence of labels of vertices

corresponding to each \hat{Z}_k as the underlying Markov chain $\hat{R} = (\hat{R}_n)_n$ on \hat{E} does not allow to construct a hidden Markov walk anymore. The reason is simple: the increments of $(\hat{Z}_n)_n$ are not independent under the law $\mathbb{P}(\bullet|\sigma(R))$.

We notice that the law of each F_k can be read from figure 4.1 from the direction and the colour of the vector corresponding to \hat{F}_{k-1} , i.e. the orientation of \hat{F}_{k-1} and the probability corresponding to this orientation. We then identify the set of all possible orientations of the edges with the correspondent probabilities

$$E' = \{(1, p), (i, p), (1, 1-p), (i, 1-p), (-1, p), (-i, p), (-1, 1-p), (-i, 1-p)\}$$

with

$$\hat{E} = \mathbb{Z}/8\mathbb{Z}$$

and define $(\hat{R}_n)_n$ as a random walk on \hat{E} whose stochastic matrix

$$Q = \begin{pmatrix} 0 & p & 1-p & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & p & 1-p & 0 & 0 & 0 \\ p & 0 & 0 & 0 & 0 & 0 & 0 & 1-p \\ 0 & p & 1-p & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & p & 1-p & 0 \\ p & 0 & 0 & 0 & 0 & 0 & 0 & 1-p \\ 0 & 0 & 0 & p & 1-p & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & p & 1-p & 0 \end{pmatrix}$$

is determined by the correlations between the \hat{F}_k s. We thus get a representation of the "swirl" model as a HMW $(\hat{R}_n, \hat{Z}_n)_n$ on $\hat{E} \times V$.

In this case, the pseudo-excursions are of random length, as the underlying Markov chain $(\hat{R}_n)_n$ is not deterministic anymore.

4.1.2 The two models as Markov chains on periodic graphs

We use here the notions and notations from chapter 2.

Rotating sum of Bernoulli i.i.d.r.v. The decomposition of the model doesn't change, what changes is only its interpretation. The \mathbb{Z}^2 network corresponding

to this model in figure 4.1 is now the graph G , and we search for the underlying graph G_0 and the Markov chain on it which determine the periodic behaviour of $(Z_n)_n$ on G .

We easily identify the set of vertices of G_0 with the set $E = \mathbb{Z}/4\mathbb{Z}$ we have determined for the corresponding HMW in section 4.1.1. We label the vertices of G from figure 4.1 with the elements from $\mathbb{Z}/4\mathbb{Z}$ (vertices of G_0). The underlying Markov chain on G_0 is deterministic and can be easily constructed using $(Z_n)_n$.

All of this has been represented in figure 2.1 from chapter 2.

"Swirl" model. In the case of the "swirl" model, we can considerably simplify its representation by choosing to see it as a Markov chain on a periodic graph rather than as a HMW.

As before, the \mathbb{Z}^2 network corresponding to this model in figure 4.1 is the graph G , and we want to determine \hat{G}_0 and the Markov chain which describes the periodic behaviour of $(\hat{Z}_n)_n$.

We notice that, as in the case of the rotating sum of Bernoulli i.i.d.r.v., the orientation $(\hat{Z}_n)_n$ gives to the edges of G yields four equivalence classes of vertices. Thus, we can this time identify the set of vertices of \hat{G}_0 with $\mathbb{Z}/4\mathbb{Z}$. Moreover, the orientation of edges that $(\hat{Z}_n)_n$ induces on G determines the underlying Markov chain on \hat{G}_0 whose stochastic matrix is given by

$$\hat{Q}_0 = \begin{pmatrix} 0 & p & 0 & 1-p \\ 1-p & 0 & p & 0 \\ 0 & 1-p & 0 & p \\ p & 0 & 1-p & 0 \end{pmatrix}$$

In figure 4.2, we present the graph G oriented by $(\hat{Z}_n)_n$ (as in figure 4.1) and whose vertices are labelled with elements of $\mathbb{Z}/4\mathbb{Z}$, as well as the graph \hat{G}_0 oriented by the underlying Markov chain.

4.1.3 Computing the area anomaly for the two models

Since the limits of the models, and in particular the area anomaly, do not depend on their interpretation, we choose to see both models as Markov chains on periodic

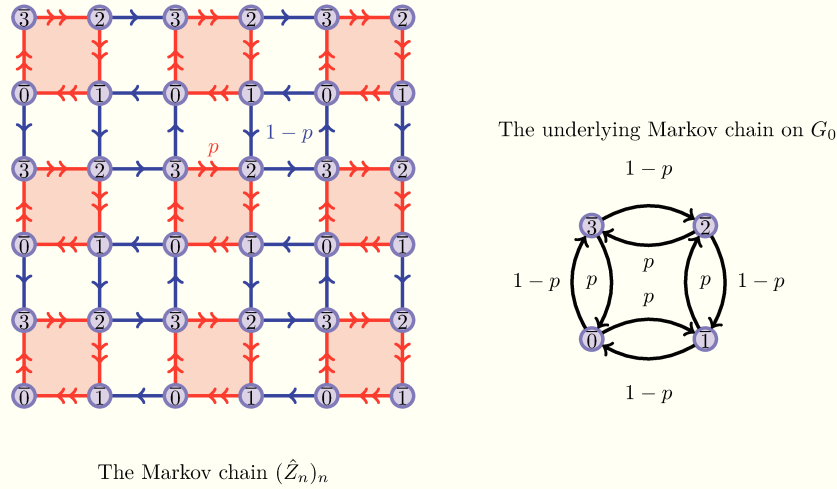


Figure 4.2: The "swirl" model represented as a Markov chain on a periodic graph.

graphs. We have seen that, for the trajectory $(X_n)_n$ of a HMW in general, and for a Markov chain on periodic graph in particular, under the condition $\mathbb{E}[X_{T_1}] = 0$ with T_1 as in (3.6) (i.e. the pseudo-excursions are centred), the area anomaly Γ is given by

$$\Gamma = C^{-1} \mathbb{E}[A_{T_1}(X)]$$

where $C > 0$ comes from the covariance matrix of X_{T_1} and $\mathbb{E}[A_{T_1}(X)]$ is the expectation of the stochastic area of a pseudo-excursion. While it is not difficult to compute C , the computations related to the stochastic area are not very common. We will propose an approach based on pseudo-excursions and the properties of the signed area that gives the possibility to treat both examples. It can also be generalized to other Markov chains on periodic graphs, especially the ones presenting certain symmetry properties.

Rotating sum of Bernoulli i.i.d.r.v. In section 2.1.3 of chapter 2 we have given directly the value of the area anomaly for this model. Let us now explain how to compute it.

We have already seen that the pseudo-excursions of $(Z_n)_n$ are of constant length

$T_1 = 4$ a.s. Moreover, we can see that the set of all possible trajectories of a pseudo-excursion can be decomposed, as presented in figure 4.3, in

- four loops B_1, B_2, B_3 and B_4 that do not influence the trajectory but generate an extra area term that contributes to the area anomaly and
- a symmetrical part, in which each path has a (unique) correspondent symmetrical path of the same associated probability, whose expectation is zero.

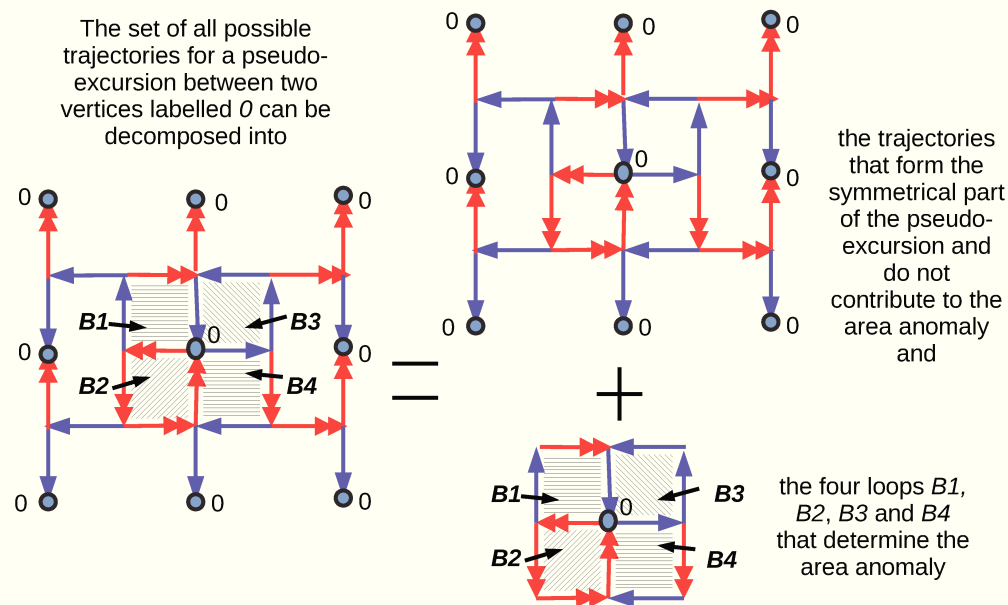


Figure 4.3: The pseudo-excursion can be decomposed in a symmetrical part, in which each path has a (unique) correspondent symmetrical path of the same associated probability, and four loops B_1, B_2, B_3 and B_4 . The symmetrical part is of centred increment and stochastic area, while the loops are the ones that generate the area anomaly.

We can thus conclude that

$$\mathbb{E}[Z_{T_1}] = \mathbb{E}[Z_4] = 0$$

and deduce at the same time that

$$C = 8p(1 - p).$$

Moreover, using the relation between the signed areas of two symmetrical paths (their signed areas are identical in absolute value but have opposite signs) stated in property 1.3.7 of chapter 1, we state the following easy property.

Property 4.1.1. *Consider a process in V whose possible trajectories can be split into two finite sets which are symmetrical, i.e. to every trajectory from one set corresponds a unique symmetrical trajectory from the other such that both have the same probability. Then the stochastic area of this process is centred (see figure 4.4).*

By property 4.1.1, we thus deduce from the decomposition in figure 4.3 that only the four loops B_1 , B_2 , B_3 and B_4 contribute to the area anomaly, the stochastic area resulting from the symmetrical part of the pseudo-excursion being centred. We can therefore conclude that

$$\mathbb{E}[A_{T_1}(Z)] = \mathbb{E}[A_4(Z)] = p^2 + (1 - p)^2 - 2p(1 - p) = (2p - 1)^2$$

p^2 being the area of B_2 , $(1 - p)^2$ the area of B_3 and $p(1 - p)$ the area of both B_1 and B_4 respectively. Consequently, the area anomaly is worth

$$\Gamma = \frac{(2p - 1)^2}{8p(1 - p)}$$

which is the result announced in section 2.1.3 of chapter 2.

The "swirl" model. Things get more complicated in this case since the pseudo-excursions are of random length and not centred. This latter problem can be solved by imposing as initial law for the Markov walk $(\hat{Z}_n)_n$ the law induced by the invariant probability of its underlying Markov chain. We will focus here only on computing the expectation of the stochastic area of a pseudo-excursion.

In order to palliate the first problem, we will use a decomposition of the pseudo-excursion that can be useful for many models that have certain symmetry properties.

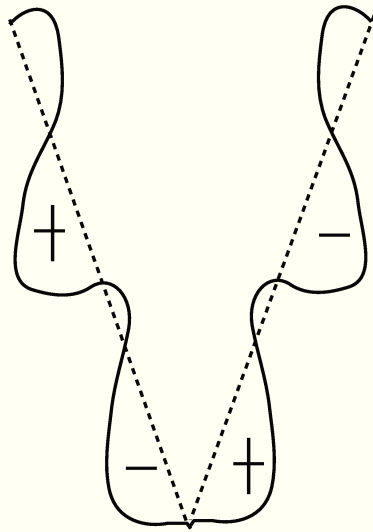


Figure 4.4: The signed areas of two symmetrical and equiprobable trajectories yield a zero mean.

The idea is to devise the pseudo-excursion into a series of pseudo-excursions *that do not contain its starting point*.

Consider a pseudo-excursion of $(\hat{Z}_n)_n$ between vertices labelled 1 (we can start with any vertex, the results are equivalent under the invariant measure).

- We notice that the paths $1 \rightarrow 0 \rightarrow 1$ and $1 \rightarrow 2 \rightarrow 1$ do not generate area anomaly.
- We then consider the two possibilities of getting to a vertex labelled 3: $1 \rightarrow 0 \rightarrow 3$ and $1 \rightarrow 2 \rightarrow 3$. We compute the signed areas of these triangles.
- Before getting again into 1, we have to follow several times either the path $3 \rightarrow 2 \rightarrow 3$ or $3 \rightarrow 0 \rightarrow 3$.
- We then have the two possibilities of getting back to a vertex labelled 1: $3 \rightarrow 0 \rightarrow 1$ and $3 \rightarrow 2 \rightarrow 1$.

Thus, the part of the pseudo-excursion between two vertices labelled 1 that contributes to the area anomaly can be decomposed in three main steps, which are

shown in figure 4.5 The pseudo-excursions between vertices labelled 3 that are

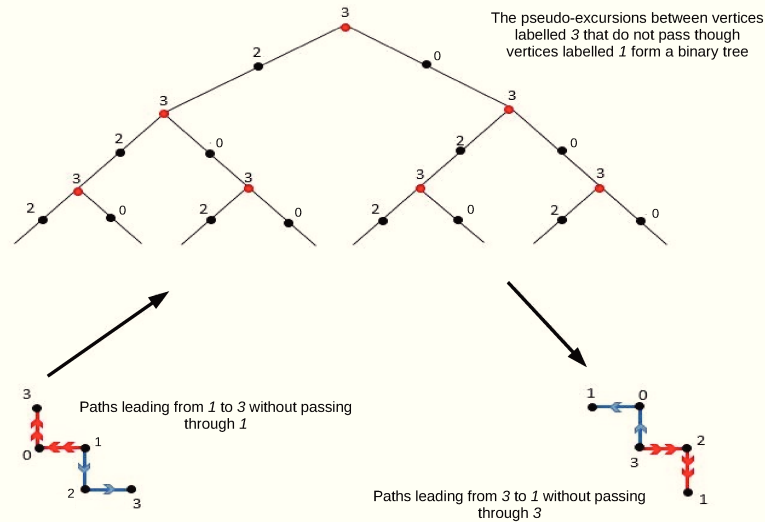


Figure 4.5: The decomposition of a pseudo-excursion between two vertices labelled 1 of the "swirl" model.

contained in pseudo-excursion we are studying can be seen as a random walk on a binary tree: in figure 4.5 we present the "unfolded" version of this tree, as on the network some branches cross or coincide. The advantage of this representation is that it allows to point out the symmetries present in this part of the excursion and it becomes obvious that property 4.1.1 applies here.

Consequently, the only parts that are relevant for computing the signed area of a pseudo-excursion are

- the path to get from 1 to 3,
- the increment of \hat{Z} between this vertex labelled 3 and the last vertex labelled 3 it gets to before going back to 1 and
- the path to get from 3 back to 1.

This is summarized in figure 4.6. However, even at this point we can find extra

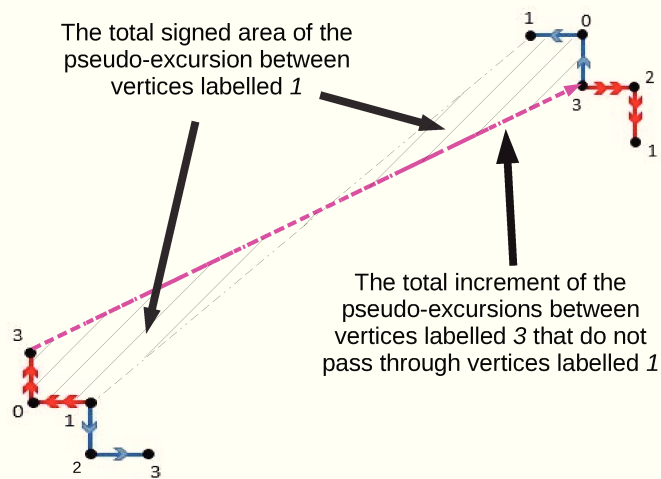


Figure 4.6: The stochastic area of a pseudo-excursion between vertices labelled 1 of the "swirl" model.

symmetries, as shown in figure 4.7: for any path that starts with $1 \rightarrow 2$ (or $1 \rightarrow 0$) and ends with $2 \rightarrow 1$ (with $0 \rightarrow 1$ resp.) we can find a symmetrical one. So this means that the only areas that have to be taking into account are those of paths beginning with blue end ending with red, like the ones in figure 4.8. We set $T^{(3)}$ to be the moment when \hat{Z} reaches a vertex labelled 3 for the last time during its pseudo-excursion from 1 to 1 and using the formula (1.20) to deduce the area of

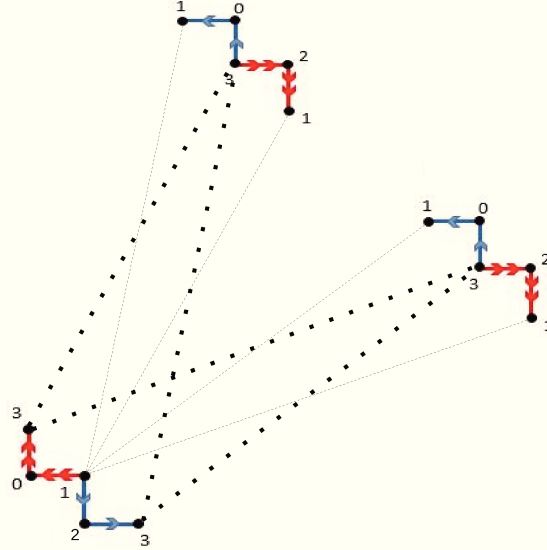


Figure 4.7: Some symmetries of a pseudo-excursion between vertices labelled 1 of the "swirl" model: each increment of $(\hat{Z}_n)_n$ corresponding to a pseudo-excursion $3 \rightarrow 3$ that does not pass through 1 can be coupled with exactly one other symmetrical and equally probable increment of the same type.

an oriented triangle, we get the following computations.

$$\begin{aligned}
 \mathbb{E} \left[A_{T_1}^{12}(\hat{Z}) \right] &= \frac{1}{2} p^4 \left((1, -1) \wedge \mathbb{E} \left[\hat{Z}_{T(3)} \right] + \left((1, -1) + \mathbb{E} \left[\hat{Z}_{T(3)} \right] \right) \wedge (-1, 1) \right) \\
 &+ \frac{1}{2} (1-p)^4 \left((-1, 1) \wedge \mathbb{E} \left[\hat{Z}_{T(3)} \right] + \left((-1, 1) + \mathbb{E} \left[\hat{Z}_{T(3)} \right] \right) \wedge (1, -1) \right) \\
 &= (p^4 - (1-p)^4) \left((1, -1) \wedge \mathbb{E} \left[\hat{Z}_{T(3)} \right] \right) \\
 &= (2p-1)(p^2 + (1-p)^2) \left((1, -1) \wedge \mathbb{E} \left[\hat{Z}_{T(3)} \right] \right)
 \end{aligned}$$

Computing $\mathbb{E} \left[\hat{Z}_{T(3)} \right]$ requires some extra work but is not too complicated, as it is

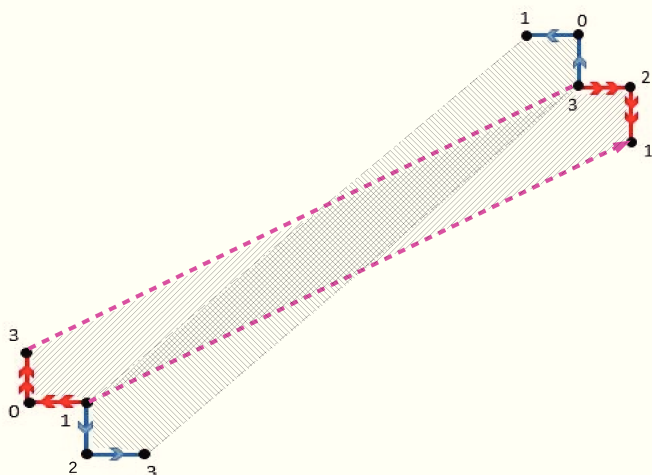


Figure 4.8: The areas to be taken into account for the area anomaly of the "swirl" model.

a sum of i.i.d.r.v. (increments of pseudo-excursions). It is worth

$$\mathbb{E} \left[\hat{Z}_{T^{(3)}} \right] = \frac{2p(1-p) + 1}{(p^2 + (1-p)^2)^2}$$

so that finally we get for the stochastic area of a pseudo-excursion the value

$$\mathbb{E} \left[A_{T_1}^{12}(\hat{Z}) \right] = (2p-1) \frac{2p(1-p) + 1}{p^2 + (1-p)^2}.$$

However, we can already state that computing the area anomaly for the "swirl" model is more complicated than for the rotating sum of Bernoulli i.i.d.r.v. In the next section, we will present a formula that allows to compute the mean of the stochastic area of a pseudo-excursion for models that have certain symmetry properties (corollary 4.2.1).

4.2 Methods for computing the area anomaly

4.2.1 A construction using the loop-erased random walk

A general presentation. We have already stressed the idea that, under certain conditions, the limit in the uniform topology of a hidden Markov walk (or of a Markov chain on periodic graph) is mainly determined by the vectors of the pseudo-excursions of this process, whereas the fluctuations during the pseudo-excursions, invisible in the uniform topology, appear at the limit in the rough topology as the area drift.

We can reformulate the above idea as follows: the convergence in the uniform topology is blind to certain fluctuations, loops in particular, that take place during an excursion of our Markov process $(R_n)_n$. At the same time, these fluctuations create the area anomaly. It is consequently interesting to study the structures of the pseudo-excursions, or pseudo-loops, of $(X_n)_n$, and this is where a loop-erased random walk approach can be useful.

In general, a *loop-erased random walk* (LERW) is obtained by erasing loops in a simple random walk ([56]). However, we can generalize it to other processes in which loops may play an important role.

We construct a loop-erased random walk from an irreducible Markov chain $(R_n)_n$, $LERW((R_n)_n)$, as the process $(W_n^E)_n$ defined by

$$\forall n \geq 1, W_n^E = (a_0, a_1, \dots, a_{m_n})$$

where $R_0 = a_0$ a.s. and $a_{k+1} = R_{T_{\kappa(n)}^{(a_k)}+1}$ for $T_i^{(a)}$ a stopping time as in (3.6) for an excursion from a to a .

The sequence $(W_n^E)_n$ is a sequence of paths in E , and it is obvious that $W_{T_n^{(a_0)}}^E = (a_0)$. W_n^E is the path of the Markov chain until time n from which we have erased all the loops.

In its geometrical interpretation, the coefficients of the area anomaly Γ are given by the expectation of the signed area corresponding to the (random) path from X_0 to X_{T_1} , which depends on the excursion (R_0, \dots, R_{T_1-1}) . The problem is that this excursion contains many intermediate excursions (or loops) for points other than R_0 , which can be iterated a great number of times, making direct computations of

the path complicated. The idea is then to identify all the possible types of loops contained in (R_0, \dots, R_{T_1-1}) , study them separately and then put together again.

We can thus construct an algorithm based on the LERW as follows:

- For $R_0 = a_0 \in E$ a.s., we construct all the possible paths $W_{T_1^{(a_0)}-1}^E$, i.e. all the paths in E from R_0 to R_{T_1} that only contain distinct points.
- For every $b_j \in W_{T_1^{(a_0)}-1}^E$, we construct all the possible paths $W_{T_1^{(b_j)}-1}^{E \setminus \{a_0\}}$, i.e. all paths in $E \setminus \{a_0\}$ from b_j to b_j that only contain distinct points.
- We start all over for every $c_j \in W_{T_1^{(b_j)}-1}^{E \setminus \{a_0\}}$.
- Since E is finite, the algorithm stops at some point.
- We then compute the increments and stochastic areas of the process (X_n) which correspond to each $W_{T_1^u-1}^{E'}$ that has appeared in the algorithm about, and we put them together.

The final difficulty is precisely the way we combine the paths and stochastic areas corresponding to every path we take from the LERW. Below we present a combinatorial formula which achieves this on a slightly different basis.

A combinatorial formula inspired by the LERW algorithm. We will now give a formula for computing the area anomaly inspired from the above algorithm. Although we have developed it based on the geometrical properties of the Markov chains on periodic graphs (from chapter 2), it can be applied to the hidden Markov walks as well, as it uses the decomposition into pseudo-excursions.

For any $a, b \in E$, consider the sets

$$\mathcal{C}_{aa} = \{aa_1 \dots a_n a : a_1, \dots, a_n \in E \setminus \{a, b\}, n \in \mathbb{N}\}$$

$$\mathcal{C}_{ab} = \{aa_1 \dots a_n b : a_1, \dots, a_n \in E \setminus \{a, b\}, n \in \mathbb{N}\}$$

$$\mathcal{C}_{ba} = \{bb_1 \dots b_n a : b_1, \dots, b_n \in E \setminus \{a, b\}, n \in \mathbb{N}\}$$

$$\mathcal{C}_{bb} = \{bb_1 \dots b_n b : b_1, \dots, b_n \in E \setminus \{a, b\}, n \in \mathbb{N}\}$$

If we add a cemetery point o to E , the elements of the above sets can be identified with $E \cup \{o\}$ -valued sequences $s = (c_n)_n$ such that $c_0 c_1 \dots c_{n(o)} \in \mathcal{C}_{uv}$ for $u, v \in \{a, b\}$

and a certain $n(o) \in \mathbb{N}$, and $c_n = o$ for $n > n(o)$. In order to avoid re-centring, we suppose that the pseudo-excursions of X are centred ($\mathbb{E}[X_{T_1}] = 0$).

We have the following result:

Proposition 4.2.1. *Let $(X_n)_n$ be a Markov chain on a periodic graph $G \subset V$ such that $X_0 = 0$ a.s. and set $(R_n)_n = (\pi_0(X_n))_n$. We denote by $A_n(X)$ the stochastic area corresponding to the geodesic embedding up to n of $(X_k)_k$. We set, for $x, y \in E$,*

- $p(s) = \mathbb{P}(R_0 = s_0, R_1 = s_1, \dots, R_{n(o)} = s_{n(o)}) \in [0, 1]$, $s \in \mathcal{C}_{xy}$: *the probability of a given path that goes from x to y without passing through x or y in between;*
- $P(x, y) = \sum_{s \in \mathcal{C}_{xy}} p(s) \in [0, 1]$: *the probability to go from x to y without passing through x or y in between;*
- $S_X(x, y) = \sum_{s \in \mathcal{C}_{xy}} p(s) X_{n(o)}(s) \in V$: *the expectation of a path of X going from x to y without passing through x or y in between;*
- $S_A(x, y) = \sum_{s \in \mathcal{C}_{xy}} p(s) A_{n(o)}(X(s)) \in V \otimes V$ (ou $V \wedge V$): *the mean value of the stochastic area of the path described above.*

We then have:

$$\begin{aligned} \mathbb{E}_a[A_{T_1}(X)] &= \frac{P(a, b)}{P(b, a)} \left(\frac{1}{2} \left(\frac{S_X(a, b)}{P(a, b)} - \frac{S_X(b, a)}{P(b, a)} \wedge S_X(b, b) \right) + S_A(b, b) \right) \\ &\quad + \frac{1}{2} \left(\frac{S_X(a, b)}{P(a, b)} \wedge \frac{S_X(b, a)}{P(b, a)} \right) P(a, b) \\ &\quad + \left(\frac{S_A(a, b)}{P(a, b)} + \frac{S_A(b, a)}{P(b, a)} \right) P(a, b) + S_A(a, a) \end{aligned}$$

where \wedge is the antisymmetric tensor product.

Remark. We will see that this formula is explicit enough to perform computations for a large class of models.

Proof. As it has already been mentioned, this method is inspired from the LERW algorithm: in order to get from a vertex labelled a to another one, with the same label, we

- go from a point labelled a to a point labelled b without passing through a ,
- count in all the pseudo-excursions between two points labelled b that do not pass through a ,
- return at some point back to a .

This is the principle applied to the "swirl" model in the section 4.1 of this chapter.

We start by defining, for any $a, b \in E$, the stopping time

$$S_{a,b} = \inf\{n > 0 : R_n = b \text{ when } R_0 = a\}$$

Furthermore, we denote by $N(b, a)$ the number of excursions through b in E of the chain $(R_n)_n$ under \mathbb{P}_b before it reaches the point a , i.e.

$$N(b, a) = \sup\{k : T_k^{(b)} < S_{b,a}\} = \inf\{k : T_{k+1}^{(b)} > S_{b,a}\}$$

where $T_k^{(b)}$ is T_k for $R_0 = b$ a.s. Since the stochastic area is invariant by translation, we can fix $R_0 = 0$ a.s. without loss of generality. For $R_0 = a$ a.s. and for $b \in E$, we can decompose $A_{T_1}(X)$ as follows

$$\begin{aligned} A_{T_1^{(a)}}(X) &= (A_{S_{a,b}}(X) + \frac{1}{2}X_{S_{a,b}} \wedge (X_{T_{N(b,a)}^{(b)}+S_{a,b}} - X_{S_{a,b}})) + \sum_{k=0}^{N(b,a)} A_{T_{k+1}^{(b)}-T_k^{(b)}}(X_{S_{a,b}+T_k^{(b)}+\bullet}) \\ &+ \frac{1}{2} \left(X_{T_{N(b,a)}^{(b)}+S_{a,b}} \right) \wedge \left(X_{T_{N(b,a)}^{(b)}+S_{a,b}+S_{b,a}} - X_{T_{N(b,a)}^{(b)}+S_{a,b}} \right) + A_{S_{b,a}} \left(X_{T_{N(b,a)}^{(b)}+S_{a,b}+\bullet} \right) \mathbf{1}_{S_{a,b} < T_1^{(a)}} \\ &+ A_{T_1} \mathbf{1}_{S_{a,b} > T_1^{(a)}}. \end{aligned}$$

Applying strong Markov property, we obtain:

$$\begin{aligned} \mathbb{E}_a \left[A_{T_1^{(a)}}(X) \right] &= (\mathbb{E}_a \left[A_{S_{a,b}}(X) \mid S_{a,b} < T_1^{(a)} \right] + \frac{1}{2} \mathbb{E}_a \left[X_{S_{a,b}} \mid S_{a,b} < T_1^{(a)} \right] \wedge \mathbb{E}_b \left[X_{T_{N(b,a)}^{(b)}} \right] \\ &+ \mathbb{E}_b \left[\sum_{k=0}^{N(b,a)} A_{T_{k+1}^{(b)}-T_k^{(b)}}(X) \right] + \frac{1}{2} (\mathbb{E}_b \left[X_{T_{N(b,a)}^{(b)}} \right] + \mathbb{E}_a \left[X_{S_{a,b}} \mid S_{a,b} < T_1^{(a)} \right]) \\ &\wedge \mathbb{E}_b \left[X_{S_{b,a}} \mid S_{b,a} < T_1^{(b)} \right] + \mathbb{E}_b \left[A_{S_{b,a}}(X) \mid S_{b,a} < T_1^{(b)} \right] \mathbb{P}_a \left(S_{a,b} < T_1^{(a)} \right) \\ &+ \mathbb{E}_a \left[A_{T_1}(X) \mathbf{1}_{S_{a,b} > T_1^{(a)}} \right]. \end{aligned}$$

Since the length of an excursion from b to b and its stochastic area do not depend on the number of excursions and are invariant by translation, we can rewrite the above formula as

$$\begin{aligned}
\mathbb{E}_a \left[A_{T_1^{(a)}}(X) \right] &= (\mathbb{E}_a \left[A_{S_{a,b}}(X) \mid S_{a,b} < T_1^{(a)} \right] \\
&+ \frac{1}{2} \mathbb{E}_a \left[X_{S_{a,b}} \mid S_{a,b} < T_1^{(a)} \right] \wedge \mathbb{E} [N(b, a)] \mathbb{E}_b \left[X_{T_1^{(b)}} \right] \\
&+ \mathbb{E} [N(b, a)] \mathbb{E}_b \left[A_{T_1^{(b)}}(X) \mid S_{b,a} > T_1^{(b)} \right] + \frac{1}{2} (\mathbb{E} [N(b, a)] \mathbb{E}_b \left[X_{T_1^{(b)}} \mid S_{b,a} > T_1^{(b)} \right] \\
&+ \mathbb{E}_a \left[X_{S_{a,b}} \mid S_{a,b} < T_1^{(a)} \right]) \wedge \mathbb{E}_b \left[X_{S_{b,a}} \mid S_{b,a} < T_1^{(b)} \right] + \mathbb{E}_b \left[A_{S_{b,a}}(X) \mid S_{b,a} < T_1^{(b)} \right]) \\
&\times \mathbb{P}_a \left(S_{a,b} < T_1^{(a)} \right) + \mathbb{E}_a \left[A_{T_1^{(a)}} \mathbf{1}_{S_{a,b} > T_1^{(a)}} \right]
\end{aligned}$$

which becomes further on

$$\begin{aligned}
\mathbb{E}_a \left[A_{T_1^{(a)}}(X) \right] &= (\mathbb{E} [N(b, a)] \left(\frac{1}{2} (\mathbb{E}_a \left[X_{S_{a,b}} \mid S_{a,b} < T_1^{(a)} \right] - \mathbb{E}_b \left[X_{S_{b,a}} \mid S_{b,a} < T_1^{(b)} \right]) \right. \\
&\quad \left. \wedge \mathbb{E}_b \left[X_{T_1^{(b)}} \mid S_{b,a} > T_1^{(b)} \right] + \mathbb{E}_b \left[A_{T_1^{(b)}}(X) \mid S_{b,a} > T_1^{(b)} \right] \right) \\
&\quad + \frac{1}{2} \mathbb{E}_a \left[X_{S_{a,b}} \mid S_{a,b} < T_1^{(a)} \right] \wedge \mathbb{E}_b \left[X_{S_{b,a}} \mid S_{b,a} < T_1^{(b)} \right] \\
&\quad + \mathbb{E}_b \left[A_{S_{b,a}}(X) \mid S_{b,a} < T_1^{(b)} \right] + \mathbb{E}_a \left[A_{S_{a,b}}(Y) \mid S_{a,b} < T_1^{(a)} \right]) \\
&\quad \mathbb{P}_a \left(S_{a,b} < T_1^{(a)} \right) + \mathbb{E}_a \left[A_{T_1^{(a)}}(X) \mid S_{a,b} > T_1^{(a)} \right] \mathbb{P}_a \left(S_{a,b} > T_1^{(a)} \right).
\end{aligned} \tag{4.2}$$

Furthermore, direct computations show that

- $\mathbb{P}_a \left(S_{a,b} < T_1^{(a)} \right) = \sum_{s \in \mathcal{C}_{ab}} p(s)$, $\mathbb{P}_a \left(S_{a,b} > T_1^{(a)} \right) = \sum_{s \in \mathcal{C}_{aa}} p(s)$
- $\mathbb{E}_a \left[X_{S_{a,b}} \mid S_{a,b} < T_1^{(a)} \right] = \frac{\sum_{s \in \mathcal{C}_{ab}} p(s) X_{n(o)}(s)}{\sum_{s \in \mathcal{C}_{ab}} p(s)}$
- $\mathbb{E}_b \left[X_{T_1^{(b)}} \mid S_{b,a} > T_1^{(b)} \right] = \frac{\sum_{s \in \mathcal{C}_{bb}} p(s) X_{n(o)}(s)}{\sum_{s \in \mathcal{C}_{bb}} p(s)}$
- $\mathbb{E}_a \left[A_{T_1^{(a)}}(X) \mid S_{a,b} > T_1^{(a)} \right] = \frac{\sum_{s \in \mathcal{C}_{aa}} p(s) A_{n(o)}(X(s))}{\sum_{s \in \mathcal{C}_{aa}} p(s)}$
- $\mathbb{E}_a \left[A_{S_{a,b}}(X) \mid S_{a,b} < T_1^{(a)} \right] = \frac{\sum_{s \in \mathcal{C}_{ab}} p(s) A_{n(o)}(X(s))}{\sum_{s \in \mathcal{C}_{ab}} p(s)}$

Second, we compute

$$\begin{aligned} \mathbb{E}[N(b, a)] &= \sum_{n \geq 1} n \mathbb{P}_b(S_{b,a} > T_1^{(b)})^n (1 - \mathbb{P}_b(S_{b,a} > T_1^{(b)})) = \sum_{n \geq 1} n \left(\sum_{s \in \mathcal{C}_{bb}} p(s) \right)^n \sum_{s \in \mathcal{C}_{ba}} p(s) \\ &= \frac{(\sum_{s \in \mathcal{C}_{bb}} p(s))(\sum_{s \in \mathcal{C}_{ba}} p(s))}{(1 - \sum_{s \in \mathcal{C}_{bb}} p(s))^2} = \frac{\sum_{s \in \mathcal{C}_{bb}} p(s)}{1 - \sum_{s \in \mathcal{C}_{bb}} p(s)} = \frac{\sum_{s \in \mathcal{C}_{bb}} p(s)}{\sum_{s \in \mathcal{C}_{ba}} p(s)}. \end{aligned}$$

Finally, by replacing all the terms in the formula (4.2) by the corresponding expressions, we get the desired result. \square

A particular and very important class of graphs/networks are the ones presenting properties of symmetry (especially by rotation by a deterministic angle θ). In practice (for instance, in statistical mechanics), they are of great interest. The expression for the area anomaly is then considerably simplified and given by

Corollary 4.2.1. *If we can find $a, b \in E$ such that there exists a bijective application $\phi : \mathcal{C}_{ab} \rightarrow \mathcal{C}_{ba}$ (resp. an application $\psi : \mathcal{C}_{aa} \rightarrow \mathcal{C}_{bb}$) such that, for any $s \in \mathcal{C}_{ab}$ (resp. $s \in \mathcal{C}_{aa}$) either $X(s) = -X(\phi(s))$ and $p(s) = p(\phi(s))$ (resp. $X(s) = -X(\psi(s))$ and $p(s) = p(\psi(s))$) or $p(s) = 0$ (resp. idem) then*

$$\mathbb{E}_a[A_{T_1}(X)] = -\frac{S_X(a, b)}{P(a, b)} \wedge S_X(a, a) + 2S_A(a, a) + 2S_A(a, b)$$

Proof. The assumptions of the corollary imply that

- $\sum_{s \in \mathcal{C}_{ab}} p(s) = \sum_{s \in \mathcal{C}_{ba}} p(s)$
- $\sum_{s \in \mathcal{C}_{aa}} p(s) = \sum_{s \in \mathcal{C}_{bb}} p(s)$
- $\sum_{s \in \mathcal{C}_{ab}} p(s)X(s) = -\sum_{s \in \mathcal{C}_{ba}} p(s)X(s)$
- $\sum_{s \in \mathcal{C}_{aa}} p(s)X(s) = -\sum_{s \in \mathcal{C}_{bb}} p(s)X(s)$
- $\sum_{s \in \mathcal{C}_{ab}} p(s)A(X(s)) = \sum_{s \in \mathcal{C}_{ba}} p(s)A(X(s))$
- $\sum_{s \in \mathcal{C}_{aa}} p(s)A(X(s)) = \sum_{s \in \mathcal{C}_{bb}} p(s)A(X(s))$

as the signed area is invariant by rotation. It is then enough to recalculate the formula from proposition 4.2.1. \square

Application: the "swirl" model. In this example, we will illustrate how the method based on the loop-erased random walk algorithm can be applied to the "swirl" model interpreted as a Markov chain on periodic graph.

The "swirl" model being invariant by rotation by $\pi/2$, it can be easily shown that if the initial law is the uniform distribution, $\mathbb{E}[\hat{Z}_{T_1}] = 0$. We thus want to compute $\mathbb{E}[A_{T_1}(\hat{Z})]$. Since we have a bidimensional process, it will be enough to compute $\mathbb{E}[A_{T_1}^{12}(\hat{Z})]$.

We notice that the points 1 and 3 satisfy the conditions of corollary 4.2.1, therefore we will only need to gather the informations from Table 4.1.

When inserting the values from Table 4.1 in the formula from corollary 4.2.1, we get

$$\begin{aligned} \mathbb{E}[A_{T_1}^{12}(\hat{Z})] &= -\frac{(1,-1)p^2 + (-1,1)(1-p)^2}{p^2 + (1-p)^2} \wedge (2,2)p(1-p) \\ &\quad + 2\left(-\frac{1}{2}p^2 + \frac{1}{2}(1-p)^2\right) \\ &= 4\frac{(1-2p)p(1-p)}{p^2 + (1-p)^2} + (1-2p) \\ &= (1-2p)\frac{2p(1-p) + 1}{p^2 + (1-p)^2} \end{aligned}$$

4.2.2 A formula using the stochastic matrix of $(R_n)_n$

In the previous section we have given a formula that allows to compute the area anomaly directly but without exploiting the fact that $(R_n)_n$ is positive recurrent. We will now try to exploit this particularity by constructing a method of computation which uses the transition matrix. First, we want to show that for our class of processes, the area anomaly can also be calculated as the limit of an expectation.

Lemma 4.2.1. *We still suppose that $(R_n, F_n)_n$ is a HMC on $E \times V$ with the F_k s uniformly bounded and $(R_n)_n$ irreducible, and denote by $(X_n)_n$ the corresponding HMW. Set $\Gamma = \mathbb{E}[A_{T_1}(X)]$. With the above notations, we have*

$$\lim_{n \rightarrow \infty} \beta C^{-1} \frac{\mathbb{E}[A_n(X)]}{n} = \Gamma$$

	s	$\hat{Z}_{n(o)}(s)$	$A_{n(o)}(\hat{Z}(s))$	$p(s)$
\mathcal{C}_{11}	121	(1, 0)	0	$p(1-p)$
	141	(0, 2)	0	$p(1-p)$
\mathcal{C}_{13}	123	(1, -1)	$-\frac{1}{2}$	p^2
	143	(-1, 1)	$\frac{1}{2}$	$(1-p)^2$
\mathcal{C}_{33}	343	(-2, 0)	0	$p(1-p)$
	323	(0, -2)	0	$p(1-p)$
\mathcal{C}_{31}	341	(-1, 1)	$-\frac{1}{2}$	p^2
	321	(1, -1)	$\frac{1}{2}$	$(1-p)^2$

Table 4.1: Path information for the "swirl" model.

Proof. We decompose:

$$\begin{aligned}\mathbb{E}[A_n(X)] &= \mathbb{E}[A_{\kappa(n)}(X_{T_{\kappa(n)}})] + \mathbb{E}\left[\sum_{k=0}^{\kappa(n)-1} A_{T_k, T_{k+1}}(X)\right] + \mathbb{E}[A_{T_{\kappa(n)}, n}(X)] \\ &\quad + \frac{1}{2}\mathbb{E}[X_{T_{\kappa(n)}} \wedge (X_n - X_{T_{\kappa(n)}})] \\ &= \mathbb{E}\left[\sum_{k=0}^{\kappa(n)-1} A_{T_k, T_{k+1}}(X)\right] + \mathbb{E}[A_{T_{\kappa(n)}, n}(X)]\end{aligned}$$

as the stochastic area of a sum of i.i.d. centred variables is of mean 0 and $\mathbb{E}[X_{T_{\kappa(n)}}] = 0$, with $X_{T_{\kappa(n)}}$ and $X_n - X_{T_{\kappa(n)}}$ being independent. Furthermore, since we can decompose $X_n = \sum_{k=1}^n F_k$ with the F_k s as in property 3.3.1, we have

$$\begin{aligned}\mathbb{E}[A_{T_{\kappa(n)}, n}(X)] &= \mathbb{E}\left[\sum_{1 \leq k_1 < k_2 \leq T_1} F_{k_1} \otimes F_{k_2} - F_{k_2} \otimes F_{k_1}\right] \\ &\leq K^2 \mathbb{E}\left[\frac{T_1(T_1 - 1)}{2}\right]\end{aligned}$$

The variable T_1 has a finite second moment as E is finite. Using the dominated convergence theorem, we obtain

$$\lim_{n \rightarrow \infty} \frac{\mathbb{E}[A_n(X)]}{n} = \mathbb{E}[A_{T_1}(X)] \mathbb{E}\left[\lim_{n \rightarrow \infty} \frac{\kappa(n)}{n}\right] = \beta^{-1} \mathbb{E}[A_{T_1}(X)] = \beta^{-1} C \Gamma$$

which achieves the proof. □

Consider $(R_n, F_n)_n$ and $(X_n)_n$ as in lemma 4.2.1. Suppose that the transition matrix Q of $(R_n)_n$ is diagonalizable. Then we can write Q as

$$Q = \sum_{\lambda \in Sp(Q)} \lambda v_\lambda v_\lambda^* \tag{4.3}$$

where $Sp(Q)$ is the set of eigenvalues of Q , v_λ is a right eigenvector associated to λ , i.e. such that

$$Qv_\lambda = \lambda v_\lambda$$

v_λ^* a left eigenvector related to λ , i.e. such that

$$v_\lambda^* Q = \lambda v_\lambda^*$$

and consequently $v_\lambda v_\lambda^*$ is a projector in $\mathcal{M}_{Card(E)}(\mathbb{R})$.

We have the following formula for computing the expectation of the stochastic area of a pseudo-excursion.

Proposition 4.2.2. *Let $\mu = (\mu_x)_{x \in E}$ be a probability law on E and ν is as defined in chapter 3. Define the matrix $\hat{Q} \in \mathcal{M}_{dim(E)}(V)$ as*

$$\forall (x, y) \in E^2, \hat{Q}(x, y) = Q(x, y) \mathbb{E}_\nu [F_1 | y]$$

We then have the following formula for computing the area anomaly:

$$\mathbb{E} [A_{T_1}^{(ij)}(X)] = \sum_{\lambda \in Sp(Q)} \frac{1}{1 - \lambda} \mu^* M^{ij}(\lambda) \mathbf{1} \quad (4.4)$$

where μ^* denotes the transposed vector μ and

$$\begin{aligned} M^{ij}(\lambda) &= (v_\lambda v_\lambda^*) \hat{Q}^{(i)} (v_1 v_1^*) \hat{Q}^{(j)} - (v_\lambda v_\lambda^*) \hat{Q}^{(j)} (v_1 v_1^*) \hat{Q}^{(i)} \\ &\quad + (v_1 v_1^*) \hat{Q}^{(i)} (v_\lambda v_\lambda^*) \hat{Q}^{(j)} - (v_1 v_1^*) \hat{Q}^{(j)} (v_\lambda v_\lambda^*) \hat{Q}^{(i)} \end{aligned}$$

with v_1 (resp. v_1^*) is the right (resp. the left) eigenvector associated to the eigenvalue 1 of the matrix Q .

Proof. Since $X_n = \sum_{k=1}^n F_k$, we have

$$\begin{aligned} \mathbb{E} \left[\sum_{1 \leq k < l \leq n} F_k^{(i)} F_l^{(j)} \right] &= \sum_{1 \leq k < l \leq n} \sum_{x_0, x_{k-1}, x_k, x_{l-1}, x_l \in E} \mu(x_0) \\ &\quad \times Q^{k-1}(x_0, x_{k-1}) Q(x_{k-1}, x_k) Q^{l-k-1}(x_k, x_{l-1}) Q(x_{l-1}, x_l) \mathbb{E}_\nu [F_1^{(i)} | x_k] \mathbb{E}_\nu [F_1^{(j)} | x_l]. \end{aligned}$$

Applying the formula 4.3, we get

$$\begin{aligned}
\mathbb{E} \left[\sum_{1 \leq k < l \leq n} F_k^{(i)} F_l^{(j)} \right] &= \sum_{1 \leq k < l \leq n} \sum_{x_l \in E} \sum_{x_0 \in E} \mu(x_0) \\
&\times \sum_{x_k \in E} \left(\sum_{x_{k-1} \in E} Q^{k-1}(x_0, x_{k-1}) \hat{Q}^{(i)}(x_{k-1}, x_k) \right) \left(\sum_{x_{l-1} \in E} Q^{l-k-1}(x_k, x_{l-1}) \hat{Q}^{(j)}(x_{l-1}, x_l) \right) \\
&= \sum_{1 \leq k < l \leq n} \sum_{x_l \in E} \sum_{x_0 \in E} \mu(x_0) Q^{k-1} \hat{Q}^{(i)} Q^{l-k-1} \hat{Q}^{(j)}(x_0, x_l) \\
&= \sum_{1 \leq k < l \leq n} \mu^* Q^{k-1} \hat{Q}^{(i)} Q^{l-k-1} \hat{Q}^{(j)} \mathbf{1} \\
&= \sum_{\lambda, \lambda' \in Sp(Q)} \left(\sum_{1 \leq k < l \leq n} \lambda^{k-1} \lambda'^{l-k-1} \right) \mu^*(v_\lambda v_\lambda^*) \hat{Q}^{(i)}(v_{\lambda'} v_{\lambda'}^*) \hat{Q}^{(j)} \mathbf{1}.
\end{aligned}$$

In order to evaluate the expression $S_n(\lambda, \lambda') = \sum_{1 \leq k < l \leq n} \lambda^{k-1} \lambda'^{l-k-1}$, we prove the following lemma.

Lemma 4.2.2. *For $a, b \in \mathbb{R}$ such that $|a| \leq 1$ and $|b| \leq 1$, and for $n \geq 2$, we have*

$$S_n(a, b) = \sum_{1 \leq k < l \leq n} a^{k-1} b^{l-k-1} = \begin{cases} n(n-1)/2 & \text{if } a = b = 1 \\ \frac{a^n - na + n - 1}{(a-1)^2} & \text{if } a \neq 1, b = 1 \\ \frac{b^n - nb + n - 1}{(b-1)^2} & \text{if } b \neq 1, a = 1 \\ \frac{1}{a-b} \left(\frac{a^n - a}{a-1} - \frac{b^n - b}{b-1} \right) & \text{if } b \neq 1, a \neq 1 \end{cases} \quad (4.5)$$

Thus, setting $S(\lambda, \lambda') = \lim_{n \rightarrow \infty} S_n(\lambda, \lambda')/n$ when this limit exists and applying lemma 4.2.1, we have the following expression for the area anomaly:

$$\mathbb{E} \left[A_{T_1}^{(ij)}(X) \right] = \sum_{\lambda, \lambda' \in Sp(Q)} S(\lambda, \lambda') \mu^* [(v_\lambda v_\lambda^*) \hat{Q}^{(i)}(v_{\lambda'} v_{\lambda'}^*) \hat{Q}^{(j)} - (v_\lambda v_\lambda^*) \hat{Q}^{(j)}(v_{\lambda'} v_{\lambda'}^*) \hat{Q}^{(i)}] \mathbf{1}. \quad (4.6)$$

In particular, if the term $S_n(1, 1)$ is multiplied by a non-zero coefficient, we need to subtract it from the sum before going to the limit, as it is of order n^2 . Moreover,

we notice that all the other cases are presented by

$$S(\lambda, \lambda') = \begin{cases} \frac{1}{1-\lambda} & \text{if } \lambda' = 1, \lambda \neq 1 \\ \frac{1}{1-\lambda'} & \text{if } \lambda = 1, \lambda' \neq 1 \\ 0 & \text{otherwise} \end{cases}$$

which combined with the formula (4.6) gives us the desired result. \square

We can deal in this way with sums on arbitrary multiplicity. For an iterated sum of order n , the numerical coefficients of the sum we will have to compute will be of the form $\sum_{\lambda_1, \dots, \lambda_n \in Sp(Q)} \lambda_1^{k_1-1} \lambda_2^{k_2-k_1-1} \dots \lambda_n^{k_n-k_{n-1}-1}$.

4.3 Computations related to combinatorial structures of hidden Markov walks

4.3.1 A connection between the iterated occupation times of hidden Markov walks and the resolvent

Let $(R_n, X_n)_n$ be, once again, a HMW on $E \times V$. In section 4.2.2, we have seen how the stochastic matrix Q of $(R_n)_n$ can be used to compute the area anomaly. Now we will see that the iterated occupation times of $(R_n)_n$ have an interesting connection with the *resolvent* of Q , i.e. the function $\mathcal{R} : \mathbb{R}_+ \setminus Sp(Q) \rightarrow \mathcal{M}_{\text{card}(E)}(\mathbb{R})$ such that

$$\forall \lambda \in \mathbb{R}_+ \setminus Sp(Q), \quad \mathcal{R}(\lambda) = (\lambda I - Q)^{-1} \quad (4.7)$$

where I is the identity matrix and $Sp(Q)$ the set of eigenvalues of the matrix Q .

More precisely, we can show that the power series whose coefficients are expectations of iterated times converges and its value can be expressed using the resolvent.

Proposition 4.3.1. *For all $(u_1, \dots, u_k) \in E^k$ and $0 < |z| < 1$, we have*

$$\sum_{n \geq 0} \mathbb{E}_v [L_{u_1, \dots, u_k; n}(R)] z^n = \frac{z^{-(k-1)}}{1-z} (\mathcal{R}(z^{-1}) - zI)(v, u_1) \dots (\mathcal{R}(z^{-1}) - zI)(u_{k-1}, u_k)$$

where \mathcal{R} is the resolvent defined in (4.7).

Proof. We first make the computations for

$$\hat{L}_{u_1, \dots, u_k; N}(R) = \sum_{0 \leq l_1 \leq \dots \leq l_k \leq N} \mathbf{1}_{R_{l_1} = u_1} \dots \mathbf{1}_{R_{l_k} = u_k}$$

and our main result for $L_{u_1, \dots, u_k; n}(R)$ will follow immediately.

$$\begin{aligned} \mathbb{E}_v \left[\hat{L}_{u_1, \dots, u_k; N}(R) \right] &= \sum_{0 \leq l_1 \leq \dots \leq l_k \leq N} \mathbb{E}_v \left[\mathbf{1}_{R_{l_1} = u_1} \dots \mathbf{1}_{R_{l_k} = u_k} \right] \\ &= \sum_{w \in E} \sum_{0 \leq l_1 \leq \dots \leq l_k \leq N} Q^{l_1}(v, u_1) Q^{l_2 - l_1}(u_1, u_2) \dots Q^{l_k - l_{k-1}}(u_{k-1}, u_k) Q^{N - l_k}(u_k, w) \\ &= \sum_{\substack{m_1 + \dots + m_{k+1} = N \\ m_1, \dots, m_{k+1} \geq 0}} Q^{m_1}(v, u_1) Q^{m_2}(u_1, u_2) \dots Q^{m_k}(u_{k-1}, u_k). \end{aligned}$$

For $(u_1, \dots, u_k) \in E^k$ and $0 < |z| < 1$, we have

$$\begin{aligned} \sum_{n \geq 0} \mathbb{E} \left[\hat{L}_{u_1, \dots, u_k; n}(R) \right] z^n &= \sum_{m_1, \dots, m_{k+1} \geq 0} Q^{m_1}(v, u_1) Q^{m_2}(u_1, u_2) \dots Q^{m_k}(u_{k-1}, u_k) \\ &\times z^{m_1 + \dots + m_{k+1}} = \sum_{m_1 \geq 0} Q^{m_1}(v, u_1) z^{m_1} \dots \sum_{m_k \geq 0} Q^{m_k}(u_{k-1}, u_k) z^{m_k} \\ &\times \sum_{m_{k+1} \geq 0} z^{m_{k+1}} = (1 - z)^{-1} (I - zQ)^{-1}(v, u_1) \dots (I - zQ)^{-1}(u_{k-1}, u_k) \\ &= \frac{z^{-k}}{1 - z} \mathcal{R}(z^{-1})(v, u_1) \dots \mathcal{R}(z^{-1})(u_{k-1}, u_k). \end{aligned}$$

Thus, we deduce that

$$\begin{aligned} \sum_{n \geq 0} \mathbb{E} [L_{u_1, \dots, u_k; n}(R)] z^n &= \sum_{m_1 \geq 1} Q^{m_1}(v, u_1) z^{m_1} \dots \sum_{m_k \geq 1} Q^{m_k}(u_{k-1}, u_k) z^{m_k} \sum_{m_{k+1} \geq 1} z^{m_{k+1}} \\ &= \frac{z^{-(k-1)}}{1 - z} (\mathcal{R}(z^{-1}) - zI)(v, u_1) \dots (\mathcal{R}(z^{-1}) - zI)(u_{k-1}, u_k) \end{aligned}$$

which achieves the proof. \square

Remarks.

- Proposition 4.3.1 can give new possibilities of studying the asymptotics of

$\mathbb{E}[L_{u_1, \dots, u_k; n}(R)]$ using results from complex analysis (for example, tauberian theorems).

- It would be interesting to see whether there is a connection between the resolvent and the quasi-shuffle product which determines the rules of multiplication for occupation times.

4.3.2 Asymptotics of iterated sums/occupation times of hidden Markov walks

Throughout this section, $(R_n, X_n)_n$ will denote a HMW on $E \times V$, $L_{u_1, \dots, u_k; N}(R)$ an iterated time of order N as in definition 3.2.5, $\tilde{S}_{X, k; j_1, \dots, j_k}(N)$ a component of an iterated sum of order k of $(X_n)_n$ as in (3.24) and T_r a stopping time as defined in (3.6).

A decomposition using the excursions of $(R_n)_n$ We have seen how decomposition into pseudo-excursions of a HMW $(R_n, X_n)_n$ on $E \times V$ can be used to prove its convergence in the rough path topology. Based on the excursions of $(R_n)_n$, we can similarly decompose the iterated occupation times of any order. The main idea is to count the number of factors of each product $\mathbf{1}_{R_{k_1}=u_1} \dots \mathbf{1}_{R_{k_l}=u_l}$ that correspond to the same excursion and thus decompose the iterated sum into a combination of iterated sums of lower order. The scheme illustrating an example of repartition in excursions is presented in figure 4.9. Figure 4.9 states that the decomposition of

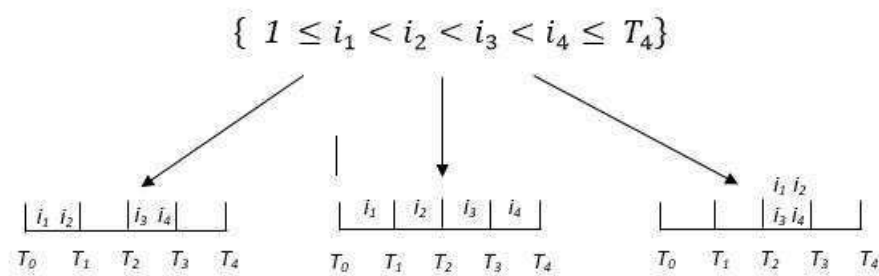


Figure 4.9: An example of repartition of indices in excursions

the iterated occupation time of order 4, $\sum_{1 \leq k_1 < k_2 < k_3 < k_4 \leq T_4} \mathbf{1}_{R_{k_1}=u_1} \cdots \mathbf{1}_{R_{k_4}=u_4}$, with T_4 as in (3.6), contains in particular the terms

- $\sum_{T_0+1 \leq k_1 < k_2 \leq T_1} \mathbf{1}_{R_{k_1}=u_1} \mathbf{1}_{R_{k_2}=u_2} \sum_{T_2+1 \leq k_3 < k_4 \leq T_3} \mathbf{1}_{R_{k_3}=u_3} \mathbf{1}_{R_{k_4}=u_4}$,
- $\sum_{T_2+1 \leq k_1 < k_2 < k_3 < k_4 \leq T_3} \mathbf{1}_{R_{k_1}=u_1} \cdots \mathbf{1}_{R_{k_4}=u_4}$,
- $\sum_{T_0+1 \leq k_1 \leq T_1} \mathbf{1}_{R_{k_1}=u_1} \cdots \sum_{T_3+1 \leq k_4 \leq T_4} \mathbf{1}_{R_{k_4}=u_4}$.

Decomposition of mean values of iterated sums of HMW. We will now use the idea expressed above to find the asymptotics for iterated sums of HMW. Let $(R_n, X_n)_n$ be a HMW on $E \times V$ as before, T_n as in (3.6) and $\tilde{S}_{X,k;j_1,\dots,j_k}(T_r)$ the iterated sum as defined in (3.15).

Lemma 4.3.1. *With the notations from above, the following decomposition holds for all $r, k \in \mathbb{N}^*$, $r > k$:*

$$\mathbb{E} \left[\tilde{S}_{X,k;i_1,\dots,i_k}(T_r) \right] = \sum_{l=1}^k \binom{r}{l} \sum_{1=j_0 \leq j_1 < \dots < j_l=k} \mathcal{I}_{j_0,j_1} \cdots \mathcal{I}_{j_{l-1},j_l} \quad (4.8)$$

where for $p \leq j$,

$$\mathcal{I}_{p,j} = \mathbb{E} \left[\sum_{1 \leq l_1 < \dots < l_{p-j+1} \leq T_1} (\Delta X_{l_1})^{(i_p)} (\Delta X_{l_2})^{(i_{p+1})} \cdots (\Delta X_{l_{p-j+1}})^{(i_j)} \right]. \quad (4.9)$$

Proof. Each of the r intervals $[1, T_1], [T_1 + 1, T_2]$, etc. contains a certain amount of the variables ΔX_m , which goes from 0 to k . So, for any $l = 1, \dots, k$, we have to

- choose l intervals among r that will be occupied (this is what gives the factor $\binom{r}{l}$);
- for any $(m_1, \dots, m_k) \in \Delta_k(T_r)$ (where $\Delta_k(n)$ is as in (3.22)), we split $(\Delta X_{m_1})^{(i_1)}, \dots, (\Delta X_{m_k})^{(i_k)}$ into l groups (one per interval) without changing the order in all the possible ways;
- Since the pseudo-excursions are i.i.d., the expectation of the product $(\Delta X_{m_1})^{(i_1)} \cdots (\Delta X_{m_k})^{(i_k)}$ will be a product of l factors, each of them representing the expectation of an iterated sum until time T_1 and such that the

orders of the l iterated sums sum up to k (which gives the products of type $\mathcal{I}_{j_0, j_1} \cdots \mathcal{I}_{j_{l-1}+1, j_l}$).

□

Limit drift in the case of centred excursions. We now want to see what kind of convergence we have for iterated sums of any order when we have a hidden Markov walk to which theorem 3.2.1 applies. We suppose that E is finite, and we will also need to have centred excursions, i.e. $\mathbb{E}[X_{T_1}] = 0$.

Using lemma 4.3.1, we will see that in this particular case the asymptotics of drifts of iterated sums point surprisingly to rough paths, as the limits of these drifts *will entirely depend on the area anomaly*.

Proposition 4.3.2 (limit drift in the case of centred excursions). *We adopt the notations from lemma 4.3.1 and suppose in addition that $\mathbb{E}[X_{T_1}] = 0$. Then we have the following convergence*

$$\frac{\mathbb{E}[\tilde{S}_{X, m; i_1, \dots, i_m}(T_r)]}{r^{m/2}} \xrightarrow{r \rightarrow \infty} \begin{cases} \frac{1}{(m/2)!} \mathcal{I}_{1,2} \cdots \mathcal{I}_{m-1, m} & \text{if } m \text{ is even} \\ 0 & \text{otherwise} \end{cases}$$

where in particular, for l odd,

$$\mathcal{I}_{l, l+1} = \mathbb{E} \left[\sum_{1 \leq k_1 < k_2 \leq T_1} (\Delta X_{k_1})^{i_l} (\Delta X_{k_2})^{i_{l+1}} \right].$$

Proof. We suppose as before $r \geq m$ (since we want to pass to the limit, it doesn't matter). We will use the result from lemma 4.3.1. For all $k \leq m$

$$L_k(\mathbb{E}[\tilde{S}_{X, m; i_1, \dots, i_m}(T_r)]) = \binom{r}{k} \sum_{1=j_0 \leq j_1 < \dots < j_k = m} \mathcal{I}_{j_0, j_1} \cdots \mathcal{I}_{j_{k-1}+1, j_k}.$$

so that

$$\mathbb{E}[\tilde{S}_{X, m; i_1, \dots, i_m}(T_r)] = \sum_{k=1}^m L_k(\mathbb{E}[\tilde{S}_{X, m; j_1, \dots, j_m}(T_r)]).$$

Suppose first m even. There are then three possibilities.

- For any $m/2 < k \leq m$, among j_1, \dots, j_k , there will always be at least one couple $j_l + 1 = j_{l+1}$ (following the principle: we want to put m rabbits into k boxes), so that

$$\mathcal{I}_{j_l, j_{l+1}} = \mathbb{E}[X_{T_1}] = 0$$

and consequently $L_k(\mathbb{E}[\tilde{S}_{X, m; j_1, \dots, j_m}(T_r)]) = 0$.

- For any $k < m/2$, we have that $L_k(\mathbb{E}[\tilde{S}_{X, m; j_1, \dots, j_m}(T_r)]) = o(r^{m/2})$, as it is of order $\binom{r}{k}$.
- For $k = m/2$, we have either exactly two terms per non-empty excursion or at least one couple of indices $j_l + 1 = j_{l+1}$ (put m rabbits into $m/2$ boxes) and thus terms containing $\mathbb{E}[X_{T_1}]$. Since all the terms containing $\mathbb{E}[X_{T_1}]$ will be 0, we deduce that

$$\begin{aligned} \lim_{r \rightarrow \infty} \frac{\mathbb{E}[\tilde{S}_{X, m; j_1, \dots, j_m}(T_r)]}{r^{m/2}} &= \lim_{r \rightarrow \infty} \frac{L_{m/2}(\mathbb{E}[\tilde{S}_{X, m; j_1, \dots, j_m}(T_r)])}{r^{m/2}} \\ &= \frac{1}{(m/2)!} \mathcal{I}_{j_1, j_2} \cdots \mathcal{I}_{j_{m-1}, j_m}. \end{aligned}$$

Now suppose m is odd.

- For any $k < m/2$, $L_k(\mathbb{E}[\tilde{S}_{X, m; j_1, \dots, j_m}(T_r)]) = o(r^{m/2})$ as before.
- For any $k > m/2$, $L_k(\mathbb{E}[\tilde{S}_{X, m; j_1, \dots, j_m}(T_r)]) = 0$ for the same reason as before.

If in addition $\mathbb{E}[X_{T_1}^{\otimes 2}] = cI$, where I is the identity matrix and $c > 0$, $\mathcal{I}_{j_l, j_{l+1}}$ can be expressed using the area anomaly and the identity matrix (see the relation between the geometric and the non-geometric drift (see (1.48) from the introduction).

□

Bibliographie

- [1] G. Alexopoulos. Random walks on discrete groups of polynomial volume growth. *The Annals of Probability*, 30(2) :723–801, 2002.
- [2] G. Alexopoulos. Sub-laplacians with drift on Lie groups of polynomial volume growth. *American Mathematical Society*, 155(739), 2002.
- [3] J. Angst, I. Bailleul, and C. Tardif. Kinetic Brownian motion on Riemannian manifolds. *Electronic Journal of Probability*, 20 :1–40, 2015.
- [4] I. Bailleul. Flows driven by rough paths. *Revista Iberoam. Math.*, 31(3) :29, 2015.
- [5] E. Baur. An invariance principle for a class of non-ballistic random walks in random environment. *Probability Theory and Related Fields*, 166(1) :463–514, 2016.
- [6] Ch. Bayer and P. Friz. Cubature on Wiener space : Pathwise convergence. *Applied Mathematics & Optimization*, 67(2) :261–278, 2013.
- [7] Ph. Biane. Calcul stochastique non-commutatif. *Lectures on probability theory (Saint-Flour, 1993)*, pages 1–96, 1995.
- [8] Ph. Biane and M. Yor. Variations sur une formule de Paul Lévy. *Annales de l'I.H.P. Probabilités et statistiques*, 23(S2) :359–377, 1987.
- [9] P. Billingsley. *Convergence of Probability Measures*. Wiley, 1999.
- [10] J.-M. Bony. Calcul symbolique et propagation des singularités pour les équations aux dérivées partielles non linéaires. *Annales scientifiques de l'École Normale Supérieure*, 14(2) :209–246, 1981.

- [11] E. Breuillard. Random walks on Lie groups. <http://www.math.u-psud.fr/~breuilla/part0gb.pdf>, 2007.
- [12] E. Breuillard, P. Friz, and M. Huesmann. From random walks to rough paths. *Proc. Amer. Math. Soc.*, 137(10) :3487–3496, 2009.
- [13] Y. Bruned, I. Chevyrev, and P. Friz. Examples of renormalized SDEs. arXiv :1701.01158, 2017.
- [14] Y. Bruned, I. Chevyrev, P. Friz, and R. Preiss. A rough path perspective on renormalization. arXiv :1701.01152, 2017.
- [15] Y. Bruned, M. Hairer, and L. Zambotti. Algebraic renormalisation of regularity structures. arxiv :1610.08468, 2017.
- [16] M. Capitaine and C. Donati-Martin. The Lévy area process for the free Brownian motion. *Journal of Functional Analysis*, 179(1) :153 – 169, 2001.
- [17] R. Catellier. Rough linear transport equation with an irregular drift. *Stochastics and Partial Differential Equations : Analysis and Computations*, 4(3) :477–534, 2016.
- [18] K. Chen. Iterated path integrals. *Bulletin of the American Mathematical Society*, 83(5) :831–879, 1977.
- [19] Y. Chen and W. Geng. Wavelet method for nonlinear partial differential equations of fractional order. *Computer and Information Science*, 4(5), 2011.
- [20] I. Chevyrev. Random walks and Lévy processes as rough paths. *Probab. Theory Relat. Fields*, pages 1–42, 2017.
- [21] L. Coutin and Z. Qian. Stochastic analysis, rough path analysis and fractional Brownian motions. *Probability Theory and Related Fields*, 122(1) :108–140, Jan 2002.
- [22] W. Dahmen, A. Kurdila, and P. Oswald. *Multiscale Wavelet Methods for Partial Differential Equations*. Academic Press, 1997.

- [23] A. Davie. Differential equations driven by rough paths : an approach via discrete approximation. *Appl. Math. Res. Express*, pages 009–40, 2007.
- [24] A. Deya, M. Gubinelli, and S. Tindel. Non-linear rough heat equations. *Probability Theory and Related Fields*, 153(1) :97–147, 2012.
- [25] A. Deya, A. Neuenkirch, and S. Tindel. A milstein-type scheme without Lévy area terms for sdes driven by fractional Brownian motion. *Annales de l’Institut Henri Poincaré, Probabilités et Statistiques*, 48(2) :518–550, 2012.
- [26] A. Deya and R. Schott. On the rough-paths approach to non-commutative stochastic calculus. *Journal of Functional Analysis*, 265 :594–628, 2013.
- [27] J. Diehl, P. Friz, and P. Gassiat. Stochastic control with rough paths. *Applied Mathematics & Optimization*, 75(2) :285–315, 2017.
- [28] S. Eilenberg and S. Mac Lane. On the groups $H(\Pi, n)$, I. *Annals of Mathematics*, 58(1) :55–106, 1953.
- [29] N. Enriquez and Y. Kifer. Markov chains on graphs and Brownian motion. *Journal of Theoretical Probability*, 14(2), 2001.
- [30] S. Fang. Canonical Brownian motion on the diffeomorphism group of the circle. *Journal of Functional Analysis*, 196(1) :162 – 179, 2002.
- [31] P. Friz, P. Gassiat, and T. Lyons. Physical Brownian motion in a magnetic field as a rough path. *Trans. Amer. Math. Soc.*, 367 :7939–7955, 2015.
- [32] P. Friz and M. Hairer. *A Course on Rough Paths - With an introduction to regularity structures*. Springer, 2014.
- [33] P. Friz and N. Victoir. Approximations of the Brownian rough path with applications to stochastic analysis. *Annals of Institute Henri Poincaré*, 41 :703–724, 2005.
- [34] P. Friz and N. Victoir. *Multidimensional Stochastic Processes as Rough Paths : Theory and Applications*. Cambridge University Press, 2010.

- [35] C. Garban, R. Rhodes, and V. Vargas. Liouville Brownian motion. *The Annals of Probability*, 44(4) :3076–3110, 2016.
- [36] Z. Ghahramani. An introduction to hidden Markov models and Bayesian networks. *International journal of pattern recognition and artificial intelligence*, 15(1) :9–42, 2001.
- [37] D. Griffiths and D. Higham. *Numerical Methods for Ordinary Differential Equations*. Springer Undergraduate Mathematics Series, 2010.
- [38] M. Gubinelli. Controlling rough paths. *Journal of Functional Analysis*, 216 :86–140, 2004.
- [39] M. Gubinelli. Ramification of rough paths. *Journal of Differential Equations*, 248(4) :693–721, 2010.
- [40] M. Gubinelli, P. Imkeller, and N. Perkowski. Paracontrolled distributions and singular PDEs. *Forum of Mathematics, Pi*, 3, 2015.
- [41] M. Hairer. Solving the KPZ equation. *Annals of Mathematics*, 178(2) :559–664, 2013.
- [42] M. Hairer. A theory of regularity structures. *Inventiones mathematicae*, 198(2) :269–504, 2014.
- [43] M. Hairer and D. Kelly. Geometric versus non-geometric rough paths. *Annals of Institute Henri Poincaré*, 51(1) :207–251, 2015.
- [44] M. Hairer and H. Weber. Rough burgers-like equations with multiplicative noise. *Probability Theory and Related Fields*, 155(1) :71–126, 2013.
- [45] M. Hoffman. Quasi-shuffle products. *Journal of Algebraic Combinatorics*, 11 :49–68, 2000.
- [46] M. Hoffman. Combinatorics on rooted trees and Hopf algebras. *Trans. Amer. Math. Soc.*, 355 :3795–3811, 2003.
- [47] M. Hoffman and K. Ihara. Quasi-shuffle products revisited. *Journal of Algebra*, 481 :293–326, 2017.

- [48] S. Ishiwata, H. Kawabi, and M. Kotani. Long time asymptotics of non-symmetric random walks on crystal lattices. *Journal of Functional Analysis*, 272(4), 2015.
- [49] K. Itô. Differential equations determining a Markoff process. *Journ. Pan-Japan Math. Coll. (in Japanese)*, 1077, 1942.
- [50] S. Jaffard. Wavelet techniques for pointwise regularity. *Annales de la faculté des sciences de Toulouse Mathématiques*, 15(1) :3–33, 2006.
- [51] J. Norris. *Markov Chains*. Cambridge University Press, 1997.
- [52] G. Jones. On the markov chain central limit theorem. *Probability Surveys*, 1 :299–320, 2004.
- [53] T. Kazami and K. Uchiyama. Random walks on periodic graphs. *Trans. Amer. Math. Soc.*, 360(11) :6065–6087, 2008.
- [54] P. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*. Springer Science and Business Media, 2011.
- [55] M. Kotani and T. Sunada. Albanese maps and off diagonal long time asymptotics for the heat kernel. *Commun. Math. Phys.*, 209(3) :633–670, 2000.
- [56] G. Lawler. *Loop-Erased Random Walk*, pages 197–217. Birkhäuser Boston, 1999.
- [57] M. Ledoux, T. Lyons, and Z. Qian. Lévy area of Wiener processes in Banach spaces. *The Annals of Probability*, 30(2) :546 – 578, 2002.
- [58] A. Lejay. *Introduction to Rough Paths*, volume 1832, pages 1–59. Springer, 2003.
- [59] A. Lejay. *Yet another introduction to rough paths*, volume 42, pages 1–101. Springer, 2009.
- [60] A. Lejay and T. Lyons. On the importance of the Lévy area for studying the limits of functions of converging stochastic processes. application to homogenization. *Current Trends in Potential Theory*, page 63–84, 2003.

- [61] D. Levin and M. Wildon. A combinatorial method for calculating the moments of Lévy area. *Trans. Amer. Math. Soc.*, 360(12), 2007.
- [62] Y. Liu and S. Tindel. Discrete rough paths and limit theorems. arXiv :1707.01783, 2017.
- [63] O. Lopusanschi and D. Simon. Area anomaly and generalized drift of iterated sums for hidden Markov walks. Submitted on HAL : hal-01586794, 2017.
- [64] O. Lopusanschi and D. Simon. Lévy area with a drift as a renormalization limit of Markov chains on periodic graphs. *Accepted for publication in "Stochastic Processes and their Applications"*, 2017.
- [65] T. Lyons. Differential equations driven by rough signals. *Revista Matemática Iberoamericana*, 14(2) :215–310, 1998.
- [66] T. Lyons, M. Caruana, and Th. Lévy. *Differential Equations Driven by Rough Paths (Ecole d'Eté de Probabilités de Saint-Flour XXXIV)*. Springer, 2004.
- [67] T. Lyons and Z. Qian. *System Control and Rough Paths*. Clarendon Press, 2002.
- [68] P. Lévy. *Processus Stochastiques et Mouvement Brownien*. Gauthier-Villars, 1948.
- [69] P. Malliavin. The canonic diffusion above the diffeomorphism group of the circle. *Comptes Rendus de l'Académie des Sciences - Series I - Mathematics*, 329(4) :325 – 329, 1999.
- [70] P. Meyer. Fock space and probability theory. *Stochastic Processes — Mathematics and Physics II. Lecture Notes in Mathematics*, 1250 :160–170, 1987.
- [71] Y. Meyer. *Ondelettes et opérateurs : Ondelettes*. Hermann, 1990.
- [72] D. Nualart. *Application of Malliavin Calculus to Stochastic Partial Differential Equations*. Springer Berlin Heidelberg, 2009.

- [73] D. Nualart and B. Saussereau. Malliavin calculus for stochastic differential equations driven by a fractional Brownian motion. *Stochastic Processes and their Applications*, 119(2) :391–409, 2009.
- [74] S. Qian and J. Weiss. Wavelets and the numerical solution of partial differential equations. *Journal of Computational Physics*, 106(1) :155–175, 1993.
- [75] L. Rabiner. A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2) :257–286, 1989.
- [76] D. Stroock and S. Varadhan. Limit theorems for random walks on Lie groups. *Sankhyā Ser. A*, 35(3) :277–294, 1973.
- [77] T. Sunada. *Topological Crystallography (With a View Towards Discrete Geometric Analysis)*, volume 6 of *Surveys and Tutorials in the Applied Mathematical Sciences*. Springer, 2013.
- [78] N. Victoir. Lévy area for the free Brownian motion : existence and non-existence. *Journal of Functional Analysis*, 208(1) :107 – 121, 2004.
- [79] S. Watanabe. *Lectures on Stochastic Differential Equations and Malliavin Calculus*. Springer-Verlag, 1984.
- [80] B.-J. Yoon. Hidden Markov models and their applications in biological sequence analysis. *Curr Genomics*, 10(6) :402–415, 2009.
- [81] M. Yor. *Remarques sur une formule de Paul Lévy*, pages 343–346. Springer-Verlag, 1979.
- [82] L. Young. *The Theory of Integration*. The University Press, 1927.