HAL
open science

La dynamique des difféomorphismes du cercle selon le point de vue de la mesure

Michele Triestino

To cite this version:

Michele Triestino. La dynamique des difféomorphismes du cercle selon le point de vue de la mesure. Mathématiques générales [math.GM]. Ecole normale supérieure de lyon - ENS LYON, 2014. Français. NNT : 2014ENSL0900 . tel-01065468

HAL Id: tel-01065468

https://theses.hal.science/tel-01065468v1

Submitted on 18 Sep 2014

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

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

THÈSE

en vue de l'obtention du grade de

Docteur de l'Université de Lyon
délivré par l'École Normale Supérieure de Lyon

Discipline : Mathématiques

Laboratoire : Unité de Mathématiques Pures et Appliquées – UMR 5669

École Doctorale : InfoMaths – ED 512

présentée et soutenue publiquement le 21 mai 2014

par M. Michele TRIESTINO

La dynamique des difféomorphismes du cercle selon le point de vue de la mesure

---

Directeur de thèse : M. Étienne GHYS.

Après l'avis de : M. François BÉGUIN ;
M. Nicolas CURIEN.

Devant la commission d'examen formée de :

M. François BÉGUIN,	*membre et rapporteur* ;
M. Nicolas CURIEN,	*membre et rapporteur* ;
M. Étienne GHYS,	*membre et directeur de thèse* ;
M. Livio FLAMINIO,	*membre* ;
M. Grégory MIERMONT,	*membre*.

A papà, perché si commuove sempre.

Prendi la posizione più comoda: seduto sdraiato, raggomitolato, coricato. Coricato sulla schiena, su un fianco, sulla pancia. In poltrona, sul divano, sulla sedia a dondolo, sulla sedia a sdraio, sul pouf. Sull'amaca, se hai un'amaca. Sul letto, naturalmeente, o dentro il letto. Puoi anche metterti a testa in giù, in posizione yoga. Col libro capovolto, si capisce.

Italo Calvino, *Se una notte d'inverno un viaggiatore*

Remerciements

À venir... merci à tout le monde !

Introduction

1 Avant-propos et plan de la thèse

Ce manuscrit est constitué de trois parties différentes qui ont peu en commun, sauf peut-être une idée de fond générale, à savoir la synergie entre systèmes dynamiques et probabilités.

La dynamique peut être bien comprise en regardant les changement des mesures de probabilité au cours de l'évolution, c'est une certitude bien ancrée chez les mathématiciens, et depuis très longtemps ! Il est clair que ces aspects ergodiques peuvent être encore plus finement compris, à partir du moment où le système possède une régularité suffisante. Les complications majeures surgissent lorsque l'on ne dispose pas de compacité : l'existence de mesures intéressantes qui se comportent bien sous les itérations de la dynamique est un événement qui est vu plutôt comme un miracle et est souvent dû à quelque symétrie remarquable.

Voici donc notre point de départ : une mesure et un cocycle, à la fois simples mais énigmatiques sous plusieurs points de vue. Puis leur emploi dans un contexte dynamique, qui ouvre les portes à des retours sur quelques résultats classiques, avec des approches probabilistes. Ceci fait l'objet des chapitres I et II, et le protagoniste principal sera une mesure de probabilité sur le groupe des difféomorphismes C^1 du cercle. Ces parties s'adressent au dynamicien, mais aussi au probabiliste : nous traitons pourtant un côté des systèmes dynamiques qui n'est sûrement pas le plus populaire chez les mathématiciens qui s'intéressent à la théorie des probabilités. Pour cette raison les résultats présentés sont accompagnés par des rappels de notions et faits classiques, de manière peut-être excessivement pédagogique.

Cette plongée dans le monde de la dynamique uni-dimensionnelle et la recherche d'outils appropriés nous a permis de revisiter un beau résultat de théorie ergodique lisse dû à Ávila et Kocsard. Le chapitre III contient essentiellement l'article [Nav-Tri] écrit avec Andrés Navas.

Nous changeons de registre dans le quatrième et dernier chapitre. Inspirés par l'un des problèmes les plus intéressants actuellement en physique statistique, nous montrons comment la compréhension d'un système dynamique sur l'espace de mesures de probabilité peut permettre d'obtenir une géométrie fractale aléatoire riche et intéressante. Bien que le modèle étudié n'est qu'un divertissement mathématique, nous espérons que les outils développés pourront être d'utilité dans le problème originel. Il s'agit d'un travail très récent, poursuivi avec la collaboration de Mikhail Khristoforov et Victor Kleptsyn. Nous nous excusons auprès du lecteur francophone de n'avoir pas traduit la prépublication [Khr-Kle-Tri] pour l'inclure dans ce manuscrit : ce choix accompagne le changement de style et sujet.

Je tiens à remercier profondément mes collaborateurs Andrés, Mikhail et Victor pour avoir donné leur accord à présenter les résultats de [Nav-Tri] et [Khr-Kle-Tri] dans ce manuscrit, ainsi que Sylvain Crovisier et Andrés (à nouveau) pour m'avoir communiqué le résultat qui apparaît dans la partie 2.3 du chapitre III.

2 Présentation des travaux

2.1 Les mesures de Malliavin-Shavgulidze

La première partie de ce travail voit son origine dans un article présentant une preuve erronée, témoignant le fait que même les « non-théorèmes » ont leur utilité en mathématiques. En 2009, E. T. Shavgulidze croyait avoir résolu le fameux problème de la moyennabilité du groupe de Thompson [Sha 5]. Désormais les spécialistes ont compris que la preuve de Shavgulidze est fausse et qu'en outre il n'est pas possible de la modifier pour qu'elle soit correcte. En revanche, Shavgulidze a eu l'occasion de faire connaître à une plus grande communauté, une très belle mesure de probabilité sur le groupe des difféomorphismes C^1 du cercle.

Cette mesure joue le rôle de la mesure de Haar définie sur un groupe topologique localement compact : il existe toujours, sur un groupe topologique localement compact, une mesure de Radon qui est invariante sous l'action du groupe par multiplication à gauche et elle est unique à multiplication scalaire près. Elle prend le nom de mesure de Haar et son existence caractérise les groupes topologiques localement compacts : il est impossible

de trouver une mesure invariante par l'action de composition à gauche, sur un groupe topologique qui n'est pas (un sous-groupe dense dans) un groupe localement compact.

Nous aurons la possibilité de faire des rappels plus détaillés dans ce texte. Expliquons plutôt ici l'avantage théorique et pratique de disposer d'une mesure de Haar sur un groupe. Il est bien connu que tout groupe fini est isomorphe à un sous-groupe du groupe symétrique sur un nombre d'objets égal à son ordre. Donc le groupe symétrique sur n éléments contient une copie de tous les groupes de cardinal au plus n. Notons que l'hypothèse de finitude n'est pas nécessaire : tout groupe s'injecte dans le groupe des bijections d'un ensemble de cardinal égal à l'ordre du groupe. En revanche, lorsque le cardinal n'est plus fini, le groupe symétrique est trop grand pour être compris : nous nous en sortons avec l'emploi de l'analyse. Si un groupe G possède une mesure de Haar μ, on peut regarder l'espace de Hilbert des fonctions L^2 sur G :

$$L^2(G,\mu) = \left\{ f : G \longrightarrow \mathbf{C} \text{ mesurable} \,\middle|\, \int_G f^2 \, d\mu < \infty \right\}.$$

Or, l'action de G par multiplication à gauche induit une action linéaire sur $L^2(G,\mu)$: étant donné $g \in G$, on associe à $f \in L^2(G,\mu)$ la fonction $L_g f$ qui fait correspondre à tout $x \in G$ la valeur $f(g^{-1}x)$. Puisque la mesure μ est invariante, cette action est aussi unitaire. La conséquence que l'on en tire est la suivante : tout groupe localement compact est isomorphe à un sous-groupe du groupe des opérateurs unitaires dans un espace de Hilbert.

Une telle vision est utile quand on s'intéresse à la théorie des représentations linéaires d'un groupe, mais elle est rarement d'utilité lorsque l'on veut faire de la dynamique non-linéaire. Une (rare ?) exception arrive quand on se place dans le monde des groupes de Lie (on pense ici aux grands résultats ergodiques obtenus pour les réseaux depuis Margulis). Mais encore, le dynamicien peut-il être satisfait si un problème apparemment simple comme celui de comprendre l'image de la mesure de Haar par l'application d'élévation à puissance $g \mapsto g^k$, est difficilement abordable ? Même pour le groupe linéaire $GL_n\mathbf{R}$ cette question n'est pas si simple. Les mesures de Haar ne sont pas proprement dynamiques : les objets naturels en dynamique sont ceux qui sont invariant par conjugaison, non par composition d'un coté ![1]

Cependant, une mesure « homogène » comme la mesure de Haar, est assez appropriée pour décrire le comportement d'un élément typique sous un aspect différent de la généricité au sens de Baire : très souvent on rencontre les mêmes phénomènes sous ces deux points de vue, mais de temps en temps les deux descriptions peuvent être contradictoires. Pour donner un exemple fortement relié à la dynamique sur le cercle, les nombres diophantiens forment une partie de mesure de Lebesgue totale dans $\mathbf{R}$, alors que les nombres de Liouville contiennent un G_δ dense.

On trouve beaucoup de groupes qui ne sont pas localement compacts mais dignes d'intérêt. Les plus importants sont les groupe de Lie de dimension *infinie*, comme les groupes des lacets dans un groupe de Lie [Pre-Seg], les opérateurs unitaires d'un espace de Hilbert ou les groupes de difféomorphismes de variétés [Arn 2, Khe-Wen]. Des mathématiciens comme Arnol'd et Kirillov ont souligné comme ces groupes peuvent être la source de belles géométries. Il ne faudra non plus négliger le rôle que les groupes de difféomorphismes jouent dans la théorie de feuilletages, puisque leur cohomologie permet d'obtenir des invariants [Fuc, Mor, Ghy-Ser]. En conclusion du premier chapitre nous essayerons de nous approcher de ce côté, en discutant notamment de l'*invariant de Godbillon-Vey* (voir par exemple [Ghy 1] pour une introduction).

Malliavin et Shavgulidze trouvent une bonne façon de choisir uniformément au hasard un difféomorphisme du cercle et de cette manière ils récupèrent une mesure qui est *quasi*-invariante sous une certaine action de multiplication [Mal-Mal 1, Mal-Mal 2, Sha 3]. L'idée est brillante, l'explication très simple ! Depuis Wiener [Wie 1], nous savons choisir uniformément une fonction continue sur le cercle de moyenne nulle : nous appelons une telle fonction aléatoire *pont brownien*. Ensuite on définit le difféomorphisme aléatoire par intégration, en choisissant uniformément par rapport à la mesure de Lebesgue l'image d'un point donné (par exemple 0) : en termes plus précis, si $B = (B_t)_{t\in[0,1]}$ est un pont brownien et λ une variable aléatoire uniforme sur $[0,1]$ indépendante de B, nous définissons le difféomorphisme aléatoire de Malliavin-Shavgulidze par

$$f(t) = \frac{\int_0^t \exp(B_s)\, ds}{\int_0^1 \exp(B_s)\, ds} + \lambda.$$

Par suite, nous disposons d'une mesure μ_{MS} sur le groupe $\mathrm{Diff}^1_+(\mathbf{S}^1)$ qu'on appelle *mesure de Malliavin-Shavgulidze*.

Un resultat de Cameron et Martin classique [Cam-Mar 1] montre que la mesure de Wiener sur l'espace des fonctions continues $C_0(\mathbf{S}^1)$ est de *type Haar* : par cela nous voulons dire qu'il existe un sous-groupe dense de $C_0(\mathbf{S}^1)$ (contenant les fonctions dérivables) qui préserve la classe de la mesure de Wiener lorsqu'il agit par

[1] On pourrait argumenter qu'une mesure de Haar μ invariante à gauche permet de construire une mesure bi-invariante (et donc invariante par conjugaison) : si $\iota : G \longrightarrow G^{op}$ est le morphisme vers le groupe opposé, $\iota(g) = g^{-1}$, alors $\mu * \iota_*\mu$ est bi-invariante. Mais il s'agit d'une construction très artificielle.

translations sur le groupe (à savoir, il préserve les ensembles de mesure nulle). Il existe un autre théorème, moins connu, de Cameron et Martin [Cam-Mar 2], qui implique entre autres que la mesure de Malliavin-Shavgulidze est de type Haar pour $\mathrm{Diff}^1_+(\mathbf{S}^1)$: le sous-groupe des difféomorphismes C^2 préserve la classe de la mesure de Malliavin-Shavgulidze.

Dans le premier chapitre de cette thèse, nous avons recueilli les résultats principaux sur les mesures de Malliavin-Shavgulidze, en rajoutant des détails aux preuves qui en nécessitent. En particulier nous présentons le théorème de Shavgulidze qui explicite la dérivée de Radon-Nikodym associée à l'action par multiplication à gauche de $\mathrm{Diff}^3_+([0,1])$ sur $(\mathrm{Diff}^1_+(\mathbf{S}^1), \mu_{MS})$:

$$\frac{d(L_\varphi)_*\mu_{MS}}{d\mu_{MS}}(f) = \exp\left\{\int_{\mathbf{S}^1} \mathcal{S}_\varphi(f(t))\,(f'(t))^2\,dt\right\},$$

puis nous rappelons le théorème de Kosyak qui montre l'ergodicité de cette action [Kos].

Bien évidemment, lorsque l'on dispose d'une mesure quasi-invariante à gauche comme μ_{MS}, on peut définir la mesure $d\tilde{\mu}_{MS}(f) = d\mu_{MS}(f^{-1})$, qui est naturellement quasi-invariante *à droite*. Ensuite, par convolution on peut obtenir aussi une mesure *quasi-bi-invariante* $\mu_{MS} * \tilde{\mu}_{MS}$: elle est peut-être plus intéressante d'un point de vue dynamique, puisqu'elle est quasi-invariante par l'action de conjugaison aussi. On perd cependant la possibilité d'exprimer facilement le difféomorphisme aléatoire, à l'aide du pont brownien.

Il ne faudra pas cacher que beaucoup d'autres mathématiciens ont construit de mesures sur les groupes de difféomorphismes des variétés. Malgré tous les efforts, il est très difficile d'obtenir de belles mesures lorsque les variétés sont de dimension plus grande que 1 : l'une des difficultés majeures est la nature algébrique de ces groupes de Lie de dimension infinie, qui devient trop compliquée. Il serait très intéressant de disposer d'une mesure quasi-invariante sur les difféomorphismes du tore ou du disque, qui préservent l'aire.

Nous n'essayerons pas de décrire d'autres exemples de mesures, car chacune porte ses caractéristiques spécifiques et il est presque impossible de les présenter sous une théorie unificatrice. La simplicité, la propriété de quasi-invariance et les propriétés de régularité nous ont convaincus que la mesure de Malliavin-Shavgulidze mérite quelques attentions supplémentaires.

Dans le chapitre II, nous entreprenons l'étude *dynamique*. Comme d'habitude, on essaye de décrire le comportement d'un difféomorphisme à conjugaison près. Depuis le travail fondateur de Poincaré [Poi], il est très naturel de comprendre, en premier lieu, quelle pourrait être la distribution du *nombre de rotation*, qui est l'invariant de conjugaison le plus célèbre.[2] Il s'avère que ce problème semble difficile d'accès.

À cause de la présence de difféomorphismes avec orbites périodiques hyperboliques, les difféomorphismes avec orbites périodiques (et donc nombre de rotation *rationnel*), constituent une partie de mesure strictement positive.

Que peut-on dire sur les orbites d'un difféomorphisme choisi uniformément par rapport à la mesure de Malliavin-Shavgulidze ? Existe-t-il presque toujours une orbite périodique ?

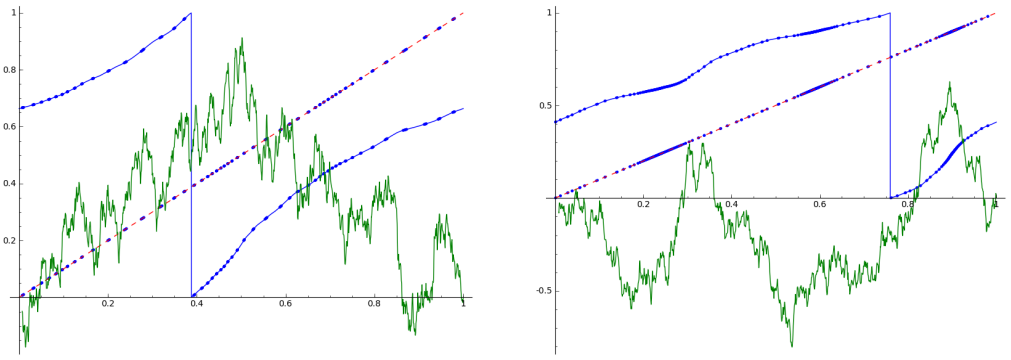

Figura V.1 – Deux exemples de difféomorphismes aléatoires, les ponts browniens qui les définissent et la trace d'une orbite.

Nos résultats donnent une bonne description conditionnement à l'événement qu'il y ait des orbites périodiques :

Théorème 2.1. *Les difféomorphismes du cercle qui possèdent un nombre fini d'orbites périodiques forment une partie de mesure de Malliavin-Shavgulidze totale parmi ceux qui ont des points périodiques.*

[2] Pour le lecteur qui n'est pas initié à la théorie des difféomorphismes du cercle, nous parcourons les aspects principaux dans la partie 3 du chapitre introductif.

La preuve s'appuie fortement sur les propriétés du mouvement brownien et elle s'inspire de la démonstration du fait que le mouvement brownien planaire ne passe presque sûrement par aucun point donné (que nous avons trouvée dans [Mör-Per]).

Plus précisément, nous montrons que presque sûrement les orbites périodiques sont *hyperboliques* : s'il existe une orbite périodique, alors on en trouve un nombre pair, une moitié est répulsive, l'autre est attractive. Les points sur le cercle qui n'appartiennent pas à l'une de ces orbites sont attirés vers deux d'entre elles, vers l'une dans le futur et vers l'autre dans le passé. Ce genre de difféomorphismes prend le nom de *Morse-Smale*. Une telle propriété permet d'étudier de manière satisfaisante la dynamique, même en faible régularité : en généralisant un résultat classique de Sternberg, les difféomorphismes Morse-Smale de classe $C^{1+\alpha}$ peuvent être localement linéarisés, on peut borner leur distorsion et, en suivant N. Kopell [Kop], on peut facilement décrire leur centralisateur (et donc leur classe de conjugaison).

Théorème 2.2. *Les difféomorphismes du cercle qui ont centralisateur C^1 trivial forment une partie de mesure de Malliavin-Shavgulidze totale parmi ceux qui ont des points périodiques.*

Cette description confirme l'intuition que l'on a au niveau topologique : dans la topologie C^1, les difféomorphismes de type Morse-Smale, avec centralisateur C^1 trivial forment une partie ouverte et dense de $\mathrm{Diff}^1_+(\mathbf{S}^1)$ [Kop, Yoc 2]. Les théorèmes précédents montrent que le complémentaire est invisible d'un point de vue de la mesure, au moins parmi les difféomorphismes avec points périodiques.

La preuve du théorème 2.1 peut être adaptée pour démontrer un résultat qui est très intéressant en soi :

Théorème 2.3. *Soient $f_1, \dots, f_n$ des difféomorphismes de Malliavin-Shavgulidze indépendants. Alors le sous-groupe qu'ils engendrent est presque-sûrement libre (de rang n).*

Ce théorème confirme l'observation topologique : É. Ghys démontre dans [Ghy 2] que génériquement dans $\mathrm{Homéo}_+(\mathbf{S}^1)$, deux homéomorphismes engendrent un sous-groupe libre. Il est aussi possible de voir le théorème 2.3 comme la généralisation du théorème de Epstein [Eps] dans le contexte des groupes de Lie au groupe de Lie de dimension infinie $\mathrm{Diff}^1_+(\mathbf{S}^1)$.

Remarque 2.1. L'espace des actions fidèles du groupe libre sur le cercle se sépare qualitativement en deux sous-parties : les actions peuvent minimales ou laisser un ensemble de Cantor invariant. Par rapport à la topologie naturelle, ces deux sous-parties contiennent des ouverts non-vides. On en déduit que les deux différentes dynamiques sont observables avec probabilité strictement positive.

Pour ce qui concerne la présence des difféomorphismes sans orbites périodiques, nous pensons qu'ils forment un ensemble négligeable : une partie du deuxième chapitre est dédiée à nos conjectures à ce propos. Sans rentrer dans le détail, nous terminons cette introduction par une image qui montrent la statistique du nombre de rotation, sur un échantillonnage d'un millier de difféomorphismes « aléatoires ».

2.2 Distributions invariantes pour les difféomorphismes du cercle

L'étude des équations cohomologiques fait partie des incontournables des systèmes dynamiques : les exemples les plus célèbres [Sie, Kol, Arn 1] viennent de la théorie KAM et plus généralement des problèmes de linéarisation (reliés à la théorie des petits diviseurs). Dans les cas les plus classiques, on cherche les solutions u des équations de la forme

$$u \circ f - u = h, \tag{V.1}$$

étant donné une fonction continue h sur une variété compacte orientée M et f difféomorphisme de cette variété qui préserve l'orientation.

On remarque facilement qu'une condition nécessaire pour résoudre (V.1) est que la fonction h ait μ-moyenne nulle pour toute mesure f-invariante μ sur M. En d'autres termes, les fonctions de la forme $u \circ f - u$ (les *cobords*) sont contenues dans l'espace vectoriel qui est défini par l'intersection de tous les noyaux des formes linéaires obtenues par intégration contre une mesure f-invariante.

En appliquant le théorème de Hahn-Banach, on trouve que les cobords sont denses dans cette intersection : si $\int h\, d\mu = 0$ pour toute mesure invariante μ, alors l'équation (V.1) ne possède pas forcément de solution, mais h peut être approché par des cobords. Pour donner un exemple d'application récente de cette observation, A. Navas a démontré que tout difféomorphisme C^1 du cercle avec nombre de rotation irrationnel α, peut être rapproché aussi près que l'on veut de la rotation R_α, dans la topologie C^1, par des conjugaisons [Nav 2].

On peut trouver des exemples de difféomorphismes f tels que l'espace des cobords $u \circ f - u$ soit fermé : lorsque f est une rotation diophantienne du tore il s'agit d'un petit exercice d'analyse de Fourier, si f est un difféomorphisme d'Anosov transitif il s'agit du théorème classique de Livsič (voir par exemple [Kat-Has]).

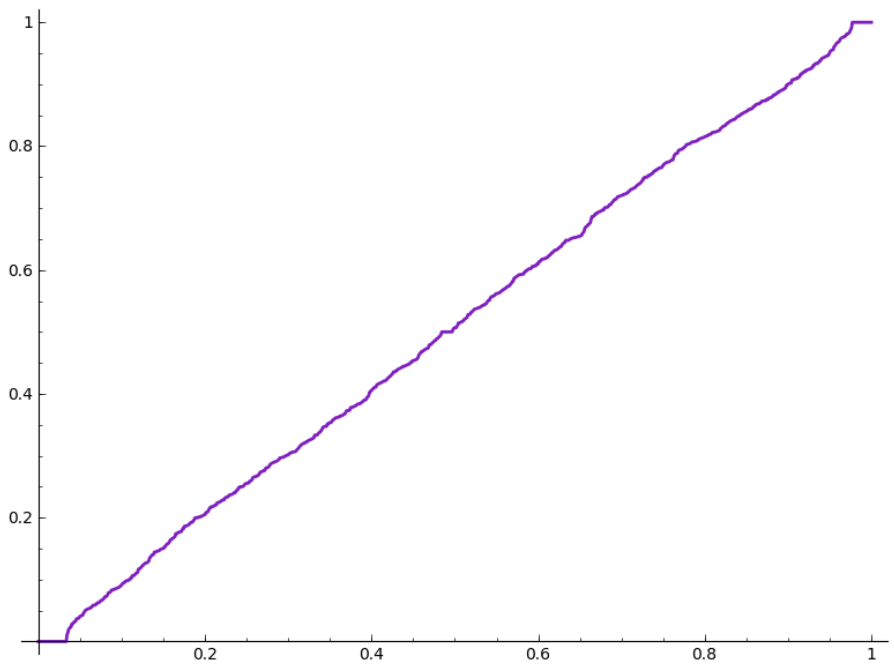

Figura V.2 – Cet escalier représente la fonction de répartition de la variable *nombre de rotation*, selon la mesure de Malliavin-Shavgulidze.

Supposons maintenant que f et h soient de classe C^r, $r \in [1, \infty]$. On remarquera que si l'on cherche des solutions C^r de l'équation (V.1), il faut modifier la condition nécessaire expliquée précédemment en demandant que h appartienne au noyau de toute *distribution* f-invariante d'ordre r (une distribution d'ordre r est une fonctionnelle linéaire sur $C^r(M)$). L'étude des distributions invariantes est donc la première étape pour résoudre (V.1).

Vers 1980, en réfléchissant autour du théorème de Livsič dans ces termes, E. Ghys pose le problème de déterminer quelles sont les distributions invariantes par un difféomorphisme C^2 minimal du cercle.

Remarque 2.2. Plus ou moins vers la même époque, des problèmes concernant les feuilletages minimaux des variétés riemanniennes induisent Haefliger à étudier les courants invariants par l'holonomie (les courants sont l'analogue des distributions pour les formes différentielles). Lorsque le feuilletage est obtenu par suspension d'une action sur le cercle d'un groupe de type fini Γ, on recherche en effet une distribution Γ-invariante sur le cercle [Hae 2, Hae-Ban].

Le résultat remarquable contenu dans [Hae-Ban] montre que certains groupes fuchsiens possèdent des distributions invariantes (alors qu'ils ne préservent pas de mesures !).

La réponse à la question de Ghys arrive 30 ans plus tard :

Théorème 2.4 (Ávila – Kocsard [Ávi-Koc 1]). *Soit $f : \mathbf{S}^1 \longrightarrow \mathbf{S}^1$ un difféomorphisme C^∞ du cercle avec nombre de rotation irrationnel. Alors, toute distribution f-invariante est un multiple scalaire de l'unique mesure de probabilité f-invariante.*

En utilisant certaines idées dans [Nav 2] et les résultats de [Dou-Yoc], il est possible d'obtenir une preuve simple et géométrique dans l'esprit du théorème 2.4, en affaiblissant les hypothèses de régularité :

Théorème 2.5. *Soit f un difféomorphisme du cercle de classe C^{1+vb} et nombre de rotation irrationnel. Alors, toute distribution d'ordre 1 qui est f-invariante est un multiple de la mesure invariante μ_f.*

En revanche, on observe que l'hypothèse de régularité C^{1+vb} est cruciale :

Théorème 2.6. *Pour tout nombre irrationnel α, il existe un difféomorphisme C^1 avec nombre de rotation α qui possède une distribution d'ordre 1 invariante qui n'est pas proportionnelle à la mesure invariante.*

De plus, le difféomorphisme peut être choisi minimal ou bien avec un ensemble de Cantor comme compact minimal invariant.

Les exemples pour le théorème 2.6 apparaissent déjà dans les travaux de R. Douady et J.-C. Yoccoz [Dou-Yoc] et de H. Kodama et S. Matsumoto [Kod-Mat].

Un autre aspect que l'étude des distributions invariantes permet de comprendre concerne les *déviations ergodiques* : si μ est une mesure invariante ergodique pour f, d'après le théorème de Birkhoff, pour toute fonction

$u \in C(M)$ la suite des moyennes

$$\frac{1}{n}\left(u + u \circ f + \ldots + u \circ f^{n-1}\right)$$

converge presque partout vers l'intégrale

$$\int_M u \, d\mu.$$

Cependant les quantités

$$u + u \circ f + \ldots + u \circ f^{n-1} - n \cdot \int_M u \, d\mu$$

peuvent ne pas être uniformément bornées. Dans le cas des systèmes hyperboliques ou paraboliques ceci est assuré par des théorèmes fameux (voir le théorème de Ratner pour les flots unipotents, le théorème central limite pour les flots géodésiques en courbure négative et les résultats en dynamique de Teichmüller). Moins d'exemples sont connus pour la dynamique elliptique et le théorème de Ávila et Kocsard en est l'un des premiers qui permettent de s'éloigner des cas linéaires.

Le résultat suivant est une amélioration en régularité d'un corollaire présent dans [Ávi-Koc 1].

Corollaire 2.7. *Soit f un difféomorphisme du cercle de classe C^{1+vb} et nombre de rotation irrationnel α. On note par (p_n/q_n) la suite des approximations rationnelles de α. Alors, pour toute fonction $u \in C^1(\mathbf{S}^1)$ la suite*

$$u + u \circ f + \ldots + u \circ f^{q_n - 1} - q_n \int_{\mathbf{S}^1} u \, d\mu_f$$

converge uniformément vers 0.

2.3 Gravité quantique et graphes hiérarchiques

Ces dernières années, un problème fascinant a réuni les attentions de physiciens, probabilistes et géomètres : d'après des célèbres travaux en physique théorique (en particuliers en théorie de cordes) [Pol 1,Pol 2,Kni-Pol-Zam], on voudrait comprendre

1. comment définir rigoureusement la métrique aléatoire $e^{\gamma X}|dz|$ sur une surface, avec X *champ libre gaussien*,
2. comprendre ensuite ses caractéristiques typiques (sa dimension d'Hausdorff, par exemple).

L'une des motivations physiques provient de la résolution de ces problèmes, qui permettrait de décrire beaucoup de phénomènes en régime critique (le modèle de Ising, par exemple). Pour plus d'explications, nous renvoyons à l'exposition de C. Garban [Gar].

Parmi les différentes approches de ce problème, on peut en trouver au moins deux qui ont donné lieu à de très beaux résultats mathématiques. Essayons de les décrire brièvement..

La première se fait à travers l'étude des *cartes planaires*. Pour fixer les idées, supposons que la surface que l'on étudie soit la sphère $\mathbf{S}^2$. Une carte planaire est un objet topologique et combinatoire : un plongement propre d'un graphe fini dans $\mathbf{S}^2$, à homéomorphisme de la sphère près. Particulièrement intéressantes sont les cartes planaires avec nombre et géométrie des faces fixés : grâce à la bijection de Cori-Vauquelin-Schaeffer on sait calculer explicitement le nombre de cartes planaires avec faces quadrangulaires et n sommets (voir [LeG-Mie]). Il s'agit d'un ingrédient fondamental dans les résultats de J.-F. Le Gall [LeG-Pau,LeG] et G. Miemont [Mie 1,Mie 2] qui décrivent la limite d'échelle des cartes (munies de leur métriques de graphe) lorsque l'on fait tendre le nombre de sommets n vers l'infini : en rénormalisant les distances par un facteur proportionnel à $n^{-1/4}$, on converge en loi vers un espace métrique aléatoire qui a presque sûrement la topologie d'une sphère mais qui est fractal sous le point de vue métrique, en étant de dimension de Hausdorff 4.

La deuxième approche (introduite indépendamment dans les travaux des physiciens [Dav,Dis-Kaw]) est issue de l'*exponentielle du champ libre gaussien*. Pour définir une famille de variables aléatoires gaussiennes $\{X(x)\}_{x\in\Omega}$ il est nécessaire et suffisant de connaitre les espérances $\{\mathbf{E}[X(x)]\}_{x\in\Omega}$ et les covariances $\{\mathrm{Cov}(X(x), X(y))\}_{x,y\in\Omega}$. Par exemple, on peut voir le pont brownien sur $[0,1]$ comme une famille de variables aléatoires gaussiennes indépendantes $\{X(t)\}_{t\in[0,1]}$, centrées et telles que

$$\mathrm{Cov}(X(s), X(t)) = \min\{s,t\} - s\,t.$$

Il s'agit d'un premier exemple de champ libre gaussien (*GFF* dans la littérature anglophone).

On va procéder de manière similaire pour définir le champ libre gaussien sur un domaine Ω du plan : pour des raisons d'invariance conforme, on définit la famille de variables gaussiennes indépendantes $\{X(x)\}_{x\in\Omega}$, centrées et telles que

$$\mathrm{Cov}(X(x), X(y)) = G_\Omega(x,y),$$

où G_Ω est la fonction de Green sur Ω. Or, on remarquera quelque souci : la fonction de Green prend la valeur infinie sur la diagonale et on peut montrer que le processus $\{X(x)\}_{x\in\Omega}$ définit en vérité une *distribution*. Pour des détails plus précis, nous recommandons l'article introductif de S. Sheffield [She 1].

Une manière alternative de définir *GFF* sur Ω (avec condition de Dirichlet au bord) est la suivante : soit $(\Phi_n)_{n\in\mathbf{N}}$ une base de l'espace de Sobolev $H_0^1(\Omega)$ (orthogonal aux constantes), alors la variable

$$h = h_\Omega = \sum_{n\in\mathbf{N}} a_n \Phi_n$$

définit presque sûrement le champ libre gaussien sur Ω, lorsque l'on prend des variables gaussiennes standards indépendantes pour les coefficients $(a_n)_{n\in\mathbf{N}}$.

Il s'avère qu'il est possible de donner du sens à l'exponentielle de *GFF*, même si l'on est en train de prendre l'exponentielle d'une distribution (en utilisant la théorie de Kahane [Kah]). Il est conjecturé que la « métrique »

$$g_h = e^{\gamma\, h(x)}|dx| \tag{V.2}$$

sur une surface Ω donne la métrique aléatoire « uniforme » (avec $0 < \gamma < 2$ fixé). Cette conjecture a été renforcée par les récents résultats de R. Rhodes et V. Vargas [Rho-Var 1] et B. Duplantier et S. Sheffield [Dup-She], qui ont étudié la *mesure*

$$\mu_h = e^{\gamma\, h(x)}dx$$

dans le contexte de la gravité quantique : ils obtiennent une expression qui relie de manière quadratique la dimension de Hausdorff euclidienne d'une partie de Ω avec sa dimension de Hausdorff par rapport à la mesure « quantique » μ_h (conjecturée dans [Kni-Pol-Zam] et connue sous le nom de *relation KPZ*).

On peut apporter une légère modification au problème précédent, reliée à la version *dyadique* du champ libre gaussien.

Supposons que Ω soit un carré et g une métrique aléatoire sur Ω. On peut prendre quatre réalisations indépendantes de (Ω, g) et les recoller de manière naturelle pour obtenir à nouveau un carré, avec une nouvelle métrique g' qui est définie en prenant le plus court chemin qui minimise la distance entre deux points (voir la figure V.3). Puis on perturbe g' en multipliant toutes les distances par un facteur aléatoire ξ (par exemple de loi log-normale) et on trouve une métrique $\tilde{g}$ en se permettant une remise à l'échelle par un facteur $\lambda > 0$ fixé. Ce modèle peut se reconduire à l'exponentielle de GFF, en sélectionnant une base orthonormée plus « géométrique » pour l'espace de Sobolev.

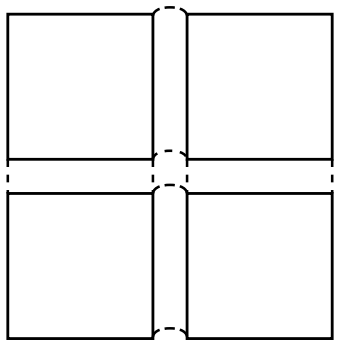

Figura V.3 – Recoller quatre carrés pour en obtenir un autre

Question 1. Existe-t-il une loi (non-atomique) sur les métriques définies sur le carré qui soit stationnaire par le processus de recollement, pour un certain λ ?

L'une des majeures difficultés que l'on rencontre est due aux recollement des métriques le long du bord : sans aucun contrôle *a priori* sur la régularité (Hölder) des géodésiques, il pourrait s'avérer que d'un point de vue métrique les bords s'évanouissent en un point.

Dans le travail fait avec Mikhail Khristoforov et Victor Kleptsyn, nous avons choisi de contourner ce problème : nous considérons des modèles jouets qui présentent des points de passage obligatoires pour les chemins géodésiques. En suivant une suggestion de I. Benjamini [Ben], nous avons étudié les métriques aléatoires stationnaires sur des objets auto-similaires obtenus comme limite de graphes finis : les *graphes hiérarchiques*. Comme pour le carré, on peut prendre un nombre fini de copies d'un graphe hiérarchique et puis les recoller en un nombre fini de points pour obtenir un nouvel espace qui a la même topologie. Dans la figure V.4, on peut observer le processus itératif pour la construction du graphe hiérarchique à forme de huit, que nous avons choisi comme modèle pour notre étude.

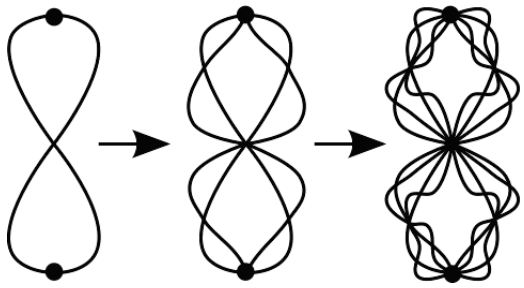

Figura V.4 – Premières étapes de la construction du huit hiérarchique

En se limitant dans un premier temps à regarder la variable aléatoire qui donne la distance entre les deux points marqués, en haut et en bas, on recherche un point fixe de l'opérateur Φ_λ qui agit sur l'espace des mesures de probabilité sur $(0, +\infty)$, et qui à une mesure μ associe la loi de la variable

$$\lambda\, \xi \cdot (\min(X^1, X^2) + \min(X^3, X^4)),$$

où les (X^i) sont indépendantes et identiquement distribuées selon μ, la variable ξ est une autre variable indépendante des précédentes et de loi m donnée (par exemple log-normale), et $\lambda > 0$ un facteur positif arbitraire.

Dans [Khr-Kle-Tri], nous avons démontré le résultat suivant:

Théorème 2.8. *Pour toute mesure de probabilité non-atomique et de support total m sur $\mathbf{R}_+$, il existe un unique $\lambda_{cr} > 0$ et une mesure de probabilité non-atomique $\bar{\mu}$ sur $\mathbf{R}_+$ telle que $\bar{\mu}$ est un point fixe de l'opérateur $\Phi_{\lambda_{cr}}$.*

Le preuve de ce théorème utilise une méthode de cut-off assez fine qui nous permet de contrôler les endroits où la métrique explose. On définit l'ensemble sur-critique Λ comme étant la réunion des λ pour lesquels on peut trouver une métrique « coupée » stationnaire non-triviale. On peut facilement vérifier que Λ est une demi-droite infinie bornée à gauche par un certain λ_{cr}. Le « lemme clef » assure que λ_{cr} n'appartient pas à Λ et ceci permet d'obtenir la loi stationnaire en faisant une double limite appropriée ($\lambda \downarrow \lambda_{cr}$ et $A \uparrow \infty$).

La structure hiérarchique permet ensuite de définir une distance aléatoire entre tous les points « dyadiques » du graphe hiérarchique. Cependant, il n'est pas toujours possible de définir une distance presque sûrement finie d'un point quelconque vers presque tout autre point du huit : des boules infinies peuvent être présentes « presque partout » dans le graphe (ce phénomène fait penser à ce qui se passe pour l'exponentielle du champ libre gaussien dans le cas *sur-critique* $\gamma > 2$).

Lorsque la mesure m vérifie des conditions supplémentaires, nous pouvons compléter l'énoncé du théorème 2.8, en montrant que la mesure stationnaire $\bar{\mu}$ est *unique* à facteur près :

Théorème 2.9. *Soit m une mesure de probabilité sur $\mathbf{R}_+$ absolument continue et avec densité continue et strictement positive, telle que la densité de $\log_* m$ tend vers 0 en $\pm\infty$.*

Soit $\bar{\mu}$ une mesure de probabilité sur $\mathbf{R}_+$ telle que $\bar{\mu} = \Phi_{\lambda_{cr}}[\bar{\mu}]$. Alors, pour toute mesure μ sur $\mathbf{R}_+$ il existe une constante $c > 0$ telle que les itérées de cette mesure convergent vers la mesure $\bar{\mu}$ remise à l'échelle par le facteur c :

$$\Phi^n_{\lambda_{cr}}[\mu] \longrightarrow \bar{\mu}(c\,\cdot) \quad \text{as } n \longrightarrow \infty.$$

En particulier, la mesure de probabilité $\Phi_{\lambda_{cr}}$-stationnaire est unique à remise à l'échelle près.

Remarque 2.3. Tout récemment, nous avons pu améliorer le résultat précédent en démontrant que la convergence est à une vitesse exponentielle. Cela nous permet de déduire qu'il est possible de définir presque sûrement une métrique sur le huit si l'on connaît toute la famille (dénombrable) des facteurs ξ que l'on utilise pour définir la distance entre tous les points dyadiques. La preuve envisagée est assez technique et fera l'objet d'un futur travail.

Revenons maintenant à la question 1 : comment utiliser les idées provenant du théorème 2.8 pour trouver une métrique aléatoire stationnaire sur le carré ?

La première étape consiste à étudier des espaces pour lesquels connaître la loi de la distance entre deux seuls points n'est pas suffisante pour reconstruire la distance entre tous les points dyadiques : on peut considérer en exemple le *triangle de Sierpinski*. Des modifications appropriés de nos arguments permettent encore de trouver une distance aléatoire stationnaire non-triviale.

Ensuite on peut considérer des graphes auto-similaires qui ressemblent à des carrés (voir la figure V.5) et obtenir une suite de métriques aléatoires desquelles, si on démontre qu'elles appartiennent à un espace compact,

on pourrait extraire une sous-suite qui converge vers une métrique aléatoire non-triviale sur le carré. Cette approche, certainement fascinante, demande un grand nombre de passages techniquement assez difficiles, que pour l'instant nous ne savons pas maîtriser et qui sont assez proches des difficultés rencontrées dans les méthodes classiques de résolution du problème de la gravité quantique.

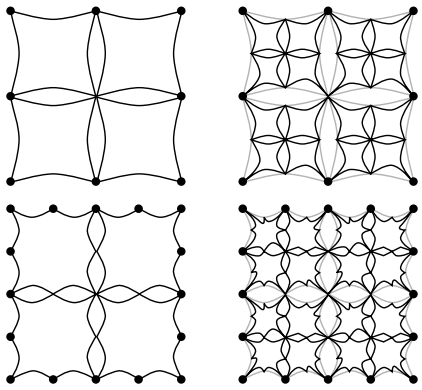

Figura V.5 – Une suite de limites de carrés discrets qui s'approche (?) d'un carré usuel

Indice

Introduction **ii**
1 Avant-propos et plan de la thèse . . . ii
2 Présentation des travaux . . . ii
2.1 Les mesures de Malliavin-Shavgulidze . . . ii
2.2 Distributions invariantes pour les difféomorphismes du cercle . . . v
2.3 Gravité quantique et graphes hiérarchiques . . . vii

Préliminaires **2**
Notations . . . 2
1 La mesure de Wiener . . . 3
1.1 Rappels de probabilité . . . 3
1.2 Mouvement brownien et régularité . . . 4
1.3 Martingales et intégration stochastique . . . 6
1.4 L'espace de Wiener et le théorème de Cameron-Martin . . . 6
2 Mesures quasi-invariantes et groupes localement compacts . . . 9
2.1 Actions sur espaces mesurés . . . 9
2.2 Actions continues sur espaces topologiques mesurés . . . 9
2.3 Mesures quasi-invariantes sur les groupes topologiques . . . 10
3 Difféomorphismes du cercle . . . 11
3.1 Poincaré et le nombre de rotation . . . 12
3.2 Questions de régularité : de Denjoy à nos jours . . . 13
3.2.1 Espaces de difféomorphismes . . . 13
3.2.2 Arithmétique du nombre de rotation . . . 14
3.2.3 Le théorème de conjugaison différentiable . . . 15

I Les mesures de Malliavin-Shavgulidze **17**
1 Un panorama sur les mesures de Malliavin-Shavgulidze . . . 17
1.1 Petite histoire . . . 17
1.2 Définition des mesures . . . 18
1.2.1 Les mesures sur le groupe des difféomorphismes de l'intervalle . . . 18
1.2.2 Les mesures sur le groupe des difféomorphismes du cercle . . . 20
1.3 Modules de continuité . . . 21
2 Quasi-invariance des mesures de Malliavin-Shavgulidze . . . 22
2.1 Le théorème de Cameron-Martin-Shavgulidze . . . 22
2.2 Ergodicité . . . 24
2.3 Preuve de la quasi-invariance de la mesure . . . 25
3 Cohomologie essentielle de groupes . . . 29
3.1 Définition . . . 30
3.2 Cohomologie essentielle d'une action . . . 30
3.3 Cohomologie L^p (réduite) . . . 30
3.4 Notre exemple : les difféomorphismes du cercle . . . 31
3.4.1 Le cocycle de Godbillon-Vey . . . 31
3.4.2 Quelques comparaisons avec les résultats classiques . . . 32

II Décrire la dynamique aléatoire **34**
1 Points périodiques et nombre de rotation aléatoire . . . 34
1.1 Difféomorphismes de l'intervalle . . . 34
1.1.1 Cadre et notations . . . 34
1.1.2 Quelques lemmes préliminaires . . . 35
1.1.3 Preuve du théorème 1.1 . . . 36
1.1.4 Sur le nombre de points fixes . . . 38
1.2 Difféomorphismes du cercle . . . 38
1.2.1 Nombre de rotation rationnel . . . 38
1.2.2 Nombre de rotation irrationnel . . . 40
1.2.3 Renormalisation et mesures . . . 41
2 Centralisateurs . . . 43
2.1 Quelques rappels sur les centralisateurs des difféomorphismes de l'intervalle . . . 43
2.2 Centralisateur C^1 d'un difféomorphisme hyperbolique du cercle . . . 46
3 Exemples de Denjoy . . . 47
3.1 Rappels sur les exemples de Denjoy . . . 48
4 Appendice : Simulations numériques . . . 50
5 Appendice : Sur la dimension des ensembles de Cantor invariants par des groupes abéliens de difféomorphismes du cercle . . . 57

III Équation cohomologique et distributions invariantes d'ordre 1 **59**
1 Mesures invariantes et cobords . . . 59
2 Distributions invariantes par les difféomorphismes avec nombre de rotation irrationnel . . . 60
2.1 Unicité des distributions d'ordre 1 pour les difféomorphismes C^{1+vb} minimaux . . . 61
2.2 Appendice : Le théorème de Douady et Yoccoz pour $s = 1$ . . . 63
2.3 Quelque remarque sur les suites qui approximent les cocycles . . . 64
2.4 Exemples de difféomorphismes C^1 avec plusieurs distributions invariantes . . . 67

IV Distances aléatoires stationnaires sur le graphes hiérarchiques **69**
1 Introduction . . . 69
2 Existence of a stationary law . . . 72
2.1 Preliminaries, notations and definitions . . . 72
2.2 Cut-off process . . . 73
2.3 Openness of the supercritical set Λ . . . 82
3 Convergence in law and uniqueness of the stationary measure . . . 83
3.1 The class $C_{\alpha,\delta}$ and contraction . . . 83
3.2 Asymptotically upper bounds and convergence . . . 87
4 Random metric on the limit object . . . 89
5 Hierarchical graphs . . . 91
5.1 Notations . . . 91
5.2 (min, +)-type recursive distributional equations . . . 92
5.3 Bond percolation on hierarchical graphs and stationary random distances . . . 92
5.4 Appendix: The theorem of Moore and Shannon . . . 93

Références **95**

Préliminaires

Notations

Nombres :

$\mathbf{N}$	nombres naturels, 0 inclus
$\mathbf{N}^*$	nombres naturels, 0 exclu
$\mathbf{Z}$	nombres entiers relatifs
$\mathbf{Q}$	nombres rationnels
$\mathbf{R}$	nombres réels
$\mathbf{R}_+$	nombres réels strictement positifs
$\mathbf{C}$	nombres complexes
$\lfloor x \rfloor$	partie entière de $x \in \mathbf{R}$
$[a_0, a_1, a_2, \ldots]$	développement en fraction continue
$\frac{p_n}{q_n}$	approximation rationnelle de $\alpha \in \mathbf{R} - \mathbf{Q}$

Fonctions :

id	identité
$\mathbf{1}_A$	fonction indicatrice de l'ensemble A
δ_{ij}	symbole de Kronecker
I_n	matrice identité $n \times n$
f^n	itération n-ième de f
$S_n h$	n-ième somme de Birkhoff : $S_n h := \sum_{k=0}^{n-1} h \circ f^k$
R_α	rotation d'angle α
ρ	nombre de rotation
L_g	action régulière gauche
R_g	action régulière droite

Dérivées :

Df, f'	dérivée de f
$D^k f$, $f^{(k)}$	dérivée k-ième de f
$\mathcal{S}_f, \mathcal{S}f$	dérivée schwarzienne de f : $\mathcal{S}f := D^2 \log Df - \frac{1}{2}(D \log Df)^2 = \frac{f'''}{f'} - \frac{3}{2}\left(\frac{f''}{f'}\right)^2$

Espaces :

I	intervalle fermé $[0, 1]$ de la droite réelle
$\mathbf{S}^1$, $\mathbf{R}/\mathbf{Z}$	cercle
$C^r(M)$	fonctions C^r sur M ($r \in \mathbf{R} \cup \{\infty\}$)
$C^\omega(M)$	fonctions analytiques sur M
$C^w(X)$	fonctions continues sur X avec module de continuité w
$C^{vb}(X)$	fonctions à variation bornée sur X
$\text{Homéo}_+(\mathbf{S}^1)$	homéomorphismes de $\mathbf{S}^1$ homotopes à l'identité
$\text{Diff}^r_+(M)$	difféomorphismes C^r qui préservent l'orientation ($r \in \mathbf{R} \cup \{\infty, \omega\}$)
$\text{Diff}^{k+w}_+(M)$	difféomorphismes C^k qui préservent l'orientation, $D^k f \in C^w(M)$ ($k \in \mathbf{N}$)
$\text{Diff}^{k+vb}_+(M)$	difféomorphismes C^k qui préservent l'orientation, $D^k f \in C^{vb}(M)$ ($k \in \mathbf{N}$)
$\text{Diff}^r_{+,\alpha}(\mathbf{S}^1)$	difféomorphismes C^r du cercle avec nombre de rotation α.
$C_0(X)$	fonctions continues sur X qui valent 0 en $0 \in X$ ($X = I, \mathbf{S}^1$)
$L^p(X, \mu)$	fonctions L^p sur (X, μ) ($p \in [1, \infty]$)
$L^q(X, \mu)$	q est le conjugué de p : $\frac{1}{p} + \frac{1}{q} = 1$
$(\Omega, \mathcal{B}, \mathbf{P})$	espace de probabilité

Normes :

$\|\cdot\|_r$	norme C^r
$\|\cdot\|_w$	norme C^w
$\mathrm{Var}(f)$	variation de f : $\mathrm{Var}(f) = \sup_\Pi \sum_{\alpha=1}^{\#\Pi} \sup_{x_\alpha, y_\alpha \in I_\alpha} \lvert f(x_\alpha) - f(y_\alpha)\rvert < \infty$ où Π est une partition de X en un nombre fini d'intervalles disjoints I_α

Mesures :

$\mathbf{2}^X$	parties de X
μ	mesure
$\mu \otimes \nu$	mesure produit
$\mathbf{P}$	mesure de probabilité
$\mathbf{E}$	espérance
Var	variance
Cov	covariance
$f_*\mu$	mesure image
Leb	mesure de Lebesgue
$\mathbf{W}_\sigma$, $\mathbf{W}$	mesure de Wiener sur $C_0(I)$
$\mathbf{W}_{0,\sigma}$, $\mathbf{W}_0$	mesure de Wiener sur $C_0(\mathbf{S}^1)$
ν_σ, ν_{MS}	mesure de Malliavin-Shavgulidze sur $\mathrm{Diff}^1_+(I)$
μ_σ, μ_{MS}	mesure de Malliavin-Shavgulidze sur $\mathrm{Diff}^1_+(\mathbf{S}^1)$
$\int_X f\,d\mu$, $\mu(f)$	intégrale de f par rapport à μ, μ-moyenne
$\ker \mu$	fonctions de μ-moyenne nulle
$\mathcal{P}(X, f)$	mesures de probabilité sur X, f-invariantes
μ_f	mesure de probabilité f-invariante
$\mathcal{B}(X)$	tribu borélienne de X
$\mathcal{C}$	tribu cylindrique
$(\Omega, \mathcal{B}, \mathbf{P})$	espace de probabilité
$\mathcal{F} = (\mathcal{F}_s)$	filtration

1 La mesure de Wiener

Le *mouvement brownien* représente, dans les mathématiques modernes, l'idée de fonction continue choisie au hasard. Le problème de la définition mathématique de cet objet a nécessité beaucoup d'efforts, achevés principalement grâce aux contributions de N. Wiener [Wie 1].

Nous présentons les outils probabilistes nécessaires à la compréhension de ce travail. Le lecteur pourra consulter le texte [Mör-Per] (voir aussi [Rev-Yor, Kar-Shr]) comme référence plus que complète sur ce sujet.

1.1 Rappels de probabilité

Soit Ω un ensemble. Une famille non-vide $\mathcal{B} \subset \mathbf{2}^\Omega$ de parties de Ω est une *tribu* (ou *σ-algèbre*) sur Ω si $\mathcal{B}$ est fermée par les opérations suivantes :

- si $E \in \mathcal{B}$ alors $\Omega - E \in \mathcal{B}$,
- si $(E_n)_{n\in\mathbf{N}} \subset \mathcal{B}$ alors $\bigcup_{n\in\mathbf{N}} E_n \in \mathcal{B}$.

Un couple $(\Omega, \mathcal{B})$ est dit *espace mesurable*. Une fonction f entre deux espaces mesurables $(\Omega, \mathcal{B})$ et $(\Omega', \mathcal{B}')$ est *mesurable* si $f^{-1}(B') \in \mathcal{B}$ pour tout $B \in \mathcal{B}'$.

Exemple 1.1. La droite réelle $\mathbf{R}$ munie de sa tribu borélienne $\mathcal{B}(\mathbf{R})$ est un espace mesurable. Plus généralement, on peut associer un espace mesurable $(X, \mathcal{B}(X))$ à tout espace topologique X, en considérant la tribu des boréliens (*i. e.* engendrée par les ouverts) sur X.

Un *espace de probabilité* $(\Omega, \mathcal{B}, \mathbf{P})$ est la donnée d'un espace Ω, d'une tribu $\mathcal{B}$ et d'une mesure de probabilité $\mathbf{P}$ sur $(\Omega, \mathcal{B})$. Une *variable aléatoire* X sur $(\Omega, \mathcal{B}, \mathbf{P})$ est une fonction mesurable $X : (\Omega, \mathcal{B}) \longrightarrow (\mathbf{R}, \mathcal{B}(\mathbf{R}))$. Nous noterons $\mathbf{E}[X]$ son espérance, $\mathbf{E}[X] = \int_\Omega X(\omega)\,d\mathbf{P}(\omega)$, et $\mathrm{Var}(X)$ sa variance, $\mathrm{Var}(X) = \mathbf{E}[X^2] - \mathbf{E}[X]^2$. Une variable aléatoire X définit une mesure de probabilité $X_*\mathbf{P}$ sur $(\mathbf{R}, \mathcal{B}(\mathbf{R}))$, que l'on appelle *distribution* de la variable: $X_*\mathbf{P}(A) = \mathbf{P}(X \in A)$.

Exemple 1.2. Une variable X sur $(\Omega, \mathcal{B}, \mathbf{P})$ est dite *gaussienne* si sa distribution est une loi gaussienne : autrement dit, il existe $\sigma > 0$, $m \in \mathbf{R}$ tels que

$$\mathbf{P}(X \in A) = \frac{1}{\sqrt{2\pi\sigma}} \int_A \exp\left(-\frac{(t-m)^2}{2\sigma}\right) dt.$$

Les paramètres m est σ^2 sont, respectivement, l'*espérance* et la *variance* de la variable X.

L'analyse de Fourier caractérise les variables de distribution gaussienne : une variable X est une variable gaussienne d'espérance m et variance σ^2 si et seulement si

$$\mathbf{E}\left[e^{i\xi X}\right] = \exp\left(i\xi m - \frac{\sigma^2}{2}\xi^2\right) \quad \text{pour tout } \xi \in \mathbf{R}. \tag{1}$$

Il s'ensuit qu'une variable gaussienne appartient à l'espace $L^p(\Omega, \mathbf{P})$ pour tout $p \in [1, \infty[$.

Une *filtration* $\mathcal{F}$ (d'une tribu $\mathcal{B}$) est une suite monotone croissante de sous-tribus $(\mathcal{F}_s)_{s\geq 0}$: $\mathcal{F}_s \subset \mathcal{F}_t$ si $s < t$. Une filtration est dite *régulière* si pour tout $s \geq 0$,

$$\bigcap_{t>s} \mathcal{F}_t = \mathcal{F}_s.$$

On dit qu'une filtration sur $(\Omega, \mathcal{B}, \mathbf{P})$ définit un *espace de probabilité filtré*.

Un *processus* $X = (X_s)_{s\geq 0}$ sur un espace de probabilité filtré $(\Omega, \mathcal{B}, \mathbf{P}, \mathcal{F})$ est la donnée, pour chaque $s \geq 0$, d'une variable aléatoire $X_s : (\Omega, \mathcal{F}_s, \mathbf{P}) \longrightarrow (\mathbf{R}, \mathcal{B}(\mathbf{R}))$. On dit aussi que le processus X est *adapté* à la filtration $\mathcal{F}$.

Remarque 1.1. Tout processus X est adapté par rapport à sa filtration *naturelle* $(\sigma(X_s, s \leq t))_{t\geq 0}$, où $\sigma(X_s, s \leq t)$ est la plus petite tribu pour laquelle les variables $(X_s, s \leq t)$ sont mesurables, à savoir engendrée par les pré-images $X_s^{-1}(A)$ des ouverts $A \subset \mathbf{R}$.

Exemple 1.3. Un processus $X = (X_s)$ est dit *gaussien*, si toute combinaison linéaire des X_s est une variable gaussienne par rapport à la loi conjointe. Les processus gaussiens sont déterminés par les espérances et les fonctions de *covariance* $\Gamma(s,t) = \mathrm{Cov}(X_s, X_t) = \mathbf{E}[X_s X_t] - \mathbf{E}[X_s]\,\mathbf{E}[X_t]$.

1.2 Mouvement brownien et régularité

Définition 1.1. *Soit* $(\Omega, \mathcal{B}, \mathbf{P})$ *un espace de probabilité. Un* mouvement brownien (standard) *sur l'intervalle* $I = [0,1]$ *est un processus* $B = (B_t)_{t\in I}$ *défini sur* Ω *qui satisfait les propriétés suivantes :*

1. *pour presque tout* $\omega \in \Omega$, *l'application* $t \in I \longmapsto B_t(\omega)$ *est une fonction continue,*
2. *pour tout choix de* $t_0 = 0 < t_1 < \ldots < t_{n-1} < t_n = 1$, *les variables aléatoires* $B_{t_i} - B_{t_{i-1}}$ *sont deux à deux indépendantes,*
3. *pour tout choix* $s < t$ *de temps dans* I, *la variable* $B_t - B_s$ *suit une loi gaussienne* $\mathcal{N}(0, t-s)$, *à savoir*

$$\mathbf{P}(a < B_t - B_s < b) = \frac{1}{\sqrt{2\pi(t-s)}} \int_a^b \exp\left(-\frac{1}{2(t-s)}x^2\right) dx,$$

4. *pour presque tout* $\omega \in \Omega$, *on a* $B_0(\omega) = 0$.

Tout mouvement brownien B induit une mesure de probabilité $\mathbf{W} := B_*\mathbf{P}$, sur l'espace $C_0(I)$ des fonctions continues sur l'intervalle I qui valent 0 en 0, muni de la tribu cylindrique $\mathcal{C}$: un *cylindre* $C \in \mathcal{C}$ est une partie de la forme

$$C = \{x \in C_0(\mathbf{S}^1) \,|\, x(t_i) - x(t_{i-1}) \in A_i,\ i = 1, \ldots n\}, \quad 0 = t_0 < t_1 < \ldots < t_{n-1} < t_n = 1 \text{ et } A_i \in \mathcal{B}(\mathbf{R}). \tag{2}$$

La mesure $\mathbf{W}$ est dite *mesure de Wiener*. La mesure d'un cylindre est déterminée par les propriétés du mouvement brownien (définition 1.1) : pour C comme à la ligne (2),

$$\mathbf{W}(C) = \frac{1}{\sqrt{(2\pi)^n (t_n - t_{n-1}) \cdots (t_1 - t_0)}} \int_{A_1} \cdots \int_{A_n} \exp\left(-\frac{1}{2}\sum_{i=1}^n \frac{y_i^2}{t_i - t_{i-1}}\right) dy_1 \cdots dy_n. \tag{3}$$

Théorème 1.1 (Wiener)**.** *Les tribus cylindrique* $\mathcal{C}$ *et borélienne* $\mathcal{B}(C_0(I))$ *sur* $C_0(I)$ *coïncident.*

Par conséquent, l'*espace de Wiener (classique)* $(C_0(I), \mathcal{B}(C_0(I)), \mathbf{W})$ représente le choix *intrinsèque* d'espace de probabilité sur lequel définir concrètement le mouvement brownien : pour $x \in C_0(I)$, B_t est la variable aléatoire qui correspond à la valeur de la fonction x au temps t,

$$B_t(x) := x(t).$$

Corollaire 1.2. *Le mouvement brownien existe et il est uniquement déterminé par ses propriétés.*

Nous choisirons comme réalisation des trajectoires browniennes, celle du mouvement brownien donné par la mesure de Wiener.

Il s'avère que les trajectoires du mouvement brownien ne sont pas différentiables mais, cependant, elle sont un peu plus que continues. Plus précisément nous avons les résultats suivants (voir [Mör-Per, chapitres 1 et 10]) :

Théorème 1.3 (Différentiabilité en aucun point)**.** *Presque sûrement, le mouvement brownien n'est différentiable en aucun point* $t \in [0,1]$.

Théorème 1.4 (Variation infinie)**.** *Presque sûrement, la variation (voir* (18) *) du mouvement brownien est infinie.*

Théorème 1.5 (Exposant d'Hölder 1/2)**.**

1. *Presque sûrement, le mouvement brownien est une fonction* α*-höldérienne, pour tout* $\alpha < 1/2$.
2. *Presque sûrement, le mouvement brownien est une fonction qui n'est pas* α*-höldérienne, pour tout* $\alpha > 1/2$.

Remarque 1.2. Pour compléter l'énoncé du théorème 1.5, il est intéressant de savoir qu'il existe presque sûrement des points *lents*, auxquels le mouvement brownien est exactement 1/2-höldérien, mais ils sont rares (au sens de la dimension de Hausdorff).

Théorème 1.6 (Module de continuité de Lévy)**.** *Soit* $w(t) = \sqrt{2t|\log t|}$. *Soit* B *un mouvement brownien standard. On a*

$$\mathbf{P}\left(\overline{\lim_{\varepsilon\to 0}} \sup_{\substack{s<t\leq 1,\\ t-s\leq\varepsilon}} \frac{|B_t - B_s|}{w(\varepsilon)} = 1\right) = 1. \tag{4}$$

L'expression (4) dit en particulier que le mouvement brownien est une fonction avec *module de continuité* w. Pour des rappels sur la notion de module de continuité, voir la partie 1.3 du premier chapitre.

Remarque 1.3. Rappelons que le mouvement brownien est un processus gaussien centré et par conséquent caractérisé par la famille des covariances $\Gamma(s,t)$.

Dans le cas du mouvement brownien, on a $\Gamma(s,t) = \min\{s,t\}$, pour tous s, t.

Le pont brownien

Il est aussi possible de définir le mouvement brownien sur le cercle $\mathbf{S}^1$. Dans ce cas on parle de *pont brownien*. Après avoir identifié le cercle au tore $\mathbf{R}/\mathbf{Z}$, l'espace $C_0(\mathbf{S}^1)$ des fonctions continues sur le cercle qui prennent la valeur 0 en 0 peut se voir comme l'espace des fonctions continues sur $\mathbf{R}$ de période 1 qui valent 0 en 0. Puisque l'intervalle $[0,1]$ est un domaine fondamental pour l'action de translation de $\mathbf{Z}$, on pourra écrire

$$C_0(\mathbf{S}^1) = \{x \in C_0(I) \,:\, x(0) = x(1) = 0\}$$

et on ne fera pas de distinction entre les domaines de définition I ou $\mathbf{S}^1$.

On remarquera que l'application π définie par

$$y \in C_0(I) \longmapsto \{x : t \longmapsto y(t) - t\,y(1)\} \in C_0(\mathbf{S}^1) \tag{5}$$

est une projection (elle est linéaire et vérifie $\pi^2 = \pi$) continue et on peut définir la *mesure de Wiener* sur $C_0(\mathbf{S}^1)$ par

$$\mathbf{W}_0 := \pi_*\mathbf{W}.$$

Le processus canonique $x = (x(t))_{t\in\mathbf{S}^1}$ sur $\left(C_0(\mathbf{S}^1), \mathcal{B}\left(C_0(\mathbf{S}^1)\right), \mathbf{W}_0\right)$, définit un processus gaussien que l'on appelle *pont brownien* et porte comme covariances

$$\Gamma(s,t) = \min\{s,t\} - s\,t. \tag{6}$$

Puisque la projection π modifie y d'une fonction linéaire, les propriétés de régularité presque sûres du pont brownien sont les mêmes que celles du mouvement brownien.

Remarque 1.4. Dans la littérature, le pont brownien est aussi présenté en conditionnant le mouvement brownien à revenir en 0 au temps 1. On retrouve un processus gaussien, dont les fonctions de covariance sont données par (6).

1.3 Martingales et intégration stochastique

Il est souvent pratique, et il le sera en effet pour nous, de pouvoir définir en quelque sorte une intégrale d'une fonction par rapport au mouvement brownien. Si f est une fonction dans $L^2(I)$, ceci se fait sans trop de complications au sens de Stieltjes. Si $f = \mathbf{1}_{[s,t[}$, on pose $\int_0^1 \mathbf{1}_{[s,t[}(u)\, dB_u := B(t) - B(s)$. On étend ensuite la fonctionnelle $\int_0^1 \cdot\, dB_u$ par linéarité à toutes les fonctions qui s'écrivent sous la forme

$$f = \sum_{i=1}^{k} a_i\, \mathbf{1}_{[s_i,t_i[}, \quad a_i \in \mathbf{R}. \tag{7}$$

On voit alors que pour presque tout $\omega \in \Omega$, $\int_0^1 \cdot\, dB_u$ définit une fonctionnelle linéaire continue (bornée) sur une partie dense dans $L^2(I)$, et qu'elle possède donc une extension à $L^2(I)$ entier.

De plus, pour tout f comme dans (7), $\int_0^1 f\, dB$ est une somme de variables aléatoires gaussiennes, donc variable gaussienne elle même. Il est classique qu'une limite de variables gaussiennes est encore gaussienne, donc pour tout $f \in L^2(I)$,

$$\int_0^1 f\, dB \tag{8}$$

est une variable gaussienne centrée, de variance $\int_0^1 f^2$ (ceci peut se voir grâce à la caractérisation (1)). La variable $\int_0^1 f\, dB$ est connue sous le nom d'*intégrale de Wiener*.

Il est cependant parfois nécessaire d'étendre la définition d'*intégrale stochastique* à une classe plus large d'« intégrands » : on veut, par exemple, définir

$$\int_0^1 B_2\, dB_1 \tag{9}$$

quand B_1 et B_2 sont deux mouvements browniens indépendants. Ici on rajoute une difficulté : la fonction que l'on est en train d'intégrer est elle même un processus aléatoire. Dans ce cas, la variable (9) prend le nom d'*intégrale d'Itō*. Nous ne rentrerons pas dans les détails, et nous renvoyons le lecteur aux références déjà citées en début de partie.

Il nous suffira de savoir que si l'on veut bien définir la variable aléatoire

$$\int_0^1 A_s\, dM_s,$$

certaines hypothèses s'imposent sur les processus A et M : M doit bien se comporter par rapport à l'évolution du temps et A doit être compatible avec l'évolution de M. En termes techniques, il faut que M soit une *semi-martingale* continue et A un processus intégrable, *adapté* à la filtration naturelle définie par M.

Soit $\mathcal{F} = (\mathcal{F}_s)_{s\in I}$ une filtration. Rappelons qu'un processus $M = (M_s)_{s\in I}$, avec $M_s \in L^1(\Omega, \mathbf{P})$ pour tout $s \in I$, est une $\mathcal{F}$-martingale s'il vérifie, pour tout $s < t$,

$$\mathbf{E}\left[M_t \,|\mathcal{F}_s\right] = M_s, \tag{10}$$

où l'expression de gauche dans (10) dénote l'*espérance conditionnelle* par rapport à la tribu $\mathcal{F}_s$; en d'autres termes, lorsque $M_s \in L^2(\Omega, \mathbf{P})$ pour tout $s \in I$, elle désigne la projection orthogonale de M_s sur le sous-espace de $L^2(\Omega, \mathbf{P})$, engendré par les fonctions caractéristiques $\mathbf{1}_A$, avec $A \in \mathcal{F}_s$.

Exemple 1.4. Soit $\mathcal{F}$ la filtration naturelle associée à un mouvement brownien standard B. Alors B est une $\mathcal{F}$-martingale.

Exemple 1.5. Pour tout $\sigma > 0$, le processus $\left(\exp\left(\sigma B_s - \frac{\sigma^2}{2} s\right)\right)_{s\in I}$ est une $\mathcal{F}$-martingale ($\mathcal{F}$ est la filtration associée à B).

1.4 L'espace de Wiener et le théorème de Cameron-Martin

En suivant [Mör-Per], nous définissons l'*espace de Dirichlet* $\mathbf{D}_0(I)$ des fonctions absolument continues $x \in C_0(I)$ avec dérivée dans $L^2(I)$: il existe $\xi \in L^2(I)$ tel que pour presque tout t, $x(t) = \int_0^t \xi(s)\, ds$.

Cet espace s'insère naturellement dans le contexte de la mesure de Wiener, puisqu'il représente exactement le sous-espace de $C_0(I)$ des fonctions x qui préservent la classe de la mesure de Wiener sous l'action de translation $T_x(y) := y + x$. Plus précisément :

Théorème 1.7 (Cameron-Martin). *Soit x une fonction dans $C_0(I)$. Alors la classe de la mesure de Wiener $\mathbf{W}$ est préservée par T_x si et seulement si $x \in \mathbf{D}_0(I)$ et, sous cette hypothèse, le cocycle de Radon-Nikodym*[3] *est*

$$\frac{d(T_x)_*\mathbf{W}}{d\mathbf{W}}(y) = \exp\left\{-\frac{1}{2}\int_0^1 x'(t)^2\,dt + \int_0^1 x'(s)\,dy(s)\right\}. \tag{11}$$

L'expression $\int_0^1 x'(s)\,dy(s)$ dans (11) dénote l'intégrale de Wiener. Pour montrer ce théorème, nous allons d'abord donner une construction alternative de la mesure de Wiener.

Dimostrazione. Soit $\mathcal{H}$ l'espace de Hilbert réel séparable $L^2(I)$ et $(g_n)_{n\in\mathbf{N}}$ une base orthonormée, par exemple, en suivant Wiener [Wie 2], on pourra prendre $g_0 = 1$ et, pour $n > 0$,

$$g_n(t) = \sqrt{2}\cos(\pi n t). \tag{12}$$

Ce choix, naturel d'un point de vue de l'analyse de Fourier, est cependant moins adapté quand on veut faire des calculs. Un choix bien meilleur, comme McKean l'a observé [McK], est la *base de Haar* : pour $m \geq 1$ et $1 \leq k \leq 2^{m-1}$ on définit

$$h_{m,k} := 2^{\frac{m-1}{2}}\left(\mathbf{1}_{[\frac{2k-2}{2^m},\frac{2k-1}{2^m}]} - \mathbf{1}_{[\frac{2k-1}{2^m},\frac{2k}{2^m}]}\right)$$

et on pose $g_0 = 1$ et $g_n = h_{m,k}$, pour $n = 2^{m-1} - 1 + k$.

L'espace $\mathbf{D}_0(I)$ est un espace de Hilbert réel quand on définit le produit scalaire

$$\langle x, y\rangle_{\mathbf{D}_0(I)} := \langle x', y'\rangle_{L^2}.$$

Une base orthonormée Φ_n est définie par intégration des g_n :

$$\Phi_n(t) := \int_0^t g_n(s)\,ds.$$

Observons que pour toutes fonctions x et y dans L^2, l'inégalité de Cauchy-Schwarz donne

$$\left|\int_0^t (x(s) - y(s))\,ds\right| \leq \|x - y\|_{L^1} \leq \|x - y\|_{L^2},$$

donc si $\lim_{n\to\infty} \|x_n - x\|_{\mathbf{D}_0(I)} = 0$ alors la suite (x_n) tend uniformément vers x. Par conséquent, pour tout $x \in \mathbf{D}_0(I)$, la série

$$x = \sum_{n\in\mathbf{N}} \langle x', g_n\rangle\,\Phi_n$$

converge dans $\mathbf{D}_0(I)$ et donc uniformément.

Soit (Z_n) une suite de variables aléatoires gaussiennes centrées de variance 1, indépendantes, identiquement distribuées. Pour tout $t \in I$, on a

$$\sum_{n\in\mathbf{N}} \Phi_n^2(t) = \sum_{n\in\mathbf{N}} \langle \mathbf{1}_{[0,t]}, g_n\rangle_{L^2}^2 = \|\mathbf{1}_{[0,t]}\|_{L^2}^2 = t, \tag{13}$$

en utilisant l'identité de Parseval.

Pour chaque t fixé, le processus à temps discret $(\sum_{k=1}^n Z_k\,\Phi_k)_{n\in\mathbf{N}}$ est une martingale vérifiant, d'après (13),

$$\sup_{n\in\mathbf{N}} \left\|\sum_{k=1}^n Z_k\,\Phi_k\right\|_{L^2} < \infty,$$

donc pour tout t fixé, la série

$$x(t) = \sum_{n\in\mathbf{N}} Z_n\,\Phi_n(t) \tag{14}$$

converge presque sûrement et dans L^2 : c'est une conséquence du théorème de convergence des martingales (voir [Mör-Per]). Cependant, presque sûrement, la série ne converge pas dans $\mathbf{D}_0(I)$, puisque $\sum_{n\in\mathbf{N}} Z_n^2 = +\infty$ avec probabilité 1. Par contre, une telle convergence est assurée dans $C_0(I)$, une fois la base (Φ_n) fixée.

Lemme 1.8. *Si g_n est la base de Haar, alors, presque sûrement, la série* (14) *converge uniformément dans $C_0(I)$.*

[3] voir la partie 2.1 pour la définition.

Il est assez clair que x défini par (14) est un mouvement brownien : il s'agit d'un processus gaussien, de covariance

$$\Gamma(s,t) := \mathrm{Cov}\,(x(s),x(t)) = \mathbf{E}\left[\sum_{n\in\mathbf{N}} Z_n^2 \Phi_n(t)\,\Phi_n(s)\right] =$$

$$= \sum_{n\in\mathbf{N}} \langle \mathbf{1}_{[0,t]}, g_n\rangle_{L^2}^2\, \langle \mathbf{1}_{[0,s]}, g_n\rangle_{L^2}^2 = \langle \mathbf{1}_{[0,s]}, \mathbf{1}_{[0,t]}\rangle_{L^2}^2 = \min\{s,t\}\,.$$

Revenons maintenant sur le théorème de Cameron-Martin 1.7 . Soit $x \in \mathbf{D}_0(I)$, on va montrer que $\mathbf{W}$ est absolument continue par rapport à $(T_x)_*\mathbf{W}$. On se fixe la base (Φ_n) donnée par les primitives des g_n. Soit (a_n) une suite dans $\ell^2(\mathbf{N})$, on définit

$$x := \sum_{n\in\mathbf{N}} a_n\,\Phi_n,$$

alors, si $y = \sum_{n\in\mathbf{N}} Z_n\,\Phi_n$ est un mouvement brownien,

$$y + x = \sum_{n\in\mathbf{N}} (a_n + Z_n)\,\Phi_n,$$

donc, il nous suffit de montrer que la loi des vecteurs gaussiens (Z_n) est absolument continue par rapport à la loi des $(Z_n + a_n)$.

Pour calculer la dérivée de Radon-Nikodym de $(T_x)_*\mathbf{W}$ par rapport à $\mathbf{W}$ on procède par approximation de la mesure. Soit $N \in \mathbf{N}$ assez grand, on pose

$$x_N := \sum_{n\le N} a_n\,\Phi_n,$$

et on note $\gamma_N = \gamma^{\otimes(N+1)}$ la mesure gaussienne produit sur $\mathbf{R}^{N+1}$. Un simple changement de variable montre que la dérivée de Radon-Nikodym de $(T_{x_N})_*\gamma_N$ par rapport à γ_N est

$$\frac{d(T_{x_N})_*\gamma_N}{d\gamma_N}(z_0,\dots,z_N) = \prod_{j=0}^{N} \frac{\exp\left(-(z_j-a_j)^2/2\right)}{\exp\left(-z_j^2/2\right)} = \exp\left(\sum_{n=0}^{N} a_n\,z_n - \frac{1}{2}\sum_{n=0}^{N} a_n^2\right).$$

Puisque l'on a fait l'hypothèse que la suite $(a_n)_{n\in\mathbf{N}}$ appartient à l'espace $\ell^2(\mathbf{N})$, la limite sur N de la suite $\frac{d(T_{x_N})_*\gamma_N}{d\gamma_N}$ existe presque sûrement et est égale en $y = \sum_{n\in\mathbf{N}} Z_n\,\Phi_n$ à $\exp\left(\langle x', y'\rangle_{L^2} - \frac{1}{2}\|x'\|_{L^2}^2\right)$, où l'on doit interpréter $\langle x', y'\rangle_{L^2} = \langle x, y\rangle_{\mathbf{D}_0(I)}$ comme l'intégrale de Wiener $\int_0^1 x(s)\,dy(s)$ définie à la ligne (8). On en conclut finalement (cf. avec (11) dans l'énoncé)

$$\frac{d(T_x)_*\mathbf{W}}{d\mathbf{W}}(y) = \exp\left(\langle x, y\rangle_{\mathbf{D}_0(I)} - \frac{1}{2}\,\|x\|_{\mathbf{D}_0(I)}^2\right).$$

Nous avons aussi la relation symétrique $(T_x)_*\mathbf{W} \le \mathbf{W}$: il suffit de remarquer que $(T_x)_*\left((T_{-x})_*\mathbf{W}\right)$ est absolument continue par rapport à $(T_{-x})_*\mathbf{W}$. □

Remarque 1.5. Cette construction du mouvement brownien marche aussi, sans surprise, pour définir le pont brownien : il suffit juste de modifier les fonctions Φ_n en posant (confronter avec (5))

$$\tilde{\Phi}_n(t) := \Phi_n(t) - t\,\Phi_n(1).$$

On remarquera que $\tilde{g}_0 \equiv 0$, et on doit donc retirer $\tilde{\Phi}_0$ de la base. Observons que

$$\sum_{n\in\mathbf{N}^*} \tilde{\Phi}_n^2(t) = \sum_{n\in\mathbf{N}^*} \Phi_n^2(t) + t^2 \sum_{n\in\mathbf{N}} \Phi_n^2(1) - 2t \sum_{n\in\mathbf{N}^*} \Phi_n(t)\,\Phi_n(1) = t - t^2.$$

En posant $\tilde{x}(t) = \sum_{n\in\mathbf{N}^*} Z_n\,\tilde{\Phi}_n(t)$, on alors

$$\tilde{\Gamma}(s,t) := \mathrm{Cov}\,(\tilde{x}(s),\tilde{x}(t)) = \mathbf{E}\sum_{n\in\mathbf{N}^*} Z_n^2\,\tilde{\Phi}_n(t)\,\tilde{\Phi}_n(s) = \min\{s,t\} - s\,t.$$

Le théorème de Cameron-Martin pour la mesure de Wiener sur $C_0(\mathbf{S}^1)$ s'énonce ainsi :

Théorème 1.9. *Soit x une fonction dans $C_0(\mathbf{S}^1)$. Alors la classe de la mesure de Wiener $\mathbf{W}_0$ est préservée par T_x si et seulement si $x \in \mathbf{D}_0(\mathbf{S}^1)$ et, sous cette hypothèse, le cocycle de Radon-Nikodym est*

$$\frac{d(T_x)_*\mathbf{W}_0}{d\mathbf{W}_0}(y) = \exp\left(\langle x, y\rangle_{\mathbf{D}_0(\mathbf{S}^1)} - \frac{1}{2}\,\|x\|_{\mathbf{D}_0(\mathbf{S}^1)}^2\right).$$

2 Mesures quasi-invariantes et groupes localement compacts

Le théorème de Cameron-Martin 1.7 rapproche la nature de l'espace de Wiener (vu comme groupe abélien) à celle d'un groupe localement compact. Cette vision nous guidera dans la suite, et il nous semble donc approprié de faire des rappels, dans cette introduction, sur les interactions entre groupes, mesures et topologie. Nous conseillons, à ce propos, le très beau livre de Halmos, [Hal], qui porte très bien ses 60 ans !

2.1 Actions sur espaces mesurés

Ne voulant pas rentrer dans toutes les subtilités de la théorie de la mesure, nous supposerons, dorénavant, que tout espace mesuré X est un espace mesurable *standard* avec une mesure *σ-finie*. Cela signifie que l'espace $(X, \mathcal{B}, \mu)$ est mesurablement isomorphe à l'espace $([0,1], \mathcal{B}([0,1]), \mathrm{Leb})$. Cette restriction est plutôt fictive pour nous, car dans les exemples que nous traiterons, ces hypothèses seront toujours vérifiées.

Définition 2.1. *Soit G un groupe qui agit (à gauche) sur un espace mesuré $(X, \mathcal{B}, \mu)$, avec $\mathcal{B}$ invariante par l'action de G et μ mesure σ-finie non-identiquement nulle. On dit que la mesure μ est* quasi-invariante *pour l'action si pour tout $g \in G$, l'image gA d'un ensemble $A \in \mathcal{B}$ est de mesure nulle si et seulement si la mesure de A est nulle.*

Lorsque $\mu(gA) = \mu(A)$ pour tout $A \in \mathcal{B}$ et tout $g \in G$, la mesure μ est invariante.

Si G agit sur $(X, \mathcal{B}, \mu)$ et μ est quasi-invariante, il existe une unique fonction

$$\rho : G \longrightarrow L^1(X, \mu)$$

telle que pour tout $g \in G$, $\int_X f(x)\, d\,(g_*\mu)\,(x) = \int_X f(g^{-1}x)\, \rho_g(x)\, d\mu(x)$ pour tout $f \in L^1(X, g_*\mu)$. Ce résultat est le bien connu *théorème de Radon-Nikodym* et la fonction ρ_g est plus souvent notée $\frac{dg_*\mu}{d\mu}$; une telle notation est surtout appropriée si l'on veut souligner sa nature de *cocycle* (de Radon-Nikodym) : pour tous $g, h \in G$,

$$\frac{d(gh)_*\mu}{d\mu}(x) = \frac{dg_*\mu}{d\mu}(hx) \cdot \frac{dh_*\mu}{d\mu}(x)$$

en dehors d'un ensemble μ-négligeable de X.

Remarque 2.1. Si μ et ν sont deux mesures sur $(X, \mathcal{B})$ avec la propriété que A est μ-négligeable si et seulement si A est ν-négligeable, on dit que μ et ν définissent la même *classe* de mesure $[\mu] = [\nu]$, ou, également, que μ est *équivalente* à ν, et on écrit $\mu \sim \nu$.

Lorsqu'une seule des implications est vérifiée, par exemple si A est ν-négligeable seulement lorsque A est μ-négligeable, on dit que ν est *absolument continue* par rapport à μ.

Une mesure μ est donc quasi-invariante sous l'action de G si et seulement si les éléments de G préservent la classe de μ.

2.2 Actions continues sur espaces topologiques mesurés

Lorsque l'on suppose que le groupe G agit de façon continue sur X, il est préférable (et plus intéressant) de considérer des mesures sur X qui s'adaptent à la topologie. Dans la suite, on supposera toujours que les espaces topologiques en question sont séparables.

Définition 2.2. *Soit $(X, \tau, \mathcal{B})$ un* espace topologique mesurable *(ici $\mathcal{B}$ est la tribu borélienne*[4]*). Une mesure μ sur $(X, \tau, \mathcal{B})$ est une* mesure de Borel *si μ est non-identiquement nulle et finie sur les parties compactes de X.*[5]

Rappelons qu'une mesure de Borel est *régulière à l'extérieur* si pour tout $E \in \mathcal{B}$,

$$\mu(E) = \inf\{\mu(U) \,:\, U \text{ ouvert qui contient } E\}.$$

De manière similaire, μ est dite *régulière à l'intérieur* si pour tout $E \in \mathcal{B}$,

$$\mu(E) = \sup\{\mu(K) \,:\, K \text{ compact contenu dans } E\}.$$

Une mesure régulière dans le sens précédent est essentiellement déterminée par ses valeurs sur les ouverts ou les compacts de X et une telle mesure est donc appropriée dans un contexte topologique. Ces deux notions sont presque équivalentes, mais il s'avère que demander la régularité *intérieure* est préférable.

[4] Normalement on demande que $\mathcal{B}$ *contient* la tribu borélienne.

[5] Dans la littérature, la définition de mesure de Borel peut changer selon les différents auteurs : la propriété qu'elle soit finie sur les compacts n'est pas toujours demandée.

Définition 2.3. *Soit $(X, \tau, \mathcal{B})$ un espace topologique mesurable. Une mesure de Borel μ est une* mesure de Radon *si elle est finie sur les compacts et régulière à l'intérieur (par rapport aux compacts).*

Les mesures que nous utiliserons seront de Radon.

Exemple 2.1. La mesure de Wiener (voir la partie 1) est une mesure de Radon et donne une mesure positive à tout ouvert : d'après le théorème de Wiener 1.1 il suffit de le vérifier sur les cylindres.

Lorsque G agit de manière continue sur X, on peut se demander s'il existe une mesure de Radon (quasi-)invariante. Sans hypothèses supplémentaires sur X, la réponse à cette question est, en général, négative. En contrepartie, dès que l'on suppose que X est *localement compact*, on trouve des résultats surprenants !

Théorème 2.1. *Soit X un espace topologique séparable localement compact et soit G un groupe qui agit sur X de façon continue. On fait les deux hypothèses suivantes :*

1. *toute orbite est dense,*
2. *lorsque K et L sont deux compacts disjoints dans X, il existe un ouvert U non vide dans X tel que, pour tout $g \in G$, le translaté gU n'intersecte pas K et L en même temps.*

Alors il existe une mesure de Radon μ sur X, invariante *par l'action de G.*

Ce théorème est dû à Steinlage [Ste] et nous l'avons appris sous cette forme dans l'œuvre (monumentale) de D. H. Fremlin [Fre, 441C]. En fait, cet énoncé généralise le très célèbre théorème d'existence de la *mesure de Haar* pour les groupes localement compacts, dont nous allons discuter dans la prochaine partie.

2.3 Mesures quasi-invariantes sur les groupes topologiques

Les actions qui nous intéresserons le plus seront celles des groupes topologiques sur eux-mêmes. Tout groupe agit naturellement sur lui-même par translations, de deux cotés différents : il peut agir *à gauche*

$$\begin{aligned} L : G \times G &\longrightarrow G \\ (g, x) &\longmapsto g^{-1}x \end{aligned}$$

et *à droite*

$$\begin{aligned} R : G \times G &\longrightarrow G \\ (g, x) &\longmapsto xg \end{aligned}$$

et ces actions sont dites *régulières*.

Remarque 2.2. Les actions L et R commutent toujours. On remarquera aussi que les actions sont *libres* et *transitives* : si $g^{-1}x = h^{-1}x$ alors $g = h$, et si $x, y \in G$ alors, en posant $g = xy^{-1}$, on a $g^{-1}x = y$.

Une autre action dynamiquement intéressante est l'*action par conjugaison* :

$$\begin{aligned} \gamma : G \times G &\longrightarrow G \\ (g, x) &\longmapsto gxg^{-1} \end{aligned}$$

Remarque 2.3. Contrairement aux actions régulières, l'action par conjugaison n'est ni libre, ni transitive (si le groupe n'est pas trivial !). Les orbites forment les *classes de conjugaison* dans G et le stabilisateur d'un élément $x \in G$ coïncide avec le *centralisateur* de x dans G.

Lorsque le groupe G est un groupe topologique, ces actions sont continues.

Définition 2.4. *Une* mesure de Haar *invariante à gauche (resp. à droite) sur G est une mesure de Radon sur G, invariante par l'action régulière gauche (resp. droite).*

Exemple 2.2. Soit Γ un groupe dénombrable discret. La mesure de comptage définit une mesure de Haar sur Γ qui est invariante à gauche et à droite en même temps. Il est facile d'observer que toute autre mesure de Haar sur Γ doit être proportionnelle à la mesure de comptage. En outre, la mesure de Haar est finie si et seulement si le groupe Γ est fini.

Cet exemple, même un peu naïf, représente pleinement la situation pour les groupes *localement compacts*. Nous avons déjà annoncé que le théorème de Steinlage 2.1 est une version « décorée » du théorème de Haar que nous énonçons ici.

Théorème 2.2. *Sur un groupe localement compact G, il existe toujours une mesure de Haar invariante à gauche, et deux telles mesures sont proportionnelles.*

De plus, la mesure est finie si et seulement si le groupe G est compact.

Remarque 2.4. Toute mesure de Haar μ invariante à gauche définit une mesure de Haar ν invariante à droite : si $E \in \mathcal{B}$ et un borélien dans G, on pose $\nu(E) = \mu(E^{-1})$. Alors pour tout $g \in G$,

$$\nu(Eg) = \mu(g^{-1}E^{-1}) = \mu(E^{-1}) = \nu(E).$$

En général, les mesures de Haar invariantes à gauche ne sont pas invariantes à droite.

De manière encore plus étonnante et remarquable, l'existence d'une mesure de Haar caractérise essentiellement les groupes localement compacts.

Définition 2.5. *Un groupe* mesurable (au sens de Weil) *est un espace mesuré* $(G, \mathcal{B}, \mu)$ *tel que*

1. *μ est σ-finie est non-identiquement nulle,*
2. *G est un groupe,*
3. *la tribu $\mathcal{B}$ et la mesure μ sont invariantes par translations à gauche,*
4. *la transformation de « déchirement » $S : G \times G \longrightarrow G \times G$ définie par $S(x, y) = (x, xy)$ est mesurable, et d'inverse mesurable.*

Remarque 2.5. La dernière condition est équivalente à demander que la fonction qui à $(x, y) \in G \times G$ associe $x^{-1}y \in G$ soit mesurable.

Définition 2.6. *Un groupe mesuré* $(G, \mathcal{B}, \mu)$ *est dit* séparé (au sens de Weil) *si pour tout élément $g \in G$ différent de l'identité, il existe $E \in \mathcal{B}$ de mesure finie strictement positive tel que*

$$\mu(gE \triangle E) > 0. \tag{15}$$

Remarque 2.6. De manière équivalente, on pourra remplacer la condition (15) par $\mu(gE \cap E) < \mu(E)$.

Lorsque $(G, \mathcal{B}, \mu)$ est un groupe mesuré et séparé, il est possible de définir une topologie de groupe topologique sur G, la *topologie de Weil*. En effet, sous ces hypothèses, la famille

$$\mathcal{N} := \{\{g \in G \,:\, \mu(gE \triangle E) < \varepsilon\} \text{ t. q. } E \in \mathcal{B} \text{ est de mesure finie strictement positive, } \varepsilon \in]0, 2\,\mu(E)[\}$$

définit un système de voisinages de l'identité dans G.

Le groupe topologique mesuré $(G, \tau_{\text{Weil}}, \mathcal{B}, \mu)$ est un groupe localement borné. Lorsque l'intérieur de $E \in \mathcal{B}$ n'est pas vide, $\mu(E) > 0$; lorsque $E \in \mathcal{B}$ est borné, $\mu(E) < \infty$. Donc G est « presque » un groupe localement compact et μ sa mesure de Haar. Le « presque » est précisé par le théorème de Mackey et Weil [Mac, Wei] que nous allons énoncer.

Définition 2.7. *Un sous-groupe G dans un groupe mesuré* $(\hat{G}, \hat{\mathcal{B}}, \hat{\mu})$ *est dit* épais *si pour tout $\hat{E} \in \hat{\mathcal{B}}$,*

$$\sup\left\{\hat{\mu}(\hat{K}) \,:\, \hat{K} \textit{ est un compact dans } \hat{G} \textit{ contenu dans } \hat{E} - G\right\} = 0.$$

Théorème 2.3 (Mackey-Weil)**.** *Soit G un groupe mesuré, séparé. Alors il existe un groupe localement compact $\hat{G}$ avec une mesure de Haar $\hat{\mu}$ sur la tribu de Borel $\hat{\mathcal{B}}$ de $\hat{G}$, tel que G soit un sous-groupe épais de $\hat{G}$, $\mathcal{B} \supset \hat{\mathcal{B}} \cap G$ et $\mu(E) = \hat{\mu}(\hat{E})$ lorsque $\hat{E} \in \hat{\mathcal{B}}$ et $E = \hat{E} \cap G$.*

Pour une preuve de ce théorème, le lecteur pourra consulter [Wei], aussi bien que le déjà cité [Hal].

Remarque 2.7. Dans [Wei], une version plus fine de ce théorème est montrée. En effet, on peut affaiblir les hypothèses dans 2.3 pour montrer que le même résultat est valable lorsque l'on suppose seulement que la mesure sur $(G, \mathcal{B})$ est *quasi-invariante*.

Le phénomène que l'on observe avec le théorème de Cameron-Martin 1.7 est par conséquent presque optimal : puisque $C_0(I)$ est un groupe qui n'est pas localement compact, il n'y a aucun espoir de définir une mesure quasi-invariante, non seulement sur $C_0(I)$, mais sur un sous-espace quelconque de $C_0(I)$ de mesure strictement positive. Cependant, la mesure de Wiener est quasi-invariante par l'action d'un sous-groupe dense, qui n'est pas localement compact, mais en même temps de mesure nulle !

3 Difféomorphismes du cercle

Il ne sera pas possible de résumer en quelques pages la théorie des difféomorphismes du cercle : leur étude, inaugurée par Poincaré dans son célèbre article [Poi], constitue une partie remarquable dans la théorie des systèmes dynamiques, et des mathématiques en général. Nous renvoyons le lecteur aux textes généraux [Kat-Has, Chapitres 11 et 12], [dMe-vSt, Chapitre I], et plus spécifiques [Yoc 2, Her 1] – pour les aspects analytiques – et [Ghy 2, Nav 1] – pour une approche plus « géométrique ».

Nous nous limiterons ici à introduire les notions et résultats fondamentaux qui seront utilisés dans la suite de ce travail.

3.1 Poincaré et le nombre de rotation

Soit f un *homéomorphisme* du cercle $\mathbf{S}^1$, que l'on voit comme le quotient de $\mathbf{R}$ par la translation $x \longmapsto x+1$. Il existe alors un homéomorphisme F de $\mathbf{R}$ dans $\mathbf{R}$ qui vérifie $F(x+1) = F(x)+1$ pour tout $x \in \mathbf{R}$ et $F(x) = f(x) \pmod 1$. On dira que F *relève* l'homéomorphisme f.[6]

De manière réciproque, tout homéomorphisme $F : \mathbf{R} \longrightarrow \mathbf{R}$ qui commute avec la translation par 1, définit par réduction $\pmod 1$ un homéomorphisme du cercle $\bar{F}$.

En termes plus savants, le morphisme π de revêtement universel

$$\mathbf{0} \longrightarrow \mathbf{Z} \longrightarrow \mathbf{R} \xrightarrow{\pi} \mathbf{S}^1 \longrightarrow \mathbf{1}$$

induit le morphisme de revêtement universel

$$\mathbf{0} \longrightarrow \mathbf{Z} \longrightarrow \widetilde{\text{Homeo}}_+(\mathbf{S}^1) \xrightarrow{\pi} \text{Homéo}_+(\mathbf{S}^1) \longrightarrow \mathbf{1} \tag{16}$$

et $\widetilde{\text{Homeo}}_+(\mathbf{S}^1)$ s'identifie aux fonctions $\{F \in \text{Homéo}_+(\mathbf{R}) \ : \ F(x+1) = F(x)+1\}$, qui est un groupe topologique lorsqu'on le munit de la topologie induite par la norme C^0 : un homéomorphisme F est ε-proche de l'identité id si

$$\|F - id\|_0 := \sup_{x \in [0,1[} |F(x) - x| < \varepsilon.$$

Parfois nous ne faisons pas de distinction entre un homéomorphisme du cercle et ses relevés. Pour limiter cette ambiguïté, nous supposerons souvent que l'homéomorphisme F qui relève $f \in \text{Homéo}_+(\mathbf{S}^1)$ sera déterminé par la condition $F(0) \in [0,1[$.

Remarque 3.1. Si F_1 et F_2 relèvent f, leur différence $F_2 - F_1$ se projette $\pmod 1$ sur la fonction identiquement nulle. Deux relevés diffèrent ainsi d'un nombre entier.

Proposition 3.1 (Poincaré). *Pour tout $f \in \text{Homéo}_+(\mathbf{S}^1)$ donné et choix du relevé F, il existe un unique $\alpha \in \mathbf{R}$ tel que pour tout $x \in \mathbf{R}$*

$$|F^n(x) - x - n\alpha| < 1. \tag{17}$$

En d'autres termes, il existe une unique translation R_α sur $\mathbf{R}$ telle que la fonction F reste à distance C^0 bornée de R_α, ainsi que toutes ses itérées F^n de $R_{n\alpha}$, avec une borne ne dépendant pas de n. Cette remarque justifie la terminologie : le nombre α donné par la proposition 3.1 est appelé le *nombre de translation* de F.

Il suit de la remarque 3.1 que le nombre de translation d'un homéomorphisme f est bien défini $\pmod 1$: l'application ρ qui à $F \in \widetilde{\text{Homeo}}_+(\mathbf{S}^1)$ associe son nombre de translation $\alpha = \rho(F)$ induit l'application

$$\begin{array}{rcl} \bar{\rho} : \text{Homéo}_+(\mathbf{S}^1) & \longrightarrow & \mathbf{S}^1 \\ f & \longmapsto & \rho(F) \pmod 1 \end{array}$$

qu'on appelle *application de nombre de rotation* et qu'on notera, avec un abus de notation, plus souvent par ρ.

Lorsqu'on munit le groupe $\text{Homéo}_+(\mathbf{S}^1)$ de sa topologie naturelle C^0, l'application ρ est une fonction continue : elle est la limite uniforme de la suite de fonctions continues

$$\frac{1}{n}(F^n(x) - x) \pmod 1.$$

L'application nombre de rotation donne beaucoup d'informations sur la dynamique induite par un homéomorphisme f : son nombre de rotation $\rho(f)$ est rationnel si et seulement si f possède une orbite périodique. Plus précisément, si $\rho(f) = \frac{p}{q}$, alors toute orbite périodique de f est de période q.

Proposition 3.2. *Soient f et g deux homéomorphismes du cercle tels qu'il existe h, fonction sur $\mathbf{S}^1$ continue et croissante*[7] *qui* semi-conjugue *f à g,*

$$h \circ f = g \circ h.$$

Alors, les nombres de rotations de f et g coïncident :

$$\rho(f) = \rho(g).$$

[6] On supposera toujours que f est un homéomorphisme qui préserve l'orientation de $\mathbf{S}^1$.

[7] On dit qu'une fonction h sur le cercle est croissante si les fonctions H qui relèvent h à $\mathbf{R}$ sont croissantes.

Quand l'homéomorphisme f ne possède pas d'orbite périodique, deux cas peuvent se produire. La première possibilité est que f soit un homéomorphisme *minimal*, à savoir que toute orbite est dense : dans ce cas il existe un homéomorphisme h, unique à rotation près, qui *conjugue* f à la rotation d'angle $\alpha = \rho(f)$,

$$h \circ f = h + \alpha.$$

Sinon, les orbites s'accumulent sur un unique ensemble minimal invariant $K = K_f$, qui est homéomorphe à un Cantor. Dans ce cas, il existe une fonction continue et croissante h qui *semi-conjugue* f à la rotation d'angle $\alpha = \rho(f)$,

$$h \circ f = h + \alpha.$$

Ce cas est dit *exceptionnel* et l'ensemble minimal K est également dit exceptionnel. Ce terme dérive du fait que, selon Poincaré, de tels exemples devaient être bien rares... Hélas, Poincaré n'a pas vécu assez longtemps pour que A. Denjoy [Den] lui confirme et contredise à la fois cette affirmation : il construit des exemples, que l'on appelle aujourd'hui *de Denjoy*, d'homéomorphismes et difféomorphisme C^1 qui sont exceptionnels[8]. D'un autre coté, il montre que la vision de Poincaré n'était pas entièrement erronée :

Théorème 3.3 (Denjoy). *Tout difféomorphisme du cercle de classe C^{1+vb} et de nombre de rotation irrationnel est minimal.*

Rappelons qu'une fonction f est dite de classe C^{1+vb} si elle est de classe C^1 et la dérivée f' est une fonction à *variation bornée* :

$$\mathrm{Var}(f') = \sup_{\Pi} \sum_{\alpha=1}^{\#\Pi} \sup_{x_\alpha, y_\alpha \in I_\alpha} |f'(x_\alpha) - f'(y_\alpha)| < \infty, \tag{18}$$

où Π est une partition de $\mathbf{S}^1$ en un nombre fini d'intervalles disjoints I_α, $\alpha = 1, \dots, \#\Pi$.

3.2 Questions de régularité : de Denjoy à nos jours

Le théorème de Denjoy 3.3 signe le début de l'étude systématique des *difféomorphismes* du cercle. Quelle est la régularité maximale de l'homéomorphisme h, donné par le théorème de Denjoy, qui conjugue f à R_α ? Cette question a été au cœur de travaux modernes sur les difféomorphismes du cercle parmi les plus fascinants : les résultats à ce propos de Siegel, Arnol'd, Moser, Herman, Yoccoz, Katznelson et Ornstein, Khanin, Sinai et beaucoup d'autres [Sie, Arn 1, Her 1, Mos, Yoc 1, Katz-Orn, Kha-Sin 1, Sta, Kha-Sin 2, Kha-Tep] ont initié des développements majeurs de la théorie. Nous nous occuperons de ce problème un peu plus loin (voir 3.2.3), après quelques rappels pour fixer notations et définitions.

3.2.1 Espaces de difféomorphismes

Soit $r \in [0, \infty]$ un nombre réel, éventuellement infini, on notera $\mathrm{Diff}^r_+(\mathbf{S}^1)$ le groupe des difféomorphismes du cercle f qui préservent l'orientation et sont C^r réguliers[9]. Quand r est un entier, cela signifie que f est dérivable r fois et que sa dérivée r-ième est continue. Quand r n'est pas un nombre entier, cela signifie que f est dérivable $\lfloor r \rfloor$ fois et que sa dérivée $\lfloor r \rfloor$-ième est une fonction $(r - \lfloor r \rfloor)$-höldérienne.

Par analogie au cas continu, on définit le groupe $\widetilde{\mathrm{Diff}}^r_+(\mathbf{S}^1) = \{F \in \mathrm{Diff}^r_+(\mathbf{R}) \,:\, F(x+1) = F(x) + 1\}$, revêtement universel du groupe $\mathrm{Diff}^r_+(\mathbf{S}^1)$:

$$\mathbf{0} \longrightarrow \mathbf{Z} \longrightarrow \widetilde{\mathrm{Diff}}^r_+(\mathbf{S}^1) \xrightarrow{\pi} \mathrm{Diff}^r_+(\mathbf{S}^1) \longrightarrow \mathbf{1}.$$

Lorsque $r \in [0, \infty]$, on munit l'espace $\mathrm{Diff}^r_+(\mathbf{S}^1)$ de la topologie induite par la norme C^r : si r est un nombre entier ou bien l'infini,

$$\|f\|_r := \max_{s \le r} \sup_{x \in \mathbf{S}^1} |D^s f(x)|$$

et plus généralement, en posant $\tau := r - \lfloor r \rfloor$,

$$\|f\|_r := \|f\|_{\lfloor r \rfloor} + \sup_{x, y \in \mathbf{S}^1} \frac{|D^{\lfloor r \rfloor} f(x) - D^{\lfloor r \rfloor} f(y)|}{|x - y|^\tau}.$$

Remarque 3.2. Soit f un difféomorphisme de classe C^r. Alors f est un difféomorphisme de classe C^s pour tout $s < r$. Notamment, par définition des normes, $\|f\|_s \le 2\,\|f\|_r$: l'inclusion de $\mathrm{Diff}^r_+(\mathbf{S}^1)$ dans $\mathrm{Diff}^s_+(\mathbf{S}^1)$ est continue.

Malheureusement, cette topologie n'est pas toujours compatible avec la structure de groupe.

[8] La communauté hésite entre l'emploi du terme *exemple* ou *contre-exemple*. Peut-être, parler de *Cantor-exemples* surmonterait ce problème...
[9] Par abus de langage, un difféomorphisme de classe C^0 est simplement un homéomorphisme.

Proposition 3.4. *Le groupe* $\mathrm{Diff}^r_+(\mathbf{S}^1)$ *des difféomorphismes* C^r *du cercle, muni de la topologie de la norme* C^r*, est un groupe topologique si est seulement si* r *est un entier ou* $r = \infty$*.*

Remarque 3.3. Dorénavant, lorsque l'on souhaitera voir $\mathrm{Diff}^r_+(\mathbf{S}^1)$ comme un groupe, la topologie sur $\mathrm{Diff}^r_+(\mathbf{S}^1)$ sera toujours la topologie $C^{\lfloor r \rfloor}$.

Nous donnerons une preuve de la proposition 3.4 en appendice à ce chapitre.

Un autre espace, très important, est l'espace des difféomorphismes analytiques. Dans ce contexte, le cercle est identifié à l'ensemble des nombres complexes de module 1. Rappelons qu'une fonction du cercle f est dite *analytique* s'il existe un nombre $\rho > 1$ tel que f possède un prolongement analytique sur la couronne circulaire $A_r := \{z \in \mathbf{C} \,:\, 1/\rho < |z| < \rho\}$. L'espace des difféomorphismes analytiques sur le cercle est noté $\mathrm{Diff}^\omega_+(\mathbf{S}^1)$ et forme un groupe, qui est de plus topologique, lorsque l'on le munit de la topologie qui est la limite inductive des topologies de la convergence uniforme C^0 sur les fonctions analytiques des couronnes A_r.

3.2.2 Arithmétique du nombre de rotation

Soit $\alpha \in \mathbf{R} - \mathbf{Q}$ un nombre irrationnel. On notera par $\left(\frac{p_n}{q_n}\right)_{n\in\mathbf{N}}$ la suite des *approximations rationnelles* de α : pour tout n,

$$0 < \left|\alpha - \frac{p_n}{q_n}\right| < \frac{1}{q_n^2}, \tag{19}$$

et pour tout p et tout $0 < q < q_n$

$$\left|\alpha - \frac{p}{q}\right| > \left|\alpha - \frac{p_n}{q_n}\right|. \tag{20}$$

Si $[a_0, a_1, a_2, \ldots]$ est le développement en fraction continue de α, à savoir

$$\alpha = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \frac{1}{a_3 + \cdots}}}, \tag{21}$$

la fraction $\frac{p_n}{q_n}$ coïncide avec la troncature au niveau n de (21) :

$$\frac{p_n}{q_n} = [a_0, a_1, \ldots, a_n] = a_0 + \cfrac{1}{a_1 + \cfrac{1}{\ddots + \frac{1}{a_n}}}.$$

Les conditions (19) et (20) se traduisent dynamiquement : le point $q_n\alpha \pmod 1$ est le point de l'orbite de 0 par R_α qui se rapproche le mieux de 0.

Rappelons qu'une mesure μ sur $\mathbf{S}^1$ est *invariante* par un homéomorphisme f si elle est invariante par l'action de $\mathbf{Z} \cong \langle f \rangle$: $f_*\mu = \mu$. La théorie des espaces de Banach assure toujours l'existence d'une mesure de probabilité invariante, que l'on peut obtenir comme valeur d'adhérence de la suite $\left(\frac{1}{n}\sum_{k=0}^{n-1} \delta_{f^k(x)}\right)_{n\in\mathbf{N}}$, pour un point $x \in \mathbf{S}^1$ fixé. Toute mesure invariante a pour support une partie compacte invariante de S^1 : une réunion d'orbites périodiques, lorsque f en possède, ou le compact minimal invariant, dans tout autre cas.

Les mesures invariantes donnent beaucoup d'informations sur la dynamique. Par exemple, si l'on pose $h(x) = \mu([x_0, x[)$, pour μ mesure f-invariante et $x_0 \in \mathbf{S}^1$ point arbitraire, on a

$$h \circ f = h + \mu([x_0, f(x_0)[)\,.$$

Autrement dit, la fonction h semi-conjugue f à la rotation d'angle $\mu([x_0, f(x_0)[)$ et, en particulier, le nombre de rotation de f est égal à $\mu([x_0, f(x_0)[)$, pour tout choix de μ et x_0.

Le lien entre mesures invariantes et dynamique est bien plus profond.

Théorème 3.5 (Inégalité de Denjoy-Koksma)**.** *Soit* f *un difféomorphisme du cercle de classe* C^1 *et de nombre de rotation irrationnel* α*. Soit* $q = q_n$ *le dénominateur d'une approximation rationnelle de* α*. Alors, pour toute fonction* $u \in C^{1+vb}(\mathbf{S}^1)$ *et* μ *mesure de probabilité* f*-invariante sur* $\mathbf{S}^1$*, on a l'inégalité*

$$\left\| S_q u - q \int_{\mathbf{S}^1} u \, d\mu \right\|_0 \leq \mathrm{Var}(u). \tag{22}$$

Voir aussi le corollaire 2.3 dans le chapitre III, pour une inégalité plus précise lorsque f et u sont un peu plus réguliers.

Remarque 3.4. L'inégalité (22) permet de voir facilement qu'il y a une unique mesure f-invariante, qu'on note μ_f.

Lemme 3.6 (Denjoy). *Soit f un difféomorphisme du cercle de classe C^1, μ une mesure f-invariante sur $\mathbf{S}^1$. On a*

$$\int_{\mathbf{S}^1} \log Df \, d\mu = 0.$$

Corollaire 3.7 (Inégalité de Denjoy). *Soit f un difféomorphisme du cercle de classe C^{1+vb} et de nombre de rotation irrationnel α. Soit $q = q_n$ le dénominateur d'une approximation rationnelle de α. Alors,*

$$\left\|\log Df^q\right\|_0 \leq \mathrm{Var}(\log Df). \tag{23}$$

Ces inégalités montrent que si l'on connaît la vitesse de divergence de (q_n), ou en d'autres termes la qualité d'approximation de α par des rationnels, on peut envisager de bonnes estimations dans la topologie C^1. Pour cela, on rappellera des notions arithmétiques.

Un nombre irrationnel α vérifie une condition *diophantienne* d'ordre δ s'il existe une constante $C > 0$ telle que pour tout $p/q \in \mathbf{Q}$,

$$0 < \left|\alpha - \frac{p}{q}\right| \geq \frac{C}{q^{2+\delta}}. \tag{24}$$

Lorsque α vérifie une condition diophantienne d'ordre 0, on dit que α est de *type constant*. Un nombre α est de *type Roth* s'il satisfait une condition diophantienne d'ordre ε pour tout $\varepsilon > 0$. Un nombre α qui vérifie une condition diophantienne quelconque est dit *diophantien*. Les nombres irrationnels qui ne sont pas diophantiens sont de *type Liouville* et forment une partie résiduelle (au sens de Baire) dans $\mathbf{R}$. En contrepartie, presque tout nombre (au sens de la mesure de Lebesgue) est de type Roth.

3.2.3 Le théorème de conjugaison différentiable

Après 50 années écoulées depuis les premiers résultats, la compréhension de la régularité du difféomorphisme conjuguant à la rotation est presque optimale. Le résultat le plus avancé s'énonce ainsi [Katz-Orn] (voir aussi [dMe-vSt]) :

Théorème 3.8 (Conjugaison différentiable). *Soit f un difféomorphisme du cercle de classe C^r, soit $\alpha = \rho(f)$ son nombre de rotation. On suppose que α vérifie une condition diophantienne d'ordre δ. Si $\delta + 2 < r$, alors l'homéomorphisme h qui conjugue f à R_α est un difféomorphisme de classe $C^{r-1-\delta-\varepsilon}$ pour tout $\varepsilon > 0$.*

Remarque 3.5. Lorsque $\delta + 2 < r < 4$, Khanin et Teplins'kyi [Kha-Tep] et Teplins'kyi [Tep] ont montré que le difféomorphisme h est plus précisément de classe $C^{r-1-\delta}$.

Essayons de comprendre, de manière un peu naïve, pourquoi une condition diophantienne sur α influence autant la régularité. Résoudre l'équation $h \circ f = h + \alpha$ dans l'espace C^1, revient à s'intéresser à l'*équation cohomologique*

$$u \circ f - u = -\log Df, \tag{25}$$

dans l'espace C^0. Le théorème suivant, dû à Gottschalk et Hedlund [Got-Hed] donne une condition nécessaire et suffisante pour que (25) ait des solutions.

Théorème 3.9 (Gottschalk — Hedlund). *Soit (X, d) un espace métrique compact et f un homéomorphisme minimal de X sur X. Soit φ une fonction continue sur X. Les deux affirmations suivantes sont équivalentes :*

1. *il existe $u \in C^0(X)$ tel que $u \circ f - u = \varphi$,*
2. *il existe un point x_0 dans X tel que $\sup_{n\in\mathbf{N}} \left|\sum_{k=0}^{n-1} f\left(f^k(x_0)\right)\right| < +\infty$.*

Puisque pour tout $n \in \mathbf{N}$, $\log Df^n = \sum_{k=0}^{n-1} \log Df \circ f^k$, une condition nécessaire et suffisante pour qu'il existe une solution continue pour (25) est

$$\sup_{n\in\mathbf{N}} \|\log Df^n\|_0 < +\infty. \tag{26}$$

La fonction qui à $f \in \mathrm{Diff}^1_+(\mathbf{S}^1)$ associe $H_1(f) = \sup_{n\in\mathbf{N}} \|\log Df^n\|_0$ est appelée *invariant de conjugaison* C^1 par Herman. On confrontera avec (23) que si les dénominateurs q_n sont trop écartés l'un de l'autre, l'inégalité de Denjoy n'influence pas trop l'invariant $H_1(f)$. Au contraire, les inégalités du type (23) sont fondamentales pour la preuve du théorème de conjugaison différentiable. Pour expliquer de manière limpide ce dernier commentaire, nous citons les résultats que l'on peut obtenir grâce à la *méthode de conjugaison rapide* due à Anosov et Katok [Kat-Has].

Théorème 3.10. *Pour tout entier $r \geq 1$, il existe un difféomorphisme minimal C^∞ du cercle qui est C^{r-1}-conjugué à la rotation, mais qui n'est pas conjugué de manière C^r.*

Démonstration de la proposition 3.4

Montrons d'abord un lemme intermédiaire.

Lemme 3.11. *Il existe un plongement du groupe des difféomorphismes C^r de l'intervalle ouvert* $]0,1[$ *dans le groupe des difféomorphismes C^r du cercle.*

Dimostrazione. Soit $\varphi : [0,1] \longrightarrow [0,1]$ un difféomorphisme de classe C^∞, infiniment tangent à l'identité en 0 et 1 et vérifiant la condition suivante : pour tout f, difféomorphisme de classe C^r de $]0,1[$, le difféomorphisme conjugué

$$\tilde{f} := \varphi \circ f \circ \varphi^{-1}$$

est un difféomorphisme de l'intervalle fermé $[0,1]$, C^r-tangent à l'identité en 0 et 1.

L'application qui à f associe $\tilde{f}$ définit un homomorphisme continu de $\mathrm{Diff}^r_+(]0,1[)$ dans le groupe

$$H := \{g \in \mathrm{Diff}^r_+([0,1]) \ : \ D^s g(0) = D^s g(1) = 1 \text{ pour tout } 0 < s \leq r\}$$

et ce groupe s'identifie naturellement au sous-groupe $\bar{H}$ de $\mathrm{Diff}^r_+(\mathbf{S}^1)$ des difféomorphismes du cercle C^r-tangents à l'identité en $0 \in \mathbf{S}^1$. □

La démonstration qui suit est une adaptation de la preuve du lemme 3.4 dans [Bri+].

Démonstration de la proposition 3.4. D'après le lemme 3.11, il suffit de montrer que le groupe des difféomorphismes de classe C^r de l'intervalle ouvert $]-1,1[$ n'est pas un groupe topologique, lorsque r n'est pas un nombre entier.

Soient $k := \lfloor r \rfloor$, $\tau := r - \lfloor r \rfloor$; pour tout $\delta > 0$, il existe un nombre rationnel p/q qui vérifie :

1. $|\tau - p/q| < \delta$,
2. p est un nombre impair, $q > k$ est un nombre pair.

Soit $\varepsilon > 0$, on définit les fonctions g et f_ε

$$g(x) = \frac{x + x^{p/q+k}}{2}, \quad f_\varepsilon(x) = x - \varepsilon + \varepsilon x^{q-1}.$$

La suite f_ε tend vers *id* dans la topologie C^r, lorsque $\varepsilon \to 0$. On se propose de montrer que la norme C^r de $g \circ f_\varepsilon - g$ ne tend pas vers 0, ce qui impliquera que le groupe $\mathrm{Diff}^r_+(\mathbf{S}^1)$ n'est pas topologique.

On a

$$D^k g(x) = C_1\, x^{p/q}, \quad D^k f_\varepsilon(x) = C_2\, x^{q-1-k},$$

d'où

$$D^k g(f_\varepsilon(x)) = C_1\, C_2 \left(x - \varepsilon + \varepsilon x^{q-1}\right)^{p/q} x^{q-1-k}$$

avec constantes $C_1 = (p/q + k)\,(p/q + k - 1) \cdots p/q$ et $C_2 = (q-1)\,(q-2) \cdots (q-1-k)$. Si l'on pose

$$\varphi(x) := D^k g(f_\varepsilon(x)) - D^k g(x),$$

alors

$$\varphi(x) = C_1\, C_2 \left(x - \varepsilon + \varepsilon x^{q-1}\right)^{p/q} x^{q-1-k} - C_1\, x^{p/q}.$$

La norme C^r de $g \circ f_\varepsilon - g$ est minorée par

$$\left| \frac{\varphi(\varepsilon) - \varphi(0)}{\varepsilon^{p/q}} \right| = \left| \frac{C_1\, C_2\, \varepsilon^{p+q-1-k} - C_1\, \varepsilon^{p/q} - C_1\, C_2\, (-\varepsilon)^{p/q}}{\varepsilon^{p/q}} \right|$$

qui est plus grande que $\frac{1}{2} C_1\,(1 + C_2) > 0$, lorsque ε est assez petit.

Quand r est un nombre entier, la continuité de la composition à gauche et à droite suit de la *formule de Faà di Bruno* [Yoc 2]. □

Capitolo I

Les mesures de Malliavin-Shavgulidze

1 Un panorama sur les mesures de Malliavin-Shavgulidze

1.1 Petite histoire

Les mathématiciens des années 1980 recherchaient des *mesures quasi-invariantes* sur $\mathrm{Diff}^1_+(I)$ et $\mathrm{Diff}^1_+(\mathbf{S}^1)$ afin de construire des représentations irréductibles pour ces groupes, comme on peut le faire pour un groupe localement compact grâce à sa mesure de Haar. Le théorème de Mackey et Weil 2.3 du chapitre précédent caractérise les groupes localement compacts comme étant les groupes qui possèdent une mesure quasi-invariante par l'action régulière gauche [Mac, Wei] : nous n'avons aucune possibilité de trouver une telle mesure sur $\mathrm{Diff}^1_+(I)$ (et $\mathrm{Diff}^1_+(\mathbf{S}^1)$), pourtant E. T. Shavgulidze donne des exemples de mesures quasi-invariantes pour les actions *restreintes* à des sous-groupes (notamment par les difféomorphismes de classe C^3).

Les premiers exemples dans [Sha 1] ne sont guère utilisables, mais on y voit déjà les outils que l'auteur utilisera dans ses futurs travaux. C'est une dizaine d'années plus tard que Shavgulidze reprendra son travail en définissant des mesures analogues pour les difféomorphismes de variétés riemanniennes arbitraires [Sha 2].

Probablement vers la fin des années 1980, les mathématiques de Shavgulidze rencontrent celles de Paul Malliavin qui s'intéresse alors aux processus de diffusion sur les variétés riemanniennes. Ils parviennent à la construction des mesures que nous appelons aujourd'hui mesures de Malliavin-Shavgulidze sur les difféomorphismes de l'intervalle et du cercle. Même si chacun est capable de construire des mesures boréliennes avec des propriétés d'invariance dans des contextes plus généraux, ils s'aperçoivent que seules ces mesures possèdent des propriétés vraiment remarquables [Mal-Mal 1, Mal-Mal 2, Sha 3].

En 1990, il n'est pas encore clair (pas encore écrit) que les mesures de Malliavin-Shavgulidze sont quasi-invariantes pour l'action des difféomorphismes C^3 : P. Malliavin et M. P. Malliavin montrent une quasi-invariance *infinitésimale* (avec des techniques d'analyse stochastique). Shavgulidze voit probablement déjà la quasi-invariance par translations (par exemple une preuve formelle pour les mesures sur $\mathrm{Diff}^1_+(I)$ est faite dans l'article de A. V. Kosyak [Kos]). Une véritable preuve apparaît en 1997 : elle est d'abord écrite dans [Sha 3] (en russe), et, avec un peu plus de détails, dans [Sha 4] (en anglais).

Nous avons fait le choix d'attribuer également le théorème de quasi-invariance pour les mesures de Malliavin-Shavgulidze à Cameron et Martin : dans l'un de leurs articles [Cam-Mar 2], on trouve un critère bien plus général pour qu'une action nonlinéaire sur $C_0(I)$ quasi-préserve la mesure de Wiener. Un simple calcul montre que l'action induite par la translation sur $\mathrm{Diff}^1_+(I)$ vérifie ces conditions !

Cependant, nous ne voulons pas réduire l'importance du travail de Malliavin et Shavgulidze : à partir de [Cam-Mar 2] il n'est pas du tout évident que le cocycle de Radon-Nikodym puisse prendre une forme si simple ![1]

Grâce à la lecture de [Cam-Mar 2], on augmente la compréhension des pages de calcul dans [Sha 4], qui autrement restent un peu mystérieuses.

Récemment P. A. Kuzmin [Kuz] reprend le résultat de Shavgulidze et l'utilise pour affaiblir le plus possible la régularité des fonctions qui quasi-préservent la mesure.

D'autres mesures quasi-invariantes sur $\mathrm{Diff}^1_+(\mathbf{S}^1)$ ont été introduites par Yu. A. Neretin [Ner] ; celui-ci cite les travaux de Malliavin et Shavgulidze, mais n'arrive pas à décrire le lien avec le sien.

Malliavin a décrit dans [Mal] une construction différente de processus de diffusion sur $\mathrm{Diff}^1_+(\mathbf{S}^1)$, mais cela n'amène pas à une mesure possédant les propriétés que nous souhaitons.

[1] Voir (I.7) et (I.8).

Remarque 1.1. P. Jones a défini une mesure sur les homéomorphismes du cercle et K. Astala, Jones, A. Kupiainen, E. Saksman [Ast-Jon-Kup-Sak] (voir aussi S. Sheffield [She 2]) ont montré que, presque sûrement cette mesure donne un homéomorphisme höldérien. La construction étant assez proche de celle de Malliavin et Shavgulidze, nous croyons fortement que la mesure de Jones est quasi-invariante par l'action des difféomoprhismes C^2.

1.2 Définition des mesures

Construire une mesure quasi-invariante sur un groupe topologique est un problème assez compliqué en général. Rappelons que, d'après ce que nous avons vu (partie 2.3 du chapitre précédent), lorsque G est un groupe non localement compact, il ne peut pas exister une mesure de Radon sur G, quasi-invariante par rapport à l'action de translation d'un sous-groupe de mesure strictement positive. Une mesure de Radon μ sur G sera dite *quasi-invariante* si elle est quasi-invariante par rapport à l'action par translation d'un sous-groupe $G_0 \leq G$, qui est de mesure nulle mais pourtant dense. Une telle définition amène la propriété suivante :

Proposition 1.1. *Soit μ une mesure de Radon non identiquement nulle sur le groupe topologique G. On suppose qu'il existe un sous-groupe dense $G_0 \leq G$ qui préserve la classe de la mesure μ. Alors pour tout ouvert $U \subset G$ arbitraire, $\mu(U) > 0$.*

Dimostrazione. En effet, le support de la mesure μ (défini comme étant le complémentaire du plus grand ouvert de mesure nulle) est un fermé G_0-invariant et coïncide donc avec G entier. □

D'après le théorème de Cameron-Martin 1.7, la mesure de Wiener sur $C_0(I)$ ou $C_0(\mathbf{S}^1)$ est quasi-invariante. En déhors de cet exemple, on n'en connait pas beaucoup d'autres (et ils sont certainement moins célèbres).

La construction des mesures de Malliavin-Shavgulidze s'appuie si fortement sur la mesure de Wiener, que d'un point de vue mesurable elles sont presque les mêmes. Mais laissons les généralités et rentrons dans le vif du sujet !

1.2.1 Les mesures sur le groupe des difféomorphismes de l'intervalle

La première étape consiste à identifier topologiquement le groupe de difféomorphismes de l'intervalle de classe C^1 qui préservent l'orientation, à l'espace de Wiener $C_0(I)$. Il y a certainement beaucoup de manières de le faire, puisque si $h : \mathrm{Diff}^1_+(I) \longrightarrow C_0(I)$ est un homéomorphisme, alors en composant h à la source par $\phi \in \mathrm{Homéo}_+(I)$, on obtient un autre homéomorphisme entre les deux espaces. Pourtant, il existe un choix particulièrement naturel.

En effet, pour connaître un difféomorphisme de l'intervalle f, il suffit de connaître sa dérivée Df, qui est une fonction continue strictement positive, de moyenne égale à 1. Réciproquement, soit h une fonction continue strictement positive sur l'intervalle I, on peut alors lui associer un difféomorphisme f défini par

$$f(t) = \frac{\int_0^t h(s)\,ds}{\int_0^1 h(s)\,ds}.$$

On remarquera que dans l'expression précédente un multiple constant de la fonction h définit le même f. On résume ceci.

Proposition 1.2. *L'espace $\mathrm{Diff}^1_+(I)$ des difféomorphismes C^1 de l'intervalle qui préservent l'orientation s'identifie à l'espace $C_1(I, \mathbf{R}_+)$ des fonctions continues positives sur l'intervalle vérifiant $h(0) = 1$. Cette application induit un homéomorphisme entre les deux espaces, munis respectivement des topologies C^1 et C^0.*

Dimostrazione. Nous avons déjà démontré la première partie de l'énoncé.

Soit h une fonction fixée et soit $\varepsilon > 0$. Nous voulons montrer qu'il existe $\delta > 0$ tel que si h et k se trouvent à une distance C^0 inférieure à δ, alors les difféomorphismes associés sont à distance C^1 inférieure à ε. Nous avons

$$\begin{aligned}
&\sup_{t\in I}\left|\frac{\int_0^t h(s)\,ds}{\int_0^1 h(s)\,ds} - \frac{\int_0^t k(s)\,ds}{\int_0^1 k(s)\,ds}\right| \\
&\qquad = \sup_{t\in I}\left|\frac{\int_0^t [h(s)-k(s)]\,ds}{\int_0^1 h(s)\,ds} + \int_0^t k(s)\,ds\left(\frac{1}{\int_0^1 h(s)\,ds} - \frac{1}{\int_0^1 k(s)\,ds}\right)\right| \\
&\qquad\qquad \leq \sup_{t\in I}\frac{t}{\int_0^1 h(s)\,ds}\delta + \sup_{t\in I}\frac{\int_0^t k(s)\,ds}{\int_0^1 h(s)\,ds\cdot\int_0^1 k(s)\,ds}\delta \leq \frac{2}{\int_0^1 h(s)\,ds}\delta,
\end{aligned}$$

et pour l'estimation sur la dérivée

$$\sup_{t\in I}\left|\frac{h(t)}{\int_0^1 h(s)\,ds}-\frac{k(t)}{\int_0^1 k(s)\,ds}\right|$$
$$\leq \sup_{t\in I}\left|\frac{h(t)-k(t)}{\int_0^1 k(s)\,ds}+k(t)\left(\frac{1}{\int_0^1 h(s)\,ds}-\frac{1}{\int_0^1 k(s)\,ds}\right)\right| \leq \left(1+\sup_{t\in I}\frac{k(t)}{\int_0^1 k(s)\,ds}\right)\frac{\delta}{\int_0^1 h(s)\,ds}$$
$$\leq \left(1+\frac{\sup_{t\in I} h(t)+\delta}{\left|\int_0^1 h(s)\,ds-\delta\right|}\right)\frac{\delta}{\int_0^1 h(s)\,ds} = \left(1+\sup_I \frac{h}{\int_0^1 h(s)\,ds}\right)\frac{\delta}{\left|\int_0^1 h(s)\,ds-\delta\right|}.$$

En posant

$$\delta = \frac{\varepsilon\int_0^1 h(s)\,ds}{1+\sup_I \frac{h}{\int_0^1 h(s)\,ds}+\varepsilon},$$

nous déduisons des estimations précédentes que la norme C^1 de la différence des difféomorphismes associés à h et k est inférieure à 2ε.

Réciproquement, soit F un difféomorphisme fixé et $\varepsilon > 0$. Nous cherchons $\delta > 0$ tel que, pour tout F' à distance C^1 de F inférieure à δ, les fonctions $DF/DF(0)$ et $DF'/DF'(0)$ soient à distance C^0 inférieure à ε.

Nous avons

$$\sup_I\left|\frac{DF}{DF(0)}-\frac{DF'}{DF'(0)}\right|$$
$$\leq \frac{1}{DF(0)}\delta+\sup_I DF'\left|\frac{1}{DF(0)}-\frac{1}{DF'(0)}\right| \leq \left(1+\sup_I\frac{DF'}{DF'(0)}\right)\frac{1}{DF(0)}\delta$$
$$\leq \left(1+\frac{\sup_I DF+\delta}{|DF(0)-\delta|}\right)\frac{1}{DF(0)}\delta = \left(1+\sup_I\frac{DF}{DF(0)}\right)\frac{\delta}{|DF(0)-\delta|}$$

et par conséquent

$$\delta = \frac{\varepsilon DF(0)}{1+\sup_I \frac{DF}{DF(0)}+\varepsilon}$$

suffit. □

Cette proposition termine « presque » notre première étape, puisque $C_1(I,\mathbf{R}_+)$ n'est pas très différent de l'espace de Wiener $C_0(I)$: en effet, par composition à gauche par la fonction logarithme, les deux espaces s'identifient topologiquement.

Dans la suite, on notera

$$A\ :\ \mathrm{Diff}^1_+(I) \longrightarrow C_0(I) \tag{I.1}$$

l'application qui à $f \in \mathrm{Diff}^1_+(I)$ associe la fonction continue $\log Df - \log Df(0)$. D'après la proposition 1.2, la fonction A est un homéomorphisme lorsque l'on munit les espaces de leurs topologies naturelles.

Observons que l'application inverse A^{-1} est définie par

$$A^{-1}x(t) = \frac{\int_0^t e^{x(s)}\,ds}{\int_0^1 e^{x(s)}\,ds},$$

et grâce à cette application, nous arrivons à la deuxième étape de la construction : la définition des mesures de Malliavin-Shavgulidze.

Il est peut-être temps de rappeler que la mesure de Wiener définie dans la partie 1.2 de l'introduction fait partie d'une famille à un paramètre de mesures : en notant $\sigma > 0$ ce paramètre, nous retrouvons avec $\sigma = 1$ la mesure de Wiener classique $\mathbf{W}$. Les autres mesures $\mathbf{W}_\sigma$ diffèrent par la *variance* que l'on considère pour les accroissements du mouvement brownien : si $x \in C_0(I)$ est distribué selon la mesure $\mathbf{W}$, la fonction $\sigma x \in C_0(I)$ est distribuée selon la mesure $\mathbf{W}_\sigma$. Plus précisément, la mesure $\mathbf{W}_\sigma$ d'un cylindre

$$C = \{x \in C_0(\mathbf{S}^1)\,|\,x(t_i)-x(t_{i-1}) \in A_i,\ i = 1,\dots n\},\quad 0 = t_0 < t_1 < \dots < t_n < t_{n+1} = 1 \text{ et } A_i \in \mathcal{B}(\mathbf{R})$$

est donnée par

$$\mathbf{W}_\sigma(C) = \frac{1}{\sqrt{(2\pi\sigma)^n (t_n - t_{n-1}) \cdots (t_1 - t_0)}} \int_{A_1} \cdots \int_{A_n} \exp\left(-\frac{1}{2\sigma} \sum_{i=1}^{n} \frac{y_i^2}{t_i - t_{i-1}} \right) dy_1 \cdots dy_n.$$

Remarque 1.2. Le paramètre σ mesure la distance entre la fonction aléatoire et la fonction identiquement nulle. Par exemple, lorsque σ est proche de zéro, les fonctions aléatoires se trouvent assez proches de la fonction nulle. Pour rendre cette affirmation plus formelle rappelons le *principe de réflexion* [Rev-Yor] : pour tout $r > 0$, on a

$$\mathbf{W}_\sigma\left(\{\|x\|_0 \leq r\}\right) \leq \frac{4}{\sqrt{2\pi}} \int_{r/\sigma}^{\infty} e^{-s^2/2}\, ds.$$

Mais σ mesure aussi la distance locale : en effet, pour σ proche de 0, les oscillations locales d'une fonction typique s'affaiblissent, comme l'explique la *loi des logarithmes itérés*

$$\mathbf{W}_\sigma\left(\left\{\overline{\lim_{\varepsilon\to 0}} \frac{x(t+\varepsilon) - x(t)}{\sqrt{2\varepsilon \log\log(1/\varepsilon)}} = \sqrt{\sigma}\right\}\right) = 1 \quad \text{pour tout } t \in I. \tag{I.2}$$

Définition 1.1. *Soit $\sigma > 0$. On définit la mesure de Malliavin-Shavgulidze de paramètre σ sur les difféomorphismes de l'intervalle, comme l'image de la mesure de Wiener par l'application A^{-1} :*

$$\nu_\sigma = (A^{-1})_* \mathbf{W}_\sigma. \tag{I.3}$$

Quand $\sigma = 1$, on parlera de la mesure de Malliavin-Shavgulidze, *que l'on notera ν_{MS} ou bien simplement ν.*

1.2.2 Les mesures sur le groupe des difféomorphismes du cercle

En modifiant la construction de la partie 1.2.1 on trouve des mesures définies sur le groupe des difféomorphismes du cercle. Dans ce contexte, le *cercle* $\mathbf{S}^1$ est paramétré par le tore euclidien $\mathbf{R}/\mathbf{Z}$.

Munis de leurs topologies naturelles, l'espace $\mathrm{Diff}^1_+(\mathbf{S}^1)$ des difféomorphismes C^1 du cercle s'identifie topologiquement à l'espace $C_0(\mathbf{S}^1) \times \mathbf{S}^1$, où $C_0(\mathbf{S}^1)$ dénote l'espace des fonctions continues $x : \mathbf{S}^1 \longrightarrow \mathbf{R}$ telles que $x(0) = 0$: à tout difféomorphisme f on associe le couple

$$B(f) = (B_0(f), f(0)), \tag{I.4}$$

avec

$$B_0(f)(t) = \log f'(t) - \log f'(0).$$

Remarque 1.3. L'application inverse est donnée par

$$B^{-1}(x, \alpha)(t) = c(x) \int_0^t e^{x(s)}\, ds + \alpha,$$

où $c(x) = \left(\int_0^1 e^{x(s)}\, ds\right)^{-1}$.

Le lecteur ayant lu la preuve de la proposition 1.2 n'aura pas de difficulté à montrer que B définit en effet un homéomorphisme.

Nous avons introduit à la ligne (5) une application π qui projette l'espace $C_0(I)$ sur $C_0(\mathbf{S}^1)$, grâce à laquelle nous avons pu définir la mesure de Wiener $\mathbf{W}_0$ sur $C_0(\mathbf{S}^1)$. Par la même méthode, on dispose des mesures $\mathbf{W}_{0,\sigma}$, définies par

$$\mathbf{W}_{0,\sigma} := \pi_* \mathbf{W}_\sigma.$$

Définition 1.2. *Soit $\sigma > 0$. On définit la mesure de Malliavin-Shavgulidze de paramètre σ sur les difféomorphismes du cercle, comme l'image du produit de la mesure de Wiener par la mesure de Lebesgue sur le cercle, par l'application B^{-1} :*

$$\mu_\sigma = B_*^{-1}\left(\mathbf{W}_{0,\sigma} \otimes \mathrm{Leb}\right). \tag{I.5}$$

Quand $\sigma = 1$, on parlera de la mesure de Malliavin-Shavgulidze, *qu'on notera μ_{MS} ou bien simplement μ.*

1.3 Modules de continuité

Nous l'avons vu dans la partie 3.2 de l'introduction, la classe de régularité d'un difféomorphisme du cercle influence les phénomènes dynamiques. Il est donc important, si l'on souhaite étudier les mesures de Malliavin-Shavgulidze d'un point de vue dynamique, de connaître la régularité d'un difféomorphisme aléatoire. Les propriétés de la mesure de Wiener que l'on a rencontrées dans la partie 1.2 nous permettent une description assez exhaustive. Nous avons de nouveau besoin de quelques rappels.

Soit $w : [0,1] \longrightarrow [0,+\infty[$ une fonction continue en 0 et vérifiant $w(s) = 0$ si et seulement si $s = 0$. On supposera aussi que w est croissante.

On définit l'espace $C^w(I)$ des fonctions avec *module de continuité* w :

$$C^w(I) = \left\{ x \in C(I) \,\middle|\, \sup_{s,t\in I, s\neq t} \frac{|x(s)-x(t)|}{w(|s-t|)} < \infty \right\}.$$

Si x est une fonction de module continuité w, on définit sa *norme* C^w :

$$\|x\|_w = \sup_{s,t\in I, s\neq t} \frac{|x(s)-x(t)|}{w(|s-t|)}.$$

Une fonction de module de continuité est automatiquement continue. Réciproquement, toute fonction sur I continue possède un module de continuité : si x est une fonction continue, on pose

$$w_x(\varepsilon) = \sup_{|s-t|\leq\varepsilon} |x(s)-x(t)|.$$

Exemple 1.1. Quand $w(s) = s^\alpha$, une fonction de module de continuité w est une fonction α-höldérienne.

Dans la suite, on notera $\mathrm{Diff}_+^{1+w}(I)$ l'espace des difféomorphismes de l'intervalle dont la dérivée est une fonction de module de continuité w. Si f est un difféomorphisme, on définit sa *norme* C^{1+w} :

$$\|f\|_{1+w} = \|f\|_1 + \|Df\|_w.$$

La proposition suivante est la clef de ce qui va suivre.

Proposition 1.3. *Soit w un module de continuité. L'application A définie dans (I.1) induit un homéomorphisme entre les difféomorphismes de l'intervalle de classe C^{1+w} et les fonctions de classe C^w, lorsque l'on munit ces espaces des topologies induites par leurs normes.*

Dimostrazione. Remarquons que si h est une fonction strictement positive définie sur l'intervalle, alors on a

$$\sup_I h\, \|h\|_w \leq \|\log h\|_w \leq \frac{1}{\min_I h} \|h\|_w,$$

puisque la fonction log est lipschitzienne sur l'intervalle compact $h(I) \subset]0,+\infty[$. Le lecteur pourra alors facilement terminer la preuve. □

Corollaire 1.4 (C^2 en aucun point)**.** *Soit $t \in I$ un point quelconque et E_t l'ensemble des difféomorphismes $f \in \mathrm{Diff}_+^1(I)$ tels que Df est continue au point t. Alors $\nu_\sigma(E_t) = 0$.*

Corollaire 1.5 (Presque jamais C^{1+vb})**.** *Soit $\mathrm{Diff}_+^{1+vb}(I)$ l'ensemble des difféomorphismes $f \in \mathrm{Diff}_+^1(I)$ tels que Df est une fonction de variation bornée sur I. Alors $\nu_\sigma\left(\mathrm{Diff}_+^{1+vb}(I)\right) = 0$.*

Corollaire 1.6 (Dérivée 1/2-höldérienne)**.**

1. *Pour tout $\tau < 1/2$, $\nu_\sigma\left(\mathrm{Diff}_+^{1+\tau}(I)\right) = 1$.*
2. *Pour tout $\tau > 1/2$, $\nu_\sigma\left(\mathrm{Diff}_+^{1+\tau}(I)\right) = 0$.*

Et pour être encore plus précis,

Corollaire 1.7 (Module de continuité de Lévy)**.** *Soit $w(t) = \sqrt{2t|\log t|}$ le module de continuité de Lévy, alors $\nu_\sigma\left(\mathrm{Diff}_+^{1+w}(I)\right) = 1$.*

Remarque 1.4. L'expression dans (4) exprime une propriété locale. Cependant, comme la fonction $\frac{x(t)-x(s)}{w(|t-s|)}$ est uniformément continue sur le carré $I \times I$ privé d'un voisinage de la diagonale, on déduit de (4) que la norme $\|x\|_w$ est finie.[2]

Remarquons que rien ne change dans les énoncés précédents si l'on s'intéresse aux mesures de Malliavin-Shavgulidze sur les difféomorphismes du cercle.

Proposition 1.8. *Soit w un module de continuité. L'application B définie dans (I.4) induit un homéomorphisme entre les difféomorphismes du cercle de classe C^{1+w} et les fonctions de classe C^w, quand on munit ces espaces des topologies induites par leurs normes.*

Corollaire 1.9. *Soit $w(t) = \sqrt{2t|\log t|}$ le module de continuité de Lévy. Pour tout $\sigma > 0$, les difféomorphismes du cercle de classe C^{1+w} ont mesure de Malliavin-Shavguldize μ_σ totale. Il en résulte qu'un difféomorphisme de Malliavin-Shavgulidze est $C^{1+\alpha}$-régulier pour tout $\alpha < 1/2$.*

2 Quasi-invariance des mesures de Malliavin-Shavgulidze

2.1 Le théorème de Cameron-Martin-Shavgulidze

Nous en venons à la propriété des mesures de Malliavin-Shavgulidze qui les rende remarquables vis-à-vis des autres mesures sur les groupes de difféomorphismes : leur *quasi-invariance*. Les bonnes hypothèses de régularité ont été montrées par Kuzmin [Kuz].

Théorème 2.1 (Cameron – Martin – Shavgulidze)**.** *Soit G l'un des groupes $\mathrm{Diff}^1_+(I)$ ou $\mathrm{Diff}^1_+(\mathbf{S}^1)$, soit μ une mesure de Malliavin-Shavgulidze sur G et soit H le sous-groupe de G constitué des difféomorphismes avec deuxième dérivée définie presque partout et dans L^∞. La mesure μ est quasi-invariante sous l'action du groupe H par translations à gauche.*

En outre, les cocycles de Radon-Nikodym sont, pour les mesures ν_σ et μ_σ respectivement :

$$\frac{d(L_\varphi)_*\nu_{MS}}{d\nu_{MS}}(f) = \exp\left\{-\int_0^1 \frac{\varphi''(f(t))}{\varphi'(f(t))}df'(t) - \frac{1}{2}\int_0^1 \left[\frac{\varphi''(f(t))}{\varphi'(f(t))}f'(t)\right]^2 dt\right\},$$

et

$$\frac{d(L_\varphi)_*\mu_{MS}}{d\mu_{MS}}(f) = \exp\left\{-\int_{\mathbf{S}^1} \frac{\varphi''(f(t))}{\varphi'(f(t))}df'(t) - \frac{1}{2}\int_{\mathbf{S}^1} \left[\frac{\varphi''(f(t))}{\varphi'(f(t))}f'(t)\right]^2 dt\right\}.$$

Remarque 2.1. Le lecteur observera la présence de la dérivée de Schwarz dans les cocycles de Radon-Nikodym (I.7) et (I.8). Nous rappelons que la dérivée schwarzienne d'un difféomorphisme φ se définit ainsi:

$$\mathcal{S}_\varphi(t) = \frac{\varphi'''(t)}{\varphi'(t)} - \frac{3}{2}\left(\frac{\varphi''(t)}{\varphi'(t)}\right)^2$$

ou bien encore

$$\mathcal{S}_\varphi(t) = D^2 \log D\varphi - \frac{1}{2}(D\log D\varphi)^2, \tag{I.6}$$

et vérifie la relation de cocycle $\mathcal{S}_{\varphi\psi} = \mathcal{S}_\varphi \circ \psi \cdot (D\psi)^2 + \mathcal{S}_\psi$. D'après cette relation, elle est naturellement définie sur l'espace des différentielles quadratiques sur le cercle. De plus, la dérivée $\mathcal{S}_\varphi$ est identiquement nulle si et seulement si φ est une transformation de Möbius. Elle pourrait suggérer une certaine invariance projective.

Lorsque $\varphi \in C^3$, on peut faire une intégration par parties dans les expressions donnant les cocycles :

$$\frac{d(L_\varphi)_*\nu_\sigma}{d\nu_\sigma}(f) = \exp\left\{-\frac{1}{\sigma}\left[\frac{\varphi''(f(1))}{\varphi'(f(1))}f'(1) - \frac{\varphi''(f(0))}{\varphi'(f(0))}f'(0) - \int_0^1 \left[\frac{\varphi'''(f(t))}{\varphi'(f(t))} - \frac{3}{2}\left(\frac{\varphi''(f(t))}{\varphi'(f(t))}\right)^2\right](f'(t))^2 dt\right]\right\}, \tag{I.7}$$

$$\frac{d(L_\varphi)_*\mu_\sigma}{d\mu_\sigma}(f) = \exp\left\{\frac{1}{\sigma}\int_{\mathbf{S}^1}\left[\frac{\varphi'''(f(t))}{\varphi'(f(t))} - \frac{3}{2}\left(\frac{\varphi''(f(t))}{\varphi'(f(t))}\right)^2\right](f'(t))^2 dt\right\}. \tag{I.8}$$

Pour compléter l'énoncé du théorème 2.1, nous observons que lorsque le sous-groupe H agit *à droite*, la classe de la mesure de μ n'est pas préservée, sauf dans un cas très particulier :

[2] Pour être plus précis, il faudrait modifier la fonction w pour la rendre croissante.

Théorème 2.2 ([Kos]). *Soit G l'un des groupes* $\mathrm{Diff}^1_+(I)$ *ou* $\mathrm{Diff}^1_+(\mathbf{S}^1)$, *soit μ une mesure de Malliavin-Shavgulidze sur G et soient $\varphi, \psi \in G$. Alors les mesures $(R_\varphi)_*\mu$ et $(R_\psi)_*\mu$ sont équivalentes si et seulement si φ et ψ diffèrent d'une rotation.*

Remarque 2.2. En particulier, sur le groupe $\mathrm{Diff}^1_+(I)$, aucune mesure de Malliavin-Shavgulidze n'est quasi-invariante par la composition à droite par aucun difféomorphisme. On en déduit qu'il existe un très grand nombre de classes de mesures différentes sur $\mathrm{Diff}^1_+(I)$ et $\mathrm{Diff}^1_+(\mathbf{S}^1)$ qui sont invariantes par l'action à gauche de H.

La preuve de ce dernier théorème utilise la propriété des logarithmes itérés (I.2) du mouvement brownien : si l'on considère la composition à droite par un difféomorphisme φ, en chaque point t on trouvera que l'événement

$$\overline{\lim_{\varepsilon\to 0}} \frac{|\log Df(t+\varepsilon) - \log Df(t)|}{\sqrt{2\varepsilon \log\log(1/\varepsilon)}} = \sqrt{\sigma\varphi'(t)} \tag{I.9}$$

est de mesure totale pour la nouvelle mesure de probabilité $(R_\varphi)_*\mu$.

L'une des différences majeures entre les actions à droite et à gauche est que la première induit un *changement de temps* pour le brownien correspondant à f, alors que l'action à gauche est simplement une « translation non-linéaire » : soit φ un difféomorphisme assez régulier et f un difféomorphisme aléatoire défini par un mouvement brownien B. En formules, l'action à droite donne au niveau de la mesure de Wiener

$$\log(f\circ\varphi)' - \log(f\circ\varphi)'(0) = B\circ\varphi + \log\varphi' - \log\varphi',$$

alors que pour celle à gauche

$$\log(\varphi\circ f)' - \log(\varphi\circ f)'(0) = B + \log\varphi'\circ f - \log\varphi'(f(0)).$$

Essayons de comprendre dans quel sens Cameron et Martin avaient démontré ce théorème 40 ans avant Shavgulidze, et avant même la définition des mesures de Malliavin-Shavgulidze. À une telle époque, l'étude de la mesure de Wiener et de ses propriétés était fortement en vogue. Rappelons que d'un point de vue purement mesurable, les mesures de Malliavin-Shavgulidze coïncident avec la mesure de Wiener et on remarquera que l'action de translation à gauche sur le groupe des difféomorphismes induit une action *non-linéaire* sur l'espace de Wiener.

Pour de telles transformations, Cameron et Martin montrent le théorème suivant [Cam-Mar 2] :

Théorème 2.3 (Cameron – Martin). *Soit $T : C_0(I) \longrightarrow C_0(I)$ une transformation non-linéaire de l'espace de Wiener $C_0(I)$. On suppose que la fonction $\Lambda = T - id$ possède une* dérivée variationnelle

$$\delta\Lambda(x) = \frac{d}{d\varepsilon}\Lambda(x+\varepsilon\cdot\delta x)\Big|_{\varepsilon=0}$$

de la forme

$$\delta\Lambda(x)(t) = \int_0^1 K(x;s)(t)\,\delta x(s)\,ds, \quad \textit{pour } t\in I.$$

Alors la mesure de Wiener $\mathbf{W}$ sur $C_0(I)$ est quasi-préservée par T et le cocycle de Radon-Nikodym est donné par

$$\frac{d(T^{-1})_*\mathbf{W}}{d\mathbf{W}}(x) = |\mathcal{D}(x)|\exp\left\{-2\int_0^1 \Lambda(x)'dx - \int_0^1 [\Lambda(x)']^2 dx\right\}, \tag{I.10}$$

où $\mathcal{D}(x)$ dénote le déterminant de Fredholm

$$\mathcal{D}(x) = 1 + \sum_{n=1}^{\infty}\frac{1}{n!}\int_0^1\cdots\int_0^1 \begin{vmatrix} K(x;t_1)(t_1) & \cdots & K(x;t_1)(t_n) \\ \vdots & & \vdots \\ K(x;t_n)(t_1) & \cdots & K(x;t_n)(t_n)\end{vmatrix} dt_1\cdots dt_n. \tag{I.11}$$

Notre but consiste maintenant à traiter la transformation non-linéaire que L_φ induit sur l'espace de Wiener. Pour des raisons de simplicité, nous considérerons seulement les mesures ν_{MS} et μ_{MS} (la variance σ ne représente pas vraiment un obstacle, mais elle alourdit les calculs).

Soit $\varphi \in \mathrm{Diff}^3_+(I)$ et posons $h := A(\varphi^{-1})$, on écrira alors

$$S_h(x) = A\circ L_\varphi\circ A^{-1}(x), \tag{I.12}$$

avec

$$S_h(x)(t) = x(t) + h\left(c(x)\int_0^t e^{x(s)}ds\right), \quad c(x) = \frac{1}{\int_0^1 e^x}.$$

Le groupe des difféomorphismes C^3 agit sur $C_0(I)$ par le S défini par (I.12).

De manière similaire, lorsque $\varphi \in \mathrm{Diff}^3_+(\mathbf{S}^1)$, on posera $h := B_0(\varphi^{-1})$ et

$$B \circ L_\varphi \circ B^{-1}(x,\alpha) = (S_h(x,\alpha), \varphi^{-1}(\alpha)),$$

avec

$$S_h(x,\alpha)(t) = x(t) + h\left(\alpha + c(x)\int_0^t e^{x(s)}ds\right) - h(\alpha), \quad c(x) = \frac{1}{\int_0^1 e^x},$$

qui définit une action sur $C_0(\mathbf{S}^1) \times \mathbf{S}^1$.

Remarque 2.3. Nous laisserons au lecteur la tâche de généraliser le théorème de Cameron-Martin 2.3 aux transformations de la mesure de Wiener $w_0 \otimes \mathrm{Leb}$ sur $C_0(\mathbf{S}^1) \times \mathbf{S}^1$.

Pour pouvoir appliquer le théorème de Cameron-Martin, il nous faut vérifier que la fonction $\Lambda_h := S_h - id$ possède une dérivée variationnelle et la calculer :

$$\begin{aligned}
&\frac{d}{d\varepsilon}h\left(c(x+\varepsilon\cdot\delta x)\int_0^t e^{x+\varepsilon\cdot\delta x}\right)\Bigg|_{\varepsilon=0} \\
&\qquad = h'\left(c(x)\int_0^t e^x\right)\left[c(x)\int_0^t (e^x\delta x) - c(x)^2\int_0^1 (e^x\delta x)\right] \\
&\qquad\qquad = \int_0^1 \left[h'\left(c(x)\int_0^t e^x\right)c(x)\left(\mathbf{1}_{[0,t]}(s) - c(x)\right)e^{x(s)}\right]\delta x(s)\,ds.
\end{aligned}$$

On pose alors

$$K_h(x;s)(t) = h'\left(c(x)\int_0^t e^x\right)c(x)\left(\mathbf{1}_{[0,t]}(s) - c(x)\right)e^{x(s)}, \tag{I.13}$$

qui représente le noyau nécessaire pour calculer le déterminant de Fredholm (I.11).

Nous pouvons suivre le même parcours pour trouver le noyau dans le cas des transformations de $C_0(\mathbf{S}^1) \times \mathbf{S}^1$ pour obtenir

$$K_h(x,\alpha;s)(t) = h'\left(\alpha + c(x)\int_0^t e^x\right)c(x)\left(\mathbf{1}_{[0,t]}(s) - c(x)\right)e^{x(s)}, \tag{I.14}$$

En revanche, il n'est pas immédiat, que les déterminants de Fredholm obtenus par (I.13) et (I.14) mènent aux dérivées de Radon-Nikodym (I.7) et (I.8).

Pour cette raison, nous avons choisi de reproduire la preuve de Shavgulidze, qui s'inspire à son tour, dans un cas très particulier, de la méthode de preuve de Cameron et Martin.

2.2 Ergodicité

La discussion précédente sur la régularité est en partie indispensable pour le sujet que nous allons aborder. En effet, la preuve de Kosyak du théorème 2.4 serait un peu plus compliquée si l'on avait seulement la quasi-invariance de la mesure de Malliavin-Shavgulidze par les difféomorphismes de classe C^3.

Soit G l'un des groupes $\mathrm{Diff}^1_+(I)$ ou $\mathrm{Diff}^1_+(\mathbf{S}^1)$, soit μ une mesure de Malliavin-Shavgulidze sur G, le couple (G,μ) définit un espace mesuré, sur lequel le groupe H agit, non en préservant la mesure, mais en préservant les ensembles négligeables (l'action de H *quasi-préserve* la mesure μ).

Définition 2.1. *Soit (X,μ) un espace mesuré standard sur lequel un groupe H agit en quasi-préservant la mesure μ.*
On dit que l'action de H est ergodique *si tout ensemble mesurable invariant est négligeable ou bien conégligeable*[3].

Dans ce contexte, Kosyak a étudié l'ergodicité de l'action de H sur (G,μ) [Kos].

[3] Un ensemble mesurable $A \subset X$ est *conégligeable* si le complémentaire dans X est négligeable.

Théorème 2.4 (Kosyak). *Soit G l'un des groupes $\mathrm{Diff}^1_+(I)$ ou $\mathrm{Diff}^1_+(\mathbf{S}^1)$, soit μ une mesure de Malliavin-Shavgulidze sur G et soit H un sous-groupe de G qui contient tous les difféomorphismes de classe $C^{2+\alpha}$ ($\alpha < 1/2$).*

L'action de H sur (G, μ) est ergodique.

Dimostrazione. La preuve du théorème 2.4 consiste à étudier l'action non-linéaire donnée par les transformations S_h (voir (I.12)) sur l'espace de Wiener, mais vu de manière « hilbertienne », comme nous l'avons fait dans la partie 1.4. Pour commencer, on s'intéressera à l'action de H sur $(C_0(I), \mathbf{W})$.

Avec le choix de la base orthonormée de Fourier $(g_n)_{n\in\mathbf{N}}$ de $L^2(I)$ (voir (12)), nous avons vu que la variable

$$x(t) = \sum_{n\in\mathbf{N}} Z_n \Phi_n(t)$$

définit un mouvement brownien, lorsque les Z_n forment une suite de variables aléatoires gaussiennes centrées de variance 1 i.i.d., et les Φ_n sont définies par intégration des g_n (voir la partie 1.4). Nous allons nous restreindre aux fonctions x appartenant à un ensemble de mesure pleine $\Omega \subset C_0(I)$ de sorte que tout $x \in \Omega$ est de classe $C^{1/2-\varepsilon}$ pour tout $\varepsilon > 0$.

L'expression définissant S_h dit que la fonction x est transformée en

$$S_h(x)(t) = x(t) + h\left(c(x)\int_0^t e^{x(s)}\,ds\right) = \sum_{n\in\mathbf{N}}\left[Z_n + \left\langle h\left(c(x)\int_0^t e^{x(s)}\,ds\right), \Phi_n\right\rangle\right]\Phi_n.$$

Fixons $x \in \Omega$ et soit $(a_n)_{n\in\mathbf{N}}$ une suite à support dans $\{0, \dots, N\}$, soit h tel que $h\left(c(x)\int_0^t e^{x(s)}\,ds\right) = \sum_{n\le N} a_n\Phi_n(t)$, alors h est une fonction de classe $C^{3/2-\varepsilon}$ (les éléments de la base de Fourier sont de classe C^∞), et si l'on définit $\varphi^{-1} := A^{-1}(h)$, alors $\varphi \in H$.

Pourquoi ce choix ? En fait, nous venons de montrer que la suite $(Z_n)_{n\in\mathbf{N}}$ peut être envoyée par un élément de H sur la suite $(Z_n + a_n)_{n\in\mathbf{N}}$, qui peut être arbitrairement choisie sur toute fenêtre $\{0, \dots, N\}$.

En regardant les coefficients de Fourier, nous identifions $(C_0(I), \mathbf{W})$ avec l'espace gaussien $(\mathbf{R}^{\mathbf{N}}, \gamma^{\otimes\mathbf{N}})$, où γ est la mesure gaussienne standard sur $\mathbf{R}$. Nous avons que toute orbite de l'action sur $(\mathbf{R}^{\mathbf{N}}, \gamma^{\otimes\mathbf{N}})$ du groupe $\mathbf{R}^{fin}$ des translations à support fini, est inclue dans une orbite de l'action non-linéaire induite par H. Or, il est classique que l'action de $\mathbf{R}^{fin}$ est ergodique (pour les probabilistes, il s'agit fondamentalement de la loi du 0–1 de Hewitt-Savage [Mör-Per, Lemma 1.25]) et on peut finalement conclure que l'action de H sur $(\mathrm{Diff}^1_+(I), \nu_{MS})$ est ergodique.

La preuve pour l'action sur $(\mathrm{Diff}^1_+(\mathbf{S}^1), \mu_{MS})$ suit les mêmes lignes : en notant $\tilde{\Phi}_n$ la base de $\mathbf{D}_0(\mathbf{S}^1)$, pour toute suite $(a_n)_{n\in\mathbf{N}}$ à support dans $\{0, \dots, N\}$ et $\theta \in \mathbf{S}^1$, on peut définir $\varphi \in H$ tel que

1. $h = B_0(\varphi^{-1})$,
2. $h\left(\alpha + c(x)\int_0^t e^{x(s)}\,ds\right) = \sum_{n\le N} a_n\tilde{\Phi}_n(t)$,
3. $\varphi^{-1}(\alpha) = \theta$,

où $(x, \alpha) \in C_0(\mathbf{S}^1) \times \mathbf{S}^1$ est fixé. □

Nous verrons plus tard que connaître l'ergodicité de l'action est intéressant à un niveau *cohomologique* (partie 3).

2.3 Preuve de la quasi-invariance de la mesure

Rentrons finalement dans les détails du théorème de Cameron-Martin-Shavgulidze 2.1. Nous avons choisi de suivre les lignes de la preuve dans [Sha 4, Sha 3] : l'âme de la démonstration étant bien classique (dans l'esprit de Cameron et Martin), il s'avère pourtant qu'elle nécessite de longs calculs. Dans [Sha 4], Shavgulidze ne montre le théorème de quasi-invariance que pour la mesure sur $\mathrm{Diff}^1_+(I)$.

Notre choix a été de réécrire la preuve de Shavgulidze pour la quasi-invariance de la mesure μ_{MS} sur $\mathrm{Diff}^1_+(\mathbf{S}^1)$, en complétant tous les détails. Pour faciliter la lecture, nous avons divisé la démonstration en plusieurs morceaux.

Approximation par des cylindres. Rappelons que d'après le théorème de Wiener 1.1 (adapté à l'espace $C_0(\mathbf{S}^1)$), les cylindres

$$C = \{x \in C_0(\mathbf{S}^1) \,|\, x(t_i) - x(t_{i-1}) \in A_i,\ i = 1, \dots n\}, \quad 0 = t_0 < t_1 < \dots < t_n < t_{n+1} = 1 \text{ et } A_i \in \mathcal{B}(\mathbf{R})$$

engendrent la tribu borélienne $\mathcal{B}\big(C_0(\mathbf{S}^1)\big)$. En particulier, par densité des nombres rationnels dans $[0,1]$, on pourra se restreindre à considérer les cylindres avec des partitions à la base données par $t_j = \frac{j}{n+1}$, $j = 0, \dots, n+1$.

Pour tout $n \in \mathbf{N}$, on note π_n la projection de $C_0(\mathbf{S}^1)$ sur $\mathbf{R}^n$ qui à une fonction continue x associe le vecteur $\big(x(\frac{1}{n+1}), \dots, x(\frac{n}{n+1})\big)$.

La transformation π_n induit une application de l'espace des fonctions essentiellement bornées $L^\infty(C_0(\mathbf{S}^1), \mathbf{W}_0)$ dans l'espace des fonctions essentiellement bornées $L^\infty(\mathbf{R}^n)$

$$F \longmapsto F_n$$

avec $F_n(x_1, \dots, x_n) = F(x^{(n)}_{\text{aff}})$, où $x^{(n)}_{\text{aff}}$ est l'unique fonction affine par morceaux sur $\mathbf{S}^1$ qui vaut x_i au temps t_i. Le théorème de Wiener équivaut au résultat suivant :

Proposition 2.5. *Soit F une fonction bornée sur $C_0(\mathbf{S}^1)$. On a*

$$\lim_{n\longrightarrow\infty} \left(\frac{n+1}{2\pi}\right)^{\frac{n+1}{2}} \int_{\mathbf{R}^n} F_n(x_1, \dots, x_n) \exp\left\{-\frac{n+1}{2}\sum_{m=1}^{n+1}(x_m - x_{m-1})^2\right\} dx_1 \cdots dx_n$$
$$= \int_{C_0(\mathbf{S}^1)} F(x)\, d\mathbf{W}_0(x).$$

Lorsque $F = F(x, \alpha)$ est une fonction essentiellement bornée sur $C_0(\mathbf{S}^1) \times \mathbf{S}^1$, on définira $F_n = F(x_1, \dots, x_n, \alpha)$ de manière analogue.

Cette approximation de dimension finie pour la mesure de Wiener nous permettra de calculer le changement de la mesure induit par la transformation $S_h \times \varphi^{-1}$, avec $h = B_0(\varphi^{-1})$. Soit $F = F(x, \alpha)$ une fonction bornée sur $C_0(\mathbf{S}^1) \times \mathbf{S}^1$, l'intégrale de F par rapport à $\left(S_h \times \varphi^{-1}\right)_* (\mathbf{W}_0 \otimes \text{Leb})$ s'écrit

$$\int_{C_0(\mathbf{S}^1)\times\mathbf{S}^1} F(x, \alpha)\, d\left(S_h \times \varphi^{-1}\right)_* (\mathbf{W}_0 \otimes \text{Leb})(x, \alpha) = \int_{C_0(\mathbf{S}^1)\times\mathbf{S}^1} F(x, \alpha)\, d\mathbf{W}_0(S_{\tilde{h}}(x, \alpha))\, d\varphi(\alpha),$$

où $\tilde{h} = B_0(\varphi)$, et on voudrait l'écrire sous la forme

$$\int_{C_0(\mathbf{S}^1)\times\mathbf{S}^1} F(x, \alpha) \cdot \frac{d\left(S_h \times \varphi^{-1}\right)_* (\mathbf{W}_0 \otimes \text{Leb})}{d(\mathbf{W}_0 \otimes \text{Leb})}(x, \alpha)\, d\mathbf{W}_0(x)\, d\alpha.$$

Notons

$$(y, \beta) = (S_{\tilde{h}}(x, \alpha), \varphi(\alpha))$$

et soit $S_{n,\tilde{h}}$ l'application de $\mathbf{R}^n \times \mathbf{S}^1$ dans $\mathbf{R}^n \times \mathbf{S}^1$ qui a $(x^{(n)}_{\text{aff}}, \alpha)$ associe $(y^{(n)}_{\text{aff}}, \beta)$.

Lorsque n tend vers l'infini, la transformation $S_{n,\tilde{h}}$ tend vers $S_{\tilde{h}}$ dans le sens suivant :

$$\lim_{n\longrightarrow\infty} \left(\frac{n+1}{2\pi}\right)^{\frac{n+1}{2}} \int_{\mathbf{R}^n\times\mathbf{S}^1} F_n(x_1, \dots, x_n, \alpha) \exp\left\{-\frac{n+1}{2}\sum_{m=1}^{n+1}(y_m - y_{m-1})^2\right\} dy_1 \cdots dy_n\, d\beta$$
$$= \int_{C_0(\mathbf{S}^1)\times\mathbf{S}^1} F(x, \alpha)\, d\mathbf{W}_0(y)\, d\beta.$$

On trouvera la densité limite $\frac{d\left(S_h\times\varphi^{-1}\right)_*(\mathbf{W}_0\otimes\text{Leb})}{d(\mathbf{W}_0\otimes\text{Leb})}(x, \alpha)$ en calculant, pour n assez grand, le changement de la mesure induit par $S_{n,h} \times \varphi^{-1}$:

$$\frac{d\left(S_h \times \varphi^{-1}\right)_* (\mathbf{W}_0 \otimes \text{Leb})}{d(\mathbf{W}_0 \otimes \text{Leb})}(x, \alpha)$$
$$= \lim_{n\longrightarrow\infty} \frac{\exp\left\{-\frac{n+1}{2}\sum_{m=1}^{n+1}(y_m - y_{m-1})^2\right\}}{\exp\left\{-\frac{n+1}{2}\sum_{m=1}^{n+1}(x_m - x_{m-1})^2\right\}} \times \left|\text{Jac}(S_{n,\tilde{h}} \times \varphi)(x_1, \dots, x_n, \alpha)\right|.$$

Nous reconnaissons ici deux facteurs distincts : l'un est relatif à la *densité gaussienne* et l'autre au *jacobien*. Nous analyserons séparément les deux morceaux.

Jacobien. Dans cette partie nous nous consacrons à calculer le jacobien de la transformation $S_{n,\tilde{h}} \times \varphi$.

$$\mathrm{Jac}\left(S_{n,\tilde{h}}\right)_{ij}(x_1,\ldots,x_n,\alpha) = \delta_{ij} + h'\left(\alpha + \frac{\sum_{m=0}^{i-1} e^{x_m}}{\sum_{m=0}^{n} e^{x_m}}\right) e^{x_j} \frac{\mathbf{1}_{[j<i]} \sum_{m=0}^{n} e^{x_m} - \sum_{m=0}^{i-1} e^{x_m}}{\left(\sum_{m=0}^{n} e^{x_m}\right)^2}. \tag{I.15}$$

Remarque 2.4. Nous recommandons au lecteur de confronter le jacobien (I.15) avec l'expression du noyau K_h dans (I.14).

On pourra écrire la matrice jacobienne sous la forme

$$\begin{pmatrix} I_n + U + V & * \\ 0 & \varphi'(\alpha) \end{pmatrix}$$

avec U matrice triangulaire inférieure nilpotente et V matrice de rang 1 :

$$U_{ij} = \mathbf{1}_{[j<i]} \tilde{h}'\left(\alpha + \frac{\sum_{m=0}^{i-1} e^{x_m}}{\sum_{m=0}^{n} e^{x_m}}\right) \frac{e^{x_j}}{\sum_{m=0}^{n} e^{x_m}},$$

$$V_{ij} = -\tilde{h}'\left(\alpha + \frac{\sum_{m=0}^{i-1} e^{x_m}}{\sum_{m=0}^{n} e^{x_m}}\right) e^{x_j} \frac{\sum_{m=0}^{i-1} e^{x_m}}{\left(\sum_{m=0}^{n} e^{x_m}\right)^2}.$$

Shavgulidze remarque alors que le fait que le déterminant de la matrice $I_n + U$ soit égal à 1 implique

$$\mathrm{Dét}\,(I_n + U + V) = \mathrm{Dét}\left(I_n + (I_n + U)^{-1} V\right).$$

Comme V est de rang 1, la matrice $(I_n + U)^{-1} V$ est encore de rang 1 et s'écrit sous la forme $(I_n + U)^{-1} a \cdot b^t$, avec $a, b \in \mathbf{R}^n$ (les définitions explicites apparaissent en (I.16)). Une telle matrice a pour valeurs propres 0 et $b^t \cdot (I_n + U)^{-1} a$ et, en diagonalisant, on trouve

$$\mathrm{Dét}\,(I_n + U + V) = 1 + b^t \cdot (I_n + U)^{-1} a.$$

La matrice inverse $(I_n + U)^{-1}$ est donnée par $I_n - U + \ldots + (-1)^{n-1} U^{n-1}$. On en déduit

$$\mathrm{Dét}\,(I_n + U + V) = 1 + b^t \cdot a - \ldots + (-1)^{n-1} b^t \cdot U^{n-1} a.$$

Il nous reste encore à écrire les vecteurs a et b :

$$a_i = -\tilde{h}'\left(\alpha + \frac{\sum_{m=0}^{i-1} e^{x_m}}{\sum_{m=0}^{n} e^{x_m}}\right) \frac{\sum_{m=0}^{i-1} e^{x_m}}{\left(\sum_{m=0}^{n} e^{x_m}\right)^2}, \quad b_j = e^{x_j}. \tag{I.16}$$

Nous écrirons aussi $U_{ij} = \widetilde{a}_i b_j$ (pour $j < i$), avec les $\widetilde{a}_i$ donnés par $a_i = -\frac{\sum_{m=0}^{i-1} e^{x_m}}{\sum_{m=0}^{n} e^{x_m}} \widetilde{a}_i$. Avec ces notations on trouve

$$\begin{aligned} &\mathrm{Dét}\,(I_n + U + V) \\ &= 1 + \sum_{i=1}^{n} b_i a_i - \sum_{i=2}^{n} b_i \sum_{j=1}^{i-1} U_{ij} a_j + \ldots + (-1)^{n-1} \sum_{i=2}^{n} b_i \sum_{i_1=1}^{i-1} U_{ii_1} \sum_{i_2=1}^{i_1-1} U_{i_1 i_2} \cdots \sum_{j=1}^{i_{n-2}} U_{i_{n-2} j} a_j \\ &= 1 + \sum_{i=1}^{n} b_i a_i - \sum_{i=2}^{n} \sum_{j=1}^{i-1} \widetilde{a}_i b_i a_j b_j \prod_{k=j+1}^{i-1} (1 - \widetilde{a}_k b_k). \end{aligned}$$

Grâce aux équivalences asymptotiques

$$a_i = o(1/n), \quad b_i = o(1)$$

on trouve

$$\mathrm{Dét}\,(I_n + U + V) = 1 + \sum_{i=1}^{n} a_i b_i - \sum_{i=2}^{n} \sum_{j=1}^{i-1} \widetilde{a}_i b_i a_j b_j \exp\left(-\sum_{k=j+1}^{i-1} \widetilde{a}_k b_k\right) + o\left(\frac{1}{n}\right).$$

Ensuite, nous prenons la limite pour $n \to \infty$:

$$\lim_{n\to\infty} \text{Dét}\,(I_n + U + V) = 1 - \int_0^1 \tilde{h}'\left(\alpha + \frac{\int_0^t e^x}{\int_0^1 e^x}\right) \frac{\int_0^t e^x}{\left(\int_0^1 e^x\right)^2} e^{x(t)} dt$$

$$+ \int_0^1 \tilde{h}'\left(\alpha + \frac{\int_0^t e^x}{\int_0^1 e^x}\right) \frac{e^{x(t)}}{\int_0^1 e^x} \int_0^t \tilde{h}'\left(\alpha + \frac{\int_0^s e^x}{\int_0^1 e^x}\right) \frac{\int_0^s e^x}{\left(\int_0^1 e^x\right)^2} e^{x(s)} \exp\left(-\int_s^t \tilde{h}'\left(\alpha + \frac{\int_0^u e^x}{\int_0^1 e^x}\right) \frac{e^{x(u)}}{\int_0^1 e^x} du\right) ds\, dt$$

$$\underset{\text{changements de variable } t' = \frac{\int_0^t e^x}{\int_0^1 e^x}}{=} 1 - \int_0^1 \tilde{h}'(\alpha + t) t\, dt + \int_0^1 \tilde{h}'(\alpha + t) \int_0^t \tilde{h}'(\alpha + s) s \exp(-\tilde{h}(\alpha + t) + \tilde{h}(\alpha + s))\, ds\, dt$$

$$\underset{\text{translation par } \alpha}{=} 1 - \int_\alpha^{1+\alpha} \tilde{h}'(t)(t - \alpha)\, dt + \int_\alpha^{1+\alpha} \tilde{h}'(t) \int_\alpha^t \tilde{h}'(s)(s - \alpha) \exp(-\tilde{h}(t) + \tilde{h}(s))\, ds\, dt$$

$$= 1 - \int_\alpha^{1+\alpha} \tilde{h}'(t) \left[(t - \alpha) - \exp(-\tilde{h}(t)) \int_\alpha^t \tilde{h}'(s) \exp(\tilde{h}(s))(s - \alpha)\, ds\right] dt$$

$$\underset{\text{intégration par parties}}{=} 1 - \int_\alpha^{1+\alpha} \tilde{h}'(t) \left[(t - \alpha) - (t - \alpha) + \exp(-\tilde{h}(t)) \int_\alpha^t \exp(\tilde{h}(s))\, ds\right] dt$$

$$= 1 - \int_\alpha^{1+\alpha} \tilde{h}'(t) \exp(-\tilde{h}(t)) \int_\alpha^t \exp(\tilde{h}(s))\, ds\, dt$$

$$\underset{\text{intégration par parties}}{=} 1 + e^{-\tilde{h}(\alpha)} \int_\alpha^{1+\alpha} e^{\tilde{h}(s)}\, ds - 1$$

$$= \int_0^1 e^{-\tilde{h}(\alpha) + \tilde{h}(s)}\, ds.$$

Nous pouvons en conclure l'égalité

$$\lim_{n\to\infty} \left|\text{Jac}\left(S_{n,\tilde{h}} \times \varphi\right)(x_1, \ldots, x_n, \alpha)\right| = \int_0^1 e^{-\tilde{h}(\alpha) + \tilde{h}(s)}\, ds \cdot \varphi'(\alpha) = \varphi'(0) \int_0^1 e^{\tilde{h}(s)}\, ds.$$

Densité gaussienne. Nous étudions ici le comportement asymptotique de la fonction de densité

$$\delta(x_1, \ldots, x_n, \alpha) = \left(\frac{n+1}{2\pi}\right)^{(n+1)/2} \exp\left\{-\frac{n+1}{2} \sum_{m=1}^{n+1} (x_m - x_{m-1})^2\right\}, \quad x_0 = x_{n+1} = 0$$

lorsque elle est transformée par $S_{n,\tilde{h}}$. Nous avons

$$\delta \circ S_{n,\tilde{h}}(x_1, \ldots, x_n, \alpha)$$

$$= \delta(x_1, \ldots, x_n, \alpha) \cdot \exp\left\{-\frac{n+1}{2} \sum_{m=1}^{n+1} \left(\tilde{h}\left(\alpha + \frac{\sum_{k=0}^{m-1} e^{x_k}}{\sum_{k=0}^{n} e^{x_k}}\right) - \tilde{h}\left(\alpha + \frac{\sum_{k=0}^{m-2} e^{x_k}}{\sum_{k=0}^{n} e^{x_k}}\right)\right)^2\right\}$$
$$\times \exp\left\{-(n+1) \sum_{m=1}^{n+1} (x_m - x_{m-1}) \left(\tilde{h}\left(\alpha + \frac{\sum_{k=0}^{m-1} e^{x_k}}{\sum_{k=0}^{n} e^{x_k}}\right) - \tilde{h}\left(\alpha + \frac{\sum_{k=0}^{m-2} e^{x_k}}{\sum_{k=0}^{n} e^{x_k}}\right)\right)\right\}$$

et par suite

$$\lim_{n\to\infty} \frac{\delta \circ S_{n,\tilde{h}}(x_1,\dots,x_n,\alpha)}{\delta(x_1,\dots,x_n,\alpha)} = \exp\left\{-\frac{1}{2}\int_0^1 \left[\left(\tilde{h}\left(\alpha + \frac{\int_0^t e^x}{\int_0^1 e^x}\right)\right)'\right]^2 dt - \int_0^1 \left(\tilde{h}\left(\alpha + \frac{\int_0^t e^x}{\int_0^1 e^x}\right)\right)' dx\right\}. \tag{I.17}$$

Remarque 2.5. Dans la dernière égalité, la deuxième intégrale est au sens de Wiener.

Remarque 2.6. Nous renvoyons le lecteur à confronter le résultat dans (I.17) avec l'expression (I.10) dans l'énoncé du théorème de Cameron-Martin 2.3.

On poursuit le calcul :

$$\begin{aligned}
&\lim_{n\to\infty} \frac{\delta \circ S_{n,\tilde{h}}(x_1,\dots,x_n,\alpha)}{\delta(x_1,\dots,x_n,\alpha)} \\
&\qquad = \exp\left\{-\frac{1}{2}\int_0^1 \left[\left(\tilde{h}\left(\alpha + \frac{\int_0^t e^x}{\int_0^1 e^x}\right)\right)'\right]^2 dt - \int_0^1 \tilde{h}'\left(\alpha + \frac{\int_0^t e^x}{\int_0^1 e^x}\right) d\frac{e^x}{\int_0^1 e^x}\right\} \\
&\qquad \underset{\text{intégration par parties}}{=} \exp\left\{-\frac{1}{2}\int_0^1 \left[\left(\tilde{h}\left(\alpha + \frac{\int_0^t e^x}{\int_0^1 e^x}\right)\right)'\right]^2 dt + \int_0^1 \tilde{h}''\left(\alpha + \frac{\int_0^t e^x}{\int_0^1 e^x}\right) \frac{e^{2x(t)}}{\left(\int_0^1 e^x\right)^2} dt\right\} \\
&\qquad = \exp\left\{\int_0^1 \left[\tilde{h}''\left(\alpha + \frac{\int_0^t e^x}{\int_0^1 e^x}\right) - \frac{1}{2}\left(\tilde{h}'\left(\alpha + \frac{\int_0^t e^x}{\int_0^1 e^x}\right)\right)^2\right] \frac{e^{2x(t)}}{\left(\int_0^1 e^x\right)^2} dt\right\}.
\end{aligned}$$

Recollons les morceaux. Les calculs précédents nous donnent :

Proposition 2.6. *Soit $\varphi \in \mathrm{Diff}^3_+(\mathbf{S}^1)$ et $h = B_0(\varphi^{-1})$, $\tilde{h} = B_0(\varphi)$, alors pour tout $(x,\alpha) \in C_0(\mathbf{S}^1) \times \mathbf{S}^1$ nous avons*

$$\begin{aligned}
&\frac{d\left(S_h \times \varphi^{-1}\right)_* (\mathbf{W}_0 \otimes \mathrm{Leb})}{d(\mathbf{W}_0 \otimes \mathrm{Leb})}(x,\alpha) \\
&\qquad = \varphi'(0)\int_0^1 e^{\tilde{h}(t)}\, dt \cdot \exp\left\{\int_0^1 \left[\tilde{h}''\left(\alpha + \frac{\int_0^t e^x}{\int_0^1 e^x}\right) - \frac{1}{2}\left(\tilde{h}'\left(\alpha + \frac{\int_0^t e^x}{\int_0^1 e^x}\right)\right)^2\right] \frac{e^{2x(t)}}{\left(\int_0^1 e^x\right)^2} dt\right\}.
\end{aligned} \tag{I.18}$$

La traduction de la proposition précédente sur $\mathrm{Diff}^1_+(\mathbf{S}^1)$ est exactement l'énoncé du théorème 2.1 : si $\tilde{h}(t) = \log\varphi'(t) - \log\varphi'(0)$, l'expression à droite dans (I.18) devient

$$\varphi'(0)\int_0^1 \frac{\varphi'(t)}{\varphi'(0)}\, dt \cdot \exp\left\{\int_0^1 \left[D^2 \log\varphi'(f(t)) - \frac{1}{2}\left[D\log\varphi'(f(t))\right]^2\right] f'(t)^2\, dt\right\}$$

et puisque $\mathcal{S}_\varphi = D^2\log\varphi' - \frac{1}{2}\left(D\log\varphi'\right)^2$, on obtient finalement l'expression (I.8)

$$\frac{d(L_\varphi)_*\mu_\sigma}{d\mu_\sigma}(f) = \exp\left\{\frac{1}{\sigma}\int_0^1 \left[\frac{\varphi'''(f(t))}{\varphi'(f(t))} - \frac{3}{2}\left(\frac{\varphi''(f(t))}{\varphi'(f(t))}\right)^2\right](f'(t))^2 dt\right\}.$$

3 Cohomologie essentielle de groupes

Étant donné un groupe G, il est possible de définir sa cohomologie $H^*(G)$. Quand le groupe est *topologique*, on définit sa cohomologie *continue* $H^*_c(G)$ (et différentiable quand le groupe est différentiable). Si le groupe porte une mesure, on définit sa *cohomologie essentielle*. Les groupes de cohomologie sont un outil algébrique assez

puissant pour distinguer les groupes et leurs actions. Dans ce contexte mesurable, beaucoup de progrès ont été fait depuis le célèbre travail de Cheeger et Gromov [Che-Gro], qui étudie la *cohomologie* L^2 pour les groupes. Les définitions qui suivront ont été adaptés à nos besoins, notamment la compréhension d'actions mesurables de groupes qui ne préservent pas une mesure de probabilité, mais une certain classe seulement.

Pour une introduction générale à la cohomologie des groupes, bien au-delà de ce dont nous avons besoin ici, le lecteur pourra trouver des références classiques (nous recommandons, par exemple, l'œuvre de Brown [Bro]).

3.1 Définition

Nous partons d'un groupe mesuré (G, μ), avec μ *mesure de probabilité*. Soit $\overline{G}$ le sous-groupe de G constitué des éléments qui préservent la classe de μ quand ils agissent, disons, à droite.

Si (X, μ) est un espace mesuré, on note $\mathcal{M}(X)$ l'espace des fonctions mesurables à valeurs réelles, modulo la relation d'égalité presque partout.

Nous définissons le complexe de cochaînes essentiellement homogènes $\mathcal{M}C^*(G)$ par

$$\mathcal{M}C^k(G) = \left\{c \in \mathcal{M}(G^{k+1}) \,:\, c(g_0\gamma, g_1\gamma, \ldots, g_k\gamma) = c(g_0, g_1, \ldots, g_k) \text{ p.p.}, \forall \gamma \in \overline{G}\right\};$$

on définit la différentielle $d = d_k : \mathcal{M}C^k(G) \longrightarrow \mathcal{M}C^{k+1}(G)$ par

$$dc(g_0, \ldots, g_k) = \sum_{i=0}^{k} (-1)^i c(g_0, \ldots, \widehat{g_i}, \ldots g_k).$$

On vérifie facilement que dc est essentiellement homogène.

Nous avons par suite une *cohomologie essentielle* de G, que l'on notera $\mathcal{M}H^*(G)$, dont le k-ième groupe est défini par le quotient

$$\ker(d_k)\big/\mathrm{im}(d_{k-1}) = \mathcal{M}Z^k(G)\big/\mathcal{M}B^k(G).$$

Exemple 3.1. Soit G un groupe quelconque et μ sa mesure de comptage. La cohomologie essentielle coïncide, dans ce cas, avec la cohomologie habituelle de G.

Plus généralement, soit G un groupe localement compact et μ sa mesure de Haar. Le groupe $\overline{G}$ est le groupe G entier. La cohomologie essentielle équivaut alors à la cohomologie borélienne $H^*_B(G)$.

3.2 Cohomologie essentielle d'une action

Remarquons que la définition précédente peut se généraliser aux actions de groupes. Soit (X, μ) un espace mesuré et G un groupe qui agit à droite sur X en préservant la classe de μ. On définit le complexe de cochaînes $\mathcal{M}C^*(X, G)$ en définissant les k-cochaînes

$$\mathcal{M}C^k(X, G) = \{c \in \mathcal{M}(X^{k+1}) \,:\, c(x_0 g, x_1 g, \ldots, x_k g) = c(x_0, x_1, \ldots, x_k) \text{ p.p.}, \forall g \in G\};$$

on définit la différentielle $d = d_k : \mathcal{M}C^k(X, G) \longrightarrow \mathcal{M}C^{k+1}(X, G)$ par

$$dc(x_0, \ldots, x_k) = \sum_{i=0}^{k} (-1)^i c(x_0, \ldots, \widehat{x_i}, \ldots x_k).$$

Nous avons par suite une *cohomologie essentielle* de l'action de G sur X, que l'on notera $\mathcal{M}H^*(X, G)$, dont le k-ième groupe est défini par le quotient

$$\ker(d_k)\big/\mathrm{im}(d_{k-1}) = \mathcal{M}Z^k(X, G)\big/\mathcal{M}B^k(X, G).$$

3.3 Cohomologie L^p (réduite)

Les définitions précédentes se prêtent mal à l'étude analytique : les groupes de cohomologie pouvant être de dimension infinie, il faut y rajouter une structure banachique. Dans les situations plus « classiques », ces groupes de cohomologie ont été introduits dans [Gro, chapitre 8].

Soit $1 \leq p < \infty$, la différentielle d préserve les sous-complexes des cochaînes L^p (car μ est une mesure de probabilité). Ce sous-complexe est en plus *fermé* pour la topologie L^p.

Nous pourrions définir ainsi la cohomologie L^p, que l'on notera $L^pH^*(X, G)$, dont le k-ième groupe est donné par

$$\left(\ker(d_k) \cap L^p(X^{k+1}, \mu^{\otimes k+1})\right)\Big/\left(\mathrm{im}(d_{k-1}) \cap L^p(X^{k+1}, \mu^{\otimes k+1})\right) = L^pZ^k(X, G)\big/L^pB^k(X, G).$$

Cependant ces quotients ne suffisent pas en général : les sous-espaces $\mathrm{im}(d_{k-1}) \cap L^p(X^{k+1}, \mu^{\otimes k+1})$ ne sont pas toujours fermés dans la topologie L^p. Pour cette raison il est naturel d'introduire la *cohomologie L^p réduite* $\overline{L^pH}^*(X,G)$, dont le k-ième groupe est donné par le quotient

$$\left(\ker(d_k) \cap L^p(X^{k+1}, \mu^{\otimes k+1})\right) \Big/ \left(\overline{\mathrm{im}(d_{k-1}) \cap L^p(X^{k+1}, \mu^{\otimes k+1})}^{L^p}\right) = L^pZ^k(X,G) \big/ \overline{L^pB}^k(X,G).$$

La norme L^p quotient rend alors $\overline{L^pH}^k$ un espace de Banach.

Remarque 3.1. Lorsque la mesure de probabilité μ est *préservée* par l'action de G, on peut définir le dual de $\overline{L^pH}^*(X,G)$ de manière très naturelle. Dans le contexte plus général d'une action qui *quasi-préserve* la mesure μ, on rencontre quelque problème : en effet, l'opérateur linéaire qu'à tout cocycle $c(x_0, \dots, x_k)$ associe le cocycle translaté $c(x_0g, \dots, x_kg)$ ne peut pas toujours être défini sur $L^p(X^{k+1}, \mu^{\otimes k+1})$ entier, car on nécessite quelque condition supplémentaire sur la dérivée de Radon-Nykodim $\frac{dg_*\mu}{d\mu}$ (notamment $\frac{dg_*\mu}{d\mu} \in L^\infty(G,\mu)$, condition qui n'est pas vérifiée dans l'exemple que l'on illustrera dans la suite).

3.4 Notre exemple : les difféomorphismes du cercle

Grâce à la mesure de Malliavin-Shavgulidze quasi-invariante à droite on peut définir la cohomologie essentielle de $\mathrm{Diff}^1_+(\mathbf{S}^1)$, par rapport au sous-groupe $\mathrm{Diff}^3_+(\mathbf{S}^1)$.

La cohomologie du groupe des difféomorphismes du cercle a été beaucoup étudiée dans le passé, surtout à cause de ses liens avec la théorie des feuilletages : d'après les travaux de Haefliger [Hae 1] (voir aussi Thurston [Thu] et Mather [Mat 1, Mat 2]) elle permet de comprendre, par exemple, le groupe de cobordisme des variétés feuilletées par des feuilletages C^r de codimension 1. C'est notamment le deuxième groupe de cohomologie de $\mathrm{Diff}^r_c(\mathbf{S}^1)$ à coefficients entiers qui joue l'un des rôles les plus intéressants (pour une introduction, voir [Ghy 1]) : la découverte de la classe de Godbillon-Vey dans les années 1970 a été fondamentale à ce propos. Cependant, la compréhension géométrique du cocycle de Godbillon-Vey n'est pas entièrement accomplie. Dans ce qui suit, nous essayons de définir une variante *stochastique* de ce cocycle, en exploitant la quasi-invariance de la mesure de Malliavin-Shavgulidze pour avoir le bon cadre algébrique.

3.4.1 Le cocycle de Godbillon-Vey

Si f, g et h sont trois difféomorphismes C^2, le cocycle de Godbillon-Vey est défini par

$$gv(f,g,h) = \int_{S^1} \begin{vmatrix} 1 & \log Df & d\log Df \\ 1 & \log Dg & d\log Dg \\ 1 & \log Dh & d\log Dh \end{vmatrix}.$$

Nous avons bien défini une cochaîne car

$$\begin{aligned} gv(f\varphi, g\varphi, h\varphi) &= \int_{S^1} \begin{vmatrix} 1 & \log Df\varphi + \log D\varphi & d(\log Df\varphi) + d\log D\varphi \\ 1 & \log Dg\varphi + \log D\varphi & d(\log Dg\varphi) + d\log D\varphi \\ 1 & \log Dh\varphi + \log D\varphi & d(\log Dh\varphi) + d\log D\varphi \end{vmatrix} \\ &= \int_{S^1} \begin{vmatrix} 1 & \log Df\varphi & d(\log Df\varphi) \\ 1 & \log Dg\varphi & d(\log Dg\varphi) \\ 1 & \log Dh\varphi & d(\log Dh\varphi) \end{vmatrix} = gv(f,g,h). \end{aligned}$$

Ceci est un 2-cocycle continu sur le groupe $\mathrm{Diff}^2_+(\mathbf{S}^1)$ car sa différentielle est

$$dgv(f,g,h,k) = \int_{S^1} \begin{vmatrix} 1 & 1 & \log Df & d\log Df \\ 1 & 1 & \log Dg & d\log Dg \\ 1 & 1 & \log Dh & d\log Dh \\ 1 & 1 & \log Dk & d\log Dk \end{vmatrix} = 0.$$

Soit $1 \le p < \infty$. Il est possible de donner un sens au cocyle gv dans la cohomologie $\overline{L^pH}^*(\mathrm{Diff}^1_+(\mathbf{S}^1))$: si f^{-1}, g^{-1}, h^{-1} sont donnés par intégration des ponts browniens indépendants B_1, B_2, B_3 (et des rotations aléatoires que nous n'allons pas expliciter), nous avons

$$gv(f,g,h) = \int_{S^1} \begin{vmatrix} 1 & -B_1f + \log\int_0^1 e^{B_1} & -d(B_1f) \\ 1 & -B_2g + \log\int_0^1 e^{B_2} & -d(B_2g) \\ 1 & -B_3h + \log\int_0^1 e^{B_3} & -d(B_3h) \end{vmatrix} = \int_{S^1} \begin{vmatrix} 1 & B_1f & d(B_1f) \\ 1 & B_2g & d(B_2g) \\ 1 & B_3h & d(B_3h) \end{vmatrix}.$$

Assurons-nous que cette intégrale est bien définie : prenons par exemple le terme $\int_{S^1} B_2 g\, d(B_1 f)$, ceci on pourra l'écrire sous la forme $\int_{S^1} B_2 g f^{-1}\, dB_1$. Afin que cette intégrale soit bien définie, il est suffisant de vérifier que $f^{-1}(t)$ est adapté par rapport à la filtration standard du pont brownien :

$$f^{-1}(t) = \frac{\int_0^t e^B}{\int_0^1 e^B} + \lambda$$

appartient évidemment à la filtration standard car pour définir $f^{-1}(t)$ on intègre une fonction qui dépend des valeurs du pont brownien jusqu'au temps t.

Théorème 3.1. *Pour tout $1 \leq p < \infty$, le cocycle de Godbillon-Vey s'étend à une 2-cochaîne essentielle de classe L^p.*

Dimostrazione. Nous venons de voir que gv est défini presque partout sur $\mathrm{Diff}^1_+(\mathbf{S}^1) \times \mathrm{Diff}^1_+(\mathbf{S}^1) \times \mathrm{Diff}^1_+(\mathbf{S}^1)$. Montrons qu'il s'agit d'une fonction L^p. Il suffit de montrer que l'espérance

$$\mathbf{E}_{w_0 \otimes w_0}\left[\left|\int B_1\, dB_2\right|^p\right]$$

est finie : on évaluera gv sur f^{-1}, g^{-1} et h^{-1}.

On se limitera ici à la preuve pour $p = 2$:

$$\mathbf{E}_{w_0 \otimes w_0}\left[\left|\int B_1\, dB_2\right|^2\right] = \mathbf{E}_{w_0}\left[\int B_1^2\right] < \infty.$$

Dans l'égalité on a utilisé le fait que pour une fonction f fixée $\int f\, dB_2$ est une variable gaussienne de variance $\int f^2$. Le fait que l'espérance de la norme L^2 du pont brownien est finie est un résultat classique (qui remonte à Smirnov [Smi]). Pour en savoir plus on pourra consulter le travail de Tolmatz [Tol].

Lorsque p est différent de 2, on peut de la même manière montrer que la norme $L^{\lfloor p \rfloor + 1}$ est finie.

Montrons que gv est essentiellement homogène :

$$\begin{aligned} gv(f\varphi, g\varphi, h\varphi) &= \int_{S^1} \begin{vmatrix} 1 & B_1 f\varphi + \log D\varphi & d(B_1 f\varphi) + d\log D\varphi \\ 1 & B_2 g\varphi + \log D\varphi & d(B_2 g\varphi) + d\log D\varphi \\ 1 & B_3 h\varphi + \log D\varphi & d(B_3 h\varphi) + d\log D\varphi \end{vmatrix} \\ &= \int_{S^1} \begin{vmatrix} 1 & B_1 f\varphi & d(B_1 f\varphi) \\ 1 & B_2 g\varphi & d(B_2 g\varphi) \\ 1 & B_3 h\varphi & d(B_3 h\varphi) \end{vmatrix} = gv(f, g, h). \end{aligned}$$

Il reste à vérifier que gv définit un 2-cocycle, c'est-à-dire $dgv = 0$ presque partout. Ceci peut être vérifié comme dans le cas classique. □

Remarque 3.2. En suivant Thurston [Bot], Hurder et Katok [Hur-Kat] interprètent le cocycle de Godbillon-Vey comme la fonction qui donne la somme algébrique des aires entourées par les trois courbes décrites dans le plan par $(\log g', \log h')$, $(\log f', \log g')$ et $(\log f', \log h')$.

Dans [Hur-Kat], on définit le cocycle de Godbillon-Vey en régularité $C^{1+\alpha}$, pour $\alpha > 1/2$, puisqu'elle s'avère être la régularité minimale nécessaire pour qu'une courbe possède une aire nulle (le premier résultat vers cette direction est dû à Duminy et Sergiescu [Dum-Ser]). Ensuite Tsuboi demontre [Tsu 1, Tsu 2] qu'une définition classique n'est pas possible en régularité $C^{1+\alpha}$ avec $\alpha < 1/2$.

Cette interprétation s'étend par contre dans ce contexte stochastique : en citant Paul Lévy [Lév],

> *« Pour qu'une aire soit remplie sans que l'oscillation brownienne soit infinie, il faut une exploration méthodique que le hasard ne peut réaliser. »*

Le cocycle de Godbillon-Vey sur $\mathrm{Diff}^1_+(\mathbf{S}^1)$ est donc la somme des aires *stochastiques* au sens de Lévy. En particulier, notre *gv stochastique* n'est pas presque sûrement constant.

3.4.2 Quelques comparaisons avec les résultats classiques

La question qui reste ouverte est savoir si gv est un cobord essentiel ou non.

Remarque 3.3. Quand le groupe G agit sur X en préservant une mesure de probabilité μ, tout cocycle c est un cobord db.

En effet, il suffit de définir $b(x_0, \dots, x_{k-1}) = \int_X c(x, x_0, \dots, x_{k-1})\, d\mu(x)$.

Dans le cas de Godbillon-Vey, nous pouvons définir une fonction b comme nous venons de le faire :

$$b(f,g) = \int_{\mathrm{Diff}^1_+(\mathbf{S}^1)} gv(x,f,g)\, d\mu(x).$$

On trouve alors $db = gv$. Godbillon-Vey est alors un cobord ? Bien sur que non : il serait alors un cobord dans la cohomologie continue aussi. En fait, b n'est pas une cochaîne :

$$b(f\varphi, g\varphi) = \int_{\mathrm{Diff}^1_+(\mathbf{S}^1)} gv(x,f,g) \frac{d\varphi_*^{-1}\mu}{d\mu}(x)\, d\mu(x).$$

Rassurant !

Pour montrer que gv n'est pas trivial, nous proposons une approche ergodique : en effet, on peut rajouter quelque considération dans notre étude de la cohomologie essentielle. Le degré 0 du complexe est étudié par des considérations *ergodiques* sur l'action de $\mathrm{Diff}^3_+(\mathbf{S}^1)$ sur $\mathrm{Diff}^1_+(\mathbf{S}^1)$.

Les 0-cochaînes sont les fonctions mesurables essentiellement constantes sur les orbites. Le théorème 2.4 de Kosyak montre que l'action est ergodique. On en déduit que les 0-cochaînes sont les fonctions essentiellement constantes. Donc, toute 0-cochaîne est un 0-cocycle. Par suite, le seul 1-cobord est la fonction essentiellement nulle.

Que peut-on dire sur les 1-cocycles ? Rappelons que d'après les résultats de J. Mather [Mat 1, Mat 2], le groupe de difféomorphismes $\mathrm{Diff}^r_+(\mathbf{S}^1)$ est simple pour tout $r \neq 2$. Il s'ensuit que son premier groupe de cohomologie (correspondant à l'abélianisé) est trivial.

Remarque 3.4. Il en est différemment pour le groupe $\mathrm{Diff}^{1+vb}_+(\mathbf{S}^1)$, comme Mather a remarqué brillamment dans [Mat 3] : si $\log Df$ est une fonction à variation bornée, on dispose de la décomposition de Radon-Nikodym de la *mesure (signée)* $d\log Df$:

$$d\log Df = \mu_{reg} + \mu_{sing},$$

avec μ_{reg} absolument continue par rapport à Leb et μ_{sing} étrangère. L'application qui à $f \in \mathrm{Diff}^{1+vb}_+(\mathbf{S}^1)$ associe la valeur $\mu_{reg}(\mathbf{S}^1)$ définit un homomorphisme vers $\mathbf{R}$ qui n'est pas trivial.

Cette construction nous donne un "morphisme" aussi dans notre contexte mesurable : si l'on pose

$$\pi(f,g) = \int_{\mathbf{S}^1} d(B_2 g) - \int_{\mathbf{S}^1} d(B_1 f), \tag{I.19}$$

avec B_1 et B_2 deux ponts browniens indépendants qui définissent f^{-1} et g^{-1}, on trouve alors que π est un 1-cocycle essentiel, mais *nul* presque partout.

L'étape suivante consiste à comprendre l'espace des 2-cobords. Ceci est plus dur, car il n'est pas réduit à 0 : soit $b = b(f,g)$ une 2-cochaîne essentielle qui ne dépend que du produit fg^{-1} – à savoir, il existe une fonction Φ telle que

$$b(f,g) = \Phi(fg^{-1}).$$

En prenant le cobord on trouve :

$$db(f,g,h) = \Phi(gh^{-1}) - \Phi(fh^{-1}) + \Phi(fg^{-1}). \tag{I.20}$$

Il est probable qu'en utilisant intelligemment la méthode de preuve du théorème de Kosyak, on puisse dire que tous les 2-cobords essentiels sont de la forme (I.20).

Pour compléter cette partie, il faut rappeler qu'il existe un autre 2-cocycle qui est non-trivial dans le cas classique, à savoir le *cocycle d'Euler*. Sa définition s'étend dans notre contexte. Si f, g et h sont trois difféomorphismes du cercle, on peut regarder l'ordre cyclique des points $f(0)$, $g(0)$ et $h(0)$ sur le cercle S^1. Supposons que ces trois points soient distincts, on assigne alors la valeur $+1$ à (f,g,h) si les trois points sont cycliquement positivement ordonnés, -1 sinon (dans notre contexte mesurable, on n'a pas besoin de définir la valeur du cocycle d'Euler lorsque deux points coïncident, qui représente un événement de mesure nulle). Nous notons cette valeur $eu(f,g,h)$. On peut vérifier facilement que eu est un cocycle essentiel, plus précisément L^∞.

Capitolo II

Décrire la dynamique aléatoire

1 Points périodiques et nombre de rotation aléatoire

Pour comprendre la dynamique décrite par un difféomorphisme, le premier pas consiste à obtenir le plus d'informations possibles sur les orbites. Dans le cas d'une dynamique en dimension 1, nous avons pu voir dans l'introduction (voir 3.1) que la connaissance des orbites périodiques permet une bonne description des phénomènes globaux. Il s'avère que si l'on considère des difféomorphismes aléatoires par rapport aux mesures de Malliavin-Shavgulidze, les orbites périodiques sont de nature assez contrôlée.

1.1 Difféomorphismes de l'intervalle

Puisque nos difféomorphismes préservent l'orientation, tout point périodique sur l'intervalle est nécessairement un point fixe. Nous voudrions donc décrire l'ensemble des points fixes d'un difféomorphisme aléatoire de $[0,1]$: est-il fini ou bien infini ? Pour nous donner une réponse exhaustive, nous étudions le *caractère* des points fixes : un point t tel que $f(t) = t$ est *parabolique* si $f'(t) = 1$, et *hyperbolique* dans le cas contraire. Pour t point fixe hyperbolique, on dira que t est *attractif* si $f'(t) < 1$ et *répulsif* autrement. Les difféomorphismes qui possèdent uniquement des points fixes hyperboliques ne sont pas trop compliqués. Le théorème suivant affirme que l'on peut négliger les phénomènes paraboliques pour les difféomorphismes de Malliavin-Shavgulidze (nous entendons par ceci qu'une certaine propriété est vérifiée par un ensemble de difféomorphismes dans $(\mathrm{Diff}^1_+(I), \nu_{MS})$ de mesure 1).

Théorème 1.1. *Pour tout $\lambda > 0$, la probabilité que le difféomorphisme f possède un point fixe avec multiplicateur λ est nulle :*

$$\nu_{MS}(\exists t \,:\, f(t) = t,\ f'(t) = \lambda) = 0.$$

En particulier un difféomorphisme de Malliavin-Shavgulidze ne possède pas de points fixes paraboliques.

Nous pouvons en déduire un premier corollaire :

Corollaire 1.2. *Un difféomorphisme de Malliavin-Shavgulidze possède un nombre fini de points fixes.*

Dimostrazione. Il suffit de montrer que tout point fixe est isolé. Si $\{t_k\}$ est une suite de points fixes qui s'accumule sur un point fixe t, par continuité de la dérivée et d'après le théorème des valeurs intermédiaire, on doit avoir $f'(t) = 1$. Donc t est parabolique. □

La preuve du théorème 1.1 utilise fortement les propriétés du mouvement brownien et la définition des difféomorphismes de Malliavin-Shavgulidze. Nous avons recherché une démonstration qui pourtant utilise le moins possible les expressions explicites des densités des variables considérées : il est assez probable que les résultats présentés dans [Mat-Hir-Yor, Man-Yor, Yor] puissent permettre de déduire l'énoncé du théorème, mais une telle preuve serait extrêmement technique.

1.1.1 Cadre et notations

Pour cette partie, nous ne précisons pas l'espace de probabilité $(\Omega, \mathcal{B}, \mathbf{P})$ qui sera arbitraire. Pour tout mouvement brownien standard $(B(t))_{t\in[0,1]}$ sur $(\Omega, \mathcal{B}, \mathbf{P})$ on définit le difféomorphisme de $[0,1]$ dans lui-même

$$f \,:\, t \longmapsto \frac{\int_0^t e^{B(s)}\,ds}{\int_0^1 e^{B(s)}\,ds} \tag{II.1}$$

qui a donc la même loi qu'un difféomorphisme de Malliavin-Shavgulidze.

La définition de f nous oblige à considérer une filtration plus grande que $\mathcal{F}_t = \sigma((B(s))_{s\leq t})$, la filtration naturelle associée à B : pour définir f au temps t, il faut connaître l'évolution de B entre 0 et t et la moyenne totale $\int_0^1 e^{B(s)}\,ds$. Il s'ensuit que si l'on définit la filtration régulière

$$\mathcal{G}_t = \bigcap_{t'>t}\left(\mathcal{F}_{t'} \vee \sigma\left(\int_0^1 e^{B(s)}\,ds\right)\right), \tag{II.2}$$

alors f définit un processus sur $(\Omega, \mathcal{B}, \mathbf{P}, \mathcal{G})$.

1.1.2 Quelques lemmes préliminaires

Invariance par inversion du temps

Il est bien connu que si B est un mouvement brownien sur $(\Omega, \mathcal{B}, \mathbf{P})$, alors le processus « inversé »

$$\overline{B}(s) := B(1-s) - B(1)$$

défini encore un mouvement brownien sur $[0,1]$, mais qui est un processus par rapport à la tribu inversée $\overline{\mathcal{F}}_t = \sigma((B(s))_{1-t\leq s\leq 1})$. On étend cette remarque aux difféomorphismes de Malliavin-Shavgulidze :

Lemme 1.3. *Si f est un difféomorphisme de Malliavin-Shavgulidze défini par l'expression (II.1), le difféomorphisme aléatoire*

$$\overline{f}(t) = 1 - \frac{\int_0^{1-t} e^{B(s)}\,ds}{\int_0^1 e^{B(s)}\,ds},$$

a la même loi que f.

Dimostrazione. En effet soit $t \in [0,1]$, on a

$$\begin{aligned}
\log \overline{f}'(t) &= B(1-t) - \log \int_0^1 e^{B(s)} ds \\
&= B(1-t) - B(1) - \log e^{-B(1)} \int_0^1 e^{B(s)} ds \\
&= \overline{B}(t) - \log \int_0^1 e^{B(s)-B(1)} ds \\
&= \overline{B}(t) - \log\left(-\int_1^0 e^{B(1-u)-B(1)} du\right) \\
&= \overline{B}(t) - \log \int_0^1 e^{\overline{B}(u)} du.
\end{aligned}$$

Comme $\overline{B}$ est encore un mouvement brownien, la preuve du lemme est terminée. □

Si l'on définit la nouvelle filtration régulière

$$\overline{\mathcal{G}}_t = \bigcap_{t'>t}\left(\overline{\mathcal{F}}_{t'} \vee \sigma\left(\int_0^1 e^{\overline{B}(u)}\,du\right)\right),$$

alors $\overline{f}$ définit un processus sur $(\Omega, \mathcal{B}, \mathbf{P}, \overline{\mathcal{G}})$.

Regarder le difféomorphisme dans un court intervalle de temps

Proposition 1.4. *Soit $\varepsilon \in]0,1[$ fixé. La probabilité qu'un difféomorphisme de Malliavin-Shavgulidze possède au moins un point fixe sur $]0,\varepsilon]$ tend vers 0 avec ε.*

Dimostrazione. Si $f \in C^1$ possède au moins un point fixe sur $]0,\varepsilon]$, le théorème des accroissements finis assure l'existence d'un point $t \in]0,\varepsilon]$ auquel $f'(t) = 1$.

Si f est défini par (II.1), la condition $f'(t) = 1$ s'écrit

$$B(t) = \log \int_0^1 e^{B(s)} ds,$$

donc par continuité höldérienne, pour tout $\alpha < 1/2$ on a

$$\left|\log f'(t) - \log f'(0)\right| = |B(t)| = \left|\log \int_0^1 e^{B(s)} ds\right| \leq C(B)\varepsilon^\alpha,$$

avec

$$C(B) \leq \sup_{x \neq y \in [0,1/2]} \frac{|B(x) - B(y)|}{|x-y|^\alpha}.$$

On obtient

$$\varepsilon \geq C(B)^{-1/\alpha} \left|\log \int_0^1 e^{B(s)} ds\right|^{1/\alpha},$$

événement qui a une probabilité qui tend vers 0 lorsque ε tend vers 0. □

Regarder le difféomorphisme après un court intervalle de temps
Soit $\varepsilon > 0$ donné, alors le processus $\widetilde{B}$ défini sur $[0, 1-\varepsilon]$ par

$$\widetilde{B}(t) := B(t+\varepsilon) - B(\varepsilon)$$

est encore un mouvement brownien, par rapport à la filtration $\widetilde{\mathcal{F}}_t = \sigma\left((\widetilde{B}(s))_{s<t}, B(\varepsilon)\right)$. Si f est un difféomorphisme de Malliavin-Shavgulidze, on veut relier sa restriction à l'intervalle $[\varepsilon, 1]$, à $\widetilde{B}$. Soit $\widetilde{f}(t) := f(t+\varepsilon)$, nous avons

$$\begin{aligned}
\log \widetilde{f}'(t) &= \log f'(t+\varepsilon) \\
&= B(t+\varepsilon) - \log \int_0^1 e^{B(s)} ds \\
&= B(t+\varepsilon) - B(\varepsilon) - \log \int_0^1 e^{B(s)-B(\varepsilon)} ds \\
&= \widetilde{B}(t) - \log\left(\int_\varepsilon^1 e^{B(s)-B(\varepsilon)} ds + e^{-B(\varepsilon)} \int_0^\varepsilon e^{B(s)} ds\right) \\
&= \widetilde{B}(t) - \log\left(\int_0^{1-\varepsilon} e^{\widetilde{B}(u)} du + e^{-B(\varepsilon)} \int_0^\varepsilon e^{B(s)} ds\right).
\end{aligned}$$

On en déduit que $\widetilde{f}$ définit un processus par rapport à la filtration régulière

$$\widetilde{\mathcal{G}}_t = \bigcap_{t'>t} \left(\widetilde{\mathcal{F}}_{t'} \vee \sigma\left(\int_0^{1-\varepsilon} e^{\widetilde{B}(u)} du, \int_0^\varepsilon e^{B(s)} ds\right)\right).$$

Les calculs précédents donnent facilement le lemme suivant :

Lemme 1.5. *Soient $C > c$ et $b \in \mathbf{R}$. Sur l'événement*

$$A(C, c, b) := \left\{B(\varepsilon) = b, \log \int_0^1 e^{B(s)} ds = C, \log \int_0^\varepsilon e^{B(s)} ds = c\right\}, \tag{II.3}$$

le processus $(f(t))_{t\in[\varepsilon,1]}$ suit la même loi que le processus $\left(\tilde{f}(t)\right)_{t\in[0,1-\varepsilon]}$ conditionné à l'événement

$$\widetilde{A}(C, c, b) := \left\{\int_0^{1-\varepsilon} e^{\widetilde{B}(s)} ds = e^{-b}\left(e^C - e^c\right)\right\}.$$

1.1.3 Preuve du théorème 1.1

La preuve suivante s'inspire de [Mör-Per, corollaire 2.26]. Notons E l'événement

$$E = \{\exists\, t \,:\, f(t) = t,\ f'(t) = \lambda\}$$

Pour tout $\varepsilon \in]0, 1[$, en définissant les événements

$$E_\varepsilon := \{\exists\, t \in]0, \varepsilon] \text{ t.q. } f(t) = t\}, \quad L_\varepsilon := \{\exists\, t \in [\varepsilon, 1] \text{ t.q. } f'(t) = \lambda, f'(f(t)) = \lambda\}$$

on a $E \subset E_\varepsilon \cup L_\varepsilon$. La proposition 1.4 assure que la probabilité de E_ε tend vers 0 lorsque ε tend vers 0. Il nous faudra montrer que la probabilité de L_ε est *nulle* pour tout $\varepsilon > 0$.

D'après (II.1),

$$\log f'(t) = B(t) - \log \int_0^1 e^{B(s)}\, ds,$$

et nous voulons démontrer que le processus planaire

$$Z(t) := \left(B(t) - \log \int_0^1 e^{B(s)}\, ds, B(f(t)) - \log \int_0^1 e^{B(s)}\, ds \right)$$

ne passe pas par le point de la diagonale $(\log \lambda, \log \lambda)$: l'événement L_ε est égal à l'événement

$$\{\exists\, t \in [\varepsilon, 1] \,:\, Z(t) = (\log \lambda, \log \lambda)\}.$$

On désintègre la mesure $\mathbf{P}$ par rapport aux événements $A(C, c, b)$ définis à la ligne (II.3) :

$$\mathbf{P}(L_\varepsilon) = \mathbf{E}\left[\mathbf{P}_{A(C,c,b)}\left(L_\varepsilon\right)\right]$$

en notant $\mathbf{P}_{A(C,c,b)}$ la mesure de probabilité conditionnelle pour l'événement $A(C, c, b)$. En utilisant le lemme 1.5, on voit que la probabilité $\mathbf{P}_{A(C,c,b)}\left(L_\varepsilon\right)$ est égale à la probabilité

$$\mathbf{P}_{\widetilde{A}(C,c,b)}\left(\widetilde{L}_\varepsilon\right), \tag{II.4}$$

avec $\widetilde{L}_\varepsilon$ défini par

$$\widetilde{L}_\varepsilon := \left\{\exists\, t \in [0, 1-\varepsilon] \,:\, (\widetilde{B}(t) + b - C, \widetilde{B}(\widetilde{f}(t) - \varepsilon) + b - C) = (\log \lambda, \log \lambda)\right\}.$$

Ce décalage de temps est une astuce, assez classique en probabilité, qui nous permettra d'échanger les rôles de $b - C$ et $\log \lambda$. On va démontrer que pour presque tous $C > c$ et $b \in \mathbf{R}$ cette probabilité est nulle.

D'après l'invariance par translations du mouvement brownien, un mouvement brownien qui part de $b - C$ et arrive en $\log \lambda$ possède la même loi qu'un mouvement brownien partant de $-\log \lambda$ et arrivant en $-b + C$ (on opère une translation par $-b + C - \log \lambda$). Par conséquent la probabilité (II.4) est égale à

$$\mathbf{P}_{\widetilde{A}(C,c,b)}\left(\exists\, t \in [0, 1-\varepsilon] \,:\, \left(\widetilde{B}(t) - \log \lambda, \widetilde{B}(\widetilde{f}(t) - \varepsilon) - \log \lambda\right) = (-b + C, -b + C)\right), \tag{II.5}$$

Fixons C, or

$$\mathbf{P}_{\widetilde{A}(C,c,b)}\left(\exists\, t \in [0, 1-\varepsilon] \,:\, \left(\widetilde{B}(t) - \log \lambda, \widetilde{B}(\widetilde{f}(t) - \varepsilon) - \log \lambda\right) = (-b + C, -b + C)\right)$$

est nul pour presque tout $x = -b + C$ si et seulement si son intégrale par rapport à $x = -b + C$ est nulle, et celle-ci est égale, au signe près, à

$$\int_{\mathbf{R}} \mathbf{P}_{\widetilde{A}(C,c,b)}\left(\exists\, t \in [0, 1-\varepsilon] \,:\, \left(\widetilde{B}(t) - \log \lambda, \widetilde{B}(\widetilde{f}(t) - \varepsilon) - \log \lambda\right) = (x, x)\right) dx. \tag{II.6}$$

et (II.6) est égal, par Fubini, à

$$\mathbf{E}_{\widetilde{A}(C,c,b)}\left[\mathrm{Leb}_\Delta\left(Im\left(\widetilde{B}(t) - \log \lambda, \widetilde{B}(\widetilde{f}(t) - \varepsilon) - \log \lambda\right) \cap \Delta\right)\right], \tag{II.7}$$

où Leb_Δ dénote la mesure de Lebesgue sur la diagonale $\Delta \subset \mathbf{R}^2$. L'espérance (II.7) est inférieure ou égale à

$$\mathbf{E}_{A(C,c,b)}\left[\mathrm{Leb}\left(\text{Zéros}\left(B(f(t)) - B(t)\right) \cap [\varepsilon, 1]\right)\right],$$

en notant $\text{Zéros}\left(B(f(t)) - B(t)\right)$ l'ensemble des t tels que $B(f(t)) = B(t)$.

On simplifie le problème encore une fois grâce au théorème de Fubini : on peut écrire

$$\begin{aligned} &\mathbf{E}_{A(C,c,b)}\left[\mathrm{Leb}\left(\text{Zéros}\left(B(f(t)) - B(t)\right) \cap [\varepsilon, 1]\right)\right] \\ &= \int_\varepsilon^1 \mathbf{P}_{A(C,c,b)}\left(B(f(t)) = B(t)\right) dt \end{aligned}$$

et il nous suffit de voir que pour presque tous $t \in [\varepsilon, 1]$, $C > c$ et $b \in \mathbf{R}$,

$$\mathbf{P}_{A(C,c,b)}\left(B(f(t)) = B(t)\right) = 0. \tag{II.8}$$

Or, la probabilité que $f(t) = t$ étant nulle, on va montrer que presque sûrement sur $C > c$, b, on a

$$\mathbf{P}_{A(C,c,b)}\left(\{B(f(t)) = B(t)\} \cap \{f(t) > t\}\right) = 0, \tag{II.9}$$

puis le fait que la probabilité

$$\mathbf{P}_{\{B(\varepsilon)=b, \log\int_0^\varepsilon e^B = c, \log\int_0^1 e^B = C\}}\left(\{B(f(t)) = B(t)\} \cap \{f(t) < t\}\right)$$

est nulle pour presque tous $C > c$ et $b \in \mathbf{R}$, suivra par méthode de couplage (en utilisant le lemme 1.3) : si $f(t) < t$, en « renversant le temps », on obtient $\overline{f}(1-t) > 1-t$, ce qui nous permet de déduire le deuxième cas à partir du premier.

Les événement $\{f(t) > t\}$ et $\{f(t) < t\}$ appartiennent à $\mathcal{G}_t$ pour tout $t \in [0,1]$. Donc on peut effectivement se restreindre au cas $f(t) > t$: l'avantage est que B peut prendre une valeur arbitraire en $f(t)$ par propriété de Markov du mouvement brownien (on remarquera que B définit un processus qui est absolument continu par rapport à un mouvement brownien sur $[0,t]$).

En effet, pour calculer $f(t)$ on doit juste avoir connaissance de B jusqu'à l'instant $t < f(t)$. On en déduit alors l'assertion (II.9), qui permet de conclure la preuve du théorème 1.1. □

1.1.4 Sur le nombre de points fixes

On peut améliorer notre description en étudiant la probabilité qu'un difféomorphisme possède exactement m points fixes. Si l'on définit la variable aléatoire $M = \#\mathrm{Fix}(f)$, on voudrait décrire la distribution de M. On peut facilement montrer que la probabilité que M soit égal à 2 est positive (0 et 1 sont toujours des points fixes) : en effet, supposons que l'intégrale totale du mouvement brownien B sur $[0,1]$ soit presque 1 et que B croisse beaucoup près de 0, en restant presque constant après ; on trouvera ainsi un difféomorphisme qui n'a pas d'autres points fixes. Plus généralement :

Proposition 1.6. *Pour chaque entier positif $m \geq 2$, la probabilité que M soit égal à m est strictement positive.*

Dimostrazione. En effet, l'ensemble des difféomorphismes avec m points fixes hyperboliques est d'intérieur non vide dans la topologie C^1. L'affirmation suit du fait que la mesure de Malliavin-Shavgulidze est strictement positive sur tout ouvert. □

Remarque 1.1. Des simulations numériques laissent penser que la loi de M est proche d'une loi géométrique de paramètre 1/2. Cependant les statistiques changent remarquablement lorsque le paramètre σ de la variance du brownien change : par exemple, pour $\sigma = 2$ la distribution de M est encore proche d'une loi géométrique de paramètre 1/2, mais elle ne l'est plus pour des valeurs de σ plus grands que 4.

Des conséquences plus remarquables du théorème 1.1 interviendront dans la suite.

1.2 Difféomorphismes du cercle

Soit $(\Omega, \mathcal{B}, \mathbf{P})$ un espace de probabilité, soit $B = (B_t)_{t\in\mathbf{S}^1}$ un pont brownien sur Ω et $\alpha \in \mathbf{S}^1$ une variable de loi uniforme et indépendante de B. Nous définissons le difféomorphisme du cercle aléatoire de Malliavin-Shavgulidze :

$$f : t \longmapsto \frac{\int_0^t e^{B(s)}\,ds}{\int_0^1 e^{B(s)}\,ds} + \alpha.$$

Il est facile de remarquer que la loi de f est invariante lorsque l'on fait une conjugaison par une rotation.

1.2.1 Nombre de rotation rationnel

Rappelons (voir partie 3.1 dans l'introduction) qu'un difféomorphisme du cercle possède une orbite périodique si et seulement si son nombre de rotation est rationnel. Nous avons

Proposition 1.7. *Soit q un entier strictement positif. La probabilité qu'un difféomorphisme de Malliavin-Shavgulidze possède une orbite de période q est positive.*

Plus précisément, la probabilité que le nombre de rotation soit égal à p/q est positive pour tout $p/q \in \mathbf{Q}$.

Dimostrazione. On fixe p/q rationnel. L'ensemble des difféomorphismes avec nombre de rotation p/q est d'intérieur non vide, puisqu'il contient les difféomorphismes avec une orbite périodique hyperbolique, dont la combinatoire est de type p/q. On conclut encore grâce aux propriétés de la mesure. □

Comme analogue à la proposition 1.2, on peut dire que les orbites périodiques sont presque toujours hyperboliques.

Proposition 1.8. *La probabilité qu'un difféomorphisme de Malliavin-Shavgulidze possède une orbite périodique* parabolique *est nulle.*

Il en résulte que les orbites périodiques sont en nombre fini, pair, et la probabilité d'en avoir exactement m est positive pour tout m pair.

Esquisse de la preuve. Rappelons qu'une orbite $\{t, f(t), \dots, f^{q-1}(t)\}$ est parabolique si et seulement si

$$f^q(t) = t \text{ et } \log Df^q(t) = \sum_{k=0}^{q-1} B(f^k(t)) - q \log \int_0^1 e^B = 0.$$

On appelle *multiplicateur* de l'orbite la quantité $Df^q(t)$ (qui ne dépend pas du point t dans l'orbite). On montrera que pour tout $\lambda > 0$, la probabilité qu'il existe une orbite périodique de multiplicateur égal à λ est nulle.

Pour cela, on peut procéder comme dans la preuve du théorème 1.1, en regardant les événements

$$\begin{aligned} E &= \{\exists t \,:\, f^q(t) = t \text{ et } Df^q(t) = \lambda\} \\ L_\varepsilon &= \{\exists t \in [\varepsilon, 1] \,:\, Df^q(t) = \lambda \text{ et } Df^q(f(t)) = \lambda\}. \end{aligned}$$

La désintégration de la mesure sera faite non seulement selon les événements $A(C, c, b)$ définis à la ligne (II.3), mais aussi sur la valeur α de l'image de 0 par f.

La preuve procède de manière semblable jusqu'à la ligne (II.8) : on veut montrer que pour presque tous $t \in [\varepsilon, 1]$, $C > c$, $b \in \mathbf{R}$, et α,

$$\mathbf{P}_{\{B(\varepsilon)=b, \log \int_0^\varepsilon e^B = c, \log \int_0^1 e^B = C, \alpha\}} (B(f^q(t)) - B(t)) = 0.$$

Au lieu de considérer seulement deux cas différents ($\{f(t) > t\}$ et $\{f(t) < t\}$), nous devons faire face à une combinatoire plus complexe. En effet, si l'on souhaite appliquer la propriété de Markov pour B, il faut que la condition suivante soit vérifiée :

$$\max_{k<q} f^k(t) < f^q(t) < 1.$$

Il est encore possible, par méthode de couplage, de se ramener toujours à ce cas : il suffit en fait d'opérer une conjugaison par une rotation appropriée (voir la figure II.1).

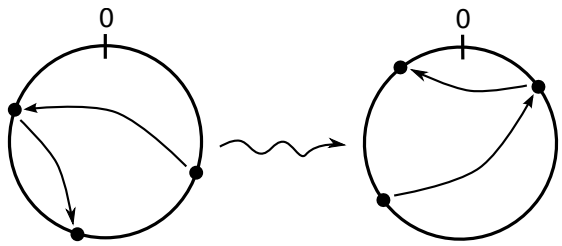

Figura II.1 – Preuve du lemme 1.8 : combinatoire d'une orbite et choix convenable de la rotation

□

Par les mêmes arguments, on peut démontrer le résultat suivant qui est intéressant en soi :

Théorème 1.9. *Soient $f_1, \dots, f_n$ des difféomorphismes de Malliavin-Shavgulidze indépendants. Alors le sous-groupe qu'ils engendrent est presque sûrement libre.*

Dimostrazione. Soit $w = w(f, g)$ un mot non-trivial en f et g. On veut montrer que la probabilité que $w(f, g) = id$ est nulle, lorsque f et g sont deux difféomorphismes de Malliavin-Shavgulidze indépendants. Soit ℓ la longueur du mot w et notons w_j le suffixe de w de longueur j (avec $j \leq \ell$). Il nous suffit de démontrer que la probabilité $\mathbf{P}(\log Dw \equiv 0)$ est égale à 0.

Or, on peut écrire $\log Dw$ sous la forme

$$\log Dw = \sum_{j=0}^{\ell-1} \epsilon_j \left[B_{i_j} \circ \psi_j \circ w_j - \log \int_{S^1} \exp(B_{i_j}) \right],$$

avec $i_j \in \{1,2\}$, $\epsilon_j \in \{\pm 1\}$, $\psi_j \in \{id, f^{-1}, g^{-1}\}$ et B_1, B_2 sont deux ponts browniens indépendants qui définissent f et g. En utilisant l'invariance par translation du processus,

$$\mathbf{P}\left(\forall\, t \in \mathbf{S}^1, \sum_{j=0}^{\ell-1} \epsilon_j \left[B_{i_j}(\psi_j(w_j(t))) - \log \int_{\mathbf{S}^1} \exp(B_{i_j})\right] = 0\right)$$
$$= \mathbf{E}\left[\mathbf{P}_{\{\int_{\mathbf{S}^1} \exp(B_i)=\exp(C_i)\}}\left(\forall t \in \mathbf{S}^1, \sum_{j=0}^{\ell-1} \epsilon_j B_{i_j}(\psi_j(w_j(t))) = \left(\#_{f^{-1}}w - \#_f w\right)\cdot C_1 + \left(\#_{g^{-1}}w - \#_g w\right)\cdot C_2,\right)\right]$$

où l'on prend l'espérance sur les valeurs possibles de $\int_{\mathbf{S}^1} \exp(B_i)$ et $\#_\gamma w$ dénote le nombre de $\gamma \in \{f^{\pm 1}, g^{\pm 1}\}$ dans le mot w.

Or, pour t fixé et presque tous les C_1, C_2,

$$\mathbf{P}_{\{\int_{\mathbf{S}^1} \exp(B_i)=\exp(C_i)\}}\left(\forall t \in \mathbf{S}^1, \sum_{j=0}^{\ell-1} \epsilon_j B_{i_j}(\psi_j w_j(t))) = \left(\#_{f^{-1}}w - \#_f w\right)\cdot C_1 + \left(\#_{g^{-1}}w - \#_g w\right)\cdot C_2,\right) = 0.$$

Ceci est facile à observer : supposons que $\psi_{\ell-1}(w_{\ell-1}(t))$ soit le dernier point sur le cercle à la gauche de 0, parmi les $\{\psi_j(w_j(t))\}$ (sinon, on fait un couplage avec le processus que l'on obtient par la conjugaison avec une rotation appropriée). Alors $B_{i_{\ell-1}}(\psi_{\ell-1}(w_{\ell-1}(t)))$ peut prendre une valeur arbitraire par propriété de Markov du mouvement brownien (le processus que l'on considère est absolument continu par rapport à un mouvement brownien pour tout $t < 1$ fixé). □

1.2.2 Nombre de rotation irrationnel

Une question plus délicate est la suivante :

Question 1. Les difféomorphismes avec nombre de rotation *irrationnel*, constituent-ils un ensemble de mesure positive ?

La topologie dans ce cas ne peut pas nous venir en aide, car l'ensemble des difféomorphismes avec nombre de rotation α irrationnel fixé forment un fermé (contractile), d'intérieur vide dans $\mathrm{Diff}^1_+(\mathbf{S}^1)$ (voir [Her 1]). En utilisant le théorème de conjugaison différentiable (théorème 3.8 dans le chapitre introductif), Herman [Her 2] démontre le résultat suivant :

Théorème 1.10. *Soit f un difféomorphisme du cercle de classe C^3, on définit $K_f \subset \mathbf{S}^1$ comme étant la réunion des paramètres $\lambda \in \mathbf{S}^1$ pour lesquels le nombre de rotation de $R_\lambda \circ f$ est irrationnel. Alors, la mesure de Lebesgue de K_f est strictement positive.*

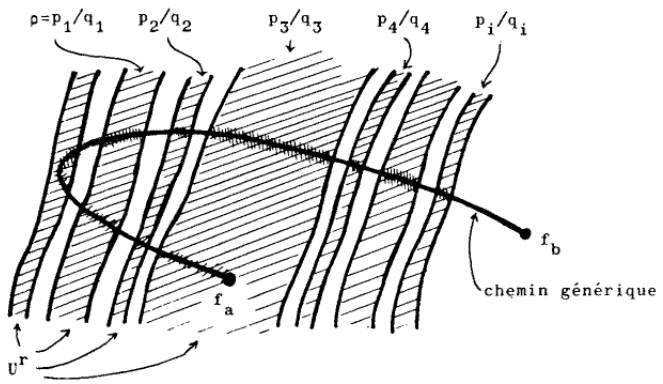

Figura II.2 – Image prise de [Her 2]. Elle représente un chemin générique dans $\mathrm{Diff}^1_+(\mathbf{S}^1)$ qui traverse des zones où le nombre de rotation est constant.

Pour montrer ce théorème, Herman explique que la fonction *nombre de rotation* ρ est différentiable (dans le sens de Gateaux) aux difféomorphismes f avec nombre de rotation irrationnel α et C^1-conjugués à la rotation, et sa différentielle est donnée par

$$v \in \mathrm{Vect}^1(\mathbf{S}^1) \longmapsto \int_{\mathbf{S}^1} (Dh \circ f \circ h^{-1})\, v \circ h^{-1}, \tag{II.10}$$

où h est le difféomorphisme conjuguant f à R_α : $f = h^{-1} \circ R_\alpha \circ h$.

Exemple 1.1. Ce phénomène s'observe par exemple dans la célèbre *famille d'Arnol'd*. Pour tout $a \in]-1/2\pi, 1/2\pi[$ et $\lambda \in \mathbf{S}^1$, on définit la fonction $f_{a,\lambda}(x) = x + \lambda + a\sin(2\pi x)$. Les ensembles $A_{p/q}$ des paramètres a et λ tels que $\rho(f_{a,\lambda}) = p/q$ ont la forme de petites langues. Pour des petites valeurs de a, les t pour lesquels $\rho(f_{a,t})$ est irrationnel est presque 1.

Cette dernière affirmation suit du fait que le bord des régions sur lesquelles le nombre de rotation est constant et rationnel, est Lipschitz au voisinage des pointes ($a \sim 0$, $\rho(f_{a,t})$ rationnel).

Nous avons fait des simulations numériques (voir appendice 4), qui montrent des langues quand on regarde les mesures de Malliavin-Shavgulidze : pour toute couple (α, σ), on peut calculer $\mu_\sigma\left(\rho^{-1}(\alpha)\right)$. Nous pensons que, contrairement à ce que l'on constate dans l'exemple d'Arnol'd, la mesure des difféomorphismes avec nombre de rotation irrationnel est nulle pour tout $\sigma > 0$. L'argument géométrique donné par Herman [Her 2] ne pourras pas être exploité, car l'on s'attend à ce que la fonction ρ n'est pas Lipschitz aux difféomorphismes avec nombre de rotation diophantien, mais $(1/2 - \varepsilon)$-Hölder.

En revanche, il est facile de voir que les difféomorphismes avec nombre de rotation irrationnel fixé, forment une partie de mesure nulle (il est en soi évident que les nombres irrationnels pour lesquels la mesure est positive, doivent être en nombre au plus dénombrable). Pour cela, nous rappellerons un résultat dû à Arnol'd (voir [Her 1]) :

Proposition 1.11. *Soit f un difféomorphisme du cercle qui fixe 0 et soit α un nombre irrationnel. Il existe un seul $\lambda = \lambda(\alpha) \in \mathbf{S}^1$ tel que le nombre de rotation de $R_\lambda \circ f$ soit égal à α. En outre, la fonction $\lambda(\alpha)$ est continue en la variable f.*

Remarque 1.2. On peut aussi montrer que pour tout nombre rationnel p/q, il existe deux fonctions λ_+ et λ_-, telles que $\rho(R_{\lambda_+} \circ f) = \rho(R_{\lambda_+} \circ f) = p/q$ et pour tout $\lambda > \lambda_+$ (resp. $< \lambda_-$) le nombre de rotation de $R_\lambda \circ f$ est strictement supérieur (resp. inférieur). Les difféomorphismes en ces paramètres sont dits *semi-stables*.

La proposition 1.8 montre en particulier que la probabilité qu'un difféomorphisme de Malliavin-Shavgulidze avec nombre de rotation p/q soit semi-stable est nulle. Nous en déduisons, par dénombrabilité des nombres rationnels, que les difféomorphismes semi-stables forment une partie négligeable dans $\mathrm{Diff}^1_+(\mathbf{S}^1)$.

D'après la proposition 1.11, on peut penser à $\rho^{-1}(\alpha)$ comme à une hypersurface dans $\mathrm{Diff}^1_+(\mathbf{S}^1)$, non lisse en général, mais avec une régularité qui dépend de la régularité du difféomorphisme et, surtout, de la condition diophantienne que α vérifie : ces deux paramètres donnent un contrôle sur la différentielle du nombre de rotation (voir (II.10)).

Pour ce qui nous concerne, il suffit de remarquer qu'on peut voir $\rho^{-1}(\alpha)$ comme l'image d'une section continue s_α du fibré en cercles $\mathrm{Diff}^1_+(\mathbf{S}^1)$ sur $\mathbf{S}^1 \backslash \mathrm{Diff}^1_+(\mathbf{S}^1)$, où $\mathbf{S}^1 \backslash \mathrm{Diff}^1_+(\mathbf{S}^1)$ s'identifie, via l'application B donnée par (I.4), à $B^{-1}\left(C_0(\mathbf{S}^1) \times \{0\}\right)$. Par définition de la mesure de Malliavin-Shavgulidze (voir (I.5)), on peut appliquer le théorème de Fubini à l'intégrale

$$\int_{C_0(\mathbf{S}^1)} \int_{\mathbf{S}^1} \mathbf{1}_{\{s_\alpha(x)\}}(t)\, dt\, d\mathbf{W}_{0,\sigma}(x) = \mu_\sigma\left(s_\alpha\left(C_0(\mathbf{S}^1)\right)\right),$$

qui est pourtant nulle.

Proposition 1.12. *Pour tout nombre irrationnel α, la mesure des difféomorphismes avec nombre de rotation α est nulle : $\mu_\sigma\left(\rho^{-1}(\alpha)\right) = 0$.* □

1.2.3 Renormalisation et mesures

L'une des clefs pour comprendre la nature des difféomorphismes du cercle est la *renormalisation*. Cachée dans les travaux de Herman et Yoccoz [Her 1, Yoc 1], elle permet de mieux expliquer pourquoi le résultat de conjugaison différentiable est vrai. Il s'agit de regarder de plus en plus localement la dynamique décrite par f. Il est possible de définir la renormalisation $\mathcal{R}f$ de tout difféomorphisme f sans point fixe : en prenant l'intervalle I_1 contenant 0 et dont les extrémités sont $f(0)$ et $f^{q_1}(0)$ (rappelons de la partie 3.2.2 que l'entier q_1 est choisi comme le plus petit k tel que $f^k(0)$ appartient à l'intervalle de gauche délimité par 0 et $f(0)$ et $f^{k+1}(0)$ à celui de droite), on considère le difféomorphisme du cercle $I_1/_{f^{q_1}(0)\sim f(0)}$ qui est défini par f sur l'intervalle d'extrémités $f^{q_1}(0)$ et 0 et par f^{q_1} sur l'autre (voir figure II.3). On opère par la suite une conjugaison affine pour envoyer $I_1/_{f^{q_1}(0)\sim f(0)}$ sur le cercle $\mathbf{R}/\mathbf{Z}$ standard : nous avons défini le renormalisé $\mathcal{R}f$.

Notons T l'homothétie $T(t) = (f(0) + 1 - f^q(0))\, t$ qui envoie $[0,1]$ sur $[-f^q(0), f(0)]$ en fixant 0. Avec une terminologie plus récente, $\mathcal{R}f$ est défini par le *pair d'applications qui commutent* $(T^{-1}fT, T^{-1}f^qT)$, qui est l'échange d'intervalles généralisé sur $[0,1]$ en figure II.3.

La théorie de Herman-Yoccoz peut s'expliquer en disant que les renormalisations successives $\mathcal{R}^n f$ se rapprochent d'une rotation lorsque le nombre de rotation est diophantien (les difféomorphismes infiniment renormalisables sont exactement ceux avec nombre de rotation irrationnel).

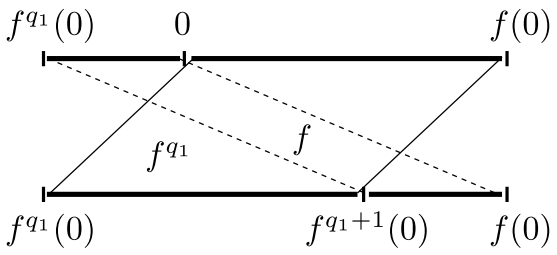

Figura II.3 – Construction de la renormalisation de f

Question 2. Quelle est l'image de la mesure μ_{MS} sous l'opérateur $\mathcal{R}$?

Nous ne savons pas répondre à cette question. Nous pouvons par contre faire une remarque facile. Nous avons vu, dans le chapitre précédent, que lorsque l'on compose à droite par un difféomorphisme φ, l'événement (I.9) est de mesure 1 pour la probabilité image. Lorsque l'on renormalise f, on fait apparaitre une itération f^q dans la définition de $\mathcal{R}f$, qui modifie cette loi des logarithmes itérés : pour tout $t \in [0,1]$ on a

$$\begin{aligned}
&\overline{\lim_{\varepsilon\to 0}} \frac{\log D\mathcal{R}f(t+\varepsilon) - \log D\mathcal{R}f(t)}{\sqrt{2\varepsilon \log\log(1/\varepsilon)}} \\
&\qquad = \begin{cases} \sqrt{T'(t)} & \text{si } t \in \left]-\dfrac{f^q(0)}{f(0)+1-f^q(0)}, 0\right[, \\[2ex] \displaystyle\sum_{k=0}^{q-1} \overline{\lim_{\varepsilon\to 0}} \frac{\log Df(f^k(T(t+\varepsilon))) - \log Df(f(T(t)))}{\sqrt{2\varepsilon\log\log(1/\varepsilon)}} & \text{si } t \in \left]0, \dfrac{f(0)}{f(0)+1-f^q(0)}\right[. \end{cases} \\
&\qquad = \begin{cases} \sqrt{f(0)+1-f^q(0)} & \text{si } t \in \left]-\dfrac{f^q(0)}{f(0)+1-f^q(0)}, 0\right[, \\[2ex] \sqrt{f(0)+1-f^q(0)} \displaystyle\sum_{k=0}^{q-1} \sqrt{Df^k\left((f(0)+1-f^q(0))t\right)} & \text{si } t \in \left]0, \dfrac{f(0)}{f(0)+1-f^q(0)}\right[. \end{cases}
\end{aligned}$$

Nous en déduisons

Proposition 1.13. *Les mesures $\mathcal{R}_*\mu_{MS}$ et μ_{MS} ne sont pas équivalentes.* □

La proposition précédente donne seulement une réponse partielle au problème : si l'on s'attend à ce que la renormalisation « lisse » les difféomorphismes, il faut croire que ce module de continuité local donné par la loi de logarithmes itérés doit tendre vers 0 lorsque l'on applique $\mathcal{R}$ une infinité de fois. En d'autres mots, on doit penser que l'opérateur $\mathcal{R}$ concentre le support de la mesure μ_{MS} de plus en plus dans un voisinage $C^{1+1/2-\varepsilon}$ des rotations : soit n, disons, un nombre impair, et regardons encore la loi des logarithmes itérés après n renormalisations d'un difféomorphisme renormalisable n fois. On a

$$\begin{aligned}
&\overline{\lim_{\varepsilon\to 0}} \frac{\log D\mathcal{R}^n f(t+\varepsilon) - \log D\mathcal{R}^n f(t)}{\sqrt{2\varepsilon \log\log(1/\varepsilon)}} \\
&\qquad = \begin{cases} \sqrt{f^{q_{n-1}}(0)+1-f^{q_n}(0)} \displaystyle\sum_{k=0}^{q_{n-1}-1} \sqrt{Df^k\left((f^{q_{n-1}}(0)+1-f^{q_n}(0))t\right)} & \text{si } t \in \left]-\dfrac{f^{q_n}(0)}{f^{q_{n-1}}(0)+1-f^{q_n}(0)}, 0\right[, \\[2ex] \sqrt{f^{q_{n-1}}(0)+1-f^{q_n}(0)} \displaystyle\sum_{k=0}^{q_n-1} \sqrt{Df^k\left((f^{q_{n-1}}(0)+1-f^{q_n}(0))t\right)} & \text{si } t \in \left]0, \dfrac{f^{q_{n-1}}(0)}{f^{q_{n-1}}(0)+1-f^{q_n}(0)}\right[. \end{cases}
\end{aligned}$$

Si le nombre de rotation de f est irrationnel, on peut regarder le comportement lorsque n tend vers l'infini. Si f n'est pas minimal, alors le facteur $\sqrt{f^{q_{n-1}}(0)+1-f^{q_n}(0)}$ converge vers un nombre strictement positif, et on devrait penser que les sommes convergent (d'après le résultat de Norton expliqué dans la partie 3.1). En revanche, si f est minimal, il faut regarder finement la compétition entre ces deux limites : le facteur $\sqrt{f^{q_{n-1}}(0)+1-f^{q_n}(0)}$ tend vers 0, mais la somme diverge.

Pourquoi sommes-nous convaincus que le nombre de rotation doit être rationnel presque sûrement ? Le fait est que pour construire $\mathcal{R}f$ on se retrouve toujours avec beaucoup de choix. Nous allons être plus précis.

Considérons le processus $(f, \mathcal{G}_t)$, où $\mathcal{G}_t$ est la filtration introduite à la ligne (II.2) (avec les modifications appropriées en considérant que l'on travaille avec un *pont* brownien). Pour définir $\mathcal{R}f$, nous n'avons pas besoin de connaitre entièrement le pont brownien entre 0 et 1.

1. Le premier point de l'orbite est $f(0) = \alpha$: il est choisi uniformément par rapport à Lebesgue et il est indépendant de $\mathcal{G}_t$ pour tout t.

2. Ensuite la définition de $f(\alpha)$ nécessite la connaissance de α et de B entre 0 et α. Supposons
$$\alpha < f(\alpha) < 1 \tag{$*$}$$
(dans le cas contraire on pourra faire un couplage avec $1 - f(1-t)$). Au lieu de considérer $(B_t)_{t\in[0,\alpha]}$, on peut se restreindre à la variable
$$\frac{\int_0^\alpha e^B}{\int_0^1 e^B} + \alpha \ :$$
nous avons que $f^2(0)$ est mesurable par rapport à la tribu
$$\sigma\left(B_1, \int_0^1 e^B, \alpha, \int_0^\alpha e^B\right).$$

3. Sous la condition précédente, q_1 définit un temps d'arrêt qui est mesurable par rapport à la tribu
$$\sigma\left(B_1, \int_0^1 e^B, \alpha, \int_0^\alpha e^B, (B_t)_{t\in[\alpha,1[}\right).$$
Le temps q_1 est fini si et seulement si f n'a pas de point fixe. On en déduit que la variable $f^{q_1+1}(0)$ est mesurable par rapport à la tribu
$$\mathcal{T}_1 = \sigma\left(B_1, \int_0^1 e^B, \alpha, \int_0^\alpha e^B, (B_t)_{t\in[\alpha, f^{q_1}(0)]}\right).$$

4. En conclusion, l'intervalle I_1 sur lequel on définit la renormalisation $\mathcal{R}f$, ainsi que le couple d'applications (f, f^{q_1}), sont mesurables par rapport à la tribu $\mathcal{T}_1$. Ce qui est important est qu'ils sont presque indépendants de l'évolution de B dans l'intervalle délimité par $f^{q_1}(0)$ et $f(0)$ qui contient 0 : il faut seulement connaitre
$$\int_{f^{q_1}(0)}^1 e^B \text{ et } \int_0^{f(0)} e^B.$$
Le point $f^{q_1+1}(0)$ joue le rôle de α dans le premier point.

5. En répétant les étapes précédentes, on en déduit que l'intervalle I_n d'extrémités $f^{q_{n-1}}(0)$ et $f^{q_n}(0)$ qui contient 0 et le couple $(f^{q_{n-1}}, f^{q_n})$ sont *presque indépendants* de l'évolution de B dans l'intervalle I_n, modulo deux conditions de bords comme dans le point précédent.

En conclusion, le fait que le difféomorphisme aléatoire f soit renormalisable une infinité de fois est indéterminé si l'on connait l'évolution de B partout, sauf en un petit intervalle autour de 0. On aimerait alors disposer d'une espèce de propriété de Markov pour pouvoir appliquer une loi du 0–1. Les images dans la figure II.12 sont assez suggestives sous ce point de vue.

2 Centralisateurs

2.1 Quelques rappels sur les centralisateurs des difféomorphismes de l'intervalle

Rappelons d'abord un résultat classique, dû originairement a Sternberg (voir [Nav 1, Yoc 2]), mais rédigé sous cette version $C^{1+\alpha}$ par M. Chaperon [Cha].

Définition 2.1. *On dit que deux difféomorphismes f et g de classe $C^{1+\alpha}$ définissent le même* germe *s'ils coïncident dans un voisinage suffisamment petit de l'origine. On définit $\mathcal{G}_+^{1+\alpha}(\mathbf{R}, 0)$ comme étant l'*espace des germes des difféomorphismes $C^{1+\alpha}$ *de la droite qui fixent* 0 *et préservent l'orientation.*

On dit qu'un germe $[g]$ est hyperbolique *si la dérivée en* 0 *de g définissant $[g]$ est différente de* 1 *(observer que cette notion est toujours bien définie).*

Théorème 2.1 (Théorème de linéarisation de Sternberg). *Soient $[g] \in \mathcal{G}_+^{1+\alpha}(\mathbf{R}, 0)$ un germe hyperbolique et $a := g'(0)$. Il existe alors un germe $[h] \in \mathcal{G}_+^{1+\alpha}(\mathbf{R}, 0)$ tel que*

$$h'(0) = 1 \text{ et } h(g(x)) = ah(x) \text{ dans un voisinage de l'origine.} \tag{$*$}$$

En outre, si $[h_1]$ est un germe C^1 qui vérifie ($*$), alors $[h_1]$ est de classe $C^{1+\alpha}$ est définit le même germe que h.

Dimostrazione. Quitte à passer à g^{-1}, on peut supposer $a < 1$. Soit $\delta > 0$ suffisamment petit tel que g soit défini sur l'intervalle $[-\delta, \delta]$, on pose $C(\delta) := \sup_{[-\delta,\delta]} |g'|$. On définit l'espace vectoriel E_δ des fonctions ψ de classe $C^{1+\alpha}$ sur $[-\delta, \delta]$, telles que $\psi(0) = \psi'(0) = 0$. On munit E_δ de la norme

$$\|\psi\| = \|\psi'\|_\alpha.$$

L'espace E_δ est alors un espace de Banach. On définit l'endomorphisme S_δ de E_δ :

$$S_\delta[\psi] = \frac{1}{a}(\psi \circ g).$$

On montre maintenant que pour δ suffisamment petit, S_δ est une contraction. Soient t et s dans $[-\delta, \delta]$, on a

$$\begin{aligned}
\left|S_\delta[\psi](t) - S_\delta[\psi](s)\right| &= \frac{1}{a}\left|(\psi \circ g)'(t) - (\psi \circ g)'(s)\right| = \left|\psi'\left(g(t)\right) g'(t) - \psi'\left(g(s)\right) g'(s)\right| \\
&\leq \frac{1}{a}\left|\psi'\left(g(t)\right) g'(t) - \psi'\left(g(t)\right) g'(s)\right| + \frac{1}{a}\left|\psi'\left(g(t)\right) g'(s) - \psi'\left(g(s)\right) g'(s)\right| \\
&\leq \frac{\left|\psi'\left(g(t)\right)\right|}{a} \|g'\|_\alpha |t-s|^\alpha + \frac{|g'(s)|}{a} \|\psi\| |g(t) - g(s)|^\alpha \\
&= \frac{\left|\psi'\left(g(t)\right) - \psi'(0)\right|}{a} \|g'\|_\alpha |t-s|^\alpha + \frac{|g'(s)|}{a} \|\psi\| |g(t) - g(s)|^\alpha \\
&\leq \frac{\|\psi\| |g(t)|^\alpha}{a} \|g'\|_\alpha |t-s|^\alpha + \frac{C(\delta)}{a} \|\psi\| C(\delta)^\alpha |t-s|^\alpha \\
&\leq \frac{\|\psi\|}{a}\left[C(\delta)^\alpha |t|^\alpha \|g'\|_\alpha + C(\delta)^{1+\alpha}\right] |t-s|^\alpha \leq \frac{\|\psi\|}{a}\left[C(\delta)^\alpha \delta^\alpha \|g'\|_\alpha + C(\delta)^{1+\alpha}\right] |t-s|^\alpha.
\end{aligned}$$

Il suffira donc de prendre δ tel que

$$C(\delta)^\alpha \delta^\alpha \|g'\|_\alpha + C(\delta)^{1+\alpha} < a,$$

ou bien

$$\delta < \frac{1}{C(\delta)}\left[\frac{a - C(\delta)^{1+\alpha}}{\|g'\|_\alpha}\right]^{1/\alpha}, \tag{II.11}$$

qui existe, s'il est suffisamment petit pour qu'il vérifie aussi

$$C(\delta) - a < a\left(\frac{1}{a^{\frac{\alpha}{1+\alpha}}} - 1\right).$$

En effet, dans ce cas, $\lim_{\delta \to 0} C(\delta) = a$ et $a < 1$, donc la quantité de droite dans (II.11) est strictement positive.

Remarquons que la fonction ψ_1 définie par $\psi_1(x) = g(x) - ax$ se restreint sur $[-\delta, \delta]$ à une fonction dans E_δ. Alors, d'après ce qu'on vient de montrer, l'équation

$$S_\delta[\psi] + \frac{1}{a}\psi_1 = \psi$$

possède une unique solution $\psi_0 \in E_\delta$. Posons $h := id + \psi_0$ et montrons que ce h convient :

$$\begin{aligned}
h\left(g(x)\right) &= g(x) + \psi_0\left(g(x)\right) \\
&= \psi_1 + ax + \psi_0\left(g(x)\right) \\
&= \psi_1(x) + ax + a\, S_\delta[\psi_0](x) \\
&= ax + a\,\psi_0(x) = ah(x)
\end{aligned}$$

et, en outre, $h(0) = 0$ et $h'(0) = 1$. Donc h définit un germe $C^{1+\alpha}$ avec les propriétés requises.

Il reste à montrer l'unicité. Soit h_1 un germe C^1 qui vérifie les bonnes propriétés, alors $g \circ h_1^{-1} = h_1^{-1} \circ M_a$ (où M_a est la multiplication par a), d'où

$$h \circ h_1^{-1} \circ M_a = h \circ g \circ h_1^{-1} = M_a \circ h \circ h_1^{-1},$$

qui veut dire que $h \circ h_1^{-1}$ commute avec M_a. Soit t un point suffisamment près de 0, alors

$$\begin{aligned} h\left(h_1^{-1}(t)\right) &= \lim_{n\to\infty} \frac{h\left(h_1^{-1}(a^n t)\right)}{a^n} \\ &= t \lim_{n\to\infty} \frac{h\left(h_1^{-1}(a^n t)\right)}{a_n t} \\ &= t\left(h \circ h_1^{-1}\right)'(0) = t. \end{aligned}$$

□

Corollaire 2.2. *Soit f un difféomorphisme $C^{1+\alpha}$ de l'intervalle avec 0 et 1 seuls points fixes. On suppose $a := f'(0) \neq 1$, alors il existe un difféomorphisme h de $[0,1[$ qui est de classe $C^{1+\alpha}$ sur $[0,1[$ et vérifie*

$$h'(0) = 1 \text{ et } h(f(x)) = ah(x) \text{ pour tout } x \in [0,1[. \tag{$**$}$$

En outre, si h_1 est un difféomorphisme C^1 qui vérifie ($**$), alors h_1 est de classe $C^{1+\alpha}$ et coïncide avec h.

Dimostrazione. Par le théorème de linéarisation de Sternberg, il existe un voisinage de 0 et un difféomorphisme $\tilde{h}$ défini sur ce voisinage tel que ($*$) soit satisfait. Ceci entraîne que la définition du difféomorphisme h souhaité est entièrement déterminée : en effet, tout voisinage de 0 contient un domaine fondamental pour l'action de f, de la forme $J_0 = [x_0, f(x_0)[$ (si l'on suppose, sans perte de généralité, $a > 1$). Soit maintenant $x \in [0,1[$ un point quelconque. Il existe un unique entier $n = n(x)$ tel que $f^{n(x)}(x) \in J_0$. On définit $h(x) := a^{-n(x)}\tilde{h}(f^{n(x)}(x))$. La fonction $n = n(x)$ vérifie $n(f(x)) = n(x) - 1$. Il en découle l'égalité

$$h(f(x)) = a^{-n(f(x))}\tilde{h}\left(f^{n(f(x))}(f(x))\right) = a^{-n(x)+1}\tilde{h}\left(f^{n(x)}(x)\right) = ah(x).$$

□

Puisque tout difféomorphisme f vérifiant les hypothèses du corollaire précédent est conjugué à M_a, la multiplication linéaire par $a = f'(0)$, il s'ensuit, en particulier, que le centralisateur C^1 de f sur $[0,1[$ est conjugué par le difféomorphisme h au centralisateur C^1 de la multiplication par $f'(0)$ sur $[0,+\infty[$. La description de ce dernier est assez élémentaire.

Soit φ un difféomorphisme sur $[0,+\infty[$ qui commute avec M_a. Supposons, sans perte de généralité, que a soit inférieur à 1. Pour tout $x \in [0,+\infty[$ et $n > 0$ entier,

$$\varphi(x) = \frac{\varphi(a^n x)}{a^n} \text{ et } \varphi'(x) = \varphi'(a^n x). \tag{II.12}$$

Puisque, pour tout x fixé, la suite $(a^n x)_{n\in\mathbf{N}}$ tend vers 0, (II.12) force la dérivée de φ à être constante.

On vient de démontrer que le centralisateur C^1 de M_a est égal à $\mathbf{R}^+$, qu'on peut voir comme étant le flot à un paramètre du champ de vecteurs $a\frac{\partial}{\partial x}$.

On en déduit que le centralisateur C^1 de f sur $[0,1[$ est égal au groupe à un paramètre $\{f_t\}$ engendré par le *champ de vecteurs de Szekeres* X_f associé à f, défini par $X_f(x) = h(x)\,Dh^{-1}(x)\,\log a$. On remarquera que le flot de X_f au temps t est défini par $f_t(x) = h^{-1}(a^t h(x))$.

Nous nous rapprochons de l'étude de l'*invariant de Mather*. Même si N. Kopell dans [Kop] avait entièrement expliqué ce dont nous avons principalement besoin, nous allons suivre l'exposition de J.-C. Yoccoz [Yoc 2].

Soit f un difféomorphisme de l'intervalle avec 0 et 1 comme uniques points fixes, supposés hyperboliques. On note $a = f'(0)$ et $b = f'(1)$. Prenons en exemple le cas $a > 1 > b$. Soient X_f et Y_f les champs de vecteurs de Szekeres associés à f sur les intervalles $[0,1[$ et $]0,1]$ respectivement. Si $g \in \mathrm{Diff}^1_+(I)$ commute avec f, d'après ce qui précède, il doit exister deux temps τ et τ' tels que $g = f_\tau = f^{\tau'}$, où $\{f_t\}_{t\in\mathbf{R}}$ est le flot engendré par X_f et $\{f^t\}_{t\in\mathbf{R}}$ celui par Y_f. Soient $x, y \in]0,1[$, on définit $W^{x,y}$ comme étant la fonction de « raccord » entre $\{f^t\}$ et $\{f_t\}$:

$$f^t(y) = f_{W^{x,y}(t)}(x). \tag{II.13}$$

Si h et k sont les difféomorphismes qui linéarisent f sur $[0,1[$ et $]0,1]$ respectivement, l'équation (II.13) devient

$$k^{-1}\left(b^t k(y)\right) = h^{-1}\left(a^{W^{x,y}(t)} h(x)\right),$$

d'où

$$W^{x,y}(t) = \log_a \frac{h\left(k^{-1}\left(b^t k(y)\right)\right)}{h(x)}.$$

La fonction $W^{x,y}(t)$ est un difféomorphisme $C^{1+\alpha}$ de $\mathbf{R}$ dans $\mathbf{R}$ qui vérifie de plus

$$\begin{aligned} W^{x,y}(t+1) &= \log_a \frac{h\left(k^{-1}\left(b\, b^t k(y)\right)\right)}{h(x)} \\ &= \log_a \frac{h\left(f\left(k^{-1}\left(b^t k(y)\right)\right)\right)}{h(x)} \\ &= \log_a \frac{ah\left(k^{-1}\left(b^t k(y)\right)\right)}{h(x)} = 1 + W^{x,y}(t) \end{aligned}$$

En d'autres termes, $W^{x,y}$ commute avec les translations entières et, par passage au quotient de cette action, il définit un difféomorphisme du cercle de classe $C^{1+\alpha}$, que l'on note encore $W^{x,y}$.

On remarquera aussi, avec des vérifications analogues, que pour tous s, t, t',

$$W^{f_t(x), f^{t'}(y)}(s) = W^{x,y}(s+t') - t.$$

Puisque $g = f_\tau = f^{\tau'}$, pour tous x, $y \in]0,1[$, $W^{x,y}$ doit vérifier

$$W^{x,y}(s+\tau') - \tau = W^{g(x),g(y)}(s). \tag{II.14}$$

Proposition 2.3 ([Yoc 2]). *Le difféomorphisme $g = f_\tau = f^{\tau'}$ commute avec f sur $[0,1]$ si et seulement si, pour tous x, $y \in]0,1[$, (II.14) est vérifiée. C'est-à-dire, le centralisateur C^1 de f dans $\mathrm{Diff}^1_+(I)$ est isomorphe au stabilisateur de $W^{x,y}$ sous l'action de $\mathbf{R}^2$:*

$$(t,t') \cdot W^{x,y}(s) = W^{x,y}(s+t') - t. \tag{II.15}$$

On peut voir que le stabilisateur de $W^{x,y}$ est toujours soit cyclique, engendré par $\left(\frac{1}{n}, \frac{1}{n}\right)$, où $\frac{1}{n}$ correspond à la sous-période de $W^{x,y} - id$, ou bien isomorphe à $\mathbf{R}$ entier.

Remarque 2.1. Les difféomorphismes de classe $C^{1+\alpha}$ qui ont un centralisateur C^1 trivial forment une partie ouverte et dense.

Étendons maintenant l'étude précédente aux difféomorphismes hyperboliques avec un nombre arbitraire de points fixes, $0 = p_0 < q_0 < p_1 < q_1 < \ldots < p_m < q_m = 1$. Le raisonnement est valide pour tout intervalle de la forme $[p_i, q_i]$ ou $[q_i, p_{i+1}]$. Remarquons que pour tout point p_i (ou q_i), le théorème de linéarisation de Sternberg 2.1 permet de trouver un difféomorphisme h_i de classe $C^{1+\alpha}$ défini sur $]q_{i-1}, q_i[$ (ou $]p_i, p_{i+1}[$) à valeurs dans $\mathbf{R}$, tel que

$$h_i'(p_i) = 1 \text{ et } h_i(f(x)) = f'(p_i)\, h_i(x) \text{ pour tout } x \in]q_{i-1}, q_i[. \tag{II.16}$$

Cela permet d'associer à f un champ de vecteurs de Szekeres sur l'intervalle $]q_{i-1}, q_i[$, de sorte que les flots f^t sur $]q_{i-1}, p_i]$ et f_t sur $[p_i, q_i[$ se raccordent en p_i.

Si g est un difféomorphisme de classe C^1 qui centralise f, il correspond à un certain f^t sur $]q_{i-1}, p_i]$ et donc, par (II.16), à f_t sur $[p_i, q_i[$. Nous en déduisons que la définition d'un tel g sur un quelconque des intervalles $[p_i, q_i]$ ou $[q_i, p_{i+1}]$ force la définition de g sur l'intervalle entier. Ainsi, la proposition 2.3 implique que le centralisateur C^1 de f sur $[0,1]$ est isomorphe au stabilisateur de $(W^{x_0,y_0}_{0,0}, W^{x_0',y_1}_{0,1}, W^{x_1,y_1'}_{1,1}, \ldots)$ par l'action diagonale de $\mathbf{R}^2$ définie sur chaque composante par (II.15) (nous avons noté $W^{x,y}_{i,i}$, $W^{x,y}_{i,i+1}$ les $W^{x,y}$ correspondant aux intervalles $[p_i, q_i]$ et $[q_i, p_{i+1}]$).

2.2 Centralisateur C^1 d'un difféomorphisme hyperbolique du cercle

Soit f un difféomorphisme de classe $C^{1+\alpha}$ qui possède uniquement des orbites périodiques hyperboliques. On suppose que le nombre de rotation de f est égal à p/q. Pour souci de simplicité on suppose que les multiplicateurs de chaque orbite sont différents l'un de l'autre. Alors le difféomorphisme $F := f^q$ est un difféomorphisme du cercle avec uniquement des points fixes hyperboliques. Un résultat de Kopell [Kop] (voir aussi [Yoc 2]) montre que les centralisateurs C^1 de f et F coïncident. La preuve de ce fait n'est pas trop difficile, mais elle ne rajoute aucun élément nouveau vraiment instructif pour nous à présent. Ainsi, nous pouvons nous restreindre au cas où le nombre de rotation de f vaut 0. Il est clair que pour tout point fixe $x_* \in \mathbf{S}^1$, le difféomorphisme f définit un difféomorphisme f_* de l'intervalle $I_* := [x_*, x_* + 1]$, et tout élément du centralisateur de f sur $\mathbf{S}^1$, centralise aussi f_* sur I_*, si l'on suppose qu'un tel élément fixe x_*. La grande différence dans l'étude des centralisateurs sur le cercle et sur l'intervalle se manifeste par une partie de torsion : il peut exister un difféomorphisme g qui permute cycliquement les points attractifs et les points répulsifs de f, tout en commutant avec f. Remarquons que sous notre hypothèse sur les multiplicateurs des orbites (donc sur les valeurs de la dérivée de f aux points fixes), ceci ne peut cependant pas se produire.

Proposition 2.4 ([Yoc 2]). *Avec les notations précédentes, le centralisateur C^1 de f est isomorphe à un sous-groupe fermé de $\mathbf{R}$. De plus, il existe un ouvert dense de f pour lequel le centralisateur C^1 est trivial.*

Théorème 2.5. *Pour presque tout difféomorphisme f du cercle avec nombre de rotation rationnel, le centralisateur C^1 de f est trivial.*

Nous utiliserons le lemme suivant dont la preuve est très simple.

Lemme 2.6. *Soit $f^q : [x_0, x_1] \longrightarrow [x_1, x_2]$ un difféomorphisme aléatoire obtenu en itérant q fois un difféomorphisme de Malliavin-Shavgulidze f avec dérivées aux extrémités fixées. Soient $(a_n)_{n\in\mathbf{N}} \subset]x_0, x_1[$ et $(b_n)_{n\in\mathbf{N}} \subset]x_1, x_2[$ deux suites (décroissantes). Alors la probabilité qu'il existe n tel que $f^q(a_n) = b_n$ est nulle.*

Pour montrer le théorème 2.5, nous suivons la preuve de Kopell dans [Kop], du fait que les contractions avec centralisateur C^1 trivial constituent une partie dense parmi les contractions.

Dimostrazione. Sans perte de généralité, on peut supposer que le nombre de rotation de f soit 0. Fixons $x_0 \in]x, y[$ et posons, pour tout $n \in \mathbf{Z}$, $x_n = f^n(x_0)$. Considérons l'ensemble des f tels que

1. f préserve un intervalle $I = [x, y]$ et sa restriction à $[x, y[$ est une contraction.
2. On connait f et Df sur $J = [x, x_1]$ et $J' = [x_0, y]$ et on sait que $f(x_0) = x_1$.

Nous allons conditionner la mesure de Malliavin-Shavgulidze par rapport à cet événement.

La deuxième condition implique que l'on connait les champs de vecteurs X et Y associés à f autour de x et de y, sur les intervalles J et J'. En particulier on connait les flots à un paramètre f_t et f^t sur les intervalles J et J' respectivement.

Soit h un difféomorphisme de I tel que $h_{|J\cup J'} = f_{|J\cup J'}$ et tel que h appartient à un flot à un paramètre $\{h^t\}$ de difféomorphismes C^1.

Tout difféomorphisme aléatoire f détermine un homéomorphisme $\psi = \psi_f$ de $]x, y]$ tel que

1. $\psi_{|[x_1,y]} = id$,
2. $\psi h = h\psi$,
3. $f = \psi^{-1} h \psi$.

En effet, il suffit de définir le difféomorphisme $\beta = h \circ f^{-1}$ de $[x_2, x_1]$ dans lui-même, tangent à l'identité en x_1 et x_2 : si l'on pose $\psi_{|[x_2,x_1]} = \beta$, on détermine ψ entièrement.

Le lemme 2.6 implique que presque sûrement, il n'existe pas de $n > 0$ tel que $f(h^{1/n}(x_0)) = h^{1/n}(x_1)$. Donc presque sûrement, il n'existe pas de $n > 0$ tel que

$$\psi(h^{1/n}(x_1)) = h^{1/n}\psi(x_1),$$

puisque $\psi = h \circ f^{-1}$ sur $[x_2, x_1]$.

Donc pour tout $k \geq 1$, il n'existe presque sûrement pas de $n > 0$ tel que

$$\psi(h^{1/n}(x_k)) = h^{1/n}\psi(x_k), \tag{II.17}$$

Supposons que le centralisateur C^1 de f ne soit pas trivial : il existe $n > 0$ tel que $f^{1/n}$ est différentiable en x et par unicité du flot à un paramètre on doit avoir $f^{1/n} = \psi^{-1}h^{1/n}\psi$. Alors $f^{1/n}$ commute avec $f = \psi^{-1}h\psi$, qui est égal à h sur $[x, x_1]$. Cela implique que sur l'intervalle $[x, x_1]$, on a $f^{1/n} = h^{1/n}$.

Donc $\psi(h^{1/n}(x)) = h^{1/n}\psi(x)$ sur $[x, x_1]$, qui est possible uniquement avec probabilité nulle d'après (II.17). □

Remarque 2.2. La preuve du théorème précédent est assez « molle » : elle ne requiert pas d'expliciter la mesure de Malliavin-Shavgulidze, en utilisant seulement ses propriétés qualitatives. Notamment, Les propriétés données par la proposition 1.8 et le lemme 2.6, avec la propriété de régularité Hölder presque sûre, suffisent.

3 Exemples de Denjoy

Nous abordons ici l'étude des difféomorphismes avec nombre de rotation irrationnel. On doit rappeler qu'à présent, on ne sait pas si les difféomorphismes forment une partie de mesure positive, par rapport à la mesure de Malliavin-Shavgulidze (question 1). Cependant, pour traiter, au sens de la mesure, l'ensemble des difféomorphismes avec nombre de rotation irrationnel α fixé, nous introduisons des mesures de Malliavin-Shavgulidze « locales ».

La proposition 1.11 nous a permis de voir $\rho^{-1}(\alpha)$ comme étant une hypersurface dans $\mathrm{Diff}^1_+(\mathbf{S}^1)$. Nous noterons aussi $\mathrm{Diff}^r_{+,\alpha}(\mathbf{S}^1)$ la réunion des difféomorphismes de classe C^r et nombre de rotation α fixé.

Définition 3.1. *La mesure de Malliavin-Shavgulidze* locale *sur* $\mathrm{Diff}^r_{+,\alpha}(\mathbf{S}^1)$ *est définie comme étant la mesure image de* $\mathbf{W}_{0,\sigma}$ *par rapport à la section* s_α *qui à* $f_0(t) = \frac{\int_0^t e^{x(s)}\,ds}{\int_0^1 e^{x(s)}\,ds}$ *associe* $s_\alpha(f_0) = R_{\lambda(\alpha)} \circ f_0$*, l'unique translatée de* f_0 *avec nombre de rotation* α*. On notera* $\mu_{\sigma,\alpha}$ *la mesure* $(s_\alpha)_*\mathbf{W}_{0,\sigma}$*.*

Par rapport aux mesures μ_σ, les mesures locales $\mu_{\sigma,\alpha}$ ont l'avantage de nous faire travailler avec des ponts browniens. Si l'on conditionnait un difféomorphisme de Malliavin-Shavgulidze à avoir nombre de rotation irrationnel, la mesure que l'on obtiendrait sur $\mathrm{Diff}^r_{+,\alpha}(\mathbf{S}^1)$ ne serait pas pour autant « homogène » : la densité en un point $x = \log Df - \log Df(0)$ dépendrait de la mesure de Lebesgue du Cantor K_f défini dans l'énoncé du théorème de Herman 1.10.

Remarque 3.1. Pour tout $\alpha \in \mathbf{R} - \mathbf{Q}$, la fonction s_α est un homéomorphisme qui préserve les classes de régularité. Il s'ensuit que presque sûrement, un difféomorphisme $f \in \mathrm{Diff}^r_{+,\alpha}(\mathbf{S}^1)$ est de classe C^{1+w} (w module de Lévy (4)), sa variation n'est pas bornée, etc.

D'après la remarque précédente, presque sûrement le théorème de Denjoy (théorème 3.3 dans l'introduction) ne peut pas s'appliquer. La question à laquelle nous souhaiterions répondre est la suivante :

Question 3. Pour chaque α irrationnel fixé, quelle est la mesure $\mu_{\sigma,\alpha}$ des difféomorphismes qui sont conjugués topologiquement à la rotation ? De quelle manière cette mesure dépend de la condition diophantienne satisfaite par α ? Est-il vrai, par exemple, que si α satisfait une condition diophantienne d'ordre δ, alors, presque sûrement, tout difféomorphisme est $C^{\frac{1}{2}-\varepsilon-\delta}$-conjugué à la rotation ?

3.1 Rappels sur les exemples de Denjoy

Pour tout α irrationnel et $\tau < 1/2$, il est possible de construire des exemples de Denjoy avec nombre de rotation α et de classe $C^{1+\tau}$. Ces exemples ont été traités systématiquement pour la première fois par Herman [Her 1, Chapitre X]. Nous les représentons ici, en nous inspirant plutôt de A. Norton [Nor 2].

On pose $\sigma = 1/\tau$ et on choisit $M = M(\sigma)$, entier suffisamment grand tel que $\left(\frac{M+1}{M}\right)^\sigma < \frac{5}{4}$. Pour tout $n \in \mathbf{Z}$, on définit $\ell_n := \frac{A}{(|n|+M)^\tau}$, où A a été choisi pour que $\sum_{n\in\mathbf{Z}} \ell_n = 1$ (cette condition assurera que le le Cantor minimal invariant est un ensemble de mesure de Lebesgue nulle).

Pour tout n fixé, soit I_n un intervalle fermé dans $\mathbf{S}^1$ et de longueur ℓ_n. On force les intervalles I_n à être ordonnés combinatoirement sur le cercle comme une orbite de la rotation d'angle α (on peut par exemple faire correspondre $n\alpha$ à I_n). Ce choix, et la densité de l'orbite $(n\alpha)_{n\in\mathbf{Z}}$, implique que $K := \mathbf{S}^1 - \bigcup_{n\in\mathbf{Z}} I_n$ est un ensemble de Cantor (de mesure de Lebesgue nulle).

On passe ensuite à la construction du difféomorphisme. Pour tout n, on définit un difféomorphisme f_n de classe $C^{1+\tau}$ de I_n sur I_{n+1} et on pose $f\big|_{I_n} = f_n$; la densité de $\bigcup_{n\in\mathbf{Z}} I_n$ implique que f s'étend à un homéomorphisme de $\mathbf{S}^1$ dans $\mathbf{S}^1$. Afin que f soit globalement de classe $C^{1+\tau}$ sur $\mathbf{S}^1$, il faut prendre quelques précautions pour définir les f_n.

Soit $\varphi : [0,1] \longrightarrow [0,4]$ une fonction C^∞ de support $\left[\frac{1}{3}, \frac{2}{3}\right]$ et de moyenne globale 1. Si $x \in I_n = [a_n, b_n]$, on définit

$$g_n(x) = 1 + \frac{\ell_{n+1} - \ell_n}{\ell_n}\,\varphi\left(\frac{x - a_n}{\ell_n}\right).$$

Alors g_n est une fonction strictement positive de classe C^∞ et $g_n(a_n) = g_n(b_n) = 1$. En outre, le choix du M implique que g_n est toujours inférieur, disons, à 2. Il s'ensuit que tout difféomorphisme

$$f_n(x) := a_{n+1} + \int_{a_n}^{b_n} g_n(y)\,dy$$

a une dérivée uniformément bornée par 2 et elle tend vers 1 lorsque $|n|$ tend vers l'infini.

Ces considérations permettent de définir globalement f, comme expliqué plus haut, et alors f sera un difféomorphisme de classe C^1 en tout point $x \in \mathbf{S}^1$. Remarquons que la dérivée de f vaut 1 sur le Cantor K.

Il reste à vérifier que f est de classe $C^{1+\alpha}$. Pour cela, il suffit de montrer que les normes C^α des g_n sont uniformément bornées. Des estimations désagréables, mais élémentaires, montrent que pour n assez grand en module, la constante de Hölder C^α de g_n est bornée par $\frac{2\beta}{A}\|\varphi\|_{\mathrm{Lip}}$.

Ces exemples ont de plus la propriété, pour nous remarquable, que la dimension de boîte supérieure du Cantor K est exactement τ. À ce propos, nous rappelons que si (X, d) est un espace métrique compact et $Y \subset X$ une partie bornée, on définit la *dimension de boîte supérieure* de Y de la manière suivante : soit, pour tout $\varepsilon > 0$, $N(\varepsilon)$

le nombre minimal de boules de rayon ε nécessaires pour recouvrir Y ; alors la dimension de boîte supérieure $\overline{\dim}_B(Y)$ de Y est définie comme étant la limite supérieure

$$\overline{\dim}_B(Y) = \overline{\lim_{\varepsilon \to 0}} \frac{\log N(\varepsilon)}{\log(1/\varepsilon)}.$$

La limite à droite dans l'expression n'existe pas forcément. Si cette limite existe, on l'appelle plus simplement *dimension de boîte* de Y.

Il est connu (voir [Nor 2]) que pour une partie compacte Y sur le cercle $\mathbf{S}^1$, de mesure de Lebesgue nulle, la dimension de boîte supérieure de Y est égale à l'*infimum* des $s > 0$ tel que $\sum \mathrm{Leb}(I)^s < \infty$, où la somme est faite sur les composantes connexes I du complémentaire de Y dans $\mathbf{S}^1$.

Dans notre exemple, la somme $\sum \mathrm{Leb}(I)^s = \sum \ell_n^s$ est finie si et seulement si $s > \tau$, d'où la validité de notre affirmation. Plus généralement, Norton montre qu'il y a des contraintes pour qu'un Cantor dans $\mathbf{S}^1$ soit l'ensemble minimal invariant d'un difféomorphisme de classe $C^{1+\tau}$:

Théorème 3.1 (Norton)**.** *Soit f un difféomorphisme du cercle avec un ensemble minimal qui est un Cantor K, tel que la dimension de boîte supérieure de K soit égale à $\sigma \in]0,1[$. Alors f n'est pas de classe $C^{1+\tau}$ pour tout $\tau > \sigma$.*

Remarque 3.2. Il est intéressant en soi de savoir que le théorème de Norton peut se généraliser aux actions libres de $\mathbf{Z}^d$ par difféomorphismes sur $\mathbf{S}^1$. Pour ceci, on renvoie à l'appendice 5.

Observons que dans notre contexte aléatoire, si f est un difféomorphisme tiré au hasard par rapport à la mesure de Malliavin-Shavgulidze locale, et si l'on suppose les deux propriétés supplémentaires suivantes :

1. f possède un Cantor invariant K,
2. $Df = 1$ sur K,

alors, la dimension de boîte supérieure de K doit être exactement $1/2$.

En effet, la dimension de boîte des zéros d'un pont brownien est presque sûrement égale à $1/2$ (voir [Mör-Per]), et donc il en est de même pour tout ensemble de niveau non-vide. Il n'est pas alors possible d'exclure l'existence d'un ensemble de probabilité non-nulle de difféomorphisme qui vérifient les propriétés 1. et 2. ci-dessus.

Un raffinement célèbre de la dimension de boîte est la dimension de Hausdorff. Dans [Kra-Sch] le résultat de Norton est amélioré, sous la condition que Df soit égal à 1 partout sur le Cantor invariant.

Théorème 3.2 (Kra – Schmeling)**.** *Pour tout $\tau > 0$, si K est un Cantor minimal invariant pour un difféomorphisme du cercle $C^{1+\tau}$ et de nombre de rotation α qui satisfait une condition diophantienne d'ordre δ, alors la dimension de Hausdorff de K est supérieure ou égale à $\tau/(1+\delta)$.*

Par suite, des exemples aléatoires de difféomorphismes de Denjoy avec $Df = 1$ sur le Cantor sont possibles seulement si le nombre de rotation α est de type Roth. Ces nombres α forment une partie de mesure pleine par rapport à la mesure de Lebesgue, et donc nous ne rencontrons encore aucune contrainte.

4 Appendice : Simulations numériques

En commençant l'étude des difféomorphismes aléatoires, nous avons eu une sensation de dépaysement : quel type de difféomorphisme devons-nous imaginer ?

Nous avons alors eu recours à une aide « numérique » : l'un des avantages à travailler avec les mouvements browniens est en effet leur simplicité quand on souhaite faire des simulations !

Dans la suite, nous présentons quelques images, en rajoutant après les lignes de code qui nous ont permis de les obtenir : elles présentent principalement deux parties, une pour générer pseudo-aléatoirement un difféomorphisme qui suit la loi de Malliavin-Shavgulidze et une autre pour calculer le nombre de rotation. Le langage est le Python. Les simulations ont été effectuées sur les serveurs du PSMN de l'École Normale Supérieure de Lyon.

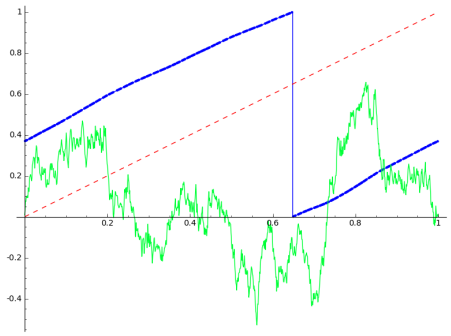

Figura II.4 – Paramètre $\sigma = 0.6$
Nombre de rotation $[0, 2, 1, 2, 2, 1, 1, 1, \ldots] \sim \frac{27}{73}$

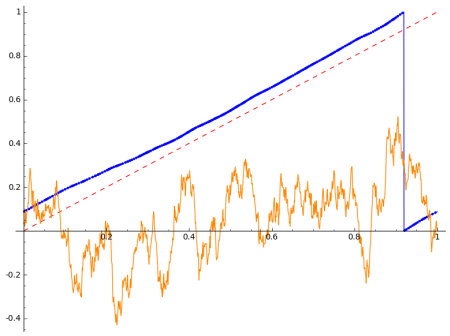

Figura II.5 – Paramètre $\sigma = 0.8$
Nombre de rotation $[0, 15, 3, 4, 1, 3, 1, 15, \ldots] \sim \frac{1216}{18619}$

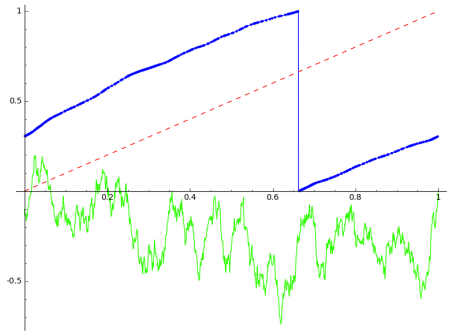

Figura II.6 – Paramètre $\sigma = 1.6$
Nombre de rotation $[0, 2, 1, 6, 2, 4, 1, 2, \ldots] \sim \frac{231}{662}$

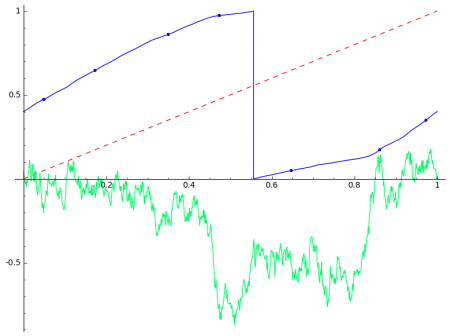

Figura II.7 – Paramètre $\sigma = 2.3$
Nombre de rotation $\frac{3}{7}$

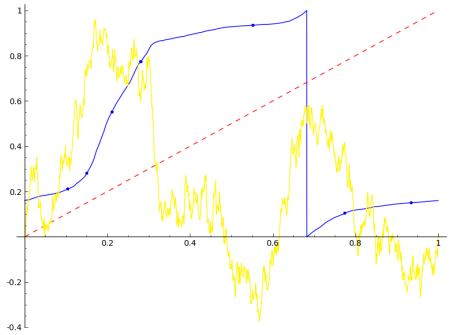

Figura II.8 – Paramètre $\sigma = 3.4$
Nombre de rotation $\frac{3}{7}$

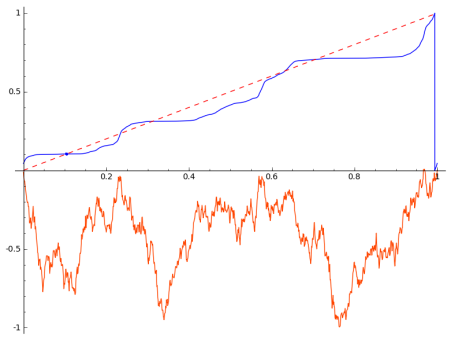

Figura II.9 – Paramètre $\sigma = 9.1$
Nombre de rotation 0

Dans les images nous avons tracé le graphe du difféomorphisme pseudo-aléatoire f, avec le pont brownien utilisé dans sa définition. Le graphe du pont brownien est colorié par une couleur entre 0 et 1 qui correspond à l'image du point 0 par f. Sur le graphe de f nous avons aussi marqué les points sur l'orbite de 0 : $f^k(0)$, $k \in \{200, \ldots, 499\}$.

Cet algorithme nous a permis de générer 1050 difféomorphismes aléatoires pour chaque loi μ_σ, avec $\sigma \in \{0.25, 0.5, \ldots, 7.75\}$, dont nous avons calculé le nombre de rotation. La figure II.10 se lit ainsi : chaque ligne correspond à un paramètre σ, pour lequel nous avons ordonné de manière croissante les nombres de rotation obtenus pour cette valeur. Pour distinguer les rationnels des irrationnels, nous avons reporté chaque nombre de rotation ρ par une couleur entre 0 et 1, qui correspond à la partie fractionnaire de ρ multiplié par un très grand nombre rationnel $(2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot 7 \cdot 8 \cdot 9 \cdot 10 \cdot 11 \cdot 13 \cdot 14 \cdot 15 \cdot 17 \cdot 19 \cdot 23)$.

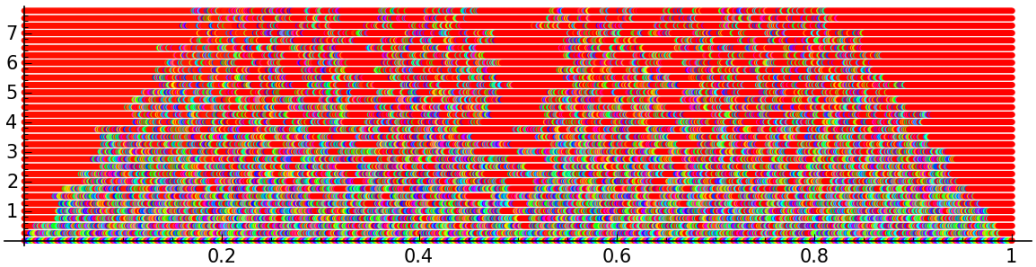

Figura II.10 – *Rouge rationnel*

Même si l'on observe beaucoup de points qui ne sont pas rouges (surtout pour des petites valeurs de σ), nous ne croyons pas que les nombres de rotation irrationnels aient μ_σ-mesure positive : il est numériquement impossible de distinguer un nombre de rotation *irrationnel* d'un nombre de rotation *rationnel* avec très grand dénominateur. Pour faire une comparaison avec ce qui se passe dans le même type de figure provenant de la famille d'Arnol'd (exemple 1.1), nous croyons que les graphes qui décrivent les bords des langues sont donnés par des fonctions non-Lipschitz mais plutôt 1/2-Hölder.

Dans la figure II.11 nous montrons les « escaliers du diable » que l'on trouve en lisant la figure II.10 ligne par ligne. On peut observer que pour $\sigma = 0.5$ la distribution du nombre de rotation est très proche de l'identité, ce qui suggérerait comme conjecture que pour des petites valeurs de σ, le fait d'avoir nombre de rotation irrationnel est un événement de mesure strictement positive.

Remarquons en outre que notre algorithme permet aussi de déterminer numériquement les distributions des $\{a_n\}_{n \leq L}$, lorsque $\sigma = 1$. Dans la figure II.12 on représente les fonctions de répartition des a_n pour $n = 1, \ldots, 6$, comparées avec la distribution classique de Gauss-Kuzmin. Il semblerait que les distributions soient presque les mêmes : un tel fait devrait impliquer que le nombre de rotation est rationnel presque sûrement. Cependant cette ressemblance est moins évidente pour des différentes valeurs de σ (dans la figure II.13 on peut voir les statistiques pour $\sigma = 2$).

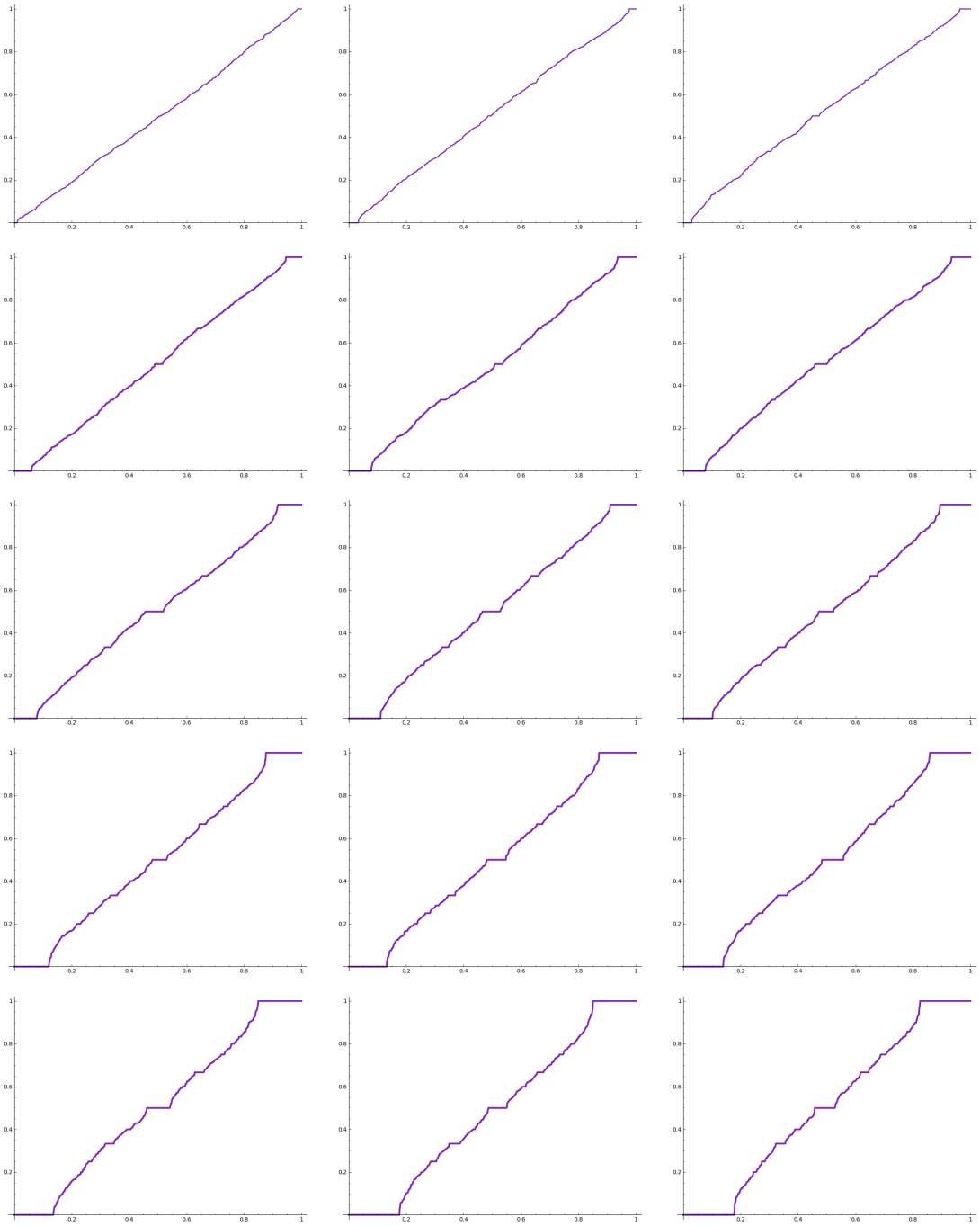

Figura II.11 – Distribution du nombre de rotation selon les mesures μ_σ, avec $\sigma \in \{0.5, 1, \ldots, 7.5\}$ (lecture ligne par ligne)

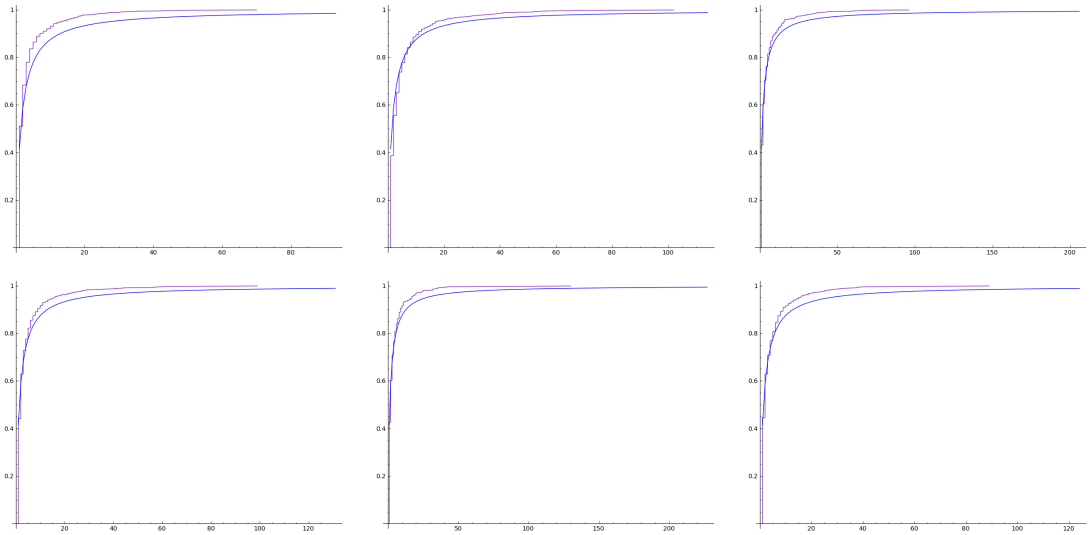

Figura II.12 – Fonctions des répartitions des a_n $(n = 1, \ldots, 6)$ lorsque $\sigma = 1$, comparées avec la distribution classique de Gauss-Kuzmin

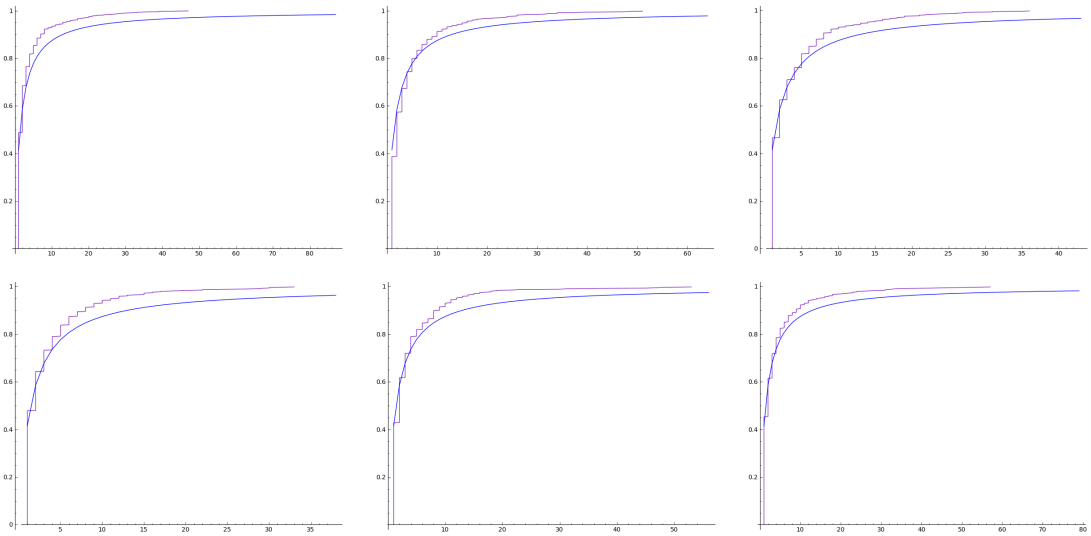

Figura II.13 – Fonctions des répartitions des a_n $(n = 1, \ldots, 6)$ lorsque $\sigma = 2$, comparées avec la distribution classique de Gauss-Kuzmin

```python
from __future__ import division,print_function
import numpy as np
from scipy.interpolate import interp1d
import matplotlib.pyplot as plt
import matplotlib.mlab as mlab
from math import exp,sqrt,floor

n = 10000 #points
K = 10 #tests
L = 8 #length of continued fraction expansion
A = 1500 #maximum size of single element in continued fraction expansion


def partfrac(x):
    return x-floor(x)

#it computes the rational number defined by a continued fraction expansion b
#of length l+1
def rational_approximation(b,l):
    p = [0,1]
    q = [1,b[1]]
    for i in range(1,l+1):
        p.append(b[i+1]*p[i]+p[i-1])
        q.append(b[i+1]*q[i]+q[i-1])
    return p[l+1]/q[l+1]

#this is for computing the rotation number of a given homeomorphism f:
#we use the classical algorithm
#(explained for example in "de Melo, van Strien : One-dimensional dynamics")
#it computes the continued fraction expansion of the rotation number
#and uses rational_approximation to return the L-th rational approximation
#of it (or it stops before if it finds a periodic orbit)
def rotation(f):
    a = [0]
    orbit = []
    orbit.append(partfrac(f(0)))
    if orbit[0]==0 :
        return 0

    def shift(x):
        if partfrac(x)>orbit[0]:
            return partfrac(x)-1
        return partfrac(x)

    def first_return(p,pre_p,y):
        x = shift(f(y))
        while x<pre_p or x>p:
            x = shift(f(x))
        return x

    a.append(1)
    x = orbit[0]

    if shift(f(orbit[0]))==0:
        return 1/2

    if shift(f(orbit[0]))<0:
        while shift(f(x))<0:
            a[1] = a[1]+1
            x = shift(f(x))
            if a[1]>A:
                return 0
            if shift(f(x))==0:
                a[1] = a[1]+1
                return 1/a[1]
        orbit.append(shift(x))
        z = shift(f(x))
        a.append(0)
        while z>0:
            y = z
            z = first_return(shift(orbit[0]),shift(orbit[1]),z)
            a[2] = a[2]+1
            if a[2]>A:
                return 1/a[1]
            if z==0:
                a[2] = a[2]+1
```

```
            return rational_approximation(a,1)
        orbit.append(y)

    if shift(f(orbit[0]))>0:
        def shift(y):
            if partfrac(y)>=orbit[0]:
                return partfrac(y)-1
            return partfrac(y)
        orbit.append(orbit[0]-1)
        a.append(0)
        while shift(f(x))>0:
            a[2] = a[2]+1
            x = shift(f(x))
            if a[2]>A:
                return 1/a[1]
            if shift(f(x))==0:
                a[2] = a[2]+1
                return rational_approximation(a,1)
        orbit.append(shift(x))
        z = shift(f(x))

    for i in range(1,L):
        a.append(0)
        if shift(orbit[i+1])<shift(orbit[i]):
            while z>0:
                y = z
                z = first_return(shift(orbit[i]),shift(orbit[i+1]),z)
                a[i+2] = a[i+2]+1
                if a[i+2]>A:
                    return rational_approximation(a,i)
                if z==0:
                    a[i+2]=a[i+2]+1
                    return rational_approximation(a,i+1)
        if shift(orbit[i+1])>shift(orbit[i]):
            while z<0:
                y = z
                z = first_return(shift(orbit[i+1]),shift(orbit[i]),z)
                a[i+2] = a[i+2]+1
                if a[i+2]>A:
                    return rational_approximation(a,i)
                if z==0:
                    a[i+2] = a[i+2]+1
                    return rational_approximation(a,i+1)
        orbit.append(y)

    return rational_approximation(a,L)


#this generates pseudo-random MS diffeomorphisms
def rotationalea(var):
    s = np.random.normal(0, var, n+2)
    ss = np.cumsum(s)
    T = np.random.rand(1,1)
    w = []
    for i in range(0,n+1):
        w.append(exp(1/sqrt(n)*(ss[i+1]-i*ss[n+1]/n)))
     #exponential of browian motion
    a = [0,w[1]]
    for j in range(2,n+1):
        a.append(a[j-1]+w[j]+w[j-1])
     #we integrate w
    p = []
    q = []
    for i in range(0,n+1):
        p.append(i/n)
        q.append(a[i]/a[n])
    g = interp1d(p,q)
     #linear intepolation on n+1 points
    def G(x):
        return g(x-floor(x)) + floor(x) + T
     #we compose by a uniform rotation T and the compute the rotation number
    return rotation(G)

#here is the "main" part: we simulate K diffeomorphisms of variance parameter h/4,
#where h ranges from 1 to 31 then it prints out a sorted vector
for h in range(1,2**5):
```

```
    v = []
    for k in range(0,K):
        v.append(rotationalea(h/2**3))

    print(h/4)
    print(sorted(v))
    print()
```

5 Appendice : Sur la dimension des ensembles de Cantor invariants par des groupes abéliens de difféomorphismes du cercle

Il est bien connu, d'après Denjoy [Den], que tout difféomorphisme du cercle de classe C^{1+vb} et sans point périodique est toujours minimal (i.e. toute orbite est dense). Il est aussi bien connu que pour tout $\varepsilon > 0$, il existe des difféomorphismes de classe $C^{2-\varepsilon}$ arbitrairement proches des rotations, sans point périodique et tels que l'adhérence de toute orbite soit un unique ensemble de Cantor (voir [Her 1, Chap. X]). Dans ce contexte, M. Herman fut le premier à demander quelles sont les contraintes pour qu'un ensemble de Cantor $K \subset \mathbf{S}^1$ soit l'ensemble minimal invariant d'un difféomorphisme de classe $C^{1+\tau}$. Certains résultats (voir par exemple [McD, Nor 1]) s'occupent des propriétés de similarité de K : en particulier le Cantor triadique classique n'est l'ensemble minimal d'aucun difféomorphisme de classe C^1. D'autres résultats concernent les dimensions fractales de ces objets. Le premier travail dans cette direction est le théorème de Denjoy « avec exposants » dû à A. Norton [Nor 2] (pour l'énoncé, voir le théorème 3.1). La proposition suivante est à la base de [Nor 2].

Proposition 5.1. *Si K est une partie compacte dans $\mathbf{S}^1$ et $\overline{\dim}_B(K) < s$, alors $G_s(K) < \infty$.*

Dans cette note, nous utiliserons cette proposition pour étendre le théorème de Norton aux actions de $\mathbf{Z}^d$ sur le cercle. Rappelons à ce propos, le théorème de B. Deroin, V. Kleptsyn et A. Navas [Der-Kle-Nav].

Théorème 5.2 (Deroin – Kleptsyn – Navas)**.** *Soit d un nombre entier supérieur ou égal à 2, soit $\varepsilon > 0$, alors toute action libre de $\mathbf{Z}^d$ sur le cercle par des difféomorphismes $C^{1+1/d+\varepsilon}$ est minimale.*

En effet, en regardant dans le détail la preuve de ce théorème, nous sommes capables de montrer un résultat « à la Norton ».

Théorème 5.3. *Soit d un nombre entier strictement positif. Si $\mathbf{Z}^d$ agit librement sur $\mathbf{S}^1$ avec ensemble minimal un Cantor K tel que $\overline{\dim}_B(K) = \sigma \in]0,1[$, alors l'action n'est pas de classe $C^{1+\tau/d}$ pour tout $\tau > \sigma$.*

Qui plus est, ce résultat est optimal : il existe des exemples d'actions libres non-minimales par des difféomorphismes de classe $C^{1+\sigma/d}$ et telles que la dimension de boîte supérieure de l'ensemble minimal soit exactement σ.

Comme nous l'avons déjà dit, la preuve du théorème n'introduit aucun élément vraiment nouveau. Il serait intéressant d'obtenir une généralisation analogue pour le théorème de Kra et Schmeling 3.2.

Dimostrazione. Pour montrer leur théorème, Deroin, Kleptsyn et Navas définissent une marche aléatoire sur le (semi-)groupe $\mathbf{Z}^d$ qui permet d'obtenir des estimations presque sûres extrêmement fondamentales.

Nous allons nous restreindre au cas $d = 2$, en laissant les autres cas au lecteur.

Soient g_1 et g_2 deux difféomorphismes de classe $C^{1+\tau}$ qui commutent. On note par Γ^+ le semi-groupe isomorphe à $\mathbf{N} \times \mathbf{N}$ engendré par g_1 et g_2. On notera par (m,n) l'élément $g_1^m\, g_2^n$.

La marche aléatoire sur Γ^+ est définie par le noyau de transition p :

$$\begin{cases} p\left((m,n) \to (m+1,n)\right) = \frac{m+1}{m+n+2} \\ p\left((m,n) \to (m,n+1)\right) = \frac{n+1}{m+n+2}. \end{cases}$$

Ce noyau induit une probabilité $\mathbf{P}$ sur l'espace des chemins Ω. Observons que la probabilité que la marche arrive en k pas sur un point de coordonnées (m,n) avec $m+n=k$, en partant de l'origine, est égale à $\frac{1}{k+1}$, et dépend donc uniquement de la somme des coordonnées k.

Soit $\omega \in \Omega$, on définit les éléments $h_n(\omega) \in \Gamma^+$ pour tout $n \in N$ par projection au pas n : $h_n(\omega) = g_1^i\, g_2^j$ si la marche ω se trouve au pas n en (i,j). Fixons $I \in C(K)$, où K est le Cantor invariant par l'action. Définissons, pour tous les $s \in]0,1]$ et $\omega \in \Omega$

$$\ell_s(\omega) := \sum_{n \geq 0} |h_n(\omega)(I)|^s.$$

Lemme 5.4. *Pour tous les $s \in]0,1]$ et $\sigma > 1/2$, il existe une constante C_σ telle que $\mathbf{E}[\ell_s] \leq C_\sigma\, G_{s/\sigma}(K)$.*

Dimostrazione. Nous avons

$$\mathbf{E}\left[\ell_s\right] = \mathbf{E}\left[\sum_{k\geq 0} |h_k(I)|^s\right] = \sum_{k\geq 0} \mathbf{E}\left[|h_k(I)|^s\right] = \sum_{k\geq 0} \sum_{m+n=k} \frac{|g_1^m g_2^n(I)|^s}{k+1}.$$

Pour tout $\sigma > 1/2$, on utilise l'inégalité de Hölder :

$$\sum_{m+n=k} \frac{|g_1^m g_2^n(I)|^s}{k+1} \leq \left[\sum_{m+n=k} |g_1^m g_2^n(I)|^{s/\sigma}\right]^\sigma \left[(k+1)\cdot \frac{1}{(k+1)^{1/(1-\sigma)}}\right]^{1-\sigma},$$

donc

$$\mathbf{E}[\ell_s] \leq \sum_{k\geq 0} \frac{\left[\sum_{m+n=k} |g_1^m\, g_2^n(I)|^{s/\sigma}\right]}{(k+1)^\sigma}.$$

Une deuxième application de l'inégalité de Hölder donne

$$\mathbf{E}[\ell_s] \leq \left[\sum_{(m,n)\in\mathbf{N}\times\mathbf{N}} |g_1^m\, g_2^n(I)|^{s/\sigma}\right]^\sigma \left[\sum_{k\geq 1}\left(\frac{1}{k^\sigma}\right)^{\frac{1}{1-\sigma}}\right]^{1-\sigma}.$$

Puisque $\sigma > 1/2$, la quantité

$$C_\sigma := \left[\sum_{k\geq 1}\left(\frac{1}{k^\sigma}\right)^{\frac{1}{1-\sigma}}\right]^{1-\sigma}$$

est finie, donc $\mathbf{E}[\ell_s] \leq C_\sigma \cdot G_{s/\sigma}(K)^\sigma$. □

Si la quantité $G_{s/\sigma}(K)$ est finie, la preuve du théorème de Deroin, Kleptsyn et Navas montre que l'action de $\mathbf{Z}^2$ doit forcement avoir un point fixe.

Si l'action n'est pas minimale mais libre, on doit avoir $G_{s/\sigma}(K) = +\infty$ et donc, d'après la proposition, $\overline{\dim}_B(K) \geq s/\sigma$. Puisque cette minoration est valable pour tout $\sigma > 1/2$, on a $\overline{\dim}_B(K) \geq 2\,s$.

Pour terminer la preuve, montrons que le résultat est optimal. En fait, comme pour le théorème de Norton, les exemples de Denjoy classiques marchent.

Rappelons qu'il est possible de construire des exemples d'actions non-minimales de $\mathbf{Z}^d$ qui possèdent comme ensemble invariant le Cantor K dont les composantes connexes du complémentaire $\mathbf{S}^1 - K$ ont pour longueurs

$$\ell_{n_1,\ldots,n_d} = \frac{A}{(|n_1| + \ldots + |n_d| + M)^{d/\tau}}$$

où M est un nombre entier suffisamment grand et A est choisi de sorte que $\sum \ell_{n_1,\ldots,n_d} = 1$.

Alors $\sum (\ell_{n_1,\ldots,n_d})^s < \infty$ si et seulement si $s > \tau$. □

Capitolo III

Équation cohomologique et distributions invariantes d'ordre 1

1 Mesures invariantes et cobords

En dynamique, il est souvent très important d'étudier les mesures qui sont préservées par l'action. Connaître quelles sont les *mesures invariantes* par un homéomorphisme est en premier lieu un problème de nature *topologique* : si f agit sur un espace compact métrisable X, le théorème de Krylov-Bogoliubov affirme qu'il existe toujours une mesure de probabilité invariante, puisqu'elle peut être obtenue comme valeur d'adhérence de la suite

$$\frac{1}{n}\sum_{k=0}^{n-1}\delta_{f^k(x)}, \quad \text{pour } x \in X$$

dans la topologie faible-$*$.

Nous noterons $\mathcal{P}(X, f)$ l'ensemble convexe des mesures de probabilité invariantes par f. Les points extrémaux de $\mathcal{P}(X, f)$ sont les mesures de probabilité *ergodiques* pour l'action ; le théorème de Krein-Milman assure qu'il existe toujours une mesure ergodique.

L'ensemble $\mathcal{P}(\mathbf{S}^1, f)$ est très bien compris lorsque f est un homéomorphisme du cercle avec nombre de rotation irrationnel : il existe une *unique* mesure de probabilité invariante, supportée sur l'ensemble compact minimal invariant par l'action. Quand $f = R_\alpha$ est une rotation d'angle irrationnel, cette affirmation est un exercice que nous montrerons tout de suite (le cas général étant moins évident, nous donnons [Her 1, proposition II.8.5] en référence).

Définition 1.1. *Soit $f : X \longrightarrow X$ un homéomorphisme d'un espace métrique compact non vide X. On dit que f est* uniquement ergodique *si l'espace des mesures de probabilité invariantes $\mathcal{P}(X, f)$ est formé d'un seul point : il existe une unique mesure de probabilité invariante, qui est ergodique.*

Le fait qu'un homéomorphisme f soit uniquement ergodique se caractérise topologiquement :

Proposition 1.1. *Les affirmations suivantes sont équivalentes :*

1. *f est uniquement ergodique, avec mesure de probabilité invariante μ ;*
2. *le sous-espace $\overline{\{u \circ f - u,\ u \in C^0(X)\}}$ est de codimension 1 dans $C^0(X)$;*
3. *pour tout $h \in C^0(X)$, la suite des moyennes de Birkhoff $\left(\frac{1}{k}S_k h\right)_{k\in\mathbf{N}}$ converge uniformément vers une constante (égale à $\int_X h\, d\mu$).*

Par le théorème de Riesz, une mesure invariante correspond à une fonctionnelle linéaire L sur $C^0(\mathbf{S}^1)$ qui vérifie $L(u \circ f - u) = 0$. Une application simple du théorème de Hahn-Banach montre que l'adhérence de $\{u \circ f - u,\ u \in C^0(\mathbf{S}^1)\}$ est égale à l'intersection des noyaux de toutes les fonctionnelles f-invariantes :

$$\overline{\{u \circ f - u,\ u \in C^0(\mathbf{S}^1)\}} = \bigcap_{\mu\in\mathcal{P}(\mathbf{S}^1, f)} \ker \mu.$$

Pour montrer l'unicité de la mesure invariante, il faut et il suffit de montrer que le sous-espace fermé $\overline{\{u \circ f - u,\ u \in C^0(\mathbf{S}^1)\}}$ est de codimension 1 dans $C^0(\mathbf{S}^1)$. Selon une terminologie commune dans la littérature, un élément $u \circ f - u$ est un *cobord* (et toute fonction $h \in C^0(\mathbf{S}^1)$ un *cocycle*).

Exemple 1.1. Quand $f = R_\alpha$ est une rotation irrationnelle, tout cocycle est approché par une suite de cobords à une constante additive près. La preuve, classique, est un exercise assez instructif.

Soit u une fonction continue sur le cercle et $(\hat{u}(k))_{k\in\mathbf{Z}}$ ses coefficients de Fourier. La composition par la rotation R_α préserve les harmoniques :

$$\widehat{u \circ f}(k) = e^{2\pi i k\alpha}\hat{u}(k),$$

comme un simple calcul de changement de variable le montre.

Par le théorème de Plancherel, il suffit de chercher pour quelles suites $(\lambda_k)_{k\in\mathbf{Z}}$ le produit scalaire

$$\sum_{k\in\mathbf{Z}} \left(e^{2\pi i k\alpha} - 1\right)\hat{u}(k)\lambda_k$$

vaut 0 pour toute fonction $(\hat{u}(k))_{k\in\mathbf{Z}}$.

On en déduit alors que seul le coefficient λ_0 peut ne pas s'annuler : en d'autres mots, la suite $(\lambda_k)_{k\in\mathbf{Z}}$ correspond dans $L^2(\mathbf{S}^1, \mathrm{Leb})$ à la fonction constante λ_0 (qui se trouve en particulier dans $C^0(\mathbf{S}^1)$) : nous pouvons conclure que toute fonctionnelle f-invariante est un multiple scalaire de l'intégration sur le cercle par rapport à la mesure de Lebesgue.

2 Distributions invariantes par les difféomorphismes avec nombre de rotation irrationnel

Beaucoup moins est connu, en général, sur ce qui concerne les *distributions invariantes* : si f est un difféomorphisme C^r d'une variété M, quelles sont les fonctionnelles linéaires f-invariantes sur $C^r(M)$? Est-il possible de trouver des distributions invariantes qui sont différentes des mesures invariantes ?

Les résultats à ce propos sont très récents et ils sont beaucoup moins précis lorsque l'on abandonne le cercle pour une autre variété. L'une des contributions majeures a été apportée par A. Ávila et A. Kocsard [Ávi-Koc 1], avec le théorème suivant qui donne une compréhension assez exhaustive des distributions invariantes par des difféomorphismes du cercle *lisses* et minimaux.

Théorème 2.1 (Ávila – Kocsard). *Soit f un difféomorphisme du cercle de classe C^∞ et nombre de rotation irrationnel. Alors, toute distribution f-invariante est un multiple de la mesure invariante μ_f.*

En d'autres termes, toute fonction de classe C^∞ et μ_f-moyenne nulle, peut être approchée dans la topologie C^∞ par une suite de cobords C^∞.

Le résultat suivant, dû à l'auteur en collaboration avec A. Navas [Nav-Tri], est l'analogue du théorème 2.1 en classe de régularité plus faible. La preuve est par contre d'un esprit plus conceptuel et géométrique. Elle s'appuie sur le travail de R. Douady et J.-C. Yoccoz, relatif aux mesures automorphes pour les difféomorphismes du cercle [Dou-Yoc].

Théorème 2.2. *Soit f un difféomorphisme du cercle de classe C^{1+vb} et nombre de rotation irrationnel. Alors, toute distribution d'ordre 1 qui est f-invariante est un multiple de la mesure invariante μ_f.*

Le théorème précédent permet d'obtenir une version améliorée de l'*inégalité de Denjoy-Koksma* (théorème 3.5 du chapitre introductif) pour des fonctions test C^1, en étendant aux difféomorphismes de classe C^{1+vb} le corollaire C de [Ávi-Koc 1] (valable pour les difféomorphismes de classe C^{11}).

Corollaire 2.3. *Soit f un difféomorphisme du cercle de classe C^{1+vb} et nombre de rotation irrationnel α. On note par (p_n/q_n) la suite des approximations rationnelles de α. Alors, pour toute fonction $u \in C^1(\mathbf{S}^1)$ la suite*

$$S_{q_n}u - q_n \int_{\mathbf{S}^1} u\, d\mu_f$$

converge uniformément vers 0.

La preuve suit exactement les mêmes lignes de [Ávi-Koc 1], le seul ingrédient nécessaire étant l'unicité de la distribution invariante d'ordre 1.

Dimostrazione. Il suffit de montrer que pour toute fonction u de classe C^1 et μ_f-moyenne nulle, la suite $(S_{q_n}u)_{n\in\mathbf{N}}$ converge uniformément vers 0.

D'après le théorème 2.2, pour tout $\varepsilon > 0$ il est possible de trouver une fonction $w \in C^1(\mathbf{S}^1)$ telle que

$$\left\|u - (w \circ f - w)\right\|_1 < \frac{\varepsilon}{2}.$$

On pose, pour $n \in \mathbf{N}$ donné

$$M_n := \sup_{x \in \mathbf{R}} |F^{q_n}(x) - x - p_n|,$$

où F est un difféomorphisme qui relève f à la droite réelle $\mathbf{R}$.

D'après la minimalité de f, nous avons que M_n tend vers 0, donc il est possible de trouver $N \in \mathbf{N}$ tel que pour tout $n > N$, on a $\|Dw\|_0 M_n < \varepsilon/2$. Soit donc $n > N$,

$$\left\|S_{q_n} u\right\|_0 \leq \left\|S_{q_n}\left(u - (w \circ f - w)\right)\right\|_0 + \left\|w \circ f^{q_n} - w\right\|_0.$$

Le terme $\left\|S_{q_n}\left(u - (w \circ f - w)\right)\right\|_0$ peut être estimé grâce à l'inégalité de Denjoy-Koksma (22) :

$$\left\|S_{q_n}\left(u - (w \circ f - w)\right)\right\|_0 \leq \operatorname{Var}\left(u - (w \circ f - w)\right) \leq \int_{\mathbf{S}^1} D\left(u - (w \circ f - w)\right) < \varepsilon/2.$$

Pour le deuxième terme on utilise le théorème des accroissements finis :

$$\left\|w \circ f^{q_n} - w\right\|_0 \leq \|Dw\|_0 M_n < \varepsilon/2.$$

Finalement on trouve que pour $n > N$ on a $\left\|S_{q_n}\right\| < \varepsilon$. □

Remarque 2.1. Lorsque f est de classe C^2 et $u = \log Df$, le corollaire 2.3 donne le résultat bien connu de M. Herman

Proposition 2.4. *La suite $(f^{q_n})_{n \in \mathbf{N}}$ converge vers l'identité dans la topologie C^1.*

□

Observons que l'hypothèse que f soit de classe C^{1+vb} est fondamentale

Théorème 2.5. *Pour tout nombre irrationnel α, il existe un difféomorphisme C^1 avec nombre de rotation α qui possède une distribution d'ordre 1 invariante qui n'est pas proportionnelle à la mesure invariante.*

En plus, le difféomorphisme peut être choisi minimal ou bien avec un ensemble de Cantor comme compact minimal invariant.

2.1 Unicité des distributions d'ordre 1 pour les difféomorphismes C^{1+vb} minimaux

Comme d'habitude, on notera $\mu = \mu_f$ l'unique mesure de probabilité f-invariante.

Comme on l'a déjà pu expliquer au point précédent, pour montrer l'unicité de la distribution d'ordre 1 invariante, il faut et il suffit de montrer que toute fonction C^1 de μ-moyenne nulle peut être approchée par des cobords dans la topologie C^1.

À ce propos, on passera par la proposition suivante

Proposition 2.6. *Soit h une fonction de classe C^1 sur le cercle. Il existe une suite de fonctions continues w_n telle que la suite*

$$(w_n \circ f)Df - w_n$$

converge uniformément vers Dh et pour tout n,

$$\int_{\mathbf{S}^1} w_n \, d\mathrm{Leb} = 0. \tag{III.1}$$

En effet, par intégration de 0 à x, on obtient que la fonction $h(x) - h(0)$ est approchée par la suite de fonctions continues définies par

$$x \longmapsto \int_0^x (w_n \circ f)Df \, d\mathrm{Leb} - \int_0^x w_n \, d\mathrm{Leb} = \int_0^{f(x)} w_n \, d\mathrm{Leb} - \int_0^x w_n \, d\mathrm{Leb} - \int_0^{f(0)} w_n \, d\mathrm{Leb}.$$

On pose

$$u_n(x) = \int_0^x w_n \, d\mathrm{Leb},$$

qui, d'après (III.1) est une fonction bien définie. Alors la fonction h est approchée dans la topologie C^1 par la suite $u_n \circ f - u_n + c_n$, où c_n est une constante :

$$c_n = h(0) - \int_0^{f(0)} w_n \, d\mathrm{Leb}.$$

Par intégration par rapport à μ, on peut conclure que c_n doit forcement converger vers 0, et donc $u_n \circ f - u_n$ est la suite de cobords cherchée pour montrer le théorème 2.2.

Pour montrer la proposition 2.6 il est crucial de passer par le cas particulier suivant (correspondant à la fonction de déplacement $f - id - R_{\rho(f)}$).

Lemme 2.7. *Il existe une suite de fonctions continues $\hat{w}_k$ telle que*

$$(\hat{w}_k \circ f)Df - \hat{w}_k$$

converge uniformément vers $Df - 1$ et pour tout k

$$\int_{\mathbf{S}^1} \hat{w}_k \, d\mathrm{Leb} = 0.$$

Dimostrazione. Soit

$$\hat{w}_k = -\frac{1}{q_k} \sum_{i=0}^{q_k-1} Df^i + 1,$$

où (q_k) dénote la suite des dénominateurs des approximations rationnelles du nombre de rotation de f. Observons que pour tout $k \geq 1$,

$$\int_{\mathbf{S}^1} \hat{w}_k \, d\mathrm{Leb} = 0.$$

En outre,

$$\begin{aligned}
\hat{w}_k\left(f(x)\right) Df(x) &= -\frac{Df(x)}{q_k} \sum_{i=0}^{q_k-1} Df^k(f(x)) + Df(x) \\
&= -\frac{1}{q_k} \sum_{i=1}^{q_k} Df^k(x) + Df(x) \\
&= \hat{w}_k(x) - 1 + \frac{1}{q_k}\left[1 - Df^{q_k}(x)\right] + Df(x),
\end{aligned}$$

d'où

$$Df - 1 = (\hat{w}_k \circ f)Df - \hat{w}_k + \frac{1}{q_k}\left[Df^{q_k} - 1\right].$$

La convergence désirée suit alors de l'inégalité de Denjoy:

$$\left\|Df^{q_k}\right\|_0 \leq e^V$$

où V est la variation totale du logarithme de Df. □

Revenons maintenant à la proposition 2.6. Pour la preuve, rappelons que pour tout $s \in \mathbf{R}$, une *mesure s-automorphe* pour f est une mesure de probabilité ν sur le cercle telle que pour toute fonction continue φ,

$$\int_{\mathbf{S}^1} \varphi \, d\nu = \int_{\mathbf{S}^1} (\varphi \circ f)(Df)^s \, d\nu.$$

Lorsque f est un difféomorphisme de classe C^{1+vb} et de nombre de rotation irrationnel, un théorème de R. Douady et J.-C. Yoccoz assure qu'une telle mesure existe et elle est unique pour chaque s (voir [Dou-Yoc]). En outre, on a l'unicité (à facteur scalaire près) même dans le contexte des mesures finies avec signe (c'est-à-dire fonctionnelles linéaires sur l'espace des fonctions continues). Lorsque $s = 1$, l'unicité entraîne que la mesure 1-automorphe est la mesure de Lebesgue (pour ce cas particulier, nous donnerons une preuve dans la partie 2.2).

Or, puisque

$$\int_{\mathbf{S}^1} Dh \, d\mathrm{Leb} = 0,$$

une application classique du théorème de Hahn-Banach (à la fonctionnelle $w \mapsto (w \circ f)Df - w$) donne une suite de fonctions continues $\bar{w}_n$ telle que

$$(\bar{w}_n \circ f)Df - \bar{w}_n$$

converge uniformément vers Dh. En effet, supposons qu'une telle suite n'existe pas. Alors Dh n'appartient pas à l'adhérence du sous-espace linéaire des fonctions de la forme $(w \circ f)Df - w$. Puisqu'un sous-espace est

convexe, il existe une fonctionnelle linéaire L qui s'annule identiquement sur ce sous-espace et qui vaut 1 au point Dh. En d'autres termes, L est, comme mesure avec signe, une mesure 1-automorphe et donc un multiple non nul de la mesure de Lebesgue. Mais ceci ne peut pas être possible puisque $L(Dh) = 1$ et Dh est de moyenne nulle par rapport à la mesure de Lebesgue.

Le problème qu'on a avec la suite $\bar{w}_n$ est qu'on ne peut pas assurer que

$$c_n = \int_{\mathbf{S}^1} \bar{w}_n \, d\mathrm{Leb}$$

vaut 0. Pour dépasser ce souci, on considère les fonctions $\widetilde{w}_n = \bar{w}_n - c_n$, qui par définition ont moyenne nulle (par rapport à la mesure de Lebesgue). Étant donné $\varepsilon > 0$, nous pouvons choisir n tel que la valeur absolue de la différence entre

$$(\widetilde{w}_n \circ f)Df - \widetilde{w}_n$$

et

$$Dh - c_n(Df - 1)$$

soit inférieure ou égale à $\varepsilon/2$. En outre, par le lemme 2.7 on peut trouver $k = k_n$ tel que

$$c_n \left\| (Df - 1) - [(\hat{w}_k \circ f)Df - \hat{w}_k] \right\|_0 \leq \frac{\varepsilon}{2}.$$

Par conséquent, on trouve

$$\left\| Dh - [(\widetilde{w}_n \circ f - c_n \hat{w}_k \circ f)Df - (\widetilde{w}_n - c_n \hat{w}_k)] \right\|_0 \leq \varepsilon,$$

qui nous permet de montrer la proposition 2.6 grâce à l'égalité

$$\int_{\mathbf{S}^1} (\widetilde{w}_n - c_n \hat{w}_k) \, d\mathrm{Leb} = 0.$$

□

2.2 Appendice : Le théorème de Douady et Yoccoz pour $s = 1$

En fait, dans le cas qui nous intéresse ($s = 1$), le résultat de Douady et Yoccoz se démontre de manière assez « classique ». Puisque l'existence d'une mesure 1-automorphe est assurée par la mesure de Lebesgue, il ne faut montrer que son unicité (pour des autres valeurs de s, dans [Dou-Yoc] l'existence est obtenue par une brillante méthode de point fixe).

On démontrera que toute mesure de probabilité 1-automorphe est ergodique : ceci suffit, car l'ensemble de ces mesures est un convexe compact (pour la topologie faible-*) et on peut appliquer le théorème de Krein-Milman.[1]

Proposition 2.8. *Soit f un difféomorphisme de classe C^{1+vb} et de nombre de rotation irrationnel et soit ν une mesure de probabilité σ-finie 1-automorphe pour f. Alors f est ν-ergodique.*

Dimostrazione. La preuve dans [Dou-Yoc] suit le célèbre argument de Katok (voir par exemple la discussion dans [Her 1, Chapitre VII]).

Soit $B \subset \mathbf{S}^1$ un borélien f-invariant et supposons $\nu(B) < 1$. On veut montrer que B est de mesure nulle. Pour combiner *mesure* et *dynamique de renormalisation* (réduction de l'étude de la dynamique à un petit voisinage d'un point), on fait appel au théorème de densité de Lebesgue classique : pour presque tout point $x \in B$, la limite des rapports des mesures

$$\lim_{n \to 0} \frac{\nu(B \cap I_n(x))}{I_n(x)}$$

tend vers 1, lorsque $\{I_n(x)\}$ est une suite décroissante de voisinages de x telle que

$$\bigcap_{n \in \mathbf{N}} I_n(x) = \{x\}.$$

Par passage au complémentaire, on a immédiatement

$$\lim_{n \to 0} \frac{\nu(B \cap I_n(x))}{\nu(I_n(x))} = 0 \tag{III.2}$$

[1] Montrer que les points extrémaux sont les mesures de probabilité ergodiques se fait de manière classique.

pour presque tout $x \in \mathbf{S}^1 - B$.

Il est naturel de choisir $I_n(x) = [x, f^{q_n}(x)]$ (nous supposons par simplicité que n est pair). Avec ce choix, il est classique que le cercle $\mathbf{S}^1$ est recouvert par la *partition dynamique*

$$\mathbf{S}^1 \subset \bigcup_{k=0}^{q_{n+1}-1} f^k(I_n(x)) \cup \bigcup_{k=0}^{q_n-1} f^k(I_{n+1}(x)).$$

Donc

$$\nu(B) \leq \sum_{k=0}^{q_{n+1}-1} \nu(B \cap f^k(I_n(x))) + \sum_{k=0}^{q_n-1} \nu(B \cap f^k(I_{n+1}(x))). \tag{III.3}$$

Puisque $f \in C^{1+vb}$, on a l'estimation (encore classique) de distorsion : pour tout $y \in I_n(x)$ et $k < q_{n+1}$ (et aussi pour $y \in I_{n+1}$, $k < q_n$), on a

$$|\log f^k(x) - \log f^k(y)| \leq V = \mathrm{Var}(\log Df).$$

On en déduit, grâce à l'invariance de B, l'inégalité

$$\begin{aligned} \nu(B \cap f^k(I_n(x))) &= \int_{\mathbf{S}^1} \mathbf{1}_{B \cap I_n(x)} \circ f^{-k}\, d\nu \\ &= \int_{\mathbf{S}^1} \mathbf{1}_{B \cap I_n(x)} \cdot Df^k\, d\nu \\ &\leq Df^k(x) e^V \cdot \nu(B \cap I_n(x)) \end{aligned}$$

et de manière similaire

$$\nu(f^k(I_n(x))) \geq Df^k(x) e^{-V} \cdot \nu(I_n(x)),$$

donc

$$Df^k(x) \leq e^V \frac{\nu(f^k(I_n(x)))}{\nu(I_n(x))}. \tag{III.4}$$

L'inégalité (III.3) devient

$$\nu(B) \leq e^V\, \nu(B \cap I_n(x)) \sum_{k=0}^{q_{n+1}-1} Df^k(x) + e^V\, \nu(B \cap I_{n+1}(x)) \sum_{k=0}^{q_n-1} Df^k(x)$$

et en utilisant (III.4) :

$$\begin{aligned} \nu(B) &\leq e^{2V} \frac{\nu(B \cap I_n(x))}{\nu(I_n(x))} \sum_{k=0}^{q_{n+1}-1} \nu(f^k(I_n(x))) + e^{2V} \frac{\nu(B \cap I_{n+1}(x))}{\nu(I_{n+1}(x))} \sum_{k=0}^{q_n-1} \nu(f^k(I_{n+1}(x))) \\ &\leq e^{2V} \left(\frac{\nu(B \cap I_n(x))}{\nu(I_n(x))} + \frac{\nu(B \cap I_{n+1}(x))}{\nu(I_{n+1}(x))} \right). \end{aligned}$$

D'après (III.2), pour presque tout $x \in \mathbf{S}^1 - B$, le terme de droite tend vers 0, donc $\nu(B) = 0$. □

2.3 Appendice : Quelque remarque sur les suites qui approximent les cocycles [2]

Nous avons expliqué, dans l'introduction à ce chapitre, comme l'équation cohomologique est reliée à l'espace des mesures invariantes $\mathcal{P}(X, f)$ en utilisant le théorème de Hahn-Banach.

Il existe une méthode alternative, et plus explicite, qui donne une suite de cobords qui approchent à un cocycle h donné (de μ-moyenne nulle pour toute $\mu \in \mathcal{P}(X, f)$). À savoir, on pose

$$u_k = -\frac{1}{k} \sum_{i=0}^{k-1} S_i h \tag{III.5}$$

et on trouve l'identité remarquable

$$u_k \circ f - u_k = h - \frac{1}{k} S_k h\,.$$

Par la proposition 1.1, nous savons que la suite $\frac{1}{k} S_k h$ tend uniformément vers 0, donc la suite u_k convient.

[2] Cette partie est essentiellement une retranscription de la note [Cro-Nav].

Il est alors naturel dans ce contexte de se poser le problème suivant. Soit X une variété et f un difféomorphisme de classe C^1 ; pour toute fonction h de classe C^1, telle que toute distribution f-invariante d'ordre 1 s'annule sur h, on sait qu'il est possible de trouver une suite de cobords C^1, qui se rapprochent à h dans la topologie C^1. Existent-ils, alors, des sous-suites de la suite $(u_k \circ f - u_k)$ (définie par (III.5)) qui convergent vers h dans la topologie C^1 ? De manière équivalente, la suite $\frac{1}{k}D(S_k h)$, converge-t-elle vers 0, au moins le long d'une sous-suite ?

Une réponse négative arrive dans le cas que nous avons affronté, quand X est le cercle $\mathbf{S}^1$ et f est un difféomorphisme minimal de classe C^{1+vb}. Le théorème 2.2 assure que l'hypothèse que l'on doit faire sur h est simplement que sa μ_f-moyenne s'annule.

Théorème 2.9 (Crovisier – Navas). *Soit α un nombre irrationnel. Pour f générique dans $\mathrm{Diff}^\infty_{+,\alpha}(\mathbf{S}^1)$ et h générique dans $\{h \in C^\infty(\mathbf{S}^1) \,|\, \mu_f(h) = 0\}$, la suite des normes C^1 des cobords $u_k \circ f - u_k$ associé à h ne converge pas.*

Par contre, pour tout $f \in \mathrm{Diff}^\infty_{+,\alpha}(\mathbf{S}^1)$, il existe une fonction h remarquable pour laquelle on peut trouver toujours une sous-suite le long de laquelle la suite converge : c'est la fonction de *déplacement* h définie par $h(x) = f(x) - x$. En effet, dans ce cas,

$$u_k(x) = x - \frac{1}{k} S_{k+1} id(x),$$

d'où

$$u_k(f(x)) - u_k(x) = h(x) - \frac{f^{k+1}(x) - x}{k}.$$

Par conséquent, si (q_n) dénote la suite des dénominateurs des approximations rationnelles de α, on trouve que $u_{q_n} \circ f - u_{q_n}$ converge vers h dans la topologie C^1, grâce à l'inégalité de Denjoy.

Remarque 2.2. On explique ainsi l'apparente nécessité de l'astuce utilisée dans la preuve de la proposition 2.6 en passant par le lemme 2.7. Il est probable qu'aucune manière plus directe n'existe pour montrer cette proposition.

Remarque 2.3. Il y a un autre cas pour lequel on peut assurer la convergence sans trop de difficultés, notamment quand le difféomorphisme f est C^1-conjugué à une rotation R. En effet, si l'on note h la conjugaison entre f et R,

$$f = h^{-1} \circ R \circ h,$$

de sorte que la mesure de probabilité f-invariante soit $\mu = \mathrm{Leb}/Dh$, alors pour tout $\varphi \in C^1(\mathbf{S}^1)$ et tout $x \in \mathbf{S}^1$ on a

$$\frac{D(S^k\varphi)(x)}{k} = \frac{1}{k}\sum_{i=0}^{k-1} D\varphi(f^i(x))Df^i(x) = \frac{1}{k}\sum_{i=0}^{k-1} D\varphi(f^i(x))Dh^{-1}(R^i h(x))Dh(x) = \frac{Dh(x)}{k}\sum_{k=0}^{k-1}\frac{D\varphi(f^i(x))}{Dh(f^i(x))}.$$

De la dernière égalité on observe que la suite $\left(\frac{D(S^k\varphi)}{k}\right)_{k\in\mathbf{N}}$ converge vers

$$Dh\int_{\mathbf{S}^1}\frac{D\varphi}{Dh}\,d\mu = Dh\int_{\mathbf{S}^1} D\varphi\, d\mathrm{Leb} = 0.$$

On obtiendra le théorème 2.9 par conséquence directe de la proposition suivante.

Proposition 2.10. *Pour $f \in \mathrm{Diff}^\infty_{+,\alpha}(\mathbf{S}^1)$ générique et $x_0 \in \mathbf{S}^1$ donné, la suite des moyennes*

$$\frac{1}{q_n}\sum_{i=0}^{q_n-1} Df^i(x_0) \tag{III.6}$$

diverge quand n tend vers l'infini. Plus précisément, on peut demander que pour tout $n > 1$,

$$\frac{1}{q_n}\sum_{i=0}^{q_n-1} Df^i(x_0) \geq \frac{n}{3}. \tag{III.7}$$

Dimostrazione. On construit d'abord un difféomorphisme f explicit qui satisfait à l'énoncé du théorème. Comme il est classique dans ces cas, on utilisera la méthode d'Anosov-Katok, et donc f apparaitra comme étant la limite d'une suite

$$f_n = h_n \circ R_{\frac{p_n}{q_n}} \circ h_n^{-1},$$

avec h_n difféomorphismes C^∞ qui fixent x_0 et $\frac{p_n}{q_n}$ nombres rationnels, définis de manière inductive.

Pour $n = 1$, on pose $h_1 = id$ et $\frac{p_1}{q_1} = 0$, donc $f_1 = id$. En ayant défini f_n, h_n et $\frac{p_n}{q_n}$, on pose $h_{n+1} = A_{n+1} \circ h_n$, où A_{n+1} est un difféomorphisme C^∞ qui vérifie :

1. A_{n+1} commute avec f_n et fixe x_0,
2. sa dérivée au point x_0 est égale à $DA_{n+1}(x_0) = \frac{C_n}{n}$ où $C_n = \inf_{x,y\in\mathbf{S}^1} \frac{Dh_n(x)}{Dh_n(y)}$,
3. il coïncide avec l'identité en dehors du ε_n-voisinage $I_1 \cup \ldots \cup I_{q_n}$ de l'orbite de x_0 par $R_{\frac{p_n}{q_n}}$, où ε_n est choisi de sorte que chaque intervalle $J_i = h_n^{-1}(I_j)$ est de longueur inférieure ou égale à $\frac{1}{nq_n}$.

Après avoir fixé un tel A_{n+1}, on pose $\frac{p_{n+1}}{q_{n+1}}$ tel que $q_{n+1} \geq 5q_n$ et les normes C^n de

$$f_{n+1}^i - f_n^i = A_{n+1} \circ h_n \circ R_{i\frac{p_{n+1}}{q_{n+1}}} \circ h_n^{-1} \circ A_{n+1}^{-1} - A_{n+1} \circ h_n \circ R_{i\frac{p_n}{q_n}} \circ h_n^{-1} \circ A_{n+1}^{-1}$$

et

$$f_{n+1}^{-i} - f_n^{-i} = A_{n+1} \circ h_n \circ R_{-i\frac{p_{n+1}}{q_{n+1}}} \circ h_n^{-1} \circ A_{n+1}^{-1} - A_{n+1} \circ h_n \circ R_{-i\frac{p_n}{q_n}} \circ h_n^{-1} \circ A_{n+1}^{-1}$$

soient inférieures à 2^{-n-1} pour tout $i = 0, \ldots, q_n$. Alors pour tous les entiers $n \geq m \geq r > 0$ et tout $i = 0, \ldots, q_n$,

$$\|f_n^i - f_m^i\|_r < \sum_{i=m}^{n-1} \frac{1}{2^{i+1}} \leq \frac{1}{2^m}, \quad \|f_m^{-i} - f_n^{-i}\|_r < \frac{1}{2^m},$$

ce qui montre en particulier que la suite (f_n) converge vers un difféomorphisme f de classe C^∞.

Pour pouvoir estimer la moyenne des dérivées des itérées de f au point x_0, observons que pour tous les entiers positifs i, n,

$$\begin{aligned} Df_{n+1}^i(x_0) &= DA_{n+1}\left(h_n\left(x_0 + i\frac{p_{n+1}}{q_{n+1}}\right)\right) Dh_n\left(x_0 + i\frac{p_{n+1}}{q_{n+1}}\right) Dh_n^{-1}(x_0) DA_{n+1}^{-1}(x_0) \\ &= \frac{Dh_n\left(x_0 + i\frac{p_{n+1}}{q_{n+1}}\right) DA_{n+1}\left(h_n\left(x_0 + i\frac{p_{n+1}}{q_{n+1}}\right)\right)}{Dh_n(x_0) DA_{n+1}(x_0)} \geq C_n \frac{DA_{n+1}\left(h_n\left(x_0 + i\frac{p_{n+1}}{q_{n+1}}\right)\right)}{DA_{n+1}(x_0)} \\ &\geq nDA_{n+1}\left(h_n\left(x_0 + i\frac{p_{n+1}}{q_{n+1}}\right)\right). \end{aligned}$$

Par conséquent, si $h_n\left(x_0 + i\frac{p_{n+1}}{q_{n+1}}\right)$ n'appartient pas à la réunion des intervalles I_j (c'est-à-dire, si $x_0 + i\frac{p_{n+1}}{q_{n+1}}$ n'appartient pas à la réunion des J_j, $j = 1, \ldots, q_n$), alors $Df_{n+1}^i(x_0) \geq n$.

Soit $k \geq q_n$, on va montrer qu'au moins pour la moitié des entiers $0 \leq i < k$ on a $x_0 + i\frac{p_{n+1}}{q_{n+1}} \notin J_j$, pour tout $1 \leq j \leq q_n$. En effet, puisque J_j est de longueur inférieure à $\frac{1}{nq_n}$, un calcul simple montre que la proportion des entiers $0 \leq i < k$ pour lesquels $x_0 + i\frac{p_{n+1}}{q_{n+1}} \in J_j$ est au plus $2\left(\frac{1}{q_{n+1}} + \frac{1}{n\,q_n}\right)$. Comme il y a q_n de tels intervalles, nous devons nous assurer que pour n suffisamment grand, $\frac{2q_n}{q_{n+1}} + \frac{2}{n} \leq \frac{1}{2}$. Ceci suit directement de nôtre construction (à savoir de la condition $q_{n+1} \geq 5q_n$).

Par suite, on a que pour tout $n > 1$,

$$\frac{1}{q_{n+1}} \sum_{i=0}^{q_{n+1}-1} Df^i(x_0) \geq \frac{1}{q_{n+1}} \sum_{\substack{i \text{ t.q.} \\ x_0 + i\frac{p_{n+1}}{q_{n+1}} \notin \bigcup_j J_j}} \left[Df_{n+1}^i(x_0) - \frac{1}{2^{n+1}}\right] \geq \frac{n}{2} - \frac{1}{2^{n+1}} \geq \frac{n}{3},$$

et on trouve ainsi que la suite (III.6) diverge.

Passons maintenant à la construction des cocycles. Soit f un difféomorphisme C^∞ fixé pour lequel les moyennes M_n définies par (III.6) divergent. Soit h une fonction C^1 quelconque de μ_f-moyenne nulle. On va montrer que pour tout $r \geq 1$, tout $\varepsilon > 0$ et tout $C > 0$, il existe un entier N et une fonction $\tilde{h} = \tilde{h}_{\varepsilon,r,C} \in C^1(\mathbf{S}^1)$ tels que :

1. $\tilde{h}$ est dans le ε-voisinage de h par rapport à la norme C^r (et donc $\tilde{h} - \int \tilde{h}\, d\mu_f$ est 2ε-proche de h),
2. $\dfrac{|D(S_{q_n}\tilde{h})(x_0)|}{q_n} = \dfrac{|D(S_{q_n}(\tilde{h} - \int \tilde{h}\, d\mu_f))(x_0)|}{q_n} > C.$

En suite, un application du théorème de Baire montrera que pour $\tilde{h} \in C^1(\mathbf{S}^1)$ de μ_f-moyenne nulle *générique*, les moyennes $\frac{D(S_{q_n}\tilde{h})}{q_n}$ divergent.

Pour montrer qu'il existe un tel $\tilde{h}$, on se fixe N_1 tel que pour tout $n \geq N_1$, on ait $M_n > \max\{4C^2, \varepsilon^{-2}\}$ (l'inégalité (III.7) en assure l'existence). Fixons N_2 tel qu'il existe une perturbation $\hat{h}$ de h d'ordre ε dans la topologie C^r, et qui vérifie $D\hat{h}\left(f^i(x_0)\right) = Dh\left(f^i(x_0)\right) + \varepsilon/2$ pour tout $i = 0, \ldots, q_{N_2-1}$. Soit $N = \max\{N_1, N_2\}$. Si

$$\frac{|D(S_{q_N}h)|}{q_N} \geq C,$$

on termine en posant $\tilde{h} = h$. Sinon, en posant $\tilde{h} = \hat{h}$, on trouve

$$\begin{aligned}\frac{D(S_{q_N}\tilde{h})}{q_N} &= \frac{1}{q_N}\sum_{i=0}^{q_N-1} D\tilde{h}\left(f^i(x_0)\right)Df^i(x_0)\\ &= \frac{1}{q_N}\sum_{i=0}^{q_N-1} Dh\left(f^i(x_0)\right)Df^i(x_0) + \frac{\varepsilon}{2q_N}\sum_{i=0}^{q_N-1} Df^i(x_0) > \frac{D(S_{q_N}h)}{q_N} + \frac{\varepsilon M_n}{2} \geq \frac{\varepsilon M_n}{2} - C\\ &\geq \frac{\sqrt{M_N}}{2} - C > C,\end{aligned}$$

comme l'on désirait. □

2.4 Exemples de difféomorphismes C^1 avec plusieurs distributions invariantes

Le premiers exemples seront des difféomorphismes qui possèdent un Cantor minimal invariant. À ce propos, on va décrire une construction de [Dou-Yoc, Section 5] et on expliquera comme elle permet d'obtenir un exemple de difféomorphisme C^1 avec nombre de rotation irrationnel qui admet une distribution d'ordre 1 invariante, différente de la mesure invariante.

Théorème 2.11 (R. Douady – J.-C. Yoccoz)**.** *Pour tout nombre irrationnel α, il existe un difféomorphisme f de classe C^1 qui est un contre-exemple de Denjoy et satisfait à la propriété suivante : il existe $x_0 \in \mathbf{S}^1 - K_f$ tel que*

$$S = \sum_{n\in\mathbf{Z}} Df^n(x_0) < \infty.$$

En particulier, la mesure de probabilité

$$\nu = \frac{1}{S}\sum_{n\in\mathbf{Z}} Df^n(x_0)\delta_{f^n(x_0)}$$

est 1*-automorphe.*

Fixons f comme dans l'énoncé du théorème, soit $\mu = \mu_f$ la mesure f-invariante. Considérons la fonctionnelle linéaire

$$L : u \longmapsto \int_{\mathbf{S}^1} u'\, d\nu,$$

définie sur l'espace des fonctions C^1 sur le cercle. On a que L est f-invariante. En effet,

$$L(u \circ f) = \int_{\mathbf{S}^1} (u \circ f)'\, d\nu = \int_{\mathbf{S}^1} (u' \circ f)Df\, d\nu = \int_{\mathbf{S}^1} u'\, d\nu = L(u),$$

où on a utilisé dans la troisième égalité le fait que ν est 1-automorphe.

Il reste à montrer que L est différente de (un multiple de) la mesure invariante μ. Soit I une composante connexe du complémentaire du minimal exceptionnel K qui contient x_0. Puisque K coïncide avec le support de μ, toute fonction u nulle en dehors de I vérifie

$$\int_{\mathbf{S}^1} u\, d\mu = 0.$$

Pour une telle fonction, on a

$$L(u) = \int_{\mathbf{S}^1} u'\, d\nu = \frac{u'(x_0)}{S},$$

et cette expression peut prendre toute valeur réelle arbitraire en faisant varier u.

La construction d'un difféomorphisme C^1 minimal avec plusieurs distributions invariantes suit une stratégie similaire. L'outil principal est un résultat récent de [Kod-Mat] qui, dans un certain sens, peut être considéré comme la contrepartie mesurable du théorème de Douady et Yoccoz.

Théorème 2.12 (H. Kodama – S. Matsumoto)**.** *Pour tout nombre irrationnel α, il existe un difféomorphisme C^1 minimal f avec nombre de rotation α qui admet un domaine fondamental mesurable C (à savoir, une partie mesurable telle que toute ses itérées sont deux à deux disjointes et recouvrent le cercle).*

En outre, pour tout $x_0 \in C$, on a

$$S = \sum_{n \in \mathbf{Z}} Df^n(x_0) < \infty.$$

Remarque 2.4. Le lecteur pourra confronter la construction astucieuse dans [Kod-Mat] avec les exemples classiques dus à Katznelson [Katz] (voir aussi [Her 1, Chapitre VII]).

Comme plus haut, la dernière condition implique que la mesure

$$\nu = \frac{1}{S} \sum_{n \in \mathbf{Z}} Df^n(x_0) \delta_{f^n(x_0)}$$

est 1-automorphe pour f, d'où on peut à nouveau définir la distribution d'ordre 1 f-invariante

$$L : u \longmapsto \int_{\mathbf{S}^1} u' \, d\nu,$$

sur les fonctions C^1 sur le cercle. Par contre, montrer que L n'est pas triviale sur le noyau de μ n'est pas géométriquement évident comme il était dans le cas précédent. Néanmoins, une méthode générale très simple montre ceci (et, en fait, le cas précédent aussi), et on l'expliquera dans la suite.

Par simple intégration, toute fonction continue sur le cercle avec moyenne nulle par rapport à la mesure de Lebesgue peut être vue comme dérivée d'une fonction C^1 de μ-moyenne nulle. Par conséquent, si L était triviale sur le noyau de μ, alors, pour toute fonction v continue sur le cercle, on aurait

$$0 = \int_{\mathbf{S}^1} \left(v - \int_{\mathbf{S}^1} v \, d\mathrm{Leb} \right) d\nu = \int_{\mathbf{S}^1} v \, d\nu - \int_{\mathbf{S}^1} v \, d\mathrm{Leb}.$$

En d'autres mots, $\nu = \mathrm{Leb}$, qui est absurde.

Remarque 2.5. Même si les exemples de Douady et Yoccoz ne peuvent pas être rendus $C^{1+\tau}$ pour $\tau > 0$, les contre-exemples de Denjoy largement décrits dans [Nav 1] possèdent des mesures atomiques 1-automorphes dans le même esprit que ceux dans [Dou-Yoc]. Le cas des difféomorphismes minimaux est plus compliqué. En effet, les exemples de Kodama et Matsumoto ne sont pas non plus $C^{1+\tau}$, et il n'est pas clair si l'on peut produire des mesures 1-automorphes en suivant une construction similaire. (En vérité, on ne connait pas des exemples de difféomorphismes $C^{1+\tau}$ minimaux qui ne sont pas ergodiques par rapport à la mesure de Lebesgue). Ceci suscite la question de savoir s'il existe un difféomorphisme du cercle de classe $C^{1+\tau}$ et minimal qui possède plusieurs distributions d'ordre 1 invariantes.

Question 4. Existe-t-il un difféomorphisme du cercle de classe C^2 et nombre de rotation irrationnel qui possède une distribution d'ordre 2 invariante, différente de la mesure invariante ?

Capitolo IV

Distances aléatoires stationnaires sur le graphes hiérarchiques

1 Introduction

The main motivation for this work is to make a little step towards the mathematical understanding of Liouville quantum gravity: the problem of giving a meaning to the "metric tensor" defined on a surface Ω by the exponential of the Gaussian Free Field (GFF for short), still remains unsolved. For planar domains Ω carrying a particular geometrical structure, it is possible to define a similar "metric tensor" using *multiplicative cascades*. To give a concrete picture, let Ω be the unit square in the plane, then using dyadic coordinates, it is possible to identify it with the ends of a rooted quaternary tree (with the little care that this identification is clearly not one-to-one). If we assign positive random weights to the edges of this tree (with the random factors that are i.i.d.), we can formally define a weight for almost any end, namely the infinite product of the factors that we read along the path connecting the root to the end, properly renormalized in such a way that its expectation stays bounded. The purpose is to understand whether it is possible to define a measurable (pseudo-)metric, whose metric tensor at a point z is the corresponding infinite product. In a different formulation, we want to understand the limit of the sequence of piecewise-flat Riemannian metrics, obtained by considering the factors on the tree up to a certain depth. There is no evident reason for this limit to be defined and non-degenerate, but observe that the resulting random metric d must be *stationary* (or self-similar): considering four different independent samplings of the random metric space (Q_i, d_i) and a new independent random factor ξ, we can construct a new random metric space $(\tilde{Q}, \tilde{d})$, which must have the same law. More precisely, we glue the four squares $(Q_i, \xi \cdot d_i)$ together along their sides and obtain a new square $\tilde{Q}$, defining the pseudo-metric $\tilde{d}$ on it by looking at the shortest path between points.

This problem is far easier in a 1D setting: multiplicative cascades on the interval have been successfully studied since the pioneering work by Kahane and Peyrière [Kah-Pey, Kah, Ben-Sch]. In this work, inspired by a question of Benjamini, we make the first steps to construct a stationary random metric on the fractal objects that are the limit of finite graphs. Namely we refer to *hierarchical graphs*, whose nice self-similar structure allows to define multiplicative cascades. This is a challenging problem: contrary to an interval, the number of possible geodesic paths joining any two points is infinite, but still they save some peculiarity of the one-dimensional world. Hierarchical graphs have been widely studied in physics and engineering for many years, but we are highly indebted to I. Benjamini who focused our attention on such examples. To make the problem more definite, he asked whether it is possible to solve it for a not too complicated object: start with the *figure eight-graph* Γ and build the associated hierarchical graph. This means the following: we mark the top and bottom vertices I and O of the eight graph and recursively build the sequence of combinatorial graphs (Γ_n, I, O), replacing every edge in Γ_{n-1} by a copy of Γ. The *hierarchical figure eight-graph* is the limit combinatorial object Γ_∞ (see Figure IV.1). We will recall the precise definition in Section 5, where we discuss hierarchical graphs in broader generality.

Our model for studying random distances on the hierarchical figure eight-graph reflects its self-similar structure, like in the example of the square. Assume that we are given a probability distribution m on the positive real numbers $\mathbf{R}_+ = (0, +\infty)$. We start with an interval of length 1, and make the process of construction step by step. At each step, for each interval I of length $|I|$, we take a random variable ξ_I, distributed with respect to m; these variables are mutually independent, for all intervals and for all steps. We then replace I by four intervals of length $|I| \cdot \xi_I$ each, forming an eight-shaped figure between the original endpoints of I. Then, it is natural to expect (and try to prove) that appropriately normalized metrics on these graphs converge to a random limit metric on the limit object.

As in the example of the square, we remark that this procedure can be reversed, at least to study the law of

random metrics. Namely, a level n figure is glued out of four independent samples of level $n-1$ figures, that are glued replacing the intervals in the eight shape, and the obtained new metric is multiplied by a random constant ξ.

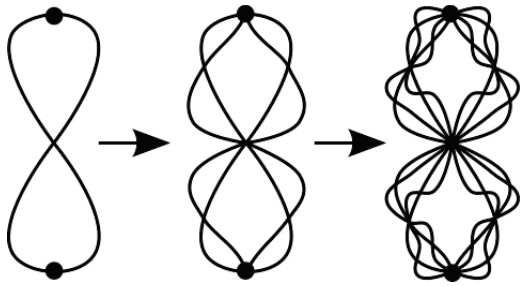

Figure IV.1 – Hierarchical figure eight construction procedures.

It is easy to see that the "top-to-bottom" IO_n distance for the level n figure depends only on ξ and on the "top-to-bottom" distances of its subfigures of level $n-1$ (any "top-to-bottom" geodesic passes through the middle point). Hence, the law $\mathcal{L}_n$ of the distance IO_n for the level n depends only on the law for the level $n-1$:

$$\mathcal{L}_n = \text{law}\left(\xi \cdot (\min(IO^1_{n-1}, IO^2_{n-1}) + \min(IO^3_{n-1}, IO^4_{n-1}))\right), \tag{IV.1}$$

where ξ and the IO^i_{n-1} are independent random variables, ξ distributed like m and each of the IO^i_{n-1}'s distributed like $\mathcal{L}_{n-1}$.

It is quite natural to conjecture that the normalization constant shall behave as λ^{-n} for some $\lambda \geq 0$: here we give a rather intuitive argument when the variable ξ has finite first moment. Indeed, we remark that the sequence $\mathbf{E}[IO_n]$ is sub-multiplicative: for any positive integers i and j:

$$\mathbf{E}[IO_{i+j}] \leq \mathbf{E}[IO_i]\,\mathbf{E}[IO_j]\,.$$

This implies that the limit $\lambda = \lim_{n\to\infty} \mathbf{E}[IO_n]^{-1/n}$ exists. The point is to show that λ is finite and no correction term appears for the exponential growth of the normalization constant.

It is natural to define the following process on $\mathcal{P}$, the space of Radon probability distributions on the compactified half-line $[0, +\infty]$ equipped with the weak-$*$ topology: we fix a random variable ξ distributed as m, then for any $\mu \in \mathcal{P}$ we define $\Phi_\lambda[\mu] \in \mathcal{P}$ which is the law of the random variable

$$R_\lambda(X^1, X^2, X^3, X^4; \xi) = \lambda\, \xi \cdot (\min(X^1, X^2) + \min(X^3, X^4)), \tag{IV.2}$$

when the random variables (X^i) are mutually independent and distributed as μ (and also independent of ξ).

Our main result is the following theorem, claiming the existence of a stationary measure for the normalized figure eight *distances* process:

Theorem 1.1. *For any non-atomic, fully supported probability measure m on $\mathbf{R}_+$ there exists a normalization constant $\lambda_{cr} > 0$ and a non-atomic probability measure $\bar{\mu}$ on $\mathbf{R}_+$ such that $\bar{\mu}$ is a fixed point for the operator $\Phi_{\lambda_{cr}}$.*

To prove it, we introduce a cut-off process, for which the existence of a stationary distribution is easy to establish (cf. [Ald-Ban, Lemma 4]). Then, passing to the double-scaling limit, we retrieve the desired stationary distribution for the original process.

Remark 1.1. Equation (IV.1) is an example of *recursive distributional equation* (RDE) of (min, +)-type. Formulated in other words, Theorem 1.1 gives an absolutely continuous solution for this RDE. Since there is no need here for introducing precisely such a setting, we shall refer to the survey by Aldous and Bandyopadhay [Ald-Ban] for more details.

We can go further when the probability distribution m is absolutely continuous with respect to the Lebesgue measure on $\mathbf{R}_+$. In this case, under some additional assumptions on the density (we write $m(dx) = \rho(x)\frac{dx}{x}$, so that the function ρ describes the density in the logarithmic coordinates), we show that any two measures $\bar{\mu}$, $\bar{\mu}'$ that we can obtain from Theorem 1.1 are essentially the same and in fact, any starting distribution converges to a stationary measure under the iterations of $\Phi_{\lambda_{cr}}$:

Theorem 1.2. *Let $m(dx) = \rho(x)\frac{dx}{x}$ be an absolutely continuous probability measure on $\mathbf{R}_+$, where the function $\rho(x)$ is strictly positive and continuous on $(0, +\infty)$ and tends to zero as $x \longrightarrow 0$ and as $x \longrightarrow +\infty$.*

Let $\bar{\mu}$ be a probability measure on $\mathbf{R}_+$ such that $\bar{\mu} = \Phi_{\lambda_{cr}}[\bar{\mu}]$. Then, for any probability measure μ on $\mathbf{R}_+$ there exists a constant $c > 0$ such that the iterations of this measure converge to the c-rescaled measure $\bar{\mu}$:

$$\Phi^n_{\lambda_{cr}}[\mu] \longrightarrow \bar{\mu}(c\,\cdot) \quad \text{as } n \longrightarrow \infty.$$

In particular the $\Phi_{\lambda_{cr}}$-stationary probability measure is unique up to a rescaling.

Remark 1.2. It is worth to detail why for any $c > 0$, the measure $\bar{\mu}(c\,\cdot)$ is a fixed point for the operator Φ_λ. This comes from the homogeneity of the expression defining Φ_λ and it is a *crucial property* for our arguments: let $\mu \in \mathcal{P}$ be a probability measure and X^i be i.i.d. random variables of law μ, then

$$R_\lambda(c^{-1}X^1, \dots, c^{-1}X^4; \xi) = c^{-1} \cdot R_\lambda(X^1, \dots, X^4; \xi),$$

so that we get $\Phi_\lambda[\mu(c\,\cdot)] = \Phi_\lambda[\mu](c\,\cdot)$.

Remark 1.3. It is interesting to compare this result to the classical contraction method for RDEs (see [Ald-Ban]). We do not believe that it could be applied in our case, but it is clear that *a posteriori*, the operator $\Phi_{\lambda_{cr}}$ defines a contraction on the space of probability distribution modulo rescaling, for this operation commutes with $\Phi_{\lambda_{cr}}$.

The results we are presenting here do not show directly the existence of a stationary random *metric*. This is a more delicate problem and we will discuss our intuition on it in an improved version of this paper.

To conclude this introduction, it is important to say that such a procedure can be adapted to some other situations, and perhaps opens another (though not at all easy or immediate!) road towards the construction of 2D random metric. A very simple generalization is indeed the content of Section 5: our results extend to a huge class of hierarchical graphs.

The next, still finite-dimensional, case that we have handled is the one of the hierarchical model on the Sierpinsky gasket: the information that we are keeping on the metric in this case are three distances between the vertices of the triangle. However, working with the distributions on a non one-dimensional space of such distances becomes more elaborate (there are no more partition functions that sometimes simplify the coupling), so to avoid overloading the paper, we avoid discussing about it here.

A toy model for Liouville Field Theory

Mainly based on the work of Polyakov [Pol 1, Pol 2], the study of *2D-quantum gravity* has been increasingly drawing attention in physics and mathematics during the past three decades. For mathematicians, this often means studying random Riemannian surfaces: in fact, the intuition of Polyakov was to "replace the old-fashioned (and extremely useful) sums over random paths" with sums over random surfaces [Pol 1]. Presented like this, there is much ambiguity on the significance of "random". Formally it becomes more definite introducing the *Liouville action* S_L for a Riemannian surface (Ω, g_0) (g_0 is a fixed *background metric*):

$$S_L(h) = \frac{1}{4\pi} \int_\Omega \left(g_0(\nabla_{g_0} h, \nabla_{g_0} h) + Q\, R_{g_0} h + 4\pi\mu e^{\gamma h} \right) d\mathrm{vol}_{g_0},$$

where R_{g_0} is the curvature tensor, $\gamma \in (0, 2]$, $Q = \frac{2}{\gamma} + \frac{\gamma}{2}$, and μ is the *cosmological constant*. Like for path-integrals, the random surface $(\Omega, e^{\gamma h})$ will be chosen with a probability proportional to $e^{-S_L(h)} dh$. When the cosmological constant μ goes to zero, we lose the interaction with matter and enter in a *pure gravity* situation. When g_0 is a flat metric, in pure gravity the action reduces to

$$S_L(h) = \frac{1}{2} \|h\|^2_{\nabla_{g_0}}, \tag{IV.3}$$

where $\|h\|_{\nabla_{g_0}} = \left(\int_\Omega g_0(\nabla_{g_0} h, \nabla_{g_0} h) d\mathrm{vol}_{g_0} \right)^{1/2}$ is the Dirichlet energy of $h \in H^1_0(\Omega)$ (H^1-functions orthogonal to constants). Mathematicians call the random field h, distributed according to this action, the *Gaussian Free Field* (GFF for short), while it has different names in the physical literature (Euclidean bosonic - spinless - massless - free field, just to quote some).

The link between 2D-quantum gravity and GFF was made explicit in the independent works by David and Distler, Kawai [Dav, Dis-Kaw] which extended to the usual conformal gauge the previous results of Knizhnik, Polyakov and Zamolodchikov (KPZ) [Kni-Pol-Zam] (dealing with the light-cone gauge, which means very roughly that they were considering a 2D world with many symmetries). For some very good introductions to 2D-Liouville quantum gravity we suggest the reading of [Gin-Moo, Nak, Tes], although they are addressed to people with a background knowledge in quantum field theory and string theory.

Mathematically, there is a huge obstacle in considering random "functions" distributed according to the action (IV.3), as the random field h is almost surely only a distribution (in the sense of Schwartz). The GFF can be defined in a few different ways (for a good introduction, though not exhaustive, we recommend [She 1]). The one which will be relevant for our model is the following: let Δ be the Laplace operator on the (flat) Riemann surface (Ω, g_0) and let $(\Phi_n)_{n\in\mathbf{N}}$ be an orthonormal basis of the Sobolev space $H^1_0(\Omega)$ formed by the non-constant eigenfunctions of $(-\Delta)^{-1/2}$, then the series

$$h = \sum_{n\in\mathbf{N}} a_n \Phi_n$$

defines the GFF when the coefficients a_n are independent normally distributed random variables.

When Ω is the unit square, changing appropriately the Hilbert orthonormal basis of $H_0^1(\Omega)$ to a basis of *vaguelets* (see [Tec], though the author uses the vaguelets representation for the GFF on the unit *circle*), we find the definition of the dyadic GFF, already introduced "by hand" at the very beginning of this work.

Though the problem of defining rigorously the random Riemannian metric "$e^{\gamma h}|dz|$" is still out of reach, the multiplicative chaos approach [Kah-Pey, Kah] works fine for the random *measure* $e^{\gamma h}d^2z$. In particular, the celebrated results of Duplantier and Sheffield [Dup-She], as well as the work of Rhodes and Vargas [Rho-Var 1], established mathematically the famous KPZ relation for measures, relation first stated in [Kni-Pol-Zam] and relating, in a very simple formula, the scaling exponents of the random metric (*measure* here) with the Euclidean one (for further reading, we suggest the review [Gar]). These results have been the starting point for many others in recent years [Bar-Jin-Rho-Var, Dup-Rho-She-Var, Bar-Kup-Nik-Sak-Web, Mil-She] (this list is certainly non-exhaustive).

Sketch of the proof of Theorem 1.1

The main idea of the proof is to use the following cut-off process: assume that after the replacement and the multiplication, the distance IO is shortcut by an "exterior" path of length A, where $A \in \mathbf{R}_+$ is a fixed constant. In other words, instead of the map R_λ in (IV.1)–(IV.2) consider the map $R_{A,\lambda} := \min(R_\lambda, A)$, defining the corresponding operator $\Phi_{A,\lambda}$.

For sufficiently large λ, there should be a $\Phi_{A,\lambda}$-stationary measure. Indeed, for such λ, it is natural to expect that measures μ, concentrated "away" from A on the interval $(0, A)$, will –in a sense– drift to the right; meanwhile, their support stays uniformly bounded above by A, so the iterations of a starting measure will not go to $+\infty$.

Moreover one can construct it starting with a Dirac measure concentrated on the infinite distance: it is easy to see that the distribution functions of its images form a pointwise monotonously increasing sequence. The only question for such modified procedure is whether this limit becomes a Dirac measure concentrated on the zero distance, in other words, whether the metric collapses.

It turns out that there exists a critical value λ_{cr} such that for $\lambda > \lambda_{cr}$ the process $\Phi_{A,\lambda}$ admits a stationary measure $\bar{\mu}_{A,\lambda}$, while for $\lambda \le \lambda_{cr}$ the whole metric space collapses into a point. A key remark here is that the collapse happens even for $\lambda = \lambda_{cr}$. Then, the stationary measure for the process $\Phi_{\lambda_{cr}}$ can be found as a limit of $\bar{\mu}_{A(\lambda),\lambda}$ as λ approaches λ_{cr} from the right, where the length $A(\lambda)$ tending to ∞ is chosen by the normalizing restriction $\bar{\mu}_{A(\lambda),\lambda}([0,1]) = 1/2$. In other words, we are at the same time decreasing the scaling parameter λ to the critical value (that would lead to the collapse if A stayed fixed) and making the shortcut length A tend to infinity (that would explode the metric if λ stayed fixed). In the limit (that we force to be non-trivial by our normalizing condition), we find the desired stationary measure, that does not "feel" the shortcuts anymore.

2 Existence of a stationary law

2.1 Preliminaries, notations and definitions

The operator Φ_λ defined by (IV.2) induces a continuous dynamics on the space $\mathcal{P}$ of probability measures on the extended half-line $[0, \infty]$. Since we are requiring that the measure m is non-atomic, it follows that the image of Φ_λ consists of *non-atomic* probability measures. The reason for considering *probability measures with atoms* as well is that the cut-off procedure that we are going to use takes a non-atomic measure to a possibly atomic one.

The space $\mathcal{P}$ carries a natural partial order $\preccurlyeq$, which is the well-known *stochastic domination*.

Definition 2.1. *Given μ and ν in $\mathcal{P}$, we write $\mu \preccurlyeq \nu$ if there is a coupling (X, Y) of these measures (that is, a random vector (X, Y) with marginal laws μ and ν respectively), in such a way that $X \le Y$ almost surely.*

We also define $\mathcal{P}_0 \subset \mathcal{P}$ to be subspace of probability measures which have no atoms at 0 or ∞. Then every probability measure $\mu \in \mathcal{P}$ can be uniquely decomposed into the convex combination

$$\mu = p_0(\mu)\mathrm{Dirac}_0 + p_\infty(\mu)\mathrm{Dirac}_\infty + (1 - p_0(\mu) - p_\infty(\mu))\,\mu_0, \tag{IV.4}$$

with μ_0 in $\mathcal{P}_0$ and $p_0(\mu)$, $p_\infty(\mu)$, $1 - p_0(\mu) - p_\infty(\mu)$ in $[0, 1]$.

The stochastic domination turns out to be helpful in many cases when dealing with RDEs (see [Ald-Ban, §2.2]). It is a crucial remark that this partial order is well adapted to the dynamics induced by the operator Φ_λ.

Lemma 2.1.

1. *For any $\lambda > 0$, the operator Φ_λ is* order-preserving*: if $\mu \preccurlyeq \nu$ then $\Phi_\lambda[\mu] \preccurlyeq \Phi_\lambda[\nu]$.*

2. *For any $\lambda \leq \lambda'$ and $\mu \in \mathcal{P}$ we have $\Phi_\lambda[\mu] \preccurlyeq \Phi_{\lambda'}[\mu]$.*

Proof. Let us prove the first statement only, since the second one is rather evident. Suppose that X_i^μ and X_i^ν, $i = 1, \dots, 4$ are respectively four independent random variables distributed according to μ and ν and defined on the same probability space $(\Omega, \mathbf{P})$, in such a way that for every i and almost every $\omega \in \Omega$ the relation $X_i^\mu(\omega) \leq X_i^\nu(\omega)$ holds. Then the claimed inequality easily follows: given another independent random variable ξ of law m, for almost every $\omega \in \Omega$ we have

$$\begin{aligned} &\lambda\xi(\omega)\left(\min\left(X_1^\mu(\omega), X_2^\mu(\omega)\right) + \min\left(X_3^\mu(\omega), X_4^\mu(\omega)\right)\right) \\ \leq &\lambda\xi(\omega)\left(\min\left(X_1^\nu(\omega), X_2^\nu(\omega)\right) + \min\left(X_3^\nu(\omega), X_4^\nu(\omega)\right)\right). \end{aligned}$$

In other words, $\Phi_\lambda[\mu] \preccurlyeq \Phi_\lambda[\nu]$ as desired. □

What is commonly known as *Strassen's Theorem* (see for example [Wer, §7.5]), asserts that the condition $\mu \preccurlyeq \nu$ is equivalent to the fact that for any *increasing* bounded real valued-function f on $[0, +\infty]$,

$$\int_{[0,\infty]} f\, d\mu \leq \int_{[0,\infty]} f\, d\nu. \tag{IV.5}$$

Taking functions f of the form $\mathbf{1}_{(x,\infty]}$ ($x \in \mathbf{R}_+$) it is easy to see that the inequality $\mu \preccurlyeq \nu$ coincides with the *reversed* inequality $F_\mu \geq F_\nu$ for distribution functions.

Remark 2.1. However this interpretation with distribution functions is peculiar to measures on the real *line*. Under the perspective of generalizing our methods to multivariate random variables, it will be only possible to make use of condition (IV.5).

This complementary point of view will be sometimes very useful, as well as the function $\theta : [0,1] \longrightarrow [0,1]$ that we are defining, naturally associated to the figure eight-graph Γ.

Definition 2.2. *Let $\theta(p)$ be defined as the probability that a path IO exists in the eight shape, when each of the edges is drawn with probability p.*

An easy computation shows that $\theta(p) = p^2(2-p)^2$. Note that the map θ on $[0,1]$ has two attracting fixed points, 0 and 1 (moreover, these points are *superattracting*: $\theta'(0) = \theta'(1) = 0$), and one repelling fixed point $p_{cr} \in (0,1)$. Actually, it is easy to find that $p_{cr} = \varphi^{-2}$, where $\varphi = \frac{1+\sqrt{5}}{2}$ is the golden ratio.

The following lemma immediately relates this function to the dynamics of Φ_λ:

Lemma 2.2. *For any $\lambda > 0$ and $\mu \in \mathcal{P}$ we have*

$$p_0\left(\Phi_\lambda[\mu]\right) = \theta\left(p_0(\mu)\right), \quad 1 - p_\infty\left(\Phi_\lambda[\mu]\right) = \theta\left(1 - p_\infty(\mu)\right).$$

Proof. Considering the figure eight-graph Γ, we look for the probability that there exists a zero-length IO-path when assigning random distances to the edges according to the probability distribution μ. This means that we look for a path from I to O passing from zero-length edges only. The definition of the function θ then shows the first equality. For the second one, we can argue in the same way, replacing the word *zero* by *finite*. □

2.2 Cut-off process

The main difficulty in finding a stationary measure is that the metric could blow up or collapse at different scales. In order to tame such problems, we artificially impose a "regular" behaviour to the random metric, using cut-off procedures.

We recall that Φ is a monotone operator. So we introduce the following property, which is in some sense "stable" under Φ.

Definition 2.3. *We say that a probability measure $\mu \in \mathcal{P}$ is λ-*zooming out *if:*

- $\mu \preccurlyeq \Phi_\lambda[\mu]$,
- *μ is not the Dirac mass at* 0,
- *μ is supported on some closed subinterval $[0, x]$, with $x \in \mathbf{R}_+$.*

It is quite natural (and it is formally implied by the second statement in Lemma 2.1) that if a measure $\mu \in \mathcal{P}$ is λ-zooming out, then it is also λ'-zooming out for any $\lambda' \geq \lambda$.

There are non-trivial measures that are smaller than their Φ_λ-image, and we want to get from them measures that λ-zoom out. For this reason we introduce the following operator that forces the support to be included in a subinterval. For any $A \in \mathbf{R}_+$, define the cut-off operator $\Phi_{A,\lambda}$ by assigning to any probability measure $\mu \in \mathcal{P}$, the law of

$$\min(R_\lambda(X_1, \dots, X_4; \xi), A),$$

where the X_i's are i.i.d. variables with law μ, and ξ is distributed with respect to m and independent of them. Geometrically, we are adding an *IO*-shortcut of length A, when gluing together four independent samples of the space. The reader will remark that also the new operator $\Phi_{A,\lambda}$ preserves the partial order on $\mathcal{P}$ due to the same coupling arguments as those used in the proof of Lemma 2.1.

Remark 2.2. It should be evident that for any $\lambda > 0$ and $A \in \mathbf{R}_+$ we have $\Phi_\lambda \succcurlyeq \Phi_{A,\lambda}$. To check this, we construct the following natural coupling: fix a measure $\mu \in \mathcal{P}$ and let Y be a random variable of law $\Phi_\lambda[\mu]$ on the probability space $(\Omega, \mathbf{P})$. We can define the new random variable

$$Y^A(\omega) = \begin{cases} Y(\omega) & \text{if } Y(\omega) \leq A, \\ A & \text{if } Y(\omega) > A, \end{cases}$$

whose law is $\Phi_{A,\lambda}[\mu]$.

An important property of the cut-off process is that, as it is very natural to expect and as we will show later (see Lemma 2.6), it has a non-trivial stationary distribution for all sufficiently large λ's. This makes the situation more "flexible": for this particular process, we do not have to look for a precise value of λ where it does neither explode nor collapse, obtaining instead a large (in fact, open, as we will show later) set of possible values of λ.

Definition of $\nu_{A,\lambda}$. We define the following sequence in $\mathcal{P}$: starting with Dirac_∞, which is the greatest probability measure in $\mathcal{P}$, we apply $\Phi_{A,\lambda}$ repeatedly in order to get a monotone decreasing sequence in $\mathcal{P}$ that must converge (in the weak-$*$ topology) to a certain probability measure $\nu_{A,\lambda}$. Namely we set

$$\begin{cases} \mu_0^{A,\lambda} = \mathrm{Dirac}_\infty \\ \mu_{n+1}^{A,\lambda} := \Phi_{A,\lambda}[\mu_n^{A,\lambda}] \quad \text{for } n \in \mathbf{N}. \end{cases} \tag{IV.6}$$

Since $\mu_0^{A,\lambda}$ is the greatest probability measure in $\mathcal{P}$, we must have $\mu_0^{A,\lambda} \succcurlyeq \mu_1^{A,\lambda}$. Then for any $n \in \mathbf{N}$, the relation $\mu_n^{A,\lambda} \succcurlyeq \mu_{n+1}^{A,\lambda}$ recursively holds, applying Lemma 2.1. This implies that the weak-$*$ limit $\nu_{A,\lambda}$ of the sequence $\left(\nu_n^{A,\lambda}\right)_{n\in\mathbf{N}}$ exists, for the sequence of distribution functions of $\mu_{n+1}^{A,\lambda}$ is a pointwise monotone, and hence pointwise converges.

Remark 2.3. This monotone construction is somehow classical and gives the greatest invariant measure. The reader can compare for example [Ald-Ban, §2.2].

To verify that $\nu_{A,\lambda}$ is a fixed point for the operator $\Phi_{A,\lambda}$, we only need the following elementary Lemma.

Lemma 2.3. *The operator $\Phi_{A,\lambda}$ is continuous.*

Proof. It is a composition of one map, associating to a measure μ the measure $\mu \otimes \cdots \otimes \mu \otimes m$ on $[0,\infty]^4 \times \mathbf{R}_+$, and the pushforward by the continuous map

$$(X_1, \dots, X_4; \xi) \mapsto \min(R_\lambda(X_1, \dots, X_4; \xi), A).$$

Both these operations are clearly continuous in the sense of $*$-weak convergence for measures on $[0,\infty]$. (Note that the "+" and "min" operations are continuous even on the compactified half-line $[0,\infty]$.) □

Remark 2.4. In fact, the same argument shows that the map $(A, \lambda, \mu) \mapsto \Phi_{A,\lambda}[\mu]$ is continuous in all the three variables $A \in [0,\infty]$, $\lambda \in \mathbf{R}_+$, $\mu \in \mathcal{P}$, with $A = \infty$ corresponding to $\Phi_{\infty,\lambda}[\mu] = \Phi_\lambda[\mu]$.

Note that for every $n \geq 1$, the support of $\mu_n^{A,\lambda}$ and hence of $\nu_{A,\lambda}$ is contained in $[0, A]$. We remark also that since $\nu_{A,\lambda} = \Phi_{A,\lambda}[\nu_{A,\lambda}]$, the measure $\nu_{A,\lambda}$ has no atoms on the open interval $(0, A)$. Hence, when $\nu_{A,\lambda}$ is not equal to Dirac_0, it is a λ-zooming out measure.

Remark 2.5. The definition of $\nu_{A,\lambda}$ depends certainly on the cut-off value A, but the condition $\nu_{A,\lambda} = \mathrm{Dirac}_0$ does not. Indeed for any $A, A' \in \mathbf{R}_+$, repeating the argument in Remark 1.2, we obtain the identity

$$\min\left(R_\lambda\left(\tfrac{A}{A'}X_1, \dots, \tfrac{A}{A'}X^4; \xi\right), A\right) = \tfrac{A}{A'}\min\left(R_\lambda\left(X^1, \dots, X^4; \xi\right), A'\right),$$

so $\Phi_{A,\lambda}[\mu((A'/A)\,\cdot)]((A/A')\,\cdot) = \Phi_{A',\lambda}[\mu]$. Iterating this equality and taking $\mu = \mathrm{Dirac}_\infty$, we get, for every Borel set $E \subset [0,\infty]$,

$$\nu_{A,\lambda}((A/A') \cdot E) = \nu_{A',\lambda}(E). \tag{IV.7}$$

There is a nice interpretation for the construction of the measure $\nu_{A,\lambda}$ (cf. the construction of the *recursive tree process* out of a RDE, as explained in [Ald-Ban, §2.3]).

That is, let $\mathbf{S}^1$ be the rooted quaternary tree, $\bullet$ its root, and let $\mathbf{x} : \mathbf{S}^1 \longrightarrow \mathbf{R}_+$ be a function associating to any vertex a positive number. Let us put $\mathbf{x}$ into correspondence with a family of functions $D_n = D_n(\mathbf{x}) : \mathbf{S}^1 \longrightarrow (0, \infty]$ defined as

$$D_n(t) = \begin{cases} \infty & \text{if } t \text{ is at distance at least } n \text{ from the root, otherwise} \\ \\ \min(R_\lambda(D_n(t_1), \dots, D_n(t_4); \mathbf{x}(t)), A) & \text{where } t_1, \dots, t_4 \in \mathbf{S}^1 \text{ are the four descendants of } t. \end{cases}$$

In other words, $(D_n)_{n\in\mathbf{N}}$ is the family of distances obtained in the following way: everything deeper than n levels is declared to be infinite, while closer to the root we are applying the procedure indicated by the figure eight-graph with the factor $\mathbf{x}$ at the corresponding vertex, adding a shortcut of length A. Then we have the following

Proposition 2.4. *The family of functions $D_n(\mathbf{x})$ is pointwise monotonously decreasing with respect to n, and their limit $D_\infty(\mathbf{x})(t) = \lim_{n\longrightarrow\infty} D_n(\mathbf{x})(t)$ is, for every $\mathbf{x}$, a map from $\mathbf{S}^1$ to $[0, A]$, satisfying the following properties:*

- *Let $t \in \mathbf{S}^1$ be a vertex. Denoting by $\mathbf{x}_t$ the restriction of $\mathbf{x}$ on the subtree $\mathbf{S}^1_t$ of $\mathbf{S}^1$ rooted at t, then the restriction of $D_\infty(\mathbf{x})$ on $\mathbf{S}^1_t$ coincides with $D_\infty(\mathbf{x}_t)$. Moreover one has*

$$D_\infty(\mathbf{x})(\bullet) = \min(R_\lambda(D_\infty(\mathbf{x}_1)(\bullet_1), \dots, D_\infty(\mathbf{x}_4)(\bullet_4); \mathbf{x}(\bullet)), A), \tag{IV.8}$$

  where $\mathbf{x}_i$ is the function $\mathbf{x}_{\bullet_i}$, with $\bullet_1, \dots, \bullet_4$ the direct descendants of the root $\bullet$.

- *If $(\mathbf{x}_t)_{t\in\mathbf{S}^1}$ are i.i.d. random variables with law m, then the law of $D_\infty(\bullet)$ is $\nu_{A,\lambda}$.*

Proof. Let us fix $n \in \mathbf{N}$, we want to prove that $D_n(t) \le D_{n-1}(t)$ for every $t \in \mathbf{S}^1$. If the depth of t is larger than or equal to n, this is evident; proceeding inductively upwards, if t is at level $n - k$, we have

$$\min(R_\lambda(D_n(t_1), \dots, D_n(t_4); \mathbf{x}(t)), A) \le \min(R_\lambda(D_{n-1}(t_1), \dots, D_{n-1}(t_4); \mathbf{x}(t)), A),$$

as desired. Hence the pointwise limit $D_\infty(\mathbf{x})(t)$ exists. The first claimed property may be easily verified, since the definition of $D_n(t)$ only depends on the values of $\mathbf{x}$ on the subtree rooted at t.

To prove the second one, we observe that the sequence of random variables $(D_n(\bullet))_{n\in\mathbf{N}}$ is related to the sequence of measures $\left(\mu_n^{A,\lambda}\right)_{n\in\mathbf{N}}$. Indeed, the law of $D_0(\bullet)$ is the atomic mass Dirac_∞ and we can argue by induction, using the relation (IV.8), to see that $\mu_n^{A,\lambda}$ is the law of $D_n(\bullet)$. □

An immediate corollary of this representation and the 0–1 law is the following Lemma.

Lemma 2.5. *For any A, λ the mass $p_0(\nu_{A,\lambda})$ equals either zero or one.*

Proof. The event $\{D_\infty(\bullet) = 0\}$ is a tail event for the family of random variables $(\mathbf{x}_t)_{t\in\mathbf{S}^1}$; indeed, altering the values of the multipliers $\mathbf{x}_t$ on any finite set of vertices can change the value of the distance $D_\infty(\bullet)$, but not the fact that this distance vanishes. Hence, due to the Kolmogorov 0–1 law, the probability of this event, which is exactly $\nu_{A,\lambda}(\{0\}) = p_0(\nu_{A,\lambda})$, equals either 0 or 1. □

Definition of the supercritical set Λ. We now see how the notion of λ-zooming out may give us an insight on the distribution $\nu_{A,\lambda}$, which, by Remark 2.5, essentially does not depend on the cut-off value A. Indeed, we intuitively know that for large values of λ, the distances in the limit graph Γ_∞ explode (and therefore for such λ, the measure $\nu_{A,\lambda}$ is not the atomic mass Dirac_0, generated by a balance between the expansion and the cut-offs); on the other hand, the opposite situation arises when λ is very close to 0, as in this case the diameter of Γ_∞ should be equal to 0, and hence $\nu_{A,\lambda} = \mathrm{Dirac}_0$. So it is natural to look for a *phase-transition parameter* λ_{cr}. This motivates the following definition.

Consider the set Λ formed by all $\lambda > 0$ for which we can obtain a non-trivial limit measure $\nu_{A,\lambda}$, for some (and hence every, cf. Remark 2.5) $A \in \mathbf{R}_+$:

$$\Lambda = \{\lambda \mid \exists A \in \mathbf{R}_+ \text{ s.t. } \nu_{A,\lambda} \ne \mathrm{Dirac}_0\} = \{\lambda \mid \forall A \in \mathbf{R}_+ \text{ s.t. } \nu_{A,\lambda} \ne \mathrm{Dirac}_0\}.$$

Lemma 2.6. *The set Λ is non-empty, as well as its complement. In other words, Λ is a non-empty half-line, starting at a positive real number λ_{cr}.*

Proof. The proof requires different definitions for the set Λ: as the reader will see, the non-emptiness of Λ and that of $\mathbf{R}_+ \setminus \Lambda$ are easier to prove dealing with the following sets Λ_1 and Λ_2.

Let us denote by Λ_1 the set of $\lambda > 0$ such that there exists a measure $\mu \in \mathcal{P}$ that λ-zooms out:

$$\Lambda_1 = \{\lambda \mid \exists\, \mu \in \mathcal{P} \text{ s.t. } \Phi_\lambda[\mu] \succcurlyeq \mu,\ \mu \neq \mathrm{Dirac}_0 \text{ and } \exists\, x \in \mathbf{R}_+ \text{ s.t. } \mu \text{ is supported on } [0, x]\}.$$

Due to Lemma 2.1, if $\lambda \in \Lambda_1$ then $\lambda' \in \Lambda_1$ for any $\lambda' > \lambda$. Hence Λ_1 can be either empty, or $(0, +\infty)$, or $[\lambda_{cr}, \infty)$ or (λ_{cr}, ∞) for some $\lambda_{cr} > 0$. We will show that the latter possibility takes place, actually. To do so, we define another set, *a priori* contained in Λ_1, formed by the factors $\lambda > 0$ so that there is a probability measure $\mu \in \mathcal{P}_0$ (so $p_0(\mu) = 0$) which is λ-zooming out:

$$\Lambda_2 = \{\lambda \mid \exists\, \mu \in \mathcal{P}_0 \text{ s.t. } \Phi_\lambda[\mu] \succcurlyeq \mu \text{ and } \exists\, x \in \mathbf{R}_+ \text{ s.t. } \mu \text{ is supported on } [0, x]\}.$$

Both sets actually coincide with Λ, as we shall prove later.

Non-emptiness of Λ_1. Choose any $p \in (0, p_{cr})$ and consider the probability measure $\mu \in \mathcal{P}$ defined as $\mu = p \cdot \mathrm{Dirac}_0 + (1-p) \cdot \mathrm{Dirac}_1$. We want to show that such μ is λ-zooming out for a certain $\lambda > 0$. This time it is more convenient to work with the distribution function of μ

$$F_\mu = p \cdot \mathbf{1}_{x<1} + \mathbf{1}_{x \geq 1}.$$

In order to keep notations not too heavy, we prefer writing F for F_μ and $\Phi_\lambda[F]$ instead of $F_{\Phi_\lambda[\mu]}$.

Remark 2.6. This abuse of notation will appear often in the rest of the paper: we will specify the probability measure associated to a distribution function only when this choice could generate some confusion.

Using Lemma 2.2 we remark that $\Phi_1[F](0) = \theta(F(0)) = \theta(p)$; since p is smaller than p_{cr} and 0 is an attracting fixed point for θ, we get $\Phi_1[F](0) < F(0)$, for $F(0)$ is equal to p. We observe also that for any $\lambda > 0$, $\Phi_\lambda[F](1)$ is equal to $\Phi_1[F](\lambda^{-1})$ and as λ goes to infinity, $\Phi_\lambda[F](1)$ tends to $\theta(p) = \Phi_1[F](0)$. This means that for some sufficiently large λ we must have $\Phi_\lambda[F](1) < p$. Taking such a λ, we notice that for $x < 1$ we have

$$\Phi_\lambda[F](x) \leq \Phi_\lambda[F](1) < p = F(x),$$

and similarly, when $x \geq 1$, we find

$$\Phi_\lambda[F](x) \leq 1 = F(x).$$

Therefore we conclude that λ is in Λ_1 and Λ_1 is non-empty.

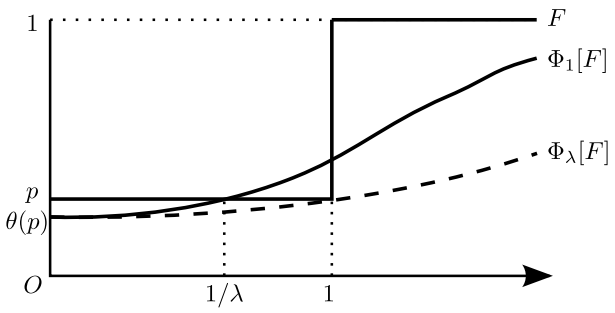

Figure IV.2 – How to show the existence of a λ such that there is a measure λ-zooming out.

Non-emptiness of $(0, +\infty) \setminus \Lambda_2$. Take any $q \in (p_{cr}, 1)$ and consider the probability measure

$$\mu_- = q \cdot \mathrm{Dirac}_1 + (1-q) \cdot \mathrm{Dirac}_\infty,$$

whose distribution function is $F_- = q \cdot \mathbf{1}_{x \geq 1} + (1-q) \cdot \mathbf{1}_\infty$.

Again Lemma 2.2 gives $\lim_{x \to +\infty} \Phi_1[F_-](x) = \theta(q)$ and so $\lim_{x \to +\infty} \Phi_1[F_-](x)$ is larger than q. Proceeding as before, we see that for some sufficiently small λ we have $\Phi_\lambda[F_-](1) > q$. Hence, by the very expression of F_-, we have just shown that F_- moves up under Φ_λ:

$$\Phi_\lambda[F_-] \geq F_- \quad (\text{or } \Phi_\lambda[\mu_-] \preccurlyeq \mu_-).$$

Moreover due to the scale-invariance (Remark 1.2), given any positive c, we have $\Phi_\lambda[F_c] \geq F_c$ for the function F_c defined by $F_c(x) = F_-(x/c)$ (and associated to the probability measure μ_c defined by $\mu_c(E) = \mu_-(c^{-1}E)$):

$$\Phi_\lambda[\mu_c] = \Phi_\lambda[\mu_-(c^{-1}\,\cdot)] = \Phi_\lambda[\mu_-](c^{-1}\,\cdot) \preccurlyeq \mu_-(c^{-1}\,\cdot) = \mu_c.$$

We want to show that any sufficiently small λ does not belong to Λ_2. So we fix any probability measure $\mu \in \mathcal{P}_0$ and denote by $F = F_\mu$ its distribution function. We shall show that μ is not λ-zooming out. Define c to be its q-quantile:

$$c = \min\{x \,:\, F(x) = \mu([0,x]) \geq q\}.$$

Then for such c, we tautologically have $F(x) \geq F_c(x)$; by the monotonicity of Φ_λ, we get the relation $\Phi_\lambda[F] \geq \Phi_\lambda[F_c]$. Now, if μ is λ-zooming out, we shall have $q \geq \Phi_\lambda[F](c^-)$. But this is not possible, since $\Phi_\lambda[F](c^-)$ is larger than $\Phi_\lambda[F_c](c^-)$ which is larger than q.

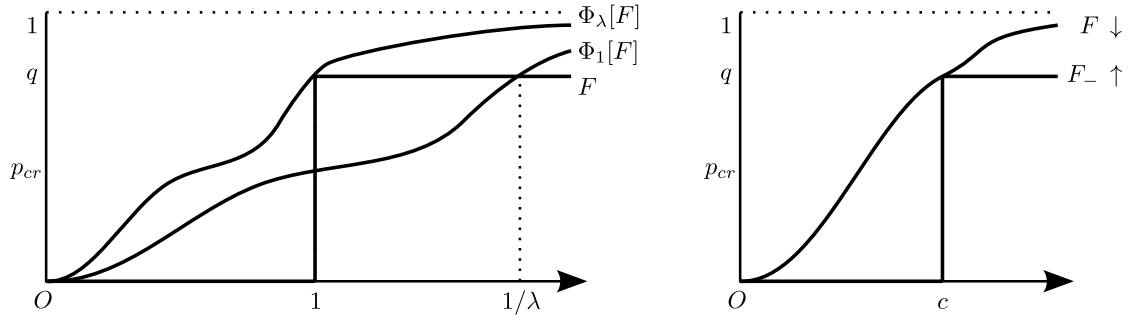

Figure IV.3 – How to show that the complementary set of Λ_2 is non-empty.

Equality $\Lambda_1 = \Lambda_2 = \Lambda$. We already remarked that we have the inclusion $\Lambda_2 \subset \Lambda_1$ directly from the definitions. Moreover Lemma 2.5 gives $p_0(\nu_{A,\lambda}) = 0$, so that the inclusion $\Lambda \subset \Lambda_2$ holds as well. It remains to prove $\Lambda_1 \subset \Lambda$.

Assume that $\lambda \in \Lambda_1$. By definition there exist a measure $\mu \in \mathcal{P}$ and $A \in \mathbf{R}_+$ such that μ is λ-zooming out, supported on $[0, A]$. Let us prove that $\mu \preccurlyeq \nu_{\lambda,A}$.

We first show by induction that for the sequence of measures $\mu_n^{A,\lambda}$ defined by (IV.6), for any n the inequality

$$\mu \preccurlyeq \mu_n^{A,\lambda} \tag{IV.9}$$

holds Indeed, for $n = 0$ we have $\mu \preccurlyeq \mathrm{Dirac}_\infty = \mu_0^{A,\lambda}$. Now, if (IV.9) holds for some n, applying $\Phi_{A,\lambda}$ we get

$$\mu \preccurlyeq \Phi_\lambda[\mu] \preccurlyeq \Phi_\lambda[\mu_n^{A,\lambda}] = \mu_{n+1}^{A,\lambda}.$$

The induction argument then shows that (IV.9) holds for all n, and passing to the limit as $n \longrightarrow \infty$ we obtain $\mu \preccurlyeq \nu_{A,\lambda}$. Hence $\nu_{A,\lambda} \neq \mathrm{Dirac}_0$. □

The following lemma completes the description of the set Λ.

Lemma 2.7 (Key Lemma). *The set Λ is open; thus $\Lambda = (\lambda_{cr}, +\infty)$ for some $\lambda_{cr} > 0$.*

Let us deduce Theorem 1.1 from it:

Proof of Theorem 1.1. The proof consists of two parts: first we show the existence of a non-trivial stationary measure (and the technical arguments are hidden in the Key Lemma 2.7), then we explain that such non-trivial stationary measure cannot have atoms on $[0, +\infty]$. The latter part of the proof is longer and rather technical, but contains essentially the arguments that we shall use when proving Lemma 2.7.

Stage 1: *Existence of a non-trivial stationary measure.* Note first that the following statement holds:

Lemma 2.8. *For any given $A \in \mathbf{R}_+$ we have*

$$\nu_{A,\lambda} \longrightarrow \mathrm{Dirac}_0 \text{ as } \lambda \searrow \lambda_{cr}. \tag{IV.10}$$

Proof. As $\lambda_{cr} \notin \Lambda$, for any given $A \in \mathbf{R}_+$ we have $\nu_{A,\lambda_{cr}} = \mathrm{Dirac}_0$. On the other hand, for any given $A \in \mathbf{R}_+$ the family of measures $\mu_n^{A,\lambda}$ is $\preccurlyeq$-decreasing *both* as n tends to ∞ and as λ tends to λ_{cr} from the right. Hence, the double limit $\lim_{n\longrightarrow\infty,\,\lambda\searrow\lambda_{cr}} \mu_n^{A,\lambda}$ exists, and is equal to each of the two repeated limits:

$$\lim_{n\longrightarrow\infty,\,\lambda\searrow\lambda_{cr}} \mu_n^{A,\lambda} = \lim_{n\longrightarrow\infty} \lim_{\lambda\searrow\lambda_{cr}} \mu_n^{A,\lambda} = \lim_{\lambda\searrow\lambda_{cr}} \lim_{n\longrightarrow\infty} \mu_n^{A,\lambda}. \tag{IV.11}$$

Recall that $\mu_n^{A,\lambda} = \Phi_{A,\lambda}^n[\mathrm{Dirac}_\infty]$. Due to the continuity of the transformation $\Phi_{A,\lambda}[\mu]$ in both λ and μ, for any fixed n one thus have $\mu_n^{A,\lambda} \longrightarrow \mu_n^{A,\lambda_{cr}}$ as $\lambda \searrow \lambda_{cr}$. Hence, the first of the two repeated limits is equal to $\nu_{A,\lambda_{cr}} = \mathrm{Dirac}_0$:

$$\lim_{n\longrightarrow\infty} \lim_{\lambda\searrow\lambda_{cr}} \mu_n^{A,\lambda} = \lim_{n\longrightarrow\infty} \mu_n^{A,\lambda_{cr}} = \nu_{A,\lambda_{cr}}$$

The other repeated limit in (IV.11) is

$$\lim_{\lambda \searrow \lambda_{cr}} \lim_{n \longrightarrow \infty} \mu_n^{A,\lambda} = \lim_{\lambda \searrow \lambda_{cr}} \nu_{A,\lambda}.$$

Hence, for any given A we have $\nu_{A,\lambda} \longrightarrow \mathrm{Dirac}_0$ as $\lambda \searrow \lambda_{cr}$. □

Now, let us properly rescale these measures, in order to keep them non-trivial in the (subsequential) limit. Namely, consider these measures for a particular choice $A = 1$, and take the $\frac{1}{2}$-quantile

$$\kappa_{\frac{1}{2}}(\lambda) := \min\left\{x \,:\, \nu_{1,\lambda}([0,x]) \geq \tfrac{1}{2}\right\}.$$

Remark 2.7. The value $\frac{1}{2}$ was chosen here for its simplicity; as the reader will see later, we could take any value in the interval $(p_{cr}, 1)$.

From Lemma 2.8 we have

Corollaire 2.9.

$$\lim_{\lambda \searrow \lambda_{cr}} \kappa_{\frac{1}{2}}(\lambda) = 0. \tag{IV.12}$$

Proof. For any $\varepsilon > 0$, the weak convergence (IV.10) implies that $\nu_{1,\lambda}([0,\varepsilon]) \longrightarrow 1$ as $\lambda \searrow \lambda_{cr}$, and hence $\kappa_{\frac{1}{2}}(\lambda) < \varepsilon$ for all λ in some right neighbourhood of λ_{cr}. □

The absence of atoms for $\nu_{A,\lambda}$ on $[0,1)$ implies that $\kappa_{\frac{1}{2}}(\lambda) > 0$ for any λ. Therefore, once $\kappa_{\frac{1}{2}}(\lambda) < 1$ (which holds in some right neighbourhood of λ_{cr}), we have the equality $\nu_{1,\lambda}([0, \kappa_{\frac{1}{2}}(\lambda)]) = \frac{1}{2}$.

Take $A(\lambda) := \frac{1}{\kappa_{\frac{1}{2}}(\lambda)}$; the scaling relation (IV.7) implies that for λ in the same right neighbourhood of λ_{cr},

$$\nu_{A(\lambda),\lambda}([0,1]) = \nu_{1,\lambda}([0, \tfrac{1}{A(\lambda)}]) = \nu_{1,\lambda}([0, \kappa_{\frac{1}{2}}(\lambda))]) = \tfrac{1}{2}, \tag{IV.13}$$

while (IV.12) implies that $A(\lambda) \longrightarrow +\infty$ as $\lambda \searrow \lambda_{cr}$.

The family $\nu_{A(\lambda),\lambda}$ is a family of probability measures on the compactified half-line $[0,\infty]$, hence there exists a convergent subsequence $\nu_{A(\lambda_j),\lambda_j}$ for some subsequence $\lambda_j \searrow \lambda_{cr}$.

Remark 2.8. Note that this family is no longer $\preccurlyeq$-monotone, as increasing A and decreasing λ lead to $\preccurlyeq$-inequalities in opposite directions; apparently, there is no easy and direct way to prove the convergence of the family $\left(\nu_{A(\lambda),\lambda}\right)_\lambda$.

We conclude the first step of this proof with the following

Lemma 2.10. *Any subsequential limit $\bar{\mu} = \lim_{j \longrightarrow \infty} \nu_{A(\lambda_j),\lambda_j}$, with $\lambda_j \searrow \lambda_{cr}$, is $\Phi_{\lambda_{cr}}$-stationary:*

$$\bar{\mu} = \Phi_{\lambda_{cr}}[\bar{\mu}].$$

Proof. Recall that for any λ the measure $\nu_{A(\lambda),\lambda}$ is $\Phi_{A(\lambda),\lambda}$-stationary

$$\nu_{A(\lambda),\lambda} = \Phi_{A(\lambda),\lambda}[\nu_{A(\lambda),\lambda}]. \tag{IV.14}$$

Passing in (IV.14) to the limit along the subsequence λ_j, and recalling that due to Remark 2.4 the operator on the right hand side is continuous in all the three arguments, we have

$$\bar{\mu} = \Phi_{\infty,\lambda_{cr}}[\bar{\mu}] = \Phi_{\lambda_{cr}}[\bar{\mu}].$$

That is, the measure $\bar{\mu}$ is $\Phi_{\lambda_{cr}}$-stationary. □

Since the measure $\bar{\mu}$ has no atoms on $(0,+\infty)$ due to its $\Phi_{\lambda_{cr}}$-stationarity, the relation (IV.13) gives us $\bar{\mu}([0,1]) = \frac{1}{2}$ when passing to the weak limit. This shows the non-triviality of $\bar{\mu}$.

Stage 2: *The constructed non-trivial stationary measure has no atoms.* As we have just noticed, the measure $\bar{\mu}$ has no atoms on $(0,+\infty)$, so in order to conclude the proof of Theorem 1.1, we have to show that the constructed measure $\bar{\mu}$ is supported on $(0,+\infty)$ (and does not charge neither 0 nor ∞).

Lemma 2.2 implies that both $p_0(\bar{\mu})$ and $1 - p_\infty(\bar{\mu})$ (*i.e.* the probabilities of zero and finite lengths respectively) are fixed points of the map θ, and thus belong to $\{0, p_{cr}, 1\}$. The two inequalities

$$\begin{cases} p_0(\bar{\mu}) \leq \bar{\mu}([0,1]) \leq 1 - p_\infty(\bar{\mu}), \\ \bar{\mu}([0,1]) = \frac{1}{2} > p_{cr}, \end{cases}$$

imply $1 - p_\infty = 1$ and hence the measure $\bar{\mu}$ does not charge $+\infty$.

More involved is to prove that the measure $\bar{\mu}$ does not charge the point 0. The above argument implies that there are two possible values for $p_0(\bar{\mu})$: either $p_0(\bar{\mu}) = 0$ or $p_0(\bar{\mu}) = p_{cr}$. We shall assume for the rest of this proof that $p_0(\bar{\mu}) = p_{cr}$, and try to get a contradiction.

Under this standing assumption we can decompose the probability law $\bar{\mu}$ according to the convex combination (IV.4):

$$\bar{\mu} = (1 - p_{cr})\, \bar{\mu}_0 + p_{cr}\, \mathrm{Dirac}_0\,.$$

We want to find a measure μ which λ_{cr}-zooms out, contradicting the Key Lemma. We will construct the measure μ in two steps:

i) We reduce the weight of the atom at zero: that is, we consider a family of measures

$$\bar{\mu}_p = (1 - p)\, \bar{\mu}_0 + p\, \mathrm{Dirac}_0.$$

We prove then that for sufficiently small p, this measure satisfies the condition $\bar{\mu}_p \preccurlyeq \Phi_{\lambda_{cr}}[\bar{\mu}_p]$. Moreover, we obtain here a stronger strict inequality for their partition functions (see Lemma 2.12 below):

$$F_{\bar{\mu}_p}(x) > F_{\Phi_{\lambda_{cr}}[\bar{\mu}_p]}(x) \quad \text{for every } x \in [0, +\infty). \tag{IV.15}$$

In particular, for any $x \in \mathbf{R}_+$, there exists some $\delta = \delta(x)$ such that

$$F_{\Phi_{\lambda_{cr}}[\bar{\mu}_p]}(y) + \delta < F_{\bar{\mu}_p}(y) \quad \text{for every } y \in [0, x]. \tag{IV.16}$$

If the measure $\bar{\mu}_p$ was supported on some finite interval, this would immediately mean that $\bar{\mu}_p$ zooms out; as it is not, we have to modify it to make it compactly supported (while not losing the inequality $\bar{\mu} \preccurlyeq \Phi_{\lambda_{cr}}[\bar{\mu}]$ during this process). From the inequality (IV.16) one can figure out why such a modification can be done.

ii) We modify the measure $\bar{\mu}_p$ "near infinity", so that it becomes compactly supported, without destroying the inequality (IV.15). The problem here is that the partition functions approach each other at infinity, so there is no immediate cut-off-like argument. Such a modification is given by Lemma 2.13 below, and it provides us with a λ_{cr}-zooming out measure and thus with the desired contradiction.

Following the road-map above, let us decompose the $\Phi_{\lambda_{cr}}$-image of the measure $\bar{\mu}$ into several components. Namely we have the following three possibilities for the four lengths $X_1, \dots, X_4$:

(a) There is an *IO*-path of zero length. This happens exactly with probability $\theta(p_{cr}) = p_{cr}$.

(b) All of them are non-zero. This happens with probability $(1 - p_{cr})^4$ and conditionally on this, the law of $R_{\lambda_{cr}}(X_1, \dots, X_4; \xi)$ is the $\Phi_{\lambda_{cr}}$-image of the conditional measure $\bar{\mu}_0$.

(c) There is no *IO*-path of zero length but there is at least one collapsed length. Conditionally on this case, the law of $R_{\lambda_{cr}}(X_1, \dots, X_4; \xi)$ is given by the law of $\lambda_{cr}\xi \cdot \min(X_1', X_2')$ where the independent random variables X_1', X_2' are both distributed with respect to $\bar{\mu}_0$.

The last case suggests to define an operator Φ'_λ on $\mathcal{P}_0$ ($\lambda > 0$), assigning to any measure $\mu_0 \in \mathcal{P}_0$ the law of $\lambda\xi \cdot \min(X_1, X_2)$, where the independent random variables X_1, X_2 are both distributed with respect to μ_0.

Immediately, we observe:

Lemma 2.11. *We have*

$$\Phi_{\lambda_{cr}}[\mu_0] \preccurlyeq \Phi'_{\lambda_{cr}}[\mu_0]$$

for any probability measure $\mu_0 \in \mathcal{P}_0$.

Proof. Indeed, this means geometrically that we almost always have a longer path if we have to pass along two edges instead of one only. More formally, we can use a coupling argument: we draw four i.i.d. random variables $X_1, \dots, X_4$ distributed with respect to μ_0 and ξ distributed with respect to m and independent of the previous variables, defined on a probability space $(\Omega, \mathbf{P})$. Then there is an evident inequality

$$\lambda\, (\min(X_1(\omega), X_2(\omega)) + \min(X_3(\omega), X_4(\omega))) \geq \lambda\, \min(X_1(\omega), X_2(\omega))$$

for almost every $\omega \in (\Omega, \mathbf{P})$ and hence

$$\lambda\xi(\omega)\, (\min(X_1(\omega), X_2(\omega)) + \min(X_3(\omega), X_4(\omega))) \geq \lambda\xi(\omega)\, \min(X_1(\omega), X_2(\omega)).$$

□

We concentrate the remaining arguments for the first step in the next Lemma.

Lemma 2.12. *Assume that*

$$\bar{\mu} = p_{cr}\,\mathrm{Dirac}_0 + (1-p_{cr})\,\bar{\mu}_0$$

is $\Phi_{\lambda_{cr}}$-invariant, and let $\bar{\mu}_p := p\,\mathrm{Dirac}_0 + (1-p)\,\bar{\mu}_0$. Then, for any sufficiently small p, we have $\bar{\mu}_p \preccurlyeq \Phi_{\lambda_{cr}}[\bar{\mu}_p]$, and moreover there is a strict inequality for the partition functions, bounded away from zero on compact intervals: for any $x \in \mathbf{R}_+$ there exists $\delta = \delta(x) > 0$ such that

$$F_{\Phi_{\lambda_{cr}}[\bar{\mu}_p]}(y) + \delta < F_{\bar{\mu}_p}(y) \quad \textit{for every } y \in [0,x].$$

Proof. The analysis of the different possibilities (a-c) shows the equality

$$\Phi_{\lambda_{cr}}[\bar{\mu}_p] = \theta(p)\,\mathrm{Dirac}_0 + (1-p)^4\,\Phi_{\lambda_{cr}}[\bar{\mu}_0] + (1-\theta(p)-(1-p)^4)\,\Phi'_{\lambda_{cr}}[\bar{\mu}_0]. \tag{IV.17}$$

Let us rewrite (IV.17) in the following way:

$$\Phi_{\lambda_{cr}}[\bar{\mu}_p] = \theta(p)\,\mathrm{Dirac}_0 + (1-\theta(p))\cdot\Big((1-q(p))\,\Phi'_{\lambda_{cr}}[\bar{\mu}_0] + q(p)\,\Phi_{\lambda_{cr}}[\bar{\mu}_0]\Big), \tag{IV.18}$$

with $q(p) = \frac{(1-p)^4}{1-\theta(p)}$. The stationarity of $\bar{\mu} = \bar{\mu}_{p_{cr}}$ then implies

$$p_{cr}\,\mathrm{Dirac}_0 + (1-p_{cr})\bar{\mu}_0 = \theta(p_{cr})\,\mathrm{Dirac}_0 + (1-\theta(p_{cr}))\cdot\Big((1-q(p_{cr}))\,\Phi'_{\lambda_{cr}}[\bar{\mu}_0] + q(p_{cr})\,\Phi_{\lambda_{cr}}[\bar{\mu}_0]\Big),$$

and as $p_{cr} = \theta(p_{cr})$, we have the convex combination

$$\bar{\mu}_0 = (1-q(p_{cr}))\,\Phi'_{\lambda_{cr}}[\bar{\mu}_0] + q(p_{cr})\,\Phi_{\lambda_{cr}}[\bar{\mu}_0]. \tag{IV.19}$$

Due to Lemma 2.11, we have

$$\Phi'_{\lambda_{cr}}[\bar{\mu}_0] \preccurlyeq \bar{\mu}_0 \preccurlyeq \Phi_{\lambda_{cr}}[\bar{\mu}_0].$$

Now, using the expression (IV.19),we see that $q(p) \longrightarrow 1 > q(p_{cr})$ as $p \longrightarrow 0$; hence, for any sufficiently small p we have

$$\bar{\mu}_0 \preccurlyeq (1-q(p))\,\Phi'_{\lambda_{cr}}[\bar{\mu}_0] + q(p)\,\Phi_{\lambda_{cr}}[\bar{\mu}_0], \tag{IV.20}$$

and thus, using (IV.18) and (IV.20),

$$\Phi_{\lambda_{cr}}[\bar{\mu}_p] \succcurlyeq \theta(p)\,\mathrm{Dirac}_0 + (1-\theta(p))\,\bar{\mu}_0 = \bar{\mu}_{\theta(p)}. \tag{IV.21}$$

As $\theta(p) < p$ for $0 < p < p_{cr}$, we have $\bar{\mu}_p \preccurlyeq \bar{\mu}_{\theta(p)} = \Phi_{\lambda_{cr}}[\bar{\mu}_p]$ for all sufficiently small p. Moreover, for any sufficiently small p, using (IV.21), we have for any $x \in \mathbf{R}_+$

$$\begin{aligned}\inf_{y\in[0,x]}(F_{\bar{\mu}_p}(y) - F_{\Phi_{\lambda_{cr}}[\bar{\mu}_p]}(y)) &\geq \inf_{y\in[0,x]}(p-\theta(p))\cdot(F_{\mathrm{Dirac}_0}(y) - F_{\bar{\mu}_0}(y))\\ &= (p-\theta(p))\cdot(1-F_{\bar{\mu}_0}(x)) =: \delta(x) > 0.\end{aligned}$$

This concludes the proof of the lemma. □

Let us fix $p > 0$ given by Lemma 2.12, and let $\mu := \bar{\mu}_p$. The second step of the proof is to modify μ so that it becomes compactly supported, without destroying the inequality $\mu \preccurlyeq \Phi_{\lambda_{cr}}[\mu]$, and thus to obtain a λ_{cr}-zooming out measure. The following construction allows to do so.

Lemma 2.13. *Let $\mu \neq \mathrm{Dirac}_0$ be a probability measure on $[0,+\infty)$ with no atoms on $(0,+\infty)$ such that $\mu \preccurlyeq \Phi_\lambda[\mu]$ for some $\lambda > 0$. Let X be a random variable (defined on some probability space $(\Omega,\mathbf{P})$) distributed according to μ and denote by x_1 its $\frac{3}{4}$-quantile. Assume that there exists $\delta > 0$ such that*

$$F_{\Phi_\lambda[\mu]}(y) + \delta < F_\mu(y) \quad \textit{for all } y \in [0,x_1],$$

where $F_{\Phi_\lambda[\mu]}$ and F_μ stay for the partition functions of $\Phi_\lambda[\mu]$ and μ respectively.

Take x_2 to be the $(1-\delta)$-quantile of μ, and take the measure $\widetilde{\mu}$ to be the law of the random variable $\widetilde{X}$ defined (on the same probability space) as

$$\widetilde{X}(\omega) = \begin{cases} X(\omega) & \textit{if } X(\omega) \leq x_2,\\ x_1 & \textit{otherwise.}\end{cases} \tag{IV.22}$$

Then the measure $\widetilde{\mu}$ is λ-zooming out: $\Phi_\lambda[\widetilde{\mu}] \succcurlyeq \widetilde{\mu}$.

Proof. We start with the technical observation that $x_2 \geq x_1$: the inequality $F_{\Phi_\lambda[\mu]}(x_1) > F_\mu(x_1) + \delta$ implies that $F_\mu(x_1) < 1 - \delta$. In particular, this implies $\widetilde{X} \leq X$ almost surely, and hence $\widetilde{\mu} \preccurlyeq \mu$. It is also not difficult to find an explicit expression for the partition function of $\widetilde{\mu}$:

$$F_{\widetilde{\mu}}(x) = \begin{cases} F_\mu(x) & \text{if } x \in [0, x_1), \\ F_\mu(x) + \delta & \text{if } x \in [x_1, x_2), \\ 1 & \text{if } x \in [x_2, +\infty). \end{cases}$$

Indeed, the coupling (IV.22) says that the measure μ is transformed in the following way: all the mass from $[x_2, +\infty)$ is collapsed into a single atom put at the point x_1. Thus, the partition function of $\widetilde{\mu}$ stays unchanged on $[0, x_1)$, changes by a constant on $[x_1, x_2)$, and is identically equal to 1 on $[x_2, +\infty)$; this easily implies the above representation.

Since $\widetilde{\mu} \preccurlyeq \mu$, applying Φ_λ gives the relation $\Phi_\lambda[\widetilde{\mu}] \preccurlyeq \Phi_\lambda[\mu]$. Though, the direction of the latter inequality does not help, so we want to *quantify* it:

$$F_{\Phi_\lambda[\widetilde{\mu}]}(x) - F_{\Phi_\lambda[\mu]}(x) \leq \delta \quad \text{for every } x \in [0, +\infty). \tag{IV.23}$$

To prove such an estimate, take four independent random vectors $(X_i, \tilde{X}_i)$, $i = 1, \dots, 4$, coupling the measures μ and $\widetilde{\mu}$ in the way described by (IV.22), as well as a random variable ξ, independent of them all. The random variables $R_\lambda(X_1, \dots, X_4; \xi)$ and $R_\lambda(\tilde{X}_1, \dots, \tilde{X}_4; \xi)$ are distributed according to $\Phi_\lambda[\mu]$ and $\Phi_\lambda[\widetilde{\mu}]$ respectively; in particular, the probability of the event

$$\left\{R_\lambda(X_1, \dots, X_4; \xi) \neq R_\lambda(\tilde{X}_1, \dots, \tilde{X}_4; \xi)\right\} \tag{IV.24}$$

is an upper bound for the left hand side of the inequality in (IV.23). A first remark is that, as for any $i = 1, \dots, 4$ one has

$$\mathbf{P}(X_i \neq \tilde{X}_i) = \mathbf{P}(X_i > x_2) = \delta,$$

the probability of the event (IV.24) does not exceed

$$\mathbf{P}(\exists\, i : \; X_i \neq \tilde{X}_i) \leq \sum_{i=1}^{4} \mathbf{P}(X_i \neq \tilde{X}_i) = 4\delta.$$

This estimate is weaker than the one we want to prove (and with such a weaker estimate, the arguments at the end of this proof would not work). So we need an additional idea: for the shortest *IO* paths to have different lengths in graphs with edges of length X_j's and $\tilde{X}_j$'s respectively, not only there should exist j such that $X_j \neq \tilde{X}_j$, but also the shortest path in one of the graphs should pass through such edge j. And as the inequality $X_j \neq \tilde{X}_j$ implies that both these lengths are relatively big (no less than x_1 which was chosen as the $\frac{3}{4}$-quantile), this is even less probable.

Formalizing this idea, we note that the event (IV.24) is covered by a union of four events of the kind: "on a given edge j we have $X_j > \tilde{X}_j$, and on the edge i parallel to j the associated X_i-length is not smaller than x_1". Indeed, if the length of the parallel edge is smaller than x_1, its presence erases both the greater lengths X_j and $\widetilde{X}_j$.

The probability of each of these events is equal to

$$\mathbf{P}(X_j \neq \tilde{X}_j) \cdot \mathbf{P}(X_i \geq x_1) = \tfrac{1}{4} \cdot \delta.$$

Hence, the total probability does not exceed $4 \cdot \frac{1}{4}\delta = \delta$, as claimed.

Let us now deduce from (IV.23) the conclusion of the lemma. Within the interval $[0, x_1)$ we have the chain of inequalities

$$F_{\widetilde{\mu}}(y) = F_\mu(y) > F_{\Phi_\lambda[\mu]}(y) + \delta \geq F_{\Phi_\lambda}[\widetilde{\mu}](y)$$

and similarly on the interval $[x_1, x_2]$ we have

$$F_{\widetilde{\mu}}(y) = F_\mu(y) + \delta > F_{\Phi_\lambda[\mu]}(y) + \delta \geq F_{\Phi_\lambda}[\widetilde{\mu}](y)$$

Lastly, for $y > x_2$ we have $F_{\widetilde{\mu}}(y) = 1 > F_{\Phi_\lambda[\widetilde{\mu}]}(y)$.

Joining the three estimates together, we obtain $F_{\widetilde{\mu}}(y) > F_{\Phi_\lambda[\widetilde{\mu}]}(y)$ for all $y \in \mathbf{R}_+$, hence $\widetilde{\mu} \preccurlyeq \Phi_\lambda[\widetilde{\mu}]$. By its definition, $\widetilde{\mu}$ is supported on a finite interval $[0, x_2]$, so we conclude that $\widetilde{\mu}$ is a λ-zooming out probability measure. □

With the previous lemma we have shown that λ_{cr} is a supercritical value, providing us with the desired contradiction, that comes from the assumption $p_0(\bar{\mu}) \neq 0$. We can thus conclude that $p_0(\bar{\mu}) = \bar{\mu}(\{0\})$ is equal to 0 and hence $\bar{\mu}$ is non-atomic. This concludes the proof of the theorem. □

2.3 Openness of the supercritical set Λ

Let us now prove the Key Lemma. The main idea here will be the following: assume $\lambda \in \Lambda$ and take the measure $\nu_{A,\lambda}$ which is not trivial, then increasing slightly the cut-off value A we obtain a measure $\widetilde{\nu}$ verifying $F_{\Phi_\lambda[\widetilde{\nu}]}(x) < F_{\widetilde{\nu}}(x)$ for every $x \in \mathbf{R}_+$. Hence the previous strict inequality may be used to find a $(\lambda - \varepsilon)$-zooming out measure; the main problem is to control the inequality in a neighbourhood of 0 and this will require arguments which are analogous to those of Lemma 2.12.

Proof of Lemma 2.7. Assume $\lambda \in \Lambda$. Our goal is to show that for sufficiently small $\varepsilon > 0$ we have $(\lambda - \varepsilon) \in \Lambda$, thus establishing the openness of Λ.

The measure $\nu_{A,\lambda}$ is a fixed point for the cut-off operator $\Phi_{A,\lambda}$ and hence, denoting by $G_{A,\lambda}$ its distribution function, we have

$$G_{A,\lambda} = \max(\Phi_\lambda[G_{A,\lambda}], \mathbf{1}_{x \geq A}).$$

Moreover $G_{A,\lambda} \not\equiv 1$ ($\nu_{A,\lambda} \neq \text{Dirac}_0$) implies that $\Phi_\lambda[G_{A,\lambda}]$ is everywhere strictly less than one (due to the full support assumption for the measure m). Hence $G_{A,\lambda}(x) = \Phi_\lambda[G_{A,\lambda}](x)$ for $x < A$ and $G_{A,\lambda}(x) > \Phi_\lambda[G_{A,\lambda}](x)$ for $x \geq A$ (see the picture on the left in Figure IV.4).

Take now $A' > A$ and define $\widetilde{G} := \max(\Phi_\lambda[G_{A,\lambda}], \mathbf{1}_{x \geq A'})$ (corresponding to $\widetilde{\nu} = \Phi_{A',\lambda}[\nu_{A,\lambda}]$). Then we have $\widetilde{G} \leq G_{A,\lambda}$ and moreover $\widetilde{G}(x) < G_{A,\lambda}(x)$ for $x \in (A, A')$. Again due to the full support assumption for the measure m, this implies

$$\Phi_\lambda[\widetilde{G}](x) < \Phi_\lambda[G_{A,\lambda}](x) \leq \Phi_{A',\lambda}[G_{A,\lambda}](x) = \widetilde{G}(x) \quad \text{for every } x \in \mathbf{R}_+. \tag{IV.25}$$

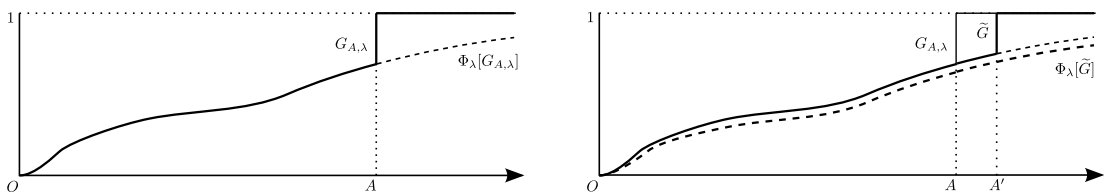

Figure IV.4 – The fixed function $G_{A,\lambda}$ for the cut-off operator $\Phi_{A,\lambda}$, its modification $\widetilde{G}$ and their Φ_λ-images

We would like to decrease λ, keeping a zooming out measure. For any compact interval $J \subset (0, +\infty)$ the strict inequality in (IV.25) between the continuous functions $\Phi_\lambda[\widetilde{G}]$ and $\Phi_\lambda[G_{A,\lambda}]$ on J is preserved under a small perturbation: considering $J = [0, A']$, it is possible to find ε_0 such that for any $\varepsilon < \varepsilon_0$ and $x \leq A'$,

$$\Phi_{\lambda-\varepsilon}[\widetilde{G}](x) < \Phi_\lambda[G_{A,\lambda}](x) \leq \widetilde{G}(x).$$

The inequality $\Phi_{\lambda-\varepsilon}[\widetilde{G}](x) \leq \widetilde{G}(x)$ also automatically holds for all $x \geq A'$, as $\widetilde{G}(x) = 1$ for any such x. The only problem is thus to handle the neighbourhood of $x = 0$.

To do so, let us modify the measure $\widetilde{\nu}$ by mixing it with a Dirac mass at zero: consider the family of measures $\widetilde{\nu}_p := (1-p)\widetilde{\nu} + p\,\text{Dirac}_0$. We claim that for all sufficiently small p, we still save the strict inequality between the partition functions of $\widetilde{\nu}_p$ and of its image $\Phi_\lambda[\widetilde{\nu}_p]$: there exists p_0 such that for every $p < p_0$ and $x \in [0, +\infty)$,

$$F_{\Phi_\lambda[\widetilde{\nu}_p]}(x) < F_{\widetilde{\nu}_p}(x). \tag{IV.26}$$

Once (IV.26) is established (and this can be done following the proof of Lemma 2.12), concluding the proof of the lemma is almost immediate. Indeed, observe that for any $p \in (0, p_{cr})$ both partition functions $F_{\widetilde{\nu}_p}$ and $F_{\Phi_\lambda[\widetilde{\nu}_p]}$ are continuous on $[0, +\infty)$ (in particular right-continuous at $x = 0$), moreover the strict inequality $F_{\widetilde{\nu}_p}(0) = p > \theta(p) = F_{\Phi_\lambda[\widetilde{\nu}_p]}(0)$ between their values is satisfied. Hence, for any such p, the strict inequality between the two continuous functions allows to find ε_1 and p_1 such that for any $\varepsilon < \varepsilon_1$ and $p < p_1$,

$$F_{\Phi_{\lambda-\varepsilon}[\widetilde{\nu}_p]}(x) < F_{\widetilde{\nu}_p}(x) \quad \text{for every } x \in [0, A'].$$

Once again the inequality $F_{\widetilde{\nu}_p}(x) \geq F_{\Phi_{\lambda-\varepsilon}[\widetilde{\nu}_p]}(x)$ is automatically satisfied as $x \geq A'$. So

$$\widetilde{\nu}_p \preccurlyeq \Phi_{\lambda-\varepsilon}[\widetilde{\nu}_p].$$

We focus now on deriving the inequality (IV.26). The equality can be obtained outside any arbitrarily small neighbourhood of 0, using the arguments of the last paragraph: for any $x_0 > 0$, the strict inequality $F_{\widetilde{\nu}} > F_{\Phi_\lambda[\widetilde{\nu}]}$

between the two continuous functions on the compact interval $[x_0, A']$ implies the same inequality for $\widetilde{\nu}_p$, for any $p < p_1 = p_1(x_0)$, while for $x > A'$ there is nothing to check.

Therefore the most important part of the proof consists in handling the inequality (IV.26) over a neighbourhood of the origin. In the same way as in (IV.17), we have

$$\Phi_\lambda[\widetilde{\nu}_p] = \theta(p)\,\mathrm{Dirac}_0 + (1-p)^4\,\Phi_\lambda[\widetilde{\nu}] + (1-\theta(p)-(1-p)^4)\,\Phi'_\lambda[\widetilde{\nu}].$$

and hence, as $p \longrightarrow 0$, we have

$$\begin{aligned} F_{\Phi_\lambda[\widetilde{\nu}_p]}(x) &= \theta(p) + (1-p)^4 F_{\Phi_\lambda[\widetilde{\nu}]}(x) + (1-\theta(p)-(1-p)^4)\,F_{\Phi'_\lambda[\widetilde{\nu}]}(x) \\ &= F_{\Phi_\lambda[\widetilde{\nu}]}(x) + 4p\,(F_{\Phi'_\lambda[\widetilde{\nu}]}(x) - F_{\Phi_\lambda[\widetilde{\nu}]}(x)) + O(p^2), \end{aligned}$$

where the $O(p^2)$ is uniform in $x \in [0, +\infty)$. At the same time,

$$\begin{aligned} F_{\widetilde{\nu}_p}(x) &= p + (1-p)F_{\widetilde{\nu}}(x) \\ &= F_{\widetilde{\nu}}(x) + p\cdot(1 - F_{\widetilde{\nu}}(x)) \\ &\geq F_{\Phi_\lambda[\widetilde{\nu}]}(x) + p\cdot(1 - F_{\widetilde{\nu}}(x)). \end{aligned}$$

Subtracting, we get

$$F_{\widetilde{\nu}_p}(x) - F_{\Phi_\lambda[\widetilde{\nu}_p]}(x) \geq p\cdot\Big((1 - F_{\widetilde{\nu}}(x)) - 4(F_{\Phi'_\lambda[\widetilde{\nu}]}(x) - F_{\Phi_\lambda[\widetilde{\nu}]}(x))\Big) + O(p^2). \tag{IV.27}$$

Note that both $F_{\widetilde{\nu}}(x)$ and $F_{\Phi_\lambda[\widetilde{\nu}]}(x)$ tend to 0 as $x \longrightarrow 0$. Hence there exists $x_0 > 0$ such that

$$F_{\widetilde{\nu}}(x_0), F_{\Phi_\lambda[\widetilde{\nu}]}(x_0) < \tfrac{1}{6}.$$

Substituting it into the right hand side of (IV.27), we get

$$(1 - F_{\widetilde{\nu}}(x)) - 4(F_{\Phi'_\lambda[\widetilde{\nu}]}(x) - F_{\Phi_\lambda[\widetilde{\nu}]}(x)) \geq 1 - \tfrac{5}{6} = \tfrac{1}{6},$$

which gives

$$F_{\widetilde{\nu}_p}(x) - F_{\Phi_\lambda[\widetilde{\nu}_p]}(x) > \tfrac{1}{6}p + O(p^2).$$

As $O(p^2)$ is uniform in x, this implies that for sufficiently small p we have $F_{\widetilde{\nu}_p}(x) > F_{\Phi_\lambda[\widetilde{\nu}_p]}(x)$ for all $x \in [0, x_0]$. The proof of the Key Lemma is now over. □

3 Convergence in law and uniqueness of the stationary measure

The purpose of this section is to show that, under suitable assumptions, the operator $\Phi_{\lambda_{cr}}$ defines a contraction on $\mathcal{P}$. This is the content of Theorem 1.2.

Fix a stationary measure $\bar{\mu}$ from the statement of Theorem 1.2, let $F_{\bar{\mu}}$ be its distribution function. For any $\alpha \in (0,1)$ denote by κ_α the α-quantile of the measure $\bar{\mu}$: let $\kappa_\alpha := F_{\bar{\mu}}^{-1}(\alpha)$. Note that these quantiles are well-defined. Indeed, $\bar{\mu}$ is a stationary measure and hence $F_{\bar{\mu}} = \Phi_{\lambda_{cr}}[F_{\bar{\mu}}]$; as the measure m is absolutely continuous with continuous positive density, $F_{\bar{\mu}}$ is a C^1 function with positive derivative on $\mathbf{R}_+$ (we operate a multiplicative convolution with m as the last calculation needed for obtaining its Φ_λ-image).

Throughout this section, we prefer to work with the space of distribution functions (cf. Remark 2.6).

Definition 3.1. *Denote by $\mathcal{F}$ the space of monotonic non-decreasing right-continuous functions on the extended half-line $[0, +\infty]$, satisfying $0 \leq F \leq 1$ and $F(+\infty) = 1$. We will denote by $\mathcal{F}_0$ the subspace of distribution functions of probability measures in $\mathcal{P}_0$ (and so verifying $F(0) = 0$ and $\lim_{x\to\infty} F(x) = 1$).*

3.1 The class $C_{\alpha,\delta}$ and contraction

Definition 3.2. *A distribution function $F \in \mathcal{F}$, is of* class (α, δ) *(where $\delta > 0$ and $\alpha \in (0, 1/2)$) if*

- *$F(x) \leq F_{\bar{\mu}}(x)$ on $[\kappa_\alpha, \kappa_{(1-\alpha)}]$,*
- *$F(x) \leq F_{\bar{\mu}}(x) + \delta$ on $[0, \kappa_\alpha) \cup (\kappa_{(1-\alpha)}, +\infty]$.*

We denote the set of such functions by $C_{\alpha,\delta}$.

We also consider the family of rescaling operators:

Definition 3.3. *For any $r \in \mathbf{R}$ let*

$$T_r : \mathcal{F} \longrightarrow \mathcal{F}, \quad T_r[F](x) = F(e^{-r}x).$$

The main result in this part is the following key proposition which gives a good family of rescaling operators "improving" the class of a given distribution function:

Proposition 3.1. *There exist constants $L > 0$ and $\alpha \in (0, 1/2)$ such that for every $\delta \in (0, \alpha]$ the operator $T_{L\delta}\Phi_{\lambda_{cr}}$ sends $C_{\alpha,\delta}$ in $C_{\alpha,\delta/2}$.*

Remark 3.1. Although not every distribution function $F \in \mathcal{F}$ belongs to some class $C_{\alpha,\delta}$, every distribution function in the smaller space $\mathcal{F}_0$ can be rescaled by some T_r to belong to a certain class. Moreover, for any such distribution function F and for *any* class $C_{\alpha,\delta}$ there exists a rescaling $T_r[F]$ that belongs to this class. It is also clear that there is no reason to consider $\delta > \alpha$: the class $C_{\alpha,\delta}$ coincides with $C_{\alpha,\alpha}$.

Using Proposition 3.1, we fix corresponding constants α and L (note that the value of L does not depend on δ). The following remark gives us a way of obtaining upper bounds for the asymptotic behaviour of the iterates $\Phi^n_{\lambda_{cr}}[F]$, comparing them to a rescaling of the distribution function of a stationary measure.

Corollaire 3.2. *Let $F \in \mathcal{F}$ be a function of class (α, δ), then for any n the function $T_{2L\delta}\Phi^n_{\lambda_{cr}}[F]$ belongs to $C_{\alpha,\delta/2^n}$.*

Proof. We have

$$\begin{aligned} T_{L\delta}\Phi_{\lambda_{cr}}[F] &\in C_{\alpha,\delta/2}, \\ T_{L\delta/2}\Phi_{\lambda_{cr}} T_{L\delta}\Phi_{\lambda_{cr}}[F] &\in C_{\alpha,\delta/4}, \\ &\text{etc.}, \\ T_{L\delta/2^n}\Phi_{\lambda_{cr}} \dots T_{L\delta}\Phi_{\lambda_{cr}}[F] &\in C_{\alpha,\delta/2^n}. \end{aligned}$$

Since the operator $\Phi_{\lambda_{cr}}$ commutes with the rescaling ones we have

$$T_{\frac{L\delta}{2^n}+\dots+\frac{L\delta}{2}+L\delta}\, \Phi^n_{\lambda_{cr}}[F] \in C_{\alpha,\delta/2^n}$$

and the result easily follows, for $T_{\frac{L\delta}{2^n}+\dots+\frac{L\delta}{2}+L\delta}\, \Phi^n_{\lambda_{cr}}[F] \geq T_{2L\delta}\, \Phi^n_{\lambda_{cr}}[F]$. □

Described in a different way, the previous corollary reads

Corollaire 3.3. *In the assumptions of Corollary 3.2, for any $x \in \mathbf{R}_+$ and $n \geq 1$ we have*

$$\Phi^n_{\lambda_{cr}}[F](x) \leq T_{-2L\delta}[F_{\bar{\mu}}](x) + \delta/2^n.$$

The rest of this section is devoted to the proof of Proposition 3.1.

The function $F_{\bar{\mu},\alpha,\delta}$. Any class $C_{\alpha,\delta}$ has a greatest element $F_{\bar{\mu},\alpha,\delta}$ (corresponding to the smallest probability measure with respect to the stochastic domination) defined by

$$F_{\bar{\mu},\alpha,\delta}(x) = \begin{cases} F_{\bar{\mu}}(x) + \delta & \text{on } [0, \kappa_{(\alpha-\delta)}) \cup [\kappa_{(1-\alpha)}, \kappa_{(1-\delta)}), \\ \alpha & \text{on } [\kappa_{(\alpha-\delta)}, \kappa_\alpha), \\ F_{\bar{\mu}}(x) & \text{on } [\kappa_\alpha, \kappa_{(1-\alpha)}), \\ 1 & \text{on } [\kappa_{(1-\delta)}, +\infty). \end{cases}$$

The useful remark is that a function $F \in \mathcal{F}$ verifies $F \leq F_{\bar{\mu},\alpha,\delta}$ if and only if it belongs to $C_{\alpha,\delta}$.

By monotonicity of the operator $\Phi_{\lambda_{cr}}$, it is sufficient to prove Proposition 3.1 for the function $F_{\bar{\mu},\alpha,\delta}$: we have to show the existence of a constant L such that the image $T_{L\delta}\Phi_{\lambda_{cr}}[F_{\bar{\mu},\alpha,\delta}]$ is of class $C_{\alpha,\delta/2}$. This requires an accurate description of how far the image $\Phi_{\lambda_{cr}}[F_{\bar{\mu},\alpha,\delta}]$ is from $F_{\bar{\mu}}$. We will first give a *global* upper bound and then study the neighbourhoods of 0 and $+\infty$.

Coupling $F_{\bar{\mu}}$ and $F_{\bar{\mu},\alpha,\delta}$. Yet another time, it will be more convenient to work with a good coupling between the laws of $F_{\bar{\mu}}$ and $F_{\bar{\mu},\alpha,\delta}$.

So for the rest of this section, we fix a probability space $(\Omega, \mathbf{P})$. Given a random variable X whose law is $\bar{\mu}$, we define a new random variable $X^{\alpha,\delta}$ by

$$X^{\alpha,\delta}(\omega) = \begin{cases} 0 & \text{if } X(\omega) \in [\kappa_{(\alpha-\delta)}, \kappa_\alpha), \\ \kappa_{(1-\alpha)} & \text{if } X(\omega) \in [\kappa_{(1-\delta)}, +\infty), \\ X(\omega) & \text{otherwise.} \end{cases} \tag{IV.28}$$

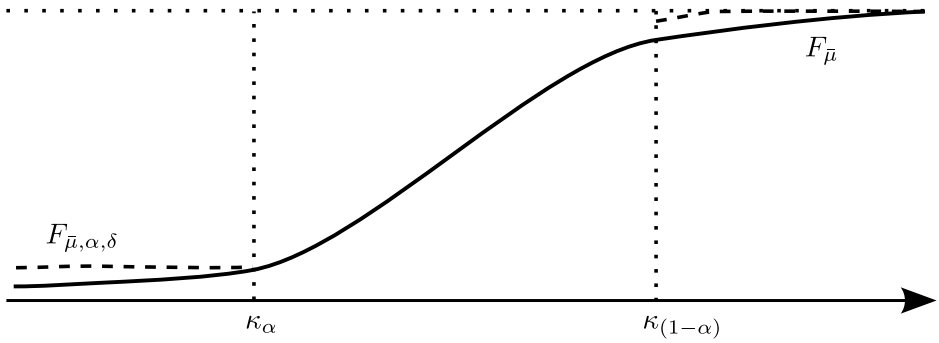

Figure IV.5 – The distribution functions $F_{\bar\mu}$ and $F_{\bar\mu,\alpha,\delta}$.

It is easy to remark that the distribution function of $X^{\alpha,\delta}$ is exactly the function $F_{\bar\mu,\alpha,\delta}$: the δ-jumps of $F_{\bar\mu,\alpha,\delta}$ at 0 and $\kappa_{(1-\alpha)}$ correspond to the first two lines in (IV.28). Moreover the event $\{X^{\alpha,\delta} \neq X\}$ is given by the first two possibilities in (IV.28) and its probability equals 2δ.

Let us apply the operator $\Phi_{\lambda_{cr}}$ to the distribution function $F_{\bar\mu,\alpha,\delta}$. To do so, draw four independent random variables $X_1, \dots, X_4$ distributed with respect to the law $\bar\mu$ and define four further random variables $X_1^{\alpha,\delta}, \dots, X_4^{\alpha,\delta}$ according to (IV.28). Take also a random variable ξ of law m and independent of the eight previous ones. We set

$$Y = R_{\lambda_{cr}}(X_1, \dots, X_4; \xi),$$
$$Y^{\alpha,\delta} = R_{\lambda_{cr}}(X_1^{\alpha,\delta}, \dots, X_4^{\alpha,\delta}; \xi).$$

Since the law $\bar\mu$ is stationary, the law of Y also coincides with $\bar\mu$. On the other hand the probability of $Y^{\alpha,\delta} \neq Y$ is upper bounded by the sum of the probabilities of four events $\{X_j \neq X_j^{\alpha,\delta}\}$ and does not exceed $4 \cdot 2\delta = 8\delta$.

Then for any $\alpha \in (0, 1/2)$ and $\delta \in (0, \alpha]$ we have the inequality

$$\Phi_{\lambda_{cr}}[F_{\bar\mu,\alpha,\delta}] \leq \min(F_{\bar\mu} + 8\delta, 1). \tag{IV.29}$$

However this bound is quite rough. For our purposes we need to be a little more careful: we know that the function $\Phi_{\lambda_{cr}}[F_{\bar\mu,\alpha,\delta}]$ takes the value $\theta(\delta) = O(\delta^2)$ at zero, which is pretty better than 4δ. Henceforth, we will give a bound on the values s_0 and t_0 for which the distance between $F_{\bar\mu}$ and $\Phi_{\lambda_{cr}}[F_{\bar\mu,\alpha,\delta}]$ is less than $\delta/2$ outside the interval $[s_0, t_0]$.

Lemma 3.4. *For every $\alpha \in (0, 1/40)$ there exist two points s_0, $t_0 \in \mathbf{R}_+$ such that the following estimates hold. For any $\delta \in (0, \alpha]$, $s \in [0, s_0)$ and $t \in (t_0, +\infty)$ we have*

$$\mathbf{P}(Y^{\alpha,\delta} \leq s < Y) \leq \delta/2, \tag{IV.30}$$
$$\mathbf{P}(Y^{\alpha,\delta} < t < Y) \leq \delta/2. \tag{IV.31}$$

Proof. We only prove the estimate (IV.30), since the estimate (IV.31) can be obtained in a similar way.

The condition $Y^{\alpha,\delta} < Y$ implies that at least one of the $X_i^{\alpha,\delta}$ is different from X_i. By symmetry, we have

$$\mathbf{P}(Y^{\alpha,\delta} \leq s < Y) \leq 4 \cdot \mathbf{P}(Y^{\alpha,\delta} \leq s < Y,\ X_4^{\alpha,\delta} < X_4).$$

As we have already discussed, the event $\{X_4^{\alpha,\delta} < X_4\}$ is covered by the two events

$$E_1 := \{X_4 \in [\kappa_{(\alpha-\delta)}, \kappa_\alpha)\} \text{ and } E_2 := \{X_4 \in [\kappa_{(1-\delta)}, +\infty), X_3 > \kappa_{(1-\alpha)}\},$$

whose probabilities are respectively $\mathbf{P}(E_1) = \delta$ and $\mathbf{P}(E_2) = \alpha\delta$.

Since the measures m and $\bar\mu$ have no atoms, we can choose s_0 such that

$$\mathbf{P}(\lambda_{cr}\xi \min(X_1, X_2) \leq s_0) \leq \tfrac{1}{8} - 5\alpha. \tag{IV.32}$$

We want to show that this choice for s_0 is fine enough.

When E_1 holds, we have

$$Y^{\alpha,\delta} = \lambda_{cr}\xi \min(X_1^{\alpha,\delta}, X_2^{\alpha,\delta}).$$

Using the same argument as before, we observe that the probability of the event

$$B = \left\{\lambda_{cr}\xi \min(X_1^{\alpha,\delta}, X_2^{\alpha,\delta}) \neq \lambda_{cr}\xi \min(X_1, X_2)\right\}$$

is at most 4δ. We can then obtain the estimate

$$\begin{aligned}\mathbf{P}(Y^{\alpha,\delta} \le s < Y \mid E_1) &= \mathbf{P}(Y^{\alpha,\delta} \le s < Y \mid E_1, B)\,\mathbf{P}(B) + \mathbf{P}(Y^{\alpha,\delta} \le s < Y \mid E_1, B^c)\,\mathbf{P}(B^c)\\ &\le 4\delta + \mathbf{P}(\lambda_{cr}\xi \min(X_1^{\alpha,\delta}, X_2^{\alpha,\delta}) \le s_0 \mid B^c)\\ &\le 4\delta + \mathbf{P}(\lambda_{cr}\xi \min(X_1, X_2) \le s_0)\end{aligned}$$

Thus, using s_0 such that (IV.32) holds, we have

$$\mathbf{P}(Y^{\alpha,\delta} \le s < Y \mid E_1) \le 4\delta + \tfrac{1}{8} - 5\alpha. \tag{IV.33}$$

We can now conclude:

$$\begin{aligned}\mathbf{P}(Y^{\alpha,\delta} \le s < Y) &\le 4 \cdot \mathbf{P}(Y^{\alpha,\delta} \le s < Y,\ X_4^{\alpha,\delta} < X_4)\\ &= 4 \cdot \mathbf{P}(Y^{\alpha,\delta} \le s < Y \mid E_1)\,\mathbf{P}(E_1) + 4 \cdot \mathbf{P}(Y^{\alpha,\delta} \le s < Y \mid E_2)\,\mathbf{P}(E_2)\\ &\le 4\delta \cdot \mathbf{P}(Y^{\alpha,\delta} \le s < Y \mid E_1) + 4\alpha\delta,\end{aligned}$$

and, according to (IV.33), this is not greater than

$$4\delta\left(4\delta + \tfrac{1}{8} - 4\alpha\right) < \tfrac{\delta}{2}$$

for every $\delta \in (0, \alpha]$, as desired. □

The previous lemma implies that for any $z \in [0, s_0] \cup [t_0, +\infty)$, the functions $\Phi_{\lambda_{cr}}[F_{\bar{\mu},\alpha,\delta}]$ and $F_{\bar{\mu}}$ differ at most by $\delta/2$ at the point z. The key fact is that the points s_0 and t_0 do not depend on δ. We use this result to end the proof of Proposition 3.1, which follows directly from the next easy lemma, when we set $\varepsilon = 8\delta$ and $\beta < \min(F_{\bar{\mu}}(s_0(\alpha)), 1 - F_{\bar{\mu}}(t_0(\alpha)), \alpha/16)$.

Lemma 3.5. *There exists a constant L_0 with the following property. Take $\delta < \varepsilon < \beta/2$ and assume that the distribution function F satisfies*

$$F(x) \le \begin{cases} F_{\bar{\mu}}(x) + \delta/2 & \text{on } [0, \kappa_\beta) \cup (\kappa_{(1-\beta)}, +\infty),\\ F_{\bar{\mu}}(x) + \varepsilon & \text{on } [\kappa_\beta, \kappa_{(1-\beta)}].\end{cases}$$

Then $T_{L_0\varepsilon}[F] \in C_{\beta,\delta/2}$.

Proof. We have already observed that the function $F_{\bar{\mu}}$ is C^1 regular with positive derivative on $(0, +\infty)$. Hence the function $g := \log F_{\bar{\mu}}^{-1}$ is well defined and locally Lipschitz on $(0, 1)$. Let L_0 be the Lipschitz constant of g on the interval $[\beta/2, 1 - \beta]$. We shall show that this choice is fine enough.

Let us remark that the hypotheses imply a further inequality for the function F, namely

$$F(x) \le 1 - \beta + \delta/2 \quad \text{on } [\kappa_{(1-\beta+\varepsilon+\delta/2)}, \kappa_{(1-\beta)}), \tag{IV.34}$$

for F is a non-decreasing function. Let us verify that $T_{L_0\varepsilon}[F]$ is in $C_{\beta,\delta/2}$.

- On the interval $[0, \kappa_\beta)$ the estimate is very plain:

  $$F(e^{-L_0\varepsilon}x) \le F_{\bar{\mu}}(e^{-L_0\varepsilon}x) + \delta/2 \le F_{\bar{\mu}}(x) + \delta/2,$$

  for $F_{\bar{\mu}}$ is non-decreasing.

- On the interval $[\kappa_{(1-\beta)}, e^{L_0\varepsilon}\kappa_{(1-\beta)})$ we use the inequality (IV.34) to find

  $$F(e^{-L_0\varepsilon}x) \le 1 - \beta + \delta/2 = F_{\bar{\mu}}(\kappa_{(1-\beta)}) + \delta/2 \le F_{\bar{\mu}}(x) + \delta/2,$$

  whereas on the interval $[e^{L_0\varepsilon}\kappa_{(1-\beta)}, +\infty)$ we can simply repeat the argument used at the previous point.

- On the interval $[\kappa_\beta, \kappa_{(1-\beta)})$ we need to verify the inequality

  $$F(e^{-L_0\varepsilon}x) \le F_{\bar{\mu}}(x). \tag{IV.35}$$

  Since $F(x) \le F_{\bar{\mu}}(x) + \varepsilon$, the inequality (IV.35) reads as

  $$e^{-L_0\varepsilon}x \le F_{\bar{\mu}}^{-1}(F_{\bar{\mu}}(x) - \varepsilon)$$

  or

  $$-L_0\varepsilon + g(F_{\bar{\mu}}(x)) \le g(F_{\bar{\mu}}(x) - \varepsilon).$$

  This means that for every $\xi \in [\beta - \varepsilon, 1 - \beta] \subset [\beta/2, 1 - \beta]$ we want

  $$g(\xi) - g(\xi - \varepsilon) \le L_0\varepsilon,$$

  which clearly holds as a consequence of our choice for L_0.

□

3.2 Asymptotically upper bounds and convergence

The result in Corollary 3.3 leads naturally to the following

Definition 3.4. *The rescaling* $T_r[F_{\bar{\mu}}]$ asymptotically upper bounds F *if for every* $\varepsilon > 0$ *it is possible to find* $n_0 \in \mathbf{N}$ *such that for any* $n \geq n_0$ *and* $x \in \mathbf{R}_+$ *we have*

$$T_r[F_{\bar{\mu}}](x) + \varepsilon \geq \Phi^n_{\lambda_{cr}}[F](x).$$

In the same way, the rescaling $T_r[F_{\bar{\mu}}]$ asymptotically lower bounds F *if for every* $\varepsilon > 0$ *it is possible to find* $n_0 \in \mathbf{N}$ *such that for any* $n \geq n_0$ *and* $x \in \mathbf{R}_+$ *we have*

$$T_r[F_{\bar{\mu}}](x) - \varepsilon \leq \Phi^n_{\lambda_{cr}}[F](x).$$

Definition 3.5. *Given any function* $F \in \mathcal{F}_0$ *we define the sets*

$$R_+ = R_+(F) = \{r \in \mathbf{R} \mid T_r[F_{\bar{\mu}}] \text{ asymptotically upper bounds } F\},$$
$$R_- = R_-(F) = \{r \in \mathbf{R} \mid T_r[F_{\bar{\mu}}] \text{ asymptotically lower bounds } F\}.$$

We have already observed in Remark 3.1 that for any $F \in \mathcal{F}_0$ there exists a rescaling of F that belongs to the class (α, α). So Corollary 3.3 implies that $R_+(F)$ is non-empty. We can repeat the arguments in Section 3.1 with the class of functions $F \in \mathcal{F}$ verifying the reversed conditions:

1. $F(x) \geq F_{\bar{\mu}}(x)$ on $[\kappa_\alpha, \kappa_{(1-\alpha)}]$,
2. $F(x) \geq F_{\bar{\mu}}(x) - \delta$ on $[0, \kappa_\alpha) \cup (\kappa_{(1-\alpha)}, +\infty]$.

Then we see that the set $R_-(F)$ is non-empty as well. Finally, for any $r_+ \in R_+$ and $r_- \in R_-$ it is easy to see that $r_+ \leq r_-$, hence the set R_+ is non-empty and right bounded.

This allows to consider the value

$$r_0 = r_0(F) := \sup R_+(F).$$

The proof of Theorem 1.2 will be concluded once we prove the following

Proposition 3.6. *For any* $F \in \mathcal{F}_0$ *the iterates* $\Phi^n_{\lambda_{cr}}[F]$ *converge to* $T_{r_0(F)}[F_{\bar{\mu}}]$ *as* n *tends to* ∞.

Proof. The upper bound is easy. Indeed for any $r < r_0$, the definition of r_0 itself gives

$$\limsup_{n \longrightarrow \infty} \Phi^n_{\lambda_{cr}}[F](x) \leq T_r[F_{\bar{\mu}}](x),$$

and passing to the limit as $r \uparrow r_0$, we obtain the desired

$$\limsup_{n \longrightarrow \infty} \Phi^n_{\lambda_{cr}}[F](x) \leq T_{r_0}[F_{\bar{\mu}}](x). \tag{IV.36}$$

To prove the equality we will need the following

Lemma 3.7. *For any* $\varepsilon > 0$ *there exist* $\delta, r > 0$ *such that if* $F \in \mathcal{F}$ *is of class* (α, δ) *and for some* $x \in \mathbf{R}_+$ *one has* $F(x) < F_{\bar{\mu}}(x) - \varepsilon$, *then* $T_r[F_{\bar{\mu}}]$ *asymptotically upper bounds* F.

Indeed, suppose there is a *strict* inequality in (IV.36), then we can find $x \in \mathbf{R}_+$ and $\varepsilon > 0$ such that

$$\limsup_{n \longrightarrow \infty} \Phi^n_{\lambda_{cr}}[T_{-r_0}[F]](x) < F_{\bar{\mu}}(x) - \varepsilon.$$

Then the Lemma gives a contradiction: there is some $r_0 + r$, with $r > 0$ which belongs to $R_+(F)$.

Proof of Lemma 3.7. Let us first reduce (at the cost of replacing ε by $\varepsilon/2$) the possible set of values of x that we have to consider from $(0, \infty)$ to some interval bounded away from 0 and from ∞. Indeed if $x > \kappa_{(1-\varepsilon/2)}$ we have

$$F(\kappa_{1-\varepsilon/2}) < F(x) < F_{\bar{\mu}}(x) - \varepsilon < 1 - \varepsilon = (1 - \varepsilon/2) - \varepsilon/2 = F_{\bar{\mu}}(\kappa_{1-\varepsilon/2}) - \varepsilon/2,$$

so the assumptions of the lemma are also satisfied for $\varepsilon' := \varepsilon/2$ and $x' := \kappa_{(1-\varepsilon')}$.

On the other hand, $F(x) < F_{\bar{\mu}}(x) - \varepsilon$ implies that $F_{\bar{\mu}}(x) \geq \varepsilon$ and hence $x \geq \kappa_\varepsilon$. It is therefore sufficient to prove the lemma under the additional assumption that x stays in the interval $[\kappa_\varepsilon, \kappa_{(1-\varepsilon)}]$.

Before pursuing the proof in full generality, we establish a weaker statement. Namely, let us assume that the map F satisfies $F(x) \leq F_{\bar{\mu}}(x)$ for all x (in other words, that $\delta = 0$) and let us show the existence of a positive $r = r(F)$ for which $T_r[F_{\bar{\mu}}]$ asymptotically upper bounds F. This statement is much weaker than the desired one,

not only because it corresponds to $\delta = 0$, but also because we establish an existence of $r > 0$ for each F instead of finding a uniform r for all functions F satisfying our assumptions; nevertheless, it clarifies the arguments that we shall use later (and would be otherwise too technical).

As $\Phi_{\lambda_{cr}}[F_{\bar{\mu}}] = F_{\bar{\mu}}$ and the measure m is of full support on $\mathbf{R}_+$, after applying $\Phi_{\lambda_{cr}}$ once we get (from the convolution operation) a strict inequality

$$\Phi_{\lambda_{cr}}[F](x) < F_{\bar{\mu}}(x) \quad \text{for any } x \in \mathbf{R}_+ . \tag{IV.37}$$

Due to the strict inequality (IV.37), it is possible to find, for any compact interval $J \subset \mathbf{R}_+$ (and in particular for $[\kappa_\varepsilon, \kappa_{(1-\varepsilon)}]$), a sufficiently small $r' > 0$ such that

$$T_{-r'}[\Phi_{\lambda_{cr}}[F]](x) < F_{\bar{\mu}}(x) \quad \text{for every } x \in J.$$

Hence for such an r' the function $T_{-r'}[\Phi_{\lambda_{cr}}[F]]$ belongs to some class (α, δ_1), with $\delta_1 = \delta_1(r') < r'$ and α depending on J (for example $\alpha < \varepsilon$ if $J = [\kappa_\varepsilon, \kappa_{(1-\varepsilon)}]$).

Using Corollary 3.3, this implies that $T_{-r'}[F]$ is asymptotically upper bounded by $T_{-2L\delta_1}[F_{\bar{\mu}}]$ and hence F is asymptotically upper bounded by $T_{r'-2L\delta_1}[F_{\bar{\mu}}]$.

The final part of the argument, that will conclude the consideration of this particular case, is that it is possible to make a choice for δ_1 such that

$$\delta_1(r') = o(r') \quad \text{as } r' \longrightarrow 0. \tag{IV.38}$$

Indeed, once this estimate is established, we deduce that for a sufficiently small r' the value $r := r' - 2L\delta_1(r')$ is strictly positive, which is exactly what we required.

To establish (IV.38), let us pass to the logarithmic scale, considering the coordinate $y = \log x$. After this change of variable, we have that for the function $\varphi(y) := F_{\bar{\mu}}(e^y)$, the derivative tends to 0 on $\pm\infty$. Indeed, $\bar{\mu}$ is equal to the multiplicative convolution of m with the law of $\lambda_{cr}\,(\min(X_1, X_2) + \min(X_3, X_4))$, where the X_j's are i.i.d. with law $\bar{\mu}$. Hence, the distribution function of $\log_* \bar{\mu}$ (*i.e.* the measure $\bar{\mu}$ pushed on $\mathbf{R}$ *via* the function log) on is a convolution of the distribution function of the measure $\log_* m$ with the measure $\nu = \text{law}(\log \lambda_{cr}\,(\min(X_1, X_2) + \min(X_3, X_4)))$. By the assumptions on m, the distribution function of the measure $\log_* m$ is a function of class C^1 with derivative that tends to zero at $\pm\infty$, hence the same holds for its convolution with an arbitrary measure ν.

From the above arguments we can find that for any $s > 0$ there exist $y_1, y_2, r'_0 \in \mathbf{R}$ such that for any $r' < r'_0$ and any $y \notin [y_1, y_2]$ the shifted function $\varphi(y + r')$ is bounded from above:

$$\varphi(y + r') < \varphi(y) + sr'.$$

Coming back to the initial coordinates we get that outside the interval $J := [e^{y_1}, e^{y_2}]$ one has

$$T_{-r'}[\Phi_{\lambda_{cr}}[F]] < T_{-r'}[F_{\bar{\mu}}] < F_{\bar{\mu}} + sr'$$

for all sufficiently small $r' > 0$. Since for the interval J, for all sufficiently small $r' > 0$ one has the strict inequality $T_{-r'}[\Phi_{\lambda_{cr}}[F]]|_J < F_{\bar{\mu}}|_J$, we eventually conclude that $T_{-r'}[\Phi_{\lambda_{cr}}[F]] \in C_{sr',\alpha}$ for some α. As $s > 0$ was arbitrary, we obtain the desired $\delta_1(r') = o(r')$ as $r' \longrightarrow 0$.

Let us now modify the above argument to return to the full general assumptions of the lemma. We set $s := 1/6L$; due to the above arguments, there exist y_1, y_2, r'_1 such that for any $y \notin [y_1, y_2]$ and any $r' \in (0, r'_1)$ we have $\varphi(y + r') < \varphi(y) + sr'$. Furthermore, without loss of generality we can assume $y_1 < \log \kappa_\alpha$ and $y_2 > \log \kappa_{(1-\alpha)}$.

Observe that if for a function $F \in C_{\alpha,\delta}$ it is possible to find a point $x_0 \in [\kappa_\varepsilon, \kappa_{(1-\varepsilon)}]$ such that $F(x_0) < F_{\bar{\mu}}(x_0) - \varepsilon$, then F is bounded from above by the function

$$F_{\alpha,\delta,\varepsilon,x_0,\bar{\mu}}(x) := \min(F_{\bar{\mu},\alpha,\delta}, (F_{\bar{\mu}}(x_0) - \varepsilon) \cdot \mathbf{1}_{x \le x_0} + \mathbf{1}_{x > x_0}).$$

Note that this family of functions defines a family of measures on $[0, +\infty]$, depending continuously on δ and x_0. Hence the family of functions $\Phi_{\lambda_{cr}}[F_{\alpha,\delta,\varepsilon,x_0,\bar{\mu}}]$ depends C^0-continuously on δ and x_0 (due to the multiplicative convolution with m). Then the function

$$H(x, x_0, \delta, r') = T_{-r'}[\Phi_{\lambda_{cr}}[F_{\alpha,\delta,\varepsilon,x_0,\bar{\mu}}]](x) - F_{\bar{\mu}}(x)$$

considered on $[e^{y_1}, e^{y_2}] \times [\kappa_\varepsilon, \kappa_{1-\varepsilon}] \times [0, \alpha] \times [0, 1]$ is also continuous. It is strictly negative when δ and r' are 0 and so, due to the compactness of $[e^{y_1}, e^{y_2}] \times [\kappa_\varepsilon, \kappa_{1-\varepsilon}]$, the same holds for sufficiently small δ and r'. Thus there exist $\delta_2 > 0$ and $r'_2 > 0$ such that for every $\delta < \delta_2$ and $r' < r'_2$ one has

$$H(x, x_0, \delta, r') < 0.$$

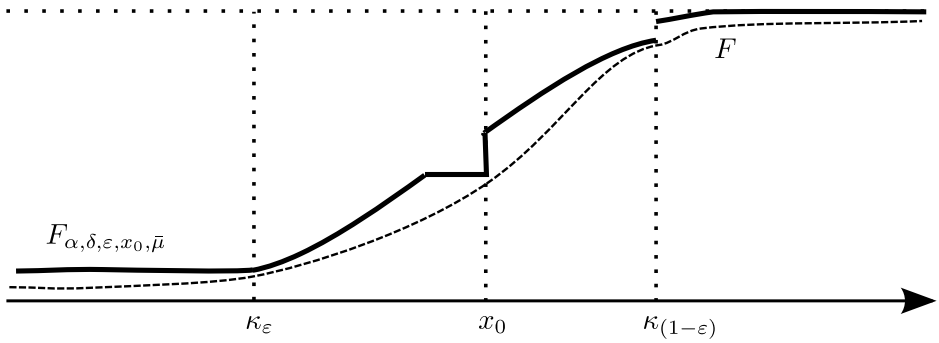

Figure IV.6 – The distribution functions F and $F_{\alpha,\delta,\varepsilon,x_0,\bar{\mu}}$.

For any such δ and r', using the inequality between $F_{\alpha,\delta,\varepsilon,x_0,\bar{\mu}}$ and F (look at Figure IV.6), we have

$$T_{-r'}[\Phi_{\lambda_{cr}}[F]](x) < F_{\bar{\mu}}(x) \quad \text{for any } x \in [e^{y_1}, e^{y_2}]. \tag{IV.39}$$

On the other hand F is of class (α, δ) and so the inequality (IV.29) holds:

$$\Phi_{\lambda_{cr}}[F] \leq F_{\bar{\mu}} + 8\delta. \tag{IV.40}$$

Take any $r' < \min(r'_1, r'_2)$ and $\delta < \min(\delta_2, r'/48L)$. Then for such r' and δ and any function F of class $C_{\alpha,\delta}$ we have the inequality

$$T_{-r'}[\Phi_{\lambda_{cr}}[F]] \leq F_{\bar{\mu}} \quad \text{on } [e^{y_1}, e^{y_2}] \supset [\kappa_\alpha, \kappa_{(1-\alpha)}]$$

due to (IV.39), while on the complement $[0, e^{y_1}] \cup [e^{y_2}, \infty]$,

$$T_{-r'}[\Phi_{\lambda_{cr}}[F]] \leq T_{-r'}[F_{\bar{\mu}} + 4\delta] \leq F_{\bar{\mu}} + \frac{r'}{6L} + 8\,\frac{1}{48L}r' \leq F_{\bar{\mu}} + \frac{r'}{3L},$$

giving that $T_{-r'}[\Phi_{\lambda_{cr}}[F]]$ is a function of class $(r'/3L, \alpha)$. Using Corollary 3.3, we obtain that $T_{-r'}[F]$ is asymptotically upper bounded by $T_{-2r'/3}[F_{\bar{\mu}}]$. Eventually we apply $T_{r'}$: then F is asymptotically upper bounded by $T_{r'}[T_{-2r'/3}[F_{\bar{\mu}}]]$ and this proves the lemma with $r = r'/3$. □

This ends the proof of Theorem 1.2. □

4 Random metric on the limit object

For any measure m, satisfying the assumptions of Theorem 1.1, we have constructed a scaling constant λ_{cr} and a stationary measure $\bar{\mu}$ satisfying $\bar{\mu} = \Phi_{\lambda_{cr}}[\bar{\mu}]$. We can then consider the corresponding random tree process: a tree of random variables $(X_\cdot, \xi_\cdot) : \mathbf{S}^1 \longrightarrow \mathbf{R}_+$, such that for every $t \in \mathbf{S}^1$, the law of ξ_t is m, that of X_t is $\bar{\mu}$ and

$$X_t = R_{\lambda_{cr}}(X_{t_1}, \ldots, X_{t_4}; \xi_t),$$

with the property that for any n the random variables $\{\xi_t\}_{\|t\| \leq n}$ and $\{X_t\}_{\|t\|=n}$ are independent altogether.

Let $p : \mathbf{S}^1 \setminus \{\bullet\} \longrightarrow \mathbf{S}^1$ be the map that associates to each non-root vertex of the tree $\mathbf{S}^1$ its parent, and let $\|\cdot\| : \mathbf{S}^1 \longrightarrow \mathbf{Z}_+$ be the distance to the root. Then, we can put in correspondence to the family of random variables $\{(X_t, \xi_t)\}_{t \in \mathbf{S}^1}$ a family of random variables $\{Y_t\}_{t \in \mathbf{S}^1}$, defined as

$$Y_t = X_t \cdot \prod_{j=1}^{\|t\|} (\lambda_{cr} \xi_{p^j(t)}). \tag{IV.41}$$

These random variables satisfy $Y_t = R(Y_{t_1}, \ldots, Y_{t_4}; 1)$, and we can interpret them as "*IO*-distances *inside* the copy of Γ corresponding to t".

Denote by V_n the set of vertices of the graph Γ_n. There is a natural inclusion $V_n \hookrightarrow V_{n+1}$, allowing to define $V_\infty = \bigcup_n V_n$, that is naturally identifiable with a subset of Γ_∞. Borrowing the concept from the interval, one can consider V_∞ as the set of "dyadic rational" points of Γ_∞.

Now, the family of random variables $\{Y_t\}_{t \in \mathbf{S}^1}$ allows us to equip each V_n with a random distance function $d_n : V_n \times V_n \longrightarrow \mathbf{R}_+$. Indeed, the edges of Γ_n are in a natural correspondence with $\{t \in \mathbf{S}^1 : \|t\| = n\}$. Set for any $t \in \mathbf{S}^1$ with $\|t\| = n$ the length of the edge, corresponding to t, to be equal to Y_t, and consider the metric

thus induced on Γ_n. Restricting it on V_n, we obtain the desired random distance function d_n. Moreover, the relation (IV.41) implies that these distances agree with each other: $d_{n'}|_{V_n \times V_n} = d_n$ for any $n' > n$.

This allows us to define a random distance function $d_\infty : V_\infty \times V_\infty \longrightarrow \mathbf{R}_+$. The next natural idea is to try to extend this metric to all the graph Γ_∞, by completing the space (V_∞, d_∞) and asking whether the inclusion $V_\infty \subset \Gamma_\infty$ extends to the isomorphism between the completion and Γ_∞; if (almost surely) yes, this would define a random metric on Γ_∞.

Let us consider the case of log-normal factor ξ: let $\log \xi \sim \mathcal{N}(0, \sigma^2)$. It seems that, in the same way as for the construction of a random measure [Rho-Var 2], there are two different regimes, "subcritical" (with $\sigma < \sigma_c$) and "supercritical" (with $\sigma > \sigma_c$), though the supercritical regime is different from the one for the case of the measure construction. We present here a (non-rigorous!) description of these regimes. Namely, consider the branching random walk corresponding to the logarithm of the product in (IV.41):

$$S_t = \sum_{j=1}^{\|t\|} (\log \lambda_{cr} + \log \xi_{p^j(t)}).$$

The standard results of the theory of branching random walks imply that the the maximum $L_n = \max_{\|t\|=n} S_t$ has an almost-surely linear behaviour:

$$\exists \gamma_{BRW} : \quad \frac{L_n}{n} \longrightarrow \gamma_{BRW} \quad \text{as } n \longrightarrow \infty. \tag{IV.42}$$

The two regimes correspond respectively to the case $\gamma_{BRW} < 0$ and $\gamma_{BRW} > 0$.

In the case $\gamma_{BRW} < 0$, the diameters of small "level n" copies of Γ_∞ (that replace edges of Γ_n) exponentially tend to zero. This is not an immediate conclusion: an immediate conclusion is that the factor $\prod_{j=1}^{n}(\lambda_{cr}\xi_{p^j(t)})$ in (IV.41) tends to zero uniformly in t with $\|t\| = n$ as n tends to infinity. And one has to take into account the other factor, X_t. Though, it is not so difficult to show that for a log-normal law of a factor ξ, the corresponding stationary measure $\bar{\mu}$ has a law with log-normally bounded tails. And such a bound implies that the interference of X_t is neglectable compared to the exponential factors.

We see such a regime for small values of σ. Indeed, for $\sigma = 0$ we have $\lambda_{cr} = 1/2$ (even if Theorem 1.1 is, formally speaking, not applicable), and thus $\gamma_{BRW} = -\log 2$. It is quite natural to expect that λ_{cr} depends continuously on m, and hence we will have $\gamma_{BRW}(\sigma) \longrightarrow -\log 2$ as $\sigma \longrightarrow 0$.

The case $\gamma_{BRW} > 0$ is more tricky. In this case, the diameter of V_∞ is almost surely infinite! However, it seems that the completion (V_∞, d_∞) is still a connected space — in contrast to the interval case and what we will explain later. Very roughly speaking, there is an *IO* geodesic; there are four "local geodesics" joining vertices from V_1 inside what corresponds to the four edges of Γ_1 (two of them are halves of the full *IO* geodesic); there are sixteen "local geodesics" joining vertices from V_2 inside what corresponds to the sixteen edges of Γ_2 (eight of those are halves of the four local geodesics seen on the previous step); etc. And the *maximal* length of the local geodesic corresponding to the edges of Γ_n increases (exponentially) with n.

Such a behaviour is also manifested in the following example, that is once again a bit outside of our theoretical arguments (note that not only the factor ξ is not log-normal here, but also that once again, formally speaking, Theorem 1.1 is not applicable to this case), but is very easy to describe. Take

$$\xi_t = \begin{cases} 10000 & \text{with probability } \varepsilon, \\ 1 & \text{with probability } 1 - \varepsilon. \end{cases}$$

Taking ε slightly smaller than $1 - p_{cr} = \frac{\sqrt{5}-1}{2}$, where p_{cr} is the critical probability for the bond percolation for the hierarchical eight graph, one sees that almost surely there will be an *IO* path, for which all the random factors, corresponding to it (on all the levels), except for a finite number, are equal to 1. This implies, that taking $\lambda_{cr} = 1/2$ (as there are two edges of Γ_{n+1} per an edge of Γ_n), we see an almost-surely finite and positive *IO* limit distance. However, taking *any* point of Γ_∞ and considering the corresponding infinite branch of the random tree process, we see that the corresponding product of $\lambda_{cr}\xi_t$'s up to the level n behaves as $(5000^\varepsilon \cdot (1/2)^{1-\varepsilon})^n$, and in particular, if ε is not too small, such a product tends to infinity exponentially fast.

Let us now see why even in the supercritical regime for log-normal measure m in the completion $\overline{(V_\infty, d_\infty)}$ there should be a *IO* geodesic. As we have already mentioned, this is in contrast with the behavior of random measures on the interval: the latter ones are almost surely atomic in the supercritical regime, hence the completed space becomes disconnected (at the points that correspond to the atoms of such a measure).

A first observation is that for a log-normal law m, the stationary law $\bar{\mu}$ of the *IO* distance should have log-normally bounded tails. It comes from the fact that, very roughly speaking, the transformation $\Phi_{\lambda_{cr}}$ first almost-squares the density at $+\infty$, and then takes the convolution with the law of $\log \xi$. Note that there is a

normal-like density that is a fixed point of an approximating transformation, acting by squaring the density everywhere and then by taking a convolution:

$$\left(\exp\left(-\frac{x^2}{4\sigma^2}\right)\right)^2 * \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{x^2}{2\sigma^2}\right) = \exp\left(-\frac{x^2}{2\sigma^2}\right) * \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{x^2}{2\sigma^2}\right) = \exp\left(-\frac{x^2}{4\sigma^2}\right). \tag{IV.43}$$

Replacing this exact estimate with an asymptotic one transforms this argument to a log-normal estimate near infinity for the law $\bar{\mu}$.

Now, the random variables $(X_t, \xi_t)_{t\in\mathbf{S}^1}$ can be determined "top to bottom": first picking at random the value $X_\bullet$ corresponding to the root (which gives us exactly the IO distance), and then consecutively choosing values $(X_{t_1}, \dots, X_{t_4}; \xi_t)$ conditionally to the already known value of X_t, starting with $t = \bullet$ and descending further and further on the tree. In particular, we can consider the random tree process of descending only along the branches that correspond to the geodesic: in this case, given X_t, we choose values $(X_{t_{(1)}}, X_{t_{(2)}}; \xi_t)$, where $t_{(1)}$ and $t_{(2)}$ are the descendants of t along which the shortest path corresponding to t passes.

In such a descent, one meets the following two effects: first, if the value of X_t is very big, the values of $X_{t_{(1)}}$ and $X_{t_{(2)}}$ are highly probably much smaller. Indeed, the tail of $\log R(X_{t_1}, \dots, X_{t_4}; 1)$ decreases quite fast (at least as fast as a normal law), and conditionally to that the sum $\log R(X_{t_1}, \dots, X_{t_4}; 1) + \log(\lambda_{cr}\xi)$ takes some large value $\log X_t$, it is most probable that the second summand is also quite big. Thus, among the values of X_t corresponding to the *binary* tree of decomposition of the shortest IO path, we almost surely should meet infinite number of small values on every infinite branch.

Now, if X_t takes a small value (less than a given fixed constant), with some bounded from below probability the parts $\lambda_{cr}\xi_t X_{t_{(1)}}$ and $\lambda_{cr}\xi_t X_{t_{(2)}}$ of X_t are both no less than $\frac{1}{3}X_t$, and hence no greater than $\frac{2}{3}X_t$. This implies that in the same proportion, between 1/3 and 2/3, will be subdivided the corresponding part of the IO length, when we pass from its partition by vertices of $V_{\|t\|}$ to the partition by vertices of $V_{\|t\|+1}$.

Adding these two effects together, we see that for the partition of the total IO length into 2^n parts by the vertices of V_n, along each branch of the infinite descent we have infinitely many chances (with bounded from below probability) to subdivide it in a proportion between 1/3 and 2/3. Hence, almost surely the diameter of such a partition tends to zero. Thus, in the completed space there is an IO geodesic.

Surely, all the above arguments are quite far from being rigorous. However, they seems sufficiently convincing to us at least to conjecture, that this is the correct description of a scenario.

5 Hierarchical graphs

Hierarchical graphs have been studied so long due to their nice self-similar structure that it is impossible for us to give an exhaustive list of references on them (see [Ham-Kum] as a recent example, historical references can be found therein).

It should be clear that what we have done for the figure eight-graph, naturally generalizes to other hierarchical graphs built with different "bricks".

5.1 Notations

We consider a finite graph $\Gamma = (V, E)$ with two marked distinct points which we denote by I (*in*) and O (*out*).

Let us fix some further notation in the case that the graph Γ is planar. If we fix an embedding of Γ, we can also define the dual graph (Γ^*, I^*, O^*): the graph Γ^* is the dual graph of Γ, but we must be a little cautious defining I^* and O^*.

Let e_{IO} be an additional edge between I and O, in such a way that in the chosen embedding it stays outside the graph: just to fix notations we make this edge pass through the point at infinity. Then there are well defined left and right hand sides of the graph. We set them to correspond respectively to I^* and O^*.

Definition 5.1 (Hierarchical graphs). *We fix an orientation on E. We set (Γ_0, I, O) to be the graph (Γ, I, O) and we define graphs (Γ_i, I, O) by induction.*

Let i be a positive integer. We define $\Gamma_i = (V_i, E_i)$ to be the graph obtained recursively from the graph $\Gamma_{i-1} = (V_{i-1}, E_{i-1})$ where we replace each oriented edge $e = (x, y)$ with a copy of Γ in such a way that x coincides with I and y with O.

This defines a natural way of considering E_{i-1} as a subset of E_i. The generation *of a vertex x is the minimum i such that $x \in E_i$.*

The in *and* out *vertices of Γ_i are the vertices I and O of the 0th generation respectively.*

Remark 5.1. In the planar case, the construction of the hierarchical graph preserves the duality between graphs: $(\Gamma_i)^* = (\Gamma^*)_i$.

When the graph distance $d_\Gamma(I,O) > 1$, there is a nontrivial ("Euclidean") geometry on Γ_∞ which is the limit of the $(\Gamma_i, \tilde{d}_{\Gamma_i})$'s in the Gromov-Hausdorff topology, where $\tilde{d}_{\Gamma_i}$ is the distance d_{Γ_i} rescaled by a factor $1/d_{\Gamma_i}(I,O)$, in such a way that $\tilde{d}_{\Gamma_i}(I,O) = 1$. We want to show that even for *random distances*, we are able to define such a limit in law, in the same way that we managed for the figure eight-graph.

5.2 (min, +)-type recursive distributional equations

Given any finite graph (Γ, I, O) as in the previous section, we can associate to it a function ρ_Γ of $\#E$ variables $x_1, \dots, x_{\#E}$ involving only min and + operations in the way that we explain hereafter.

Definition 5.2. *A path π in Γ from I to O is a* simple path *if it crosses every edge only once.*

Let us fix a labelling $\{1, \dots, \#E\}$ for the edges, then to any simple path π we associate the linear function $\ell(\pi)(x_1, \dots, x_{\#E}) = \sum_{i=1}^{\#E} \mathbf{1}_{i\in\pi}\, x_i$.

Then we can define the function ρ_Γ as $\rho_\Gamma = \min_{\pi \text{ simple path}} \ell(\pi)$.

Sometimes the function ρ_Γ can have a nicer expression, for instance the reader may think about the function ρ given by the figure eight-graph: the definition of ρ simplifies to $\rho(x_1, x_2, x_3, x_4) = \min(x_1, x_2) + \min(x_3, x_4)$.

Given any absolutely continuous probability measure m on $(0, +\infty)$, we can define the *recursive distributional equation*

$$Y = \xi \cdot \rho_\Gamma(X_i), \tag{IV.44}$$

where ξ is a random variable of law m and Y and the X_i are i.i.d. random variables whose law $\mathcal{L}$ is the unknown of the problem.

Yet another time, we shall suggest the reader to compare (IV.44) with equations (IV.1)-(IV.2).

Remark 5.2. Conversely, it is possible to associate a "brick" graph to any (min, +)-type RDE of a certain form: if

$$Y = \xi \cdot \rho(X_i)$$

is the RDE we want to study, then a sufficient condition is that the function ρ makes the variables X_i appear only once.

5.3 Bond percolation on hierarchical graphs and stationary random distances

In the proof of Theorem 1.1, the function θ in Definition 2.2 played a crucial role. In particular, we have already remarked that a key feature is the nature of its fixed points: 0 and 1 are super-attracting and there is only another one in $(0,1)$. Indeed, this is the only property of θ that we really needed. Though, such a behaviour is not a special property of the figure eight-graph, but it is peculiar for graphs that do not possess *pivotal edges*, *i.e.* neither an IO edge, nor an edge whose removal would disconnect I from O. Indeed, under these two assumptions 0 and 1 are also super-attracting (this is rather easy to see) and a theorem of Moore and Shannon [Moo-Sha] says that after the change of variable $s = \log \frac{p}{1-p} \in \mathbf{R}$ the map θ becomes expanding and hence possesses a unique repelling fixed point.

Definition 5.3. *Given any $p \in [0,1]$, we consider the Bernoulli percolation of parameter p on Γ. This defines the function θ of variable p*

$$\theta(p) = \mathbf{P}(I \longrightarrow O), \tag{IV.45}$$

where $I \longrightarrow O$ denotes the event that I is connected to O.

It is easy to observe that θ defines a monotone continuous function, fixing 0 and 1: θ is a homeomorphism of the closed interval $[0,1]$ (defined by a polynomial expression in p).

Proposition 5.1. *Let (Γ, I, O) be any finite graph with two distinct marked points and no pivotal edges. Then* 0 *and* 1 *are super-attracting points for the dynamics induced by the map θ associated to Γ* (i.e. *the first derivative of θ at* 0 *is* 0*). This infers that θ has at least* three *fixed points on* $[0,1]$.

Proof. The weight of every subgraph $\Delta \in \{0,1\}^E$ is $p^{\#_1\Delta}(1-p)^{\#_0\Delta}$. A path connecting I to O must pass at least through $d_\Gamma(I,O)$ edges: if $\Delta \in I \longrightarrow O$ then $\#_1\Delta \geq d_\Gamma(I,O)$. Furthermore, we must also have $\#_0\Delta \geq 2$, for there is no pivotal edge. □

Example 5.1. The smaller nontrivial graph we can consider in this setting is the figure eight-graph or its dual, the diamond-graph.

Théorème (Moore – Shannon, [Moo-Sha]). *With the previous notations, if there are no pivotal edges, the function θ has exactly three fixed points.*

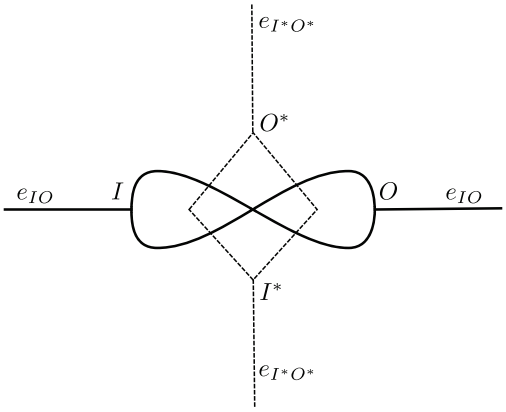

Figure IV.7 – The figure eight-graph and its dual graph, the diamond.

This directly implies the following extended version of Theorem 1.1:

Theorem 5.2. *Given any finite graph (Γ, I, O) as in the previous section, then for any non-atomic, fully supported probability distribution m on $\mathbf{R}_+$ there exists a normalization constant $\lambda_{cr} \in \mathbf{R}_+$ and a non-atomic probability distribution $\bar{\mu}$ on $\mathbf{R}_+$ such that $\bar{\mu}$ is a solution of the modified RDE (IV.44):*

$$Y = \lambda_{cr}\, \xi \cdot \rho_\Gamma(X_i).$$

Moreover, if (Γ, I, O) does not have pivotal edges, than $\bar{\mu}$ is unique up to rescaling.

Indeed, if the graph (Γ, I, O) does not have pivotal edges, everything works. A little more care is needed when *there are* pivotal edges: in such cases, either 0 or 1 are super-attracting (we suppose that there are at least two different edges in Γ), and we proceed with an *upper* or *lower cut-off* respectively in our arguments. Very different is the analysis of the convergence to the stationary measure. We describe the situation with illustrative examples.

Example 5.2. When (Γ, I, O) is the interval, the associated RDE describes a Galton-Watson process

$$R_\lambda(X_1, X_2; \xi) = \lambda \xi\, (X_1 + X_2). \tag{IV.46}$$

With some reasonable moment condition on ξ, it is classical that there is a unique (up to rescaling) stationary probability measure and the convergence is exponential (the contraction method applies [Ald-Ban]).

When (Γ, I, O) consists of two parallel edges, the RDE is associated to the extrema of a deterministic BRW:

$$R_\lambda(X_1, X_2; \xi) = \lambda \xi\, \min(X_1, X_2).$$

With some reasonable moment condition on ξ, it is classical that there is a unique (up to rescaling) stationary probability measure, but the convergence is not exponential (this is a nice result obtained with large deviations arguments, see for instance [Add-Ree]).

We expect that relevant difference between these two examples is the presence of the "+" operation in the first RDE, which intertwines the core part with the tails of the distribution. We conjecture that the convergence to the stationary measure is exponential when $d_\Gamma(I, O) \geq 2$ (and there is some reasonable moment condition on ξ).

Remark 5.3. It would be very interesting to study hierarchical graphs built out of *more than one kind of "brick"*, say $\Gamma^1, \Gamma^2, \ldots, \Gamma^d$. For example, at every step in the construction, one can choose randomly to insert one graph among the Γ^i's.

Remark 5.4. There is no apparent relation between solutions of the RDEs defined for a planar graph (Γ, I, O) and (Γ^*, I^*, O^*).

5.4 Appendix: The theorem of Moore and Shannon

We end this section giving the proof of Theorem 5.3, written in a modern language.

Let us make a change of variables, we set

$$F = \frac{\phi}{1-\phi}, \quad t = \frac{p}{1-p} \quad \text{and } \tau = \log t;$$

this gives another function $\tilde{F} : \mathbf{R} \longrightarrow \mathbf{R}$ which is the function ϕ in the new coordinate system:

$$\tilde{F}(\tau) = \log F(e^{\tau}).$$

In this setting, the theorem reads as

Proposition 5.3. *The map $\tilde{F}$ is non-strictly expanding, that is, for every $\tau \in \mathbf{R}$ we have*

$$\tilde{F}'(\tau) \geq 1.$$

Note first how the probabilities can be described in terms of t. Namely, the probability of a given subgraph $\Gamma \subset G$ with $k = |\Gamma|$ edges is

$$\mathbf{P}_p(\Gamma) = p^k (1-p)^{n-k},$$

where n denotes the number of edges in G. Thus, dividing all the probabilities by $(1-p)^n$, we see that the probability of $\Gamma \subset G$ is proportional to $\frac{p^k}{(1-p)^k} = t^k$.

Hence, $\frac{\phi(p)}{1-\phi(p)}$ can be represented as $\dfrac{S_{I \twoheadrightarrow O}(t)}{S_{I \to O}(t)}$, where, for given $X \subset \{0,1\}^E$,

$$S_X(t) = \sum_{\Gamma \in X} t^{|\Gamma|}.$$

We have

$$\begin{aligned}
\frac{d\tilde{F}}{dt}(t) = \frac{d \log F}{dt}(t) = t\, \frac{d \log \dfrac{S_{I \twoheadrightarrow O}}{S_{I \to O}}}{dt}(t) \\
= \frac{t \sum_{\Gamma \in I \to O} |\Gamma|\, t^{|\Gamma|-1}}{S_{I \to O}(t)} - \frac{t \sum_{\Gamma \in I \twoheadrightarrow O} |\Gamma|\, t^{|\Gamma|-1}}{S_{I \twoheadrightarrow O}(t)} = \frac{\sum_{\Gamma \in I \to O} |\Gamma|\, t^{|\Gamma|}}{S_{I \to O}(t)} - \frac{\sum_{\Gamma \in I \twoheadrightarrow O} |\Gamma|\, t^{|\Gamma|}}{S_{I \twoheadrightarrow O}(t)} \\
= \mathbf{E}_p\left(|\Gamma| \,\middle|\, \Gamma \in I \to O\right) - \mathbf{E}_p\left(|\Gamma| \,\middle|\, \Gamma \in I \twoheadrightarrow O\right).
\end{aligned}$$

In other words, we have to show that the mean number of edges under the conditions that there is a current exceeds by at least 1 the mean number of edges assuming that there is no current.

This is done by the following

Lemma 5.4. *Let $\varphi : \{0,1\}^n \longrightarrow \{0,1\}$ be a monotonic boolean function, $A = \varphi^{-1}(0)$, $B = \varphi^{-1}(1)$ both nonempty, and let $k : \{0,1\}^n \longrightarrow \{0,\ldots,n\}$ be the weight function* (i. e. $k(x) = x_1 + \ldots + x_n$). *Then*

$$\mathbf{E}_p(k \,|\, B) - \mathbf{E}_p(k \,|\, A) \geq 1. \tag{IV.47}$$

Proof. Let us multiply both sides of (IV.47) by $\mathbf{P}_p(A)\,\mathbf{P}_p(B)$: it becomes

$$\mathbf{P}_p(A)\,\mathbf{E}_p(k\,\mathbf{1}_B) - \mathbf{P}_p(B)\,\mathbf{E}_p(k\,\mathbf{1}_A) \geq \mathbf{P}_p(A)\,\mathbf{P}_p(B). \tag{IV.48}$$

We will prove (IV.48) by induction on n, no longer assuming that A and B are nonempty (for if A or B is empty, (IV.48) becomes trivial). Note that

$$\mathbf{E}_p(k\,\mathbf{1}_B) = \mathbf{E}_p(k) - \mathbf{E}_p(k\,\mathbf{1}_A) = np - \mathbf{E}_p(k\,\mathbf{1}_A),$$

and hence the left hand side in (IV.48) is equal to

$$np\,\mathbf{P}_p(A) - \mathbf{P}_p(A)\,\mathbf{E}_p(k\,\mathbf{1}_A) - \mathbf{P}_p(B)\,\mathbf{E}_p(k\,\mathbf{1}_A) = np\,\mathbf{P}_p(A) - \mathbf{E}_p(k\,\mathbf{1}_A).$$

So, we want to prove

$$np\,\mathbf{P}_p(A) - \mathbf{E}_p(k\,\mathbf{1}_A) \geq \mathbf{P}_p(A)\left(1 - \mathbf{P}_p(A)\right). \tag{IV.49}$$

Let us decompose A with respect to the last coordinate x_n: denote by A_0 and A_1, its intersections with $\{x_n = 0\}$ and $\{x_n = 1\}$ respectively, that we consider as subsets of $\{0,1\}^{n-1}$. Then, $A_0 \supset A_1$, due to the monotonicity of φ, and hence for the $(n-1)$-dimensional probabilities $\alpha_0 = \mathbf{P}_{p,n-1}(A_0)$, $\alpha_1 = \mathbf{P}_{p,n-1}(A_1)$ we have $\alpha_0 \geq \alpha_1$.

Applying the induction assumption, we have

$$(n-1)\,p\,\alpha_0 - \mathbf{E}_{p,n-1}\left(k_{n-1}\,\mathbf{1}_{A_0}\right) \geq \alpha_0\,(1-\alpha_0), \tag{R_0}$$
$$(n-1)\,p\,\alpha_1 - \mathbf{E}_{p,n-1}\left(k_{n-1}\,\mathbf{1}_{A_1}\right) \geq \alpha_1\,(1-\alpha_1). \tag{R_1}$$

where $k_{n-1} : \{0,1\}^{n-1} \longrightarrow \{0,\ldots,n-1\}$ is the weight function on the $(n-1)$-dimensional cube.

Adding together (R_0) with factor $(1-p)$ and (R_1) with factor p, we get

$$(n-1)\,p\,\left((1-p)\,\alpha_0 + p\,\alpha_1\right) - \mathbf{E}_p\left((x_1+\ldots+x_{n-1})\,\mathbf{1}_A\right) \geq (1-p)\,(\alpha_0-\alpha_0^2) + p\,(\alpha_1-\alpha_1^2). \tag{IV.50}$$

Compare it to the desired (IV.49): as $\mathbf{P}_p(A) = p\,\mathbf{P}_{p,n-1}(A_1) + (1-p)\,\mathbf{P}_{p,n-1}(A_0)$, in the left hand side we are missing

$$p\,\left((1-p)\,\alpha_0 + p\,\alpha_1\right) - \mathbf{E}_p\left(x_n\,\mathbf{1}_A\right),$$

while the right hand side should instead be

$$\left((1-p)\,\alpha_0 + p\,\alpha_1\right) - \left((1-p)\,\alpha_0 + p\,\alpha_1\right)^2.$$

Let us simplify these differences:

$$p\,\left((1-p)\,\alpha_0 + p\,\alpha_1\right) - \mathbf{E}_p\left(x_n\mathbf{1}_A\right) = p\,\left((1-p)\,\alpha_0 + p\,\alpha_1\right) - p\,\alpha_1 = p(1-p)\,(\alpha_0-\alpha_1).$$

For the right hand side, the difference is

$$\begin{aligned}\Big[\left((1-p)\,\alpha_0 + p\,\alpha_1\right) - \left((1-p)\,\alpha_0 + p\,\alpha_1\right)^2\Big] - \Big[(1-p)\,(\alpha_0-\alpha_0^2) + p\,(\alpha_1-\alpha_1^2)\Big]\\ = -\left((1-p)\,\alpha_0 + p\,\alpha_1\right)^2 + (1-p)\,\alpha_0^2 + p\,\alpha_1^2 = p(1-p)\,\alpha_0^2 + p(1-p)\,\alpha_1^2 - 2\,p(1-p)\,\alpha_0\alpha_1\\ = p(1-p)\,(\alpha_0-\alpha_1)^2.\end{aligned}$$

As

$$p(1-p)\,(\alpha_0-\alpha_1) \geq p(1-p)\,(\alpha_0-\alpha_1)^2,$$

we thus have

$$p\,\mathbf{P}_p(A) - \mathbf{E}_p(x_n\,\mathbf{1}_A) \geq \Big[\mathbf{P}_p(A) - \mathbf{P}_p(A)^2\Big] - \Big[(1-p\,\alpha_0-\alpha_0^2) + p\,(\alpha_1-\alpha_1^2)\Big], \tag{IV.51}$$

and adding (IV.51) to (IV.50), we obtain the desired (IV.49). □

Bibliografia

[Add-Ree] L. Addario-Berry & B. Reed. Minima in branching random walks. *Ann. Probab.* **37**, no. **3** (2009), 1044-1079.

[Ald-Ban] D. J. Aldous & A. Bandyopadhyay. A survey of max-type recursive distributional equations. *Ann. Appl. Probab.* **15**, no. **2** (2005), 1047–1110.

[Arn 1] V. I. Arnol'd. Small denominators. I. Mapping the circle onto itself. *Izv. Akad. Nauk SSSR Ser. Mat.* **25** (1961), 21–86.

[Arn 2] _____ & B. A. Khesin. *Topological Methods in Hydrodynamics.* Applied Mathematical Sciences **125**, Springer (1998).

[Ast-Jon-Kup-Sak] K. Astala, P. Jones, A. Kupiainen & E. Saksman. Random conformal weldings. *Acta Math.* **207**, no. **2** (2011), 203–254.

[Ávi-Koc 1] A. Ávila & A. Kocsard. Cohomological equations and invariant distributions for minimal circle diffeomorphisms. *Duke Math. J.* **158**, no. **3** (2011), 501–536.

[Bar-Jin-Rho-Var] J. Barral, X. Jin, R. Rhodes & V. Vargas. Gaussian Multiplicative Chaos and KPZ Duality. *Comm. Math. Phys.* **323**, no. **2** (2013), 451–485.

[Bar-Kup-Nik-Sak-Web] _____ , A. Kupiainen, M. Nikula, E. Saksman & C. Webb. Critical Mandelbrot Cascades. *Comm. Math. Phys.* **325**, no. **2** (2014), 685–711.

[Ben] I. Benjamini. Euclidean vs. graph metric. Dans *Erdös centennial*, curated by L. Lovász, V. T. Sós & D. Palvolgyi. Bolyai Society Mathematical Studies **25**, Springer (2013), 35–57.

[Ben-Sch] _____ & O. Schramm. KPZ in one dimensional random geometry of multiplicative cascades. *Comm. Math. Phys.* **289**, no. **2** (2009), 653–662.

[Bot] R. Bott. On some formulas for the characteristic classes of group-actions. Dans *Differential topology, foliations and Gelfand-Fuks cohomology (Proc. Sympos., Pontifícia Univ. Católica, Rio de Janeiro, 1976).* Lecture Notes in Math. **652**, Springer (1978).

[Bri+] M. G. Brin et al. On Shavgulidze's Proof of the Amenability of some Discrete Groups of Homeomorphisms of the Unit Interval. Preprint arXiv:0908.1353

[Bro] K. S. Brown. *Cohomology of groups.* Graduate Texts in Mathematics **87**, Springer-Verlag (1994).

[Cam-Mar 1] R. H. Cameron & W. T. Martin. Transformations of Wiener integrals under translations. *Ann. of Math. (2)* **45** (1944), 386–396.

[Cam-Mar 2] _____ & _____ . The transformation of Wiener integrals by nonlinear transformations. *Trans. Amer. Math. Soc.* **66** (1949), 253–283.

[Cha] M. Chaperon. Invariant manifolds revisited. *Proc. Steklov Inst. Math.* **236**, no. **1** (2002), 415–433.

[Che-Gro] J. Cheeger & M. Gromov. L_2-cohomology and group cohomology. *Topology* **25**, no. **2** (1986), 189–215.

[Cro-Nav] S. Crovisier & A. Navas. On the C^1 convergence of ergodic means for circle diffeomorphisms. *Communication privée.*

[Dav] F. David. Conformal field theories coupled to 2-D gravity in the conformal gauge. *Mod. Phys. Lett. A* **3**, no. **17** (1988), 1651–1656.

[Den] A. Denjoy. Sur les courbes définies par les équations différentielles à la surface du tore. *J. Math. Pures Appl. (9)* **11** (1932), 333–375.

[Der-Kle-Nav] B. Deroin, V. Kleptsyn & A. Navas. Sur la dynamique unidimensionnelle en régularité intermédiaire. *Acta Math.* **199**, no. **2** (2007), 199–262.

[Dis-Kaw] J. Distler & H. Kawai. Conformal field theory and 2D quantum gravity. *Nucl. Phys. B* **321** (1989), 509–527.

[Dou-Yoc] R. Douady & J.-C. Yoccoz. Nombre de rotation des difféomorphismes du cercle et mesures automorphes. *Regul. Chaotic Dyn.* **4**, no. **4** (1999), 19–38.

[Dum-Ser] G. Duminy & V. Sergiescu. Sur la nullité de Godbillon-Vey. *C. R. Acad. Sci. Paris* **292** (1981), 821–824.

[Dup-She] B. Duplantier & S. Sheffield. Liouville Quantum Gravity and KPZ. *Inv. Math.* **185**, no. **2** (2011), 333–393.

[Dup-Rho-She-Var] _____ , R. Rhodes, _____ & V. Vargas. Renormalization of Critical Gaussian Multiplicative Chaos and KPZ Relation. *Comm. Math. Phys.*, to appear (2014), 1–48.

[Eps] D. B. A. Epstein. Almost all subgroups of a Lie group are free. *J. Algebra* **19** (1971), 261–262.

[Fre] D. H. Fremlin. *Topological measure spaces.* Part I. Deuxième édition. Torres Fremlin (2006).

[Fuc] D. B. Fuchs. *Cohomology of infinite-dimensional Lie algebras.* Contemporary Soviet Mathematics, Consultants Bureau (1986).

[Gar] C. Garban. Quantum gravity and the KPZ formula [after Duplantier-Sheffield]. *Sém. Bourbaki, Vol. 2011/12* **1052**, Astérisque no. **352** (2013), 315–354.

[Ghy 1] É. Ghys. L'invariant de Godbillon-Vey. *Sém. Bourbaki, Vol. 1988/89* **706**, Astérisque no. **177-178** (1989), 155–181.

[Ghy 2] _____ . Groups acting on the circle. *Enseign. Math. (2)* **47**, no. **3-4** (2001), 329–407.

[Ghy-Ser] _____ & V. Sergiescu. Sur un groupe remarquable de difféomorphismes du cercle. *Comment. Math. Helv.* **62**, no. **2** (1987), 185–239.

[Gin-Moo] P. Ginsparg & G. Moore. *Lectures on 2D gravity and 2D string theory.* Dans *Recent direction in particle theory, Proceedings of the 1992 TASI*, éditeurs J. Harvey & J. Polchinski. *World Scientific* (1993).

[Got-Hed] W. H. Gottschalk & G. A. Hedlund. *Topological dynamics.* AMS Colloquium Publications, Vol. **36** (1955).

[Gro] M. Gromov. Asymptotics invariants of infinite groups. Dans *Geometric Group Theory, vol. 2*, éditeurs G. A. Niblo & M. A. Roller. *London Math. Soc. Lecture Note Series* **182**, Cambridge University Press (1993).

[Hae 1] A. Haefliger. Feuilletages sur les variétés ouvertes. *Topology* **9** (1970), 183–194.

[Hae 2] _____ . Some remarks on foliations with minimal leaves. *J. Differential Geom.* **15**, no. **2** (1980), 269–284.

[Hae-Ban] _____ & Li Banghe. Currents on a circle invariant by a Fuchsian group. *Geometric dynamics (Rio de Janeiro, 1981)*, dans *Lecture Notes in Math.* **1007**, Springer (1983), 369–378.

[Hal] P. R. Halmos. *Measure theory.* D. Van Nostrand Company (1950).

[Ham-Kum] B. M. Hambly & T. Kumagai. Diffusion on the scaling limit of the critical percolation cluster in the diamond hierarchical lattice. *Comm. Math. Phys.* **295** (2010), 26–69.

[Her 1] M. R. Herman. Sur la conjugaison différentiable des difféomorphismes du cercle à des rotations. *Publ. Math. de l'IHÉS* **49** (1979), 5–233.

[Her 2] _____ . Mesure de Lebesgue et nombre de rotation. Dans *Geometry and topology (Proc. III Latin Amer. School of Math., Inst. Mat. Pura Aplicada CNPq, Rio de Janeiro, 1976)*. pp. 271–293. Lecture Notes in Math. **597**, Springer (1977), 271–293.

[Hur-Kat] S. Hurder & A. Katok. Differentiability, rigidity and Godbillon-Vey classes for Anosov flows. *Publ. Math. de l'IHÉS* **72** (1991), 5–61.

[Kah] J.-P. Kahane. Sur le chaos multiplicatif. *Ann. Sci. Math. Québec* **9**, no. **2** (1985), 105–150.

[Kah-Pey] _____ & J. Peyrière. Sur certaines martingales de Benoit Mandelbrot. *Advances in Math.* **22**, no. **2** (1976), 131–145.

[Kar-Shr] I. Karatzas & S. E. Shreve. *Brownian motion and stochastic calculus*, deuxième édition. Graduate Texts in Mathematics **113**, Springer-Verlag (1994).

[Kat-Has] A. B. Katok & B. Hasselblatt. *Introduction to the modern theory of dynamical systems.* Enc. Math. and Appl., Cambridge University Press (1995).

[Katz] Y. Katznelson. Sigma-finite invariant measures for smooth mappings of the circle. *J. d'Analyse Math.* **31** (1977), 1–18.

[Katz-Orn] _____ & D. Ornstein. The differentiability of the conjugacy of certain diffeomorphisms of the circle. *Ergodic Theory Dynam. Systems* **9**, no. **4** (1989), 643–680.

[Kha-Sin 1] K. M. Khanin & Ya. G. Sinai. A new proof of M. Herman's theorem. *Comm. Math. Phys.* **112**, no. **1** (1987), 89–101.

[Kha-Sin 2] _____ & _____ . Smoothness of conjugacies of diffeomorphisms of the circle with rotations. *Russian Math. Surveys* **44**, no. **1** (1989), 69–99.

[Kha-Tep] _____ & A. Teplins'kyi. Herman's theory revisited. *Invent. Math.* **178**, no. **2** (2009), 333–344.

[Khe-Wen] B. A. Khesin & R. Wendt. *The Geometry of Infinite-Dimensional Groups.* Springer (2009).

[Khr-Kle-Tri] M. Khristoforov, V. Kleptsyn & M. Triestino. Stationary random distances on hierarchical graphs and (min, +)-type recursive distributional equations. Preprint arXiv:1310.6116

[Kni-Pol-Zam] V. G. Knizhnik, A. M. Polyakov & A. B. Zamolodchikov. Fractal structure of 2D quantum gravity. *Modern Phys. Lett. A* **3**, no. **8** (1988), 819–826.

[Kod-Mat] H. Kodama & S. Matsumoto. Minimal C^1-diffeomorphisms of the circle which admit measurable fundamental domains. *Proc. of the AMS* **141**, no. **6** (2013), 2061–2067.

[Kol] A. N. Kolmogorov. On conservation of conditionally periodic motions under a small variation of the Hamiltonian function (Russian). *Dokl. Akad. Nauk SSSR* **98** (1954), 527–530.

[Kop] N. Kopell. Commuting diffeomorphisms. Dans *Global Analysis, Proc. Sympos. Pure Math.* vol. **XIV** (1968),165–184.

[Kos] A. V. Kosyak. Irreducible regular Gaussian representations of the groups of the interval and circle diffeomorphisms. *J. Funct. Anal.* **125**, no. **2** (1994), 493–547.

[Kra-Sch] B. Kra & J. Schmeling. Diophantine classes, dimension and Denjoy maps. *Acta Arith.* **105**, no. **4** (2002), 323–340.

[Kuz] P. A. Kuzmin. On circle diffeomorphisms with discontinuous derivatives and quasi-invariance subgroups of Malliavin-Shavgulidze measures. *J. Math. Anal. Appl.* **330**, no. **1** (2007), 744–750.

[LeG] J.-F. Le Gall. Uniqueness and universality of the Brownian map. *Ann. Probab.* **41**, no. **4** (2013), 2880–2960.

[LeG-Mie] _____ & G. Miermont. Scaling limits of random trees and planar maps. Dans *Probability and statistical physics in two and more dimensions*, Clay Math. Proc. **15**, AMS (2012), 155–211.

[LeG-Pau] _____ & F. Paulin. Scaling limits of bipartite planar maps are homeomorphic to the 2-sphere. *Geom. Funct. Anal.* **18**, no. **3** (2008), 893–918.

[Lév] _____ . *Processus stochastiques et mouvement brownien.* Les Grands Classiques Gauthier-Villars, Éditions Jacques Gabay (1992). Réimpression de la deuxième (1965) édition.

[Mac] G. W. Mackey. Borel structure in groups and their duals. *Trans. Amer. Math. Soc.* **85** (1957), 134–165.

[Mal-Mal 1] M. P. Malliavin & P. Malliavin. Mesures quasi invariantes sur certains groupes de dimension infinie. *C. R. Acad. Sci. Paris Sér. I Math.* **311**, no. **12** (1990), 765–768.

[Mal-Mal 2] _____ & _____ . An infinitesimally quasi-invariant measure on the group of diffeomorphisms of the circle. *Special functions (Okayama, 1990)*, ICM-90 Satell. Conf. Proc., Springer (1991), 234–244.

[Mal] P. Malliavin. The canonic diffusion above the diffeomorphism group of the circle. *C. R. Acad. Sci. Paris Sér. I Math.* **329**, no. **4** (1999), 325–329.

[Man-Yor] R. Mansuy & M. Yor. *Aspects of Brownian motion.* Universitext, Springer-Verlag (2008).

[Mat 1] J. N. Mather. Commutators of diffeomorphisms. *Comment. Math. Helv.* **49** (1974), 512–528.

[Mat 2] _____ . Commutators of diffeomorphisms: II. *Comment. Math. Helv.* **50** (1975), 33–40.

[Mat 3] _____ Commutators of diffeomorphisms, III: a group which is not perfect. *Comment. Math. Helv.* **60** (1985), 122–124.

[Mat-Hir-Yor] H. Matsumoto & M. Yor. Exponential functionals of Brownian motion. I. Probability laws at fixed time. *Probab. Surv.* **2** (2005), 312–347.

[McD] D. McDuff. C^1-minimal subsets of the circle. *Ann. Inst. Fourier* **31**, no. **1** (1981), vii, 177–193.

[McK] H. P. McKean. Geometry of differential space. *Ann. Probability* **1** (1973), 197–206.

[Mie 1] G. Miermont. On the sphericity of scaling limits of random planar quadrangulations. *Electron. Commun. Probab.* **13** (2008), 248–257.

[Mie 2] _____ . The Brownian map is the scaling limit of uniform random plane quadrangulations. *Acta Math.* **210**, no. **2** (2013), 319–401.

[Mil-She] J. Miller & S. Sheffield. Quantum Loewner Evolution. Preprint arXiv:1312.5745

[dMe-vSt] W. de Melo & S. van Strien. *One-dimensional dynamics.* Erg. Math. Grenz. (3), **25**. Springer-Verlag (1993).

[Moo-Sha] E. F. Moore & C. E. Shannon. Reliable circuits using less reliable relays. I. *J. Franklin Inst.* **262** (1956), 191–208.

[Mor] S. Morita. *Geometry of Characteristic Classes.* Translations of Mathematical Monographs, AMS (2001).

[Mör-Per] P. Mörters & Y. Peres. *Brownian motion.* Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press (2010).

[Mos] J. Moser. A rapidly convergent iteration method and non-linear differential equations. II. *Ann. Scuola Norm. Sup. Pisa (3)* **20** (1966), 499–535.

[Nak] Y. Nakayama. Liouville field theory: a decade after the revolution. *Int. J. Mod. Phys. A* **19** (2004), 2771–2930.

[Nav 1] A. Navas. *Groups of circle diffeomorphisms.* Chicago Lectures in Mathematics (2011).

[Nav 2] _____ . Sur les rapprochements par conjugaison en dimension 1 et classe C^1. *Compositio Math.*, to appear (2014).

[Nav-Tri] _____ & M. Triestino. On the invariant distributions of C^2 circle diffeomorphisms of irrational rotation number. *Math. Z.* **274**, no. **1** (2013), 315–321.

[Ner] Yu. A. Neretin. Representations of Virasoro and affine Lie algebras. *Representation theory and noncommutative harmonic analysis, I*, Enc. Math. Sci. **22**, Springer (1994), 157–234.

[Nor 1] A. Norton. Denjoy minimal sets are far from affine. *Ergodic Theory Dynam. Systems* **22**, no. **6** (2002), 1803–1812.

[Nor 2] _____ . Denjoy's theorem with exponents. *Proc. Amer. Math. Soc.* **127**, no. **10** (1999), 3111–3118.

[Poi] H. Poincaré. Mémoire sur les courbes définies par les équations différentielles, IV. *J. Math. Pures Appl. (3)* **2** (1886), 151–217.

[Pol 1] A.M. Polyakov. Quantum geometry of bosonic strings. *Phys. Lett.* **103 B**, no. **3** (1981), 207–210.

[Pol 2] _____ . *Gauge fields and strings.* Harwood Academic Publishers (1987).

[Pre-Seg] A. Pressley & G. Segal. *Loop groups.* Oxford Mathematical Monographs (1988).

[Rev-Yor] D. Revuz & M. Yor. *Continuous martingales and Brownian motion.* Troisième édition, Springer-Verlag (1999).

[Rho-Var 1] R. Rhodes & V. Vargas. KPZ formula for log-infinitely divisible multifractal. *ESAIM Prob. Stat.* **15** (2011), 358–371 (2011).

[Rho-Var 2] _____ & _____ . Gaussian multiplicative chaos and applications: a review. Preprint arXiv:1305.6221

[Sha 1] E. T. Shavgulidze. An example of a measure which is quasi-invariant relative to the action of a group of diffeomorphisms of the circle. *Funktsional. Anal. i Prilozhen.* **12**, no. **3** (1978), 55–60, 96.

[Sha 2] _____ . A measure that is quasi-invariant with respect to the action of a group of diffeomorphisms of a finite-dimensional manifold. *Dokl. Akad. Nauk SSSR* **303**, no. **4** (1988), 811–814.

[Sha 3] _____ . Quasi-invariant measures on groups of diffeomorphisms. *Tr. Mat. Inst. Steklova* **217** (1997), 189–208.

[Sha 4] _____ . Some properties of quasi-invariant measures on groups of diffeomorphisms of the circle. *Russ. J. Math. Phys.* **7**, no. **4** (2000), 464–472.

[Sha 5] _____ . Amenability of discrete subgroups of the group of diffeomorphisms of the circle. *Russ. J. Math. Phys.* **16**, no. **1** (2009), 130–132.

[She 1] S. Sheffield. Gaussian free fields for mathematicians. *Probab. Theory Related Fields* **139**, no. **3-4** (2007), 521–541.

[She 2] _____ . Conformal weldings of random surfaces: SLE and the quantum gravity zipper. Preprint arXiv:1012.4797

[Sie] C. L. Siegel. Iteration of analytic functions. *Ann. of Math. (2)* **43** (1942), 607–612.

[Smi] N. V. Smirnov. Sur la distribution de ω^2 (criterium de M. von Mises). *C. R. Acad. Sci. Paris* **202** (1936), 449–452.

[Sta] J. Stark. Smooth conjugacy and renormalisation for diffeomorphisms of the circle. *Nonlinearity* **1**, no. **4** (1988), 541–575.

[Ste] R. C. Steinlage. On Haar measure in locally compact T_2 spaces. *Amer. J. Math.* **97** (1975), 291—307.

[Tec] N. Tecu. Random conformal welding towards criticality. Prépublication `http://users.math.yale.edu/~nt73/Subpages/RandomConformalWelding.pdf`.

[Tep] O. Yu. Teplins'kyi. On the smoothness of the conjugacy of the circle diffeomorphisms to rigid rotations. *Ukrainian Math. J.* **60**, no. **2** (2008), 310–326.

[Tes] J. Teschner. Liouville theory revisited. *Classical Quantum Grav.* **18**, no. **23** (2001), R153–R222.

[Thu] W. Thurston. Foliations and groups of diffeomorphisms. *Bull. Amer. Math. Soc.* **80** (1974), 304–307.

[Tol] L. Tolmatz. On the distribution of the square integral of the Brownian bridge. *Ann. Probab.* **30**, no. **1** (2002), 253–269.

[Tsu 1] T. Tsuboi. On the Hurder-Katok extension of the Godbillon-Vey invariant. *J. Fac. Sci. Univ. Tokyo Sect. IA Math.* **37**, no. **2** (1990).

[Tsu 2] _____ . Area functionals and Godbillon-Vey cocycles. *Ann. Inst. Fourier* **42**, no. **2** (1992), 421–447.

[Wei] A. Weil. *L'intégration dans les groupes topologiques et ses applications.* Actual. Sci. Ind., no. **869**, Hermann et Cie., Paris (1940).

[Wer] W. Werner. *Percolation et modèle d'Ising.* Cours spécialisés, no. **16**, SMF (2009).

[Wie 1] N. Wiener. Differential space. *Journal Math. Phys.* **2** (1923), 131–174.

[Wie 2] _____ . Generalized harmonic analysis. *Acta Math.* **55**, no. **1** (1930), 117–258.

[Yoc 1] J.-C. Yoccoz. Conjugaison différentiable des difféomorphismes du cercle dont le nombre de rotation vérifie une condition diophantienne. *Ann. Sci. École Norm. Sup. (4)* **17**, no. **3** (1984), 333–359.

[Yoc 2] _____ . Petits diviseurs en dimension 1. *Astérisque* no. **231**, SMF (1995).

[Yor] M. Yor. *Exponential functionals of Brownian motion and related processes.* Springer Finance, Springer-Verlag (2001).