Thèse de doctorat de l'Université de Neuchâtel

Spécialité :

Mathématiques

Présentée par

Aline KURTZMANN

Pour obtenir le grade de docteur ès sciences de l'Université de Neuchâtel

Sujet de la thèse :

COMPORTEMENT ASYMPTOTIQUE de DIFFUSIONS RENFORCÉES SUR $\mathbb{R}^d$

Soutenue le 22 mai 2007 devant le jury composé de :

M. Michel BENAIM (directeur de thèse)
M. Patrick CATTIAUX (rapporteur, Toulouse)
M. Thomas MOUNTFORD (examinateur, E.P.F. Lausanne)
M. Olivier RAIMOND (rapporteur, Paris Sud Orsay)
M. Alain VALETTE (président, Neuchâtel)
M. Cédric VILLANI (examinateur, E.N.S. Lyon)

Université de Neuchâtel

FACULTE DES SCIENCES
Secrétariat-Décanat de la faculté
- Rue Emile-Argand 11
- CP 158
- CH-2009 Neuchâtel

IMPRIMATUR POUR LA THESE

Comportement asymptotique de diffusions renforcées sur R^n

Aline KURTZMANN

UNIVERSITE DE NEUCHATEL

FACULTE DES SCIENCES

La Faculté des sciences de l'Université de Neuchâtel,
sur le rapport des membres du jury

MM. M. Benaïm (directeur de thèse),
A. Valette, T. Mountford (EPF Lausanne),
P. Cattiaux (Paris), C. Villani (Lyon)
et O. Raimond (Paris)

autorise l'impression de la présente thèse.

Neuchâtel, le 5 juin 2007

Le doyen :
T. Ward

UNIVERSITE DE NEUCHATEL
FACULTE DES SCIENCES
Secrétariat-décanat de la faculté
Rue Emile-Argand 11 - CP 158
CH-2009 Neuchâtel

■ Téléphone : +41 32 718 21 00 ■ Fax : +41 32 718 21 03 ■ E-mail : secretariat.sciences@unine.ch ■ www.unine.ch/sciences

Remerciements

Je tiens tout d'abord à remercier Michel Benaïm pour avoir accepté d'encadrer ma thèse et m'avoir proposé un joli sujet. Merci également de m'avoir permis de travailler dans des conditions fort agréables à l'université de Neuchâtel.

Je suis très reconnaissante à Patrick Cattiaux et Olivier Raimond d'avoir accepté d'être les rapporteurs de ce travail. Je remercie en particulier Patrick Cattiaux de m'avoir décrypté le monde, alors inconnu pour moi, des inégalités fonctionnelles ainsi que Olivier Raimond pour m'avoir proposé le travail qui constitue la première partie de cette thèse.

Je suis également très honorée de la participation de Thomas Mountford, Alain Valette ainsi que Cédric Villani au jury. Je remercie tout particulièrement Cédric Villani pour son enthousiasme débordant et pour les discussions que nous avons pu avoir.

La première partie de cette thèse fait suite à un travail en commun avec Sébastien Chambeu. Ce fut un réel plaisir que de faire ces petites virées parisiennes afin de venir travailler avec toi.

J'aimerais également exprimer toute ma reconnaissance à Victor Kleptyn (bien que notre travail ne figure pas dans ce manuscrit). Travailler avec toi est extrêmement passionnant, enrichissant (et fatiguant aussi !) Je te remercie pour ton soutien, ton optimisme et tes conseils éclairés.

Je tiens aussi à remercier Pierre Tarrès ; à la fois pour ta disponibilité permanente lors de ton année à Neuchâtel, nos discussions mathématiques (entre autres !), ton amitié et pour m'avoir constamment encouragée. Je me réjouis de venir travailler avec toi.

Merci à toute l'équipe de mathématiques, profs, post-docs et thésards de l'université de Neuchâtel pour leur bonne humeur. Je pense en particulier à Tatiana et Ola pour nos sorties sportives à la piscine, Erwann pour m'avoir aidé à faire mes premiers pas dans la recherche, ainsi que Roger, puis Kola pour avoir mis une ambiance à la fois si chaleureuse et studieuse dans le bureau. J'exprime également ici toute ma reconnaissance à Christine Vuilleumier pour sa gentillesse et sa capacité à régler dans la bonne humeur tous les problèmes, petits ou gros.

Un grand merci également à tous mes amis. Je pense à ceux de très longue date, Marie, Delphine et Nicolas, mais aussi aux "Strasbourgeois" comme Béatrice, Marie et Manu, et bien entendu aux "Parisiens" à l'image d'Ashkan et Delia, Najat, Marc, Sylvain et Anne, ou encore à Adeline et Jérôme qui forment un savant mélange des deux.

Enfin, je tiens à exprimer toute mon affection envers mes grand-parents, parents, frères et nièce. Merci surtout à mes parents, à qui je dois tout.

Pour finir, toutes mes pensées sont pour Jean-Frédéric. Merci pour ta patience et tout ce que tu m'as apporté.

“- Et alors, cela va durer encore longtemps ? interrogea Woland. Echec au roi.
- J’ai sans doute mal entendu, mon maître, lui répondit le chat, il n’y a pas et il ne peut y avoir échec au roi.
- Je répète : échec au roi.
- Messire, fit le chat sur un ton faussement inquiet, vous devez être surmené : il n’y a pas échec au roi !
- Le roi est sur la case D2”, dit Woland sans regarder l’échiquier.
“Messire, vous me faites peur !” gémit le chat en affectant une mine épouvantée. “Il n’y a pas de roi sur cette case !
- Que dis-tu là ?” et Woland, interloqué, regarda l’échiquier où l’officier qui se tenait sur la case du roi détournait la tête et se dissimulait le visage derrière sa main.
“Tu es un beau gredin, dit pensivement Woland.
- Messire ! J’en réfère derechef à la logique’, dit le chat en pressant ses pattes contre son poitrail. “Si un joueur annonce échec au roi alors que le roi, depuis longtemps, n’est plus sur l’échiquier, l’échec est déclaré nul.
- Tu abandonnes, oui ou non ? cria Woland d’une voix terrible.
- Laissez-moi le temps de réfléchir”, lui répondit humblement le chat ; il s’accouda sur la table, se boucha les oreilles avec ses pattes et réfléchit profondément. Au bout d’un long moment il finit par déclarer : “J’abandonne.”
“Cette tête de mule est à tuer, murmura Azazello.
- Oui, j’abandonne, dit le chat, mais uniquement parce que je ne peux pas jouer dans une atmosphère de harcèlement, entretenue par des envieux !” Il se leva et les pièces du jeu rentrèrent d’elles-mêmes dans leur boîte.

Le Maître et Marguerite, M. Boulgakov

Table des matières

Remerciements i

Introduction 1
1. Notions provenant de la théorie des systèmes dynamiques 4
2. Application aux algorithmes stochastiques 10
3. Quelques inégalités fonctionnelles 17
4. Diffusions renforcées par la moyenne de la mesure d'occupation normalisée 19
5. Diffusions renforcées par la mesure d'occupation normalisée 24

Bibliographie 35

partie 1. Diffusions renforcées par la moyenne de la mesure d'occupation normalisée 41

Chapitre 1. Comportement ergodique et convergence presque sûre de certaines diffusions auto-interactives 43
1. Introduction 43
2. Notation, hypotheses and existence 45
3. A motivating example: the quadratic case 47
4. Study of the process Y 53
5. Asymptotic behavior of Y 59
6. Behavior of X in the case of a general potential 65

Bibliography 69

Chapitre 2. Convergence en loi de certaines diffusions auto-interactives : la méthode du recuit simulé 71
1. Introduction 71
2. Notation, hypothesis and existence 73
3. Former results 74
4. Asymptotic behavior of Y 74
5. The process X: convergence in law 86
6. Concluding remarks 87

Bibliography 89

partie 2. Diffusions renforcées par la mesure d'occupation normalisée 91

Chapitre 3. Panorama des processus avec renforcement 93
1. Introduction 93
2. Processus renforcé par ses extrema 96
3. Diffusions renforcées par la mesure d'occupation non normalisée 101
4. Diffusions renforcées par la mesure d'occupation normalisée 106

Bibliographie 115

Chapitre 4. Diffusions renforcées sur un espace non compact 119
1. Introduction 119
2. Motivation 122
3. Preliminaries and Tools 125
4. Main results 130
5. Study of the dynamical system Φ 133
6. Study of the family of semi-groups $(P_t^\mu, t \geq 0, \mu \in \mathcal{M}_\beta(\mathbb{R}^d; V))$ 140
7. Behavior of the occupation measure 147
8. Some ideas for diffusions in a Riemannian manifold 156
9. Conclusion 158

Bibliography 161

Chapitre 5. Exemple de diffusion renforcée sur $\mathbb{R}^2$ 163
1. Introduction 163
2. Notation and background 165
3. Example 168

Bibliography 179

Annexe A. Diffusions attirées par leur maximum 181

Annexe B. Décomposition d'une fonction strictement convexe à l'infini 185

Résumé 187

Introduction

Ce mémoire porte sur l'étude des diffusions renforcées sur des espaces non compacts, via la méthode dite de "l'Equation Différentielle Ordinaire" (E.D.O.). Cette méthode fut introduite par Ljung [**82**] et, depuis, étudiée intensément. Elle s'avère être la source de bon nombre de travaux importants, comme ceux de Kushner et ses co-auteurs (voir le livre écrit avec Clark [**76**] ou celui avec Yin [**77**]) et plus récemment ceux de Benveniste, Métivier et Priouret [**16**] ou Duflo [**47**]. Pourtant, la plupart de ces travaux sont conduits sous des dynamiques forts simples. Il a fallu attendre les travaux de Benaïm et Hirsch [**6**] pour une étude plus complète de cette méthode. Elle s'est ensuite beaucoup développée depuis les années 90, afin d'étudier des processus renforcés (qui dépendent de leur trajectoire passée). Pour le moment, il n'existe pas encore de méthode unifiée pour traiter ce genre de processus. Notre étude est basée sur l'utilisation de la méthode de l'E.D.O. afin de décrire le comportement asymptotique de certains processus, satisfaisant de manière générale l'équation :

$$\mathrm{d}X_t = \mathrm{d}B_t - F(t, X_t, \mu_t)\mathrm{d}t, \tag{0.1}$$

où B est un mouvement brownien standard de dimension d et μ_t est la mesure d'occupation normalisée du processus X. De plus, la mesure d'occupation normalisée est définie par

$$\mu_t = \frac{r}{r+t}\mu + \frac{1}{r+t}\int_0^t \delta_{X_s}\mathrm{d}s, \tag{0.2}$$

où nous supposons donnés le poids initial $r > 0$ et la probabilité initiale μ. En fait, cela revient à supposer que le processus X suit la loi μ durant le temps r et nous le considérons après le temps r. Cet "artifice" nous permet simplement d'éviter les problèmes de définition de μ_t en $t = 0$ (que l'on aurait si on supposait $r = 0$). Remarquons également que μ_t est une mesure *aléatoire*.

Afin d'étudier le comportement asymptotique d'un tel processus, nous envisageons celui-ci (ou sa mesure d'occupation) comme étant un schéma (continu) d'approximation stochastique d'un certain système dynamique *déterministe*. C'est ce que nous appelons la "méthode de l'E.D.O.". En fait, nous montrons que μ_t, ou X_t selon les cas, est une *pseudo-trajectoire asymptotique* du système déterministe en question. Cette notion, définie un peu plus loin, permet de faire bénéficier le processus étudié des propriétés asymptotiques du système déterministe.

Les processus que nous traitons ici se divisent en deux catégories : les processus renforcés par la moyenne de leur mesure d'occupation normalisée ; et les processus renforcés par leur mesure d'occupation normalisée via une fonction d'interaction dans le terme de dérive. En résumé, nous allons dans un premier temps étudier le processus Z solution de l'E.D.S. suivante

$$\mathrm{d}Z_t = \mathrm{d}B_t - g(t)\nabla W(Z_t - \overline{\mu}_t)\mathrm{d}t, \;\; Z_0 = z. \tag{0.3}$$

Ici, B est un mouvement brownien standard de dimension d et $W : \mathbb{R}^d \to \mathbb{R}$ est un potentiel suffisamment lisse strictement convexe en dehors d'un compact. Nous définissons la moyenne (aléatoire) de μ_t par

$$\overline{\mu}_t = \frac{r}{r+t}\overline{\mu} + \frac{1}{r+t}\int_0^t Z_s \mathrm{d}s.$$

Le processus Z est inhomogène, car nous mettons un poids $g(t)$ dans le terme de dérive. Cela implique que, selon la fonction g, l'impact de la trajectoire passée sur le présent varie avec le temps. La différence $Z_t - \overline{\mu}_t$ est tout de même un processus de Markov inhomogène de dimension finie. En fait, cette différence est relativement proche d'une diffusion ordinaire (car il est markovien bien qu'inhomogène) et donc son étude sera assez poussée.

Un élément de la deuxième famille présentée ici est un processus X satisfaisant l'E.D.S.

$$\mathrm{d}X_t = \mathrm{d}B_t - \left(\nabla V(X_t) + \nabla W * \mu_t(X_t)\right)\mathrm{d}t, \;\; X_0 = x, \tag{0.4}$$

où nous avons conservé les notations de la diffusion précédente. Une différence notoire entre les processus Z et X est que, pour le premier, la moyenne de μ_t entre en jeu et non μ_t lui-même. Ici, la fonction $V : \mathbb{R}^d \to \mathbb{R}$ est un potentiel de confinement (strictement convexe à l'infini) suffisamment lisse, $W : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ est un potentiel d'interaction suffisamment lisse contrôlé par V. De plus, $W * \mu_t(x)$ est le "produit de convolution"

$$W * \mu_t(x) = \int_{\mathbb{R}^d} W(x,y)\mu_t(\mathrm{d}y).$$

Cette fois, le couple (X_t, μ_t) est markovien, mais contrairement au cas précédent, il est en général impossible de se ramener à un processus de dimension finie.

Cette étude généralise complètement celle faite précédemment, sur un espace compact, par Benaïm, Ledoux et Raimond [**10, 13**]. Le manque de compacité engendre bon nombre de problèmes techniques. Tout d'abord, il faut trouver la "bonne topologie" qui tient compte de la non compacité de l'espace. C'est pourquoi nous introduisons la V-norme (voir paragraphe 5). Les plus importantes difficultés consistent en l'existence d'un trou spectral uniforme et l'ultracontractivité (uniforme) du semi-groupe de Markov associé à la vraie diffusion X^μ qui satisfait l'équation (0.4) avec μ_t fixé à μ

(pour tout t) dans le terme de dérive. Pour ce faire, le potentiel de confinement V jour un rôle primordial. Ces notions fonctionnelles sont introduites dans le troisième paragraphe de cette introduction. Elles sont essentielles pour prouver que μ_t est pseudo-trajectoire asymptotique d'un certain système dynamique.

Pour étudier ces deux familles de processus à mémoire longue, nous utilisons la méthode de l'E.D.O. A la fin des années 1990, M. Benaïm et ses co-auteurs ont amorcé une nouvelle approche à l'approximation stochastique basée sur des techniques de système dynamique. Considérons le processus (à temps discret) *d'approximation stochastique* $(X_n, n \geq 0)$ à valeurs dans $\mathbb{R}^d$ et adapté à la filtration $\mathcal{F} = (\mathcal{F}_n, n \geq 0)$ donnée. Supposons que X_n satisfait l'équation

$$X_{n+1} - X_n = \frac{1}{n}(F(X_n) + \xi_{n+1} + r_n), \tag{0.5}$$

où F est un champ de vecteurs sur $\mathbb{R}^d$, le processus (ξ_n) satisfait $\mathbb{E}(\xi_{n+1}|\mathcal{F}_n) = 0$ et le terme (aléatoire) r_n tend vers zéro.

Ce processus d'approximation (X_n) peut être considéré comme une version discrète des processus renforcés que nous étudions dans cette thèse. C'est pourquoi nous commençons par donner quelques propriétés de ces processus, plus faciles à appréhender dans le cas discret, afin d'introduire les notions clés de notre étude (en temps continu).

Dans le cas particulier où $\xi_n \equiv 0$ et $r_n \equiv 0$, alors la version continue (modulo un changement de temps exponentiel) de (0.5) est l'E.D.O.

$$\frac{\mathrm{d}}{\mathrm{d}t} X_t = F(X_t). \tag{0.6}$$

Deux idées viennent alors à l'esprit. La première est que les trajectoires de la suite aléatoire (0.5) approchent celles de l'E.D.O. (0.6). En effet, la suite aléatoire est une sorte de schéma d'Euler afin de simuler (numériquement) l'E.D.O. avec un pas de $1/n$. Comme le pas est relativement petit, on peut espérer que le bruit n'aura pas de grande influence et donc les comportements asymptotiques de l'E.D.O. et de la suite aléatoire sont très semblables. La seconde remarque concerne le bruit. En effet, certaines trajectoires sont évitées par le processus aléatoire, comme par exemple les orbites périodiques linéairement instables.

Des problèmes peuvent cependant intervenir. Effectivement, les trajectoires de l'E.D.O. (0.6) peuvent, elles aussi, être fort compliquées à comprendre. Ainsi, en dimension supérieure ou égale à deux, il se peut que les trajectoires soient attirées par une orbite périodique (et donc "s'enroulent en spirale"). En dimension supérieure à trois, il arrive aussi que les trajectoires soient chaotiques. Même dans ces cas difficiles à traiter, la méthode de l'E.D.O. relie de manière très précise les comportements asymptotiques du système dynamique et du processus.

L'introduction se compose ensuite de cinq parties. Chacune d'elle présente les différents outils (dynamiques ou probabilistes) et thèmes composant cette thèse, qui est quant à elle constituée de deux parties indépendantes. Dans la première partie, qui comprend deux chapitres, nous nous intéressons à la première famille de diffusions dont le terme de dérive dépend de la moyenne de leur mesure d'occupation normalisée. Dans ce cas, il est possible de faire une étude complète d'un tel processus. Nous donnerons des conditions nécessaires et suffisantes pour la convergence (presque sûre et en probabilité) de la trajectoire du processus ainsi que son comportement ergodique. Ensuite, dans la seconde partie, qui comprend cette fois trois chapitres, nous commençons par situer brièvement nos travaux en décrivant les différents types de processus renforcés non markoviens connus jusqu'à présent. Nous étudions, dans le chapitre 4, les processus renforcés par leur mesure d'occupation normalisée. En particulier, nous donnons des critères de convergence de la mesure empirique. Dans le dernier chapitre, nous exhibons sur un exemple la diversité des comportements possibles. Deux annexes sont enfin ajoutées à cette étude. La première contient une étude d'un cas particulier d'un processus qui est une généralisation possible du mouvement brownien perturbé. La suivante exhibe la décomposition d'une fonction strictement convexe en dehors d'un compact. Remarquons que certaines définitions données dans cette introduction sont répétées dans le manuscrit, car les différents chapitres se veulent indépendants. Signalons également qu'à la fin de chaque chapitre se trouve une bibliographie spécifique, en plus de la bibliographie générale de l'ouvrage qui se trouve à la fin de cette introduction.

1. Notions provenant de la théorie des systèmes dynamiques

Commençons par présenter les techniques venant de la théorie des systèmes dynamiques. Dans toute cette partie, nous considérons un espace métrique (E, d).

1.1. Attracteurs. Nous avons besoin de rappeler quelques notions venant de la théorie des systèmes dynamiques, qui seront manipulées dans la suite de ce manuscrit. Définissons tout d'abord ce qu'est un flot sur cet espace métrique.

DÉFINITION 1.1. *Un flot $\Phi : \mathbb{R} \times E \to E$; $(t, x) \mapsto \Phi(t, x) = \Phi_t(x)$ est une fonction continue telle que $\Phi_0 = id$ et $\Phi_{t+s} = \Phi_t \circ \Phi_s$ pour tout $s, t \in \mathbb{R}$. Si on remplace $\mathbb{R}$ par $\mathbb{R}_+$, alors Φ est un semi-flot.*

Maintenant, nous pouvons passer aux notions topologiques relatives à un flot.

DÉFINITION 1.2. *a) Un sous-ensemble A de E est dit positivement invariant (respectivement negativement invariant, respectivement invariant) pour le flot Φ si $\Phi_t(A) \subset A$ (respectivement $A \subset \Phi_t(A)$, respectivement $\Phi_t(A) = A$) pour $t \geq 0$.*

b) Un sous-ensemble A de E est dit attractif (respectivement attracteur) pour le flot Φ si les conditions suivantes sont vérifiées :

(1) *A est non vide, compact et positivement invariant (respectivement invariant),*

(2) *A admet un voisinage $\mathcal{N} \subset E$ tel que $d(\Phi_t(x), A) \to 0$ lorsque $t \to +\infty$ uniformément en $x \in \mathcal{N}$.*

c) Le bassin d'attraction d'un attracteur A pour le flot Φ est l'ouvert de E positivement invariant comprenant tous les points x tels que $d(\Phi_t(x), A) \to 0$ lorsque $t \to +\infty$.

d) Un ensemble attractif global (respectivement attracteur global) est un ensemble attractif (respectivement attracteur) dont le bassin est l'espace tout entier E.

e) Un attracteur pour le flot Φ est dit propre s'il ne vaut ni $\emptyset$, ni E.

f) Un ensemble E est dit libre d'attracteur pour le flot Φ s'il n'admet aucun attracteur propre.

La notion d'attracteur est fondamentale. Il s'agit d'un ensemble vers lequel toutes les trajectoires aboutissent. Prenons l'exemple de l'oscillateur amorti, dont l'équation linéarisée est donnée par

$$\frac{\mathrm{d}^2\theta}{\mathrm{d}t^2} + \gamma\frac{\mathrm{d}\theta}{\mathrm{d}t} + \omega^2\theta = 0,$$

où γ est le coefficient d'amortissement et ω^2 est le carré de la pulsation. Les trajectoires solution de cette équation sont tangentes en chaque point au vecteur vitesse local dont les composantes pour un point de coordonnées $(x = \theta, y = \dot{\theta})$ sont par définition

$$\begin{cases} \frac{\mathrm{d}x}{\mathrm{d}t} = \dot{\theta} = y, \\ \frac{\mathrm{d}y}{\mathrm{d}t} = \ddot{\theta} = -\gamma\dot{\theta} - \omega^2\theta = -\gamma y - \omega^2 x. \end{cases}$$

L'origine $(\theta = \dot{\theta} = 0)$ est le seul point du plan où le champ de vecteurs est nul. Il s'agit du seul point singulier pour le système dynamique, correspondant à une solution stationnaire de l'équation du mouvement. Si $\gamma > 0$, alors cet état est stable et toute évolution aboutit à l'origine. On trouve bien que 0 est un attracteur.

1.2. Récurrence par chaîne. Ce paragraphe est basé sur le cours de Benaïm [**8**].

Présentons maintenant la notion de récurrence par chaîne. Il s'agit d'une notion dynamique introduite par Conley [**37**]. Elle généralise la notion habituelle de trajectoire, afin de tenir compte de la perturbation. Cette notion est très utile pour comprendre le comportement de processus d'approximation stochastique ou de pseudo-trajectoire asymptotique (dont la définition est donnée dans la sous-section suivante). La récurrence par chaîne est de plus fortement liée à la notion d'attracteur comme nous le verrons par la suite.

Définition 1.3. *Soit $\delta > 0$, $T > 0$. Une (δ, T)-pseudo-orbite de $a \in E$ à $b \in E$ est une suite finie de portions de trajectoires $\{\Phi_t(y_i); 0 \leq t \leq t_i\}$ qui relient "presque" a à b, (pour $0 \leq i \leq k-1$ et $t_i \geq T$) telle que :*
- $d(y_0, a) < \delta$,
- $d(\Phi_{t_i}(y_i), y_{i+1}) < \delta$,
- $y_k = b$.

Nous notons $a \to b$ s'il existe une (δ, T)-pseudo-orbite de a à b pour tout $\delta > 0$, $T > 0$.

Définition 1.4. *1) Un point a est récurrent par chaîne si $a \to a$.*

2) Si, de plus, chaque point de E est récurrent par chaîne, alors on dit que Φ est un flot récurrent par chaîne.

3) Si pour tous $a, b \in E$, on a $a \to b$, alors le flot Φ est dit transitif par chaîne.

Une façon imagée de voir qu'un point a est récurrent par chaîne est la suivante. Supposons donnée une pseudo-trajectoire qui suit le flot, mais qui peut "faire un écart" (ou sauter) pour aller suivre une autre trajectoire ε-proche dès qu'elle est restée durant le temps T sur la même trajectoire. Supposons que le flot partant de a ne revient pas en a. Alors, grâce aux "sauts" de la pseudo-trajectoire, cette dernière peut revenir en a et ce point est ainsi récurrent par chaîne.

Notons $R(\Phi)$ l'ensemble des points de E récurrents par chaîne pour le flot Φ. On peut montrer que l'ensemble $R(\Phi)$ est fermé, positivement invariant. Il est de plus invariant lorsqu'il s'agit d'un compact.

Définition 1.5. *Soit A un ensemble invariant par le flot Φ. Le flot $\Phi|_A = (\Phi_t|A)$ est le flot restreint à A.*

Définition 1.6. *Un ensemble compact invariant sur lequel Φ est récurrent (respectivement transitif) par chaîne est appelé ensemble intrinséquement récurrent par chaîne (respectivement ensemble intrinséquement transitif par chaîne).*

Voici l'ensemble des relations entre ces différentes notions. Ce résultat est dû à Bowen [**22**].

Proposition 1.7. *([**8**] proposition 5.3) Soit un ensemble $A \subset E$. Les assertions suivantes sont équivalentes :*

(1) *A est transitif par chaîne ;*
(2) *A est récurrent par chaîne et connexe ;*
(3) *A est un ensemble fermé invariant et A est un ensemble libre d'attracteur.*

Considérons par exemple le flot sur le cercle $\mathbb{S}^1$. Supposons qu'il tourne (strictement) dans le sens indirect, sauf en deux points fixes, les pôles sud S et nord N. Voir la figure ci-dessous. A priori, on ne restera pas bloqué en l'un des points N ou S. Le flot est récurrent par chaîne et le seul attracteur est l'espace tout entier.

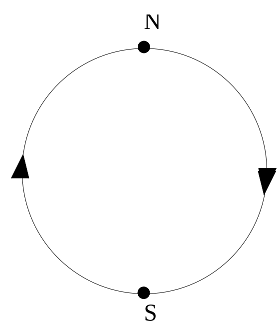

Fig. 1. Flot sur le cercle.

En revanche, si le point S est un attracteur et N est répulsif (voir la deuxième figure ci-dessous), alors l'ensemble récurrent par chaîne est maintenant réduit à $\{N, S\}$. Tous les points de $\mathbb{S}^1$ sont transitifs par chaîne, alors

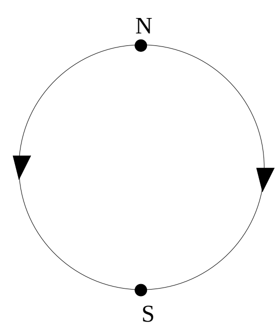

Fig. 2. Le point N est répulsif.

que le segment $[N, S]$ n'est pas intrinsèquement transitif par chaîne, bien qu'il soit composé de points transitifs par chaîne.

1.3. Pseudo-trajectoire asymptotique. Passons maintenant à la notion de pseudo-trajectoire asymptotique, qui généralise la relation de chaîne introduite par Conley. En effet, il faut ici tenir compte à la fois des perturbations et de la continuité temporelle. On considère un flot Φ sur un espace métrique (E, d). Nous renvoyons à l'article [**6**] pour plus de détails.

Définition 1.8. *Une fonction continue $\xi : \mathbb{R} \to E$ est une pseudo-trajectoire asymptotique du flot Φ si pour tout temps $T > 0$, on a*

$$\lim_{t\to\infty} \sup_{0\leq s\leq T} d(\xi_{t+s}, \Phi_s(\xi_t)) = 0.$$

Par exemple, considérons l'équation différentielle suivante (à valeurs $\mathbb{R}$ ou $\mathbb{R}^d$) :

$$\dot{\xi} = f(\xi) + g(t), \tag{1.1}$$

où $f : \mathbb{R} \to \mathbb{R}$ est lipschitzienne et $g : \mathbb{R}_+ \to \mathbb{R}$ est une fonction continue telle que $\lim\limits_{t\to\infty} g(t) = 0$. Considérons alors la solution de l'équation

$$\dot{x} = f(x) \tag{1.2}$$

et montrons que ξ est pseudo-trajectoire asymptotique du flot engendré par (1.2). Nous avons en effet l'égalité

$$\xi_{t+s} - \Phi_s(\xi_t) = \int_t^{t+s} \left(f(\xi_u) - f(\Phi_u(\xi_t))\right) \mathrm{d}u + \int_t^{t+s} g(u)\mathrm{d}u.$$

En utilisant le fait que f est lipschitzienne de constante c et le résultat de Gronwall, nous obtenons facilement

$$\sup_{0\leq s\leq T} |\xi_{t+s} - \Phi_s(\xi_t)| \leq e^{cT} \int_t^{t+T} |g(u)|\mathrm{d}u.$$

Comme $g(t)$ converge vers 0, on obtient bien le résultat attendu : $\lim\limits_{t\to\infty} \sup\limits_{0\leq s\leq T} |\xi_{t+s} - \Phi_s(\xi_t)| = 0$, donc ξ est une pseudo-trajectoire asymptotique pour le flot engendré par (1.2).

Nous pouvons donner une nouvelle caractérisation de la notion de pseudo-trajectoire asymptotique. Soit $\mathcal{C}^0(\mathbb{R}, E)$ l'ensemble des fonctions continues de $\mathbb{R}$ dans E muni de la topologie de la convergence uniforme sur les intervalles compacts. Il s'agit d'un espace métrisable. Nous définissons également le flot de translation $\theta : \mathcal{C}^0(\mathbb{R}, E) \times \mathbb{R} \to \mathcal{C}^0(\mathbb{R}, E)$ par $\theta_t(\xi)(s) = \xi(t+s)$. Le sous-ensemble $S_\Phi \subset \mathcal{C}^0(\mathbb{R}, E)$ est composé de toutes les trajectoires Φ^p où $p \in E$.

Nous sommes maintenant en mesure d'énoncer un théorème donnant une autre caractérisation des pseudo-trajectoires asymptotiques.

Théorème 1.9. *([**8**] théorème 3.2) Soit $\xi : \mathbb{R}_+ \to E$ une fonction continue dont l'image a une adhérence compacte dans E. Les deux assertions suivantes sont équivalentes :*

(1) *ξ est une pseudo-trajectoire asymptotique pour Φ,*

(2) *ξ est uniformément continue et tout point limite de $\theta_t(\xi)$ appartient à S_Φ.*

De plus, elles impliquent que la suite $(\theta_t(\xi); t \geq 0)$ est relativement compacte dans $\mathcal{C}^0(\mathbb{R}, E)$.

La notion de pseudo-trajectoire asymptotique est essentielle dans notre étude, comme nous le verrons dans la partie suivante de cette introduction. En particulier, elle fait le lien entre les processus d'approximation stochastique et les ensembles transitifs par chaîne. Elle nous sert également à établir des résultats d'ergodicité pour la première famille de processus considérée ici.

1.4. Ensemble ω-limite. Par abus de langage, nous définissons la précompacité comme suit.

DÉFINITION 1.10. *Une fonction continue $X : \mathbb{R}_+ \to E$ est précompacte si son image a une adhérence compacte dans E.*

Introduisons maintenant l'objet d'étude de ce paragraphe : l'ensemble ω-limite d'une trajectoire.

DÉFINITION 1.11. *Soit $X : \mathbb{R}_+ \to E$ une fonction continue. L'ensemble ω-limite de X est*

$$\omega(X) := \bigcap_{t \geq 0} \overline{X([t, \infty))}. \tag{1.3}$$

Dans certains cas, nous pouvons exhiber des propriétés intéressantes de ω, comme le montre le résultat suivant, issu de l'article [**6**].

THÉORÈME 1.12. *([**8**] théorème 5.7)*

(1) *Si X est une pseudo-trajectoire asymptotique précompacte, alors $\omega(X)$ est intrinséquement transitif par chaîne,*

(2) *Soit $L \subset E$ un ensemble intrinséquement transitif par chaîne et supposons que E est localement connexe. Alors il existe une pseudo-trajectoire asymptotique précompacte X telle que $\omega(X) = L$.*

Dans le cas des systèmes plans dont les points d'équilibre sont isolés, on peut aller beaucoup plus loin. En particulier, on montre que, dans ce cas, il n'existe pas de trajectoires "spiralées" tournant autour d'une orbite périodique.

THÉORÈME 1.13. *([**8**] théorème 6.12) Supposons que Φ est un flot défini sur $\mathbb{R}^2$, dont les équilibres sont isolés. Soit L un ensemble récurrent par chaîne interne. Alors, pour tout $p \in L$, nous sommes dans l'un des cas suivants :*

(1) *p est un équilibre,*

(2) *p est périodique,*

(3) *il existe une orbite chaînée cyclique* $\Gamma \subset L$ *contenant* p.

Ce théorème induit une généralisation du théorème de Poincaré-Bendixson pour les pseudo-trajectoires asymptotiques :

Corollaire 1.14. *([**8**] corollaire 6.14) Soit* Φ *un flot défini sur* $\mathbb{R}^2$, $K \subset \mathbb{R}^2$ *un sous-ensemble compact ne contenant aucun équilibre de* Φ *et* X *une pseudo-trajectoire asymptotique pour* Φ. *S'il existe une constante* $T > 0$ *telle que* $X_t \in K$ *pour tout* $t \geq T$, *alors* $\omega(X)$ *est soit une orbite périodique, soit un cylindre d'orbites périodiques.*

Bien entendu, si X est une trajectoire de Φ, alors par le théorème de Poncaré-Bendixson nous savons que l'ensemble ω-limite $\omega(X)$ est une orbite périodique.

Proposition 1.15. *([**8**] théorème 6.15) Soit* Φ *un flot défini sur un ouvert du plan et* X *une pseudo-trajectoire asymptotique de* Φ. *Supposons que pour toute région du plan* A *dont la superficie est strictement positive, la superficie de* $\Phi_t(A)$ *décroît strictement pour* $t > 0$[1]. *Alors*

(1) $\omega(X)$ *est un ensemble connexe d'équilibres, qui est nulle part dense et qui ne sépare pas le plan,*

(2) *si* Φ *a au plus un nombre dénombrable de points stationnaires, alors* $\omega(X)$ *est l'un de ces points stationnaires.*

Remarque 1.16. *Pour que les conditions de la proposition précédente soient remplies, il suffit que le flot* Φ *soit induit par un champ de vecteurs* F *tel que* $\mathrm{div}(F) < 0$.

2. Application aux algorithmes stochastiques

Cette partie est à nouveau basée sur le cours de Benaïm [**8**].

2.1. Notion probabiliste de pseudo-trajectoire asymptotique. Pour des processus, nous adaptons naturellement la définition de pseudo-trajectoire asymptotique précédente en deux notions distinctes. D'un côté, nous donnons une version presque sûre de la définition ; et de l'autre, nous considérons une version plus faible, avec une convergence en probabilité.

Définition 2.1. *a) Un processus continu* $\xi : \mathbb{R} \to E$ *est une pseudo-trajectoire asymptotique presque sûre du flot* Φ *si pour tout temps* $T > 0$, *on a*

$$\mathbb{P}\left(\lim_{t\to\infty} \sup_{0\leq s\leq T} d(\xi_{t+s}, \Phi_s(\xi_t)) = 0\right) = 1.$$

[1]Le système défini par ce flot est un système dissipatif. Pour une introduction aux systèmes dissipatifs déterministes, nous renvoyons le lecteur au livre de Bergé, Pomeau et Vidal [**17**]

b) Un processus continu $\xi : \mathbb{R} \to E$ est une pseudo-trajectoire asymptotique en probabilité du flot Φ si pour tout temps $T > 0$ et tout $\alpha > 0$, on a

$$\lim_{t\to\infty} \mathbb{P}\left(\sup_{0\leq s\leq T} d(\xi_{t+s}, \Phi_s(\xi_t)) \geq \alpha \right) = 0.$$

2.1.1. *Approximation stochastique et pseudo-trajectoire asymptotique.* Le premier résultat important est qu'un processus d'approximation stochastique est une pseudo-trajectoire asymptotique pour le flot engendré par l'E.D.O. (0.6). Plus précisément, nous avons :

PROPOSITION 2.2. *([**8**] proposition 4.4) Soit (X_n) un processus d'approximation stochastique. Rappelons qu'il vérifie l'équation suivante*

$$X_{n+1} - X_n = \frac{1}{n}\left(F(X_n) + \xi_{n+1} + r_n\right).$$

Nous supposons de plus que ξ_n est borné pour tout n et que F est une fonction lipschitzienne.
Définissons $X_t := X_n + (t-n)(X_{n+1} - X_n)$ pour $n \leq t < n+1$. Alors (X_t) est presque sûrement une pseudo-trajectoire asymptotique pour le flot Φ (induit par F) dont les courbes intégrales correspondent à celles de F.

REMARQUE 2.3. *Au lieu de considérer le processus d'approximation de la proposition ci-dessus, nous pouvons généraliser l'étude au processus*

$$X_{n+1} - X_n = \gamma_n\left(F(X_n) + \xi_{n+1} + r_n\right),$$

*avec $\gamma_n = o(1/\log n)$ et $\sum_{n\geq 1} \frac{|r_n|}{n} < \infty$ presque sûrement. Plus généralement encore, il est possible de relaxer l'hypothèse de bornitude sur la suite ξ_n en exigeant l'existence d'une constante $q \geq 2$ telle que $\sup_n \mathbb{E}|\xi_n|^q < \infty$ et $\sum_n \gamma_n^{1+q/2} < \infty$. Le même résultat est alors encore vrai (voir [**8**]).*

2.1.2. *Exemple : le recuit simulé.* Nous avons déjà vu qu'un processus d'approximation stochastique est une pseudo-trajectoire asymptotique (presque sûre). Voici un autre exemple : considérons le processus Y, solution de

$$\mathrm{d}Y_t = \varepsilon(t)\mathrm{d}B_t - Y_t\mathrm{d}t, \;\; Y_0 = 0$$

où B est un mouvement brownien standard de dimension d. Remarquons que dans ce cas précis, nous connaissons la solution explicite de Y. Montrons que Y est pseudo-trajectoire asymptotique (presque sûre ou en probabilité) du flot Φ engendré par le système dynamique

$$\dot{y} = -y.$$

Nous connaissons les solutions explicites de ces deux équations. En effet, Y est un processus d'Ornstein-Uhlenbeck général, qui s'exprime donc sous la

forme
$$Y_t = e^{-t}\int_0^t e^s \varepsilon(s)\mathrm{d}B_s.$$
Quant au flot engendré par la solution y de l'E.D.O. linéraire, il s'agit tout simplement de
$$\Phi_t(y) = ye^{-t}.$$

Nous avons alors pour tout $t, s \geq 0$
$$|Y_{t+s} - \Phi_s(Y_t)| = e^{-s}\left|\int_0^s e^u \varepsilon(t+u)\mathrm{d}B_{t+u}\right|.$$
On conclut alors trivialement que
$$\sup_{0\leq s\leq T}|Y_{t+s} - \Phi_s(Y_t)| \leq e^T \sup_{0\leq s\leq T}\left|\int_t^{t+s} \varepsilon(u)\mathrm{d}B_u\right|.$$
Par l'inégalité de Burkholder-Davis-Gundy, il est alors clair que si ε appartient à $L^2(\mathbb{R}_+)$, nous avons une pseudo-trajectoire asymptotique en probabilité. Pour qu'il s'agisse d'une pseudo-trajectoire asymptotique presque sûre, il faut en outre supposer que pour tout $c > 0$, la fonction $e^{-c/\varepsilon(t)}$ est intégrable sur $\mathbb{R}_+$ (voir [**8**], proposition 4.6). Dans le cas d'une pseudo-trajectoire asymptotique presque sûre, le processus Y converge presque sûrement vers 0 (qui est l'unique minimum global de $V(x) = |x|^2$).

Donnons maintenant un exemple de fonction ε telle que Y est une pseudo-trajectoire asymptotique en probabilité mais pas presque-sûre du flot Φ. Supposons que $\varepsilon(t) := (\log(2+t))^{-1/2}$. Supposons que Y est également une pseudo-trajectoire presque sûre. Comme 0 est l'unique minimum de la fonction V, cela impliquerait alors que Y_t converge presque sûrement vers 0 (car l'ensemble des points limites de Y est un ensemble libre d'attracteur, inclus dans les points critiques de V par le théorème 2.7). Pourtant, si on applique la loi du logarithme itéré à Y, on trouve, en posant $T(t) :=< e^{\cdot}Y_{\cdot} >_t = \int_0^t \frac{\mathrm{d}s}{\log(2+s)}e^{2s}$, que $e^{-t}\sqrt{T(t)\log\log T(t)}$ est équivalent à 1 pour t grand. Ainsi, il existe une sous-suite (t_n) telle que $t_n \to \infty$ et $Y_{t_n} \xrightarrow[n\to\infty]{p.s.} 1$. Donc Y ne converge pas vers 0 presque sûrement. Cela prouve que Y est une pseudo-trajectoire asymptotique en probabilité, mais pas presque sûre.

2.1.3. *Exemple : processus renforcé.* Nous reprenons l'exemple traité dans le chapitre 4 du manuscrit. Considérons ici un exemple simple de diffusion renforcée sur $\mathbb{R}$, appartenant à la seconde famille de processus considérés dans cette thèse. Nous travaillons avec les potentiels suivants :
$$V(x) := \frac{1}{4}x^4 + \frac{1}{5}x^2 + 1 \text{ et } W(x,y) := \cos(x)\varphi_1(y) + \sin(x)\varphi_2(y)$$
où les fonctions φ_i sont lisses et dominées par V (cela signifie qu'il existe une constante $C > 0$ telle que $|\varphi_i(y)| \leq CV(y)$). Soit (X_t, μ_t) la solution de

l'E.D.S.

$$\begin{cases} \mathrm{d}X_t = -\left(V'(X_t) + W' * \mu_t(X_t)\right)\mathrm{d}t + \mathrm{d}B_t \\ \mathrm{d}\mu_t = (\delta_{X_t} - \mu_t)\frac{\mathrm{d}t}{r+t} \end{cases}$$

Comme les variables du potentiel d'interaction W sont séparées, on peut réduire cette E.D.S. de dimension infinie en une équation en dimension trois, avec les nouvelles variables $Y_t := \int_{\mathbb{R}} \varphi_1(y)\mu_t(\mathrm{d}y)$ et $Z_t := \int_{\mathbb{R}} \varphi_2(y)\mu_t(\mathrm{d}y)$, soit :

$$\begin{cases} \mathrm{d}X_t = \mathrm{d}B_t - \left[V'(X_t) - Y_t \sin(X_t) + Z_t \cos(X_t)\right]\mathrm{d}t, \ \ X_0 = x \\ \mathrm{d}Y_t = \left[\varphi_1(X_t) - Y_t\right](r+t)^{-1}\mathrm{d}t, \ \ Y_0 = y \\ \mathrm{d}Z_t = \left[\varphi_2(X_t) - Z_t\right](r+t)^{-1}\mathrm{d}t, \ \ Z_0 = z \end{cases}$$

Nous espérons que le couple de variables (Y_t, Z_t) converge presque sûrement vers une limite (disons déterministe pour fixer les idées) (α, β). Ainsi, le processus $(X_t, t \geq 0)$ doit avoir un comportement asymptotique proche de celui d'une diffusion réelle classique, avec pour mesure ergodique :

$$\mu_{\alpha,\beta}(\mathrm{d}x) := Z_0(\alpha,\beta)^{-1} \exp\left\{-2V(x) - 2\alpha\cos(x) - 2\beta\sin(x)\right\}\mathrm{d}x$$

où $Z_0(\alpha,\beta) := \int \exp\left\{-2V(x) - 2\alpha\cos(x) - 2\beta\sin(x)\right\}\mathrm{d}x$ est la constante de normalisation.

Le théorème ergodique ponctuel classique implique alors que (Y_t, Z_t) converge presque sûrement vers $(\int_{\mathbb{R}} \varphi_1(y)\mu_{\alpha,\beta}(\mathrm{d}y), \int_{\mathbb{R}} \varphi_2(y)\mu_{\alpha,\beta}(\mathrm{d}y))$. Les constantes (α, β) sont donc des points fixes de l'application

$$(\alpha,\beta) \mapsto \frac{1}{Z_0(\alpha,\beta)} \begin{pmatrix} Z_1(\alpha,\beta) \\ Z_2(\alpha,\beta) \end{pmatrix}$$

avec les notations $Z_i(\alpha,\beta) := \int_{\mathbb{R}} \varphi_i(y) e^{\{-2V(y) - 2\alpha\cos(y) - 2\beta\sin(y)\}}\mathrm{d}y$. Dans le but d'étudier ces points fixes, considérons l'E.D.O. suivante :

$$\begin{cases} \dot{\alpha_t} = -\alpha_t + Z_1(\alpha_t, \beta_t)/Z_0(\alpha_t, \beta_t) \\ \dot{\beta_t} = -\beta_t + Z_2(\alpha_t, \beta_t)/Z_0(\alpha_t, \beta_t) \end{cases} \tag{2.1}$$

Si nous retournons maintenant à l'étude du processus renforcé $(X_t, t \geq 0)$ et oublions l'espace d'un instant l'hypothèse (éventuellement fausse) que (Y_t, Z_t) converge nécessairement vers une limite déterministe, l'E.D.O. (2.1) est tout de même fort utile pour étudier le comportement asymptotique de (Y_t, Z_t). Grâce aux techniques développées dans le chapitre 4 (qui suivent les idées introduites dans le papier [**10**]), on montre que (Y_{e^t}, Z_{e^t}) est une *pseudo-trajectoire asymptotique presque sûre* du système dynamique non linéaire (2.1), soit presque sûrement

$$\lim_{t\to+\infty} \sup_{0\leq s\leq T} \left(\left|Y(e^{t+s}) - \alpha_s(Y(e^t))\right| + \left|Z(e^{t+s}) - \beta_s(Z(e^t))\right|\right) = 0$$

pour $T > 0$.

Comme application numérique (voir la figure ci-dessous), nous choisissons $\varphi_1(y) := 1.3y^3$ et $\varphi_2(y) := -1.3y$. Cela revient à considérer le processus de dimension trois suivant :

$$\begin{cases} \mathrm{d}X_t = \mathrm{d}B_t - \left[2X_t(1+2X_t^2) - Y_t\sin(X_t) + Z_t\cos(X_t)\right]\mathrm{d}t, \;\; X_0 = x \\ \mathrm{d}Y_t = \left[1.3X_t^3 - Y_t\right](r+t)^{-1}\mathrm{d}t, \;\; Y_0 = y \\ \mathrm{d}Z_t = -\left[1.3X_t + Z_t\right](r+t)^{-1}\mathrm{d}t, \;\; Z_0 = z \end{cases}$$

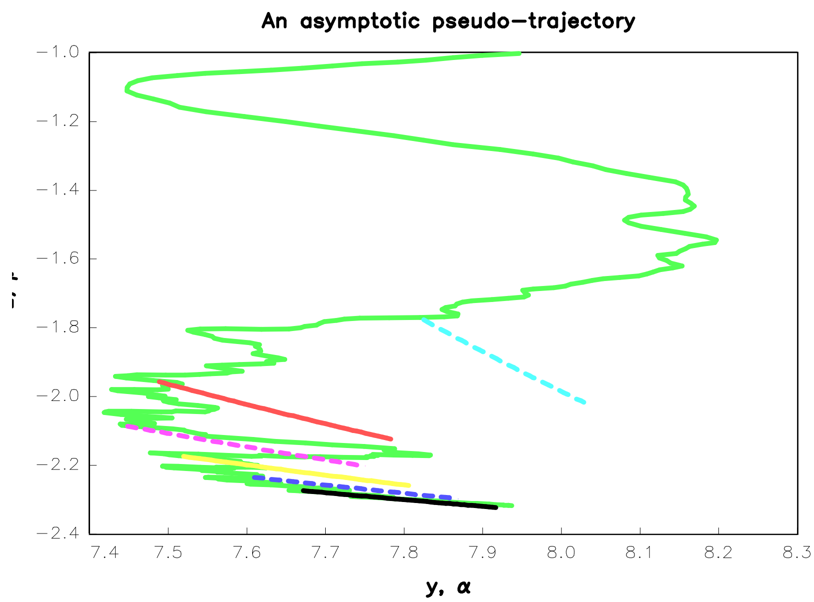

FIG. 3. Une trajectoire de (Y_{e^t}, Z_{e^t}) en vert et des orbites du flot.

2.2. L'ensemble limite "aléatoire" d'une pseudo-trajectoire asymptotique. Ce qui nous intéresse ici est essentiellement de décrire le comportement asymptotique de certains processus à mémoire longue. Nous cherchons en particulier à décrire la mesure d'occupation normalisée μ_t. Lorsque nous parlons d'objets probabilistes, nous notons $L(\mu_t)$ (ou L) l'ensemble ω-limite du processus μ_t (et non de $\Phi_t(\mu)$, qui correspond à la notation $\omega(\mu)$) afin de marquer la différence entre L qui est aléatoire (et dépend de toute la trajectoire) et ω qui est déterministe.

Nous sommes en mesure d'énoncer un résultat essentiel, qui décrit plus précisément l'ensemble limite d'un processus d'approximation stochastique. Il s'agit en fait d'un corollaire de la proposition 2.2.

Théorème 2.4. *Soit $X = (X_n; n \geq 0)$ un processus d'approximation stochastique de bruit ξ_n borné et dont la fonction F est lipschitzienne. Alors, l'ensemble limite $L(X)$ est presque sûrement transitif par chaîne. Il est donc invariant, connexe et n'admet pas d'attracteur propre.*

Théorème 2.5. *([**8**] théorème 7.3) Soit $X = (X_n; n \geq 0)$ un processus d'approximation stochastique dont la fonction F est lipschitzienne. Considérons un attracteur A pour le flot associé à X. Alors nous sommes dans l'une des deux situations suivantes :*

(1) *soit il existe un temps t pour lequel $(X_{t+s}; s \geq 0)$ évite presque sûrement tout voisinage de A ;*
(2) *soit $\mathbb{P}(L(X) \subset A) > 0$.*

Nous avons donc un résultat de convergence vers un attracteur. Obtenir des résultats de non convergence est singulièrement plus difficile. En effet, il faut alors exiger de l'instabilité linéaire, c'est-à-dire la partie réelle d'au moins l'une des valeurs propres de dF est strictement positive.

Théorème 2.6. *([**8**] théorème 9.1) Soit $X = (X_n; n \geq 0)$ un processus d'approximation stochastique sur une variété E compacte, dont le bruit ξ_n est borné pour tout n. Soit Γ un équilibre linérairement instable, ou une orbite périodique linérairement instable, pour le flot induit par F tel qu'il existe un voisinage $\mathcal{N}(\Gamma)$ de Γ et $b > 0$ tels que pour tout vecteur unité $v \in \mathbb{R}^d$, on a*

$$\mathbb{E}[(\xi_{n+1}, v)^+ | \mathcal{F}_n] \geq b \mathbb{1}_{X_n \in \mathcal{N}(\Gamma)}.$$

Supposons de plus qu'il existe une constante $\frac{1}{2} < \alpha \leq 1$ telle que la fonction F est de classe $\mathcal{C}^{1+\alpha}$. Alors

$$\mathbb{P}\left(\lim_{n\to\infty} d(X_n, \Gamma) = 0\right) = 0.$$

Reprenons ici l'exemple du cercle $\mathbb{S}^1$, en supposant le pôle S est attractif, alors que N est répulsif. Le flot a alors cinq ensembles connexes, fermés et invariants, qui sont $\mathbb{S}^1, \{N\}, \{S\}$ et les deux demi-cercles. Les ensembles $\{N\}$ et $\{S\}$ sont les seuls ensembles intrinséquement transitifs par chaîne. Ce sont donc les deux seuls ensembles limites possibles d'un processus d'approximation stochastique. Si le point $\{N\}$ est linéairement instable et s'il y a suffisamment de bruit, alors on ne va pas converger vers ce point, comme le prouve le théorème 2.6.

2.2.1. *Mesure d'occupation et pseudo-trajectoire asymptotique.* Dans le cas d'une pseudo-trajectoire asymptotique en probabilité, nous pouvons également décrire le comportement asymptotique de la mesure d'occupation normalisée du processus. Ce travail a été mené dans l'article [**11**].

Théorème 2.7. *([**8**] théorème 10.1) Soit X une pseudo-trajectoire asymptotique en probabilité pour le flot Φ. Alors, presque sûrement, l'ensemble des* *points limites de la mesure d'occupation normalisée* $\left\{\mu_t = \frac{1}{t}\int_0^t \delta_{X_s}\mathrm{d}s\right\}$ *est un sous-ensemble des mesures invariantes par* Φ*.*

Prenons par exemple le cas du recuit simulé, traité précédemment, où nous considérons le processus Y, solution de

$$\mathrm{d}Y_t = \varepsilon(t)\mathrm{d}B_t - Y_t\mathrm{d}t,\ \ Y_0 = 0.$$

Nous avons vu dans ce cas simplifié que Y est une pseudo-trajectoire en probabilité (au moins) du flot engendré par l'E.D.O. $\dot{y} = -y$. Dans ce cas, le résultat précédent nous dit que μ_t converge (au sens de la convergence faible des mesures) vers δ_0 presque sûrement. Un cas plus intéressant est de considérer la solution de

$$\mathrm{d}Y_t = \varepsilon(t)\mathrm{d}B_t - \nabla V(Y_t)\mathrm{d}t,\ \ Y_0 = 0,\ \text{ et }\ \dot{y} = -\nabla V(y)$$

où $V = W + \chi$ a un nombre fini de points critiques et est telle que W est strictement convexe et $\nabla\chi$ est lipschitzienne (de constante 1 pour simplifier). Dans ce cas, si nous notons $(c_i)_{1\leq i\leq N}$ les points critiques de V, alors le théorème dit que μ_t converge faiblement vers une combinaison convexe de mesures de Dirac : $\sum\limits_{1\leq i\leq N} a_i\delta_{c_i}$ (avec $a_i \geq 0$). Un exemple est étudié plus en détails dans le premier chapitre de cet ouvrage.

2.2.2. *Fonction de Lyapunov.* Dans la théorie des systèmes dynamiques, il existe une notion très importante, appelée fonction de Lyapunov. Lorsqu'on parvient à exhiber une telle fonction, l'étude se simplifie considérablement. Tout d'abord, définissons ce qu'est cette fonction.

Définition 2.8. *Une fonction de Lyapunov pour le flot* Φ *par rapport à l'ensemble compact invariant* $\Lambda \subset E$ *est une fonction continue* $V : E \to \mathbb{R}$ *qui est constante le long des trajectoires de* Λ *(i.e. pour tout* $x \in \Lambda, t \geq 0$*,* $V(\Phi_t(x)) = V(x)$*) et strictement décroissante ailleurs (i.e. pour tout* $x \in E\backslash\Lambda, t > 0$*,* $V(\Phi_t(x)) < V(x)$*).*

Remarque 2.9. *Lorsque le flot* Φ *est associé à l'équation* $\dot{x} = -\nabla V(x)$*, alors* V *est une fonction de Lyapunov par rapport à l'ensemble des points critiques de* V*. Les flots de gradient sont plus faciles à étudier que les flots généraux.*

L'existence d'une fonction de Lyapunov pour le flot associé à F implique la convergence du processus d'approximation stochastique X vers un ensemble sur lequel cette fonction est constante.

Proposition 2.10. *(*[**8**] *proposition 6.4) Supposons que* V *est une fonction de Lyapunov sur un ensemble* Λ *tel que* $V(\Lambda)$ *est d'intérieur vide. Alors chaque ensemble* L *intrinsèquement transitif par chaîne est contenu dans* Λ *et* V *restreinte à* L *est constante.*

3. Quelques inégalités fonctionnelles

La méthode de l'E.D.O. que nous présentons ici a déjà été utilisée dans l'étude de diffusions renforcées par Benaïm, Ledoux et Raimond [**10**], mais uniquement pour des processus à valeurs dans un espace compact. Nous généralisons dans les chapitres 4 et 5 toute cette étude à des espaces non compacts ($\mathbb{R}^d$ ou plus généralement une variété riemannienne lisse, complète, connexe, sans bord et à courbure de Ricci minorée). Pour ce faire, nous avons besoin d'utiliser certaines inégalités fonctionnelles et en particulier la notion d'ultracontractivité.

Deux inégalités fonctionnelles sont en effet essentielles dans ce travail : l'inégalité de Sobolev logarithmique et l'ultracontractivité. L'inégalité de Sobolev logarithmique implique en particulier l'existence d'un trou spectral et cette inégalité sera essentielle dans le chapitre 2 pour prouver la convergence de l'algorithme du recuit simulé. L'ultracontractivité quant à elle signifie que, étant donné le semi-groupe markovien $(P_t, t \geq 0)$ de probabilité invariante μ, P_t est un opérateur borné de $L^2(\mu)$ dans $L^\infty(\mu)$ pour tout $t \geq 0$. Ici, nous aurons besoin d'une version uniforme de cette notion.

3.1. Inégalité de Sobolev logarithmique. Une première notion importante est la fonctionnelle d'*entropie*, définie comme suit. Soit μ une mesure sur un espace polonais E. Pour toute fonction $f \in L^1(\mu)$, on pose

(3.1)
$$\mathrm{Ent}_\mu(f) := \mathbb{E}_\mu\left(f \log(f/\mathbb{E}_\mu f)\right) = \int f \log f \mathrm{d}\mu - \left(\int f \mathrm{d}\mu\right) \log\left(\int f \mathrm{d}\mu\right).$$

Considérons un semi-groupe markovien $(P_t, t \geq 0)$, de générateur infinitésimal L et de mesure invariante μ. Nous définissons alors la fonctionnelle d'*énergie* pour toute fonction $f \in L^2(\mu)$ par :

$$\mathcal{E}_\mu(f) := -\mathbb{E}_\mu(fLf) \tag{3.2}$$

DÉFINITION 3.1. *Soit L un générateur infinitésimal de mesure invariante μ. On dit que μ satisfait une inégalité de Sobolev logarithmique de constante $c > 0$ si pour toute fonction $f \in L^2(\mu)$, on a*

$$\mathrm{Ent}_\mu(f^2) \leq c\mathcal{E}_\mu(f).$$

Nous avons besoin d'obtenir de telles inégalités, mais uniformes. Comme nous ne considérons que des diffusions sur $\mathbb{R}^d$, dont le terme de dérive s'exprime essentiellement par le gradient d'une fonction W, nous avons que $\mathcal{E}_\mu(f) = \int |\nabla f|^2 \mathrm{d}\mu$. Dans notre cas, concernant l'algorithme du recuit simulé au chapitre 2, cela signifie que nous considérons une mesure invariante Π_t d'un processus de Markov (inhomogène) et nous cherchons les constantes $c(t)$ donnant une décroissance de l'énergie libre associée à notre processus. Nous obtiendrons ainsi un critère de convergence en probabilité du processus Z, correspondant à la première famille étudiée dans cette thèse.

L'inégalité de Sobolev logarithmique est également importante car elle implique l'exitence d'un trou spectral.

PROPOSITION 3.2. *(Rothaus [**99**]) Soit un semi-groupe markovien* $(P_t, t \geq 0)$*, de générateur infinitésimal* L *et de mesure invariante* μ*. Si* μ *satisfait une inégalité de Sobolev logarithmique de constante* $c > 0$*, alors elle vérifie l'inégalité de trou spectral suivante : pour tout* $f \in L^2(\mu)$*, on a*

$$||P_t f - \mathbb{E}_\mu(f)||_{L^2(\mu)} \leq e^{-ct}||f - \mathbb{E}_\mu(f)||_{L^2(\mu)}.$$

Cette dernière inégalité est équivalente à l'inégalité de Poincaré :

$$\mathbb{E}_\mu(f^2) - (\mathbb{E}\mu f)^2 \leq \frac{1}{c}\mathcal{E}_\mu(f).$$

Pour obtenir une inégalité de Sobolev logarithmique, on peut utiliser le critère de Bakry et Emery [**3**]. Nous considérons toujours un semi-groupe markovien $(P_t, t \geq 0)$, de générateur infinitésimal L et de mesure invariante μ. Pour $f, g \in L^2(\mu)$, nous définissons l'*opérateur carré du champ* associé à L par :

$$\Gamma(f,g) := \frac{1}{2}\left(L(fg) - fLg - gLf\right). \tag{3.3}$$

On appelle *opérateur* Γ_2 la forme bilinéaire symétrique suivante :

$$\Gamma_2(f,g) := \frac{1}{2}\left(L\Gamma(f,g) - \Gamma(Lf,g) - \Gamma(f,Lg)\right). \tag{3.4}$$

DÉFINITION 3.3. *Nous disons que l'opérateur* L *vérifie une inégalité de courbure-dimension* $CD(\rho, m)$*, avec* $\rho \in \mathbb{R}$ *et* $m \in [1, \infty]$*, si pour toute fonction* $f \in L^2(\mu)$*, on a*

$$\Gamma_2(f,f) \geq \rho\Gamma(f,f) + \frac{1}{m}(Lf)^2.$$

L'inégalité en dimension infinie $\Gamma_2(f,f) \geq \rho\Gamma(f,f)$ *est appelée critère de courbure ou* Γ_2*.*

THÉORÈME 3.4. *(Bakry-Emery [**3**]) Supposons que le générateur infinitésimal* L *satisfait le critère de courbure* $CD(\rho, \infty)$ *avec* $\rho > 0$*. Alors la mesure* μ*, qui est supposée invariante par* L*, satisfait une inégalité de Sobolev logarithmique de constante* $2/\rho$*.*

Remarquons que si $L = \Delta - (\nabla V, \nabla)$, alors $\Gamma(f,f) = |\nabla f|^2$ et $\Gamma_2(f,f) = |\nabla^2 f|^2 + (\nabla f, \nabla^2 V \nabla f)$, donc la condition V est strictement uniformément convexe suffit à obtenir une inégalité de Sobolev logarithmique pour la mesure invariante e^{-V}.

Nous nous servirons de ce critère afin de prouver l'existence d'un trou spectral (uniforme) pour le générateur infinitésimal correspondant à la diffusion renforcée X. Cela est fait au chapitre 4.

3.2. Ultracontractivité. Une deuxième notion fonctionnelle essentielle est l'ultracontractivité. La définition usuelle est la suivante. Un semi-groupe markovien $(P_t, t \geq 0)$ est dit ultracontractif de $L^2(\mu)$ dans $L^\infty(\mu)$ si P_t est un opérateur borné de $L^2(\mu)$ dans $L^\infty(\mu)$ pour tout t. Cela signifie qu'il existe une constante $c(t)$ telle que

$$||P_t f||_\infty \leq c(t)||f||_2. \tag{3.5}$$

Un critère pour obtenir de l'ultracontractivité est donné par Röckner et Wang [**97**] :

Proposition 3.5. *Soit $(P_t, t \geq 0)$ un semi-groupe de Markov, de générateur infinitésimal $A := \frac{1}{2}\Delta - (\nabla U, \nabla)$, où $\nabla^2 U \geq -K$. Supposons qu'il existe une fonction $\chi : \mathbb{R}_+ \to \mathbb{R}_+^*$ continue et strictement croissante telle que :*

(1) $\lim\limits_{r\to\infty} \frac{\chi(r)}{r} = \infty$,

(2) *l'application $g_\chi(r) := r\chi(m \log r)$ est convexe sur $[1, \infty)$ pour tout $m > 0$,*

(3) *il existe $b > 0$ tel que $A|x|^2 \leq b - \chi(|x|^2)$.*

Alors P_t admet une unique probabilité invariante. Si de plus $\int_2^\infty \frac{1}{\chi(m \log r)} \frac{\mathrm{d}r}{r} < \infty$ pour $m > 0$, alors P_t est ultracontractif. Si en outre, on a $\chi(r) = \chi r^\delta$ avec $\chi > 0$ et $\delta > 1$, alors il existe une constante $c(b, \chi) > 0$ telle que pour tout $0 < t \leq 1$, on a

$$||P_t||_{2\to\infty} \leq \exp\{ct^{-\delta/(\delta-1)}\}.$$

Dans notre cas, nous avons une nouvelle fois besoin d'uniformité.

Définition 3.6. *Soit $\varepsilon > 0$. Soit $(P_t^\lambda, t \geq 0, \lambda \in \Lambda)$ une famille de semi-groupes de Markov. Supposons que pour tout $\lambda \in \Lambda$, le semi-groupe $(P_t^\lambda, t \geq 0)$ admet une unique probabilité invariante μ_λ. Nous disons que la famille $(P_t^\lambda, t \geq 0, \lambda \in \Lambda)$ est uniformément ultracontractive si les opérateurs P_t^λ sont uniformément bornés de $L^2(\mu_\lambda)$ dans $L^\infty(\mu_\lambda)$ for tout $0 < t \leq \varepsilon$: il existe une constante $c > 0$, indépendante de λ, telle que*

$$||P_t^\lambda||_{2\to\infty} \leq c(t), \ \ \forall 0 < t \leq \varepsilon, \ \ \forall \lambda \in \Lambda, \tag{3.6}$$

avec la notation $||P_t^\lambda||_{2\to\infty} := \sup\limits_{f \in L^2(\mu_\lambda)\setminus\{0\}} \frac{||P_t^\lambda f||_\infty}{||f||_2}$.

L'uniforme ultracontractivité sera utilisée dans le chapitre 4 de la thèse, afin de contrôler le "noyau fondamental" (il s'agit de l'inverse du générateur infinitésimal) de la diffusion renforcée X. Cela nous permettra de montrer que la mesure d'occupation normalisée μ_t, correspondant à la seconde famille de processus, est une pseudo-trajectoire asymptotique d'un certain flot déterministe.

4. Diffusions renforcées par la moyenne de la mesure d'occupation normalisée

Cette partie correspond à la première partie, soit les deux premiers chapitres, de cette thèse. Il s'agit d'un travail effectué en commun avec Sébastien Chambeu (Université d'Orsay).

Soit un potentiel $W : \mathbb{R}^d \to \mathbb{R}$, suffisamment lisse (de classe $\mathcal{C}^2$) et une fonction "mémoire" $g : \mathbb{R}_+ \to \mathbb{R}_+$ strictement positive, de primitive G. Nous supposons que W est strictement convexe en dehors d'un compact et qu'elle admet un nombre fini de points critiques. Notons $Min = \{m_1, m_2, \cdots, m_n\}$ ses minima locaux et $Max = \{M_1, M_2, \cdots, M_p\}$ l'ensemble de ses points instables, soit ses maxima locaux et points selles. De plus pour chaque i, pour tout $\xi \in \mathbb{R}^d$, nous supposons $(\nabla^2 W(m_i)\xi, \xi) > 0$ et pour tout M_i, $\nabla^2 W$ admet une valeur propre strictement négative. Nous nous intéressons au processus à mémoire longue Z, solution de l'équation

$$\left\{ \begin{array}{l} \mathrm{d}Z_t = \mathrm{d}B_t - g(t)\nabla W(Z_t - \overline{\mu}_t)\mathrm{d}t \\ Z_0 = z \end{array} \right. \tag{4.1}$$

où B est un mouvement brownien standard et nous rappelons que $\overline{\mu}_t$ est la moyenne de la mesure empirique du processus Z, définie par

$$\overline{\mu}_t = \frac{1}{r+t}\left(r\bar{\mu} + \int_0^t Z_s \mathrm{d}s \right), \ \ \overline{\mu}_0 = \bar{\mu}. \tag{4.2}$$

Ici, μ est une probabilité initiale (donnée) sur $\mathbb{R}^d$, $\bar{\mu}$ représente la moyenne de μ et $r > 0$ est un poids initial.

Nous nous intéressons au comportement asymptotique d'un tel processus. Le premier réflexe naturel est de considérer le processus "centré" suivant :

$$Y_t := Z_t - \overline{\mu}_t. \tag{4.3}$$

Il est aisé de vérifier que $(Y_t, t \geq 0)$ satisfait l'équation

$$\left\{ \begin{array}{l} \mathrm{d}Y_t = \mathrm{d}B_t - g(t)\nabla W(Y_t)\mathrm{d}t - Y_t\frac{\mathrm{d}t}{r+t}; \\ \ \ Y_0 = z - \overline{\mu}. \end{array} \right. \tag{4.4}$$

Nous remarquons de plus que nous pouvons exprimer $\overline{\mu}_t$ et X_t en fonction de Y_t uniquement, car

$$\mathrm{d}\overline{\mu}_t = Y_t \frac{\mathrm{d}t}{r+t}.$$

L'idée est alors d'étudier le processus markovien inhomogène Y, car ce dernier est plus simple que Z qui n'est pas markovien.

4.1. Cas quadratique. Bien entendu, dans le cas quadratique $W(x) = (x, cx)$ (où c est une matrice symétrique définie positive), nous avons une solution explicite de l'équation et nous connaissons donc l'écriture de Z, $\overline{\mu}_t$

et Y en fonction du mouvement brownien uniquement :

$$\begin{aligned}
Z_t &= z + rc(\overline{\mu} - z)F(t) + \int_0^t \Big[1 - (r+s)ce^{cG(s)}(F(t) - F(s))\Big] \mathrm{d}B_s, \\
\overline{\mu}_t &= z + r\,(\overline{\mu} - z)\left(\frac{1}{r+t}e^{-cG(t)} + cF(t)\right) \\
&+ \int_0^t \left[1 - (r+s)ce^{cG(s)}\left(F(t) - F(s) + \frac{1}{r+t}\frac{e^{-cG(t)}}{c}\right)\right] \mathrm{d}B_s, \\
Y_t &= \frac{1}{r+t}e^{-cG(t)}\left(\int_0^t (r+s)e^{cG(s)}\mathrm{d}B_s + r(z - \overline{\mu})\right).
\end{aligned}$$

Cela nous permet d'obtenir tous les comportements de Z en fonction de la fonction g. Ainsi, nous avons une certaine intuition sur le comportement asymptotique de Z avec des potentiels plus complexes.

4.2. Comportement asymptotique.

4.2.1. *Ergodicité (ponctuelle).* Tout d'abord, il est possible de montrer que le processus Y vérifie le théorème ergodique ponctuel, ce qui s'exprime de la façon suivante. (Pour fixer les idées, on peut supposer pour simplifier que les points critiques de W ne sont pas dégénérés.)

THÉORÈME 4.1. *Il existe des variables $a_i > 0$ et $b_i \geq 0$ telles que*

$$\frac{1}{t}\int_0^t \delta_{Y_s}\mathrm{d}s \underset{t\to\infty}{\to} \sum_{i=1}^{n} a_i\delta_{m_i} + \sum_{i=1}^{p} b_i\delta_{M_i} \ \ p.s.,$$

au sens de la convergence faible des mesures : pour toute fonction f continue bornée sur $\mathbb{R}^d$, on a presque sûrement $\lim\limits_{t\to\infty}\frac{1}{t}\int_0^t f(Y_s)\mathrm{d}s = \sum\limits_{1\leq i\leq n} a_i f(m_i) + \sum\limits_{1\leq i\leq p} b_i f(M_i)$.

Pour démontrer un tel résultat, la théorie des pseudo-trajectoires asymptotiques est essentielle ! En effet, l'idée est de prouver que Y est une pseudo-trajectoire asymptotique en probabilité d'un certain flot de gradient Φ. Ensuite, nous savons par [**11**] que la mesure d'occupation normalisée limite est un sous-ensemble des probabilités invariantes pour Φ. En fait, il s'agit de combinaison linéaire de mesures de Dirac en les points critiques de W. De plus, nous conjecturons que $b_i \equiv 0$. Maintenant, nous nous demandons si ce théorème reste encore valable pour le processus (plus complexe) Z ? La réponse est donnée par les deux résultats suivants.

PROPOSITION 4.2. *Le processus $\overline{\mu}_t$ converge presque sûrement si et seulement si $\frac{1}{t}\int_0^t Y_s\mathrm{d}s = o\left((\log t)^{-1}\right)$.*
En particulier, si $Y_t \xrightarrow[t\to\infty]{p.s.} 0$ à une vitesse polynomiale alors le processus $\overline{\mu}_t$ converge presque sûrement.

Corollaire 4.3. *Le processus Z satisfait le théorème ergodique ponctuel si et seulement si $\overline{\mu}_t$ converge presque sûrement.*

Remarque 4.4. *Si par exemple g est polynomiale ($g(t) = t^\beta$ avec $\beta > 0$) ou exponentielle et si W a un unique minimum en 0, alors $\overline{\mu}_t$ converge par la proposition précédente et ainsi Z satisfait le théorème ergodique ponctuel. En revanche, si W n'admet pas de minimum en 0, alors Y ne converge pas p.s. vers 0, ce qui implique une divergence de $\overline{\mu}_t$ et donc Z n'est pas ergodique.*

4.2.2. *Convergence presque sûre des trajectoires.* Nous pouvons également en dire plus sur la trajectoire de Z, comme le prouve le résultat suivant.

Théorème 4.5. *Supposons que $g(t)^{-1}\log G(t) = o((\log t)^{-2})$. Alors la variable limite Y_∞ est un des points critiques de W. Pour $Y_\infty = 0$, posons $Z_\infty := \overline{\mu}_0 + \int_0^\infty Y_s \frac{\mathrm{d}s}{r+s}$ (lorsqu'elle est finie). Nous sommes alors dans l'un des cas suivants :*

(1) *Si 0 est l'unique minimum de W, alors*

$$\mathbb{P}(Z_t \to Z_\infty) = \mathbb{P}(\bar{\mu}_t \to Z_\infty) = 1;$$

(2) *Si 0 est un minimum local de W et s'il existe d'autres minima locaux, alors*

$$\mathbb{P}(Z_t \to Z_\infty) + \mathbb{P}(|Z_t| \to \infty) = 1$$

et

$$1 > \mathbb{P}(Z_t \to Z_\infty) > 0, \;\; 1 > \mathbb{P}(\bar{\mu}_t \to Z_\infty) > 0;$$

(3) *Si 0 n'est pas un minimum local de W, alors*

$$\mathbb{P}(|Z_t| \to \infty) = \mathbb{P}(\bar{\mu}_t \to \infty) = 1.$$

Remarquons de plus que sur $\{Y_\infty \neq 0\}$, on a

$$\frac{X_t}{\log t} \xrightarrow{p.s.} Y_\infty.$$

4.3. Convergence en loi et recuit simulé. Bien entendu, lorsque $g(t)^{-1}\log G(t)$ ne converge pas vers 0, nous ne pouvons en aucun cas espérer une convergence presque sûre. En revanche, une convergence dans un sens plus faible (en probabilité ou en loi) est possible. Pour ce faire, nous utilisons l'inégalité de Sobolev logarithmique et nous inspirons des techniques fonctionnelles de recuit simulé. Expliquons brièvement en quoi consiste le recuit simulé.

Le problème lié au recuit simulé consiste en la localisation des minima globaux d'un certain potentiel W. Ce problème est issu de la chimie. Notons ε la température et considérons la mesure de Gibbs suivante sur un espace $\mathcal{U}$:

$$\Pi(\mathrm{d}x) = \frac{1}{Z} e^{-\frac{1}{k\varepsilon}W(x)}\mathrm{d}x, \tag{4.5}$$

où k est la constante de Boltzmann et W est le potentiel (énergie) du système. Z est la constante de normalisation. Nous cherchons alors l'ensemble $\underline{\mathcal{U}} = \{x \in \mathcal{U}; W(x) = \inf(W(y); y \in \mathcal{U})\}$. Le recuit simulé, introduit par Kirkpatrick, Gelatt et Vecchi [**71**] et indépendamment par Cerny [**30**], est conçu pour résoudre des problèmes d'optimisation globale comme celuici. Le recuit simulé que nous présentons ici est l'interprétation mathématique des remarques suivantes. Lorsque la température ε tend vers 0, alors la distribution de Gibbs (4.5) est concentrée sur les minima globaux $\underline{\mathcal{U}}$ de W. Donc les états de basse énergie du système (proches de $\underline{\mathcal{U}}$) contrôlent le comportement du système en basse température. Expérimentalement, on atteint ces états en mélangeant une substance puis en la refroidissant lentement, en faisant attention aux températures qui figent le système. Certains cristaux sont obtenus par cette méthode. Si on refroidit le système trop rapidement, alors on peut le figer en un état "métastable", c'est-à-dire s'arrêter en un minimum local (non global). Si en revanche, on refroidit le système trop lentement, alors le système va se figer en un minimum global, mais très lentement. La température optimale de refroidissement (la température de recuit) est obtenue par la compétition entre ces deux phénomènes. L'idée est alors "d'imiter mathématiquement" cette expérience, en supposant que la température $\varepsilon(t)$ dépend du temps et tend vers 0 à l'infini.

Dans le cas discret, le premier probabiliste à traiter ce problème, est Hajek [**57**]. Il génère une chaîne de Markov inhomogène (sur un espace d'état fini) dont la probabilité invariante est la mesure de Gibbs. En temps continu, cela revient à étudier la diffusion markovienne X^ε dans $\mathbb{R}^d$, dont le terme de dérive est le gradient du potentiel W et dont le coefficient de diffusion décroît vers 0, soit

$$\mathrm{d}X_t^\varepsilon = \varepsilon(t)\mathrm{d}B_t - \nabla W(X^\varepsilon)\mathrm{d}t, \quad X_0^\varepsilon = x. \tag{4.6}$$

Ici, ε^2 représente le schéma de décroissance de la température. Sous certaines conditions, le système aura tendance à se concentrer au voisinage des minima globaux du potentiel. Néanmoins, lorsque la décroissance de température est trop rapide, il pourra se geler en des minima locaux. Il faut donc trouver les bons taux de refroidissement pour éviter cette dernière situation.

Il faut voir X^ε comme une perturbation de la trajectoire X^0 du système dynamique $\frac{\mathrm{d}X_t^0}{\mathrm{d}t} = -\nabla W(X_t^0)$. Chiang, Hwang et Sheu [**5**] ont étudié la vitesse de convergence de la quantité $\mathbb{P}_x(X_t^\varepsilon \in \cdot)$ via la théorie des grandes déviations pour la densité des fonctions de transition de X^ε. Il s'avère qu'il existe un nombre $c > 0$ tel que pour tout $k > c$, une décroissance de la température en $k(\log t)^{-1}$ entraîne le système vers les minima globaux. Cette constante c est liée au comportement asymptotique du trou spectral (de la seconde valeur propre) de certains opérateurs, qui ne sont autres que les générateurs infinitésimaux des diffusions X^ε.

Notons que Chiang, Hwang et Sheu [**5**] font partie des premiers probabilistes à avoir prouvé la convergence de l'algorithme du recuit simulé pour $\varepsilon(t)^2 = k/\log t$ avec $k > c$ suffisamment grand. Ensuite, Royer [**100**] parvient à trouver la constante c optimale (si $k \leq c$, alors il prouve que l'algorithme ne converge pas). Enfin, Miclo [**88**] utilise et généralise une méthode fonctionnelle, initiée par Holley et Stroock [**63**]. Il montre par des inégalités fonctionnelles (Sobolev logarithmique) que l'énergie libre (c'est-à-dire l'entropie relative de la loi du processus au temps t par rapport à la probabilité invariante au même instant t) satisfait une certaine inégalité différentielle, et tend vers 0 à l'infini (sous certaines hypothèses). Ceci implique la convergence du processus vers les minima globaux du potentiel. C'est de cette dernière étude qu'est inspiré le chapitre 2.

Théorème 4.6. *Supposons que* $\lim_{t\to\infty} g(t)^{-1}\log G(t) = k$*, où* $k > 2osc(W)$ *est une constante liée aux oscillations de la fonction* W*. Alors le processus* Y *converge en loi vers une variable aléatoire* Y_∞*, dont la loi de probabilité est concentrée sur les minima globaux de* W*.*

Remarque 4.7. *Nous obtenons alors immédiatement que si nous supposons que* $g(t)^{-1}\log G(t)$ *converge vers une constante positive* $k > c$*. Alors les assertions suivantes sont équivalentes :*

(1) *le processus* Z *converge en loi vers* $\overline{\mu}_\infty$*,*

(2) $\overline{\mu}_t$ *converge presque sûrement vers* $\overline{\mu}_\infty$*,*

(3) $\frac{1}{t}\int_0^t Y_s \mathrm{d}s = o\left((\log t)^{-1}\right)$.

Cela nous implique alors le comportement suivant du processus Z, grâce au théorème de Slutsky.

Théorème 4.8. *Supposons que* $g(t)^{-1}\log G(t)$ *converge vers une constante positive* $k > c$*. Alors nous avons le résultat suivant :*

- *Si* W *est telle que* $\sum_{1\leq i\leq n} a_i m_i + \sum_{1\leq i\leq p} b_i M_i = 0$*, alors* Z_t *converge en loi vers* $Y_\infty + \int_0^\infty Y_s \frac{\mathrm{d}s}{r+s}$ *;*
- *Sinon, alors* Z_t *diverge.*

Remarque 4.9. *1) Une condition suffisante pour que* $\overline{\mu}_t$ *converge est que* W *a un unique minimum en 0. Ainsi, lorsque* $g(t) = 2c\log t$ *et* W *admet un unique minimum en 0, alors* Z *converge en loi vers 0. De plus, lorsque* W *est symmétrique, alors* Z *converge en loi.*
2) Remarquons également que ce dernier résultat n'est absolument pas en contradiction avec celui concernant la convergence presque sûre. Effectivement, nous conjecturons que si $k \leq c$*, alors le processus* Y *converge en loi vers une combinaison convexe de minima locaux de* W*. Cette étude n'est pas faite ici.*

5. Diffusions renforcées par la mesure d'occupation normalisée

Ce paragraphe correspond à la seconde partie de la thèse (chapitres 3, 4 et 5). Le chapitre 3 ne contient aucun nouveau résultat. Il s'agit en fait d'une introduction très générale sur les diffusions renforcées, énonçant bon nombre de résultats sans démonstration. Il permet de situer notre travail (c'est un court "survey"). Les chapitres 4 et 5 quant à eux généralisent les articles de Benaïm, Ledoux et Raimond [**10**], ainsi qu'en partie Benaïm et Raimond [**13**]. Ce sont les premiers résultats obtenus pour des diffusions renforcées par la mesure d'occupation normalisée dans des espaces non compacts.

Dans les chapitres 4 et 5 de ce manuscrit, nous nous plaçons sur l'espace de dimension infinie $\mathcal{P}(\mathbb{R}^d)$, l'ensemble des probabilités sur $\mathbb{R}^d$, muni de la topologie de la convergence faible, qui est métrisable. Ainsi, toutes les définitions (topologiques) précédentes d'ensemble limite, d'attracteur et d'ensemble transitif par chaîne sont encore valables, de même que le théorème annonçant que les ensembles limites de pseudo-trajectoires asymptotiques sont transitifs par chaîne. Le point très délicat ici est de prouver qu'une diffusion renforcée est encore pseudo-trajectoire asymptotique d'un certain système déterministe.

5.1. Etude théorique. Décrivons tout d'abord le modèle étudié. Supposons donnés deux potentiels $V : \mathbb{R}^d \to \mathbb{R}_+$ et $W : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}_+$. V est appelé "potentiel de confinement" et W est le "potentiel d'interaction", supposés de classe $\mathcal{C}^2$ et positifs. Nous étudions ici le processus X solution de l'E.D.S. suivante :

$$\begin{cases} \mathrm{d}X_t = \mathrm{d}B_t - (\nabla V(X_t) + \nabla W * \mu_t(X_t))\,\mathrm{d}t \\ \mathrm{d}\mu_t = (\delta_{X_t} - \mu_t)\frac{\mathrm{d}t}{r+t} \\ X_0 = x, \;\; \mu_0 = \mu \end{cases} \tag{5.1}$$

où (B_t) est un mouvement brownien de dimension d et (μ_t) est la mesure d'occupation normalisée du processus, avec un poids initial $r > 0$ et une probabilité initiale μ. Rappelons qu'il s'agit donc de la suite $(\mu_t; t \geq 0)$ définie par

$$\mu_t := \frac{r\mu + \int_0^t \delta_{X_s}\mathrm{d}s}{r+t}. \tag{5.2}$$

Le problème est de décrire le comportement asymptotique de μ_t. On se pose en particulier la question de savoir si ce processus vérifie le théorème ergodique ponctuel.

Les hypothèses essentielles sur les potentiels V et W sont les suivantes :

(1) V est une fonction strictement convexe et il existe des constantes $c > 0$ et $\delta > 1$ telles que on a $\lim_{|x|\to\infty} |x|^{-2\delta}(\nabla V(x), x) \geq c$;

(2) il existe $\kappa > 0$ telle que pour tout $x, y \in \mathbb{R}^d$, on a

$$W(x,y) + |\nabla_x W(x,y)| + |\nabla^2_{xx} W(x,y)| \leq \kappa(V(x) + V(y));$$

(3) $W(x,y) = W_1(x,y) + W_2(x,y)$ où W_2 ainsi que ses deux premières dérivées (en x) sont bornées par rapport à la variable x (ce qui

signifie que $W_2(x,y)+|\nabla_x W_2(x,y)|+|\nabla^2_{xx}W_2(x,y)| \leq \kappa V(y)$) et il existe $M>0$ telle que pour tout $x,y,\xi \in \mathbb{R}^d$, on a

$$((\nabla^2 V(x) + \nabla^2_{xx} W_1(x,y))\xi, \xi) \geq M|\xi|^2.$$

Remarque 5.1. *1) L'hypothèse de convexité de V sert à confiner le processus μ_t (afin d'obtenir de la tension), alors que la condition sur $(\nabla V(x), x)$ sert à appliquer le résultat de Röckner et Wang, cité dans le paragraphe sur les inégalités fonctionnelles (ultracontractivité uniforme).*
2) Nous sommes obligés de contrôler W par V, à la fois pour bien définir les objets considérés, mais surtout pour des raisons techniques.
3) L'hypothèse de courbure sert à obtenir un trou spectral, indispensable à notre étude. Elle signifie de plus que V peut compenser un manque de convexité de W.

Considérer des diffusions renforcées est singulièrement plus compliqué sur un espace non compact que sur un compact. En effet, le premier problème de taille est d'obtenir l'existence d'un trou spectral. Dans le cas compact, cela ne pose pas de problème majeur. En revanche, sur un espace non compact, cette existence n'est pas évidente du tout ! C'est pourquoi nous avons besoin au chapitre 4 d'introduire des inégalités fonctionnelles uniformes, inutiles dans le cas compact mais indispensables ici. L'existence du trou de spectre ainsi que l'uniforme ultracontractivité sont intimement liées au comportement asymptotique (stricte convexité) du potentiel de confinement V. Le manque de compacité entraîne également d'autres difficultés techniques. Effectivement, cela engendre des problèmes topologiques, que nous parvenons à surmonter grâce à la norme V définie un peu plus loin. Cela explique le besoin de dominer le potentiel d'interaction W par V. Enfin, nous obtenons un système dynamique qui engendre un flot local et non global. Le dernier point est le plus délicat : il faut montrer que le processus μ_t est pseudo-trajectoire asymptotique presque sûre du système dynamique en question.

5.1.1. *Tension.* La proposition suivante montre que, grâce au potentiel de confinement V qui est supposé strictement convexe, nous obtenons que le processus μ_t est tendu. Nous pouvons ainsi restreindre l'espace des probabilités sur $\mathbb{R}^d$ à un ensemble compact (pour la topologie faible* de la convergence des mesures).

Proposition 5.2. *i) La famille $(\mu_t, t \geq 0)$ est presque sûrement tendue.*
ii) L'ensemble limite de μ_t est compact pour la topologie faible.*

5.1.2. *Intuition.* Notre but est d'étudier le comportement asymptotique de μ_t. Si nous considérons le processus de diffusion X^μ, satisfaisant l'équation

$$\mathrm{d}X^\mu_t = \mathrm{d}B_t - (\nabla V(X^\mu_t) + \nabla W * \mu(X^\mu_t))\mathrm{d}t, \;\; X_= x,$$

qui correspond à l'équation (2.5), où μ_t est fixé à μ dans le terme de dérive, alors nous savons par un résultat d'ergodicité classique que μ_t converge presque sûrement vers la probabilité invariante de la diffusion X^μ. La probabilité invariante $\Pi : \mathcal{P}(\mathbb{R}^d) \to \mathcal{P}(\mathbb{R}^d)$ de la nouvelle diffusion est donnée par :

$$\Pi(\mu)(\mathrm{d}x) := \frac{1}{Z(\mu)} e^{-2(V(x)+W*\mu(x))}\mathrm{d}x. \tag{5.3}$$

Nous cherchons alors un équivalent à ce résultat classique et avons l'intuition suivante : lorsque μ_t converge, la mesure limite est un point fixe de Π. En fait, similairement aux processus d'approximation stochastique dont nous avons discuté précédemment, nous pouvons espérer que les trajectoires de μ_t sont approchées par un flot déterministe. Ceci se traduit par le fait que la mesure empirique μ_t est une pseudo-trajectoire asymptotique pour le flot du système dynamique induit par $\Pi(\mu_t) - \mu_t$. En réalité, nous allons montrer que la mesure changée de temps $\mu_{r(e^t-1)}$ est pseudo-trajectoire asymptotique du flot précédent.

Expliquons d'où vient le changement de temps. Remarquons tout d'abord que la dérivée temporelle de μ_t vaut $\frac{\mathrm{d}}{\mathrm{d}t}\mu_t = \frac{\delta_{X_t}-\mu_t}{r+t}$. Afin d'obtenir un résultat homogène temporellement, considérons le changement de temps $h(t) = r(e^t - 1)$. Nous obtenons alors $\frac{\mathrm{d}}{\mathrm{d}t}\mu_{h(t)} = \delta_{X_{h(t)}} - \mu_{h(t)}$. Il s'agit d'une mesure singulière par rapport à la mesure de Lebesgue. Néanmoins, nous parvenons à montrer que la différence entre cette dérivée et $\Pi(\mu_{h(t)}) - \mu_{h(t)}$ tend presque sûrement vers zéro lorsque t croît vers l'infini. Cela signifie que le processus μ_t est pseudo-trajectoire asymptotique presque sûre du système dynamique. Nous pouvons alors utiliser les techniques introduites dans les premiers paragraphes de cette introduction.

5.1.3. *Système dynamique et principaux résultats.* Comme nous travaillons sur un espace de mesures, nous héritons de différentes notions de convergence. Ainsi, nous parlons de pseudo-trajectoires asymptotiques (p.s. ou en probabilité) *faibles* (respectivement *fortes*) lorsque la distance d choisie correspond à la topologie de la convergence faible (respectivement forte) des mesures.

Nous sommes ici obligés de continuellement jongler entre la topologie faible* et la topologie forte des mesures. Ceci est dû au fait que la première topologie est tout à fait adaptée au domaine des probabilités, alors que les résultats liés au système dynamique sont connus pour la topologie forte.
Soit l'espace des mesures signées $\mathcal{M}(\mathbb{R}^d)$. Nous introduisons l'espace suivant :

$$\mathcal{M}(\mathbb{R}^d; V) := \{\mu \in \mathcal{M}(\mathbb{R}^d); \int_{\mathbb{R}^d} V(y)|\mu|(\mathrm{d}y) < \infty\}.$$

Soit $\mathcal{C}^0(\mathbb{R}^d, \mathcal{M}(\mathbb{R}^d; V))$ l'ensemble des fonctions continues de $\mathbb{R}^d$ dans $\mathcal{M}(\mathbb{R}^d; V)$ muni de la topologie de la convergence uniforme sur les intervalles compacts. Cet espace est métrisable et donc tous les résultats précédents concernant les systèmes dynamiques et les processus d'approximation stochastique sont encore vrais.

Afin de tenir compte du fait que l'espace ambiant n'est pas compact, nous introduisons la V-norme de la façon suivante : pour toute fonction continue f,

$$||f||_V := \sup\left\{\frac{|f(x)|}{V(x)}; x \in \mathbb{R}^d\right\}.$$

La convergence au sens de la topologie faible* se définit de la manière suivante. Considérons une suite de mesures $(\mu_n; n \geq 0)$. On dit que la suite μ_n *tend faiblement* vers la mesure μ_∞ si pour toute fonction continue $f \in \mathcal{C}^0(\mathbb{R}^d; V)$, de V-norme finie, on a

$$\lim_{n\to\infty} \int_{\mathbb{R}^d} f(x)\mu_n(\mathrm{d}x) = \int_{\mathbb{R}^d} f(x)\mu_\infty(\mathrm{d}x).$$

De plus, on dit que la suite (μ_n) tend vers μ_∞ pour la topologie forte si on a

$$\lim_{t\to\infty} ||\mu_n - \mu_\infty||_V = 0,$$

où $||\mu||_V := \sup\left\{\left|\int_{\mathbb{R}^d} f(x)\mu(\mathrm{d}x)\right|; \quad f \in \mathcal{C}^0(\mathbb{R}^d; V), |f| \text{ majorée par } V\right\}$.

Introduisons maintenant le système dynamique Φ. Soit $\Phi : \mathbb{R}\times\mathcal{M}(\mathbb{R}^d; V) \to \mathcal{M}(\mathbb{R}^d; V)$ le flot défini par

$$\Phi_0(\mu) = \mu, \quad \frac{\mathrm{d}}{\mathrm{d}t}\Phi_t(\mu) = \Pi(\Phi_t(\mu)) - \Phi_t(\mu). \tag{5.4}$$

C'est afin d'étudier ce flot que nous avons besoin des deux différentes topologies. Le résultat suivant annonce le lien entre le système déterministe et le processus aléatoire. Pour les conditions initiales x, r, μ, nous notons $\mathbb{P}_{x,r,\mu}$ la loi de la solution de (2.5).

THÉORÈME 5.3. *Sous la mesure $\mathbb{P}_{x,r,\mu}$, la fonction $t \mapsto \mu_{h(t)}$ est presque sûrement une pseudo-trajectoire asymptotique pour le flot Φ pour la topologie faible.*

Nous pouvons, grâce au théorème précédent combiné avec le théorème 1.12, décrire l'ensemble limite de la mesure d'occupation normalisée.

THÉORÈME 5.4. *L'ensemble limite de $(\mu_t, t \geq 0)$ est presque sûrement un ensemble libre d'attracteur pour le flot Φ.*

5.1.4. *Symétrie de W.* Dans certains cas particuliers, nous pouvons caractériser en détails l'ensemble limite de la mesure empirique.

THÉORÈME 5.5. *Supposons que W est symétrique. Alors, l'ensemble limite de $(\mu_t, t \geq 0)$ est $\mathbb{P}_{x,r,\mu}$-presque sûrement un sous-ensemble compact connexe (pour la topologie faible*) des points fixes de Π.*

COROLLAIRE 5.6. *Supposons W symétrique. Si l'ensemble des points fixes de Π est fini $(\mu_i, 1 \leq i \leq n)$, alors $(\mu_t; t \geq 0)$ converge presque sûrement vers l'un de ces points fixes.*

Remarque 5.7. *On ne peut pas expliciter la valeur des probabilités du corollaire, car elles dépendent du point de départ (cela est montré simplement sur un exemple au début du chapitre 4).*

5.2. Application : exemple sur $\mathbb{R}^2$. La mesure empirique μ_t peut avoir un comportement asymptotique fort complexe. Nous étudions un cas particulier, inspiré du cas compact ([**10**]), dans le chapitre 5. Décrivons ici ce qui se passe. Nous nous plaçons sur $\mathbb{R}^2$. Soit γ la probabilité définie par $\gamma(\mathrm{d}x) = \frac{1}{Z}e^{-2V(x)}\mathrm{d}x$, où Z est juste la constante de normalisation. Nous supposons que le potentiel de confinement V est tel que $V(x) = V(|x|)$ et que le potentiel d'interaction s'écrit sous la forme d'un produit scalaire $W(x,y) = (x, Ry)$ où R est une matrice de rotation sur $\mathbb{R}^2$. Dans la suite, nous étudions en premier lieu le cas $R = -Id$ (qui est déjà intéressant malgré la symétrie de W), puis le cas général. Evidemment, nous pouvons facilement généraliser cette étude au potentiel $W(x,y) = a(x, Ry)$ où a est un réel quelconque.

Le but de cet exemple est de montrer que la dynamique asymptotique de μ_t est très riche. Nous allons en particulier exhiber un cas où la mesure d'occupation μ_t converge vers une mesure aléatoire et un autre où elle tourne, ce qui signifie que son ensemble limite est un cercle de mesures.

Tout d'abord, commençons par étudier le comportement de la moyenne empirique $\overline{\mu}_t$. En effet, nous montrons –dans le chapitre 5– qu'il est suffisant d'étudier $\overline{\mu}_t$ pour obtenir des informations sur μ_t. Cela est dû au résultat suivant :

Lemme 5.8. *([**10**], corollaire 3.10) Soit (E,d) un espace métrique, $\bar{\Phi} : E \times \mathbb{R} \to E$ un flot sur E et $G : \mathcal{P}(\mathbb{R}^d) \to E$ une fonction continue. Supposons que $G \circ \Phi_t = \bar{\Phi}_t \circ G$. Soit L l'ensemble limite de μ_t. Alors presque sûrement, $G(L)$ est un ensemble libre d'attracteur pour $\bar{\Phi}$.*

Cela justifie le fait d'étudier $\overline{\mu}_t$. Il faut ensuite considérer l'application G et le flot $\bar{\Phi}$ qui conviennent. Il est facile de montrer que la moyenne empirique satisfait l'équation

$$\frac{\mathrm{d}}{\mathrm{d}t}\bar{\mu}_t = \bar{F}(\bar{\mu}_t), \;\; \bar{\mu}_0 = \bar{\mu},$$

avec $\bar{F}(\mu) := -\mu + \int_{\mathbb{R}^2} x\bar{\Pi}(\mu)(\mathrm{d}x)$ et $\bar{\Pi}(\bar{\mu})(\mathrm{d}x) := \frac{e^{-2(x,R\bar{\mu})}}{Z(\bar{\mu})}\gamma(\mathrm{d}x)$. Dans notre cas, nous choisissons $G(\mu) = \overline{\mu}$ et $\bar{\Phi}$ le flot engendré par l'équation précédente.

Définissons également la fonction $H : \mathbb{R}_+ \to \mathbb{R}_+$ qui joue un rôle primordial dans la suite de l'étude :

$$H(\alpha) := \int_0^\infty \mathrm{d}\rho\gamma(\rho) \int_0^{2\pi} \mathrm{d}v e^{-\alpha\rho\cos v}.$$

Remarquons que $H(\alpha)$ s'exprime comme intégrale de la fonction de Bessel I_0 par $H(\alpha) = \int_0^\infty \mathrm{d}\rho\gamma(\rho)I_0(\alpha\rho)$. Tout d'abord, étudions l'équation simplifiée (où p est le point cardinal $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ et $F : \mathbb{R}_+ \to \mathbb{R}$) :

$$\frac{\mathrm{d}}{\mathrm{d}t}\alpha = F(\alpha) = -\alpha + 2\frac{\mathrm{d}}{\mathrm{d}\alpha} \log\left(\int_{\mathbb{R}^2} e^{-\alpha(x,Rp)}\gamma(\mathrm{d}x)\right). \tag{5.5}$$

Nous pouvons alors énoncer le résultat intermédiaire suivant.

Proposition 5.9. *Si $\int_0^\infty \rho^2\gamma(\rho)\mathrm{d}\rho \leq 1$, alors 0 est l'unique équilibre de l'équation (5.5) et 0 est stable. De plus, le bassin d'attraction de 0 est $\mathbb{R}_+$.*

Si $\int_0^\infty \rho^2\gamma(\rho)\mathrm{d}\rho > 1$, alors 0 est un point linéairement instable et il existe un autre point d'équilibre (strictement positif), noté α_1, pour l'équation (5.5) qui est stable. En outre, le bassin d'attraction de α_1 est $\mathbb{R}_+^*$.

De cette proposition, nous déduisons le comportement asymptotique de μ_t dans le cas particulier qui suit, où W est symétrique.

Théorème 5.10. *Considérons la diffusion renforcée sur $\mathbb{R}^2$ associée aux potentiels $V(x) = V(|x|)$ et $W(x,y) = -(x,y)$. Nous avons alors deux comportements possibles :*

(1) *Si $\int_0^\infty \gamma(\rho)\rho^2\mathrm{d}\rho \leq 1$, alors p.s. $\mu_t \xrightarrow{(w)} \gamma$;*

(2) *Si $\int_0^\infty \gamma(\rho)\rho^2\mathrm{d}\rho > 1$, alors il existe une variable aléatoire $v \in \mathbb{S}^1$ telle que p.s. $\mu_t \xrightarrow{(w)} \mu_\infty^v$ avec*

$$\mu_\infty^v(\mathrm{d}x) = \frac{e^{\alpha_1(x,v)}}{Z_1}\gamma(\mathrm{d}x),$$

où Z_1 est la constante de normalisation et α_1 est l'unique solution strictement positive de l'équation $F(\alpha) = -\alpha + 2\frac{H'(\alpha)}{H(\alpha)} = 0$.

Voici une illustration dynamique de ce résultat. La première figure montre que la mesure γ est stable et attire toutes les trajectoires dans le cas où le moment quadratique de γ est inférieur à 1.

La figure suivante représente le point γ, qui devient instable lorsque l'intégrale est strictement supérieure à 1 :

Le principal résultat est le théorème suivant, qui généralise le théorème précédent.

Théorème 5.11. *Considérons la diffusion auto-interagissante sur $\mathbb{R}^2$ associée aux potentiels $V(x) = V(|x|)$ et $W(x,y) = (x, Ry)$, où R est la matrice de rotation bi-dimensionnelle $R = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$, avec $0 \leq \theta < 2\pi$. Alors le comportement asymptotique de μ_t est l'un des trois cas suivants :*

(1) *Si V est tel que $\int_0^\infty \gamma(\rho)\rho^2\mathrm{d}\rho\cos\theta + 1 > 0$, alors presque sûrement, nous avons $\mu_t \xrightarrow{(w)} \gamma$;*

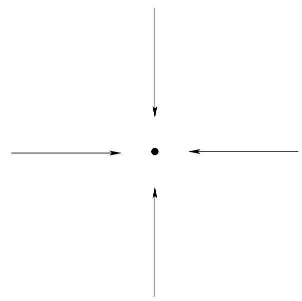

FIG. 4. La mesure d'occupation converge p.s. vers γ.

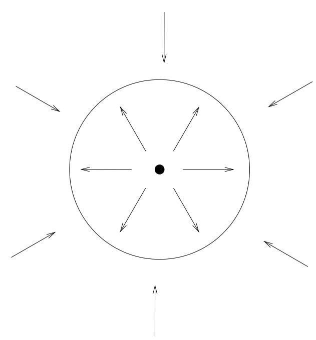

FIG. 5. La mesure d'occupation converge p.s. vers un point μ_∞^v du cercle.

(2) *Si V est tel que $\int_0^\infty \gamma(\rho)\rho^2 \mathrm{d}\rho \cos\theta + 1 \leq 0$, alors il y a deux différents comportements possibles :*

a) si $\theta = \pi$, alors il existe une variable aléatoire $v \in \mathbb{S}^1$ telle que presque sûrement $\mu_t \xrightarrow{(w)} \mu_\infty^v$ avec

$$\mu_\infty^v(\mathrm{d}x) = \frac{e^{\alpha_1(x,v)}}{Z_1}\gamma(\mathrm{d}x),$$

où Z_1 est la constante de normalisation et α_1 est l'unique solution strictement positive de l'équation $-\alpha + 2\frac{H'(\alpha)}{H(\alpha)} = 0$, *b) si $\theta \neq \pi$, alors l'ensemble limite de (μ_t) est $\{\nu(\delta), 0 \leq \delta < 2\pi\}$ avec probabilité 1, où*

$$\nu(\delta) = \frac{1}{e^{T_\theta} - 1} \int_0^{T_\theta} \mathrm{d}s e^s \mu_\infty^{v(s \tan\theta + \delta), \theta},$$

avec $T_\theta = 2\pi(\tan\theta)^{-1}$ et $\mu_\infty^{v,\theta}$ est l'unique solution strictement positive de $-\alpha + 2\cos\theta \frac{H'(\alpha)}{H(\alpha)} = 0$.

Nous remarquons immédiatement que dans le cas où $-\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2}$, alors presque sûrement, la mesure μ_t converge (au sens de la convergence faible des mesures) vers la probabilité γ. De plus, le moment quadratique de γ est essentiel dans cette étude : le comportement asymptotique de la mesure empirique normalisée dépend complétement de V, même pour des potentiels W relativement simples. Expliquons un peu ce dernier résultat.

Ce théorème signifie que la probabilité γ joue le même rôle que celui joué habituellement par la mesure de Lebesgue. De ce fait, si nous supposons que $\int_0^\infty \mathrm{d}\rho\gamma(\rho)\rho^2 \cos\theta > -1$, alors la mesure empirique se comporte comme un mouvement brownien (mais construit par rapport à la mesure γ).

En revanche, si $\theta = \pi$ et $\int_0^\infty \mathrm{d}\rho\gamma(\rho)\rho^2 \geq 1$, alors l'attraction est suffisamment forte pour que μ_t converge (au sens de la topologie faible). Dans ce cas, nous pouvons en fait contrer les variations dûes au mouvement brownien et la mesure d'occupation empirique converge presque sûrement vers une probabilité, qui est proche d'une gaussienne.

Si on considère enfin que $\theta \neq \pi$ et qu'il y a suffisamment d'attraction, alors le biais induit par θ oblige la mesure μ_t à tourner et l'ensemble limite de (μ_t) est un cercle de mesures $\{\nu(\delta), 0 \leq \delta < 2\pi\}$. Ce dernier comportement est illustré par la figure suivante.

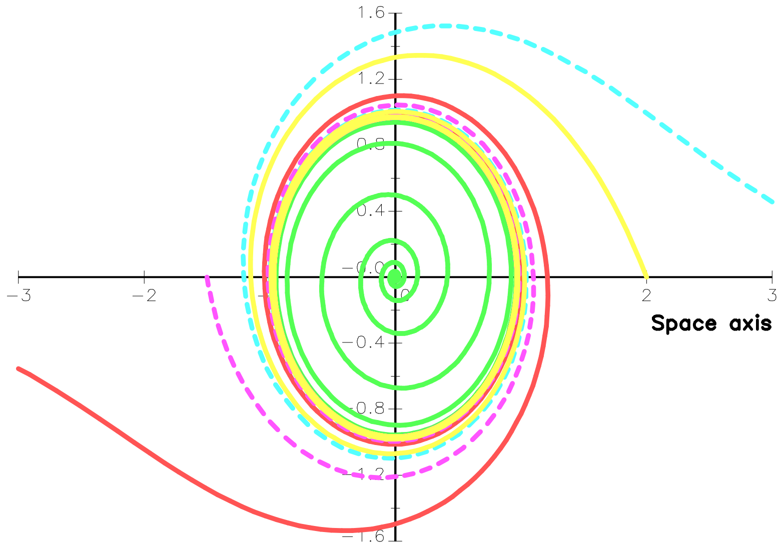

FIG. 6. L'ensemble limite est un cercle de mesures.

Bibliographie

[1] Ané, C., Blachère, S., Chafaï, D., Fougères, P., Gentil, I., Malrieu, F., Roberto, C., and Scheffer, G. *Sur les inégalités de Sobolev logarithmiques*, vol. 10 of *Panoramas et Synthèses [Panoramas and Syntheses]*. Société Mathématique de France, Paris, 2000. With a preface by Dominique Bakry and Michel Ledoux.

[2] Arthur, W. B., Ermoliev, Y. M., and Kaniovski, Y. M. Strong laws for a class of path-dependent stochastic processes with applications. In *Stochastic optimization (Kiev, 1984)*, vol. 81 of *Lecture Notes in Control and Inform. Sci.* Springer, Berlin, 1986, pp. 287–300.

[3] Bakry, D. L'hypercontractivité et son utilisation en théorie des semigroupes. In *Lectures on probability theory (Saint-Flour, 1992)*, vol. 1581 of *Lecture Notes in Math.* Springer, Berlin, 1994, pp. 1–114.

[4] Bakry, D., and Émery, M. Diffusions hypercontractives. In *Séminaire de probabilités, XIX, 1983/84*, vol. 1123 of *Lecture Notes in Math.* Springer, Berlin, 1985, pp. 177–206.

[5] Benachour, S., Roynette, B., Talay, D., and Vallois, P. Nonlinear self-stabilizing processes. I. Existence, invariant probability, propagation of chaos. *Stochastic Process. Appl. 75*, 2 (1998), 173–201.

[6] Benachour, S., Roynette, B., and Vallois, P. Nonlinear self-stabilizing processes. II. Convergence to invariant probability. *Stochastic Process. Appl. 75*, 2 (1998), 203–224.

[7] Benaïm, M. Sur la nature des ensembles limites des trajectoires des algorithmes d'approximation stochastiques de type Robbins-Monro. *C. R. Acad. Sci. Paris Sér. I Math. 317*, 2 (1993), 195–200.

[8] Benaïm, M. Dynamics of stochastic approximation algorithms. In *Séminaire de Probabilités, XXXIII*, vol. 1709 of *Lecture Notes in Math.* Springer, Berlin, 1999, pp. 1–68.

[9] Benaïm, M., and Hirsch, M. W. Asymptotic pseudotrajectories and chain recurrent flows, with applications. *J. Dynam. Differential Equations 8*, 1 (1996), 141–176.

[10] Benaïm, M., Ledoux, M., and Raimond, O. Self-interacting diffusions. *Probab. Theory Related Fields 122*, 1 (2002), 1–41.

[11] Benaim, M., and Raimond, O. On self attracting/repelling diffusions. *C. R. Math. Acad. Sci. Paris 335*, 6 (2002), 541–544.

[12] Benaïm, M., and Raimond, O. Self-interacting diffusions. II. Convergence in law. *Ann. Inst. H. Poincaré Probab. Statist. 39*, 6 (2003), 1043–1055.

[13] Benaïm, M., and Raimond, O. Self-interacting diffusions. III. Symmetric interactions. *Ann. Probab. 33*, 5 (2005), 1717–1759.

[14] Benaïm, M., and Schreiber, S. J. Ergodic properties of weak asymptotic pseudotrajectories for semiflows. *J. Dynam. Differential Equations 12*, 3 (2000), 579–598.

[15] Benveniste, A., Goursat, M., and Ruget, G. Robust identification of a nonminimum phase system : blind adjustment of a linear equalizer in data communications. *IEEE Trans. Automat. Control 25*, 3 (1980), 385–399.

[16] Benveniste, A., Métivier, M., and Priouret, P. *Adaptive algorithms and stochastic approximations*, vol. 22 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, 1990. Translated from the French by Stephen S. Wilson.

[17] Bergé, P., Pomeau, Y., and Vidal, C. *L'ordre dans le chaos*. A Wiley-Interscience Publication. Hermann, éditeurs des sciences et des arts., New York, 1986. Vers une approche déterministe de la turbulence, Préface de David Ruelle.

[18] Billingsley, P. *Convergence of probability measures*, second ed. Wiley Series in Probability and Statistics : Probability and Statistics. John Wiley & Sons Inc., New York, 1999. A Wiley-Interscience Publication.

[19] Bismut, J.-M. A generalized formula of Itô and some other properties of stochastic flows. *Z. Wahrsch. Verw. Gebiete 55*, 3 (1981), 331–350.

[20] Bolley, F., Guillin, A., and Villani, C. Quantitative concentration inequalities for empirical measures on non compact spaces. *Probab. Theory Related Fields 137* (2007), 541–593.

[21] Borodin, A. N., and Salminen, P. *Handbook of Brownian motion—facts and formulae*, second ed. Probability and its Applications. Birkhäuser Verlag, Basel, 2002.

[22] Bowen, R. ω-limit sets for axiom A diffeomorphisms. *J. Differential Equations 18*, 2 (1975), 333–339.

[23] Brandière, O., and Duflo, M. Les algorithmes stochastiques contournent-ils les pièges ? *Ann. Inst. H. Poincaré Probab. Statist. 32*, 3 (1996), 395–427.

[24] Carmona, P., Petit, F., and Yor, M. Some extensions of the arc sine law as partial consequences of the scaling property of Brownian motion. *Probab. Theory Related Fields 100*, 1 (1994), 1–29.

[25] Carmona, P., Petit, F., and Yor, M. Beta variables as times spent in $[0, \infty[$ by certain perturbed Brownian motions. *J. London Math. Soc. (2) 58*, 1 (1998), 239–256.

[26] Carrillo, J. A., McCann, R. J., and Villani, C. Kinetic equilibration rates for granular media and related equations : entropy dissipation and mass transportation estimates. *Rev. Mat. Iberoamericana 19*, 3 (2003), 971–1018.

[27] Cartan, H. *Calcul différentiel*. Hermann, Paris, 1967.

[28] Cattiaux, P., and Guillin, A. Deviation bounds for additive functionnals of markov processes, preprint, 2006.

[29] Cattiaux, P., Guillin, A., and Malrieu, F. Probabilistic approach for granular media equations in the non uniformly convex case, *Probab. Theory Related Fields*, 2007 (to appear).

[30] Černý, V. Thermodynamical approach to the traveling salesman problem : an efficient simulation algorithm. *J. Optim. Theory Appl. 45*, 1 (1985), 41–51.

[31] Chambeu, S., and Kurtzmann, A. Convergence in law of some particular self-interacting diffusions : the simulated annealing method. preprint, 2007.

[32] Chambeu, S., and Kurtzmann, A. Some particular self-interacting diffusions : ergodic behavior and almost sure convergence. preprint, 2007.

[33] Chaumont, L., and Doney, R. A. Pathwise uniqueness for perturbed versions of Brownian motion and reflected Brownian motion. *Probab. Theory Related Fields 113*, 4 (1999), 519–534.

[34] Chaumont, L., and Doney, R. A. Some calculations for doubly perturbed Brownian motion. *Stochastic Process. Appl. 85*, 1 (2000), 61–74.

[35] Chaumont, L., Doney, R. A., and Hu, Y. Upper and lower limits of doubly perturbed Brownian motion. *Ann. Inst. H. Poincaré Probab. Statist. 36*, 2 (2000), 219–249.

[36] Chiang, T.-S., Hwang, C.-R., and Sheu, S. J. Diffusion for global optimization in $\mathbf{R}^n$. *SIAM J. Control Optim. 25*, 3 (1987), 737–753.

[37] Conley, C. *Isolated invariant sets and the Morse index*, vol. 38 of *CBMS Regional Conference Series in Mathematics*. American Mathematical Society, Providence, R.I., 1978.

[38] Cranston, M., and Le Jan, Y. Self-attracting diffusions : two case studies. *Math. Ann. 303*, 1 (1995), 87–93.

[39] Cranston, M., and Mountford, T. S. The strong law of large numbers for a Brownian polymer. *Ann. Probab. 24*, 3 (1996), 1300–1323.

[40] Davies, E. B. *Heat kernels and spectral theory*, vol. 92 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 1990.

[41] Davies, E. B., and Simon, B. Ultracontractivity and the heat kernel for Schrödinger operators and Dirichlet Laplacians. *J. Funct. Anal. 59*, 2 (1984), 335–395.

[42] Davis, B. Weak limits of perturbed random walks and the equation $Y_t = B_t + \alpha \sup\{Y_s\colon\ s \le t\} + \beta \inf\{Y_s\colon\ s \le t\}$. *Ann. Probab. 24*, 4 (1996), 2007–2023.

[43] Davis, B. Brownian motion and random walk perturbed at extrema. *Probab. Theory Related Fields 113*, 4 (1999), 501–518.

[44] Demazure, M. *Bifurcations and catastrophes*. Universitext. Springer-Verlag, Berlin, 2000. Geometry of solutions to nonlinear problems, Translated from the 1989 French original by David Chillingworth.

[45] Doney, R. A. Some calculations for perturbed Brownian motion. In *Séminaire de Probabilités, XXXII*, vol. 1686 of *Lecture Notes in Math.* Springer, Berlin, 1998, pp. 231–236.

[46] Down, D., Meyn, S. P., and Tweedie, R. L. Exponential and uniform ergodicity of Markov processes. *Ann. Probab. 23*, 4 (1995), 1671–1691.

[47] Duflo, M. *Algorithmes stochastiques*, vol. 23 of *Mathématiques & Applications (Berlin) [Mathematics & Applications]*. Springer-Verlag, Berlin, 1996.

[48] Durrett, R. T., and Rogers, L. C. G. Asymptotic behavior of Brownian polymers. *Probab. Theory Related Fields 92*, 3 (1992), 337–349.

[49] Ethier, S. N., and Kurtz, T. G. *Markov processes*. Wiley Series in Probability and Mathematical Statistics : Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1986. Characterization and convergence.

[50] Gallot, S., Hulin, D., and Lafontaine, J. *Riemannian geometry*, third ed. Universitext. Springer-Verlag, Berlin, 2004.

[51] Geman, S., and Hwang, C.-R. Diffusions for global optimization. *SIAM J. Control Optim. 24*, 5 (1986), 1031–1043.

[52] Ghomi, M. The problem of optimal smoothing for convex functions. *Proc. Amer. Math. Soc. 130*, 8 (2002), 2255–2259 (electronic).

[53] Gidas, B. Nonstationary Markov chains and convergence of the annealing algorithm. *J. Statist. Phys. 39*, 1-2 (1985), 73–131.

[54] Gidas, B. The Langevin equation as a global minimization algorithm. In *Disordered systems and biological organization (Les Houches, 1985)*, vol. 20 of *NATO Adv. Sci. Inst. Ser. F Comput. Systems Sci.* Springer, Berlin, 1986, pp. 321–326.

[55] Gidas, B. Metropolis-type Monte Carlo simulation algorithms and simulated annealing. In *Topics in contemporary probability and its applications*, Probab. Stochastics Ser. CRC, Boca Raton, FL, 1995, pp. 159–232.

[56] Gross, L. Logarithmic Sobolev inequalities. *Amer. J. Math. 97*, 4 (1975), 1061–1083.

[57] Hajek, B. Cooling schedules for optimal annealing. *Math. Oper. Res. 13*, 2 (1988), 311–329.

[58] Harrison, J. M., and Shepp, L. A. On skew Brownian motion. *Ann. Probab. 9*, 2 (1981), 309–313.

[59] Herrmann, S., and Roynette, B. Boundedness and convergence of some self-attracting diffusions. *Math. Ann. 325*, 1 (2003), 81–96.

[60] Herrmann, S., and Scheutzow, M. Rate of convergence of some self-attracting diffusions. *Stochastic Process. Appl. 111*, 1 (2004), 41–55.

[61] Hirsch, M. W., and Smale, S. *Differential equations, dynamical systems, and linear algebra*. Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], New York-London, 1974. Pure and Applied Mathematics, Vol. 60.

[62] Holley, R., and Stroock, D. Logarithmic Sobolev inequalities and stochastic Ising models. *J. Statist. Phys. 46*, 5-6 (1987), 1159–1194.

[63] Holley, R., and Stroock, D. Simulated annealing via Sobolev inequalities. *Comm. Math. Phys. 115*, 4 (1988), 553–569.

[64] Hwang, C.-R. Laplace's method revisited : weak convergence of probability measures. *Ann. Probab. 8*, 6 (1980), 1177–1182.

[65] Hwang, C.-R., and Sheu, S. J. Large-time behavior of perturbed diffusion Markov processes with applications to the second eigenvalue problem for Fokker-Planck operators and simulated annealing. *Acta Appl. Math. 19*, 3 (1990), 253–295.

[66] Jacod, J., and Protter, P. *Probability essentials*. Universitext. Springer-Verlag, Berlin, 2000.

[67] Jacod, J., and Shiryaev, A. N. *Limit theorems for stochastic processes*, second ed., vol. 288 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2003.

[68] Kallenberg, O. *Foundations of modern probability*, second ed. Probability and its Applications (New York). Springer-Verlag, New York, 2002.

[69] Kavian, O., Kerkyacharian, G., and Roynette, B. Quelques remarques sur l'ultracontractivité. *J. Funct. Anal. 111*, 1 (1993), 155–196.

[70] Kendall, M., and Stuart, A. *The advanced theory of statistics. Vol. 1*, fourth ed. Macmillan Publishing Co., Inc., New York, 1977. Distribution theory.

[71] Kirkpatrick, S., Gelatt, Jr., C. D., and Vecchi, M. P. Optimization by simulated annealing. *Science 220*, 4598 (1983), 671–680.

[72] Kontoyiannis, I., and Meyn, S. P. Spectral theory and limit theorems for geometrically ergodic Markov processes. *Ann. Appl. Probab. 13*, 1 (2003), 304–362.

[73] Kunita, H. *Stochastic flows and stochastic differential equations*, vol. 24 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1990.

[74] Kurtzmann, A. An example of self-interacting diffusion on $\mathbb{R}^2$. preprint, 2007.

[75] Kurtzmann, A. The ode method for some self-interacting diffusions on non-compact spaces. preprint, 2007.

[76] Kushner, H. J., and Clark, D. S. *Stochastic approximation methods for constrained and unconstrained systems*, vol. 26 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1978.

[77] Kushner, H. J., and Yin, G. G. *Stochastic approximation algorithms and applications*, vol. 35 of *Applications of Mathematics (New York)*. Springer-Verlag, New York, 1997.

[78] Lang, S. *Real and functional analysis*, third ed., vol. 142 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1993.

[79] Ledoux, M. The geometry of Markov diffusion generators. *Ann. Fac. Sci. Toulouse Math. (6) 9*, 2 (2000), 305–366. Probability theory.

[80] Ledoux, M., and Talagrand, M. *Probability in Banach spaces*, vol. 23 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]*. Springer-Verlag, Berlin, 1991. Isoperimetry and processes.

[81] Lépingle, D. Sur le comportement asymptotique des martingales locales. In *Séminaire de Probabilités, XII (Univ. Strasbourg, Strasbourg, 1976/1977)*, vol. 649 of *Lecture Notes in Math.* Springer, Berlin, 1978, pp. 148–161.

[82] Ljung, L. Analysis of recursive stochastic algorithms. *IEEE Trans. Automatic Control AC-22*, 4 (1977), 551–575.

[83] Malrieu, F. Convergence to equilibrium for granular media equations and their Euler schemes. *Ann. Appl. Probab. 13*, 2 (2003), 540–560.

[84] McCann, R. J. Existence and uniqueness of monotone measure-preserving maps. *Duke Math. J. 80*, 2 (1995), 309–323.

[85] Meyn, S. P., and Tweedie, R. L. *Markov chains and stochastic stability.* Communications and Control Engineering Series. Springer-Verlag London Ltd., London, 1993.

[86] Meyn, S. P., and Tweedie, R. L. Stability of Markovian processes. II. Continuous-time processes and sampled chains. *Adv. in Appl. Probab. 25*, 3 (1993), 487–517.

[87] Meyn, S. P., and Tweedie, R. L. Stability of Markovian processes. III. Foster-Lyapunov criteria for continuous-time processes. *Adv. in Appl. Probab. 25*, 3 (1993), 518–548.

[88] Miclo, L. Recuit simulé sur $\mathbf{R}^n$. *C. R. Acad. Sci. Paris Sér. I Math. 310*, 11 (1990), 783–786.

[89] Norris, J. R., Rogers, L. C. G., and Williams, D. Self-avoiding random walk : a Brownian motion model with local time drift. *Probab. Theory Related Fields 74*, 2 (1987), 271–287.

[90] Norton, D. E. The fundamental theorem of dynamical systems. *Comment. Math. Univ. Carolin. 36*, 3 (1995), 585–597.

[91] Otto, F., and Villani, C. Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. *J. Funct. Anal. 173*, 2 (2000), 361–400.

[92] Pemantle, R. A survey of random processes with reinforcement. *Probab. Surveys 4* (2007), 1–79.

[93] Perman, M., and Werner, W. Perturbed Brownian motions. *Probab. Theory Related Fields 108*, 3 (1997), 357–383.

[94] Raimond, O. Self-attracting diffusions : case of the constant interaction. *Probab. Theory Related Fields 107*, 2 (1997), 177–196.

[95] Raimond, O. Self-interacting diffusions : a simulated annealing version. preprint, 2006.

[96] Revuz, D., and Yor, M. *Continuous martingales and Brownian motion*, third ed., vol. 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1999.

[97] Röckner, M., and Wang, F.-Y. Supercontractivity and ultracontractivity for (non-symmetric) diffusion semigroups on manifolds. *Forum Math. 15*, 6 (2003), 893–921.

[98] Rogers, L. C. G., and Williams, D. *Diffusions, Markov processes, and martingales. Vol. 2*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 2000. Itô calculus, Reprint of the second (1994) edition.

[99] Rothaus, O. S. Diffusion on compact Riemannian manifolds and logarithmic Sobolev inequalities. *J. Funct. Anal. 42*, 1 (1981), 102–109.

[100] Royer, G. A remark on simulated annealing of diffusion processes. *SIAM J. Control Optim. 27*, 6 (1989), 1403–1408.

[101] Tarrès, P. Pièges répulsifs. *C. R. Acad. Sci. Paris Sér. I Math. 330*, 2 (2000), 125–130.

[102] Tarrès, P. *Pièges des algorithmes stochastiques et marches aléatoires renforcées par sommet*. PhD thesis, Ecole Normale Supérieure de Cachan, 2001.

[103] Tóth, B. Self-interacting random motions—a survey. In *Random walks (Budapest, 1998)*, vol. 9 of *Bolyai Soc. Math. Stud.* János Bolyai Math. Soc., Budapest, 1999, pp. 349–384.

[104] Tóth, B., and Werner, W. The true self-repelling motion. *Probab. Theory Related Fields 111*, 3 (1998), 375–452.

[105] Tromba, A. J. The Morse-Sard-Brown theorem for functionals and the problem of Plateau. *Amer. J. Math. 99*, 6 (1977), 1251–1256.

[106] Villani, C. *Topics in optimal transportation*, vol. 58 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2003.

[107] Yor, M. *Some aspects of Brownian motion. Part I.* Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 1992. Some special functionals.

[108] Yor, M. *Some aspects of Brownian motion. Part II.* Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 1997. Some recent martingale problems.

Première partie

Diffusions renforcées par la moyenne de la mesure d'occupation normalisée

CHAPITRE 1

Comportement ergodique et convergence presque sûre de certaines diffusions auto-interactives

Il s'agit d'un travail en commun avec Sébastien Chambeu. Le titre original de cet article, écrit en anglais, est "Some particular self-interacting diffusions : ergodic behavior and almost sure convergence".

1. Introduction

Self-interacting diffusions have been first introduced by Durrett and Rogers [**10**] under the name of Brownian polymers. The study of such processes (with path-interaction) has been an intensive research area since the seminal work of Norris, Rogers and Williams [**15**]. Cranston and Le Jan [**7**], Raimond [**17**] and Herrmann and Roynette [**12**] have studied some self-interacting diffusions and have obtained that the sample paths of the solution converge a.s. or at least are a.s. bounded (see also Cranston & Mountford [**8**] and Mountford & Tarrès [**14**]). Self-interacting diffusions depending on the normalized occupation measure have been recently studied since the work of Benaïm, Ledoux and Raimond [**2**]. In the compact case (the process is living on a Riemannian compact manifold), they have proved that the asymptotic behavior of the normalized occupation measure can be related to the analysis of some deterministic dynamical flow defined on the space of the Borel probability measures. Benaïm and Raimond [**3, 4**] went further in this study and in particular, they gave sufficient conditions for the a.s. convergence of the normalized occupation measure.

Very recently, Raimond [**18**] has studied the asymptotic properties of some processes X living on a Riemannian compact manifold M, solution to the Stochastic Differential Equation (SDE)

$$\mathrm{d}X_t = \mathrm{d}B_t - g(t)\nabla V * \mu_t(X_t)\mathrm{d}t,$$

with $V * \mu_t(x) = \frac{1}{t}\int_0^t V(x, X_s)\mathrm{d}s$, $\mu_t = \frac{1}{t}\int_0^t \delta_{X_s}\mathrm{d}s$ and $g(t) = a\log(1+t)$ (or more generally $|g(t)| \leq a\log(t)$ and $g'(t) = O(t^{-\gamma})$ with $0 < \gamma \leq 1$). He has more particularly investigated the example $M = \mathbb{S}^n$ and $V(x,y) = -\cos d(x,y)$ (where d is the geodesic distance on $\mathbb{S}^n$) and proved that μ_t converges a.s. towards a Dirac measure.

In this paper, we are concerned with some self-interacting processes living on $\mathbb{R}^d$. Let us consider a potential $V : \mathbb{R}^d \to \mathbb{R}_+$ smooth enough and a non decreasing application $g : \mathbb{R}_+ \to \mathbb{R}_+$. Our goal is to study the ergodic behavior of the self-interacting diffusion X solution to the SDE

$$\begin{cases} \mathrm{d}X_t = \mathrm{d}B_t - g(t)\nabla V(X_t - \overline{\mu}_t)\mathrm{d}t \\ X_0 = x \end{cases} \tag{1.1}$$

where B is a standard Brownian motion and $\overline{\mu}_t$ denotes the empirical mean of the process X, defined by

$$\overline{\mu}_t = \frac{1}{r+t}\left(r\bar{\mu} + \int_0^t X_s \mathrm{d}s\right), \quad \overline{\mu}_0 = \overline{\mu}. \tag{1.2}$$

Here μ is an initial (given) probability measure on $\mathbb{R}^d$, $\bar{\mu}$ denotes the mean of μ and $r > 0$ is an initial weight.

We first get interested in this kind of processes when we were trying to extend the results of Benaïm, Ledoux and Raimond [**2**] and Raimond [**18**] to non-compact spaces. We began to notice that for a quadratic interaction potential V, the SDE satisfied by the process is exactly of the previous form and then we managed to go further when penalizing the occupation measure. Afterwards, it was natural to study the self-interacting diffusions discussed here.

The first important (and natural!) thing to do is to introduce the process Y defined by

$$Y_t = X_t - \overline{\mu}_t. \tag{1.3}$$

It appears that the process $(Y_t, t \geq 0)$ is the solution to the SDE

$$\begin{cases} \mathrm{d}Y_t = \mathrm{d}B_t - g(t)\nabla V(Y_t)\mathrm{d}t - Y_t \frac{\mathrm{d}t}{r+t}; \\ Y_0 = x - \overline{\mu}; \\ \mathrm{d}\overline{\mu}_t = Y_t \frac{\mathrm{d}t}{r+t}. \end{cases} \tag{1.4}$$

The study of Y is easier than the study of X and, moreover, it has an interesting asymptotic behavior. Actually, we will prove that, depending on g, the process Y converges a.s. and satisfies the pointwise ergodic theorem (the convergence in probability or in law will be studied in a forthcoming paper). Nevertheless, we are interested in the process X which does not satisfy the pointwise ergodic theorem in general, because the empirical mean does not converge. This explains the difficulty of the study of self-interacting diffusions in non-compact spaces that is $\mathrm{d}X_t = \mathrm{d}B_t - \int_{\mathbb{R}^d} \nabla V(X_t - x)\mathrm{d}\mu_t(x)\mathrm{d}t$.

We manage to give here a complete description to the asymptotic behavior of both μ_t and X (for this latter, we are just looking for the almost sure convergence). For simplicity, we suppose that the potential V does not admit any degenerate critical point. This assumption will be weakened in the following. First, we state the ergodic result:

Theorem 1.1. (1) *The process Y satisfies the pointwise ergodic theorem. This means that with probability 1, the normalized occupation measure of Y converges weakly to a random measure, and* *what is more, this last measure is a convex combination of Dirac measures taken in the critical points of V.*

(2) *The process X satisfies the pointwise ergodic theorem if and only if V has a unique minimum in 0 and converges almost surely.*

The second and main result of this paper is the following description of the asymptotic behavior of X:

THEOREM 1.2. *Suppose that $g(t)^{-1}\log G(t) = o((\log t)^{-2})$, where G is a primitive of g.*

(1) *Then the process Y_t converges almost surely to Y_∞, where Y_∞ belongs to the set of the local minima of V. Moreover, for all local minimum m of V, one has $\mathbb{P}(Y_\infty = m) > 0$.*
(2) *On one hand, on the set $\{Y_\infty = 0\}$, we have that both X_t and $\overline{\mu}_t$ converge almost surely to $\overline{\mu}_\infty := \overline{\mu} + \int_0^\infty Y_s \frac{\mathrm{d}s}{r+s}$. On the other hand, on the set $\{Y_\infty \neq 0\}$, we get that $\lim\limits_{t\to\infty} X_t/\log t = Y_\infty$.*

The remainder of the paper is organized in the following way. In Section 2, we introduce the notations and hypotheses, and add the proof of the existence of a global solution to the SDE studied. Afterwards, we motivate our study by the basic case V quadratic, for which we have an explicit expression of X and Y. Later, we describe in details the behavior of Y around the local extrema of V. Then, we describe in Section 5 the ergodic behavior of Y and give conditions (on g) for the almost sure convergence of Y. Finally, Section 6 is divided in two parts. The first one is devoted to the main results (namely we give necessary and sufficient conditions for the ergodic result for X), whereas in the second one we deal with conditions for the almost sure convergence of X (depending on g).

2. Notation, hypotheses and existence

We give briefly the notations we will use in the following. We denote by G the function $G(t) = \int_0^t g(s)\mathrm{d}s$. In the whole following, $(\cdot,\cdot)$ denotes the Euclidian scalar product. We denote by $\mathcal{P}(\mathbb{R}^d)$ the set of probability measure on $\mathbb{R}^d$.

In the sequel, the technical assumptions on the potential $V : \mathbb{R}^d \to \mathbb{R}_+$ are the following:

(1) (*regularity and positivity*) $V \in \mathcal{C}^2(\mathbb{R}^d)$ and $V \geq 0$;
(2) (*convexity*) V is strictly uniformly convex out of a compact set K;
(3) (*growth*) there exist $a, b > 0$ such that for all $x \in \mathbb{R}^d$, we have

$$\Delta V(x) \leq a + bV(x). \tag{2.1}$$

We also assume that V has a finite number of critical points. Let $Max = \{M_1, M_2, \cdots, M_p\}$ be the set of saddle points and local maxima of V and

$Min = \{m_1, m_2, \cdots, m_n\}$ be the set of the local minima of V. We assume that $\forall i$, $\forall \xi \in \mathbb{R}^d$, $(\nabla^2 V(m_i)\xi, \xi) > 0$ and for all M_i, $\nabla^2 V$ admits a negative eigenvalue. The convexity assumption implies that we can decompose $V = W + \chi$ where W is a strictly convex function and χ is a compactly supported function: there exists $c > 0$ and $\tilde{C} >$ such that $\nabla^2 W(x) \geq cId > 0$ and $\nabla\chi$ is Lipschitz with a constant $\tilde{C} > 0$.

Remark 2.1. *We can for instance suppose that V is (asymptotically) a polynomial ($V(x) = |x|^4$) or exponential.*

We also need to have some (not restrictive) hypothesis for the application $g : \mathbb{R}_+ \to \mathbb{R}_+$.

1. $g \in \mathcal{C}^1(\mathbb{R}_+)$ and $g > 0$ is positively bounded by below at the infinity;
2. for all $T > 0$, $G^{-1}(t+T) - G^{-1}(t) \underset{t\to\infty}{\longrightarrow} 0$ where G^{-1} is the generalized inverse of G (that is $G^{-1}(t) = \inf\{x | G(x) \geq t\}$);
3. $t\frac{g'(t)}{g(t)}$ converges when t tends to the infinity.

Remark 2.2. *Without any lost of generality, we can suppose that $g \geq 2$. For instance, $g(t) = 2 + t$ satisfies the preceding conditions.*

The main goal of this paper is to study the asymptotic behavior of μ_t and show that X satisfies the pointwise ergodic theorem, that is

Definition 2.3. *The process X satisfies the pointwise ergodic theorem if there exists a measure μ_∞ such that $\mu_t := \frac{1}{r+t}\left(r\mu + \int_0^t \delta_{X_s}\mathrm{d}s\right) \xrightarrow{w} \mu_\infty$ a.s.. That is for all continuous bounded f, $\frac{1}{t}\int_0^t f(X_s)\mathrm{d}s \xrightarrow{a.s.} \int f\mathrm{d}\mu_\infty$.*

Let us begin by showing that the SDE studied admits a unique global strong solution:

Proposition 2.4. *For any $x \in \mathbb{R}^d$, $\mu \in \mathcal{P}(\mathbb{R}^d)$ and $r > 0$, there exists a unique global strong solution $(X_t, t \geq 0)$ of (1.1).*

Proof. We want to emphasize that the local existence and uniqueness of all the SDEs studied in this paper is standard (see for instance [**20**] theorem 11.2). We will just prove here that Y, hence X (because $X_t = Y_t + \int_0^t Y_s \frac{\mathrm{d}s}{r+s}$), does not explode in a finite time.

We apply the Itô formula to the asymptotically increasing function $x \mapsto V(x)$:

$$\mathrm{d}V(Y_t) = (\nabla V(Y_t), \mathrm{d}B_t) + \left(\frac{1}{2}\Delta V(Y_t) - g(t)|\nabla V(Y_t)|^2 - \frac{1}{r+t}(\nabla V(Y_t), Y_t)\right)\mathrm{d}t.$$

Let us introduce the family of stopping times $\tau_n = \inf\{t \geq 0; V(Y_t) + \int_0^t g(s)|\nabla V(Y_s)|^2\mathrm{d}s > n\}$. We note that $\int_0^{t\wedge\tau_n}(\nabla V(Y_s), \mathrm{d}B_s)$ is a true martingale. Now the growth condition (3.1) on V implies that there exists C such that

$$\mathbb{E}V(Y_{t\wedge\tau_n}) \leq V(y) + C\int_0^t (1 + \mathbb{E}V(Y_{s\wedge\tau_n}))\mathrm{d}s.$$

Applying the Gronwall lemma to $\alpha(t) = \mathbb{E}V(Y_{t\wedge\tau_n})$, we get $\mathbb{E}V(Y_{t\wedge\tau_n}) \leq (V(y)+C+1)e^{Ct}$. It just remains to prove that for all $t \geq 0$, the probability $\mathbb{P}(\forall m, \tau_m \leq t)$ vanishes. The Markov inequality joint with the preceding inequality implies the following for any $n \in \mathbb{N}$:

$$\begin{aligned}
\mathbb{P}(\forall m, \tau_m \leq t) &\leq \mathbb{P}(\tau_n \leq t) = \mathbb{P}\left(\sup_{0\leq s\leq t}\left\{V(Y_s) + \int_0^s g(u)|\nabla V(Y_u)|^2 \mathrm{d}u\right\} > n\right) \\
&\leq \mathbb{P}\left(\sup_{0\leq s\leq t}\int_0^s (\nabla V(Y_u), \mathrm{d}B_u) > n/2\right) \\
&+ \mathbb{P}\left(V(y) + C(t + \int_0^t V(Y_u)\mathrm{d}u) > n/2\right) \\
&\leq \frac{2}{n}C\left(t + \int_0^t \mathbb{E}V(Y_s)\mathrm{d}s\right) = O(n^{-1}).
\end{aligned}$$

Therefore, we conclude that there exists m such that $V(Y_t) \leq m$ for all $t > 0$ and as $\lim_{|x|\to\infty} V(x) = \infty$, it implies that Y does not explode in a finite time. □

3. A motivating example: the quadratic case

We consider $V(x) = \frac{1}{2}(x, cx)$, where c is a symmetric positive definite matrix. In this case, the SDE studied becomes:

$$\begin{cases} \mathrm{d}X_t = \mathrm{d}B_t - g(t)\nabla V * \mu_t(X_t)\mathrm{d}t \\ X_0 = x \end{cases} \tag{3.1}$$

where $V * \mu_t(x) := \int V(x-y)\mu_t(\mathrm{d}y)$ and μ_t is the normalized occupation measure of the process, namely

$$\mu_t = \frac{r}{r+t}\mu + \frac{1}{r+t}\int_0^t \delta_{X_s}\mathrm{d}s.$$

Remark 3.1. *For the sake of simplicity, we restrict our attention to the case $d = 1$, because the method is exactly the same when $d \geq 1$. The only thing to do is to diagonalize the matrix c and to remember that for an orthogonal matrix U, the process $(U \cdot B_s, s \geq 0)$ is also a Brownian motion. Therefore we consider*

$$\nabla V * \mu_t(x) = cx - \frac{1}{r+t}\int_0^t cX_s \mathrm{d}s - \frac{r}{r+t}c\overline{\mu}.$$

3.1. Explicit expression of X. When the interaction function is quadratic, we can express X in terms of Brownian martingales and a deterministic part. Moreover, we can prove the convergence of the normalized occupation measure "directly" only with the expression of X_t and $\overline{\mu}_t$.

Proposition 3.2. *If X is the solution to (3.1) and $Y_t = X_t - \bar{\mu}_t$ then we have*

$$Y_t = \frac{1}{r+t} e^{-cG(t)} \left(\int_0^t (r+s) e^{cG(s)} \mathrm{d}B_s + r(x - \overline{\mu}) \right).$$

Proof. We introduce the process Y. This process satisfies

$$\mathrm{d}Y_t = \mathrm{d}B_t - \left(cg(t) + \frac{1}{r+t} \right) Y_t \mathrm{d}t, \;\; Y_0 = x - \overline{\mu}. \tag{3.2}$$

We want to show that $Y_t = \frac{1}{r+t} e^{-cG(t)} \left(\int_0^t (r+s) e^{cG(s)} \mathrm{d}B_s + r(x - \overline{\mu}) \right)$. Our strategy is to consider the process U, which is a modification of Y, defined by $U_t := (r+t) e^{cG(t)} Y_t$. Then the Itô formula implies

$$\mathrm{d}U_t = (r+t) e^{cG(t)} \mathrm{d}B_t \; ; \; U_0 = r(x - \overline{\mu})$$

and we deduce from U the expression of Y. □

Corollary 3.3. *The solution to the SDE (3.1) is given by*

$$X_t = x + rc(\overline{\mu} - x) F(t) + \int_0^t \left[1 - (r+s) c e^{cG(s)} (F(t) - F(s)) \right] \mathrm{d}B_s$$

and furthermore

$$\begin{aligned} \overline{\mu}_t &:= \frac{1}{r+t} \int_0^t X_s \mathrm{d}s + \frac{r}{r+t} \overline{\mu} = x + r\left(\overline{\mu} - x\right) \left(\frac{1}{r+t} e^{-cG(t)} + cF(t) \right) \\ &+ \int_0^t \left[1 - (r+s) c e^{cG(s)} \left(F(t) - F(s) + \frac{1}{r+t} \frac{e^{-cG(t)}}{c} \right) \right] \mathrm{d}B_s \end{aligned}$$

where $F(t) = \int_0^t e^{-cG(s)} \frac{g(s)}{r+s} \mathrm{d}s$.

Proof. We already know the expression of $Y_t := X_t - \bar{\mu}_t$. We also easily get that $\mathrm{d}\overline{\mu}_t = \frac{Y_t}{r+t} \mathrm{d}t$, and therefore, by Fubini's theorem for stochastic integrals

$$\overline{\mu}_t = \int_0^t (r+u) e^{cG(u)} (H(t) - H(u)) \mathrm{d}B_u + r(x - \bar{\mu}) H(t) + \bar{\mu}$$

with $H(t) := \int_0^t \frac{e^{-cG(u)}}{(r+u)^2} \mathrm{d}u$. This last result implies

$$\begin{aligned} X_t &= Y_t + \overline{\mu}_t \\ &= \int_0^t (r+s) e^{cG(s)} \left(\frac{e^{-cG(t)}}{r+t} + H(t) - H(s) \right) \mathrm{d}B_s \\ &+ \left(\frac{e^{-cG(t)}}{r+t} + H(t) \right) r(x - \overline{\mu}) + \overline{\mu} \end{aligned}$$

Using an integration by parts we get $H(t) - H(s) = \frac{e^{-cG(s)}}{r+s} - \frac{e^{-cG(t)}}{r+t} - \int_s^t \frac{cg(y) e^{-cG(y)}}{r+y} \mathrm{d}y$ and the result follows. □

Remark 3.4. *According to the expression of X, we find that $(X_t, t \geq 0) \overset{(d)}{=} (-X_t, t \geq 0)$ if and only if $\overline{\mu} = x = 0$.*

3.2. Ergodic result. We now begin to prove the pointwise ergodic theorem for the following non-homogeneous (Gauss-)Markov process Y.

Lemma 3.5. *Let $a : \mathbb{R}_+ \to \mathbb{R}_+$ be a continuous function, positively bounded by below at the infinity, such that $A(t) := \int_0^t a(s)\mathrm{d}s$ increases to $A(\infty) = \infty$, and for $V(t) := e^{-2A(t)} \int_0^t e^{2A(s)}\mathrm{d}s$, $V(\infty) < \infty$. Consider the process defined by*

$$\mathrm{d}Y_t = -a(t)Y_t\mathrm{d}t + \mathrm{d}B_t, \;\; Y_0 = y.$$

Then, denoting by γ the centered Gaussian measure with variance $V(\infty)$[1]*, we have for all continuous bounded function φ*

$$\frac{1}{t}\int_0^t \varphi(Y_s)\mathrm{d}s \xrightarrow[t\to\infty]{a.s.} \int \varphi(y)\gamma(\mathrm{d}y).$$

Proof. We will prove the result for the Fourier transform. We begin to note that we can give an explicit expression of this process, that is

$$Y_t = e^{-A(t)}\left(\int_0^t e^{A(s)}dB_s + y\right).$$

Let $\mathcal{F}_s := \sigma(B_u, 0 \leq u \leq s)$. It is obvious that, knowing the filtration $\mathcal{F}_s$, Y_t has a Gaussian law of mean $m(s,t) := e^{-(A(t)-A(s))}Y_s$ and variance $V(s,t) := e^{-2A(t)}\int_s^t e^{2A(u)}\mathrm{d}u$. Fix $t, u \in \mathbb{R}$ and define the martingale $M_s^{t,u} := \mathbb{E}\left(e^{iuY_t}|\mathcal{F}_s\right) = \exp\left\{iuY_s m(s,t) - \frac{u^2}{2}V(s,t)\right\}$. We use the Itô formula for the martingale $s \mapsto M_s^{t,u}$ to find that $\mathrm{d}M_s^{t,u} = iue^{-(A(t)-A(s))}M_s^{t,u}\mathrm{d}B_s$. As a consequence, we find (by integrating the preceding martingale between 0 and t) that

$$e^{iuY_t} = \mathbb{E}e^{iuY_t} + \int_0^t iue^{-(A(t)-A(s))}M_s^{t,u}\mathrm{d}B_s.$$

Then it is easily shown, while applying the Fubini theorem for stochastic integrals (see [**19**] p.175),

$$\int_0^t e^{iuY_s}\mathrm{d}s = \int_0^t \mathbb{E}e^{iuY_s}\mathrm{d}s + \int_0^t \mathrm{d}B_s \int_s^t iue^{-(A(r)-A(s))}M_s^{r,u}\mathrm{d}r. \tag{3.3}$$

But, as Y is a Gaussian process with variance $V(0,t)$, it converges in distribution to a Gaussian variable of law $\gamma = \mathcal{N}(0, V(\infty))$ and we have because of the Cesàro result,

$$\frac{1}{t}\int_0^t \mathbb{E}[e^{iuY_s}]\mathrm{d}s \underset{t\to\infty}{\to} e^{-u^2V(\infty)^2/2}.$$

It remains to find an equivalent to the stochastic part of (3.3). On the set $\{\int_0^\infty \left|\int_s^\infty iue^{-(A(r)-A(s))}M_s^{r,u}\mathrm{d}r\right|^2 \mathrm{d}s < \infty\}$, the stochastic part of (3.3)

[1] we consider here, by an abuse of notations, that $\mathcal{N}(0,0) = \delta_0$.

converges a.s. to a finite variable and therefore it is of the order of $o(t)$. We decompose the stochastic part of the equation(3.3) in

(3.4)
$$\int_0^t \mathrm{d}B_s \int_s^{+\infty} iue^{-(A(r)-A(s))} M_s^{r,u} \mathrm{d}r - \int_0^t \mathrm{d}B_s \int_t^{+\infty} iue^{-(A(r)-A(s))} M_s^{r,u} \mathrm{d}r.$$

On $\{\int_0^\infty \left|\int_s^\infty iue^{-(A(r)-A(s))} M_s^{r,u} \mathrm{d}r\right|^2 \mathrm{d}s = \infty\}$, we use the law of large numbers for martingales to get

$$\int_0^t \mathrm{d}B_s \int_s^\infty iue^{-(A(r)-A(s))} M_s^{r,u} \mathrm{d}r = o\left(\int_0^t \left|\int_s^\infty iue^{-(A(r)-A(s))} M_s^{r,u} \mathrm{d}r\right|^2 \mathrm{d}s\right).$$

Actually, we find the rough upper bound by using the initial definition of $M_s^{r,u}$:

$$\left|\int_s^t iue^{-(A(r)-A(s))} M_s^{r,u} \mathrm{d}r\right| \leq |u| \int_s^t e^{-(A(r)-A(s))} \mathrm{d}r = |u| e^{A(s)} (K(t) - K(s))$$

where we have defined $K(t) := \int_0^t e^{-A(s)} \mathrm{d}s$. We now need the following development of K: $K(t) = K(\infty) - \frac{e^{-A(t)}}{a(t)} + o\left(\frac{e^{-A(t)}}{a(t)}\right)$ and we thus find

$$\begin{aligned} \int_0^t e^{2A(s)} (K(t) - K(s))^2 \mathrm{d}s &\leq 2(K(t) - K(\infty))^2 \int_0^t e^{2A(s)} \mathrm{d}s \\ &+ 2\int_0^t e^{2A(s)} (K(\infty) - K(s))^2 \mathrm{d}s = O(t). \end{aligned}$$

For the second part of (3.4) we have

$$\left|\int_0^t \mathrm{d}B_s \int_t^\infty iue^{-(A(r)-A(s))} M_s^{r,u} \mathrm{d}r\right| = O\left(|u|(K(\infty) - K(t)) \int_0^t e^{A(s)} \mathrm{d}B_s\right).$$

We finally use the law of the iterated logarithm for the Brownian motion in order to obtain:

$$\int_0^t \mathrm{d}B_s \int_t^\infty iue^{-(A(r)-A(s))} M_s^{r,u} \mathrm{d}r = O\left(|u| a(t)^{-3/2} \log A(t)\right) = o(t)$$

and the result follows. $\square$

Lemma 3.6. *The variable $\overline{\mu}_t = \frac{1}{r+t} \int_0^t X_s \mathrm{d}s + \frac{r}{r+t} \overline{\mu}$ converges almost surely.*

Proof. We recall that $H(t) := \int_0^t \frac{e^{-cG(u)}}{(r+u)^2} \mathrm{d}u = \frac{1}{r} - cF(t) - \frac{e^{-cG(t)}}{r+t}$. We begin to decompose the process $\overline{\mu}_t = \overline{\mu}_t^1 + \overline{\mu}_t^2 + \overline{\mu}_t^3$ with

$$\begin{aligned} \overline{\mu}_t^1 &= \overline{\mu} + r(x - \overline{\mu}) H(t); \\ \overline{\mu}_t^2 &= (H(t) - H(\infty)) \int_0^t (r+s) e^{cG(s)} \mathrm{d}B_s; \\ \overline{\mu}_t^3 &= \int_0^t (r+s) e^{cG(s)} (H(\infty) - H(s)) \mathrm{d}B_u. \end{aligned}$$

Obviously the deterministic part of $\overline{\mu}_t$, namely $\overline{\mu}_t^1$, converges because of the convergence of H.
We need the following development of H:

$$H(t) = H(\infty) - \frac{1}{cg(t)(r+t)^2} e^{-cG(t)} + o\left(\frac{e^{-cG(t)}}{t^2 g(t)}\right). \tag{3.5}$$

The deterministic part of $\overline{\mu}_t^2$ is equivalent to $\frac{1}{cg(t)(r+t)^2} e^{-cG(t)}$. Moreover the quadratic variation of the stochastic part of $\overline{\mu}_t^2$ is equal to $\int_0^t (r+s)^2 e^{2cG(s)} \mathrm{d}s = O\left(\frac{(r+t)^2 e^{2cG(t)}}{\frac{G(t)}{t}}\right)$. The law of the iterated logarithm ([**13**] theorem 3) implies then that $\overline{\mu}_t^2 \xrightarrow[t\to\infty]{a.s.} 0$.
For the last part of $\overline{\mu}_t$, we remark that it is a local martingale, and actually a L^2-bounded-martingale. Thus $\overline{\mu}_t^3$ converges a.s. to $\bar{\mu}_\infty^3$.
We conclude that $\overline{\mu}_t \xrightarrow[t\to\infty]{a.s.} \overline{\mu}_\infty$ with $\overline{\mu}_\infty = \overline{\mu} + H(\infty) r(x - \overline{\mu}) + \overline{\mu}_\infty^3$. □

THEOREM 3.7. *With probability 1, the normalized occupation measure μ_t converges weakly to a random measure μ_∞, and the previous limit $\overline{\mu}_\infty$ is the mean of μ_∞.*

PROOF. We point out that the deterministic part of X_t is convergent (because of the formula 3.3).
We decompose the process X into three parts: $X_t = \overline{\mu}_\infty + \phi(t) U_t + o(1)$ where

$$\begin{aligned}
\overline{\mu}_\infty &= x + cr(\overline{\mu} - x)F(\infty) + \int_0^\infty \left[1 - (r+s)ce^{cG(s)}(F(\infty) - F(s))\right] \mathrm{d}B_s \\
U_t &= \frac{e^{-cG(t)}}{r+t} \int_0^t (r+s) e^{cG(s)} \mathrm{d}B_s \\
\phi(t) &= c(r+t)\left(F(\infty) - F(t)\right) e^{cG(t)}
\end{aligned}$$

One more time, we will prove the result for the Fourier transform of the process. We have the following:

$$\frac{1}{t} \int_0^t e^{iuX_s} \mathrm{d}s = \frac{e^{iu(\overline{\mu}_\infty + o(1))}}{t} \int_0^t e^{iu\phi(s)U_s} \mathrm{d}s.$$

As it was shown in the lemma 3.6, the random variable $\overline{\mu}_\infty$ is well-defined. Moreover, we also know (see the lemma 3.5) that the random variable $\phi(t)U_t$ satisfies the pointwise ergodic theorem. It then implies the ergodic result for this process: $\frac{1}{t}\int_0^t e^{iu\phi(s)U_s} \mathrm{d}s$ converges a.s.. As a consequence, the Fourier transform of μ_t converges a.s. and we conclude that $\mu_t \xrightarrow[t\to\infty]{w} \mu_\infty$ a.s. □

COROLLARY 3.8. *Suppose that g converges. Then the limit μ_∞ is a Gaussian measure with a random mean, $\mu_\infty = \mathcal{N}\left(\overline{\mu}_\infty, \frac{1}{2g(\infty)c}\right)$.*

PROOF. Straightforward. □

3.3. Asymptotic behavior of X. In the preceding subsection, we have still shown that the process X satisfies the pointwise ergodic theorem. We prove here that, depending on g, the process X exhibits two different asymptotic behaviors: either X converges in probability, or it converges almost surely.

First, we give a result describing roughly the asymptotic behavior of X.

PROPOSITION 3.9. *Suppose that g converges $\left(\lim_{t\to\infty} g(t) = a\right)$. Then we get $\mathbb{P}\left(\limsup_{t\to\infty} X_t = +\infty\right) = \mathbb{P}\left(\liminf_{t\to\infty} X_t = -\infty\right) = 1$.*

PROOF. The measure μ_∞ is diffusive. Let A be a subset of $\mathbb{R}$. We have that

$$\int_0^t \delta_{X_s}(A)\mathrm{d}s \underset{t\to\infty}{\sim} tl$$

where l is a positive constant depending on A. Therefore $\int_0^\infty \delta_{X_s}(A)\mathrm{d}s = \infty$ a.s.. It implies that for all constant $K > 0$ $\int_0^\infty \delta_{X_s}([K,\infty[)\mathrm{d}s = \infty$ a.s. and thus

$$\mathbb{P}\left(\bigcap_{K\geq 1}\left\{\int_0^\infty \mathrm{d}s\mathbf{1}_{\{X_s\geq K\}} = \infty\right\}\right) = 1.$$

We conclude that $\mathbb{P}(\limsup_{t\to\infty} X_t = +\infty) = 1$. The proof is exactly the same for $\underline{\lim}_{t\to\infty} X_t$. $\square$

PROPOSITION 3.10. *Suppose that $\lim_{t\to\infty} g(t) = \infty$. Then X_t converges in probability and μ_t converges a.s. to δ_{X_∞}.*

PROOF. We know that Y is a Gaussian process and $\mathbb{E}(Y_t^2) = O(g(t)^{-1}))$ thus Y converges in L^2 and therefore in probability to 0 (by the inequality of Markov). Writing $X_t = Y_t + \int_0^t Y_s \frac{\mathrm{d}s}{r+s}$, we get that $(\mathbb{E}|X_t|^2, t \geq 0)$ is a Cauchy sequence and thus converges. As consequence, X converges in L^2. We then easily get that μ_t converges toward δ_{X_∞} in probability. But we still know that μ_t converges a.s. and then we conclude by uniqueness of the limit. $\square$

PROPOSITION 3.11. *Suppose that $g(t)^{-1}\log G(t)$ is bounded for all $t \geq 0$. Then there exists $M > 0$ such that*

$$\mathbb{P}(\limsup_{t\to\infty}|Y_t| = M) = 1.$$

PROOF. The process Y satisfies the SDE (3.2). Therefore, we can rewrite Y as a Brownian local martingale: $Y_t = \frac{1}{f(t)}\left(Y_0 + \int_0^t f(s)\mathrm{d}B_s\right)$ where $f(t) := (r+t)e^{cG(t)}$. We point out the following asymptotic result

$$\int_0^t e^{2cG(s)}\mathrm{d}s = O\left(g(t)^{-1}e^{2cG(t)}\right).$$

The law of iterated logarithm permits us to conclude that there exists $M > 0$ such that with probability 1, we have $\overline{\lim}|Y_t| = M$. $\square$

Corollary 3.12. *Suppose that $g(t)^{-1}\log G(t)$ is bounded for all $t \geq 0$. Then the process X_t is bounded a.s., converges in probability (but not a.s.!) and a.s. μ_t converges weakly to δ_{X_∞}.*

Proof. We can express the process X in the following way: $X_t = Y_t + \overline{\mu}_t$. We know that Y is a.s. bounded and $\overline{\mu}_t$ converges a.s. so that X is also a.s. bounded. Moreover, Y is a Gaussian process and thus Y converges (in law) to a centered Gaussian variable, which is bounded, that is Y converges in probability to 0. But we remark that Y does not converge a.s. because of the law of the iterated logarithm. As a consequence, the process X converges in probability to $\overline{\mu}_\infty$. We conclude by the uniqueness of the limit that $\mu_t \xrightarrow[t\to\infty]{a.s.} \delta_{X_\infty}$. $\square$

Remark 3.13. *We meet the condition of the preceding result for instance with $g(t) = \log(1+t)$.*

Proposition 3.14. *Suppose that $\lim\limits_{t\to\infty} g(t)^{-1}\log G(t) = 0$. Then the following holds:*
i) The process $Y_t := X_t - \overline{\mu}_t$ converges to 0 a.s.
ii) The process X_t converges to $\overline{\mu}_\infty$ a.s. and μ_t converges a.s. to $\delta_{\overline{\mu}_\infty}$.

Proof. We only have to prove that $Y_t := X_t - \overline{\mu}_t$ converges a.s. to 0. We have seen that $Y_t = \frac{e^{-cG(t)}}{r+t}\int_0^t (r+s)e^{cG(s)}\mathrm{d}B_s + r(x-\overline{\mu})\frac{e^{-cG(t)}}{r+t} =: Y_t^1 + Y_t^2$. The deterministic part of Y_t, namely Y_t^2, converges obviously to 0. Then, the law of the iterated logarithm implies that Y_t converges a.s. to 0. $\square$

Remark 3.15. *For instance we can choose $g(t) = t^\beta$ with $\beta > 0$ or $g(t) = e^t$.*

4. Study of the process Y

We study the process Y, which is the solution to the following SDE

$$\mathrm{d}Y_t = \mathrm{d}B_t - \left(g(t)\nabla V(Y_t) + \frac{Y_t}{r+t}\right)\mathrm{d}t;\ \ Y_0 = x - \overline{\mu}. \tag{4.1}$$

We recall that V is a general potential of class $\mathcal{C}^2$ and strictly uniformly convex out of a compact set. From now on, we suppose that $\lim\limits_{t\to\infty} g(t) = +\infty$.

4.1. The process Y_t gets close to the critical points of V.

Proposition 4.1. *Let Y be the solution to (1.4). Then with probability 1, the process Y gets as close as we want to the set $Min \cup Max$, that is $\forall\varepsilon > 0$, let $T_t^\varepsilon := \inf\{s \geq t; d(Y_s, Min \cup Max) < \varepsilon\}$, then for all t, we get $T_t^\varepsilon < \infty$ a.s.*

Proof. Let $\varepsilon > 0$. We apply the Itô formula to the function $x \mapsto V(x)$ and we obtain

$$\mathrm{d}V(Y_t) = \left(\nabla V(Y_t), \mathrm{d}B_t - g(t)\nabla V(Y_t)\mathrm{d}t - \frac{Y_t}{r+t}\mathrm{d}t\right) + \frac{1}{2}\Delta V(Y_t)\mathrm{d}t;$$

$$\mathrm{d}V(Y_t) = (\nabla V(Y_t), \mathrm{d}B_t) - D(t, Y_t)\mathrm{d}t. \tag{4.2}$$

where we have introduced

$$D(t,y) \quad = \quad g(t)|\nabla V(y)|^2 + \frac{1}{r+t}(y, \nabla V(y)) - \frac{1}{2}\Delta V(y). \tag{4.3}$$

We recall that $Min \cup Max$ is the set of the local extrema of V. Then it follows from the hypotheses on the function V that on the set $\{z; d(z, Min \cup Max) > \varepsilon\}$ and for $t \geq 0$, the applications $y \mapsto \frac{1}{r+t}(y, \nabla V(y)) + \frac{1}{2}|\nabla V(y)|^2 - \frac{1}{2}\Delta V(y)$ and $y \mapsto (g(t) - \frac{1}{2})|\nabla V(y)|^2$ are bounded from below (recall the growth assumption $\Delta V \leq a + bV$, and because V is strictly convex out of a compact set, there exists a positive constant k such that $\Delta V(y) \leq k(1 + |\nabla V(y)|^2)$. Actually, for the strictly convex function W, there exists a positive constant k such that $W(y) \leq k|\nabla V(y)|^2)$. Moreover, the second one, that is $y \mapsto (g(t) - \frac{1}{2})|\nabla V(y)|^2$, is positive for t arbitrarily large. Therefore, there exists $t_0 = t_0(\varepsilon)$ such that: $\forall\ t > t_0$, $\forall\ y\ \in \{z; d(z, Min \cup Max) > \varepsilon\}$ we have

$$g(t)|\nabla V(y)|^2 + \frac{1}{r+t}(y, \nabla V(y)) - \frac{1}{2}\Delta V(y) \geq \frac{g(t)}{2}|\nabla V(y)|^2 > 0. \tag{4.4}$$

We also introduce the stopping time $T_t^\varepsilon = \inf\{s \geq t; d(Y_s, Min \cup Max) < \varepsilon\}$ and we want to prove that for all $t > t_0$, we get $\mathbb{P}(T_t^\varepsilon < +\infty) = 1$. Then, it follows from (4.2) and (4.4) that, for $t > t_0$, we have two super-martingales:

$$\left(V(Y_{s\wedge T_t^\varepsilon})\right)_{s\geq t} \quad \text{and} \quad \left(V(Y_{s\wedge T_t^\varepsilon}) + \frac{1}{2}\int_0^{s\wedge T_t^\varepsilon} g(u)|\nabla V(Y_{u\wedge T_t^\varepsilon})|^2\mathrm{d}u\right)_{s\geq t}.$$

These two super-martingales are nonnegative and therefore they converge a.s. as $s \to \infty$. As a consequence, the process $\left(\int_0^{s\wedge T_t^\varepsilon} g(u)|\nabla V(Y_{u\wedge T_t^\varepsilon})|^2 du\right)_{s\geq t}$ also converges a.s. If we suppose that we are on the set $\{T_t^\varepsilon = +\infty\}$ it then follows that

$$|\nabla V(Y_{s\wedge T_t^\varepsilon})|^2 \xrightarrow[s\to\infty]{a.s.} 0.$$

Thus $Y_{s\wedge T_t^\varepsilon}$ gets close to $Min \cup Max$ and there is a contradiction. Finally, $\mathbb{P}(T_t^\varepsilon < +\infty) = 1$ for all $t \geq t_0$ and the result follows. □

Corollary 4.2. *Let $(Y_t)_{t\geq 0}$ be the solution to (1.4). Then a.s. the process Y gets close to the set $Min \cup Max$ infinitely often, i.e. $\forall \varepsilon > 0$, there exists a sequence of stopping times $(T_n)_{n\geq 1}$ such that T_n goes to the infinity and*

$$\forall n \geq 1,\ \mathbb{P}(T_n < +\infty) = 1 \quad \textit{and} \quad d(Y_{T_n}, Min \cup Max) < \varepsilon.$$

Proof. We apply the proof of the preceding proposition 4.1, with $T_n^\varepsilon = \inf\{s > n; d(Y_s, Min \cup Max) < \varepsilon\} < \infty$ a.s. We conclude the proof by taking $T_n = T_n^\varepsilon$. □

4.2. Case of a stable critical point: local minimum. We want to study the behavior of the process Y near a local minimum m. We will prove that if the process Y is near a local minimum m then the hypothesis on V and g imply that the probability such that the set $\{Y_s; s \geq 0\}$ is included in a neighborhood of m is positive. Actually, a second-order Taylor expansion permits us to compare $(y-m, \nabla V(y))$ and $|y-m|^2$ and we use a comparison theorem for the associate SDE.

Let m be a local minimum. We assume that m is a local minimum of V with $\nabla^2 V(m) > 0$. Taylor's formula implies that there exists $a > 0$ and $\varepsilon_0 > 0$ such that

$$\forall y \text{ such that } |y-m| \leq \varepsilon_0 \text{ we have } (y-m, \nabla V(y)) \geq a|y-m|^2.$$

PROPOSITION 4.3. *Suppose that* $t = o(g(t))$. *Let* $\varepsilon_0 > \varepsilon > 0$ *and* Y *the solution to (1.4). If there exists* $T_0 > 0$ *such that, for all* $T > T_0$, $|Y_T - m| < \varepsilon$, *then the event* $\{\forall s \geq T; |Y_s - m| < \varepsilon\}$ *has a positive probability to occur. Moreover, almost surely, on the event* $\{\forall s \geq T; |Y_s - m| < \varepsilon\}$, *we have*

$$|Y_{t+T} - m| = O\left(\sqrt{g(t+T)^{-1}\log G(t+T)}\right).$$

PROOF. We will show during the proof that the event $\{\forall s \geq T; |Y_s - m_i| < \varepsilon\}$ has a positive probability to occur. Suppose for the moment that this event has a positive probability.

Suppose that $d = 1$. Let us consider the process $\widetilde{Y}$ defined by $\widetilde{Y}_t = Y_{t+T} - m$. Then, $V''(m) > 0$ implies that there exists $a > 0$ such that

$$\forall y \text{ such that } y \in [0;\varepsilon],\ V'(y+m) \geq ay. \tag{4.5}$$

Let us introduce the non-negative process U solution to the SDE

$$\mathrm{d}U_t = \mathrm{d}B_t^T - ag(t+T)U_t\mathrm{d}t + \mathrm{d}L_t,\ U_0 = \widetilde{Y_0}, \tag{4.6}$$

where L corresponds to the local time of U in 0. We now proceed in several steps.

Step 1: we show that the equation (4.6) has a unique solution. Let U_t be a solution of (4.6). Let Z be the process defined by $Z_t = e^{aG(t+T)}U_t$. By definition of U, we easily obtain that

$$\mathrm{d}Z_t = e^{aG(t+T)}\mathrm{d}B_t^T + e^{aG(t+T)}\mathrm{d}L_t.$$

Let $\alpha(t)$ be the function such that $\int_0^{\alpha(t)} e^{2aG(s+T)}\mathrm{d}s = t$ and define the process A by $A_t := \int_0^t e^{aG(s+T)}\mathrm{d}L_s$. Then, if we consider the time-changed process $Z_{\alpha(t)} = W_t + A_{\alpha(t)}$ (where $W_t = \int_0^{\alpha(t)} e^{aG(s+T)}\mathrm{d}B_s^T$ is a Brownian motion), we remark that A_t increases if and only if L_t increases and consequently, $A_{\alpha(t)}$ increases if and only if $Z_{\alpha(t)}$ vanishes. $A_{\alpha(t)}$ is the local time at zero of the standard Brownian motion W. The Skorokhod lemma entails that the process $Z_{\alpha(t)}$ is uniquely defined (see [**11, 19**]) by $Z_{\alpha(t)} = W_t^+$

where W_t^+ is the reflected Brownian motion associated to W. Therefore the SDE (4.6) has a unique (strong) solution given by

$$U_t = e^{-aG(t+T)} W^+_{\alpha^{-1}(t)}. \tag{4.7}$$

We point out that the process U_t is nonnegative.

Step 2: Now by the law of the iterated logarithm, we obtain that there exists a positive constant C such that

$$U_t \leq C e^{-aG(t+T)} \sqrt{\alpha^{-1}(t) \log(\log(\alpha^{-1}(t)))} \quad a.s.$$

If we manage to prove that $C < \infty$, then this proves directly that $U_t = O\left(g(t+T)^{-1/2}\sqrt{\log G(t+T)}\right)$. We know that there exists a one dimensional Brownian motion β such that

$$(U_t, t \geq 0) \overset{(d)}{=} (e^{-aG(t+T)} \left(\int_0^t e^{2aG(s+T)} \mathrm{d}s\right)^{1/2} |\beta_t^T|, t \geq 0).$$

As a consequence, we get that the process $\sup\limits_{t\geq 0} U_t$ has the same law as $\sup\limits_{t\geq 0} \left(\int_0^t e^{2a(G(s+T)-G(t+T))}\mathrm{d}s\right)^{1/2} |\beta_t^T|$. But we remark that we have the following upper bound:

$$\sup_{t\geq 0}\left(\int_0^t e^{2a(G(s+T)-G(t+T))}\mathrm{d}s\right)^{1/2} |\beta_t^T| \leq \sup_{t\geq 0}\left(\int_0^t e^{2a(G(s+T)-G(t+T))}\mathrm{d}s\right)^{1/2} \sup_{0\leq s\leq t} \beta_s^T.$$

Moreover, by the identity of Lévy, the process $\sup\limits_{0\leq s\leq t} \beta_s^T$ has the same law as L_t^T, that is the local time of β in 0. As a corollary to the identity of Lévy (we use the scaling property for the Brownian motion) one can prove that $L_t/\sqrt{t} \overset{(d)}{=} L_1 \overset{(d)}{=} \sup\limits_{0\leq s\leq 1} \beta_s = |Z|$, where Z is a standard Gaussian variable. Thus, $L_t/\sqrt{t}$ converges to Z. We recall that $t = o(g(t))$. Therefore, for all $\eta > 0$, we have $\mathbb{P}\left(\sup\limits_{t\geq 0} \frac{L_t^T}{\sqrt{g(t+T)}} < \eta\right) > 0$. Moreover, we easily find (with an integration by parts) that $\int_0^t e^{2aG(s+T)}\mathrm{d}s$ is asymptotically equivalent to $e^{2aG(t+T)}/g(t+T)$. We conclude that for all $\eta > 0$, we get

$$\mathbb{P}\left(\sup_{t\geq 0}\left(\int_0^t e^{2a(G(s+T)-G(t+T))}\mathrm{d}s\right)^{1/2} \sup_{0\leq s\leq t} \beta_s^T < \eta\right) > 0.$$

Thus, the result follows and C is a finite non-random constant.

Step 3: we will now prove a martingale comparison theorem who enables us to show that $\widetilde{Y}_t \leq U_t$ a.s. Let l be a function of class $\mathcal{C}^2$ such that:

$$\begin{cases} \forall x > 0,\ l(x) > 0 \text{ and } l'(x) > 0, \\ \forall x \leq 0,\ l(x) = 0 \end{cases}$$

According to the Itô formula added to (4.6), we have

$$\begin{aligned} l(\widetilde{Y}_t - U_t) &= -\int_0^t l^{'}(\widetilde{Y}_s - U_s)\left(g(s+T)V^{'}(\widetilde{Y}_s + m) - g(s+T)aU_s + \frac{\widetilde{Y}_s + m}{r+s+T}\right)\mathrm{d}s \\ &- \int_0^t l^{'}(\widetilde{Y}_s - U_s)\mathrm{d}L_s. \end{aligned}$$

We recall that the process U is nonnegative. On the event $\{\widetilde{Y}_s > U_s\}$, the process $\widetilde{Y}$ is positive, and therefore by the lower bound (4.5), we find that $g(s+T)V^{'}(\widetilde{Y}_s+m) - g(s+T)aU_s \geq g(s+T)(a\widetilde{Y}_s - aU_s)$. We then have the almost sure bound $l(\widetilde{Y}_t - U_t) \leq 0$ and this added to the definition of l leads to

$$\widetilde{Y}_t \leq U_t \quad a.s. \tag{4.8}$$

Using the same argument on $[-\varepsilon, 0]$ we get the lower bound $V_t \leq \widetilde{Y}_t$, where V_t is a non-positive process.

Finally, the processes V and U satisfy, by the law of the iterated logarithm

$$\limsup_{t\to\infty} U_t = -\liminf_{t\to+\infty} V_t = \lim_{t\to\infty}\sqrt{g(t+T)^{-1}\log G(t+T)}.$$

Suppose that $d > 1$. We have seen previously that, without any lost of generality, we can reduce to the case $m = 0$. We introduce the process $\widetilde{Y}_t = Y_{t+T}$. The Itô formula implies (recall that $d > 1$)

$$\mathrm{d}|\widetilde{Y}_t| = \mathrm{d}W_t - g(t+T)\left(\frac{\widetilde{Y}_t}{|\widetilde{Y}_t|}, \nabla V(\widetilde{Y}_t)\right)\mathrm{d}t - \frac{|\widetilde{Y}_t|}{r+t+T}\mathrm{d}t + \frac{d-1}{2|\widetilde{Y}_t|}\mathrm{d}t \tag{4.9}$$

where $W_t = \int_0^t \left(\frac{\widetilde{Y}_s}{|\widetilde{Y}_s|}, \mathrm{d}B_s^T\right)$ is a standard Brownian motion. Then, the condition $\nabla^2 V(0) > 0$ implies that there exists $a > 0$ such that

$$\forall y \in [0;\varepsilon],\ (y, \nabla V(y)) \geq a|y|^2. \tag{4.10}$$

Let us introduce the $d-1$ dimensional Bessel process R. Consider the time-changed process $U_t := e^{-aG(t+T)}R_{\int_0^t e^{2aG(s+T)}\mathrm{d}s}$, which is the nonnegative strong solution to the SDE

$$\mathrm{d}U_t = \mathrm{d}\beta_t^T - ag(t+T)U_t\mathrm{d}t + \frac{d-1}{2U_t}\mathrm{d}t, \tag{4.11}$$

where β_t is a Brownian motion. Now, applying the comparison theorem (we have still proved it) to the nonnegative processes $\widetilde{Y}_t$ and U_t we obtain, on the event $\{\forall s \geq T; |Y_s| < \varepsilon\}$, that $\widetilde{Y}_t \leq U_t$. On the other hand R_t is the radial part of a d-dimensional Brownian motion. With the same argument of scaling as in the one dimensional case, the law of the iterated logarithm implies that $R_t = O(\sqrt{(t+T)\log\log(t+T)})$, and consequently $U_t = O\left(\sqrt{g(t+T)^{-1}\log G(t+T)}\right)$.

It remains to prove that the event $\{\forall s \geq T; |Y_s - m| < \varepsilon\}$ has a positive probability to occur. Let $\tau_T := \inf\{s > T; |Y_s - m| > \varepsilon\}$. We know that for $t < \tau_T$, we have $|Y_{t+T} - m| \leq U_t + V_t$. Then, we find that almost surely, $\overline{\lim_{t\to\infty}} |Y_{t\wedge\tau_T} - m| \leq \lim_{t\to\infty} \sqrt{g(t)^{-1}\log G(t)} < \varepsilon$. As a consequence, we get that $\tau_T = \infty$ almost surely. This concludes the proof. $\square$

Corollary 4.4. *Suppose that $g(t)^{-1}\log G(t) \underset{t\to\infty}{\longrightarrow} 0$. Then, the event $\{\forall s \geq T; |Y_s - m| < \varepsilon\}$ the process $Y_t - m$ converges almost surely to 0.*

Proof. We follow the previous proof and recall that $|Y_t - m| \leq U_t + V_t$. We conclude by the law of the iterated logarithm. $\square$

4.3. Case of an unstable critical point.

4.3.1. *Case of a local maximum.* Let M_i be a local maximum of V. The fact that $\Delta V(M_i) < 0$ and the hypothesis on V imply that $\varepsilon_1 := \sup\{\varepsilon; \forall |y| < \varepsilon,\ \Delta V(M_i + y) < 0\}$ exists and is finite.

Proposition 4.5. *Let $0 < \varepsilon < \varepsilon_1$, M_i a local maximum of V and T a positive stopping time such that, for Y the solution of (1.4), we have $|Y_T - M_i| < \varepsilon$. Then*

$$\mathbb{P}\left(\forall s \geq T;\ |Y_s - M_i| < \varepsilon\right) = 0.$$

Proof. For the sake of simplicity, we restrict our attention to the case $M_i = 0$, because the method is exactly the same when $M_i \neq 0$.

We recall the Itô formula :

$$\mathrm{d}V(Y_{t+T}) \quad = \quad \mathrm{d}M_{t+T} - D(t+T, Y_{t+T})\mathrm{d}t$$

where M_t is the local martingale $\int_0^t (\nabla V(Y_s), \mathrm{d}B_s)$ and $D(t,y)$ is defined by (4.3). On the event $A := \{\forall s \geq T;\ |Y_s| < \varepsilon\}$ we immediately obtain the bound

$$\begin{aligned} D(t+T, Y_{t+T}) \quad &= \quad g(t+T)|\nabla V(Y_{t+T})|^2 + \frac{1}{r+t+T}(Y_{t+T}, \nabla V(Y_{t+T})) \\ &- \quad \frac{1}{2}\Delta V(Y_{t+T}) \\ &\geq \quad \frac{C_1}{r+t+T} + \frac{C_2}{2} \end{aligned}$$

where $C_1 = \inf\{(y, \nabla V(y)); |y| < \varepsilon\}$ and $C_2 = -\sup\{\Delta V(y); |y| < \varepsilon\} > 0$. We thus find for t large enough that $D(t+T, Y_{t+T}) \geq C > 0$ and therefore, with $M_t \mathbb{1}_A = o(t)$, we get

$$\mathbb{E}\left(V(Y_{t+T})\mathbb{1}_A\right) \leq \mathbb{E}\left(V(Y_T)\mathbb{1}_A\right) - Ct\mathbb{P}(A) + o(t). \tag{4.12}$$

Finally, this last inequality is impossible since V is a nonnegative function. To conclude, we obtain $\mathbb{P}(A) = 0$. $\square$

Remark 4.6. *If $M_i \neq 0$ we have an additional term $M_i \log(t+T)$ and the proof is exactly the same.*

4.3.2. *Case of a saddle point.* Let M_i be a saddle point of V. First of all, we remark that, if $\Delta(M_i) < 0$, then we can follow the proof of the proposition 3.3 to conclude. But we prefer to give here a general proof.

Let e be an unstable direction (that is $\partial^2_{ee}V(M_i) < 0$) associate to the saddle point M_i and $P_e : \mathbb{R}^d \mapsto \mathbb{R}e$ the projection on $\mathbb{R}e$. We know by the hypotheses on V that such a direction exists (because for all i, $\nabla^2 V$ admits a negative eigenvalue in M_i). As $\partial^2_{ee}V(M_i) < 0$ and the hypothesis on V imply that $\varepsilon_2 := \sup\{\varepsilon;\ \forall |y| < \varepsilon, \partial^2_{ee}V(M_i + y) < 0 \text{ and } (\partial_e V(P_e(y)), \partial_e V(y)) > 0\}$ exists and is finite.

PROPOSITION 4.7. *Let $0 < \varepsilon < \varepsilon_2$, M_i a saddle point of V and T a positive stopping time such that, for Y the solution of (1.4), we have $|Y_T - M_i| < \varepsilon$. Then*

$$\mathbb{P}\left(\forall s \geq T;\ |Y_s - M_i| < \varepsilon\right) = 0$$

PROOF. One more time, for the sake of simplicity, we restrict our attention to the case $M_i = 0$, because the method is exactly the same when $M_i \neq 0$. The Itô formula applied to the function $x \mapsto V(P_e(x))$ implies that

$$\mathrm{d}V\left(P_e(Y_{t+T})\right) = \mathrm{d}M_{t+T} - \widetilde{D}(t+T, Y_{t+T})\mathrm{d}t,$$

where M is the local martingale $\int_0^t \left(\partial_e V(P_e(Y_s)), P_e(\mathrm{d}B_s)\right)$ and the drift term is defined by

$$\widetilde{D}(t, Y_t) := g(t)\left(\partial_e V(P_e(Y_t)), \partial_e V(Y_t)\right) + \frac{1}{r+t}\left(\partial_e V(P_e(Y_t)), P_e(Y_t)\right) - \frac{1}{2}\partial^2_{ee}V(P_e(Y_t)).$$

On the event $A := \{\forall s \geq T;\ |Y_s| < \varepsilon\}$ we immediately obtain the bound

$$\widetilde{D}(t+T, Y_{t+T}) \quad \geq \quad \frac{C_3}{r+t+T} + \frac{C_4}{2},$$

where $C_3 := \inf\{(P_e(y), \partial_e V(P_e(y))); |y| < \varepsilon\}$ and $C_4 := -\sup\{\partial^2_{ee}V(P_e(y)); |y| < \varepsilon\} > 0$. We thus find for t large enough that $\widetilde{D}(t+T, Y_{t+T}) \geq 2C > 0$ and therefore, with $M_t \mathbb{1}_A = o(t)$, we get

$$\mathbb{E}\left(V(P_e(Y_{t+T}))\mathbb{1}_A\right) \leq \mathbb{E}\left(P_e(V(Y_T))\mathbb{1}_A\right) - Ct\mathbb{P}(A) + o(t).$$

Finally, as V is a nonnegative function, this last inequality is impossible for $\mathbb{P}(A) > 0$. $\square$

5. Asymptotic behavior of Y

5.1. Pointwise ergodic theorem. The aim of this paragraph is to prove that Y satisfies the pointwise ergodic theorem. We begin to show that Y is bounded in L^2.

LEMMA 5.1. *The process Y is L^2-bounded.*

PROOF. The Itô formula implies, with $V = W + \chi$

$$\mathrm{d}|Y_t|^2 = 2(Y_t, \mathrm{d}B_t) - 2g(t)(Y_t, \nabla W(Y_t))\mathrm{d}t - 2g(t)(Y_t, \nabla\chi(Y_t))\mathrm{d}t - \frac{2|Y_t|^2}{r+t}\mathrm{d}t + d\mathrm{d}t.$$

But by hypothesis W is strictly convex everywhere, with a constant of convexity C, and χ a compactly supported function, therefore $y \mapsto (y, \nabla\chi(y))$ is bounded (by a positive constant M). For all $n \in \mathbb{N}$, define the stopping time $\tau_n = \inf\{t; |Y_t| > n\}$. Then we get by localization (and because g is a non decreasing function):

$$\mathbb{E}|Y_{t\wedge\tau_n}|^2 \leq \mathbb{E}|Y_0|^2 + dt + MG(t) < \infty.$$

Thus, we let n goes to the infinity, we use the lemma of Fatou and for all $t \geq 0$, we get $Y_t \in L^2$. By application of the Itô formula, we find the following inequality

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbb{E}|Y_t|^2 \leq -2\left(Cg(t) + \frac{1}{r+t}\right)\mathbb{E}|Y_t|^2 + d + Mg(t).$$

Now we solve this inequality by solving $\dot{u} = -2(Cg(t) + \frac{1}{r+t})u$ and then

$$\mathbb{E}|Y_t|^2 \leq u(t)\left(\mathbb{E}|Y_0|^2 + \int_0^t (d + Mg(s))u(s)^{-1}\mathrm{d}s\right) = O(1).$$

□

Remark 5.2. *The same result holds for $V(Y_t)$ by adapting the proof.*

The idea (in order to obtain the ergodic result for Y) is to introduce a dynamical system ϕ for which Y is an asymptotic pseudo-trajectory in probability, that is

Definition 5.3. *The process Y is an asymptotic pseudo-trajectory in probability for the flow ϕ if $\forall T, \alpha > 0$, it holds*

$$\lim_{t\to\infty}\mathbb{P}\left(\sup_{0\leq h\leq T}|Y_{t+h} - \phi_h(Y_t)| \geq \alpha|\mathcal{F}_t\right) = 0. \tag{5.1}$$

We refer the reader to [**1**] for more details on the notion of asymptotic pseudo-trajectory.

Let us consider the time-changed process $Y_{G^{-1}(t)}$. This process satisfies in particular (for all $h \geq 0$)

$$\begin{aligned} Y_{G^{-1}(t+h)} - Y_{G^{-1}(t)} &= B_{G^{-1}(t+h)} - B_{G^{-1}(t)} - \int_0^h \nabla V(Y_{G^{-1}(t+s)})\mathrm{d}s \\ &\quad - \int_0^h Y_{G^{-1}(t+s)}\frac{\mathrm{d}s}{\kappa(t+s)} \end{aligned}$$

where we have defined $\kappa(t) := (r + G^{-1}(t))g(G^{-1}(t))$.

Proposition 5.4. *Let $\phi : \mathbb{R}\times\mathbb{R}^d \to \mathbb{R}^d$ be the flow defined by*

$$\frac{\mathrm{d}}{\mathrm{d}t}\phi_t(x) = -\nabla V(\phi_t(x)); \quad \phi_0(x) = x. \tag{5.2}$$

Then Y is an asymptotic pseudo-trajectory in probability for the flow ϕ, that is we have for all $T > 0$ and $\alpha > 0$:

$$\lim_{t\to\infty} \mathbb{P}\left(\sup_{0\le h\le T} |Y_{G^{-1}(t+h)} - \phi_h(Y_{G^{-1}(t)})| \ge \alpha | \mathcal{F}_t \right) = 0.$$

PROOF. A simple computation enables us to find, with the notation $\widetilde{Y}_t = Y_{G^{-1}(t)}$ and $\widetilde{B}_t = B_{G^{-1}(t)}$, that

$$\begin{aligned} \widetilde{Y}_{t+h} - \phi_h(\widetilde{Y}_t) &= \widetilde{B}_{t+h} - \widetilde{B}_t + \int_0^h \left(\nabla V(\phi_s(\widetilde{Y}_t)) - \nabla V(\widetilde{Y}_{t+s}) \right) \mathrm{d}s \\ &- \int_0^h \widetilde{Y}_{t+s} \frac{\mathrm{d}s}{\kappa(t+s)}. \end{aligned}$$

If we now consider the square process $|\widetilde{Y}_{t+h} - \phi_h(\widetilde{Y}_t)|^2$, we can apply the Itô formula to the function $h \mapsto e^{-2\tilde{C}h}|\widetilde{Y}_{t+h} - \phi_h(\widetilde{Y}_t)|^2$. We recall that V is a strictly uniformly convex function out of a compact set, that is the sum of a uniform convex function W (with constant $C > 0$) and a compactly supported function χ such that $\nabla\chi$ is $\tilde{C}$-Lipschitz. It then implies

$$\begin{aligned} \frac{1}{2}\mathrm{d}(e^{-2\tilde{C}h}|\widetilde{Y}_{t+h} - \phi_h(\widetilde{Y}_t)|^2) &= e^{-2\tilde{C}h}\left(\widetilde{Y}_{t+h} - \phi_h(\widetilde{Y}_t), \mathrm{d}\widetilde{B}_{t+h} \right) \\ &+ e^{-2\tilde{C}h}\left(\widetilde{Y}_{t+h} - \phi_h(\widetilde{Y}_t), \nabla V(\phi_h(\widetilde{Y}_t)) - \nabla V(\widetilde{Y}_{t+h}) \right) \mathrm{d}h \\ &+ e^{-2\tilde{C}h}\frac{1}{\kappa(t+h)}\left(\widetilde{Y}_{t+h} - \phi_h(\widetilde{Y}_t), \widetilde{Y}_{t+h} \right) \mathrm{d}h \\ &+ \frac{1}{2}e^{-2\tilde{C}h}\mathrm{d} < \widetilde{Y}_{t+\cdot} - \phi_\cdot(\widetilde{Y}_t) >_h \\ &- \tilde{C}e^{-2\tilde{C}h}|\widetilde{Y}_{t+h} - \phi_h(\widetilde{Y}_t)|^2\mathrm{d}h \\ &\le e^{-2\tilde{C}h}(\widetilde{Y}_{t+h} - \phi_h(\widetilde{Y}_t), \mathrm{d}\widetilde{B}_{t+h}) \\ &- Ce^{-2\tilde{C}h}|\widetilde{Y}_{t+h} - \phi_h(\widetilde{Y}_t)|^2\mathrm{d}h + \frac{1}{2g(G^{-1}(t+h))}\mathrm{d}h \\ &+ \frac{e^{-2\tilde{C}h}}{\kappa(t+h)}\left(\widetilde{Y}_{t+h} - \phi_h(\widetilde{Y}_t), \widetilde{Y}_{t+h} \right) \mathrm{d}h. \end{aligned}$$

As a consequence, we have the following upper bound for the square process:

$$\begin{aligned} \frac{1}{2}\sup_{0\le h\le T} |\widetilde{Y}_{t+h} - \phi_h(\widetilde{Y}_t)|^2 &\le \sup_{0\le h\le T} e^{2\tilde{C}h}\int_0^h e^{-2\tilde{C}s}(\widetilde{Y}_{t+s} - \phi_s(\widetilde{Y}_t), \mathrm{d}\widetilde{B}_{t+s}) \\ &+ \frac{e^{2\tilde{C}T}}{2}\int_0^T \frac{1}{g(G^{-1}(t+s))}\mathrm{d}s \\ &+ \sup_{0\le h\le T} e^{2\tilde{C}h}\int_0^h e^{-2\tilde{C}s}\frac{\left(\widetilde{Y}_{t+s} - \phi_s(\widetilde{Y}_t), \widetilde{Y}_{t+s} \right)}{\kappa(t+s)}\mathrm{d}s. \end{aligned}$$

We can now deduce a upper bound for the mean of the preceding process. By the inequality of Burkholder-Davis-Gundy for the local martingale $\int_0^h e^{-2\tilde{C}s}(\widetilde{Y}_{t+s} - \phi_s(\widetilde{Y}_t), \mathrm{d}\widetilde{B}_{t+s})$ and a rough upper bound for its quadratic variation, there exists a positive constant C_2 such that:

$$\begin{aligned}
\mathbb{E}\left(\sup_{0\le h\le T} |\widetilde{Y}_{t+h} - \phi_h(\widetilde{Y}_t)|^2\right) &\le C_2 e^{4\tilde{C}T}(G^{-1}(t+T) - G^{-1}(t))\mathbb{E}\left(\sup_{0\le h\le T} |\widetilde{Y}_{t+h} - \phi_T(\widetilde{Y}_t)|^2\right) \\
&+ e^{2\tilde{C}T}(G^{-1}(t+T) - G^{-1}(t)) \\
&+ e^{2\tilde{C}T}\mathbb{E}\left(\sup_{0\le h\le T}\int_0^h \left(\widetilde{Y}_{t+s} - \phi_s(\widetilde{Y}_t), \widetilde{Y}_{t+s}\right)\frac{\mathrm{d}s}{\kappa(t+s)}\right).
\end{aligned}$$

We now need to estimate the last mean of the previous inequality. We have:

$$\int_0^h \left(\widetilde{Y}_{t+s} - \phi_s(\widetilde{Y}_t), \widetilde{Y}_{t+s}\right)\frac{\mathrm{d}s}{\kappa(t+s)} \le \frac{1}{2}\int_0^h \frac{|\widetilde{Y}_{t+s} - \phi_s(\widetilde{Y}_t)|^2}{\kappa(t+s)}\mathrm{d}s + \frac{1}{2}\int_0^h \frac{|\widetilde{Y}_{t+s}|^2}{\kappa(t+s)}\mathrm{d}s.$$

The next step is to use the lemma 5.1 and the fact that the function κ is non-decreasing in order to find the bounds:

$$\begin{aligned}
\mathbb{E}\left(\sup_{0\le h\le T}\int_0^h \frac{|\widetilde{Y}_{t+s}|^2}{\kappa(t+s)}\mathrm{d}s\right) &\le \frac{MT}{\kappa(t)}; \\
\mathbb{E}\left(\sup_{0\le h\le T}\int_0^h \frac{|\widetilde{Y}_{t+s} - \phi_s(\widetilde{Y}_t)|^2}{\kappa(t+s)}\mathrm{d}s\right) &\le \frac{T}{\kappa(t)}\mathbb{E}\left(\sup_{0\le h\le T} |\widetilde{Y}_{t+h} - \phi_h(\widetilde{Y}_t)|^2\right).
\end{aligned}$$

But by hypothesis, we recall that $(G^{-1}(t+T) - G^{-1}(t))$ and $\kappa(t)^{-1}$ converge to 0 when t increases to the infinity. As a consequence we obtain for t large enough:

$$\mathbb{E}\left(\sup_{0\le h\le T} |\widetilde{Y}_{t+h} - \phi_h(\widetilde{Y}_t)|^2\right) \le 2e^{4\tilde{C}T}(G^{-1}(t+T) - G^{-1}(t)) + 2Me^{2\tilde{C}T}\frac{T}{\kappa(t)}.$$

To conclude, we just need to use the inequality of Markov:

$$\mathbb{P}\left(\sup_{0\le h\le T} |\widetilde{Y}_{t+h} - \phi_h(\widetilde{Y}_t)| \ge \alpha|\mathcal{F}_t\right) \le \frac{1}{\alpha^2}\mathbb{E}\left(\sup_{0\le h\le T} |\widetilde{Y}_{t+h} - \phi_h(\widetilde{Y}_t)|^2\right)$$

and the result follows. □

LEMMA 5.5. *Let us consider* $(\mu_t^{G^{-1}}, t \ge 0)$ *the family of the normalized occupation measure of the time-changed process* $Y_{G^{-1}}$*. Then* $(\mu_t^{G^{-1}}, t \ge 0)$ *is a tight family of measures.*

PROOF. We will show that a.s. $\varphi(t) := \int_0^t V(Y_{G^{-1}(s)})\mathrm{d}s = O(t)$ and the result follows (indeed let $A > 0$ and K a compact set such that $\forall x \in K^c$ we have $V(x) \ge A$ then $\mu_t^{G^{-1}}(V) \ge A\mu_t^{G^{-1}}(K^c)$). From the growth assumption (3.1) on V, we know that there exist $a, b > 0$ and for all $\varepsilon > 0$, the convexity assumption implies that there exists a constant k_ε such that

$$\Delta V \le a + bV \text{ and } V \le k_\varepsilon + \varepsilon|\nabla V|^2.$$

It then easily implies that

$$\varphi(t) \leq k_\varepsilon t + \varepsilon \int_0^t |\nabla V(Y_{G^{-1}(s)})|^2 \mathrm{d}s \tag{5.3}$$

and $\int_0^t \Delta V(Y_{G^{-1}(s)})\mathrm{d}s \leq at + b\varphi(t)$. If we apply the Itô formula to the process $t \mapsto V(Y_{G^{-1}(t)})$, we obtain that

$$\begin{aligned} V(Y_{G^{-1}(t)}) - V(Y_{G^{-1}(0)}) &= \int_0^t (\nabla V(Y_{G^{-1}}(s)), \mathrm{d}B_s) - \int_0^t g(s)|\nabla V(Y_{G^{-1}(s)})|^2 \mathrm{d}s \\ &- \int_0^t \frac{(Y_{G^{-1}(s)}, \nabla V(Y_{G^{-1}(s)})\mathrm{d}s}{(r+G^{-1}(s))g(G^{-1}(s))} + \frac{1}{2}\int_0^t \Delta V(Y_{G^{-1}(s)})\mathrm{d}s. \end{aligned} \tag{5.4}$$

We have two cases: on the set $\{< \int_0^\cdot (\nabla V(Y_{G^{-1}}(s)), \mathrm{d}B_s) >_\infty < \infty\}$, this (local) martingale is bounded in L^2 and thus converges, whereas on the other set $\{< \int_0^\cdot (\nabla V(Y_{G^{-1}}(s)), \mathrm{d}B_s) >_\infty = \infty\}$, we have the a.s. equality $\int_0^t (\nabla V(Y_{G^{-1}}(s)), \mathrm{d}B_s) = o\left(< \int_0^\cdot (\nabla V(Y_{G^{-1}}(s)), \mathrm{d}B_s) >_t\right)$. Therefore for t large enough, we get

$$\int_0^t (\nabla V(Y_{G^{-1}}(s)), \mathrm{d}B_s) \leq \frac{1}{2}\int_0^\cdot |\nabla V(Y_{G^{-1}}(s))|^2 \mathrm{d}s$$

and the Itô formula (5.4) implies for t large enough

$$\begin{aligned} \int_0^t |\nabla V(Y_{G^{-1}(s)})|^2 \mathrm{d}s &\leq 2\int_0^t g(s)|\nabla V(Y_{G^{-1}(s)})|^2 \mathrm{d}s \\ &\leq \int_0^t \Delta V(Y_{G^{-1}(s)})\mathrm{d}s - 2V(Y_{G^{-1}(t)}) + 2V(Y_{G^{-1}(0)}) \\ &- 2\int_0^t \frac{(Y_{G^{-1}(s)}, \nabla V(Y_{G^{-1}(s)}))}{(r+G^{-1}(s))g(G^{-1}(s))}\mathrm{d}s \\ &\leq at + b\varphi(t) + 2V(Y_{G^{-1}(0)}) = O(t) + b\varphi(t). \end{aligned}$$

We use this result for the inequality (6.4) and we choose ε small enough such that there exist $C_1, C_2 > 0$ with $\varphi(t) \leq C_1 t + C_2 V(Y_{G^{-1}(0)})$ that is $\varphi(t) = O(t)$. □

THEOREM 5.6. *The process Y satisfies the pointwise ergodic theorem. More precisely, there exists some deterministic $a_i, b_i \geq 0$ such that μ_t converges (for the weak convergence of measures) toward* $\sum_{1\leq i\leq n} a_i\delta_{m_i} + \sum_{1\leq i\leq p} b_i\delta_{M_i}$.

PROOF. The proof of this result is divided in several parts. We consider the time-changed process $Y_{G^{-1}(t)}$. We have introduce the dynamical system ϕ of the proposition 5.4 and prove that $Y_{G^{-1}(t)}$ is close to the flow induced by the dynamical system.
We recall the result of the lemma 5.4: for all $T > 0$ and $\alpha > 0$ that

$$\lim_{t\to\infty} \mathbb{P}\left(\sup_{0\leq h\leq T} |Y_{G^{-1}(t+h)} - \phi_h(Y_{G^{-1}(t)})| \geq \alpha\right) = 0.$$

The main result of Benaïm & Schreiber [**5**] implies that the limit points of the normalized occupation measure of $Y_{G^{-1}(\cdot)}$ are included in the set of all the “invariant measures” for the equation $\frac{\mathrm{d}}{\mathrm{d}t}\phi_t(x) = -\nabla V(\phi_t(x))$ with the initial condition $\phi_0(x) = x$. But all these invariant measures are included in $Vect\{\delta_{m_1}, \cdots, \delta_{m_n}, \delta_{M_1}, \cdots, \delta_{M_p}\}$. Therefore, if we have only a local minimum, we are done and we have the convergence. Else, let $\mu_t^{G^{-1}}$ be the normalized occupation measure of the time-changed process $Y_{G^{-1}}$ defined at the beginning of the section. The lemma 5.5 asserts that: $(\mu_t^{G^{-1}}, t \geq 0)$ is a tight family of measures.
So we now have proved that the normalized occupation measure of $Y_{G^{-1}}$ converges a.s. to $\sum_{i=1}^{n} a_i\delta_{m_i} + \sum_{i=1}^{p} b_i\delta_{M_i}$ (where a_i, b_i are some constants) and the last step is to show that the same result holds for Y.
For all continuous bounded function ψ and $t > q$, by an integration by parts, we have

$$\int_q^t \psi(Y_s)\mathrm{d}s = \frac{G(t)}{g(t)}\mu_{G(t)}^{G^{-1}}\psi - \frac{G(q)}{g(q)}\mu_{G(q)}^{G^{-1}}\psi + \int_q^t \frac{g'(s)G(s)}{g^2(s)}\mu_{G(s)}^{G^{-1}}\psi\mathrm{d}s.$$

But we have $\int_q^t \frac{g'(s)G(s)}{g^2(s)}\mathrm{d}s = -\frac{G(t)}{g(t)} + \frac{G(q)}{g(q)} + t - q$. It implies that

$$\begin{aligned}\int_q^t \psi(Y_s)\mathrm{d}s &= (t-q)\mu_{G(t)}^{G^{-1}}\psi + \frac{G(q)}{g(q)}\left(\mu_{G(t)}^{G^{-1}}\psi - \mu_{G(q)}^{G^{-1}}\psi\right) \\ &+ \int_q^t \frac{g'(s)G(s)}{g^2(s)}(\mu_{G(s)}^{G^{-1}}\psi - \mu_{G(t)}^{G^{-1}}\psi)\mathrm{d}s.\end{aligned}$$

As $\mu_{G(t)}^{G^{-1}}\psi$ converges a.s., we deduce that

$$\mu_t\psi = o(1) + \mu_{G(t)}^{G^{-1}}\psi + \frac{1}{t}\int_q^t \frac{g'(s)G(s)}{g^2(s)}(\mu_{G(s)}^{G^{-1}}\psi - \mu_{G(t)}^{G^{-1}}\psi)\mathrm{d}s.$$

Using the integration by parts, we easily see that $\frac{1}{t}\int_q^t \frac{g'(s)G(s)}{g^2(s)}\mathrm{d}s$ is bounded and we are done. □

Remark 5.7. *We deeply believe that $b_i = 0$ for all i.*

5.2. Almost sure convergence. We will prove that the process Y converges a.s. towards a minimum. Let $0 < \varepsilon < \varepsilon_0$ and $T > T_0$ be as in the previous part. Let m be a local minimum of V such that $|Y_T - m| < \varepsilon$.

Lemma 5.8. *If $\lim_{t\to\infty} g(t)^{-1}\log G(t) = 0$, then for all $c > 0$, we get $\int_0^\infty e^{-cg(t)}\mathrm{d}t < +\infty$.*

Proof. For all $\varepsilon > 0$, there exists t large enough, such that for all $s \geq t$, we have $g(s)/\log G(s) \geq \varepsilon^{-1}$. Moreover, we know that there exists a positive constant a such that for t large enough $g(t) \geq a$ and then $G(t) \geq at$. As a consequence, we get $g(t) \geq \varepsilon^{-1}\log(at)$. We now conclude the proof $\int_1^\infty e^{-(c\log(at))/\varepsilon}\mathrm{d}t < \infty$ (for example, we choose $\varepsilon = c/2$). □

Proposition 5.9. *If $g(t)^{-1}\log G(t)$ converges to 0, then Y_t converges a.s. and for all i, we have $\mathbb{P}\left(\lim_{t\to\infty} Y_t = m_i\right) > 0$ and $\mathbb{P}\left(\lim_{t\to\infty} Y_t = M_i\right) = 0$.*

Proof. We begin to prove that Y converges a.s. by using the result of Benaïm ([**1**] proposition 4.6). It asserts that if $F(x) = -\nabla V(x)$ is a continuous globally integrable vector field, if for all $c > 0$, we have that $\int_0^\infty e^{-cg\circ G^{-1}(t)}\mathrm{d}t < +\infty$ and $\mathbb{P}(\sup_t |Y_t| < \infty) = 1$, then Y is almost surely an asymptotic pseudo-trajectory for the flow induced by F. Actually, the first and last conditions are easily satisfied under our hypothesis. Moreover, as G^{-1} is a nondecreasing function, the integral $\int_0^\infty e^{-cg(t)}\mathrm{d}t$, which is finite, is a upper bound for the preceding integral. As a consequence, the process Y is an asymptotic pseudo-trajectory for the flow Φ defined by (2.8). Thus the set of the limit points of Y is an attractor free set. Finally, Y converges almost surely and the limit points of Y are included in the set $\{x; \nabla V(x) = 0\}$.

If Y converges to Y_∞, then the limit-process Y_∞ is not a local maximum M_i because of the proposition 3.3. We work on the event $\{\forall s \geq T; |Y_s - m_i| < \varepsilon\}$. We recall the proposition 4.3. We have a.s. that $V_t \leq Y_{t+T} - m_i \leq U_t$ with $\limsup_{t\to\infty} U_t = -\liminf_{t\to+\infty} V_t = \lim_{t\to\infty}\sqrt{g(t)^{-1}\log G(t)}$. It then holds $\overline{\lim_{t\to\infty}} U_t / \sqrt{\frac{\log G(t)}{g(t)}} = 1$ a.s. and thus $U_t \xrightarrow[t\to\infty]{a.s.} 0$. The same holds for V_t and the result follows. $\square$

Corollary 5.10. *If V is a strictly uniformly convex function everywhere (with a unique minimum m), then $Y_t \xrightarrow[t\to\infty]{a.s.} m$.*

Proof. One more time we consider the one-dimensional case and $m = 0$. In this case we mimic step by step the proof of the proposition 4.3 in order to find the a.s. inequalities for all $t \geq 0$

$$|Y_t - m| \leq |U_t| + |V_t|, \tag{5.5}$$

and we use the law of the iterated logarithm. $\square$

Remark 5.11. *On one hand, if there exists a local minimum $m \neq 0$, then $\mathbb{P}(\overline{\mu}_t$ converges$) < 1$. On the other hand, if 0 is a local minimum, then $\overline{\mu}_t$ converges on the set $\{\left|\int_0^t Y_s \frac{\mathrm{d}s}{r+s}\right| < \infty\}$.*

6. Behavior of X in the case of a general potential

6.1. Ergodicity of X. For the moment, we have proved that Y satisfies the pointwise ergodic theorem. The main question of this paper is to know whether X also satisfies the pointwise ergodic theorem or not. The answer is naturally: it depends on the function g! Nevertheless, the following results give necessary and sufficient conditions for the ergodic theorem for X.

Remark 6.1. *The process $\overline{\mu}_t$ converges a.s. if and only if $\int_0^t Y_s \frac{\mathrm{d}s}{r+s}$ converges. In particular, if $Y_t \xrightarrow[t\to\infty]{a.s.} 0$ polynomially fast then the process $\overline{\mu}_t$ converges a.s.*

A necessary condition for the almost sure convergence of Y to 0 is to consider a potential V with a unique minimum 0 (for instance symmetric and convex).

PROPOSITION 6.2. *The process X satisfies the pointwise ergodic theorem if and only if $\overline{\mu}_t$ converges a.s.*

PROOF. We recall that $X_t = Y_t + \overline{\mu}_t$. We have shown in the preceding section that Y always satisfies the pointwise ergodic theorem (as we have proved it before). To conclude the proof, we will use the Fourier transform of the normalized occupation of X. We have for all $u \in \mathbb{R}^d$:

$$\begin{aligned}\frac{1}{t}\int_0^t e^{i(u,X_s)}\mathrm{d}s &= \frac{e^{i(u,\overline{\mu}_\infty)}}{t}\int_0^t e^{i(u,Y_s)}\mathrm{d}s \\ &+ \frac{1}{t}\int_0^t e^{i(u,Y_s)}\left(e^{i(u,\overline{\mu}_s)} - e^{i(u,\overline{\mu}_\infty)}\right)\mathrm{d}s\end{aligned}$$

The first right member converges a.s. to $e^{i(u,\overline{\mu}_\infty)}\int e^{i(u,y)}\gamma(\mathrm{d}y)$. For the second right member, we use the Cesàro result to prove that it converges a.s. to 0. We can now conclude that X satisfies the pointwise ergodic theorem. □

6.2. Almost sure convergence. In order to study the asymptotic behavior of $(X_t, t \geq 0)$, we will consider the process Y defined by $Y_t = X_t - \overline{\mu}_t$.

THEOREM 6.3. *Suppose that $\sqrt{g(t)^{-1}\log G(t)} = o((\log t)^{-1})$. One of the following holds:*

(1) *If 0 is the unique local minimum of V then*
$$\mathbb{P}\left(X_t \to \overline{\mu} + \int_0^\infty Y_s \frac{\mathrm{d}s}{r+s}\right) = \mathbb{P}\left(\overline{\mu}_t \to \overline{\mu} + \int_0^\infty Y_s \frac{\mathrm{d}s}{r+s}\right) = 1;$$

(2) *If 0 is a local minimum of V and there exists other local minima, then*
$$\mathbb{P}\left(X_t \to \overline{\mu} + \int_0^\infty Y_s \frac{\mathrm{d}s}{r+s}\right) + \mathbb{P}(|X_t| \to \infty) = 1$$
and
$$1 > \mathbb{P}\left(X_t \to \overline{\mu} + \int_0^\infty Y_s \frac{\mathrm{d}s}{r+s}\right) > 0,\ \ 1 > \mathbb{P}\left(\overline{\mu}_t \to \overline{\mu} + \int_0^\infty Y_s \frac{\mathrm{d}s}{r+s}\right) > 0;$$

(3) *If 0 is not a local minimum of V, then*
$$\mathbb{P}(|X_t| \to \infty) = \mathbb{P}(\overline{\mu}_t \to \infty) = 1.$$

PROOF. We only have to recall that $Y_t = X_t - \overline{\mu_t}$ and apply the proposition 4.2 added to the inequalities (5.5) to find:

- If $m = 0$ then both $\overline{\mu}_t$ and X_t converge a.s.

- If $m \neq 0$ then $\overline{\mu}_t \overset{a.s.}{\to} sgn(m)\infty$ and X_t does not converge a.s.

$\square$

REMARK 6.4. *The* $\lim_{t\to\infty} g(t)^{-1}\log G(t) = 0$, *is for instance satisfied for* $g(t) = t^{\beta}$, *with* $\beta > 0$, *or* $g(t) = e^{t}$.

Bibliography

[1] Benaïm M. [1999], Dynamics of stochastic approximation algorithms, *Sém. Prob. XXXIII, Lecture Notes in Math.* **1709**, 1-68, Springer.

[2] Benaïm M., Ledoux M. *&* Raimond O. [2002], Self-interacting diffusions, *Prob. Theory Relat. Fields* **122**, 1-41.

[3] Benaïm M. *&* Raimond O. [2003], Self-interacting diffusions II: convergence in law, *An. Inst. H. Poincaré, Prob. & Stat.* **39**(6), 1043-1055.

[4] Benaïm M. *&* Raimond O. [2005], Self-interacting diffusions III: symmetric interactions, *Ann. Prob.* **33**(5), 1716-1759.

[5] Benaïm M. *&* Schreiber S.J. [2000], Weak asymptotic pseudotrajectories for semiflows: ergodic properties, *J. Dynamics and Diff. Eq.* **12**(3).

[6] Cattiaux P. *&* Guillin A. [2006], Deviation bounds for additive functionnals of Markov processes, *Preprint.*

[7] Cranston M. *&* Le Jan Y. [1995], Self-attracting diffusions: two cases studies, *Math. Ann.* **303**, 87-93.

[8] Cranston M. *&* Mountford T.S. [1996], The strong law of large numbers for a Brownian polymer, *Ann. Prob.* **2**(3), 1300-1323.

[9] Duflo M. [1996], *Algorithmes stochastiques*, Springer "Mathématiques & Applications 23".

[10] Durrett R.T. *&* Rogers L.C.G. [1992], Asymptotic behavior of Brownian polymers, *Prob. Th. Rel. Fields* **92**(3), 337-349.

[11] Harrison J.M. *&* Shepp L.A. [1981], On the skew Brownian motion, *Ann. Prob.* **9**(2), 309-313.

[12] Herrmann S. *&* Roynette B. [2003], Boundedness and convergence of some self-interacting diffusions, *Math. Ann.* **325**(1), 81-96.

[13] Lépingle D. [1978], Sur le comportement aymptotique des martingales locales, *Sém. Prob. (Strasbourg)* **12**, 148-161.

[14] Mountford T.S. *&* Tarrès P. [2005], An asymptotic result for Brownian polymers, preprint.

[15] Norris J.R., Rogers L.C.G. *&* Williams D. [1987], Self-avoiding random walk: a Brownian motion model with local time drift, *Prob. Th. Rel. Fields* **74**(2), 271-287.

[16] Pemantle R. [2006], A survey of random processes with reinforcement, preprint.

[17] Raimond O. [1997], Self-attracting diffusions: case of constant interaction, *Prob. Theory Relat. Fields* **107**, 177-196.

[18] Raimond O. [2006], Self-interacting diffusions: a simulated annealing version, preprint.

[19] Revuz D. *&* Yor M. [1998], *Continuous Martingales and Brownian Motion*, 3rd edition, Springer.

[20] Rogers L.C.G. *&* Williams D. [2000], *Diffusions, Markov processes and Martingales*, 2nd edition, vol. 2 "Itô Calculus", Cambridge Univ. Press.

CHAPITRE 2

Convergence en loi de certaines diffusions auto-interactives : la méthode du recuit simulé

Le titre original de cet article, écrit en anglais, est "Convergence in law of some particular self-interacting diffusions : the simulated annealing method". Il s'agit d'un complément de l'étude menée au chapitre précédent.

1. Introduction

In a preceding paper ([**4**]), we have obtained some necessary and sufficient conditions for both the pointwise ergodicity and the almost sure convergence of some self-interacting diffusions. We will go further in the study of such processes. The aim of this paper is to obtain necessary and sufficient conditions for the convergence in distribution of the self-interacting diffusion X defined by

$$\mathrm{d}X_t = \mathrm{d}B_t - g(t)\nabla V(X_t - \overline{\mu}_t)\mathrm{d}t, \;\; X_0 = x \tag{1.1}$$

where B is a standard Brownian motion and $\overline{\mu}_t$ denotes the empirical mean of the process X, defined by

$$\overline{\mu}_t = \frac{1}{r+t}\left(r\bar{\mu} + \int_0^t X_s \mathrm{d}s\right), \;\; \overline{\mu}_0 = \overline{\mu}. \tag{1.2}$$

Here μ is an initial (given) probability measure on $\mathbb{R}^d$, $\bar{\mu}$ denotes the mean of μ and $r > 0$ is an initial weight.

The idea is to use the well-known theory of simulated annealing, which has been developed a lot since the 80's. An important question for physical systems is to find the globally minimum energy states of the system. Experimentally, the ground states are reached by a procedure, the chemical annealing. One first melts a substance and then cool it slowly, being careful to pass slowly through the freezing temperature. If the temperature decreases too rapidly, then the system does not end up into a ground state, but in a local (but not global) minimum. On the other hand, if the temperature decreases too slowly, then the system approaches the ground states, but very slowly. The competition between these two effects determines the optimal speed of cooling, that is the annealing schedule.

The study of simulated annealing has (at least!) involved the theory of (homogeneous and) non-homogeneous Markov chains and diffusion processes, large deviation theory, spectral analysis of operator and singular perturbation theory.

Pioneer work was done by Freidlin an Wentzell [**7**]. The initial problem consists in finding the global minima of a certain function U. Actually, one has to study the diffusion Markov process X^ε in $\mathbb{R}^d$ given by the Langevin-type Markov diffusion (we emphasize here that $\varepsilon = \varepsilon(t)$ in the SDE)

$$\mathrm{d}X_t^\varepsilon = \varepsilon \mathrm{d}B_t - \nabla U(X^\varepsilon)\mathrm{d}t. \tag{1.3}$$

Here, X^ε may be considered as a perturbation of the trajectory X^0 of the dynamical system $\frac{\mathrm{d}X_t^0}{\mathrm{d}t} = -\nabla U(X_t^0)$. The idea is the following. If the temperature ε is constant for a sufficiently large amount of time, then the process X^ε and the fixed temperature process behave approximatively the same at the end of that time interval. We denote by Min the set of all the global minima of U. The optimal annealing schedule, that is ε, for the convergence criterion $\mathbb{P}(X_t^\varepsilon \in Min) \underset{t\to\infty}{\longrightarrow} 1$ was first determined by Hajek [**11**] for a finite state space. Chiang, Hwang and Sheu [**5**] studied the convergence rate of $\mathbb{P}_x(X_t^\varepsilon \in \cdot)$ *via* the large deviations of the transition density of X^ε. It is actually strongly related to the spectral gap of the invariant measure associated to the process X^ε.

Chiang, Wang and Sheu were one of the firsts to show the convergence of the algorithm of the simulated annealing, for $\varepsilon(t)^2 = k/\log t$ for k large enough. Later, Royer [**21**] obtained the same result for $k > \Lambda$, where Λ is a constant related to the second eigenvalue λ_2^ε of the corresponding (to X^ε) infinitesimal generator. We remark that the result is optimal in the sense that X^ε does not converge in probability for $k \leq \Lambda$. Moreover, Hwang and Sheu [**13**] established (by probabilistic methods) the existence of $\Lambda := \lim\limits_{\varepsilon\to 0} -\varepsilon^2 \log \lambda_2^\varepsilon$. Finally, Holley and Stroock [**14**] initiated an other method and proved, in the discrete case, the convergence of the simulated annealing algorithm via the Sobolev inequality. They went further in their study with Kusuoka [**12**]. Later, Miclo [**20**] proved, by using some functional inequalities, that the free energy (that is the relative entropy of the distribution of the process at time t with respect to the invariant probability at that time t) satisfies a differential inequality, which implies (under some decreasing evolution of the temperature to zero) the convergence of the process to the global minima of the potential.

A natural question arises: what happens if the temperature, that is ε, decreases too fast to zero? Then, the potential can freeze in a local minimum (the "choice" of this minimum depends on the initial condition) and therefore the process converges to this local minimum.

As it is now usual, we begin to study the process $Y_t := X_t - \bar{\mu}_t$, which satisfies the following SDE

$$\begin{cases} \mathrm{d}Y_t = \mathrm{d}B_t - g(t)\nabla V(Y_t)\mathrm{d}t - Y_t\frac{\mathrm{d}t}{r+t}; \\ Y_0 = x - \overline{\mu}; \\ \mathrm{d}\overline{\mu}_t = Y_t\frac{\mathrm{d}t}{r+t}. \end{cases} \tag{1.4}$$

We will adapt the simulated annealing method to Y. We will prove that, depending on g, either the process Y converges in probability towards a variable which is concentrated on the global minima of V or converges in distribution.

The paper is organized in the following way. In Section 2, we recall the notations and hypothesis. The most important results of the preceding paper are reminded in the Section 3. The fourth section is devoted to the study of the process Y. In particular, we prove the convergence of Y in probability towards global minima thanks to the simulated annealing method and its convergence in distribution towards local minima. Finally, we deduce in the Section 5 some necessary and sufficient conditions for the convergence of the self-interacting process X.

2. Notation, hypothesis and existence

We recall briefly the notation of Chambeu *&* Kurtzmann [**4**] which will be useful in the following. We denote by G the function $G(t) = \int_0^t g(s)\mathrm{d}s$. In the whole following, $(\cdot,\cdot)$ denotes the Euclidian scalar product. We denote by $\mathcal{P}(\mathbb{R}^d)$ the set of probability measures on $\mathbb{R}^d$.

In the sequel, the technical assumptions on the potential $V : \mathbb{R}^d \to \mathbb{R}_+$ are the following:

(1) (*regularity and positivity*) $V \in \mathcal{C}^2(\mathbb{R}^d)$ and $V \geq 0$;
(2) (*convexity*) V is strictly uniformly convex out of a compact set K and $\lim_{|x|\to+\infty} V(x) = +\infty$;
(3) (*growth*) there exist $a, b > 0$ such that for all $x \in \mathbb{R}^d$, we have

$$\Delta V(x) \leq a + bV(x). \tag{2.1}$$

We also assume that V has a finite number of critical points. Let $Max = \{M_1, M_2, \cdots, M_p\}$ be the set of the saddle points and local maxima of V and $Min = \{m_1, m_2, \cdots, m_n\}$ be the set of the local minima of V, such that for all i, we have $\det\nabla^2 V(m_i) \neq 0$. Without any loss of generality, we suppose that $\min\limits_{x\in\mathbb{R}^d} V(x) = 0$. We assume that $\forall i$, $\forall \xi \in \mathbb{R}^d$, $(\nabla^2 V(m_i)\xi, \xi) > 0$. The convexity assumption implies that we can decompose $V = W + \chi$ where W is a strictly convex function and χ is a compactly supported function: there exists $c > 0$ and $\tilde{C} >$ such that $\nabla^2 W(x) \geq cId > 0$ and $\nabla\chi$ is $\tilde{C}$-Lipschitz.

We also assume that the application $g : \mathbb{R}_+ \to \mathbb{R}_+$ satisfies:

(1) $g \in \mathcal{C}^1(\mathbb{R}_+)$ is such that $\lim\limits_{t\to\infty} g(t) = +\infty$ and $g(0) > 0$;

(2) for all $T > 0$, $G^{-1}(t+T) - G^{-1}(t) \underset{t\to\infty}{\longrightarrow} 0$ where G^{-1} is the generalized inverse of G;

(3) $t\frac{g'(t)}{g(t)}$ converges when t tends to the infinity.

We have already shown that the SDE (1.1) studied admits a unique global strong solution:

PROPOSITION 2.1. *([**4**] proposition 2.4) For any $x \in \mathbb{R}^d$, $\mu \in \mathcal{P}(\mathbb{R}^d)$ and $r > 0$, there exists a unique global strong solution $(X_t, t \geq 0)$ of (1.1).*

3. Former results

It has been proved in the preceding paper [**4**] the following results. In the whole section, the process Y is the solution to (1.4).

PROPOSITION 3.1. *(proposition 4.1) With probability 1, the process Y gets close to the set $Min \cup Max$, that is $\forall \varepsilon > 0$, let $T_t^\varepsilon := \inf\{s \geq t; d(Y_s, Min \cup Max) < \varepsilon\}$. Then, for all $t > 0$, we have $T_t^\varepsilon < \infty$ almost surely.*

PROPOSITION 3.2. *(proposition 4.3) Suppose that $t = o(g(t))$. Let $\varepsilon_0 > \varepsilon > 0$ and Y the solution to (1.4). If there exists $T_0 > 0$ such that, for all $T > T_0$, $|Y_T - m| < \varepsilon$, then the event $\{\forall s \geq T; |Y_s - m| < \varepsilon\}$ has a positive probability to occur. Moreover, almost surely, on the event $\{\forall s \geq T; |Y_s - m| < \varepsilon\}$, we have*

$$|Y_{t+T} - m| = O\left(\sqrt{g(t+T)^{-1}\log G(t+T)}\right).$$

PROPOSITION 3.3. *(proposition 4.4) Let $0 < \varepsilon < \varepsilon_1$, M_i a local maximum of V and T a positive stopping time such that we have $|Y_T - M_i| < \varepsilon$. Then*

$$\mathbb{P}\left(\forall s \geq T;\ |Y_s - M_i| < \varepsilon\right) = 0$$

THEOREM 3.4. *(theorem 5.6) The process Y satisfies the pointwise ergodic theorem. This means that with probability 1, the normalized occupation measure of Y converges weakly to a random measure, and what is more, this last measure is a convex combination of Dirac measures taken in the critical points of V. More precisely, there exist $a_i, b_i \geq 0$ such that*

$$\frac{1}{t}\int_0^t \delta_{Y_s}\mathrm{d}s \underset{t\to\infty}{\longrightarrow} \sum_{i=1}^n a_i\delta_{m_i} + \sum_{i=1}^p b_i\delta_{M_i} \ \ a.s.$$

4. Asymptotic behavior of Y

Instead of considering Y, we are interested in the time-changed process $Z_t := Y_{G^{-1}(t)}$. This last process satisfies the following SDE

(4.1)
$$\mathrm{d}Z_t = \frac{1}{\sqrt{g \circ G^{-1}(t)}}\mathrm{d}W_t - \left(\nabla V(Z_t) + \frac{Z_t}{(r + G^{-1}(t))g \circ G^{-1}(t)}\right)\mathrm{d}t,\ \ Z_0 = z,$$

where W is a Brownian motion such that $\frac{1}{\sqrt{g\circ G^{-1}(t)}}\mathrm{d}W_t$ has the same law as $B_{G^{-1}(t)}$[1].

4.1. Tightness. We begin to prove that the law of the process Z is a tight family of measures.

LEMMA 4.1. *(Chiang-Hwang-Sheu) There exist two constants $R, c > 0$ such that*

$$\mathbb{E}\left(V(Z_t)1\!\!1_{V(Z_t)\geq R}\right) \leq \frac{c}{g\circ G^{-1}(t)}.$$

PROOF. We reproduce the proof given by Duflo [**6**]. There exists r such that, for $V(x) \geq r$, we get $|\nabla V(x)|^2 \geq 2cV(x)$ where $c > 0$ is a constant. Let ϕ be a non-negative and non-decreasing function of class $\mathcal{C}^2$, $\phi : \mathbb{R} \to [0,1]$ such that $\phi(x) = 0$ for $x \leq r$, $\phi(x) = 1$ for $x \geq R$ where $r < R < \infty$. Then, the continuous function $\nabla(\phi\circ V) = (\phi'\circ V)\nabla V$ is bounded (because its support is compact). Consider the application $\psi := (\phi\circ V)V$. We apply the Itô formula to the function $x \mapsto \psi(x)$ and we get

$$\begin{aligned}
\mathrm{d}\psi(Z_t) &= \sqrt{(g\circ G^{-1}(t))^{-1}}(\nabla\psi(Z_t), \mathrm{d}B_t) + \frac{1}{2g\circ G^{-1}(t)}\Delta\psi(Z_t)\mathrm{d}t \\
&- \left(\phi\circ V(Z_t) + V(Z_t)\phi'\circ V(Z_t)\right)\times \\
&\times \left(|\nabla V(Z_t)|^2 + \frac{(Z_t, \nabla V(Z_t))}{(r+G^{-1}(t))g\circ G^{-1}(t)}\right)\mathrm{d}t.
\end{aligned}$$

Let $\alpha(t) := \mathbb{E}[\phi\circ V(Z_t)V(Z_t)]$. Following the lines of the preceding paper (lemma 4.8), we obtain that $\mathbb{E}V(Z_t) = O(1)$ and therefore everything is well-defined. We recall that $0 \leq \phi \leq 1$ is bounded and ϕ' is a compactly supported function. Because of the growth assumption (3.1) on V, we easily obtain that there exists a positive constant c such that $\Delta\psi \leq c(\phi\circ V)V$. Therefore, the preceding formula leads to

$$\begin{aligned}
\alpha'(t) &= \frac{1}{2g\circ G^{-1}(t)}\mathbb{E}\Delta\psi(Z_t) \\
&- \mathbb{E}\left[\left(\phi\circ V(Z_t) + V(Z_t)\phi'(Z_t)\circ V(Z_t)\right)\times\right. \\
&\times \left.\left(|\nabla V(Z_t)|^2 + \frac{(Z_t, \nabla V(Z_t))}{(r+G^{-1}(t))g\circ G^{-1}(t)}\right)\right] \\
&\leq \frac{c}{2g\circ G^{-1}(t)}\alpha(t) + M((r+G^{-1}(t))g\circ G^{-1}(t))^{-1} \\
&- \mathbb{E}\left(\phi\circ V(Z_t)|\nabla V(Z_t)|^2\right) \\
&- \mathbb{E}\left(\phi\circ V(Z_t)\frac{(Z_t, \nabla V(Z_t))}{(r+G^{-1}(t))g\circ G^{-1}(t)}\right).
\end{aligned}$$

We recall that $V = W + \chi$, where W is strictly convex, this implies that there exists a constant $a > 0$ such that $(x, \nabla W(x)) \geq a|x|^2$. Moreover, $\nabla\chi$

[1]The Wiener processes B_t and W_t are not the same, but this does not matter because we are only interested in the probability distributions.

is supposed to be Lipschitz, with a constant $\widetilde{C} > 0$. Therefore, we get that there exists a positive constant M such that

$$\begin{aligned}
\alpha'(t) &\leq \frac{c}{2g \circ G^{-1}(t)}\alpha(t) + M((r+G^{-1}(t))g \circ G^{-1}(t))^{-1} - 2c\alpha(t) \\
&+ \widetilde{C}\mathbb{E}\left(\phi \circ V(Z_t)\frac{|Z_t|^2}{(r+G^{-1}(t))g \circ G^{-1}(t)}\right) \\
&\leq \frac{c}{2g \circ G^{-1}(t)}\alpha(t) + M((r+G^{-1}(t))g \circ G^{-1}(t))^{-1} - 2c\alpha(t) \\
&+ \frac{m}{(r+G^{-1}(t))g \circ G^{-1}(t)}\alpha(t).
\end{aligned}$$

Thus, for t large enough, we have $\alpha'(t) \leq -c\alpha(t) + \frac{\widetilde{C}}{2}(g \circ G^{-1}(t))^{-1}$. In order to solve this inequality, let $\alpha(t) := \beta(t)e^{-ct}$. This yields to $\alpha(t) \leq \frac{\widetilde{C}}{4cg \circ G^{-1}(t)}$. To conclude, we just need to remark that $\mathbf{1}_{\{V \geq R\}} \leq \phi \circ V$ and finally the result follows. $\square$

Corollary 4.2. *The processes Y and Z are tight.*

Proof. The previous proposition implies that $\mathbb{E}(V(Z_t)) = O(1)$. Indeed

$$\mathbb{E}V(Z_t) = \mathbb{E}(V(Z_t)\mathbb{1}_{V(Z_t)<R}) + \mathbb{E}(V(Z_t)\mathbb{1}_{V(Z_t)\geq R}) \leq R + \frac{c}{g(0)}.$$

Then, for a positive constant A, there exists a compact set K such that $\{V \leq A\} \subset K$ and then

$$\mathbb{P}(Z_t \in K) \geq \mathbb{P}(V(Z_t) \leq A) \geq 1 - \frac{\mathbb{E}(V(Z_t))}{A} \underset{A\to\infty}{\to} 1.$$

$\square$

4.2. Convergence in distribution towards the global minima of V. We define $\varepsilon^2(t) = \frac{1}{g \circ G^{-1}(t)}$ and $a(t) = (r + G^{-1}(t))\varepsilon^{-2}(t)$. The process Z reads

$$\mathrm{d}Z_t = \varepsilon(t)\mathrm{d}B_t - \nabla V_t(Z_t)\mathrm{d}t \tag{4.2}$$

where we have defined $V_t(x) := V(x) + \frac{|x|^2}{a(t)}$. Actually, we will prove that this non-homogeneous Markov process converges in distribution to its "invariant" probability measure. Of course, if we suppose that $a(t) \equiv a$ and $\varepsilon(t) \equiv \varepsilon$, then the convergence in distribution is obvious. It happens that a crucial constant λ, the spectral gap, appears naturally in our study. Heuristically, when the time is of order $e^{\varepsilon^{-2}\lambda}$, the transition density has a nice lower bound and the process is very close to the invariant probability

$$\Pi_{t,\varepsilon}(\mathrm{d}x) := \frac{1}{Z(t,\varepsilon)}e^{-2\varepsilon^{-2}V_t(x)}\mathrm{d}x. \tag{4.3}$$

It remains to compute the convergence of t to the infinity.

Let $L_{t,\varepsilon}$ be the operator defined by $L_{t,\varepsilon} := \frac{1}{2}\varepsilon^2\Delta - (\nabla V_t, \nabla)$. As $|\nabla V_t|^2 - \Delta V_t$ is bounded from below, the theory of Schrödinger operator implies that $L_{t,\varepsilon}$ is self-adjoint in $L^2(\Pi_{t,\varepsilon})$. Moreover, under our hypothesis ($|\nabla V_t|^2 - \Delta V_t$ bounded from below), the operator $L_{t,\varepsilon}$ admits a spectral gap. Moreover, when $|\nabla V_t|^2 - \Delta V_t$ goes to the infinity as $|x| \to \infty$ then the spectrum of $L_{t,\varepsilon}$ is discrete: $0 = \lambda_1(t,\varepsilon) < -\lambda_2(t,\varepsilon) < \ldots$. The subspace corresponding to the first eigenvalue $\lambda_1(t,\varepsilon)$ is composed by the constant functions and therefore

$$\lambda_2(t,\varepsilon) = \inf\left\{\int |\nabla\phi|^2 \mathrm{d}\Pi_{t,\varepsilon}; \quad \mathrm{Var}_{\Pi_{t,\varepsilon}}(\phi) = 1, \phi \in \mathcal{D}(\mathbb{R}^d)\right\}.$$

Our first aim is to compute the eigenvalue λ_2 and study its behavior when $t \to \infty$ (that is $\varepsilon \to 0$).

Lemma 4.3. *Let $\varepsilon > 0$ be fixed. The probability measure $\Pi_{t,\varepsilon}$ converges weakly, as $t \to \infty$, to $\Pi_{\infty,\varepsilon}(\mathrm{d}x) := \frac{1}{Z(\varepsilon)} e^{-2\varepsilon^{-2}V(x)}\mathrm{d}x$.*

Proof. We just need to recall that $\varepsilon^2 a(t) = r + G^{-1}(t)$, which diverges with t. The end of the proof is straightforward. Moreover, the convergence in uniform in ε. We recall more explicitly that

$$Z(t,\varepsilon(t)) = \int_{\mathbb{R}^d} e^{-2\varepsilon^{-2}(t)V(x)} e^{-2\frac{|x|^2}{a(t)\varepsilon^2(t)}}\,\mathrm{d}x.$$

Let K be the compact set $K := \{x | V(x) \leq 1\}$. There exists a constant $A > 0$ such that $K \subset B(0,A)$. Then, on one hand, we get

$$\begin{aligned}\int_{K^c} e^{-2\varepsilon^{-2}(t)V(x)} e^{-2\frac{|x|^2}{a(t)\varepsilon^2(t)}}\,\mathrm{d}x &\leq \int_{\mathbb{R}^d} e^{-2\varepsilon^{-2}(t)} e^{-2\frac{|x|^2}{a(t)\varepsilon^2(t)}}\,\mathrm{d}x \\ &\leq C(a(t)\varepsilon^2(t))^{-d/2} e^{-2\varepsilon^{-2}(t)}.\end{aligned}$$

On the other hand we obtain,

$$\begin{aligned}\int_K e^{-2\varepsilon^{-2}(t)V(x)}\mathrm{d}x &\geq \int_K e^{-2\varepsilon^{-2}(t)V(x)} e^{-2\frac{|x|^2}{a(t)\varepsilon^2(t)}}\,\mathrm{d}x \\ &\geq \int_K e^{-2\varepsilon^{-2}(t)V(x)} e^{-2\frac{A}{a(t)\varepsilon^2(t)}}\,\mathrm{d}x.\end{aligned}$$

But we know (see [**15**] or [**23**]), by the Laplace formula, that

$$\int_K e^{-2\varepsilon^{-2}(t)V(x)}\mathrm{d}x \underset{t\infty}{\sim} \sum_i (2\pi\varepsilon^2(t))^{d/2} (\det\nabla^2 V(x_i))^{-1/2}$$

where $(x_i)_i$ are the whole global minima of V (we recall that it is a finite set). As a consequence,

$$Z(t,\varepsilon(t)) \underset{t\infty}{\sim} \sum_i (2\pi\varepsilon^2(t))^{d/2} (\det\nabla^2 V(x_i))^{-1/2}.$$

By the same method, if ϕ is a continuous function with a compact support which contains, for example, only the global minimum x_1, we have

$$\int_{\mathbb{R}^d} \phi(x) e^{-2\varepsilon^{-2}(t)V(x)} e^{-2\frac{|x|^2}{a(t)\varepsilon^2(t)}} \mathrm{d}x \underset{t\infty}{\sim} (2\pi\varepsilon^2(t))^{d/2} (\det\nabla^2 V(x_1))^{-1/2} \phi(x_1).$$

This gives the explicit form of Π_0. □

Consider for a moment $\Pi_{\infty,\varepsilon}$. We remark that V_t converges to V when t goes to the infinity. Then, Hwang [**15**] established that $\Pi_{\infty,\varepsilon}$ converges weakly when ε converges to zero. Let N the set of the global minima of V. Then, Hwang has proved the following:

- if $\lambda(N) > 0$ (where λ is the Lebesgue measure on $\mathbb{R}^d$), then $\Pi_{\infty,\varepsilon}$ converges weakly to $\frac{1}{\lambda(N)} \mathbb{1}_N \mathrm{d}x$;
- if $N = \{x_0\}$, then $\Pi_{\infty,\varepsilon}$ converges weakly to δ_{x_0};
- if $N = \{x_1, \ldots, x_m\}$ then $\Pi_{\infty,\varepsilon}$ converges weakly to

$$\frac{1}{\sum\limits_{1\leq i\leq m} (\det\nabla^2 V(x_i))^{-1/2}} \sum_{1\leq i\leq m} (\det\nabla^2 V(x_i))^{-1/2} \delta_{x_i};$$

- more generally, if N is the finite union of some smooth manifolds ($\mathcal{C}^3$), and each component is a compact smooth manifold, connected and has a dimension strictly lower than k, and the determinant of the Hessian (normal to N in $x \in N$) $\det\left(\frac{\mathrm{d}^2 V}{\mathrm{d}t^2}(x)\right)$ is not identically zero. Then, there exists a probability measure $\mathcal{M}$, on the highest dimensional manifolds, such that $\Pi_{\infty,\varepsilon}$ converges weakly to $\frac{1}{\int (\det\left(\frac{\mathrm{d}^2 V}{\mathrm{d}t^2}(x)\right))^{-1/2}\mathcal{M}(\mathrm{d}x)} (\det\left(\frac{\mathrm{d}^2 V}{\mathrm{d}t^2}(x)\right))^{-1/2}\mathcal{M}(\mathrm{d}x)$.

We adapt to our setting the results of Hwang in the following proposition.

PROPOSITION 4.4. *The probability measure $\Pi_{t,\varepsilon(t)}$ converges weakly to Π_0 as t goes to the infinity. Moreover, the probability measure Π_0 concentrates on the global minima of V.*

PROOF. The result of Hwang shows that the probability measure $\Pi_{\varepsilon(t)}$ converges weakly to Π_0 as t goes to the infinity, and the probability measure Π_0 concentrates on the global minima of V. We combine this result with the lemma 4.3 and the result follows. □

Let $z_0 \in \mathbb{R}^d$ such that $V_t(z_0) = 0$. Let K be the compact set corresponding to the support of χ.

DEFINITION 4.5. *The maximal height of the function V_t is the non-negative function $m(t)$ defined by*

$$m(t) := \sup\{H_t(x, z_0) - V_t(x); \quad x \in K\}, \tag{4.4}$$

where

$$\begin{aligned} H_t(x,z) &:= \inf\{E_t(\gamma); \quad \gamma \in \mathcal{C}^1([0,1],K); \gamma(0)=x, \gamma(1)=z\}, \\ E_t(\gamma) &:= \sup\{V_t(\gamma(u)); \quad u \in [0,1]\}. \end{aligned}$$

REMARK 4.6. *1) The maximal height function $m(t)$ does not depend on z_0 and we will choose $z_0 = 0$ in the following.*
2) The function $m(t)$ corresponds to the maximum of all the minimal energies one needs in order to go from each point of $\mathbb{R}^d$ to z_0.
3) The function $m(t)$ is positive if and only if there exists several local minima.

LEMMA 4.7. *We have that $\lim_{t\to\infty} m(t) = m$, where m is the maximal height function corresponding to V.*

PROOF. Let us $M := \sup_{x\in K}|x|^2$. It is clear that for all path γ, we have $E_t(\gamma) \le E(\gamma) + \frac{M}{a(t)}$. Then, by definition we get

$$|H_t(x,0) - H_\infty(x,0)| \le \frac{M}{a(t)}.$$

As a consequence, we get that there exists $C>0$ such that

$$|m(t) - m(\infty)| \le \sup\left\{\left|H_t(x,0) - H_\infty(x,0) - \frac{|x|^2}{a(t)}\right|; \quad x \in K\right\} \le \frac{C}{a(t)}$$

and the result follows. □

A very important theorem permits us to relate the height function to the second eigenvalue of the infinitesimal generator of Y^ε (that is the constant involved in the spectral gap inequality).

THEOREM 4.8. *(Jacquot, theorem 1.1)* $\lim_{\varepsilon\to 0}\varepsilon^2 \log\lambda_2(\infty,\varepsilon) = -2m(\infty)$.

In order to prove that the process Z converges in distribution towards the global minima of V, we need two more technical results. Among several possibilities, we follow the approach initiated by Holley and Stroock [**14**], Holley, Kusuoka and Stroock [**12**] and Miclo [**20**]. We suppose in the following that $g\circ G^{-1}(t)$ is asymptotically equivalent to $\frac{\log(1+t)}{c}$ for c small enough (we will give explicitly the constant c in the following).

DEFINITION 4.9. *The measure μ satisfies the logarithmic Sobolev inequality, with the constant c, denoted $LSI(c)$, if for all function $h \in L^2(\mu)$, we have*

$$\int h^2 \log h^2 \mathrm{d}\mu - \left(\int h^2 \mathrm{d}\mu\right)\log\left(\int h^2\mathrm{d}\mu\right) \le c\int |\nabla h|^2 \mathrm{d}\mu.$$

Let $p(s,x,t,y)$ denote the density of the semi-group corresponding to the non-homogeneous Markov process Z.

LEMMA 4.10. *The family of probability measures* $(\Pi_{t,\varepsilon(t)}, t \geq 0)$ *satisfies a logarithmic Sobolev inequality* $LSI(C(t))$, *where* $C(t) = 2e^{\frac{2}{\varepsilon(t)^2} osc\chi}/c$ *(where we recall that* $V = W + \chi$ *and* c *is the convexity constant of* W*).*

PROOF. The idea is to use the celebrated Bakry-Emery Γ_2 criterion [**2**]. We recall that, to the operator $L_{t,\varepsilon(t)}$, we can associate the operator "carré du champ", that is (for all function $f, g \in \mathcal{C}^\infty$)

$$\Gamma_t^V(f,g) := \frac{1}{2}\left(L_{t,\varepsilon(t)}(fg) - fL_{t,\varepsilon(t)}g - gL_{t,\varepsilon(t)}f\right). \tag{4.5}$$

Then, we define the operator Γ_2 as

$$\Gamma_2^V(t)(f) := \frac{1}{2}\left(L_{t,\varepsilon(t)}\Gamma_t^V(f,f) - 2\Gamma_t^V(f, L_{t,\varepsilon(t)}f)\right). \tag{4.6}$$

The Γ_2 criterion asserts that if there exists a positive constant C such that $\Gamma_2^{V_t} \geq C\Gamma_t^{V_t}$, then $\Pi_{t,\varepsilon(t)}$ satisfies a logarithmic Sobolev inequality, with the constant $2/C$ (see for instance Ané & al [**1**], chap.5).

An easy calculation, for all function f of class $\mathcal{C}^\infty$ leads to

$$\Gamma_t^V(f,f) = \frac{\varepsilon(t)^2}{2}|\nabla f|^2$$

and

$$\Gamma_2^V(t)(f) = \frac{\varepsilon(t)^2}{2}(\nabla f, \nabla^2 V \nabla f) + \frac{\varepsilon(t)^4}{4}||\nabla^2 f||^2 + \frac{\varepsilon(t)^2}{2a(t)}|\nabla f|^2.$$

As V (and also V_t) is a strictly convex function out of a compact set, we have the decomposition $V = W + \chi$ where W is strictly convex with a constant c and χ is a compactly supported function. We apply the criterion of Bakry and Emery to the function W and we get that $\Gamma_2^W(t)(f) \geq c\Gamma_t^W(f)$. Thus, the probability measure $e^{-2\varepsilon^{-2}(t)(W(x)+|x|^2/a(t))}/Z$ satisfies the inequality $LSI(2/c)$. We conclude, by the perturbation lemma due to Holley and Stroock [**13**], that the measure $\Pi_{t,\varepsilon(t)}$ satisfies a Sobolev logarithmic inequality with a constant less than $2e^{\frac{2}{\varepsilon(t)^2} osc\chi}/c$, where $osc(\chi) = \sup\chi - \inf\chi$. Thus the result follows. □

LEMMA 4.11. *The invariant measure* $\Pi_{t,\varepsilon}$ *admits a spectral gap: there exists a constant* $\lambda_2(\varepsilon) > 0$ *such that for all* $s \geq 0$, *all continuous* $f \in L^2(\Pi_{t,\varepsilon})$

$$||P_s^{t,\varepsilon} f - \Pi_{t,\varepsilon} f||_{L^2(\Pi_{t,\varepsilon})} \leq e^{-\lambda_2(t,\varepsilon)s}\mathrm{Var}_{\Pi_{t,\varepsilon}}(f).$$

PROOF. We still know that $\Pi_{t,\varepsilon(t)}$ satisfies the inequality $LSI(C(t))$. Therefore, we get the spectral inequality with $\lambda_2(t, \varepsilon(t))$. □

We want to use some functional inequalities in order to prove the convergence of Z_t (and thus Y_t) towards the global minima of V.

DEFINITION 4.12. *We define the free energy (up to an additive constant), also named relative Kullback information, of a measure P with respect to a measure Π by the following:*

$$H(P|\Pi) := \int \mathrm{d}P \log \frac{\mathrm{d}P}{\mathrm{d}\Pi}.$$

Equivalently, if we suppose that P, respectively Π, has a density p, respectively π, with respect to the Lebesgue measure λ, then we define

$$H(p|\Pi) := \int p \log \frac{p}{\pi} \mathrm{d}\lambda. \tag{4.7}$$

PROPOSITION 4.13. *For all initial t_0, x_0, we get*

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t} H\left(p(t_0,x_0,t,\cdot)|\Pi_{t,\varepsilon(t)}\right) \quad &\leq \quad -\frac{2}{C(t)}\varepsilon(t)^2 H\left(p(t_0,x_0,t,\cdot)|\Pi_{t,\varepsilon(t)}\right) \\ &- \quad 4\dot{\varepsilon}(t)\varepsilon(t)^{-3}\int p(t_0,x_0,t,\cdot)(V_t - < V_t >_{\Pi_{t,\varepsilon(t)}})\mathrm{d}\lambda \\ &+ \quad \frac{2}{\varepsilon(t)^2}\int p(t_0,x_0,t,\cdot)(\dot{V}_t - < \dot{V}_t >_{\Pi_{t,\varepsilon(t)}})\mathrm{d}\lambda.\end{aligned}$$

PROOF. We adapt the proof of Holley and Stroock [**14**] (and Miclo [**20**]). In order to shorten notation, let $p_t := p(t_0, x_0, t, \cdot)$ be the distribution law of the process Z_t, knowing that $Z_{t_0} = x_0$. We recall that the family of probability measures $(\Pi_{t,\varepsilon(t)}, t \geq 0)$ satisfies a Sobolev logarithmic inequality $LSI(C(t))$. We have $\Pi_{t,\varepsilon(t)}(\mathrm{d}x) = \pi_{t,\varepsilon(t)}(x)\lambda(\mathrm{d}x)$. Therefore, we choose the function

$$h_t = \sqrt{\frac{p_t}{\pi_{t,\varepsilon(t)}}},$$

where we remark that $\int h_t^2 \mathrm{d}\Pi_{t,\varepsilon(t)} = 1$.

As a consequence, we get by the lemma 4.10 that there exists a constant $C(t)$ such that

$$H(p_t|\Pi_{t,\varepsilon(t)}) = \int p_t \log \frac{p_t}{\pi_{t,\varepsilon(t)}} \mathrm{d}\lambda \leq C(t) \int |\nabla h_t|^2 \mathrm{d}\Pi_{t,\varepsilon(t)}. \tag{4.8}$$

It remains to compute the derivative of h_t. We find that

$$\nabla h_t = \frac{\sqrt{p_t}}{2\pi_{t,\varepsilon(t)}} \left(\frac{\nabla p_t}{p_t} + 2\frac{\nabla V_t}{\varepsilon(t)^2} \right).$$

We put this last estimate in the preceding inequality and thus

$$H(p_t|\Pi_{t,\varepsilon(t)}) \leq \frac{C(t)}{4} \int p_t \left| \frac{\nabla p_t}{p_t} + 2\frac{\nabla V_t}{\varepsilon(t)^2} \right|^2 \mathrm{d}\lambda.$$

We recall that we are looking for an inequality including the time-derivative of the free energy H. We have

$$\frac{\mathrm{d}}{\mathrm{d}t} H(p_t|\Pi_{t,\varepsilon(t)}) = \int \dot{p}_t \log \frac{p_t}{\pi_{t,\varepsilon(t)}} \mathrm{d}\lambda - \int p_t \frac{\dot{\pi}_{t,\varepsilon(t)}}{\pi_{t,\varepsilon(t)}} \mathrm{d}\lambda. \tag{4.9}$$

Our strategy is to find a upper bound for the both terms on the right hand side. The Kolmogorov forward equation reads

$$\dot{p}_t = \frac{1}{2}\varepsilon(t)^2\Delta p_t + (\nabla p_t, \nabla V_t) = \nabla\cdot\left(\frac{1}{2}\varepsilon(t)^2\nabla p_t + p_t\nabla V_t\right). \tag{4.10}$$

We also remark that we have the following estimates:

$$\frac{\dot{\pi}_{t,\varepsilon(t)}}{\pi_{t,\varepsilon(t)}} = 4\frac{\dot{\varepsilon}(t)}{\varepsilon(t)^3}\left(V_t - < V_t >_{\pi_{t,\varepsilon(t)}}\right) - \frac{2}{\varepsilon(t)^2}(\dot{V}_t - < \dot{V}_t >_{\Pi_{t,\varepsilon(t)}}), \tag{4.11}$$

where we have used the usual notation $< f >_{\Pi_{t,\varepsilon(t)}} = \int f \mathrm{d}\Pi_{t,\varepsilon(t)}$. Moreover, we also find

$$\frac{\nabla\pi_{t,\varepsilon(t)}}{\pi_{t,\varepsilon(t)}} = -2\frac{\nabla V_t}{\varepsilon(t)^2}. \tag{4.12}$$

We put the first estimate, given by the Kolmogorov equation, in the first formula. We integrate by parts and we finally use the Sobolev logarithmic inequality (4.8) to get

$$\begin{aligned}
\int \log\frac{p_t}{\pi_{t,\varepsilon(t)}}\dot{p}_t\mathrm{d}\lambda &= \int \log\frac{p_t}{\pi_{t,\varepsilon(t)}}\nabla\cdot\left(\frac{1}{2}\varepsilon(t)^2\nabla p_t + p_t\nabla V_t\right)\mathrm{d}\lambda \\
&= -\int\left(\frac{\nabla p_t}{p_t} - \frac{\nabla\pi_{t,\varepsilon(t)}}{\pi_{t,\varepsilon(t)}}\right)\left(\frac{1}{2}\varepsilon(t)^2\nabla p_t + p_t\nabla V_t\right)\mathrm{d}\lambda \\
&= -\int\left(\frac{\nabla p_t}{p_t} + 2\frac{\nabla V_t}{\varepsilon(t)^2}, \frac{1}{2}\varepsilon(t)^2\nabla p_t + p_t\nabla V_t\right)\mathrm{d}\lambda \\
&= -\frac{\varepsilon(t)^2}{2}\int p_t\left|\frac{\nabla p_t}{p_t} + 2\frac{\nabla V_t}{\varepsilon(t)^2}\right|^2\mathrm{d}\lambda \\
&\leq -\frac{2}{C(t)}\varepsilon(t)^2 H(p_t|\Pi_{t,\varepsilon(t)}).
\end{aligned}$$

On the other hand, we obtain the following for the second integral:

$$\begin{aligned}
\int p_t\frac{\dot{\pi}_{t,\varepsilon(t)}}{\pi_{t,\varepsilon(t)}}\mathrm{d}\lambda &= 4\frac{\dot{\varepsilon}(t)}{\varepsilon(t)^3}\int p_t(V_t - < V_t >_{\Pi_{t,\varepsilon(t)}})\mathrm{d}\lambda \\
&- \frac{2}{\varepsilon(t)^2}\int p_t(\dot{V}_t - < \dot{V}_t >_{\Pi_{t,\varepsilon(t)}})\mathrm{d}\lambda.
\end{aligned}$$

We put all the pieces together and this leads to the result. □

Let us prove the classical result:

LEMMA 4.14. *(Miclo, lemme 6) Let $f : [0,\infty[\to\mathbb{R}_+$ be a continuous function such that a.s.*

$$f'(t) \leq \alpha_t - \beta_t f(t),$$

where α and β are two continuous non-negative functions such that $\int^\infty \beta_t \mathrm{d}t = \infty$ and $\lim_{t\to\infty}\alpha_t/\beta_t = 0$. Then $\lim_{t\to\infty} f(t) = 0$.

Proof. The idea is to adapt the usual inequality of Gronwall. Let $g(t) = f(t)\exp\left(\int_0^t \beta_s \mathrm{d}s\right)$ be a continuous function. We get a.s. (with respect to the time t)

$$g'(t) \leq \alpha_t \exp\left(\int_0^t \beta_s \mathrm{d}s\right).$$

We therefore obtain that

$$g(t) \leq g(0) + \int_0^t \alpha_s \exp\left(\int_0^s \beta_u \mathrm{d}u\right)\mathrm{d}s.$$

Consequently, for all initial $t_0 \geq 0$, we have

$$\begin{aligned} f(t) &\leq f(0)\exp\left(-\int_0^t \beta_u \mathrm{d}u\right) \\ &+ \exp\left(-\int_0^t \beta_u \mathrm{d}u\right)\int_0^{t_0} \mathrm{d}s\alpha_s \exp\left(\int_0^s \beta_u \mathrm{d}u\right) \\ &+ \exp\left(-\int_0^t \beta_u \mathrm{d}u\right)\int_{t_0}^{t} \mathrm{d}s\alpha_s \exp\left(\int_0^s \beta_u \mathrm{d}u\right). \end{aligned}$$

Let $\eta > 0$. We choose t_0 such that for all $t \geq t_0$, we have $\alpha_s \leq \eta\beta_s$. We thus find a upper bound, for all $t \geq t_0$, for the last term of the preceding inequality:

$$\begin{aligned} &\exp\left(-\int_0^t \beta_u \mathrm{d}u\right)\int_{t_0}^{t} \mathrm{d}s\eta\beta_s \exp\left(\int_0^s \beta_u \mathrm{d}u\right) \\ \leq\ &\eta\exp\left(-\int_0^t \beta_u \mathrm{d}u\right)\left(\exp\left(\int_0^t \beta_u \mathrm{d}u\right) - \exp\left(\int_0^{t_0} \beta_u \mathrm{d}u\right)\right) \leq \eta. \end{aligned}$$

We also remark that the two first terms go to zero when t goes to the infinity and thus $\limsup_{t\to\infty} f(t) \leq \eta$. The result follows. $\square$

We now need a technical lemma.

Lemma 4.15. *For all $t \geq 0$, the quantity $< |x|^2 >_{\Pi_{t,\varepsilon(t)}}$ is bounded.*

Proof. Let K be the compact set $K := \{x | V(x) \leq \eta\}$ where η is a positive constant. As $\Pi_{t,\varepsilon(t)}$ converges weakly to Π_0 we only need to prove that $< |x|^2 \mathbb{1}_{K^c} >_{\Pi_{t,\varepsilon(t)}}$ is bounded. We recall that

$$\Pi_t(\mathrm{d}x) = \frac{e^{-2\varepsilon^{-2}(t)V_t(x)}}{Z(\varepsilon(t),t)}\mathrm{d}x.$$

We have

$$\begin{aligned} \int_{K^c} |x|^2 e^{-2\varepsilon^{-2}(t)V(x)} e^{-2\frac{|x|^2}{a(t)\varepsilon^2(t)}}\mathrm{d}x &\leq \int_{K^c} |x|^2 e^{-2V(x)} e^{-2V(x)(\varepsilon^{-2}(t)-1)}\mathrm{d}x \\ &\leq \int_{K^c} |x|^2 e^{-2V(x)}\mathrm{d}x e^{-2\eta(\varepsilon^{-2}(t)-1)} \\ &= Ce^{-2\eta\varepsilon^{-2}(t)} \end{aligned}$$

where C is a positive constant.
But we have still proved (see the proof of the proposition 4.4) that $Z(t,\varepsilon(t)) \underset{t\infty}{\sim} \sum_i (2\pi\varepsilon^2(t))^{d/2}(\det\nabla^2 V(x_i))^{-1/2}$, therefore

$$< |x|^2 \mathbb{1}_{K^c} >_{\Pi_{t,\varepsilon(t)}} \leq \widetilde{C}\varepsilon^{-d}(t)e^{-2\eta\varepsilon^{-2}(t)} \underset{t\to\infty}{\to} 0$$

and the result follows. □

THEOREM 4.16. *Suppose that $\varepsilon^2(t) = k/\log(t)$, where the constant $k > 2osc(\chi)$. Then, for all initial t_0, x_0, the free energy $H\left(p(t_0,x_0,t,\cdot)|\Pi_{t,\varepsilon(t)}\right)$ converges to 0 as t goes to the infinity.*

PROOF. Let $t_0 \geq 0$ and $x_0 \in \mathbb{R}^d$. Consider the process Z_t, which is the solution to the SDE

$$\mathrm{d}Z_t = \varepsilon(t)\mathrm{d}W_t - \left(\nabla V(Z_t) + \frac{Z_t}{a(t)}\right)\mathrm{d}t, \;\; Z_{t_0} = x_0.$$

We can rewrite the result of the proposition 4.13 in the following way:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}H(p_t|\Pi_{t,\varepsilon(t)}) \;\; &\leq \;\; -\frac{2}{C(t)}\varepsilon(t)^2 H(p_t|\Pi_{t,\varepsilon(t)}) + \frac{2}{\varepsilon(t)^2}(\mathbb{E}\dot{V}_t(Z_t) - <\dot{V}_t>_{\Pi_{t,\varepsilon(t)}}) \\ &- \;\; 4\dot{\varepsilon}(t)\varepsilon(t)^{-3}\left(\mathbb{E}V_t(Z_t) - <V_t>_{\Pi_{t,\varepsilon(t)}}\right).\end{aligned}$$

We remind the reader that $V(x) \geq c|x|^2$ out of a compact set and we have proved in the lemma 6.4 that $\mathbb{E}V(Z_t) = O(1)$. We therefore have $\mathbb{E}V_t(Z_t) = O(1)$. Moreover, the function $t \mapsto a(t)$ is non-decreasing whereas $t \mapsto \varepsilon(t)$ is non-increasing. Thus, as $\dot{V}_t = -\frac{\dot{a}(t)}{a(t)^2}$, the two terms $\mathbb{E}(\dot{V}_t(Z_t))$ and $<V_t>_{\Pi_{t,\varepsilon(t)}}$ do not play any role in the upper bound. Therefore, it remains to find a upper bound for $<\dot{V}_t>_{\Pi_{t,\varepsilon(t)}}$. To this aim, we use the lemma 4.15. Consequently, there exist two positive constants M_1, M_2 such that

$$\frac{\mathrm{d}}{\mathrm{d}t}H(p_t|\Pi_{t,\varepsilon(t)}) \leq -\frac{2}{C(t)}\varepsilon(t)^2 H(p_t|\Pi_{t,\varepsilon(t)}) + M_1\frac{\dot{\varepsilon}(t)}{\varepsilon(t)^3} + M_2\frac{\dot{a}(t)}{\varepsilon(t)^2 a(t)^2}.$$

It now just remains to use the previous lemma 4.14. We still know that, for $\varepsilon^2(t) = k/\log t$, and we easily compute the time-derivative of $a(t)$: $\frac{\dot{a}(t)}{a(t)^2\varepsilon^4(t)} = -\frac{\dot{\varepsilon}(t)}{\varepsilon^3(t)(r+G^{-1}(t))} + \frac{1}{(g\circ G^{-1}(t))(r+G^{-1}(t))^2\varepsilon^2(t)}$. We use the explicit expression of ε, to get $C(t)\frac{\dot{a}(t)}{a(t)^2\varepsilon^4(t)} = \frac{C(t)}{2kt(r+G^{-1}(t))} + C(t)\frac{\log t}{k(g\circ G^{-1}(t))(r+G^{-1}(t))^2}$ and, as $G^{-1}(t)$ is a non-decreasing function and because of the hypothesis on k, the first term converges to 0 when t goes to the infinity. For the second term, we recall that $\log G(t)/g(t)$ is supposed to be bounded and therefore, $G(t)^{2osc(\chi)/k}\log G(t)/(kg(t)(r+t)^2) \underset{t\to\infty}{\longrightarrow} 0$, because $G(t) = o(t^2)$. (Actually, one has: there exist two positive constants m, M such that $mg(t) \leq \log G(t) \leq Mg(t)$. Thus $mg(t) \leq \log(tg(t)) = \log t + \log g(t)$. This naturally implies that $g(t) = O(\log t)$ and then $G(t) \leq tg(t) = o(t^2)$.) The lemma asserts that if ε satisfies

- $\int^\infty \varepsilon(t)^2\frac{\mathrm{d}t}{C(t)} = \infty$,

- $\frac{\dot{\varepsilon}(t)}{\varepsilon^5(t)} \underset{t\to\infty}{\longrightarrow} 0$,

then $\lim\limits_{t\to\infty} H(p_t|\Pi_{t,\varepsilon(t)}) = 0$. If we let $\varepsilon^2(t) = k/\log t$ with the given condition on the constant k, then the result follows. □

REMARK 4.17. *We believe that the constant k is not optimal here, because we have used the perturbation lemma of Holley and Stroock for the estimation of the logarithmic Sobolev constant. Actually, k should be directly related to m, as Holley, Kusuoka and Stroock* [**12**] *proved it for $a(t) \equiv \infty$.*

LEMMA 4.18. *The speed of convergence of $H(p(t_0, x_0, t, \cdot)|\Pi_{t,\varepsilon(t)})$ toward 0 is $(G^{-1}(t)\log t)^{-1}$.*

PROOF. We recall the lemma 4.14. We thus get that the speed of convergence is given by $\exp\left(-\int_0^t \beta_s \mathrm{d}s\right)\int_0^t \mathrm{d}s\alpha_s e^{\int_0^s \beta_u \mathrm{d}u}$, with $\beta_s = s^{2osc(\chi)/k}/\log s$ and $\alpha_s = (s(r + G^{-1}(s)))^{-1} + \log s(g \circ G^{-1}(s)(r + G^{-1}(s))^2)^{-1}$. By an integration by parts, we find that $\int_0^t \beta_s \mathrm{d}s$ is equivalent, when t goes to the infinity, to $t^{1+2osc(\chi)/k}/\log t$ and thus, the "speed of convergence" is of the order of $(G^{-1}(t)\log t)^{-1} + (g \circ G^{-1}(t)(G^{-1}(t))^2)^{-1}$. Finally, we obtain that the speed of convergence of the relative Kullback information to zero is $(G^{-1}(t)\log t)^{-1} = o((\log t)^{-1})$. □

REMARK 4.19. *It is known since the work of Freidlin and Wentzell* [**7**] *(see for example the chapter 5 of the book* [**6**]*), that the Gibbs measure $\Pi_{t,\varepsilon(t)}$ satisfies a large deviation principle. Therefore, the speed of convergence of $\Pi_{t,\varepsilon(t)}$ toward Π_0 is exponential ($e^{-\log t/2k}$, that is $t^{-1/2k}$).*

COROLLARY 4.20. *Suppose that $\varepsilon^2(t) = k/\log(t)$, where the constant $k > 2osc(\chi)$. Then the process Z converges in distribution to a random variable which concentrates on the global minima of V.*

PROOF. The Kullback information $H(p_t|\Pi_{t,\varepsilon(t)})$ measures the distance between p_t and $\Pi_{t,\varepsilon(t)}$ in the following way: $||p_t - \Pi_{t,\varepsilon(t)}||_{TV}^2 \leq 2H(p_t|\Pi_{t,\varepsilon(t)})$, where $||\cdot||_{TV}$ denotes the total variation norm (see for instance [**14**] or [**22**]). The result follows because $\Pi_{t,\varepsilon(t)}$ converges weakly to Π_0. □

COROLLARY 4.21. *Suppose that $\lim\limits_{t\to\infty} g(t)^{-1}\log G(t) = k > 2osc(\chi)$. Then the process Y converges in distribution to a random variable Y_∞, which concentrates on the global minima of V.*

PROOF. Straightforward. □

REMARK 4.22. *We emphasize that if $\lim\limits_{t\to\infty} g(t)^{-1}\log G(t) = k$, with k not large enough, then the previous result is false and in that case, the process Y does not converge toward the global minima of V.*

5. The process X: convergence in law

We give necessary and sufficient conditions for the convergence in probability of X. As usual, we begin to work with the process $Y_t = X_t - \overline{\mu}_t$. In order to link this section with the preceding one, we recall that $\varepsilon(t)^2 = (g \circ G^{-1}(t))^{-1} = k/\log t$. It implies that we consider the functions g such that (at least asymptotically) $\log G(t) = kg(t)$.

We are now able to conclude the study of the asymptotic behavior of the process X.

THEOREM 5.1. *Suppose that* $\lim_{t\to\infty} g(t)^{-1}\log G(t) = k$, *where* $k > 2osc(\chi)$. *Then it holds one of the following:*

1. *If* V *is a function such that* $\sum_{1\le i\le p} a_i m_i + \sum_{1\le i\le n} b_i M_i = 0$, *then* X_t *converges in law to* $Y_\infty + \int_0^\infty Y_s \frac{\mathrm{d}s}{r+s}$;
2. *Else,* X_t *diverges.*

PROOF. The main aim of this proof is to decide whether $\int_0^t Y_s \frac{\mathrm{d}s}{r+s}$ converges a.s. or not. Actually, suppose that V is such that the preceding integral converges. The celebrated Slutsky theorem asserts that for two sequences (U_t), (V_t) of $\mathbb{R}^d$ valued random variables, with $U_t \xrightarrow[t\to\infty]{(d)} U$ and $|U_t - V_t| \xrightarrow[t\to\infty]{\mathbb{P}} 0$, then $V_t \xrightarrow[t\to\infty]{(d)} U$. To prove the result, we just need to let $U_t = \overline{\mu}_t$, $V_t = X_t$ (and thus $V_t - U_t = Y_t$) and remark that

$$X_t = Y_t + \int_0^t \frac{\mathrm{d}s}{r+s} Y_s = Y_t + \overline{\mu}_t.$$

Suppose that V is such that $\sum_{1\le i\le p} a_i m_i + \sum_{1\le i\le n} b_i M_i = 0$. Then, we know that $\frac{1}{t}\int_0^t Y_s \mathrm{d}s \xrightarrow{a.s.} 0$ and it remains to find the rate of convergence in order to conclude the proof. But, in the work of Benaïm and Schreiber [**3**] (theorem 1), it is shown that for an asymptotic pseudo-trajectory (in probability) Y, the speed of convergence of the mean of the normalized occupation measure of Y is the same as the speed of the pseudo-trajectory. This means that the speed of convergence of the normalized occupation measure of the time-changed process $Y_{G^{-1}(t)}$ is $G^{-1}(1+t) - G^{-1}(t)$. But, we are looking for the speed of convergence for Y_t itself. By an integration by parts, we obtain that

$$\begin{aligned}
\frac{1}{t}\int_0^t Y_s \mathrm{d}s &= \frac{1}{t}\int_0^{G(t)} Y_{G^{-1}(u)} \frac{\mathrm{d}u}{g \circ G^{-1}(u)} \\
&= \frac{1}{tg(t)}\int_0^{G(t)} Y_{G^{-1}(s)}\mathrm{d}s + \frac{1}{t}\int_0^{G(t)} \mathrm{d}u \frac{g' \circ G^{-1}(u)}{(g\circ G^{-1}(u))^3}\int_0^u Y_{G^{-1}(s)}\mathrm{d}s.
\end{aligned}$$

The first right-hand term converges to 0 because we know that $G(t) \le tg(t)$. It remains to prove the convergence of the second term. We recall to the reader that, up to a multiplicative positive constant, $g \circ G^{-1}(u) = \log(2 +$

u). Moreover, we also know that $\frac{1}{u}\int_0^u Y_{G^{-1}(s)}\mathrm{d}s$ is a.s. bounded. As a consequence, we obtain that the second right-hand term is upper bounded (up to a multiplicative constant) by

$$\frac{1}{t}\int_0^{G(t)} \frac{\mathrm{d}u}{(\log(2+u))^2} = \frac{G(t)}{t(\log G(t))^2} + o\left(\frac{G(t)}{t(\log G(t))^2}\right) \le \frac{g(t)}{\log G(t)}\frac{1}{(\log G(t))}$$

and the result follows.

Suppose now that V is a function such that $\sum\limits_{1\le i\le p} a_i m_i + \sum\limits_{1\le i\le n} b_i M_i \neq 0$. We recall that we have proved that a.s. $\frac{1}{t}\int_0^t \delta_{Y_s}\mathrm{d}s \to \sum\limits_{1\le i\le p} a_i\delta_{m_i} + \sum\limits_{1\le i\le n} b_i\delta_{M_i}$ in sense of the weak convergence of measures. But we emphasize that since the work of Meyn and Tweedie [**19**], we know that the duality for the weak convergence is true for the dual V-norm, that is for functions f bounded by V. In our case, we just have proved that for all functions f such that $\|f\|_2 := \sup_{x\in\mathbb{R}^d}\frac{|f(x)|}{\|x\|^2} < \infty$, we have

$$\frac{1}{t}\int_0^t f(Y_s)\mathrm{d}s \xrightarrow{a.s.} \sum_{1\le i\le p} a_i f(m_i) + \sum_{1\le i\le n} b_i f(M_i).$$

Now, we can choose $f = id$ and the result follows: $\int_0^t Y_s\mathrm{d}s$ does not converge when V satisfies our hypothesis. □

Remark 5.2. *1) The condition on g in the preceding result is for instance satisfied for $g(t) = k\log t$.*
2) In particular, if V is a symmetric function, then the preceding result asserts that X_t converges in law.

6. Concluding remarks

We point out that we did not manage to obtain the optimal annealing schedule (that is the constant k of the preceding sections in not optimal). We also emphasize that the function ε was supposed to decrease slowly to zero. That is the reason why we obtained the convergence of Y to the global minima of V. But if ε goes too fast to zero, then the process Y can freeze in a local minimum, depending on the initial value Y_0. This question, corresponding to the case $0 < k \le 2osc\chi$, will be addressed later.

Moreover, some questions arise naturally:

- what happens with a function g that vanishes periodically, for instance $g(t) = \cos(\pi t)$?
- what happens if $\lim\limits_{t\to\infty} g(t) = 0$? Does the process X behave like a Brownian motion?

Bibliography

[1] Ané C., Blachère S., Chafaï D., Fougères P., Gentil Y., Malrieu F., Roberto C. & Scheffer G. [2001], *Sur les inégalités de Sobolev logarithmiques*, Panoramas et Synthèses 10, SMF.

[2] Bakry D. & Emery M. [1985], Diffusions hypercontractives, *Sém. Prob. XIX*, Springer LNM 1123, 177-206.

[3] Benaïm M. & Schreiber S.J. [2000], Weak asymptotic pseudotrajectories for semi-flows: ergodic properties, *J. Dynamics and Diff. Eq.* **12**(3).

[4] Chambeu S. & Kurtzmann A. [2007], Some particular self-interacting diffusions: ergodic behavior and almost sure convergence, preprint.

[5] Chiang T.S., Hwang C.R. & Sheu S.J. [1987], Diffusion for global optimization in $\mathbb{R}^n$, *SIAM J. Control Optim.*, **25**, 737-753.

[6] Duflo M. [1996], *Algorithmes stochastiques*, Springer "Mathématiques & Applications 23".

[7] Freidlin M.I. & Wentzell A.D. [1984], *Random perturbation of dynamical systems*, Springer.

[8] Geman S. & Hwang C.R. [1986], Diffusions for global optimization, *SIAM J. Control Optim.*, **24**, 1031-1043.

[9] Gidas B. [1987], The Langevin equation as a global minimization algorithm, *Disordered systems and biological organizations*, E. Bienenstock, F. Flogeman, G. Weisbuch eds., Springer-Verlag, Berlin.

[10] Gilbarg D. & Trudinger N.S. [1977], *Elliptic partial differential equations of second-order*, 2nd ed., Springer-Verlag, Berlin.

[11] Hajek B. [1988], Cooling schedules for optimal annealing, *Math. Oper. Res.*, **13**, 311-329.

[12] Holley R.A., Kusuoka S. & Stroock D.W. [1989], Asymptotics of the spectral gap with applications to the theory of simulated annealing, *J. Funct. Anal.*, **83**, 333-347.

[13] Holley R. & Stroock D. [1987], Logarithmic Sobolev Inequalities and Stochastic Ising Models, *J. Stat. Phys.* **46**, 1159-1194.

[14] Holley R.A. & Stroock D.W. [1988], Simulated annealing via Sobolev inequalities, *Comm. Math. Phys.*, 553-569.

[15] Hwang C.R. [1980], Laplace's method revisited: weak convergence of probability measures, *Ann. Proba.*, **8**, 1177-1182.

[16] Hwang C.R. & Sheu S.J. [1990], Large-time nehavior of perturbed diffusion Markov processes with applications to the second eigenvalue problem for Fokker-Planck operators and simulated annealing, *Acta App. Math.*, **19**, 253-295.

[17] Jacot S. [1992], Comportement asymptotique de la seconde valeur propre des processus de Kolmogorov, *J. Mult. An.*, **40**, 335-347.

[18] Lépingle D. [1978], Sur le comportement aymptotique des martingales locales, *Sém. Prob. (Strasbourg)* **12**, 148-161.

[19] Meyn S.P. & Tweedie R.L. [1993], *Markov Chains and Stochastic Stability*, Springer-Verlag.

[20] Miclo L. [1992], Recuit simulé sur $\mathbb{R}^n$. Etude de l'évolution de l'énergie libre, *Ann. Inst. H. Poincaré*, **28** (2), 235-266.

[21] Royer G. [1989], A remark on simulated annealing of diffusion processes, *SIAM J. Control Optim.*, **27**, 1403-1408.

[22] Villani C. [2003], *Topics in Optimal Transportation*, Graduate Studies in Mathematics, Vol. 58, *AMS*.

[23] Zitt P-A. [2006], *Applications d'inégalités fonctionnelles à la mécanique statistique et au recuit simulé*, thesis Université Paris X Nanterre.

Deuxième partie

Diffusions renforcées par la mesure d'occupation normalisée

CHAPITRE 3

Panorama des processus avec renforcement

Dans ce chapitre, nous allons présenter trois grandes familles de diffusions renforcées (non markoviennes) qui se différencient nettement : renforcement par la mesure d'occupation, renforcement par la mesure d'occupation normalisée et renforcement par les extrema. Leur point commun essentiel se situe en le fait que ces trois diffusions sont à mémoire longue, et leur trajectoire à l'instant présent dépend de toute la trajectoire passée.

1. Introduction

Le monde des processus de Markov (en particulier les diffusions) est un univers bien cartographié par les probabilistes, qui explorent, depuis une décennie, de nouveaux processus perturbés ou renforcés. Qu'est-ce qu'un processus renforcé ? Nous nous intéressons à des processus dépendant de leur trajectoire passée, soit par une dérive lorsque le processus est solution d'une équation différentielle stochastique, soit par un "réfléchissement" selon certains paramètres. Nous donnons plus de détails dans la suite du chapitre. Parmi les processus renforcés, on peut distinguer deux grandes familles : la première est composée des processus markoviens tandis que l'autre contient des processus non markoviens. D'un point de vue probabiliste, cela fait une différence énorme ! On ne peut donc pas les traiter par les mêmes techniques. En fait, qu'ils soient markoviens ou non, il n'existe pas de méthode globale pour analyser les processus renforcés. Pemantle [**41**] a récemment référencé les principales techniques et approches concernant les processus renforcés (discrets ou continus). Une partie de ce paragraphe se base sur ce travail.

Les processus renforcés (discrets) ont des applications en biologie, en économie, en théorie des jeux etc. Par exemple, un problème que se posent les psychologues est de savoir comment le comportement humain s'apprend. Un modèle simplifié consiste à supposer que des individus (par exemple des êtres humains) ont le choix entre deux possibilités : A ou B. Après avoir choisi, ils reçoivent une récompense (par exemple de l'argent) selon leur choix. La question est de savoir comment la récompense va influencer leurs actions futures. Une autre application consiste à envisager deux technologies (par exemple Apple contre IBM) entrant sur le marché au même moment. Aucune des deux n'est meilleure techniquement. Supposons que les nouveaux consommateurs choisissent l'une des deux proportionnellement à ce que les gens ont acheté avant eux. Il s'agit du modèle de l'urne de Pólya, qui implique qu'asymptotiquement, les parts de marché sont aléatoires. Dans le cas de Apple, le modèle de l'urne de Pólya est discuté par Arthur, Ermoliev et Kaniovskii [**1**] et la part de marché (asymptotique) de cette technologie se situe entre 10% et 15%. Pour d'autres exemples d'application des processus renforcés, nous renvoyons à nouveau le lecteur au travail de Pemantle [**41**].

En ce qui concerne les processus renforcés markoviens, nous pensons essentiellement aux processus correspondant à l'équation différentielle partielle introduite par McKean et Vlasov. Il s'agit en fait d'une équation de milieu granulaire, donnée par

$$\mathrm{d}X_t = \mathrm{d}B_t - \nabla V * \nu_t(X_t)\mathrm{d}t$$

où B est un mouvement brownien standard, V est une fonction strictement convexe (sauf éventuellement en 0) – ou convexe selon les résultats obtenus –, $*$ est le produit de convolution et ν_t représente la loi à l'instant t du processus X.

Notons $\Pi(\mu)(\mathrm{d}x) := \frac{1}{Z}e^{-2V*\mu(x)}\mathrm{d}x$ où Z est la constante de normalisation. Remarquons que si on considère le "point fixe" de Π (qui est bien unique lorsque V est strictement convexe), c'est-à-dire la mesure de probabilité μ telle que $\mu^* = \Pi(\mu^*)$, alors cette probabilité est invariante pour le processus : si on part d'un point X_0 de loi μ^*, alors pour chaque instant t, X_t est de loi μ^*. On cherche à obtenir une vitesse de convergence vers cet équilibre pour la distance de Wasserstein. Divers travaux sur ce problème ont été menés ces dernières années, en particulier ceux de Carrillo, McCann et Villani [**14**], Bolley, Guillin et Villani [**10**], ou encore Cattiaux, Guillin et Malrieu [**15**]. Malrieu [**36**] utilise quant à lui un système de particules approchant l'E.D.S. non linéaire.

Dans ce mémoire, nous nous intéressons aux diffusions renforcées non markoviennes. Nous divisons alors cette famille en trois grands sous-groupes. Un premier groupe est constitué par le mouvement brownien perturbé, étudié initialement par Carmona, Petit et Yor [**12**]. Ce processus se distingue des suivants par le fait qu'il ne satisfait aucune E.D.S. De plus, il est renforcé par son maximum et/ou par son minimum et non pas par toute sa trajectoire passée. Cela revient à étudier l'équation

$$X_t = f(t) + \alpha \sup_{0\le s\le t} X_s + \beta \inf_{0\le s\le t} X_s,$$

où α et β sont des réels et f une fonction continue, nulle en 0, aléatoire ou déterministe. Souvent, nous supposerons que f est un mouvement brownien standard. Contrairement aux autres processus, pour lesquels le résultat est plus "standard", se pose ici la question de l'existence et unicité de la solution en fonction de f. Cette étude sera menée dans la première section.

Ensuite, Durrett et Rogers [**27**] considèrent l'équation

$$\mathrm{d}X_t = \mathrm{d}B_t - f * \mu_t(X_t)\mathrm{d}t$$

où B est un mouvement brownien standard, f est une fonction continue et μ_t représente la *mesure d'occupation* à l'instant t du processus. On définit le produit de convolution par $f * \mu_t(x) = \int_{\mathbb{R}^d} f(x-y)\mu_t(\mathrm{d}y)$. On cherche alors à étudier le comportement asymptotique de X selon différentes fonctions f. Cette étude fait l'objet du premier paragraphe de ce chapitre.

Enfin, une dernière famille se distingue. Benaïm, Ledoux et Raimond [**7**] ont étudié une équation proche de la précédente, mais où μ_t est la *mesure d'occupation normalisée* à l'instant t du processus et f est le gradient d'une certaine fonction V. De plus, ils ne travaillent pas sur $\mathbb{R}^d$, mais sur une variété riemanienne compacte, ce qui simplifie un certain nombre de choses. La question naturelle est alors d'étudier le comportement asymptotique de μ_t et non plus de X. Le troisième et dernier paragraphe contient les résultats obtenus pour ce processus.

En annexe est étudié le "mouvement brownien perturbé version E.D.S.". Il s'agit de l'étude de l'équation

$$\mathrm{d}X_t = \mathrm{d}B_t + c\left(\sup_{0\leq s\leq t} X_s - X_t\right)\mathrm{d}t.$$

Il n'existe pas de technique globale pour analyser les processus renforcés. Pour avoir une vue des techniques les plus utilisées, nous renvoyons à l'article de Pemantle [**5**]. Néanmoins, ces processus ont une caractéristique commune. En élargissant l'espace d'état, on peut toujours se ramener à un processus markovien (par exemple en considérant (X_t, t, μ_t) dans le modèle renforcé par la mesure d'occupation normalisée ou encore $\left(X_t, \sup_{s\leq t} X_s, \inf_{s\leq t} X_s\right)$ pour le mouvement brownien perturbé). Pourtant, personne n'exploite cette information, car alors l'espace d'état est trop gros pour qu'on puisse en tirer des propriétés intéressantes.

Une grande différence entre ces processus renforcés et les processus markoviens est que la tribu asymptotique n'est pas triviale en général. En effet, dans le travail mené par Benaïm, Ledoux et Raimond, la mesure d'occupation normalisée μ_t, lorsqu'elle converge, converge vers différents points fixes avec une probabilité dépendant des conditions de départ, soit (x,μ). De plus, un processus renforcé n'est pas toujours ergodique. En effet, si nous considérons le modèle envisagé par Durrett et Rogers [**27**], on se rend compte que, selon la fonction f, le processus X converge presque sûrement ou non. Dans le cas où il converge, alors il est clair par le résultat de Cesàro que la mesure d'occupation normalisée μ_t de X converge presque sûrement. En revanche, lorsque X diverge, alors μ_t ne converge pas toujours. Cela se voit également dans l'étude de Benaïm, Ledoux et Raimond [**7**]. Ils donnent des conditions sous lesquelles μ_t converge presque sûrement et d'autres pour que μ_t ne converge pas.

Ce chapitre a pour but de présenter une grande partie des résultats obtenus (à notre connaissance) au cours de ces dernières années sur les diffusions renforcées. Nous nous limitons uniquement aux espaces d'état continus et ne traitons en aucun cas ce qui se passe dans le cas discret. Nous ne discutons pas non plus des processus markoviens, dont l'étude diffère quelque peu. En effet, dans ce cas, les techniques utilisées font appel aux inégalités fonctionnelles de type Sobolev logarithmique, ou Poincaré. Elles reposent plus largement sur les techniques du transport de masse. De plus, nous ne donnons que les ingrédients des démonstrations des résultats annoncés mais ne les reproduisons pas ici dans leur intégralité.

2. Processus renforcé par ses extrema

Le mouvement brownien perturbé est un processus qui se comporte comme le mouvement brownien standard entre ses extrema et qui est réfléchi suivant certains paramètres réels lorsqu'il atteint son maximum ou minimum passé. Mathématiquement parlant, nous généralisons cette notion en étudiant l'équation unidimensionnelle

$$X_t = f(t) + \alpha \sup_{0 \leq s \leq t} X_s + \beta \inf_{0 \leq s \leq t} X_s. \tag{2.1}$$

Ici, α et β sont des réels et f une fonction continue, nulle en 0. Lorsque nous parlons de mouvement brownien perturbé, nous supposons que $f(t) = B_t$

2.1. Existence d'une solution. Il est en fait plus facile de construire un mouvement brownien perturbé uniquement soit par son maximum, soit par son minimum. Supposons par exemple que $\alpha = 0$, $\beta < 1$ et $f(t) = B_t$ un mouvement brownien standard (en fait, pour $\beta \geq 1$, l'équation étudiée n'admet pas de solution avec probabilité 1, ce que nous démontrons un peu plus tard). Un tel processus se comporte alors localement comme un mouvement brownien standard, sauf lorsqu'il atteint son minimum. Remarquons que lorsque $0 < \beta < 1$, alors la perturbation est auto-répulsive, alors que pour $\beta < 0$ elle est auto-attractive. Une remarque fort importante est la suivante. Si on note L le temps local en 0 du mouvement brownien réfléchi $W_t^+ = B_t - \inf_{0 \leq s \leq t} B_s$, alors on remarque que pour tout $t \geq 0$ nous avons une solution au problème :

$$X_t = W_t^+ - \frac{1}{1-\beta} L_t.$$

Parmi les propriétés les plus intéressantes concernant X, notons que ses temps locaux pris en des temps d'arrêts bien choisis sont des processus de Bessel $BES(2)$. On peut ainsi obtenir des équivalents des théorèmes de Ray-Knight et la loi de l'arcsinus pour le processus X.

En revanche, pour $\alpha \neq 0$ et $\beta \neq 0$, on ne peut plus trouver de représentation du processus en fonction du temps local (ou autre) comme cela était possible précédemment. Carmona, Petit et Yor [**12**] sont les premiers à obtenir des résultats sur ce processus pour α et β non nuls. Ils commencent bien évidemment par se poser la question de l'existence et l'unicité de solution d'une telle équation. Commençons par le résultat de non-existence suivant

Proposition 2.1. *(Carmona-Petit-Yor, proposition 2.1) Pour $\alpha \geq 1$ ou $\beta \geq 1$, l'équation (2.1) n'admet pas de solution.*

Démonstration. Supposons pour simplifier que $f(t) = B_t$ un mouvement brownien standard. Par symétrie, il suffit de montrer que pour $\alpha \geq 1$, l'équation (2.1) n'admet pas de solution. Supposons au contraire que la solution existe. La formule de Tanaka implique que la solution, notée X, vérifie

$$X_t^+ = \int_0^t \mathbb{1}_{X_s>0} \mathrm{d}X_s + \frac{1}{2} L_t$$

où L est le temps local de X en 0. En se rappelant l'expression de X, on obtient que la martingale locale suivante est (continue et) négative

$$N_t := \int_0^t \mathbb{1}_{X_s>0} \mathrm{d}B_s = X_t^+ - \alpha \sup_{s\leq t} X_s - \frac{1}{2} L_t \leq (1-\alpha) \sup_{s\leq t} X_s \leq 0.$$

Donc $N_t \overset{p.s.}{=} 0$ et sa variation quadratique est également nulle. Le processus X reste donc négatif, ce qui donne $\sup_{s\leq t} X_s = 0$ et finalement

$$X_t = B_t + \beta \inf_{0\leq s\leq t} X_s.$$

Nous traitons maintenant séparément les cas $\beta < 1$ et $\beta \geq 1$. En effet, si $\beta \geq 1$, alors on refait la même démarche pour prouver que $X_t \geq 0$ pour tout t et $\inf_{0\leq s\leq t} X_s = 0$, ce qui donne alors $X_t = B_t = 0$, d'où la contradiction. On a donc $\beta < 1$. Par le lemme de réflexion de Skorokhod, on a $X_t = B_t + \frac{\beta}{1-\beta} \inf_{0\leq s\leq t} B_s$ et donc le processus X ne peut pas rester toujours négatif. D'où le résultat. □

Carmona, Petit et Yor prouvent également les théorèmes de Ray-Knight pour le processus X. Une conséquence du premier théorème de Ray-Knight est le résultat suivant, qui est l'équivalent de la loi de l'arcsinus lorsque $\alpha = \beta = 0$.

Proposition 2.2. *(Carmona-Petit-Yor, proposition 3.7) Supposons que $f(t) = B_t$ et notons X la solution de l'équation (2.1). Alors le processus $\frac{A_t}{t} := \frac{1}{t}\int_0^t \mathbb{1}_{X_s>0}\mathrm{d}s$ suit une loi Beta $\left(\frac{1-\alpha}{2}, \frac{1-\beta}{2}\right)$.*

Ce résultat est une conséquence du premier théorème de Ray-Knight, comme nous l'avons déjà mentionné. Perman et Werner ([**42**] proposition 4) ont observé qu'il s'agit également d'une conséquence du fait que les variables aléatoires A_τ et $t - A_\tau$, où τ est l'inverse du temps local en 0 de X, sont indépendantes. De plus, leurs inverses suivent une loi Gamma (nous renvoyons à Doney [**25**] pour plus de précisions).

2.2. Etude de la solution : adaptabilité et propriété de Markov. Si $\alpha < 1$ et $\beta < 1$ sont des constantes telles que $|\alpha\beta| < (1-\alpha)(1-\beta)$, alors Carmona, Petit et Yor ont de plus montré que l'équation (2.1) admet une unique solution forte, qui, dans le cas $f(t) = B_t$, est également adaptée à la filtration du mouvement brownien B. Mais ce résultat est amélioré par Davis [**23**]. Dans toute la suite, nous supposons donc $\alpha < 1$ et $\beta < 1$.

PROPOSITION 2.3. *(Davis, théorème 1.1) Posons* $\rho := \frac{\alpha\beta}{(1-\alpha)(1-\beta)}$. *Il y a alors deux possibilités :*

(1) *si $\rho \leq 1$, alors l'équation (2.1) admet une unique solution,*

(2) *sinon, alors l'équation (2.1) admet au moins une solution, telle que $X_0 = 0$, pour chaque fonction f et il existe des fonctions pour lesquelles il y a une infinité de solutions.*

En outre, Davis montre que dans le cas $|\rho| < 1$ et $f(t) = B_t$, la solution de (2.1) peut s'identifier à la limite (en loi) d'un processus discret : la marche aléatoire renforcée "simple" (on ajoute un poids de 1 à chaque sommet atteint). Définissons le processus discret $\mathbf{Y}_{\alpha,\beta} = Y_0, Y_1, \cdots$ par : $Y_0 = 0$,

$$\begin{aligned}\mathbb{P}(Y_{n+1} = Y_n + 1 | Y_i, 1 \leq i \leq n) &= 1 - \mathbb{P}(Y_{n+1} = Y_n - 1 | Y_i, 1 \leq i \leq n) \\ &= \begin{cases} \frac{1}{2} \text{ si } n = 0 \text{ ou } Y_n \neq \inf_{0 \leq i \leq n} Y_i, \sup_{0 \leq i \leq n} Y_i \\ \frac{1}{2-\alpha} \text{ si } n > 0 \text{ et } Y_n = \sup_{0 \leq i \leq n} Y_i \\ 1 - \frac{1}{2-\beta} \text{ si } n > 0 \text{ et } Y_n = \inf_{0 \leq i \leq n} Y_i \end{cases}\end{aligned}$$

Davis montre alors que le processus de Carmona, Petit et Yor est la limite spatiale de la marche aléatoire renforcée simple.

THÉORÈME 2.4. *(Davis, théorème 1.2) Si $|\rho| < 1$ et $f(t) = B_t$, alors* $\{\frac{\mathbf{Y}_{[nt]}}{\sqrt{n}}; t \geq 0\}$ *converge en loi vers* $\{X_t; t \geq 0\}$.

Notons que pour $f(t) = B_t$, il n'est *a priori* pas clair qu'il existe une solution mesurable par rapport à la filtration de B telle que $X_0 = 0$. De même, on ne sait pas s'il existe une solution X telle que $\left(X_t, \inf_{0 \leq s \leq t} X_s, \sup_{0 \leq s \leq t} X_s\right)$ soit markovien. Perman et Werner [**42**] répondent à ces questions dans le cas $\rho > 1$.

THÉORÈME 2.5. *(Perman-Werner, théorème 1) Pour tout $\alpha < 1$ et $\beta < 1$, il existe un processus $\tilde{X}$ continu, tel que $\tilde{X}_0 = 0$ et*

– *le processus* $\left(\tilde{X}_t, \inf_{0 \leq s \leq t} \tilde{X}_s, \sup_{0 \leq s \leq t} \tilde{X}_s\right)$ *est fortement markovien,*

– *le processus* $\tilde{X}_t - \beta \inf_{0 \leq s \leq t} \tilde{X}_s - \alpha \sup_{0 \leq s \leq t} \tilde{X}_s$ *est un mouvement brownien.*

En outre, supposons que $f(t) = B_t$. S'il existe une solution X à (2.1) telle que :

- *$X_0 = 0$,*
- *le triplet $\left(X_t, \inf_{0\le s\le t} X_s, \sup_{0\le s\le t} X_s\right)$ est fortement markovien,*

alors X et $\tilde{X}$ ont même loi.

THÉORÈME 2.6. *(Perman-Werner, théorème 2) Supposons que f est un mouvement brownien standard, et notons sa filtration $\mathcal{F}$. Alors, pour tout $0 \le \alpha < 1$ et $0 \le \beta < 1$, il existe presque sûrement une unique solution à l'équation (2.1) telle que $X_0 = 0$. De plus, X est $\mathcal{F}$-adapté et a la même loi que le processus $\tilde{X}$.*

Un peu plus tard, Chaumont et Doney [**16**] finalisent l'étude de l'existence et unicité des solutions par le résultat suivant, complétant l'étude précédente pour $\rho < -1$.

THÉORÈME 2.7. *(Chaumont-Doney, théorème 2) Supposons que $f(t) = B_t$ et $\alpha < 1$, $\beta < 1$. Alors il existe presque sûrement une unique solution à l'équation (2.1). De plus, cette solution est adaptée à la filtration de B.*

2.3. Comportement asymptotique des extrema. Le dernier travail dont nous voulons rapporter les résultats ici est celui de Chaumont, Doney et Hu [**18**]. Ils étudient entre autres les comportements asymptotiques des extrema de la solution de (2.1), notée X. De façon intuitive, les très grandes valeurs de X_t ne doivent dépendre que de la perturbation aux maxima de X et donc les limites supérieures de X_t doivent se comporter comme celles du processus Y correspondant à la solution de (2.1) avec $\beta = 0$. Bien entendu, dans le cas $\alpha = \beta = 0$, tous les résultats suivants correspondent à ceux bien connus sur le mouvement brownien standard. Nous donnons des énoncés lorsque t tend vers l'infini, mais on peut également donner des énoncés similaires correspondant au cas $t \to 0$.

Supposons dans toute la suite que $\alpha < 1$ et $\beta < 1$.

Un premier résultat concerne le comportement asymptotique de $\sup_{0\le s\le t} X_s$, qui implique la loi du logarithme itéré pour le mouvement brownien perturbé.

THÉORÈME 2.8. *(Chaumont-Doney-Hu, théorème 1.1) Nous supposons que $f(t) = B_t$. Soit $g > 0$ une fonction croissante. On a alors*

$$\mathbb{P}\left(\sup_{0\le s\le t} X_s > \sqrt{t}g(t) \text{ infiniment souvent}\right) = 0$$

si seulement si

$$\int_1^\infty \frac{\mathrm{d}t}{t} g(t) \exp\{-\frac{(1-\alpha)^2}{2} g(t)^2\} < \infty.$$

Similairement, la probabilité précédente vaut 1 si et seulement l'intégrale diverge.

Corollaire 2.9. *Une conséquence est la loi du logarithme itéré :*

$$\limsup_{t\to\infty} \frac{X_t}{\sqrt{2t\log\log t}} \overset{p.s.}{=} \frac{1}{1-\alpha}.$$

Remarquons que la loi du logarithme itéré précédente, pour $\alpha \geq 0$ et $\beta < 1$, peut se déduire de celle du mouvement brownien, avec l'appui du lemme de réflection de Skorokhod. En outre, le processus X est identiquement égal en loi à $-X$ où on a interverti les rôles de α et β. On déduit donc aisément des énoncés similaires pour $\inf\limits_{0\leq s\leq t} X_s$.

Nous pouvons également obtenir des résultats de logarithme itéré sur la limite inférieure du processus $\sup\limits_{0\leq s\leq t} X_s$, grâce au résultat suivant.

Théorème 2.10. *(Chaumont-Doney-Hu, théorème 1.2) Nous supposons que $f(t) = B_t$. Soit $g > 0$ une fonction croissante. On a alors*

$$\liminf_{t\to\infty} \left(\frac{g(t)}{\sqrt{t}} \sup_{0\leq s\leq t} X_s \right) \overset{p.s.}{=} 0$$

si et seulement si

$$\int_1^\infty t^{-1} g(t)^{\beta-1} \mathrm{d}t = \infty.$$

De plus, l'intégrale converge si et seulement si $\liminf\limits_{t\to\infty} \frac{g(t)}{\sqrt{t}} \sup\limits_{0\leq s\leq t} X_s \overset{p.s.}{=} \infty$.

En prenant $g(t) = (\log t)^{\varepsilon+1/(1-\beta)}$, on obtient la loi du logarithme itéré pour le processus $\sup\limits_{0\leq s\leq t} X_s$.

Corollaire 2.11. *Sous les hypothèses précédentes, on a presque sûrement*

$$\liminf_{t\to\infty} \left(\frac{(\log t)^{\varepsilon+\frac{1}{1-\beta}}}{\sqrt{t}} \sup_{0\leq s\leq t} X_s \right) = \begin{cases} 0 & si\ \varepsilon \leq 0 \\ \infty & si\ \varepsilon > 0 \end{cases}$$

Nous mettons en avant le fait que l'intégrale impliquée dans le théorème ne dépend pas de α. Cela signifie que les petites valeurs de $\sup\limits_{0\leq s\leq t} X_s$ impliquent une perturbation en les minima de X.

Enfin, un dernier résultat montre en quelque sorte comment les deux perturbations en le minimum et en le maximum peuvent soit s'annuler mutuellement, soit au contraire se renforcer mutuellement.

Théorème 2.12. *(Chaumont-Doney-Hu, théorème 1.3) Nous supposons que $f(t) = B_t$. Soit $g > 0$ une fonction croissante. On a alors*

$$\mathbb{P}\left(\sup_{0\leq s\leq t} |X_s| < \frac{\sqrt{t}}{g(t)} \ \textit{infiniment souvent} \right) = 0$$

si et seulement si

$$\int_1^\infty \mathrm{d}t g(t)^{2(1-\beta-\alpha)} t^{-1} \exp\{-\frac{\pi^2}{8} g(t)^2\} < \infty.$$

De plus, l'intégrale diverge si et seulement si la probabilité précédente vaut 1.

COROLLAIRE 2.13. *Sous les hypothèses précédentes, on obtient la loi du logarithme itéré de Chung :*

$$\liminf_{t\to\infty} \left\{ \sqrt{\frac{\log\log t}{t}} \sup_{0\le s\le t} |X_s| \right\} \overset{p.s.}{=} \frac{\pi}{\sqrt{8}}.$$

3. Diffusions renforcées par la mesure d'occupation non normalisée

On étudie ici un processus, en fait un mouvement brownien avec dérive, dont la dérive dépend de façon naturelle du passé *via* sa mesure d'occupation. Durrett et Rogers [**27**] proposent d'étudier un objet mathématique modélisant l'évolution d'un polymère régie selon un certain processus, appelé 'polymère brownien'. Ils suggèrent alors de choisir une fonction f et étudient la solution de l'équation différentielle stochastique (E.D.S.) suivante à valeurs dans $\mathbb{R}^d$

$$\mathrm{d}X_t = \mathrm{d}B_t + \int_0^t \mathrm{d}s f(X_t - X_s)\mathrm{d}t \tag{3.1}$$

où B est un mouvement brownien standard. Tout d'abord, remarquons que pour toute fonction f mesurable bornée, il existe une unique solution faible à l'équation (3.1). Ce résultat provient de la formule de changement de dérive. Si de plus f est continue, alors il y a au plus une solution forte. Bien entendu, si f est une fonction lipschitzienne, un résultat standard de Rogers et Williams ([**45**] théorème 11.2) montre que l'équation précédente admet une unique solution forte.

3.1. Premiers résultats. Durrett et Rogers sont les pionniers concernant l'étude de tels processus. Une façon de comprendre l'équation (3.1) est d'imaginer par exemple une particule qui, dans le cas répulsif, a plutôt tendance à s'éloigner des sites déjà visités ; et, dans le cas attractif, préfère retouner aux endroits qu'elle a déjà visité. De manière plus "physicienne", le processus X_t correspond à l'emplacement de l'extrémité d'un polymère à l'instant t. Dans leur article, Durrett et Rogers prouvent certains résultats, mais surtout émettent plusieurs conjectures. Voici une vue d'ensemble de son contenu.

THÉORÈME 3.1. *(Durrett-Rogers, théorème 1) Supposons qu'il existe une constante M telle que $||f(x)|| \le M$ et f est à support compact. Alors, il* *existe une constante $\Gamma < \infty$ telle que*

$$\limsup_{t\to\infty} \frac{||X_t||}{t} \overset{p.s.}{\leq} \Gamma.$$

THÉORÈME 3.2. *(Durrett-Rogers, théorème 2) Supposons $d = 1$, f est une fonction positive ou nulle, telle que $f(0) > 0$. Il existe alors une constante $\gamma > 0$ telle que*

$$\liminf_{t\to\infty} \frac{||X_t||}{t} \overset{p.s.}{\geq} \gamma.$$

Cranston et Mountford [**21**] ont prouvé la conjecture suivante, partiellement émise par Durrett et Rogers.

THÉORÈME 3.3. *(Cranston-Mountford, théorème 1) Soit $d = 1$. Supposons que f est une fonction continue, positive ou nulle mais non identiquement nulle, lipschitzienne à support compact. Alors il existe une constante $c > 0$ telle que*

$$\lim_{t\to\infty} \frac{X_t}{t} \overset{p.s.}{=} c.$$

DÉMONSTRATION. Nous allons uniquement donner l'intuition menant à ce résultat. Pour ce faire, fixons $T > 0$ et considérons

$$Y_t = B_t + \int_0^t \mathrm{d}s \int_{(s-T)^+}^s \mathrm{d}u f(Y_s - Y_u).$$

Il s'agit du même type de modèle que pour le processus X, mais on ne regarde l'interaction avec le passé que pendant une durée T. Pour $t \geq T$, le processus $Z_t = \{Y_t - Y_{t-s}, 0 \leq s \leq T\}$ est une chaîne de Harris exponentiellement ergodique, d'espace d'état $\mathcal{C}([0,T];\mathbb{R})$ et donc par le théorème limite quotient, Y_t/t converge presque sûrement vers la moyenne du drift. Similairement, il est raisonnable de penser que $\{X_t - X_{t-s}, 0 \leq s \leq t\}$ converge rapidement vers l'équilibre. $\square$

Parmi les autres exemples d'applications f considérées par Durrett et Rogers, la suivante est étudiée en détails. Supposons que $f : \mathbb{R} \to \mathbb{R}$ est telle que

(1) $|f(x)| \leq M$,

(2) f est strictement décroissante sur l'intervalle $[y, \infty)$,

(3) il existe une constante $0 < \beta < 1$ telle que $x^\beta f(x) \underset{x\to\infty}{\longrightarrow} l > 0$.

Posons maintenant $\alpha := 2/(1+\beta)$ et soit c_0 la solution de l'équation

$$\alpha c_0^{\beta+1} = l \int_0^1 \mathrm{d}u (1 - u^\alpha)^{-\beta}.$$

Théorème 3.4. *(Durrett-Rogers, théorème 3) Sous les hypothèses ci-dessus, on a alors presque sûrement*

$$\limsup_{t\to\infty} \frac{X_t}{t^\alpha} \leq c_0.$$

Si de plus, f est positive ou nulle et $f(0) > 0$, alors

$$\frac{X_t}{t^\alpha} \xrightarrow[t\to\infty]{p.s.} c_0.$$

La conjecture suivante, initialement émise par Durrett et Rogers, est alors prouvée par Mountford et Tarrès [**38**].

Théorème 3.5. *(Mountford-Tarrès) Supposons que $f(x) = \frac{x}{1+|x|^{\beta+1}}$ avec $0 < \beta < 1$. Alors avec une probabilité de $1/2$, on a*

$$\frac{X_t}{t^\alpha} \underset{t\to\infty}{\to} c_0.$$

Le point essentiel de la preuve consiste à montrer que $\limsup\limits_{t\to\infty} |X_t| \overset{p.s.}{=} \infty$.

Après tous les résultats démontrés précédemment, il reste une dernière conjecture, non prouvée jusqu'à présent :

Conjecture 3.6. *(Durrett-Rogers, conjecture 2) Supposons que f est une fonction à support compact, impaire. Alors*

$$\frac{X_t}{t} \xrightarrow[t\to\infty]{p.s.} 0.$$

De plus, Tóth [**47**] a conjecturé plus tard (en comparant le modèle précédent avec la marche aléatoire auto-évitante exponentielle sur $\mathbb{Z}$) que, sous les conditions précédentes, $\frac{X_t}{t^{2/3}}$ converge en loi.

3.2. Généralisation. Ensuite, Cranston et Le Jan [**20**] ont considéré le modèle unidimensionnel de Durrett et Rogers pour deux applications différentes. Lorsque $f(x) = -ax$ avec $a > 0$, on a alors une force de rappel de la localisation actuelle vers tout le passé. Ce rappel étant une fonction croissante non majorée, il n'est pas surprenant que le processus X_t converge presque sûrement vers la moyenne de la limite de la mesure d'occupation, soit

Théorème 3.7. *(Cranston-Le Jan, théorème 1) Soit $a > 0$ et $f(x) = -ax$. Il existe une variable aléatoire X_∞ telle que $X_t \underset{t\to\infty}{\longrightarrow} X_\infty$ presque sûrement et dans L^2.*

Démonstration. La démonstration de ce résultat n'est pas difficile et repose totalement sur le fait que l'E.D.S.

$$\mathrm{d}X_t = \mathrm{d}B_t - a\int_0^t \mathrm{d}s(X_t - X_s)\mathrm{d}t$$

admet une unique solution (forte) qui peut être connue explicitement (car X est un processus gaussien !)

$$X_t = \int_0^t \left(1 - ase^{as^2/2} \int_s^t e^{-au^2/2}\mathrm{d}u\right) \mathrm{d}B_s.$$

□

Cranston et Le Jan ont également traité le cas $f(x) = -a\mathrm{sgn}(x)$. Malgré la discontinuité de f, on peut aisément prouver l'existence et unicité de la solution de l'E.D.S. Pour cette fonction, la dérive représente une force de rappel vers la médiane et non plus vers la moyenne, mais le même résultat reste vrai.

THÉORÈME 3.8. *(Cranston-Le Jan, théorème 2) Soit $a > 0$ et $f(x) = -asgn(x)$. Il existe une variable aléatoire X_∞ telle que $X_t \underset{t\to\infty}{\longrightarrow} X_\infty$ presque sûrement et dans L^2.*

La preuve repose essentiellement sur l'étude de l'E.D.S. suivante

$$\mathrm{d}Y_t = \mathrm{d}B_t - \left(a\int_0^t \mathrm{sgn}(Y_t - Y_s)\mathrm{d}s + V(Y_t)\right)\mathrm{d}t, \;\; Y_0 = 0$$

où V est une fonction mesurable, positive sur $\mathbb{R}_+$. Nous remarquons que le processus couplé $(Y_t, \int_0^t \delta_{Y_s}\mathrm{d}s)$ est markovien. Les auteurs montrent alors que $\sup_{t\geq 0} Y_t < \infty$ presque sûrement et que si pour tout $\varepsilon > 0$, il existe $M(\varepsilon) > 0$ tel que $V(x) \geq M(\varepsilon)$ sur $\mathbb{R}_+$, alors $\mathbb{P}\left(\sup_{t>0} Y_t > \varepsilon\right) < \varepsilon$. Cela implique que X est borné.

Raimond [**43**] a ensuite étendu ce dernier résultat à la dimension $d \geq 2$.

THÉORÈME 3.9. *(Raimond, théorème 1) Soit $a > 0$ et $f(x) = -a\frac{x}{||x||}$. Considérons le mouvement brownien d-dimensionnel B et soit X le processus solution de*

$$\mathrm{d}X_t = \mathrm{d}B_t - a\int_0^t \mathrm{d}s\frac{X_t - X_s}{||X_t - X_s||}\mathrm{d}t, \;\; X_0 = 0.$$

Alors X_t converge presque sûrement.

La ligne directrice de la démonstration est la suivante. On utilise la même méthode que Cranston et Le Jan, en introduisant le processus Y et en supposant de plus que $(\nabla V(x), x) \geq 0$. Les choses sont cependant bien plus compliquées ici car ces derniers utilisent un argument purement unidimensionnel et de plus, en dimension $d \geq 2$, le processus peut tourner autour d'un point, ce qui implique alors que X ne converge pas ! Il faut donc étudier en détail la dérive et, pour ce faire, introduire le point H_t qui la minimise : $\int_0^t \mathrm{d}s\frac{H_t - X_s}{||H_t - X_s||} = 0$.

L'idée est maintenant la suivante. Si X_t reste proche de H_t pendant un temps grand, alors la portion de trajectoire proche de ce processus minimisant va créer une importante force de rappel vers ce dernier. Ainsi, pour les temps grands, le processus X reste proche de H et va donc converger. Les ingrédients de la preuve sont un résultat de comparaison entre X et H et le fait que le processus joint $\left(X_t, \int_0^t \delta_{X_s} \mathrm{d}s\right)$ est markovien. Nous ne donnons pas les détails de la preuve, qui est fort technique.

Finalement, Herrmann et Roynette [**29**] généralisent dans une autre direction le résultat de convergence de Cranston et Le Jan. Ils travaillent en dimension $d = 1$ et considèrent une fonction f impaire et bornée. Leur principal résultat est le suivant.

Théorème 3.10. *(Herrmann-Roynette, théorème 1) Supposons qu'il existe des constantes $C > 0$, $\rho > 0$ et $k \in \mathbb{N}^*$ telles que $|f(x)| \geq Ce^{-\rho|x|^{-k}}$ dans un voisinage de l'origine. Alors X_t converge presque sûrement.*

La preuve de ce théorème est une adaptation des démonstrations des résultats précédents de Cranston & Le Jan et Raimond.

Herrmann et Roynette étudient notamment l'unique solution faible de l'équation

$$X_t = B_t - \int_0^t \mathrm{d}s \int_0^s \mathrm{d}u \ \operatorname{sgn}(X_s - X_u) \mathbb{1}_{\{|X_s - X_u| \geq 2\}}. \tag{3.2}$$

Théorème 3.11. *(Herrmann-Roynette, théorème 2) Le processus X satisfaisant l'équation (3.2) est presque sûrement borné.*

Remarquons qu'il suffit de montrer que le processus est minoré et on conclut la majoration (et donc le résultat) par symétrie.

Soit le temps d'arrêt $\tau_n := \inf\{t \geq 0; X_t = n\}$ pour $n \geq 4$. Considérons l'événement $\Omega_n := \{\tau_n < \infty\}$ et supposons que $\mathbb{P}(\Omega_n) > 0$ (sinon, X est trivialement minoré). L'idée consiste à montrer que, conditionnellement à l'événement Ω_n, la probabilité que le processus n'atteigne jamais le niveau $n + 2$ est strictement positive ($\mathbb{P}(\tau_{n+2} < \infty | \Omega_n) < 1$) et indépendante de n. La preuve est très technique et n'est pas reproduite ici.

Enfin, Herrmann et Scheutzow [**30**] étudient la vitesse de convergence de la diffusion renforcée convergente sous les mêmes hypothèses que Herrmann et Roynette.

Théorème 3.12. *(Herrmann-Scheutzow, théorème 1) Si f est une fonction impaire, de classe $\mathcal{C}^1$ et s'il existe des constantes $\eta > 0$, $\gamma \geq 1$ et $C_\gamma > 0$ telles que pour tout $|x - y| \leq \eta$, on a $|f(x) - f(y)| \geq C_\gamma |x - y|^\gamma$, alors pour tout $\mu < \frac{1}{1+\gamma}$, on obtient presque sûrement pour le processus X satisfaisant l'équation (3.2)*

$$\lim_{t \to \infty} \left(\frac{t}{\log t}\right)^\mu \sup_{t \leq T} |X_T - X_t| = 0.$$

Démonstration. Elle est basée sur un résultat de comparaison. Nous donnons la preuve dans le cas simplifié où f est linéaire, $f(x) = -ax$ avec $a > 0$. On sait, grâce à Cranston et Le Jan, que X_t converge presque sûrement vers X_∞, dont nous connaissons l'expression explicite. Nous avons

$$X_t = \int_0^t h(t,s)\mathrm{d}B_s, \;\; X_\infty = \int_0^\infty h(s)\mathrm{d}B_s,$$

avec $h(t,s) = 1 - ase^{as^2/2}\int_s^t e^{-au^2/2}\mathrm{d}u$ et $h(s) = \lim_{t\to\infty} h(t,s)$. Ecrivons $X_t - X_\infty = S_t - V_t$ avec $V_t = \int_t^\infty h(s)\mathrm{d}B_s$ et $S_t = \int_t^\infty e^{-au^2/2}\mathrm{d}u \int_0^t ase^{as^2/2}\mathrm{d}B_s$. Par le théorème de Dubins-Schwartz, le processus S s'écrit comme un mouvement brownien changé de temps et par la loi du logarithme itéré, on obtient $\limsup_{t\to\infty}\sqrt{\frac{t}{\log t}}S_t = \sqrt{\frac{2}{a}}$. Il ne reste plus qu'à étudier le processus V. L'idée est d'inverser l'échelle de temps et de considérer $M_t := \int_{t^{-1}}^\infty h(s)\mathrm{d}B_s$, qui est une martingale pour la filtration changée d'échelle de temps. Comme $< M >_\infty$ est bornée, on peut grossir la filtration pour que M s'écrive comme un mouvement brownien changé de temps. Ainsi, on utilise la loi du logarithme itéré pour ce mouvement brownien, ce qui donne

$$\limsup_{t\to\infty}\sqrt{\frac{t}{\log t}}V_t = \limsup_{t\to 0}\frac{1}{\sqrt{-t\log t}}M_t = 0.$$

On obtient le résultat cherché en utilisant la symétrie de X. □

4. Diffusions renforcées par la mesure d'occupation normalisée

Ces diffusions ont été introduites par Benaïm, Ledoux et Raimond [**7**]. Ici, la dérive est une fonction de la mesure d'occupation *normalisée*. Celles-ci sont plus proches, dans l'idée, des marches aléatoires renforcées que le modèle précédent, proposé par Durrett et Rogers [**27**]. Une autre différence importante est que Benaïm, Ledoux et Raimond travaillent sur une variété riemannienne de dimension d, sans bord et *compacte*. Nous allons dans ce paragraphe exposer une partie des résultats contenus dans les articles de Benaïm, Ledoux et Raimond [**7**] ainsi que Benaïm et Raimond [**8**].

4.1. Outils et définition. Nous considérons donc des processus à valeurs dans une variété riemannienne notée M, qui est compacte, complète, connexe et sans bord. Nous notons $\mathcal{M}(M)$ l'ensemble des mesures bornées de M et $\mathcal{P}(M)$ l'ensemble des probabilités sur M. Entrons maintenant dans le vif du sujet et définissons les diffusions renforcées. Considérons une mesure d'occupation μ. Lorsqu'elle est normalisée, nous ajoutons un poids initial r afin d'éviter des problèmes de convergence.

Soit $W : M \times M \to M$ un potentiel d'interaction, suffisamment lisse (par exemple de classe $\mathcal{C}^2$ en la première variable). Considérons une probabilité μ sur M. Nous définissons également la 'convolution' entre μ et W de la manière suivante :

$$W * \mu(x) := \int_M W(x,y)\mathrm{d}\mu(y). \tag{4.1}$$

De plus, nous introduisons l'opérateur A_μ, défini sur $\mathcal{C}^\infty(M)$ par

$$A_\mu f = \frac{1}{2}\Delta f - (\nabla W * \mu, \nabla f) = \frac{1}{2}e^{2W*\mu}\mathrm{div}(e^{-2W*\mu}\nabla f).$$

DÉFINITION 4.1. *Un ensemble de lois $\{\mathbb{P}_{x,r,\mu}; x \in M, r > 0, \mu \in \mathcal{P}(M)\}$ définit une diffusion renforcée (encore appelée diffusion auto-interagissante ou auto-interactive) si les deux conditions suivantes sont satisfaites :*

1. $\mathbb{P}_{x,r,\mu}(X_0 = x) = 1$ *pour tout triplet* (x,r,μ) *;*

2. *pour toute fonction $f \in \mathcal{C}^\infty(M)$ et tout triplet (x,r,μ), le processus M^f est une $\mathbb{P}_{x,r,\mu}$-martingale où M^f est défini par*

$$M_t^f := f(X_t) - f(X_0) - \int_0^t A_{\mu_s} f(X_s)\mathrm{d}s.$$

La fonction considérée ici dans la dérive sera toujours $W * \mu_t$. Ainsi, nous écrivons l'E.D.S. satisfaite par le processus renforcé X comme

$$\mathrm{d}X_t = \sum_{i=1}^N F_i(X_t)\circ\mathrm{d}B_t^i - \left(\frac{r}{r+t}\int_M \nabla_x W(X_t,y)\mathrm{d}\mu(y) + \frac{1}{r+t}\int_0^t \nabla_x W(X_t,X_s)\mathrm{d}s\right)\mathrm{d}t, \tag{4.2}$$

où r est un poids donné (strictement positif), μ est la probabilité initiale et $B = (B^1, \cdots, B^N)$ est un mouvement brownien standard sur $\mathbb{R}^N$. Ici, le symbole $\circ$ signifie que l'intégration est à comprendre au sens de Stratonovich. $(F_i)_{1\le i\le N}$ est la famille vectorielle apparaissant dans la décomposition de Hörmander de l'opérateur de Laplace-Beltrami sur M :

$$\Delta = \sum_{i=1}^N F_i^2.$$

Une première étape consiste à prouver l'existence du processus X partant de n'importe quel point de M et de toute mesure d'occupation (probabilité) initiale.

PROPOSITION 4.2. *(Benaïm-Ledoux-Raimond, proposition 2.5) Pour chaque application W, il existe une unique diffusion renforcée.*

DÉMONSTRATION. Nous pouvons supposer que M est une sous-variété isométriquement plongée dans $\mathbb{R}^N$, pour N suffisamment grand (Nash, 1956). Le problème est de déterminer des champs de vecteurs F_i donnant le résultat recherché. Commençons par introduire le mouvement brownien standard sur la variété M. Soit $e = (e_1, \cdots, e_N)$ une base de $\mathbb{R}^N$. Notons $F_i(x)$ la projection orthogonale de e_i sur le plan tangent à M en $x \in M$, soit T_xM.

Prolongeons le champ de vecteurs F_i à $\mathbb{R}^N$ de façon $\mathcal{C}^\infty$. On construit le mouvement brownien sur la variété M comme solution de l'E.D.S. sur $\mathbb{R}^N$

$$\mathrm{d}Y_t = \sum_{i=1}^{N} F_i(Y_t) \circ \mathrm{d}B_t^i.$$

Comme les F_i sont par définition tangents à la variété, on peut espérer qu'une telle solution partant d'un point de M reste sur cette variété (pour le voir, on peut remarquer que la fonction $g(x) := ||x - M||^2$ a M comme ligne de niveau. Nous renvoyons le lecteur au livre de Rogers et Williams pour plus de précisions).

La construction de la diffusion renforcée est analogue à celle du mouvement brownien sur M (il faudrait également montrer que X reste bien sur la variété). Il faut étendre la fonction W à une fonction $\bar{W}$ définie sur l'espace ambiant et de classe $\mathcal{C}^\infty$. Comme pour tout $x \in M$, $\nabla W * \mu(x)$ est la projection orthogonale de $\nabla\bar{W} * \mu(x)$ sur l'espace tangent T_xM, on obtient

$$(4.3)\quad \nabla W * \mu(x) = \sum_{i=1}^{N} (\nabla W * \mu(x), e_i) F_i(x) = \sum_{i=1}^{N} (\nabla\bar{W} * \mu(x), F_i(x)) F_i(x).$$

Nous pouvons maintenant travailler sur $\mathbb{R}^N$ et considérer l'E.D.S. suivante

$$(4.4)\quad \mathrm{d}X_t = \sum_{i=1}^{N} F_i(X_t) \circ \mathrm{d}B_t^i - \sum_{i=1}^{N} (\nabla\bar{W} * \mu_t(X_t), F_i(X_t)) F_i(X_t)\mathrm{d}t,\ \ X_0 = x.$$

Cette E.D.S. admet une solution faible (X_t, B_t) et notons $\mathbb{P}_{x,r,\mu}$ la loi de X (qui est bien une probabilité sur $(\Omega, \mathcal{F}(B))$). Pour chaque $x \in M$ et chaque fonction $f \in \mathcal{C}^\infty(\mathbb{R}^N)$ à support compact, la formule d'Itô généralisée implique que $f(X_t) - f(x) - \int_0^t A_{\mu_s} f(X_s)\mathrm{d}s$ est une martingale pour la filtration du mouvement brownien. De plus, la martingale précédente est en fait M_t^f par définition du laplacien de Laplace-Beltrami. On a donc bien une diffusion renforcée, comme définie au début du paragraphe. □

Remarquons que le résultat précédent prouve également que la mesure d'occupation $(\mu_t, t \geq 0)$ est bien définie.

4.2. Système dynamique. Le point suivant est de savoir ce qu'il advient au temps $t + \delta t$ lorsqu'on fixe la dérive à $-\nabla W * \mu_t(X_t)$. Une idée naturelle est alors de généraliser l'étude d'une diffusion standard en définissant une fonction Π, qui à une mesure μ associe la densité d'une diffusion standard dont la dérive est $-\nabla W * \mu$. La probabilité invariante d'un tel processus s'écrit

$$(4.5)\qquad \Pi(\mu)(\mathrm{d}x) := \frac{e^{-2W*\mu(x)}}{\int_M e^{-2W*\mu(z)}\mathrm{d}z}\mathrm{d}x.$$

Nous mettons en avant le fait que le processus μ_t est *aléatoire* ! Le théorème ergodique standard pour les vraies diffusions (appelé théorème limite quotient) implique alors que μ_t converge presque sûrement (lorsque t tend vers l'infini) vers l'unique probabilité invariante $\Pi(\mu)$. Cela donne l'intuition que $\mu_{t+\delta t}$, pour δt infinitésimal, vaut approximativement $\mu_t + \frac{\delta t}{t}(\Pi(\mu_t) - \mu_t)$. Nous justifions ainsi l'idée d'approximer l'évolution aléatoire de μ_t par une évolution déterministe, où μ_t évolue à chaque instant vers $\Pi(\mu_t)$. Introduisons donc le champ de vecteurs $F : \mathcal{M}(M) \to \mathcal{M}(M)$, défini sur l'espace des mesures par

$$F(\mu) := \Pi(\mu) - \mu. \tag{4.6}$$

Ce champ de vecteurs induit un flot Φ régulier, comme le montre le résultat suivant.

Proposition 4.3. *(Benaïm-Ledoux-Raimond, lemme 3.1) Soit F le champ de vecteurs défini sur $\mathcal{M}(M)$ par $F(\mu) := \Pi(\mu) - \mu$. Ce champ de vecteurs est $\mathcal{C}^\infty$ et induit un flot Φ de classe $\mathcal{C}^\infty$ défini par*

$$\Phi_0(\mu) = \mu, \ \ \frac{\mathrm{d}}{\mathrm{d}t}\Phi_t(\mu) = F(\Phi_t(\mu)).$$

De plus, ce flot laisse l'ensemble $\mathcal{P}(M)$ positivement invariant.

Démonstration. Le fait que F est $\mathcal{C}^\infty$ se déduit de l'expression de $\Pi(\mu)$. De plus, la différentielle de Π en μ est donnée par l'expression

$$D\Pi(\mu)\cdot\nu(\mathrm{d}x) = -2\left(W * \nu(x) - \int_M W * \nu(y)\Pi(\mu)(\mathrm{d}y)\right)\Pi(\mu)(\mathrm{d}x).$$

En majorant la norme de F, on obtient que F est lipschitzienne pour la topologie forte. Donc F définit un flot, qui est de plus $\mathcal{C}^\infty$.

Montrons maintenant que Φ laisse l'ensemble convexe $\mathcal{P}(M)$ positivement invariant. Soit $\mu \in \mathcal{P}(M)$. Notons la distance de la mesure μ à $\mathcal{P}(M)$ pour la topologie forte par $d(\mu, \mathcal{P}(M)) := \inf\{|\mu - \nu|; \nu \in \mathcal{P}(M)\}$. On a pour tout $\nu \in \mathcal{P}(M)$

$$|\Phi_t(\mu) - \nu| = |(1-t)\mu + t\Pi(\mu) - \nu + o(t)| \leq |(1-t)\mu + t\Pi(\mu) - \nu| + o(t).$$

Remarquons que $\Pi(\mu) \in \mathcal{P}(M)$ et d est une fonction convexe. Cela implique alors

$$d(\Pi_t(\mu), \mathcal{P}(M)) \leq (1-t)d(\mu, \mathcal{P}(M)) + o(t).$$

Par convexité, l'application $t \mapsto d(\Phi_t(\mu), \mathcal{P}(M))$ est dérivable à droite, de dérivée

$$\frac{\mathrm{d}^+}{\mathrm{d}t} d(\Phi_t(\mu), \mathcal{P}(M))\Big|_{t=0} \leq -d(\mu, \mathcal{P}(M)).$$

Par la propriété de flot, cette inégalité s'étend pour tout t et on obtient finalement pour tout $t \geq 0$

$$\Phi_t(\mathcal{P}(M)) \subset \mathcal{P}(M).$$

□

Nous rappelons que la topologie *faible* sur l'ensemble des mesures bornées sur M, soit $\mathcal{M}(M)$ est la topologie induite par la famille de semi-normes $\{\mu \mapsto \left|\int_M f(x)\mathrm{d}\mu(x)\right|; f \in \mathcal{C}(M)\}$. La topologie *forte* quant à elle est induite par la famille de normes $\left\{\mu \mapsto |\mu| = \sup\{\left|\int_M f(x)\mathrm{d}\mu(x)\right|; f \in \mathcal{C}(M), ||f||_\infty = 1\}\right\}$. Notons également que $(\mathcal{M}(M), |\cdot|)$ est un espace de Banach (il s'agit de l'espace dual de $\mathcal{C}(M)$). Dans la suite de ce paragraphe, nous naviguerons constamment entre la topologie faible et la topologie forte.

Nous avons maintenant défini un flot sur M pour la topologie forte. Or l'espace $\mathcal{P}(M)$ est compact pour la topologie faible, donc il est intéressant de travailler avec cette dernière. Nous considérons alors l'application pour la topologie *faible*

$$\Psi : \mathbb{R} \times \mathcal{M}(M) \to \mathcal{M}(M); \;\; \Psi(t, \mu) = \Phi_t(\mu).$$

Benaïm, Ledoux et Raimond ont prouvé le résultat suivant, dont nous admettrons la preuve (qui repose sur une aproximation de μ) :

LEMME 4.4. *(Benaïm-Ledoux-Raimond, lemme 3.3) L'application Ψ est continue pour la topologie faible.*

En ce qui concerne les définitions relatives aux systèmes dynamiques, nous renvoyons le lecteur à l'introduction de ce document. Dans la suite, nous appellerons Φ le flot défini par Ψ. Le but de ce travail est de montrer que la mesure d'occupation μ_t est proche (dans un sens à définir) du flot Φ. Ainsi, cela nous permettra d'étudier un objet déterministe afin d'obtenir des informations sur un processus aléatoire. En fait, nous verrons qu'il faudra considérer μ_{e^t} au lieu de μ_t, ceci s'expliquant par

$$\frac{\mathrm{d}}{\mathrm{d}t}\mu_{e^t} = \frac{e^t}{r + e^t}\left(\delta_{X_e^t} - \mu_{e^t}\right).$$

4.3. Résultats. Passons maintenant à la notion de pseudo-trajectoire asymptotique. On considère un flot Φ sur un espace métrique (E, d).

DÉFINITION 4.5. *Une fonction continue $\xi : \mathbb{R} \to E$ est une pseudo-trajectoire asymptotique du flot Φ si pour tout temps $T > 0$, on a*

$$\lim_{t\to\infty} \sup_{0\le s\le T} d(\xi_{t+s}, \Phi_s(\xi_t)) = 0.$$

Définissons la famille de mesures $(\varepsilon_{t,t+s}; t, s \geq 0)$ par

$$\varepsilon_{t,t+s} := \int_t^{t+s} \left(\delta_{X_{e^u}} - \Pi(\mu_{e^u})\right) \mathrm{d}u. \tag{4.7}$$

Cette famille jouera un rôle technique important dans la suite.

Le lien entre la famille $\varepsilon_{t,t+s}$ définie par (4.7) et la notion de pseudo-trajectoire asymptotique est donnée par le résultat suivant

LEMME 4.6. *(Benaïm-Ledoux-Raimond, proposition 3.5) La fonction ξ : $\mathbb{R}_+ \to \mathcal{P}(M)$, $t \mapsto \mu_{e^t}$ est une pseudo-trajectoire asymptotique pour le flot*

Φ si et seulement si pour toute fonction $f \in \mathcal{C}^\infty(M)$, tout $T > 0$, on a $\lim_{t\to\infty} \sup_{0\le s\le T} |\varepsilon_{t,t+s} f| = 0$.

Le point essentiel du travail fourni par Benaïm, Ledoux et Raimond est de prouver le résultat suivant (dont nous ne donnons pas la preuve, qui est très technique).

Proposition 4.7. *(Benaïm-Ledoux-Raimond, théorème 3.6) Pour toute fonction $f \in \mathcal{C}^\infty(M)$ et tout $T > 0$, il existe une constant $K > 0$ (dépendant uniquement de W et M) telle que pour tout $\delta > 0$*

$$\mathbb{P}_{x,r,\mu}\left(\sup_{0\le s\le T} |\varepsilon_{t,t+s} f| \ge \delta | \mathcal{F}_{e^t}\right) \le \frac{K}{\delta^2} ||f||_\infty e^{-t}.$$

Corollaire 4.8. *La fonction $\xi : \mathbb{R}_+ \to \mathcal{P}(M)$, $t \mapsto \mu_{e^t}$ est $\mathbb{P}_{x,r,\mu}$-presque sûrement une pseudo-trajectoire asymptotique pour le flot Φ.*

Par analogie avec l'ensemble ω-limite d'une trajectoire d'un flot, introduisons l'ensemble limite d'une pseudo-trajectoire asymptotique.

Définition 4.9. *L'ensemble limite d'un processus μ_t est l'ensemble aléatoire*

$$L(\mu_t) := \left\{\bigcap_{t\ge 0} \overline{\{\mu_s ; s \ge t\}}\right\}.$$

Maintenant, le résultat primordial permettant de passer d'une information sur le flot Φ déterministe à une connaissance sur la mesure d'occupation normalisée (μ_t) aléatoire est le suivant. Benaïm et Hirsch (voir [**6**] ou [**5**]) ont prouvé que si ξ est une pseudo-trajectoire asymptotique relativement compacte pour le flot Φ sur un ensemble métrique E, alors l'ensemble limite $L(\xi)$ est compact, connexe et n'a pas d'attracteur propre pour le flot restreint à E. Dans le cas traité ici, cela implique le théorème :

Théorème 4.10. *(Benaïm-Ledoux-Raimond, théorème 3.8) Comme $t \mapsto \mu_{e^t}$ est une pseudo-trajectoire asymptotique pour le flot Φ_t sur l'espace métrique compact $\mathcal{P}(M)$, alors $L(\mu_t)$ est un ensemble compact connexe et invariant par Φ et (avec probabilité 1) n'a pas d'attracteur propre pour le flot restreint à $\mathcal{P}(M)$.*

Dans certains cas, il est possible de décrire assez précisément $L(\mu_t)$, comme le montre le résultat suivant.

Théorème 4.11. *(Benaïm-Raimond 2005, théorème 2.4) Supposons que W est symétrique (i.e. $W(x,y) = W(y,x)$). Alors l'ensemble limite $L(\mu_t)$ est un sous-ensemble compact (pour la topologie faible) connexe des points fixes de* Π.

Démonstration. Définissons l'énergie libre d'une fonction $f \in L^2$ strictement positive par

$$\mathcal{F}(f) := \frac{1}{2} < W * f, f > + < f, \log f >$$

avec les notations désormais usuelles $W * f(x) = \int_M W(x,y)f(y)\mathrm{d}y$ et $< f, g >= \int_M f(x)g(x)\mathrm{d}x$. On peut aisément vérifier que $\mathcal{F}$ est une fonction de Lyapunov pour le flot Φ et que $F(\mu) = 0$ si et seulement si μ est une mesure absolument continue par rapport à la mesure de Lebesgue sur M, de densité f, qui est un point critique de l'énergie libre : $\nabla\mathcal{F}(f) = 0$ (voir [**8**], proposition 2.9). Cela signifie que les points critiques de $\mathcal{F}$ sont les zéros de F. Or, nous avons vu dans le théorème précédent que μ_{e^t} est une pseudo-trajectoire asymptotique pour le flot Φ. Cela implique alors entre autres que $L(\mu_t)$ est un ensemble compact (pour la topologie faible) connexe. Nous savons de plus par le résultat précédent que $L(\mu_t)$ n'a pas d'attracteur propre pour Φ. Comme le flot Φ admet un fonction de Lyapunov, cela suffit à montrer que l'ensemble limite de μ_t est un sous-ensemble des points fixes de Π. $\square$

Corollaire 4.12. *Supposons que W est symétrique. Si de plus les points fixes de Π sont tous des points isolés, alors μ_t converge presque sûrement vers un de ces points fixes.*

Nous remarquons ici que si Π a un unique point fixe, alors μ_t converge presque sûrement vers ce point fixe. Une condition suffisante pour obtenir ce résultat est que $\mathcal{F}$ soit strictement convexe (ce qui est le cas par exemple si W est strictement convexe) et alors μ_t converge presque sûrement vers le minimum global de $\mathcal{F}$. Pour plus de précisions sur l'énergie libre, nous renvoyons le lecteur à McCann [**37**] ou Villani [**49**].

Nous pouvons même en dire plus dans le cas de convergence (lorsque W symétrique).

Théorème 4.13. *(Benaïm-Raimond 2005, théorème 2.24) Soit μ_∞ un point fixe de Π. Si μ_∞ est stable, alors $\mathbb{P}(\mu_t \to \mu_\infty) > 0$.*

En revanche, si μ_∞ est un point fixe de Π instable ou un point selle, alors $\mathbb{P}(\mu_t \to \mu_\infty) = 0$.

4.4. Applications. Nous allons rapidement exploiter les résultats précédents sur un exemple (tiré de [**7**]) lorsque W n'est pas symétrique. Nous ne donnons pas de démonstration des résultats. Dans ce cas, il se peut qu'il n'existe pas de fonction de Lyapunov et que l'ensemble limite L soit une orbite non triviale. Ici, comme μ_t est une pseudo-trajectoire asymptotique du flot Φ, on a l'intuition que μ_t va tourner autour de l'orbite en question à une vitesse logarithmique (à cause du changement de temps).

Exemple 4.14. *Supposons $M = S^1$ et définissons pour $c \in \mathbb{R}$ et $\phi \in [0, 2\pi[$ la fonction 2π-périodique*

$$W(x,y) := 2c\cos(x - y + \phi) = -cd^2(x + \phi, y) + 2c,$$

où d est la distance sur S^1 (vu comme un sous-ensemble de $\mathbb{C}$), soit $d(x,y) = |e^{ix} - e^{iy}|$. En posant $W_t(\alpha) := \frac{1}{t}\int_0^t d^2(\alpha, X_s + \phi)$ et $W_t'(\alpha) = \partial_\alpha W_t(\alpha)$, cela revient à étudier l'E.D.S.

$$\mathrm{d}X_t = \mathrm{d}B_t + cW_t'(X_t)\mathrm{d}t.$$

– *Lorsque* $\phi = 0$ *ou* π, *alors* $W_t(\alpha)$ *est juste la moyenne de la distance temporelle entre* α *et la trajectoire passée* $(X_s, 0 \leq s \leq t)$. *Si on suppose de plus que* $c < 0$, *alors nous avons un processus auto-attractif (si* $c > 0$, *il est alors auto-répulsif). On peut montrer plus. Si* $c\cos\phi < -1/2$, *l'attraction est suffisamment forte pour inhiber, ou tout du moins confiner, les effets du mouvement brownien et* μ_t *converge presque sûrement vers une loi gaussienne. En revanche, si* $c\cos\phi \geq -1/2$, *alors le comportement de* μ_t *est celui d'un mouvement brownien (l'attraction ne suffit plus à contrer les effets de ce dernier). Dans ce cas, on montre en fait que la mesure de Lebesgue sur* S^1 *est l'unique ensemble invariant pour le flot* Φ.

– *Supposons ici que* $c\cos\phi < -1/2$ *et* $\phi \neq 0$ *et* π. *La fonction* W *n'est pas symétrique et le comportement asymptotique de* μ_t *est plus riche. L'ensemble limite* L *est composé de la mesure de Lebesgue (qui est un point d'équilibre instable) et d'une orbite périodique* $\{\nu_\theta; \theta \in S^1\}$. *Il y a alors suffisamment d'attraction pour que* μ_t *soit périodique et que* L *soit le cercle composé des mesures* $\{\nu_\theta; \theta \in S^1\}$.

Nous pouvons également généraliser l'exemple précédent à des polynômes trigonométriques quelconques :

EXEMPLE 4.15. *Considérons* $W(x) = 2 \sum\limits_{1\leq k\leq n} (a_k \cos(kx) + b_k \sin(kx))$. *Une étude trigonométrique assez poussée permet de prouver ce qui suit.*

(1) *s'il existe* $1 \leq k \leq n$ *tel que* $a_k < -1/2$, *alors presque sûrement* μ_t *ne converge pas vers la mesure de Lebesgue,*

(2) *si pour tout* $1 \leq k \leq n$, *on a* $a_k > -1/2$, *alors* μ_t *converge vers la mesure de Lebesgue avec une probabilité positive,*

(3) *si l'une des deux conditions suivantes est vérifiée*
  a) pour tout $1 \leq k \leq n$, *on a* $b_k = 0$ *ou* $a_k \geq 0$ *;*
  b) pour tout $1 \leq k \leq n$, *on a* $b_k = 0$, $a_k \leq 0$ *et* $\sum\limits_k a_k > -1/2$.

  Alors μ_t *converge presque sûrement vers la mesure de Lebesgue.*

Bibliographie

[1] ARTHUR B., ERMOLIEV Y. & KANIOVSKII Y. [1987], Path dependent processes and the emergence of macro-structure, *Eur. J. Oper. Res.* **30**, 294-303.

[2] BAKRY D. [1994], L'hypercontractivité et son utilisation en théorie des semigroupes, *Lectures on Prob. Th. and Stat., Ecole de Prob. de St-Flour*, Springer, 1-114.

[3] BAKRY D. & EMERY M. [1985], Diffusions hypercontractives, *Sém. Prob. XIX*, Springer LNM 1123, 177-206.

[4] BENAÏM M. [1993], Sur la nature des ensembles limites des trajectoires des algorithmes d'approximation stochastique de type Robbins-Monroe, *C.R. Acad. Sc. Paris Sér. I Math.* **317**, 195-200.

[5] BENAÏM M. [1999], Dynamics of stochastic approximation algorithms, *Sém. Prob. XXXIII, Lecture Notes in Math.* **1709**, 1-68, Springer.

[6] BENAÏM M. & HIRSCH M.W. [1996], Asymptotic pseudotrajectories and chain reccurent flows, *J. Dynam. Diff. Equ.* **8**, 141-176.

[7] BENAÏM M., LEDOUX M. & RAIMOND O. [2002], Self-interacting diffusions, *Prob. Theory Relat. Fields* **122**, 1-41.

[8] BENAÏM M. & RAIMOND O. [2005], Self-interacting diffusions III : symmetric interactions, *Ann. Prob.* **33**(5), 1716-1759.

[9] BISMUT J.M. [1981], A generalized formula of Itô and some other properties of solution flows, *Z. Wahrsch. Verw. Gebiete* **55**, 331-350.

[10] BOLLEY F., GUILLIN A. & VILLANI C. [2006], Quantitative concentration inequalities for empirical measures on non compact spaces, *Prob. Th. Rel. Fields*, to appear.

[11] BORODIN A.N. & SALMINEN P. [2002], Handbook of Brownian motion - Facts and formulae, 2nd edition, *Prob. and Its Appl., Birkhauser.*

[12] CARMONA R., PETIT F. & YOR M. [1994], Some extensions of the arcsine law as (partial) consequences of the scaling property of Brownian motion, *Prob. Th. Rel. Fields* **100**, 1-29.

[13] CARMONA R., PETIT F. & YOR M. [1998], Beta variables as time spent in $[0, \infty]$ by certain perturbed reflecting Brownian motion, *J. London Math. Soc.* **58**, 239-256.

[14] CARRILLO J.A., MCCANN R.J. & VILLANI C. [2003], Kinetic equilibration rates for granular media and related equations : entropy dissipation and mass transportation estimates, *Rev. math. Iberoam.* **19**, 3, 971-1018.

[15] CATTIAUX P., GUILLIN A. & MALRIEU F. [2007], Probabilistic approach for granular media equations in the non uniformly convex case , *Prob. Th. Rel. Fields*, to appear.

[16] CHAUMONT L. & DONEY R.A. [1999], Pathwise uniqueness for perturbed versions of Brownian motion and reflected Brownian motion, *Prob. Th. Rel. Fields* **113**, 519-534.

[17] CHAUMONT L. & DONEY R.A. [2000], Some calculations for doubly perturbed Brownian motion, *Stoch. Proc. Appl.* **85**, 61-74.

[18] CHAUMONT L., DONEY R.A. & HU Y. [2000], Upper and lower limits of doubly perturbed Brownian motion, *Ann. Inst. H. Poincaré, Prob. Stat.* **36**(2), 219-249.

[19] Conley C.C. [1978], Isolated invariant sets and the Morse index, *CBMS Regional conference series in mathematics* **38**, Am. Math. Soc., Providence.

[20] Cranston M. & Le Jan Y. [1995], Self-attracting diffusions : two cases studies, *Math. Ann.* **303**, 87-93.

[21] Cranston M. & Mountford T.S. [1996], The strong law of large numbers for a Brownian polymer, *Ann. Prob.* **2**(3), 1300-1323.

[22] Davies E.B & Simon B. [1984], Ultracontractivity and the heat kernel for Schrödinger operators and Dirichlet Laplacians, *J. Func. An.* **59**, 335-395.

[23] Davis B. [1996], Weak limits of perturbed random walks and the equation $Y_t = B_t + \alpha \sup\{Y_s; s \leq t\} + \beta \inf\{Y_s; s \leq t\}$, *Ann. Prob.* **24**, 2007-2023.

[24] Davis B. [1996], Brownian motion and random walk perturbed at extrema, *Probab. Th. Relat. Fields* **113**, 501-518.

[25] Doney R.A. [1998], Some calculations for perturbed Brownian motion, *Sém. Proba. XXXII*, Springer LNM 1686, 231-236.

[26] Duflo M. [1996], *Algorithmes stochastiques*, Springer, Berlin.

[27] Durrett R.T. & Rogers L.C.G. [1992], Asymptotic behavior of Brownian polymers, *Prob. Th. Rel. Fields* **92**(3), 337-349.

[28] Gallot S., Hulin D. & Lafontaine J. [2004], *Riemannian Geometry*, 3rd edition, Springer "Universitext".

[29] Herrmann S. & Roynette B. [2003], Boundedness and convergence of some self-attracting diffusions, *Math. Ann.* **325**(1), 81-96.

[30] Herrmann S. & Scheutzow M. [2004], Rate of convergence of some self-attracting diffusions, *Stoch. Proc. Appl.* **111**, 41-55.

[31] Hirsch M.W. & Smale S. [1974], *Differential Equations, Dynamical Systems, and Linear Algebra*, Academic Press.

[32] Holley R. & Stroock D. [1987], Logarithmic Sobolev Inequalities and Stochastic Ising Models, *J. Stat. Phys.* **46**, 1159-1194.

[33] Kallenberg 0. [2001], *Foundations of Modern Probability*, 2nd edition, Probability and its Applications, Springer.

[34] Kunita H. [1990], *Stochastic flows and stochastic differential equations*, Cambridge studies in advanced mathematics.

[35] Ledoux M. [2000], The geometry of Markov diffusion generators, *Ann. Fac. Sci. Toulouse* **IX**, 305-366.

[36] Malrieu F. [2003]Convergence to equilibrium for granular media equations and their Euler schemes, *Ann. Appl. Probab.* **13**(2), 540-560.

[37] McCann R. [1997], A convexity principle for interacting gases, *Adv. Math.* **128**, 153-179.

[38] Mountford T.S. & Tarrès P. [2007], An asymptotic result for Brownian polymers, *Ann. Inst. H. Poincaré, Prob. Stat.*, to appear.

[39] Norris J.R., Rogers L.C.G. & Williams D. [1987], Self-avoiding random-walk : a Brownian motion model with local time drift, *Prob. Th. Rel. Fields* **74**, 2, 271-287.

[40] Norton D.E. [1995], The fundamental theorem of dynamical systems, *Comment. Math. Univ. Carolinae* **36**, 3, 585-597.

[41] Pemantle R. [2007], A survey of random processes with reinforcement, *Prob. Surveys* **4**, 1-79.

[42] Perman M. & Werner W. [1997], Perturbed Brownian motion, *Prob. Theory Relat. Fields* **108**, 357-383.

[43] Raimond O. [1997], Self-attracting diffusions : case of constant interaction, *Prob. Theory Relat. Fields* **107**, 177-196.

[44] Revuz D. & Yor M. [1998], *Continuous Martingales and Brownian Motion*, 3rd edition, Springer.

[45] Rogers L.C.G. & Williams D. [2000], *Diffusions, Markov processes and Martingales*, 2nd edition, vol. 2 "Itô Calculus", Cambridge Univ. Press.

[46] Rothaus O. [1981], Diffusions on compact Riemannian manifolds and logarithmic Sobolev inequalities, *J. Func. An.* **42**, 102-109.

[47] Tóth B. [1999], Self-interacting random motions - a survey, in P.Révész and B. Tóth editors, *Random Walks, Bolyai Soc. Math. St.* **9**, 349-384.

[48] Tóth B. & Werner W. [1998], The true self-repelling motion, *Prob. Theory Rel. Fields* **111**, 375-452.

[49] Villani C. [2003], *Topics in Optimal Transportation*, Graduate Studies in Mathematics, Vol. 58, *AMS*.

[50] Yor M. [1992], *Some aspects of Brownian motion. Part I : Some special functionals*, Lecture Notes, ETH Zürich, Birkhäuser, Basel.

CHAPITRE 4

Diffusions renforcées sur un espace non compact

Le titre original de l'article (écrit en anglais) dont est composé ce chapitre est "The ODE method for some self-interacting diffusions on non-compact spaces".

1. Introduction

This paper addresses the long-term behavior of a class of 'self-interacting diffusion' processes $(X_t, t \geq 0)$ on non-compact spaces. These processes are time-continuous, non-Markov and live on $\mathbb{R}^d$ (or more generally on a smooth d-dimensional, complete, connected Riemannian manifold without boundary and with a Ricci curvature bounded from below). They are solutions to a kind of diffusion SDEs, whose drift term depends on the whole past of the path through the occupation measure of the process. Despite their lack of the Markov property, they often exhibit an interesting ergodic behavior.

1.1. Previous results on self-interacting diffusions. Time-continuous self-interacting processes, also named 'reinforced processes', have already been studied in many contexts. Under the name of 'Brownian polymers', Durrett & Rogers [**14**] first introduced them as a possible correct mathematical model for the evolution of a growing polymer. They are solutions of SDEs of the form

$$\mathrm{d}X_t = \mathrm{d}B_t + \mathrm{d}t \int_0^t \mathrm{d}s f(X_t - X_s)$$

where $(B_t; t \geq 0)$ is a standard Brownian motion and f a given function. As the process $(X_t; t \geq 0)$ evolves in an environment changing with its past trajectory, this SDE defines a self-interacting diffusion, which can be either self-repelling or self-attracting, depending on the function f. In any dimension, Durrett & Rogers obtained that $|X_t|/t$ is bounded (by a deterministic variable) whenever f has a compact support. They also proved in the one-dimensional case, that $|X_t|/t$ converges a.s. to a non random limit when f is non negative and $f(0) > 0$.

Afterwards, Cranston & Mountford [**12**] proved that X_t/t converges a.s. if f has a compact support, is nonnegative in a neighborhood of 0 and Lipschitz continuous. Cranston & Le Jan [**11**] studied the self-attracting case, either for a linear interaction or for a constant interaction in dimension 1. It happens that a.s. the sample paths of the solutions converge or at least are compact. Raimond [**29**] generalized this second study in dimension greater than 2. Later, Herrmann & Roynette [**16**] obtained the same kind of results (in dimension 1) for an odd, bounded and decreasing f, with a condition on f in the neighborhood of 0. Recently, Mountford & Tarrès [**25**] were interested in the self-repelling case and solved another conjecture of Durrett & Rogers. They proved that in this case there exist positive constants α, c such that $X_t/t^\alpha \to c$ with probability 1/2.

Parallel to the previous study, Tóth & Werner [**35**] considered another self-repelling random motion, by taking $f(0) = 0$ and maintaining $\int_{-\infty}^0 f(x)\mathrm{d}x = c$. They constructed a continuous, locally self-repelling process. This process is a.s. continuous and recurrent, it has a regular occupation time density and the self-repellence of its trajectory is to be understood in the sense that the process is instantaneously pushed in the direction of the decrease of its local time. But it is neither a semimartingale, nor a solution of a SDE. For further references on the subject, we refer the reader to the survey of Tóth [**34**].

Other self-interacting diffusions, with dependence on the (convoled) *normalized* occupation measure $(\mu_t, t \geq 0)$ have been considered since the work of Benaïm, Ledoux & Raimond [**6**]. They introduced a process living in a compact smooth connected Riemannian manifold M without boundary:

$$\mathrm{d}X_t = \sum_{i=1}^{N} F_i(X_t) \circ \mathrm{d}B_t^i - \int_M \nabla_x W(X_t, y)\mu_t(\mathrm{d}y)\mathrm{d}t, \tag{1.1}$$

where W is a (smooth) interaction potential and $(B^1, \cdots, B^N)$ is a standard Brownian motion on $\mathbb{R}^N$. The symbol $\circ$ stands for the Stratonovich stochastic integration as usual and $(F_i)_{1\leq i\leq N}$ is the family of smooth vector fields on M that appears in the Hörmander 'sum of squares' decomposition of the Laplace-Beltrami operator of M:

$$\Delta = \sum_{i=1}^{N} F_i^2.$$

The normalized occupation measure of the process involved in the SDE is defined by

$$\mu_t := \frac{r}{r+t}\mu + \frac{1}{r+t}\int_0^t \delta_{X_s}\mathrm{d}s, \tag{1.2}$$

where μ is the initial probability measure and r is a positive weight. In the compact-space case, they showed that the asymptotic behavior of μ_t can be related to the analysis of some deterministic dynamical flow defined on the space of the Borel probability measures. Some convergence in law properties are given by Benaïm & Raimond [**7**]. They went further in this study in [**8**] and gave sufficient conditions for the a.s. convergence of the normalized occupation measure. It happens that with a symmetric interaction, μ_t converges a.s. to a local minimum of a nonlinear free energy functional (each local minimum having a positive probability to be chosen).

All these results are summarized in a recent survey of Pemantle [**28**]. Both the survey of Pemantle and the survey of Tóth include results concerning self-interacting random walks.

1.2. Statement of the problem. The present paper follows the same lead but tries to extend the results of Benaïm, Ledoux & Raimond [**6**] in the non-compact setting. We present all results in the Euclidean space $\mathbb{R}^d$ for the sake of simplicity, but, as will be explained in the last section, they can be extended to the case of a complete connected Riemannian manifold without boundary with no further difficulty than the use of notations and a bit of geometry.

Here we set the main definitions: let us consider a confinement potential $V : \mathbb{R}^d \to \mathbb{R}^+$ and an interaction potential $W : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^+$. For any Borel bounded measure μ, we consider the 'convoled' function

$$W * \mu : \mathbb{R}^d \to \mathbb{R},\ \ W * \mu(x) := \int_{\mathbb{R}^d} W(x,y)\mu(\mathrm{d}y).$$

Our main object of interest is the self-interacting diffusion solution to

$$\begin{cases} \mathrm{d}X_t = \mathrm{d}B_t - \left(\nabla V(X_t) + \nabla W * \mu_t(X_t)\right)\mathrm{d}t \\ \mathrm{d}\mu_t = (\delta_{X_t} - \mu_t)\frac{\mathrm{d}t}{r+t} \\ X_0 = x,\ \ \mu_0 = \mu \end{cases} \tag{1.3}$$

where (B_t) is a d-dimensional Brownian motion, and (μ_t) the normalized occupation measure of the process defined as before (1.2), with initial weight $r > 0$ and initial probability measure μ. That is the sequence $(\mu_t; t \geq 0)$ defined by

$$\mu_t := \frac{r\mu + \int_0^t \delta_{X_s}\mathrm{d}s}{r+t}. \tag{1.4}$$

Our goal is to study the long term behavior of the normalized occupation measure $(\mu_t, t \geq 0)$. We will in particular prove that it is closely related to the behavior of a deterministic flow and we will give some sufficient conditions on the interaction potential in order to have the pointwise ergodic theorem for the process $(X_t, t \geq 0)$.

The main differences between the present paper and the paper [**6**] are the following. Obviously, the non-compactness of the space arises a lot of technical problems.
First, we need to take into account the non-compactness of $\mathbb{R}^d$, that is the reason why we introduce the V-norm in Section 3 (also named "dual weighted norm"). Fortunately, we manage to show that μ_t is a tight family of measures.

Second, the dynamical system involved in the study induces only a local flow and not a global one.
Last, if we suppose that the occupation measure μ_t appearing in the drift term is fixed to a measure μ, then we obtain the Feller diffusion X^μ. Let us note by A^μ the infinitesimal generator corresponding to this diffusion and Q^μ its fundamental kernel, that is $A^\mu Q^\mu = \Pi(\mu) - Id$, where $\Pi(\mu)$ is the invariant probability measure of X^μ. An essential point of our study consists in finding an upper bound for the operator Q_μ, and it is much more difficult in our case. Actually, one has to use the notion of (uniform) ultracontractivity, which means that the family of Markov semi-groups $(P_t^\lambda, t, \lambda)$ is uniformly bounded from $L^2(\mu_\lambda)$ to $L^\infty(\mu_\lambda)$.

1.3. Outline of contents. The organization of this paper is as follows. In the next section, we motivate the study of the self-interacting diffusions by a simple example and present some of our results. In Section 3, we describe our exact framework and introduce again some notations. Section 4 is devoted to the presentation of the main results. After that, we recall all the necessary definitions about dynamical systems and analyze the deterministic flow associated to a self-interacting diffusion in Section 5. In Section 6, we study in details a certain family of Markov semi-groups for which we will prove the existence of a uniform spectral gap and ultracontractivity. The proofs of the main results are given in Section 7, which heavily relies on the spectral analysis of the preceding section. It deals with the behavior of the normalized occupation measure $(\mu_t, t \geq 0)$. Finally, we explain how to generalize our study to a non-compact complete connected smooth Riemannian manifold in Section 8.

2. Motivation

We study here a simple example of self-interacting process in one dimension to understand what is going on. We consider here the following potentials:

$$V(x) := \frac{1}{4}x^4 + \frac{1}{5}x^2 + 1 \text{ and } W(x,y) := \cos(x)\varphi_1(y) + \sin(x)\varphi_2(y)$$

where the functions φ_i are smooth and controlled by V (that is there exists a positive constant C such that $|\varphi_i(y)| \leq CV(y)$). We consider again the following SDE:

$$\begin{cases} \mathrm{d}X_t = -\left(V'(X_t) + W' * \mu_t(X_t)\right)\mathrm{d}t + \mathrm{d}B_t \\ \mathrm{d}\mu_t = (\delta_{X_t} - \mu_t)\frac{\mathrm{d}t}{r+t} \end{cases}$$

Since our interaction potential has separated variables, this infinite-dimensional SDE can be reduced to a 3-dimensional SDE with the new variables $Y_t :=$

$\int_{\mathbb{R}} \varphi_1(y)\mu_t(\mathrm{d}y)$ and $Z_t := \int_{\mathbb{R}} \varphi_2(y)\mu_t(\mathrm{d}y)$:

$$\begin{cases} \mathrm{d}X_t = \mathrm{d}B_t - [V'(X_t) - Y_t \sin(X_t) + Z_t \cos(X_t)]\, \mathrm{d}t, \;\; X_0 = x \\ \mathrm{d}Y_t = [\varphi_1(X_t) - Y_t]\,(r+t)^{-1}\mathrm{d}t, \;\; Y_0 = y \\ \mathrm{d}Z_t = [\varphi_2(X_t) - Z_t]\,(r+t)^{-1}\mathrm{d}t, \;\; Z_0 = z \end{cases}$$

We expect the new variables (Y_t, Z_t) to converge a.s. to some (say deterministic to fix the ideas but this is actually too demanding!) limits (α, β); therefore the process $(X_t, t \geq 0)$ should behave asymptotically as a classical real diffusion with ergodic measure:

$$\mu_{\alpha,\beta}(\mathrm{d}x) := Z_0(\alpha, \beta)^{-1} \exp\{-2V(x) - 2\alpha\cos(x) - 2\beta\sin(x)\}\mathrm{d}x$$

where $Z_0(\alpha, \beta) := \int \exp\{-2V(x) - 2\alpha\cos(x) - 2\beta\sin(x)\}\mathrm{d}x$ is the normalization constant.

The pointwise ergodic theorem now implies that (Y_t, Z_t) should a.s. converge to $(\int_{\mathbb{R}} \varphi_1(y)\mu_{\alpha,\beta}(\mathrm{d}y), \int_{\mathbb{R}} \varphi_2(y)\mu_{\alpha,\beta}(\mathrm{d}y))$. This shows that the constants (α, β) should be the fixed points of the map

$$(\alpha, \beta) \mapsto \frac{1}{Z_0(\alpha, \beta)} \begin{pmatrix} Z_1(\alpha, \beta) \\ Z_2(\alpha, \beta) \end{pmatrix}$$

with the notations $Z_i(\alpha, \beta) := \int_{\mathbb{R}} \varphi_i(y) e^{\{-2V(y) - 2\alpha\cos(y) - 2\beta\sin(y)\}}\mathrm{d}y$. To study these fixed points, we are lead to consider the following ODE:

$$\begin{cases} \dot{\alpha}_t = -\alpha_t + Z_1(\alpha_t, \beta_t)/Z_0(\alpha_t, \beta_t) \\ \dot{\beta}_t = -\beta_t + Z_2(\alpha_t, \beta_t)/Z_0(\alpha_t, \beta_t) \end{cases} \tag{2.1}$$

We represent below the orbits of the ODE (3.3). In the numerical applications we will choose $\varphi_1(y) := 1.3y^3$, and $\varphi_2(y) := -1.3y$. In our numerical example, we observe that we have three fixed points: $(0, 0)$ and $(8.492..., -2.424...)$ are sinks and $(0.538..., -0.527...)$ is a saddle (see figure 2).

Now if we turn back to our self-interacting motion $(X_t, t \geq 0)$ and leave aside the (maybe false!) assumption that (Y_t, Z_t) will necessarily converge to a deterministic limit, the ODE (3.3) turns nonetheless to be very useful to study the asymptotic behavior of (Y_t, Z_t). Handling the technics that are developed in this paper (and which follows the ideas first introduced in [**6**]) the reader is able to prove that (Y_{e^t}, Z_{e^t}) is an *asymptotic pseudo-trajectory* of the non-linear dynamical system (3.3) that is

$$\lim_{t \to +\infty} \sup_{0 \leq h \leq T} \left(|Y(e^{t+h}) - \alpha_h(Y(e^t))| + |Z(e^{t+h}) - \beta_h(Z(e^t))| \right) = 0$$

for all $T > 0$. For an illustration, see the figure 2.

As (Y_{e^t}, Z_{e^t}) is an asymptotic pseudo-trajectory for the non-linear dynamical system (3.3), the theory of the pseudo-trajectories implies that the limit points of the joined-process (Y, Z) are almost surely included in the critical points of the map corresponding to the ODE, that is the set

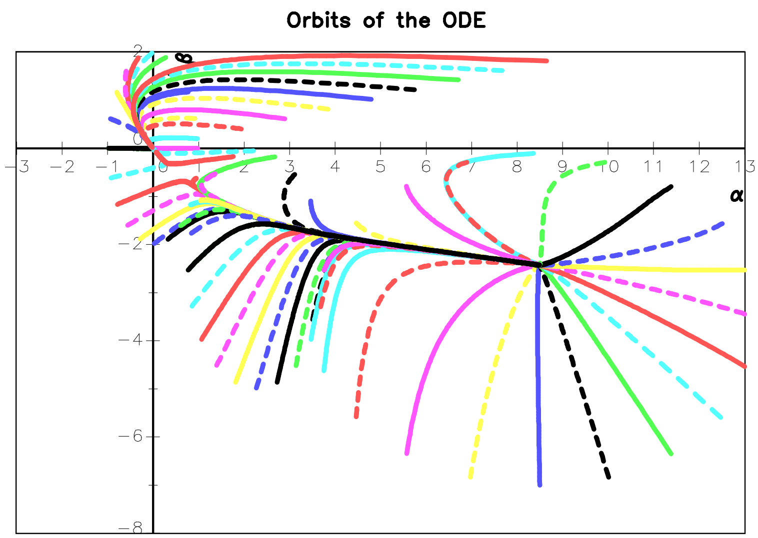

FIGURE 1. Orbits of the dynamical system (3.3).

$Fix := \left\{(\alpha,\beta); \alpha = \frac{Z_1(\alpha,\beta)}{Z_0(\alpha,\beta)}, \beta = \frac{Z_2(\alpha,\beta)}{Z_0(\alpha,\beta)}\right\}$. In addition, the preceding set contains only isolated points and therefore, the process (Y, Z) converges a.s. to one of these critical points. Obviously, if we have an unstable equilibrium (α_0, β_0), then $\mathbb{P}_{x,y,z,r}((Y_t, Z_t) \to (\alpha_0, \beta_0)) = 0$. Moreover, following the lines of Benaïm & Raimond [**8**], the reader also manages to prove that for a *saddle* point (α_1, β_1), we get the same result: $\mathbb{P}_{x,y,z,r}((Y_t, Z_t) \to (\alpha_1, \beta_1)) = 0$. Actually, the process (Y, Z) converges a.s. to one of the stable fixed points (sinks) of the ODE (3.3), that is for all *sink* $(\alpha_i, \beta_i) \in Fix$, we get

$$\mathbb{P}_{x,y,z,r}((Y_t, Z_t) \to (\alpha_i, \beta_i)) > 0.$$

Of course, if we have only one sink, then the preceding probability equals 1. In addition, the "choice" of the sink in the convergence will depend crucially on the initial parameters x, r, μ. Indeed, suppose that we have at least two sinks $(\alpha_i, \beta_i), i = 1, 2$ (as in the numerical example). Then we know that for $i = 1, 2$, we get $\mathbb{P}_{x,y,z,r}((Y_t, Z_t) \to (\alpha_i, \beta_i)) > 0$. Therefore the asymptotic σ-algebra of the quadruple-process $(X_t, Y_t, Z_t, r + t)$ is non trivial and the process is not ergodic. Nevertheless, the quadruple-process is a true diffusion

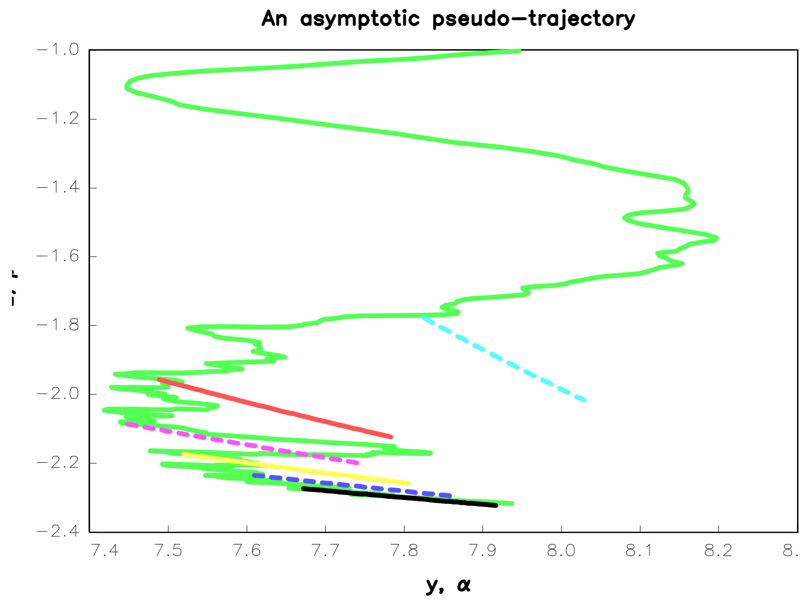

FIGURE 2. A path of (Y_{e^t}, Z_{e^t}) in green and some orbits of the flow.

and Markov homogeneous. Let

$$p_{\alpha,\beta}(x,y,z,r) := \mathbb{P}_{x,y,z,r}((Y_t, Z_t) \to (\alpha, \beta)).$$

We emphasize that the function $p_{\alpha,\beta}$ is (bounded) invariant for the process $(X_t, Y_t, Z_t, r+t)$. But the preceding Markov process is not ergodic. Thus, the bounded invariant functions are not constant. As a consequence, the probability $p_{\alpha,\beta}(x,y,z,r)$ depends on the initial value (x,y,z,r).

3. Preliminaries and Tools

3.1. Technical assumptions on the potentials. Let $(\cdot,\cdot)$ stand for the Euclidian scalar product.

In the sequel, the technical assumptions on the potentials V and W are the following:

i) (*regularity and positivity*) $V \in \mathcal{C}^2(\mathbb{R}^d)$ and $W \in \mathcal{C}^2(\mathbb{R}^d \times \mathbb{R}^d)$ and $V \geq 1,\ W \geq 0$;

ii) (*convexity*) V is a strictly uniformly convex function, *i.e.* there exists $K > 0$ such that for all $x, \xi \in \mathbb{R}^d$: $(\nabla^2 V(x)\xi, \xi) \geq K|\xi|^2$;

iii) (*growth*) there exist $R, c > 0$, $\delta > 1$ such that for all $|x| \geq R$, $(\nabla V(x), x) \geq c|x|^{2\delta}$ and there exists $C > 0$ such that for all $x, y \in \mathbb{R}^d$ we have

$$|\nabla V(x) - \nabla V(y)| \leq C(|x-y| \wedge 1)(V(x) + V(y)); \tag{3.1}$$

iv) (*domination*) there exists $\kappa > 0$ such that for all $x, y \in \mathbb{R}^d$,

$$W(x,y) + |\nabla_x W(x,y)| + |\nabla^2_{xx} W(x,y)| \leq \kappa \left(V(x) + V(y)\right); \tag{3.2}$$

v) (*curvature*) we suppose that we can decompose $W(x,y) = W_1(x,y) + W_2(x,y)$, where W_2 and its two first derivatives with respect to x are three bounded functions in the variable x and there exists $M > 0$ such that for all $x, y, \xi \in \mathbb{R}^d$ we have

$$\left((\nabla^2 V(x) + \nabla^2_{xx} W_1(x,y))\xi, \xi\right) \geq M|\xi|^2. \tag{3.3}$$

Remark 3.1. *1) The most important conditions are the uniform convexity of V and the fact that W is controlled by V.*
2) The growth condition (3.1) on V ensures that there exist $a, b > 0$ such that for all $x \in \mathbb{R}^d$, we have

$$\Delta V(x) \leq a + bV(x). \tag{3.4}$$

3) The positivity and domination conditions (3.2) on the interaction potential are not so hard to be satisfied, since the self-interacting process will be invariant by the gauge transform $W(x,y) \mapsto W(x,y) + \phi(y)$ for any function ϕ that does not grow faster than V. In the example discussed in the last section, we see that choosing $\phi(y) = 1.3(y^4 + 25/16y^2 + 65/64)$ enables us to meet the required conditions.
4) The curvature condition (3.3) means that V can offset a lack of convexity of W_1.

3.2. Some useful measure spaces. As usual, we denote by $\mathcal{M}(\mathbb{R}^d)$ the space of signed (bounded) Borel measures on $\mathbb{R}^d$ and by $\mathcal{P}(\mathbb{R}^d)$ its subspace of probability measures. We will need the following measure space:

$$\mathcal{M}(\mathbb{R}^d; V) := \{\mu \in \mathcal{M}(\mathbb{R}^d); \int_{\mathbb{R}^d} V(y)|\mu|(\mathrm{d}y) < \infty\}, \tag{3.5}$$

where $|\mu|$ is the variation of μ (that is $|\mu| := \mu^+ + \mu^-$ with (μ^+, μ^-) the Hahn-Jordan decomposition of μ). This space will enable us to always check the integrability of V (and therefore of W and its derivatives thanks to the domination condition (3.2)) with respect to the (random) measures to be considered. We endow this space with the following dual weighted supremum norm (or dual V-norm) defined by

$$||\mu||_V := \sup_{\varphi; |\varphi| \leq V} \left| \int \varphi \mathrm{d}\mu \right|, \quad \mu \in \mathcal{M}(\mathbb{R}^d; V). \tag{3.6}$$

This norm naturally arises in the approach to ergodic results for time-continuous Markov processes of Meyn & Tweedie [**24**]. It also makes $\mathcal{M}(\mathbb{R}^d;V)$ a Banach space.

Next, we consider $\mathcal{P}(\mathbb{R}^d;V) := \mathcal{M}(\mathbb{R}^d;V) \cap \mathcal{P}(\mathbb{R}^d)$. Both are not empty since they contain the measure (possibly normalized)

$$\gamma(\mathrm{d}x) := \exp\left(-2V(x)\right)\mathrm{d}x. \tag{3.7}$$

Finally for any $\beta > 0$, we introduce the subspaces

$$\mathcal{M}_\beta(\mathbb{R}^d;V) := \{\mu \in \mathcal{M}(\mathbb{R}^d;V); \int_{\mathbb{R}^d} V(y)|\mu|(\mathrm{d}y) \leq \beta\},$$

and likewise $\mathcal{P}_\beta(\mathbb{R}^d;V) := \mathcal{M}_\beta(\mathbb{R}^d;V) \cap \mathcal{P}(\mathbb{R}^d)$.

3.3. The family of semi-groups (P_t^μ). In all the following, $(\Omega, \mathcal{F}, (\mathcal{F}_t)_t, \mathbb{P})$ will be a filtered probability space satisfying the usual conditions. For any bounded Borel measure $\mu \in \mathcal{M}(\mathbb{R}^d;V)$, let $(X_t^\mu, t \geq 0)$ be the Feller diffusion defined by the following SDE

$$\begin{cases} \mathrm{d}X_t^\mu = \mathrm{d}B_t - (\nabla V(X_t^\mu) + \nabla W * \mu(X_t^\mu))\,\mathrm{d}t, \\ X_0^\mu = x. \end{cases} \tag{3.8}$$

Let also $\mathcal{C}_0^\infty(\mathbb{R}^d) := \{f : \mathbb{R}^d \to \mathbb{R}; f \in \mathcal{C}^\infty(\mathbb{R}^d), \lim_{|x|\to\infty} f(x) = 0\}$. We consider the differential operator A_μ defined on $\mathcal{C}_0^\infty(\mathbb{R}^d)$ by

$$\begin{aligned} A_\mu f &:= \frac{1}{2}\Delta f - (\nabla W * \mu, \nabla f) - (\nabla V, \nabla f), \\ A_\mu f &= \frac{1}{2}e^{2(V+W*\mu)}\mathrm{div}(e^{-2(V+W*\mu)}\nabla f). \end{aligned} \tag{3.9}$$

A_μ corresponds to the infinitesimal generator of the true diffusion $(X_t^\mu; t \geq 0)$ (3.8). We also denote by $(P_t^\mu; t \geq 0)$ the Markov semi-group associated to A_μ.

We emphasize that (X_t^μ) is a positive-recurrent (reversible) diffusion. We denote by $\Pi(\mu)$ its unique invariant probability measure:

$$\Pi(\mu)(\mathrm{d}x) := \frac{e^{-2W*\mu(x)}}{Z(\mu)}\gamma(\mathrm{d}x) \tag{3.10}$$

where $Z(\mu) := \int_{\mathbb{R}^d} e^{-2W*\mu(x)}\gamma(\mathrm{d}x) < +\infty$ is just the normalization constant.

3.4. The infinite-dimensional ODE. We introduce a dynamical system on the set of signed measures $\mathcal{M}(\mathbb{R}^d;V)$. We will assume in the next section the existence of the flow $\Phi : \mathbb{R} \times \mathcal{M}(\mathbb{R}^d;V) \to \mathcal{M}(\mathbb{R}^d;V)$ defined by

$$\Phi_0(\mu) = \mu, \ \ \frac{\mathrm{d}}{\mathrm{d}t}\Phi_t(\mu) = \Pi(\Phi_t(\mu)) - \Phi_t(\mu). \tag{3.11}$$

REMARK 3.2. *(important!) For W symmetric or bounded, we will prove the existence of the flow.*

REMARK 3.3. *It is readily seen that the two-dimensional differential system introduced for the example is semi-conjugate to this infinite-dimensional one (consider the surjection $\mu \mapsto (\int \varphi_1 \mathrm{d}\mu, \int \varphi_2 \mathrm{d}\mu)$). This situation is general when the interaction potential W has separated variables, a particular case we will not consider any more.*

In order to study the flow Φ, we will need to endow the space $\mathcal{P}(\mathbb{R}^d; V)$ with different topologies. When nothing else is stated, we will consider that this space of measures is endowed with the strong topology defined by the dual weighted supremum norm $||\cdot||_V$. But, as the reader will notice, we will frequently need to switch from the strong topology to the weak* topology of convergence of measures. We adopt here a non-standard definition compatible with possibly unbounded functions (controlled by V). We introduce the weighted supremum norm (or V-norm)

$$||f||_V := \sup_{x\in\mathbb{R}^d} \frac{|f(x)|}{V(x)}, \tag{3.12}$$

and the space of continuous V-bounded functions

$$\mathcal{C}^0(\mathbb{R}^d; V) := \{f \in \mathcal{C}^0(\mathbb{R}^d) : ||f||_V < \infty\} \tag{3.13}$$

Similarly let $\mathcal{C}^\infty(\mathbb{R}^d; V) := \mathcal{C}^\infty(\mathbb{R}^d) \cap \mathcal{C}(\mathbb{R}^d; V)$. Now for any sequence of probability measures $(\mu_n, n \geq 1)$ and any probability measure μ (all belonging to $\mathcal{P}(\mathbb{R}^d; V)$), we define the weak* convergence in the following way:

$$\mu_n \xrightarrow{w} \mu \text{ if and only if } \int_{\mathbb{R}^d} \varphi \mathrm{d}\mu_n \underset{n\to\infty}{\longrightarrow} \int_{\mathbb{R}^d} \varphi \mathrm{d}\mu,\ \forall \varphi \in \mathcal{C}^0(\mathbb{R}^d; V).$$

We point out that our definition of the weak* convergence always implies the common definition.

3.5. The self-interacting diffusion. We remind the self-interacting diffusion considered here:

$$\begin{cases} \mathrm{d}X_t = \mathrm{d}B_t - (\nabla V(X_t) + \nabla W * \mu_t(X_t))\,\mathrm{d}t \\ \mathrm{d}\mu_t = (\delta_{X_t} - \mu_t)\frac{\mathrm{d}t}{r+t} \\ X_0 = x,\ \ \mu_0 = \mu \end{cases}$$

PROPOSITION 3.4. *For any $x \in \mathbb{R}^d$, $\mu \in \mathcal{P}(\mathbb{R}^d; V)$ and $r > 0$, there exists a unique global strong solution $(X_t, \mu_t, t \geq 0)$.*

PROOF. By a standard theorem (see [**32**], chap.V theorem 11.2), we know that given r, μ, x, ∇V and $\nabla_x W$ locally Lipschitz, the equation (2.5) has a pathwise unique local strong solution, which law is denoted by $\mathbb{P}_{x,r,\mu}$. So, we need only to prove that the local solution does not explode.

First, we point out that for all $t > 0$ such that $(X_s, s \leq t)$ is defined, $\mu_t \in \mathcal{P}(\mathbb{R}^d; V)$. Indeed, we have for all initial $x \in \mathbb{R}^d$

$$\int V(y)\mu_t(\mathrm{d}y) = \frac{r}{r+t}\int V(y)\mu(\mathrm{d}y) + \frac{1}{r+t}\int_0^t V(X_s)\mathrm{d}s < +\infty.$$

In order to show that the solution never explodes, we introduce the Lyapunov functional $\mathcal{E}$ defined on $\mathbb{R}^d \times \mathcal{P}(\mathbb{R}^d;V)$ by

$$\mathcal{E}_\mu(x) := V(x) + W * \mu(x). \tag{3.14}$$

As the $\mathcal{C}^2$-valued process $(t,x) \mapsto \mathcal{E}_{\mu_t}(x)$ is of class $\mathcal{C}^2$ (in the space variable) and is a $\mathcal{C}^1$-semi-martingale (in the time variable), the generalized Itô formula, also named Itô-Ventzell formula (see [**21**]), applied to $(t,x) \mapsto \mathcal{E}_{\mu_t}(x)$ implies

$$\begin{aligned}\mathcal{E}_{\mu_t}(X_t) &= \mathcal{E}_\mu(x) + \int_0^t (\nabla\mathcal{E}_{\mu_s}(X_s), \mathrm{d}B_s) - \int_0^t |\nabla\mathcal{E}_{\mu_s}(X_s)|^2\,\mathrm{d}s \\ &+ \frac{1}{2}\int_0^t \Delta\mathcal{E}_{\mu_s}(X_s)\mathrm{d}s + \int_0^t (W(X_s,X_s) - W*\mu_s(X_s))\,\frac{\mathrm{d}s}{r+s}.\end{aligned}$$

Let us introduce the sequence of stopping times

$$\tau_n := \inf\{t \geq 0; \mathcal{E}_{\mu_t}(X_t) + \int_0^t |\nabla\mathcal{E}_{\mu_s}(X_s)|^2\,\mathrm{d}s > n\}.$$

We note that $\int_0^{t\wedge\tau_n}(\nabla\mathcal{E}_{\mu_s}(X_s), \mathrm{d}B_s)$ is a true martingale. Now the growth condition on V and the domination condition (3.2) on W imply: $\exists C > 0$ such that

$$\mathbb{E}\mathcal{E}_{\mu_{t\wedge\tau_n}}(X_{t\wedge\tau_n}) \leq \mathcal{E}_\mu(x) + C\mathbb{E}\int_0^{t\wedge\tau_n}[1 + \mathcal{E}_{\mu_s}(X_s)]\,\mathrm{d}s.$$

Now, applying the Gronwall lemma to $\mathbb{E}\mathcal{E}_{\mu_{t\wedge\tau_n}}(X_{t\wedge\tau_n})+1$, we get $\mathbb{E}\mathcal{E}_{\mu_{t\wedge\tau_n}}(X_{t\wedge\tau_n}) \leq (\mathcal{E}_\mu(x) + C + 1)e^{Ct}$.

The last step is to show that for any $t \geq 0$, the probability $\mathbb{P}_{x,r,\mu}(\forall m, \tau_m \leq t)$ vanishes. The Markov inequality joint with the preceding inequality implies the following, for any $n \in \mathbb{N}^*$

$$\begin{aligned}\mathbb{P}_{x,r,\mu}(\forall m, \tau_m \leq t) &\leq \mathbb{P}_{x,r,\mu}(\tau_n \leq t) = \mathbb{P}_{x,r,\mu}\left(\sup_{0\leq s\leq t}\mathcal{E}_{\mu_s}(X_s) + \int_0^s |\nabla\mathcal{E}_{\mu_u}(X_u)|^2\,\mathrm{d}u > n\right) \\ &\leq \mathbb{P}_{x,r,\mu}\left(\sup_{0\leq s\leq t}\int_0^s(\nabla\mathcal{E}_{\mu_u}(X_u), \mathrm{d}B_u) > \frac{n}{2}\right) \\ &+ \mathbb{P}_{x,r,\mu}\left(\mathcal{E}_\mu(x) + \frac{1}{2}\int_0^t \Delta\mathcal{E}_{\mu_u}(X_u)\mathrm{d}u + \int_0^t W(X_u,X_u)\frac{\mathrm{d}u}{r+u} > \frac{n}{2}\right) \\ &\leq \frac{2}{nC}\left(t + \int_0^t (\mathcal{E}_\mu(x) + C + 1)e^{Cu}\mathrm{d}u\right) = O\left(\frac{1}{n}\right).\end{aligned}$$

Thus by the nonnegativity of W, there exists some m such that $V(X_t) \leq \mathcal{E}_{\mu_t}(X_t) \leq m$ for all $t \geq 0$ and $\lim_{|x|\to\infty} V(x) = \infty$. Therefore, the process $(X_t, t \geq 0)$ does not explode in a finite time and the SDE (2.5) admits a global strong solution. □

4. Main results

4.1. General idea. We sketch here the general idea of the proof and explain why the tools introduced in the preliminary section arise quite naturally.

First consider that the occupation measure appearing in the drift is 'frozen' to some fixed measure μ. We obtain the Feller diffusion X_t^μ. For this diffusion, it is easy to prove the existence of a spectral gap and therefore get that the semi-group $(P_t^\mu; t \geq 0)$ is exponentially ergodic, that is

$$||P_t^\mu f - \Pi(\mu)f||_V \leq K(\mu)||f||_V e^{-c(\mu)t}, \quad f \in \mathcal{C}^0(\mathbb{R}^d; V). \tag{4.1}$$

To get, as by-product, the almost sure convergence of the empirical occupation measure of the process X_t^μ, a standard technique[1] is to consider the operator (sometimes called the 'fundamental kernel' as in Kontoyiannis & Meyn [**20**])

$$Q_\mu f := \int_0^\infty (P_t^\mu f - \Pi(\mu)f)\,\mathrm{d}t \tag{4.2}$$

for any $f \in \mathcal{C}^\infty(\mathbb{R}^d; V)$. Then it is enough to apply the Itô formula to $Q_\mu f(X_t^\mu)$ and divide both members by t to get the desired result. Indeed one has

$$Q_\mu f(X_t^\mu) = Q_\mu f(x) + \int_0^t (\nabla Q_\mu f(X_s^\mu), \mathrm{d}B_s) + \int_0^t A_\mu Q_\mu f(X_s^\mu)\mathrm{d}s.$$

Some easy bounds on the semi-group (P_t^μ) are enough to prove that almost all terms are negligible compared to t and it remains to recognize the third term since $A_\mu Q_\mu f = \Pi(\mu)f - f$.

Now when μ_t changes in time, we still can write a convenient extended form of the Itô formula (which let appear the time derivative of $Q_{\mu_t} f(x)$) but we need to improve the remainder of the argument.

First, we will prove that we can find $\beta > 0$ such that $\mu_t \in \mathcal{P}_\beta(\mathbb{R}^d; V)$ a.s. for all $t \geq 0$ and this last set is compact for the weak* convergence. Then, we will study the family of semi-groups $(P_t^\mu, t \geq 0)$ where $\mu \in \mathcal{P}_\beta(\mathbb{R}^d; V)$. Obviously when using the ergodic estimates (4.1) we would like to get bounds that are uniform in μ. Section 6 will be devoted to those uniform properties of the family of semi-groups $(P_t^\mu; t \geq 0)$.

Second, if we compute the time derivative of μ_{e^t}, we obtain that it equals $\frac{e^t}{r+e^t}(\delta_{X_{e^t}} - \mu_{e^t})$ and in particular, it is singular with respect to the Lebesgue measure. However, the distance between the time-derivative of μ_{e^t} and the term $\Pi(\mu_{e^t}) - \mu_{e^t}$ converges to zero a.s. As with stochastic approximation processes, one expects the trajectories of the process μ_t to approximate the trajectories of a deterministic flow, that is the flow of the dynamical system

[1] Of course, here X_t^μ is a Markov process and has an invariant probability and therefore one can just use the limit-quotient theorem, which implies the limit-ratio theorem.

induced by $\Pi(\mu_t) - \mu_t$ is an asymptotic pseudotrajectory of the empirical measure μ_t. It is this very last remark that conveyed to Benaïm & *al* [**6**] the idea of comparing the asymptotic evolution of $(\mu_t; t \geq 0)$ with the flow $(\Phi_t(\mu))$.

4.2. Tightness of (μ_t). In their paper Benaïm & *al* [**6**] crucially rely on the compacity of the manifold M where the self-interacting diffusion lives. The compacity of the state space M readily implies the compacity of the space of probability measures $\mathcal{P}(M)$ and therefore that the process $(\mu_t, t \geq 0)$ is tight. The tightness of the occupation measure allows to write some technical bounds that are needed to prove the pseudo-asymptotic property. Compacity is also a desired property when attractors are looked for in a dynamical system. Here due to the confinement potential V (and possibly also to the interaction potential W) it is not obvious that the process $(\mu_t, t \geq 0)$ remains in a (weakly) compact space of measures, but it is true!

PROPOSITION 4.1. *There exists $\beta > 0$ such that $\mu_t \in \mathcal{P}_\beta(\mathbb{R}^d; V)$ for all $t \geq 0$, where the space $\mathcal{P}_\beta(\mathbb{R}^d; V)$ is compact for the weak* topology of measures.*

The proof of this proposition is postponed to the Section 7.

COROLLARY 4.2. *i) The family $(\mu_t, t \geq 0)$ is a.s. tight.*
ii) The limit set of μ_t is compact for the weak topology.*

PROOF. *(of the corollary)* Straightforward. □

4.3. The "ODE method". The notion of asymptotic pseudo-trajectories was first introduced in Benaïm & Hirsch [**5**], and is particularly useful for analyzing the long-term behavior of stochastic processes, considered as approximations to solutions of ordinary differential equation (the "ODE method"). We refer to this article for more details and just give the essential and necessary definitions.

DEFINITION 4.3. *i) For every continuous function $\xi : \mathbb{R}_+ \to \mathcal{P}(\mathbb{R}^d; V)$, the $\omega-$limit set of ξ, denoted by $\omega(\xi)$, is the set of limits of weak* convergent sequences $\xi(t_k), t_k \uparrow \infty$, that is*

$$\omega(\xi) := \bigcap_{t \geq 0} \overline{\xi([t, \infty))}$$

where $\overline{\xi([t, \infty))}$ stands for the closure of $\xi([t, \infty))$ according to the weak topology.*

ii) A continuous function $\xi : \mathbb{R}_+ \to \mathcal{P}_\beta(\mathbb{R}^d; V)$ is an asymptotic pseudo-trajectory (or asymptotic pseudo-orbit) for the flow Φ (for the weak topology of measures) if for all $T > 0$, for all $f \in \mathcal{C}(\mathbb{R}^d; V)$*

$$\lim_{t \to +\infty} \sup_{0 \leq s \leq T} |\xi(t+s)f - \Phi_s(\xi(t))f| = 0. \tag{4.3}$$

The purpose here is to find a weak* asymptotic pseudo-trajectory for the flow Φ defined by (2.8). Actually, we will show here that the time-changed process $\mu_{h(t)}$ (and not μ_t) is an asymptotic pseudo-trajectory for Φ, where h is a deterministic time-change defined by

$$h(t) := r(e^t - 1) \ \ \forall t \geq 0. \tag{4.4}$$

The need for a time-change comes from the normalization of the occupation measure μ_t. The factor $(r+t)^{-1}$ disappears when we consider

$$\frac{\mathrm{d}}{\mathrm{d}t}\mu_{h(t)} = \delta_{X_{h(t)}} - \mu_{h(t)}.$$

THEOREM 4.4. *Under $\mathbb{P}_{x,r,\mu}$, the function $t \mapsto \mu_{h(t)}$ is almost surely an asymptotic pseudo-trajectory for Φ (for the weak* topology).*

The proof of this theorem is given in Section 7.

This theorem enables us to describe the limit set of (μ_t):

COROLLARY 4.5. *The limit set of $(\mu_t; t \geq 0)$ is invariant by the flow Φ, compact and attractor free. Moreover, it is a (weak*-) compact subset of $\left\{\int_{\mathcal{P}(\mathbb{R}^d;V)} \Pi(\mu)\rho(\mathrm{d}\mu); \rho \in \mathcal{P}(\mathcal{P}(\mathbb{R}^d;V))\right\}$.*

The definition of an attractor free set will be given later, in Section 5. We will prove this result in Section 7 (the compactness is a consequence of the corollary 4.2).

THEOREM 4.6. *Assume that W is symmetric. Then the $\omega-$limit set of $(\mu_t, t \geq 0)$ is $\mathbb{P}_{x,r,\mu}$-a.s. a compact connected subset –for the weak* topology– of the fixed point of Π, that is $\left\{\mu \in \mathcal{P}_\beta(\mathbb{R}^d;V); \mu = \Pi(\mu)\right\}$.*

The proof of this theorem is given in Section 7.

COROLLARY 4.7. *Suppose that W is symmetric. If Π contains only finitely many fixed points $(\mu_i, 1 \leq i \leq n)$, then $(\mu_t; t \geq 0)$ converges almost surely to one of these fixed points.*

PROOF. Straightforward. □

COROLLARY 4.8. *Suppose that W is symmetric and that the mapping Π has a unique fixed point μ_∞. Then*

$$\lim_{t\to\infty} \mu_t = \mu_\infty \ \ \mathbb{P}_{x,r,\mu} - a.s.$$

PROOF. It is a consequence of the theorem 4.6 and the corollary 4.7. □

4.4. Application: a sufficient condition for the global convergence of $(\mu_t, t \geq 0)$. We give here a sufficient condition for the convergence of the empirical occupation measure. The point is to find a criterion which ensures Π to have a unique fixed point.

For a symmetric W, we introduce the *free energy* (up to a multiplicative constant) corresponding to the ODE studied

$$\mathcal{F}(\mu) := \int_{\mathbb{R}^d} \log\left(\frac{\mathrm{d}\mu}{\mathrm{d}\gamma}\right) \mathrm{d}\mu + \int_{\mathbb{R}^d\times\mathbb{R}^d} W(x,y)\mu(\mathrm{d}x)\mu(\mathrm{d}y).$$

This functional is the sum of an internal energy (the entropy term), a potential energy $\mathcal{V}$ and an interacting energy term $\mathcal{W}$. The competition between $\mathcal{U}$, $\mathcal{V}$ and $\mathcal{W}$ can determine a unique minimizer for $\mathcal{F}$.

THEOREM 4.9. *Suppose that for all $x, y \in \mathbb{R}^d$, $u \in \mathbb{R}^d$, $v \in \mathbb{R}^d$, we have that there exists $K > 0$ such that $\nabla^2_{x,y}(V+W)((u,v),(u,v)) \geq K(|u|^2 + |v|^2)$. Then there exists a unique probability measure μ_∞ such that $\lim_{t\to\infty} \mu_t = \mu_\infty$ $\mathbb{P}_{x,r,\mu} - a.s.$.*

PROOF. Under our hypothesis, McCann has proved in his paper [**23**] that $\mathcal{F}$ is strictly displacement convex. – This means that if we consider ρ_0, ρ_1 two L^1 probability measures, the theory of mass transportation (see e.g. [**37**]) states that there exists a convex function φ such that $\nabla\varphi_\sharp\rho_0 = \rho_1$. We denote by ρ_s $(0 \leq s \leq 1)$ the measure defined by $\rho_s := ((1-s)Id + s\nabla\varphi)_\sharp\,\rho_0$. We say that the function $\mathcal{F}$ is displacement convex if for all ρ_0, ρ_1, the application $s \mapsto \mathcal{F}(\rho_s)$ is convex.– Hence $\mathcal{F}$ has a unique critical point, which is a unique global minimum. It is the *unique* fixed point, say μ_∞, of the mapping Π. Therefore $\lim_{t\to\infty} \mu_t = \mu_\infty$ $\mathbb{P}_{x,r,\mu} -$ a.s. We also refer the reader to the paper [**8**]. $\square$

Numerical simulation: The following figure has been obtained by numerical integration of (X_t) over the time interval $(0,T)$ for $T = 1000$, using a step size of 0.02. We have chosen $r = 0.2$, $x_0 = 0$ and $\mu_0 = \exp\{-\frac{1}{2}x^2\}/\sqrt{2\pi}$. Suppose that we work on $\mathbb{R}$, with the maps $V : \mathbb{R} \to \mathbb{R}; V(x) := \frac{1}{4}x^4 + \frac{1}{5}x^2 + 1$ and $W : \mathbb{R}\times\mathbb{R} \to \mathbb{R}; W(x,y) := 0.05\cos(x-y) + 0.05$. It is easily seen that, for all $y \in \mathbb{R}$, the function $x \mapsto V(x) + W(x,y)$ is strictly convex and thus satisfies the hypothesis of the preceding proposition. The figure represents the density of μ_T with respect to the Lebesgue measure (the caption with dashed lines represents the unique solution of $\mu = \Pi(\mu)$).

5. Study of the dynamical system Φ

5.1. Existence of the flow. We first recall that $\mathcal{M}(\mathbb{R}^d; V)$ equipped with the dual weighted supremum norm $||\cdot||_V$ is a Banach space. We start with an easy result that will be used many times:

LEMMA 5.1. *For any $\mu \in \mathcal{M}(\mathbb{R}^d; V)$ the function $W * \mu$ belongs to $\mathcal{C}^\infty(\mathbb{R}^d; V)$ and we have*

$$||W * \mu||_V \leq 2\kappa||\mu||_V.$$

PROOF. Straightforward thanks to the domination condition (3.2). $\square$

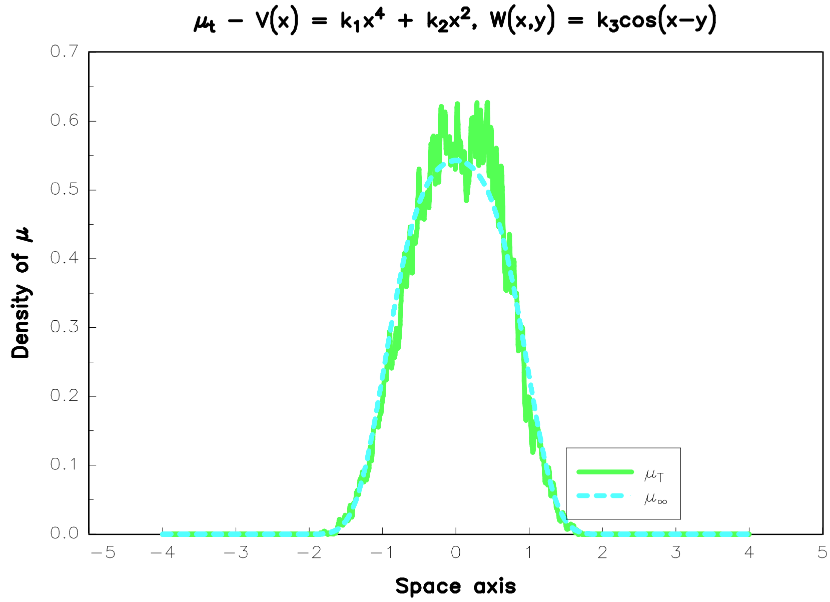

FIGURE 3. Illustration of Proposition 4.8.

Now let us consider the vector field

$$F : \mathcal{M}(\mathbb{R}^d; V) \to \mathcal{M}(\mathbb{R}^d; V), \;\; \mu \mapsto \Pi(\mu) - \mu. \tag{5.1}$$

By standard results of ordinary differential equations in Banach spaces, we can state the following result:

PROPOSITION 5.2. *F is a $\mathcal{C}^\infty$ vector field which induces the local flow $\Phi = (\Phi_t)$ in $\mathcal{M}(\mathbb{R}^d; V)$, that is the solution to the initial value problem $\dot{\mu} = F(\mu)$, $\mu(0) = \mu$ is the curve $t \mapsto \Phi_t(\mu)$ defined for t in some open interval $I_\mu := (\sigma_\mu, \tau_\mu)$, $-\infty \le \sigma_\mu < 0 < \tau_\mu \le +\infty$.*

PROOF. From the preceding lemma it is clear that the linear map $\mu \mapsto W * \mu$ is continuous from $\mathcal{M}(\mathbb{R}^d; V)$ to $\mathcal{C}^\infty(\mathbb{R}^d; V)$, hence $\mathcal{C}^\infty$. By composition, the map Π is $\mathcal{C}^\infty$ for the (strong) topology induced by the dual V-norm, and, as a result, F is a $\mathcal{C}^\infty$ vector field. The end of the proposition readily follows from the standard Cauchy theory of dynamical systems in Banach spaces. $\square$

Lemma 5.3. *For any $\beta > 0$, the application Π restricted to $\mathcal{P}_\beta(\mathbb{R}^d;V)$ is bounded and globally Lipschitz with constants depending on β only.*

Proof. First we need to show that $\mu \mapsto Z(\mu)$ is bounded from below. For $\mu \in \mathcal{P}_\beta(\mathbb{R}^d;V)$ we have, from the lemma above, $W * \mu(x) \leq 2\kappa\beta V(x)$, therefore we get:

$$Z(\mu) = \int_{\mathbb{R}^d} e^{-2W*\mu(x)}\gamma(\mathrm{d}x) \geq \int_{\mathbb{R}^d} e^{-2\kappa\beta V(x)}\gamma(\mathrm{d}x)$$

and thus we have the following bound for $\Pi(\mu)$:

$$||\Pi(\mu)||_V \leq \left(\int_{\mathbb{R}^d} e^{-2\kappa\beta V(x)}\gamma(\mathrm{d}x)\right)^{-1} \int_{\mathbb{R}^d} V(x)\gamma(\mathrm{d}x) =: C_\beta. \tag{5.2}$$

Now we know that Π is $\mathcal{C}^\infty$ on $\mathcal{M}(\mathbb{R}^d;V)$ with the strong topology. Its differential at the measure μ is the continuous linear operator $D\Pi(\mu) : \mathcal{M}(\mathbb{R}^d;V) \to \mathcal{M}(\mathbb{R}^d;V)$ defined by

$$D\Pi(\mu)\cdot\nu(\mathrm{d}x) := -2\left(W*\nu(x) - \int_{\mathbb{R}^d} W*\nu(y)\Pi(\mu)(\mathrm{d}y)\right)\Pi(\mu)(\mathrm{d}x). \tag{5.3}$$

Now we fix $\nu \in \mathcal{M}(\mathbb{R}^d;V)$. Since we have $|W*\nu(x)| \leq 2\kappa||\nu||_V V(x)$, we find that there exists a positive constant C such that

$$||D\Pi(\mu)\cdot\nu||_V \leq 4\kappa(1+C_\beta)||\nu||_V \int_{\mathbb{R}^d} V^2(x)\Pi(\mu)(\mathrm{d}x).$$

But for $\mu \in \mathcal{P}_\beta(\mathbb{R}^d;V)$, the same computation used for the bound of $\Pi(\mu)$ enables to control the last integral, hence we get a bound on the differential and Π is Lipschitz as stated. □

Corollary 5.4. *Let $\beta > 0$. The application F restricted to $\mathcal{P}_\beta(\mathbb{R}^d;V)$ is Lipschitz for the strong topology.*

Proof. Straightforward because Π is Lipschitz. □

Proposition 5.5. *We assume that W is either symmetric or bounded in the second variable (that is $W(x,y) \leq \kappa V(x)$). Then it holds:*
i) The semi-flow Φ (for $t \geq 0$) leaves $\mathcal{P}(\mathbb{R}^d;V)$ positively invariant.
ii) The semi-flow Φ_t restricted to the set $\mathcal{P}(\mathbb{R}^d;V)$ does not explode (in a finite time) and therefore $\tau_\mu = +\infty$.
iii) For all $\mu \in \mathcal{P}(\mathbb{R}^d;V)$ invariant under Φ, we have $I_\mu = \mathbb{R}$.

Proof. i) Straightforward because $\mathcal{P}(\mathbb{R}^d;V)$ is a convex set. As $\mathcal{P}(\mathbb{R}^d;V)$ is a metrizable space, let d be a metric on this set. We define, for $\nu \in \mathcal{P}(\mathbb{R}^d;V)$, the distance $d(\mathcal{P}(\mathbb{R}^d;V),\nu) := \inf\{d(\mu,\nu);\mu \in \mathcal{P}(\mathbb{R}^d;V)\}$. Let $\mu,\nu \in \mathcal{P}(\mathbb{R}^d;V)$ and $t \geq 0$. A trivial computation leads to

$$d(\Phi_t(\mu),\nu) \leq (1-t)d(\mu,\nu) + td(\Pi(\mu),\nu) + o(t),$$

and thus $d(\mathcal{P}(\mathbb{R}^d;V),\Phi_t(\mu)) = 0$ (we adapt the proof of the proposition 5.7).

ii) We first suppose that $W = W_2$ is bounded in y. We work with the V-norm. Since $W(x,y)$ is a bounded function in the second variable, that is $W(x,y) \leq \kappa V(x)$, we find that the normalization of $\Pi(\mu)$, that is $Z(\mu)$ is bounded from below in the following way:

$$Z(\mu) = \int_{\mathbb{R}^d} \exp\{-2 \int_{\mathbb{R}^d} W(x,y)\mu(\mathrm{d}y)\}\gamma(\mathrm{d}x) \geq \int_{\mathbb{R}^d} \exp\{-2\kappa V(x)\}\gamma(\mathrm{d}x).$$

As a consequence, we simply have that

$$||\Pi(\mu)||_V \leq \int_{\mathbb{R}^d} V(x)\gamma(\mathrm{d}x) \frac{1}{\int_{\mathbb{R}^d} e^{-2\kappa V(x)}\gamma(\mathrm{d}x)} =: C.$$

Therefore, we obtain the upper bound

$$||\dot{\mu}_t||_V \leq ||\mu_t||_V + C$$

and the result follows the classical theory of ODE on Banach space: F is completely integrable and generates a C^∞ semi-flow $\Phi : \mathbb{R}_+ \times \mathcal{P}(\mathbb{R}^d; V) \to \mathcal{P}(\mathbb{R}^d; V)$.

Let us suppose that W is symmetric, that is $W(x,y) = W(y,x)$. Now, we remark that the free energy introduced in the preceding section is not a Lyapunov function for the dynamical system because most of the times $\phi_t(\mu)$ is not an absolutely continuous probability measure with respect to γ and therefore, $\mathcal{F}(\Phi_t(\mu)) = \infty$. Therefore, we consider the Lyapunov function $\mathcal{E}(\mu) := \mathcal{F}(\Pi(\mu))$. Actually, $\mathcal{F}$ is a $\mathcal{C}^\infty$ function for the strong topology (the V-norm). We compute

$$D\mathcal{F}(\mu) \cdot \nu = \int_{\mathbb{R}^d} \left[\log\left(\frac{\mathrm{d}\mu}{\mathrm{d}\gamma}(x) \right) + 2W * \mu(x) \right] \mathrm{d}\nu(x). \tag{5.4}$$

But we recall that Π is $\mathcal{C}^\infty$ and we know $D\Pi$ given by the equation (5.3). As a consequence, we obtain

$$\begin{aligned} D\mathcal{E}(\mu) \cdot \nu &= D\mathcal{F}(\Pi(\mu)) \circ D\Pi(\mu) \cdot \nu \\ &= -4 \int_{\mathbb{R}^d} (W * \Pi(\mu) - W * \mu) \left(W * \Pi(\nu) - \int_{\mathbb{R}^d} W * \nu \mathrm{d}\Pi(\mu) \right) \mathrm{d}\Pi(\mu). \end{aligned}$$

Now, it remains to choose $\nu = \Pi(\mu) - \mu$ in order to obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}(\Phi_t(\mu)) = -4 \int_{\mathbb{R}^d} (W * \nu(x))^2 \mathrm{d}\Pi(\mu)(x) + 4 \left(\int_{\mathbb{R}^d} W * \nu(x) \mathrm{d}\Pi(\mu)(x) \right)^2 \leq 0.$$

As a consequence, for all positive constant $c > 0$, the set $\{\mu; \mathcal{E}(\mu) \leq c\}$ is compact. Finally, we get for all $t \geq 0$, $\mathcal{E}(\Phi_t(\mu)) \leq \mathcal{E}(\mu)$ and we conclude by applying the Cauchy-Lipschitz theorem.

iii) Let $\mu \in \mathcal{P}(\mathbb{R}^d; V)$ such that for all $t \in I_\mu$, we have $\Phi_t(\mu) \in \mathcal{P}(\mathbb{R}^d; V)$. We know by the first point that $\tau_\mu = +\infty$ and it remains to prove that $\sigma_\mu = -\infty$. But we just adapt the previous proof to show that for all $t \geq 0$, we get $\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}(\Phi_{-t}(\mu)) \geq 0$. We conclude in the same way. □

5.2. An important set. Here, we introduce a crucial object for the analysis of the dynamical system Φ:

$$\widehat{\mathrm{Im}(\Pi)} := \left\{ \int_{\mathcal{P}(\mathbb{R}^d;V)} \Pi(\mu)\rho(\mathrm{d}\mu); \rho \in \mathcal{P}_w(\mathcal{P}(\mathbb{R}^d;V)) \right\} \tag{5.5}$$

where $\mathcal{P}_w(\mathcal{P}(\mathbb{R}^d;V))$ is the topological space obtained by endowing $\mathcal{P}(\mathcal{P}(\mathbb{R}^d;V))$ with the topology of the weak* convergence. The set $\widehat{\mathrm{Im}(\Pi)}$ is well defined, because $\mathcal{P}(\mathbb{R}^d;V)$ is a Polish space and also $\mathcal{P}_w(\mathcal{P}(\mathbb{R}^d;V))$. It represents the convex hull of $\mathrm{Im}(\Pi)$. In this definition of $\widehat{\mathrm{Im}(\Pi)}$, $\int_{\mathcal{P}(\mathbb{R}^d;V)} \Pi(\mu)\rho(\mathrm{d}\mu)$ denotes a probability measure such that, for all $f \in \mathcal{C}^0(\mathbb{R}^d;V)$,

$$\int_{\mathcal{P}(\mathbb{R}^d;V)} \Pi(\mu)\rho(\mathrm{d}\mu) \cdot f = \int_{\mathcal{P}(\mathbb{R}^d;V)} \Pi(\mu) f \rho(\mathrm{d}\mu).$$

We observe that the mapping $\rho \mapsto \int_{\mathcal{P}(\mathbb{R}^d;V)} \Pi(\mu)\rho(\mathrm{d}\mu)$ maps continuously $\mathcal{P}(\mathcal{P}(\mathbb{R}^d;V))$ to $\mathcal{P}(\mathbb{R}^d;V)$.

Now we need to refresh the reader's memory with a (not too short!) list of important definitions of the theory of dynamical systems. First we recall that $\mathcal{P}(\mathbb{R}^d;V)$ is a metrizable space and we can choose the following metric: since $\mathcal{C}^0(\mathbb{R}^d;V)$ is separable, we exhibit a sequence $(f_k)_k$ dense in $\{f \in \mathcal{C}^0(\mathbb{R}^d;V)/||f||_V \leq 1\}$, and set for all $\mu, \nu \in \mathcal{P}(\mathbb{R}^d;V)$:

$$\mathrm{dist}(\mu,\nu) := \sum_{k=1}^{\infty} 2^{-k} |\mu(f_k) - \nu(f_k)|. \tag{5.6}$$

Definition 5.6. *a) A subset A of $\mathcal{P}(\mathbb{R}^d;V)$ is positively invariant (resp. negatively invariant, resp. invariant) for Φ provided $\Phi_t(A) \subset A$ (respectively $A \subset \Phi_t(A)$, respectively $\Phi_t(A) = A$) for all $t \geq 0$.*

b) A subset A of $\mathcal{P}(\mathbb{R}^d;V)$ is an attracting set (respectively attractor) for Φ provided:

1. *A is nonempty, compact for the weak* topology and positively invariant, (respectively invariant) and*
2. *A has a neighborhood $\mathcal{N} \subset \mathcal{P}(\mathbb{R}^d;V)$ such that $\mathrm{dist}(\Phi_t(\mu), A) \to 0$ as $t \to +\infty$ uniformly in $\mu \in \mathcal{N}$.*

c) The basin of attraction of an attractor $K \subset A$ for $\Phi|A = (\Phi_t|A)_t$ is the positively invariant open set (in A) comprising all points x such that $\mathrm{dist}(\Phi_t(x), A) \to 0$ as $t \to +\infty$, it means that it is the set of measures in the space $\mathcal{P}(\mathbb{R}^d;V)$ such that all initial conditions chosen in this set dynamically evolve to A

$$B(K, \Phi|A) = \{\mu \in A; \lim_{t\to\infty} \mathrm{dist}(\Phi_t(\mu), K) = 0\}.$$

d) A global attracting set (respectively global attractor) is an attracting set (respectively attractor) whose basin is the whole space $\mathcal{P}(\mathbb{R}^d;V)$.

e) Let A be a positively invariant set for Φ. An attractor for $\Phi|A = (\Phi_t|A)_t$ (defined by taking the restriction of (Φ_t) to A) is proper if it is different from $\emptyset$ and A.

f) An attractor-free set is a nonempty compact invariant set A such that $\Phi|A$ has no proper attractor.

Now we can state and prove the following:

PROPOSITION 5.7. *([**6**]) $\widehat{\mathrm{Im}(\Pi)}$ contains every subset of $\mathcal{P}(\mathbb{R}^d; V)$ negatively invariant under Φ.*

PROOF. First, by definition, $\widehat{\mathrm{Im}(\Pi)}$ is convex and contains $\Pi(\mathcal{P}(\mathbb{R}^d; V))$. We let $\mathrm{dist}(\mu, \nu)$ be the metric defined by (2.4) on $\mathcal{P}(\mathbb{R}^d; V)$ and obviously define $\mathrm{dist}(\mu, X) = \inf\{\mathrm{dist}(\mu, \nu); \nu \in X\}$. For $\mu \in \mathcal{P}(\mathbb{R}^d; V)$ and $\nu \in \widehat{\mathrm{Im}(\Pi)}$, we have for all $f \in \mathcal{C}^\infty(\mathbb{R}^d; V)$

$$|\Phi_t(\mu)f - \nu f| = |(1-t)\mu f + t\Pi(\mu)f - \nu f + o(t)| \leq |(1-t)\mu f + t\Pi(\mu)f - \nu f| + o(t).$$

By the definition of dist, it then implies

$$\mathrm{dist}(\Phi_t(\mu), \widehat{\mathrm{Im}(\Pi)}) \leq \mathrm{dist}((1-t)\mu + t\Pi(\mu), \widehat{\mathrm{Im}(\Pi)}) + o(t).$$

But the mapping $\mu \mapsto \mathrm{dist}(\mu, \widehat{\mathrm{Im}(\Pi)})$ is convex. As a consequence, and because $\Pi(\mu) \in \widehat{\mathrm{Im}(\Pi)}$,

$$\mathrm{dist}(\Phi_t(\mu), \widehat{\mathrm{Im}(\Pi)}) \leq (1-t)\mathrm{dist}(\mu, \widehat{\mathrm{Im}(\Pi)}) + o(t).$$

Since the function $\mu \mapsto \mathrm{dist}(\mu, \widehat{\mathrm{Im}(\Pi)})$ is convex, $t \mapsto \mathrm{dist}(\Phi_t(\mu), \widehat{\mathrm{Im}(\Pi)})$ admits a right derivative. It stems that

$$\left.\frac{\mathrm{d}^+}{\mathrm{d}t}\right|_{t=0} \mathrm{dist}(\Phi_t(\mu), \widehat{\mathrm{Im}(\Pi)}) \leq -\mathrm{dist}(\mu, \widehat{\mathrm{Im}(\Pi)}).$$

The invariance by time-translation shows that the same estimate must hold at any other time t that we derived at $t = 0$. Thus, we obtain for all $\mu \in \mathcal{P}(\mathbb{R}^d; V)$ and $t \geq 0$:

$$\mathrm{dist}(\Phi_t(\mu), \widehat{\mathrm{Im}(\Pi)}) \leq e^{-t}\mathrm{dist}(\mu, \widehat{\mathrm{Im}(\Pi)}).$$

This proves that $\widehat{\mathrm{Im}(\Pi)}$ contains every subset of $\mathcal{P}(\mathbb{R}^d; V)$ negatively invariant under Φ. □

REMARK 5.8. *It is easily proved that $\widehat{\mathrm{Im}(\Pi)}$ is a positively invariant set for the flow Φ.*

It is already known since Conley [**10**] that if U is an open set with compact closure, such that $\Phi_T(\overline{U}) \subset U$ for some $T > 0$, then there exists an attractor $A \subset U$ whose basin contains $\overline{U}$.

5.3. Fixed points of Π. In this subsection, we prove that for a symmetric W, the normalized occupation measure converges a.s. to a zero of a certain functional, the free energy.

We recall the free energy corresponding the ODE studied

$$\mathcal{F}(\mu) := \int_{\mathbb{R}^d} \log\left(\frac{\mathrm{d}\mu}{\mathrm{d}\gamma}\right) \mathrm{d}\mu + \int_{\mathbb{R}^d\times\mathbb{R}^d} W(x,y)\mu(\mathrm{d}x)\mu(\mathrm{d}y). \tag{5.7}$$

We note that this functional is finite if and only if the measure μ is absolutely continuous with respect to the Lebesgue measure. Therefore, for $\mu \in \mathcal{P}_\beta(\mathbb{R}^d; V)$, we will consider $\mathcal{F}(\Pi(\mu))$ or equivalent only the probability measures having a density with respect to the Lebesgue measure.

We will prove that there is a link between the zeros of F and the energy $\mathcal{F}$: for any probability μ on $\mathbb{R}^d$, when have that $F(\mu) = 0$ if and only if μ has a density and μ is a critical point for the free energy $\nabla\mathcal{F}(\mu) = 0$.

We know that if the normalized occupation measure μ_t converges, then the limit will be a probability measure which has a density with respect to γ. Thus, instead of working with the function $F(\mu) = -\mu + \Pi(\mu)$, we will only consider probability measures μ having a density f: $\mu = f\gamma$.

PROPOSITION 5.9. *Suppose that W is symmetric. Then the zeros of F are the critical points of $\mathcal{F}$.*

PROOF. We easily prove that $\mathcal{F}$ is a strictly convex function. Actually, denote by $D\mathcal{F}$ its differential function. Let $\mu, \nu \in \mathcal{P}_\beta(\mathbb{R}^d; V)$ two probability measures having a density with respect to the Lebesgue measure. We remind that the differential function is given by the following (because W is symmetric):

$$D\mathcal{F}(\mu)\cdot\nu \quad = \quad \int_{\mathbb{R}^d} \log\left(\frac{\mu(x)}{\gamma(x)}\right)\nu(x)\mathrm{d}x + 2\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} W(x,y)\nu(x)\mu(y)\mathrm{d}x\mathrm{d}y.$$

Similarly, we get for $\alpha \in \mathcal{P}_\beta(\mathbb{R}^d; V)$ a probability measure having a density with respect to the Lebesgue measure

$$D^2\mathcal{F}(\mu)\cdot(\nu,\xi) = \int_{\mathbb{R}^d} \nu(x)\xi(x)\mu(x)^{-1}\gamma(x)\mathrm{d}x + \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} W(x,y)\nu(x)\xi(y)\mathrm{d}x\mathrm{d}y$$

and the strict convexity follows. The convexity of $\mathcal{F}$ implies that for all probability measures $\mu, \nu \in \mathcal{P}_\beta(\mathbb{R}^d; V)$ absolutely continuous with respect to the Lebesgue measure,

$$\mathcal{F}(\nu) - \mathcal{F}(\mu) \geq D\mathcal{F}(\mu)\cdot(\nu-\mu).$$

We can choose $\nu = \Pi(\mu)$ (which is a global minimum for $\mathcal{F}$!) and apply the preceding inequality which gives

$$0 \geq \mathcal{F}(\Pi(\mu)) - \mathcal{F}(\mu) \geq D\mathcal{F}(\mu)\cdot(\Pi(\mu)-\mu)\,.$$

Now, by convexity of $\mathcal{F}$, $D\mathcal{F}(\mu)\cdot(\Pi(\mu)-\mu) = 0$ if and only if $\mu = \Pi(\mu)$ (*i.e.* $F(\mu) = 0$). □

Corollary 5.10. *Suppose that W is symmetric. Then the fixed points of Π are the critical points of the free energy $\mathcal{F}$.*

Proof. Straightforward. $\square$

6. Study of the family of semi-groups $(P_t^\mu, t \geq 0, \mu \in \mathcal{M}_\beta(\mathbb{R}^d; V))$

In this section, we introduce two crucial functional inequalities, namely the spectral gap and the ultracontractivity for the family of semi-groups P_t^μ. The notion of ultracontractivity and its relation to the analysis of Markov semi-groups were studied by Davies and Simon [**13**] and recently by Röckner & Wang [**31**] for more general diffusions. The need of the ultracontractivity property will impose some kind of boundedness on the convolution term in the SDE that cannot be easily removed. We recall that $\Pi(\mu)(\mathrm{d}x) = \frac{e^{-2W*\mu(x)}}{Z(\mu)}\gamma(\mathrm{d}x)$ and that we denote by $(P_t^\mu, t \geq 0, \mu \in \mathcal{P}_\beta(\mathbb{R}^d; V))$ the Feller semigroup associated to the diffusion $(X_t^\mu, t \geq 0)$ (3.8). (P_t^μ) is a symmetric hypoelliptic semi-group.
Let $L^2(\Pi(\mu))$ denote the space of Borel real-valued functions $f : \mathbb{R}^d \to \mathbb{R}$ such that $\int_{\mathbb{R}^d} |f(x)|^2 \Pi(\mu)(\mathrm{d}x) < \infty$, and

$$(f, g)_\mu := \int_{\mathbb{R}^d} f(x)g(x)\Pi(\mu)(\mathrm{d}x)$$

is the inner product on this space and $||.||_{2,\mu}$ is the associated norm.

Remark 6.1. *For any probability measure $\mu \in \mathcal{P}_\beta(\mathbb{R}^d; V)$, $L^2(\Pi(\mu)) = L^2(\gamma)$.*

We introduce here two operators : Q_μ, the "inverse" of A_μ, that is the operator defined for any function f by

$$Q_\mu f := \int_0^\infty (P_t^\mu f - \Pi(\mu)f)\,\mathrm{d}t \tag{6.1}$$

and K_μ, the orthogonal projection defined by

$$K_\mu f := f - \Pi(\mu)f. \tag{6.2}$$

They are linked together by the following relations (where $\mathcal{D}^2(\mu)$ is the domain of A_μ in $L^2(\Pi(\mu))$)

$$\forall f \in L^2(\Pi(\mu)),\ \ A_\mu \circ Q_\mu(f) = -K_\mu f,$$
$$\forall f \in \mathcal{D}^2(\mu),\ \ Q_\mu \circ A_\mu(f) = -K_\mu f.$$

Remark 6.2. *The integrability of $(P_t^\mu f - \Pi(\mu)f)$ will come from the uniform spectral gap obtained in the following subsection.*

6.1. Uniform spectral gap. We refer to Bakry [**2**], and we consider the functional algebra $\mathcal{C}^\infty(\mathbb{R}^d;V)$. Indeed, it is contained in $L^p(\Pi(\mu))$, for all $1<p<\infty$ and dense in all of them.

LEMMA 6.3. *The family of measures $e^{-2W*\mu(x)}\gamma(\mathrm{d}x)$, $\forall\mu\in\mathcal{P}_\beta(\mathbb{R}^d;V)$, satisfies a logarithmic Sobolev inequality and there exists a uniform spectral gap for the family of measures $e^{-2W*\mu(x)}\gamma(\mathrm{d}x)$, $\forall\mu\in\mathcal{P}_\beta(\mathbb{R}^d;V)$. These two facts correspond to the following inequalities: $\exists C_1, C_2$, two positive constants, independent of μ, such that $\forall f\in\mathcal{C}^\infty(\mathbb{R}^d;V)$:*

$$i)\quad \int f^2\log\left(\frac{f^2}{||f||_{2,\mu}}\right)e^{-2W*\mu}\mathrm{d}\gamma \le C_2\int|\nabla f|^2e^{-2W*\mu}\mathrm{d}\gamma.$$

$$ii)\quad \int f^2e^{-2W*\mu}\mathrm{d}\gamma-\left(\int fe^{-2W*\mu}\mathrm{d}\gamma\right)^2\le C_1\int|\nabla f|^2e^{-2W*\mu}\mathrm{d}\gamma.$$

Furthermore, $\forall t>0$, $||P_t^\mu(K_\mu f)||_{2,\mu}\le e^{-t/C_1}||K_\mu f||_{2,\mu}$.

PROOF. i) We can decompose the function W in two parts: $W=W_1+W_2$, where W_2 is a bounded function with respect to x and W_1 satisfies the curvature bound (3.3) given in the preliminaries: there exists $\alpha<1/2$ such that for all $\mu\in\mathcal{P}_\beta(\mathbb{R}^d;V)$, for all $x,\xi\in\mathbb{R}^d$

$$\left((\frac{1}{2}\nabla^2V(x)+\nabla^2W_1*\mu(x))\xi,\xi\right)\ge-\alpha K|\xi|^2.$$

We will begin to prove the first part of the lemma for the measure $\gamma_1(\mathrm{d}x)$, corresponding to $V+W_1$, thanks to the abstract Γ_2 criterion due to Bakry & Emery. Then, this measure will be "perturbed" by $W_2*\mu$ and we will prove the result for $e^{-2W_2*\mu(x)}\gamma_1(\mathrm{d}x)=e^{-2W*\mu(x)}\gamma(\mathrm{d}x)$.

To the operator $A_1^\mu:=\frac{1}{2}\Delta-(\nabla V,\nabla)-(\nabla W_1*\mu,\nabla)$, we can associate the operator "carré du champ"

$$\Gamma(f):=|\nabla f|^2$$

and the operator

$$\Gamma_2^\mu(f):=|\nabla^2f|^2+(\nabla f,\nabla^2(V+W_1*\mu)\ \nabla f),$$

as in Bakry & Emery [**3**]. The strictly uniformly (in μ) convex function $V+W_1*\mu$ satisfies the Γ_2 criterion, which in turn implies the Bakry-Emery curvature criterion: $\exists C>0$ such that for all $\mu\in\mathcal{P}_\beta(\mathbb{R}^d;V)$

$$\inf_\mu\Gamma_2^\mu(f)\ge C\Gamma(f)$$

where $C\ge\left(\frac{1}{2}-a\right)K>0$ is independent of μ.

It follows from a lemma due to Holley & Stroock [**18**] (see for instance [**1**]) that the measure $e^{-2W_2*\mu(x)}\gamma_1(dx)$ satisfies a logarithmic Sobolev inequality, with a constant independent of μ smaller than $\frac{2}{C}e^{2M}$, because W_2 is bounded (in x) and there exists a constant M such that for all $\mu\in\mathcal{P}(\mathbb{R}^d;V)$, we have $\mathrm{osc}(W_2*\mu):=\sup(W_2*\mu)-\inf(W_2*\mu)\le M(=\kappa\beta)$. Thus we are done.

ii) Rothaus [**33**] has proved that if a measure satisfies a logarithmic Sobolev inequality with a constant c, then it also satisfies a Poincaré inequality with a constant $1/c$. Moreover, it is known that a Poincaré inequality is equivalent to the existence of a spectral gap.

From ii), we find the following estimate on the semigroup $(P_t^\mu)_{t\geq 0}$ (see [**2**]): there exists some positive constant C_1, independent of μ, such that for all $t>0$, $f\in L^2(\Pi(\mu))$,

$$||P_t^\mu(K_\mu f)||_{2,\mu} \leq e^{-t/C_1}||K_\mu f||_{2,\mu}.$$

$\square$

6.2. Uniform ultracontractivity. We introduce the following definition:

Definition 6.4. *Let $(P_t^\lambda, t\geq 0, \lambda\in\Lambda)$ a family of Markov semi-groups. We assume that for each parameter $\lambda\in\Lambda$ the semi-group $(P_t^\lambda, t\geq 0)$ has a unique invariant probability measure μ_λ. We say that the family $(P_t^\lambda, t\geq 0, \lambda\in\Lambda)$ is uniformly ultracontractive if the operators P_t^λ are uniformly bounded from $L^2(\mu_\lambda)$ to $L^\infty(\mu_\lambda)$ for all $t\in]0,\varepsilon]$ (for some $\varepsilon>0$), i.e. $\exists C>0$, independent of λ, such that*

$$||P_t^\lambda||_{2\to\infty} \leq C(t), \quad \forall t\in]0,\varepsilon], \forall\lambda\in\Lambda, \tag{6.3}$$

where $||P_t^\lambda||_{2\to\infty} := \sup\limits_{f\in L^2(\mu_\lambda)\setminus\{0\}} \frac{||P_t^\lambda f||_\infty}{||f||_2}$.

Let $q:\mathbb{R}^d\to\mathbb{R}; x\mapsto q(x)=\frac{|x|^2}{2}$. In order to prove that the family of semigroups $(P_t^\mu)_{t\geq 0}$ is uniformly ultracontractive, we will use the following result due to Röckner & Wang ([**31**] corollary 2.5):

Lemma 6.5. *(Röckner & Wang) Let $(P_t, t\geq 0)$ be a Markovian semi-group, with infinitesimal generator $A := \frac{1}{2}\Delta - (\nabla U, \nabla)$, with $\nabla^2 U\geq -K$. Assume that there exists a continuous increasing map $\chi:\mathbb{R}_+\mapsto\mathbb{R}_+\setminus\{0\}$ such that*

(1) $\lim\limits_{r\to\infty}\frac{\chi(r)}{r}=\infty$,

(2) *the mapping $g_\chi(r) := r\chi(m\log r)$ is convex on $[1,\infty)$ for any $m>0$,*

(3) *$Aq(x)\leq b-\chi(q(x))$ for some $b>0$.*

Then P_t has a unique invariant probability measure. If $\int_2^\infty \frac{\mathrm{d}r}{r\chi(m\log r)}<\infty$, $m>0$, then P_t is ultracontractive. If moreover $\chi(r)=\chi r^\delta$, with $\chi>0,\delta>1$, then there exists some $c=c(b,\chi)>0$ such that for all $t\in(0,1]$

$$||P_t||_{2\to\infty}\leq \exp\{ct^{-\delta/(\delta-1)}\}.$$

COROLLARY 6.6. *The family of semigroups $(P_t^\mu, t \geq 0, \mu \in \mathcal{M}_\beta(\mathbb{R}^d; V))$ is uniformly ultracontractive and $||P_t^\mu||_{2\to\infty} \leq \exp\{ct^{-\delta/(\delta-1)}\}$, where the positive constant c is uniform.*

PROOF. We apply the result of Röckner & Wang for $U := V + W * \mu$ and use the curvature bound (3.3). We find that each $(P_t^\mu)_{t\geq 0}$ is ultracontractive. Furthermore, the curvature bound (3.3) on $1/2V + W_1 * \mu$, the boundedness of $\nabla_x W_2$ and the growth condition $(\nabla V(x), x) \geq c|x|^{2\delta}$, all together imply that there exist $a, b > 0$ such that for $|x|$ large enough and for any initial $\mu \in \mathcal{P}_\beta(\mathbb{R}^d; V)$ and any $r > 0$

$$\begin{aligned} A_{\mu_t} q(x) &= d/2 - (\nabla W * \mu_t(x), x) - (\nabla V(x), x) \\ &\leq b - aq(x)^\delta \leq b - aq(x). \end{aligned} \tag{6.4}$$

If we let $\chi(t) := t^\delta$, with $\delta > 1$, we find that the constant c is uniform in μ. Thus, we have the uniform ultracontractivity. $\square$

Now we work with the three operators: A_μ, K_μ and Q_μ. We recall that $K_\mu f = f - \Pi(\mu) f$ and $Q_\mu f = \int_0^\infty P_t^\mu (K_\mu f) \mathrm{d}t = \int_0^\infty P_t^\mu (f - \Pi(\mu) f) \, \mathrm{d}t$.

PROPOSITION 6.7. *For all $\varepsilon > 0$, there exists a positive constant $K(\varepsilon)$ such that for all $\mu \in \mathcal{P}_\beta(\mathbb{R}^d; V)$, $x \in \mathbb{R}^d$, $f \in \mathcal{C}^0(\mathbb{R}^d; V)$:*

$$|Q_\mu f(x)| \leq (\varepsilon V(x) + K(\varepsilon))||f||_V.$$

If $f \in \mathcal{C}^\infty(\mathbb{R}^d; V)$, then $Q_\mu f \in \mathcal{C}^1(\mathbb{R}^d)$ and $|\nabla Q_\mu f(x)| \leq C(\varepsilon V(x) + K)||f||_V$.

PROOF. One has clearly

$$|Q_\mu f(x)| \leq \int_0^\infty |P_t^\mu (K_\mu f)(x)| \mathrm{d}t.$$

Let t_0 be a positive constant (we will choose it precisely later). We can decompose the right-hand side of the preceding inequality:

$$\int_0^\infty |P_t^\mu (K_\mu f)(x)| \mathrm{d}t = \int_0^{t_0} |P_t^\mu (K_\mu f)(x)| \mathrm{d}t + \int_{t_0}^\infty |P_t^\mu (K_\mu f)(x)| \mathrm{d}t.$$

We begin to work with the second term. By use of the uniform ultracontractivity of the semi-group (P_t^μ) and the uniform spectral gap, we have the following

$$\begin{aligned}
\int_{t_0}^{\infty} |P_t^{\mu}(K_\mu f)(x)| \mathrm{d}t &= \int_0^{\infty} |P_{t_0}^{\mu} P_t^{\mu}(K_\mu f)(x)| \mathrm{d}t \\
&\leq \exp\{ct_0^{-\delta/(\delta-1)}\} \int_0^{\infty} ||P_t^{\mu}(K_\mu f)||_{2,\mu} \mathrm{d}t \\
&\leq \exp\{ct_0^{-\delta/(\delta-1)}\} ||K_\mu f||_{2,\mu} \int_0^{\infty} e^{-t/C_1} \mathrm{d}t \\
&\leq \exp\{ct_0^{-\delta/(\delta-1)}\} ||f||_{2,\mu} \int_0^{\infty} e^{-t/C_1} \mathrm{d}t \\
&\leq \exp\{ct_0^{-\delta/(\delta-1)}\} ||f||_V \left(\int V^2 \mathrm{d}\Pi(\mu)\right)^{1/2} \int_0^{\infty} e^{-t/C_1} \mathrm{d}t.
\end{aligned}$$

We now have to work with the first term of the equality. We easily know that

$$|P_t^{\mu} f(x)| \leq ||f||_V P_t^{\mu} V(x).$$

But we recall that for the Lyapunov function $\mathcal{E}_\mu(x) = V(x) + W * \mu(x)$, we have

$$\mathbb{E}_x \mathcal{E}_{\mu_t}(X_t) \leq (\mathcal{E}_\mu(x) + 1 + C)e^{Ct}$$

and by definition $P_t^{\mu} V(x) \leq P_t^{\mu}(\mathcal{E}_\mu)(x) = \mathbb{E}_x \mathcal{E}_\mu(X_t)$. We need a last bound for $\mathcal{E}_\mu(x)$:

$$\mathcal{E}_\mu(x) \leq V(x) + \kappa \int (V(x) + V(y))\mu(\mathrm{d}y) \leq (1 + \kappa + \kappa\beta)V(x).$$

We finally find

$$|P_t^{\mu} f(x)| \leq (1 + \kappa + \kappa\beta + 1 + C)V(x) \int_0^{t_0} e^{Cs} \mathrm{d}s ||f||_V.$$

We can choose t_0 such that $|P_t^{\mu} f(x)| \leq \varepsilon V(x) ||f||_V$. We can conclude that there exists some $K(\varepsilon)$ which satisfies the lemma:

$$|Q_\mu f(x)| \leq (\varepsilon V(x) + K(\varepsilon)) ||f||_V. \tag{6.5}$$

Suppose that f is smooth. We remind the notations: $\Gamma(f) := |\nabla f|^2$ and $\Gamma_2^{\mu}(f) := |\nabla^2 f|^2 + (\nabla f, \nabla^2(V + W * \mu)\nabla f)$ and there exists a real number α such that $\Gamma_2^{\mu}(f) \geq \alpha \Gamma(f)$. The Γ_2-criterion implies the following inequality (see for instance Ledoux [**22**] or [**1**] p.83), $\forall f \in \mathcal{C}^{\infty}(\mathbb{R}^d; V)$,

$$|\nabla P_t^{\mu}(K_\mu f)| \leq C\left(|P_t^{\mu}(K_\mu f)^2| + |P_t^{\mu}(K_\mu f)|^2\right) \quad \forall t > 0 \tag{6.6}$$

where $C = C(\alpha) > 0$. The rest of the proof follows exactly the first part of it. We conclude that there exists $C > 0$ such that $\forall \varepsilon > 0$, there exists $K(\varepsilon)$ such that

$$|\nabla Q_\mu f(x)| \leq C(\varepsilon V(x) + K(\varepsilon)) ||f||_V.$$

$\square$

6.3. Regularity with respect to the measure μ.

Proposition 6.8. *The mappings $\mu \mapsto A_\mu$ and $\mu \mapsto K_\mu$ are $\mathcal{C}^\infty$ and for any function $f \in \mathcal{C}^2(\mathbb{R}^d; V)$, the application $\mu \mapsto Q_\mu f$ is $\mathcal{C}^\infty$ for the strong topology of measures and we have for the (first) differentials (for any $\mu, \nu \in \mathcal{M}(\mathbb{R}^d; V)$):*

$$\begin{aligned}
D(A_\mu f) \cdot \nu &= -(\nabla W * \nu, \nabla f);\\
D(K_\mu f) \cdot \nu &= - < D\Pi(\mu) \cdot \nu, f >;\\
D(Q_\mu f) \cdot \nu &= < D\Pi(\mu) \cdot \nu, Q_\mu f > + Q_\mu((\nabla W * \nu), \nabla Q_\mu f),
\end{aligned}$$

with the notation $< D\Pi(\mu) \cdot \nu, f > := \int_{\mathbb{R}^d} f(x) D\Pi(\mu) \cdot \nu(\mathrm{d}x)$.

Proof. Let $\beta > 0$ and fix $\mu \in \mathcal{M}_\beta(\mathbb{R}^d; V)$. We already know that $\mu \mapsto W * \mu$ and Π are $\mathcal{C}^\infty$; so there is nothing to prove in case of A_μ or K_μ. To look at Q_μ we need to consider the resolvent operator of P_t^μ:

$$R_\lambda^\mu f := \int_0^\infty e^{-\lambda t} P_t^\mu f \,\mathrm{d}t,\ \lambda > 0.$$

We recall that for all $\lambda > 0$ we have $R_\lambda^\mu = (\lambda - A_\mu)^{-1}$ (see *e.g.* [**19**]). For all $\lambda > 0$, we define

$$Q_\mu(\lambda) := K_\mu(\lambda - A_\mu)^{-1}.$$

We now have the following equality $Q_\mu f - Q_\mu(\lambda) f = \int_0^\infty \mathrm{d}t P_t^\mu K_\mu (1 - e^{-\lambda t}) f$ and we obtain by the uniform spectral gap that there exists $C, C_1 > 0$

$$||Q_\mu f - Q_\mu(\lambda) f||_V \leq \lambda ||f||_V \int_0^\infty C(\beta) e^{-tC_1} \mathrm{d}t.$$

Besides, we have:

$$\begin{aligned}
Q_\mu f &= \lim_{\lambda \to 0+} \int_0^\infty e^{-\lambda t} \left(P_t^\mu f - \Pi(\mu) f\right) \mathrm{d}t\\
Q_\mu f &= \lim_{\lambda \to 0+} \left(R_\lambda^\mu f - \lambda^{-1} \Pi(\mu) f\right) \qquad (6.7)
\end{aligned}$$

where the convergence is readily seen to be uniform in μ. Therefore $\mu \mapsto Q_\mu f$ is continuous.

The mappings $\mu \mapsto \Pi(\mu)$ and $\mu \mapsto \nabla W * \mu$ are $\mathcal{C}^\infty$. This fact implies that $\mu \mapsto K_\mu$ and $\mu \mapsto A_\mu$ are $\mathcal{C}^\infty$. Therefore for every $\lambda > 0$ we find that the map $\mu \mapsto R_\lambda^\mu f$ is $\mathcal{C}^\infty$. We have the following differential[2]:

$$\begin{aligned}
DQ_\mu(\lambda) \cdot \nu &= (DK_\mu \cdot \nu)((\lambda - A_\mu)^{-1}) + K_\mu(\lambda - A_\mu)^{-1}(DA_\mu \cdot \nu)(\lambda - A_\mu)^{-1}\\
&= (D\Pi(\mu) \cdot \nu)((\lambda - A_\mu)^{-1}) + K_\mu(\lambda - A_\mu)^{-1}(\nabla W * \nu, \nabla)(\lambda - A_\mu)^{-1}.
\end{aligned}$$

[2]Rigorously speaking, we use here the differentiability of the operator-valued function $\mu \mapsto A_\mu$. For this to hold, we need to consider A_μ as a bounded operator from the space $L^2(\gamma)$ endowed with the norm

$$||f||_{2,\mu,1} := ||f||_{2,\mu} + ||A_\mu f||_{2,\mu}.$$

to the same space $L^2(\gamma)$ endowed with the standard quadratic norm. Then equipping the space of bounded operators with the topology induced by the operator norm, the claimed differentiability is obvious.

We will prove that each term of the preceding equality converges uniformly. For the first term, we have for all $f \in \mathcal{C}^\infty(\mathbb{R}^d; V)$:

$$(D\Pi(\mu)\cdot\nu)((\lambda - A_\mu)^{-1})f =< D\Pi(\mu)\cdot\nu, K_\mu(\lambda - A_\mu)^{-1}f >$$

and therefore

$$\lim_{\lambda\to 0+}(D\Pi(\mu)\cdot\nu)((\lambda - A_\mu)^{-1})f =< D\Pi(\mu)\cdot\nu, Q_\mu f > \tag{6.8}$$

where the convergence in uniform. Its remains to prove the convergence of the second term. We find the following

$$\begin{aligned} K_\mu(\lambda - A_\mu)^{-1}(\nabla W * \nu, \nabla)(\lambda - A_\mu)^{-1} &= K_\mu(\lambda - A_\mu)^{-1}(\nabla W * \nu, \nabla)K_\mu(\lambda - A_\mu)^{-1} \\ &= Q_\mu(\lambda)(\nabla W * \nu, \nabla Q_\mu(\lambda)f). \end{aligned}$$

If we manage to prove that $\nabla Q_\mu(\lambda)f$ converges uniformly to $\nabla Q_\mu f$, then we are done. We have by definition of $Q_\mu(\lambda)$:

$$\nabla Q_\mu(\lambda)f = \int_0^\infty \nabla(P_t^\mu f)e^{-\lambda t}\mathrm{d}t$$

and therefore

$$|\nabla Q_\mu f - \nabla Q_\mu(\lambda)f| \leq \int_0^\infty |\nabla(P_t^\mu f)|(1 - e^{-\lambda t})\mathrm{d}t.$$

We use the inequality (6.6) to prove that this family of differentials converge uniformly with respect to μ; so $\mu \mapsto Q_\mu f$ is actually $\mathcal{C}^1$ with the differential given in the statement of the proposition. Looking at this differential, we see that it is itself a $\mathcal{C}^1$ function of μ, so by induction it can be proved that $\mu \mapsto Q_\mu f$ is $\mathcal{C}^\infty$. □

Remark 6.9. *We could have also proved that $\mu \mapsto P_t^\mu f$ and $\mu \mapsto Q_\mu$ are $\mathcal{C}^\infty$. But this results will not be needed in the remainder.*

Corollary 6.10. *For every $f \in \mathcal{C}^0(\mathbb{R}^d; V)$, we have the uniform inequalities*

$$\begin{aligned} ||D(Q_\mu f)\cdot\nu||_V &\leq K||f||_V||\nu||_V, \\ |(DQ_\mu\cdot\nu)(f)(x)| &\leq C(\varepsilon + K(\varepsilon))^2||\nu||_V||f||_V. \end{aligned}$$

Proof. The first inequality is straightforward. We will prove the second one. We have the following:

$$|(DQ_\mu\cdot\nu)(f)(x)| \leq |(D\Pi(\mu)\cdot\nu)(Q_\mu f)(x)| + |Q_\mu(\nabla W * \nu(x), \nabla Q_\mu f(x))|.$$

We will treat each of the two terms on the right side of the inequality separately. If we consider the second member of the right side of the inequality, we find

$$\begin{aligned} |Q_\mu(\nabla W * \nu(x), \nabla Q_\mu f(x))| &\leq (\varepsilon V(x) + K(\varepsilon))^2||(\nabla W * \nu, \nabla Q_\mu f)||_{V^2} \\ &\leq (\varepsilon V(x) + K(\varepsilon))^2||\nabla W * \nu||_V||\nabla Q_\mu f)||_V \\ &\leq C(\varepsilon V(x) + K(\varepsilon))^2||\nu||_V||f||_V. \end{aligned}$$

We work now with the other member of the inequality.

$$\begin{aligned}|(D\Pi(\mu)\cdot\nu)(Q_\mu f)| &\leq 2\int |Q_\mu f(x)||W*\nu(x)-\int W*\nu(y)\Pi(\mu)(\mathrm{d}y)|\Pi(\mu)(\mathrm{d}x)\\ &\leq C||f||_V\int(\varepsilon V(x)+K(\varepsilon))(V(x)+2||\nu||_V+\alpha)\Pi(\mu)(\mathrm{d}x)\\ &\leq C||f||_V(\varepsilon+K(\varepsilon))||\nu||_V.\end{aligned}$$

Putting the pieces together, we are done. □

7. Behavior of the occupation measure

7.1. Tightness of $(\mu_t, t\geq 0)$**.** We prove by the following result that we almost work in a compact set:

Proposition 4.1 *Let* x, r, μ. *Then there exists* $\beta > 0$ *such that a.s.* $\mu_t \in \mathcal{P}_\beta(\mathbb{R}^d; V)$ *for all* $t \geq 0$.

Proof. We set $\phi(t) := \int_0^t V(X_s)\mathrm{d}s$. All we need to prove is there exists β' such that a.s. $\phi(t) \leq \beta' t)$. We use again the Lyapunov functional $\mathcal{E}_\mu(x) = V(x) + W*\mu(x)$. We have already shown:

$$\begin{aligned}\mathcal{E}_{\mu_t}(X_t) &= \mathcal{E}_\mu(x)+\int_0^t(\nabla\mathcal{E}_{\mu_s}(X_s),\mathrm{d}B_s)-\int_0^t|\nabla\mathcal{E}_{\mu_s}(X_s)|^2\,\mathrm{d}s\\ &+ \frac{1}{2}\int_0^t\Delta\mathcal{E}_{\mu_s}(X_s)\mathrm{d}s+\int_0^t(W(X_s,X_s)-W*\mu_s(X_s))\frac{\mathrm{d}s}{r+s}.\end{aligned}$$

The strong law of large numbers (for martingales) implies that for t large enough we will have $\int_0^t(\nabla\mathcal{E}_{\mu_s}(X_s),\mathrm{d}B_s) \leq \frac{1}{2}\int_0^t|\nabla\mathcal{E}_{\mu_s}(X_s)|^2\,\mathrm{d}s$, and we therefore get the inequality:

$$\int_0^t|\nabla\mathcal{E}_{\mu_s}(X_s)|^2\,\mathrm{d}s \leq 2\mathcal{E}_\mu(x)+\int_0^t\Delta\mathcal{E}_{\mu_s}(X_s)\mathrm{d}s+\frac{2}{r}\int_0^t W(X_s,X_s)\mathrm{d}s.$$

Now we want to find an integral inequality on ϕ. To this aim, we will control separately each of the three terms of the last inequality and let $\phi(t)$ appear.

- From the growth assumption on V, for any $\epsilon > 0$ we can find $k_\epsilon > 0$ such that $V \leq k_\epsilon + \epsilon|\nabla V|^2$, thus by means of an integration we get
$$\phi(t) \leq k_\epsilon t + \epsilon\int_0^t|\nabla\mathcal{E}_\mu(X_s)|^2\,\mathrm{d}s.$$
- From the domination condition (3.2) on W, we have $\Delta W*\mu(x) \leq \kappa(V(x)+\mu(V))$; the growth condition (3.4) on V ensures that we have the same kind of inequality $\Delta V \leq a + bV$ for some $a, b > 0$; therefore we get:
$$\Delta\mathcal{E}_{\mu_s}(X_s) \leq a + \kappa\mu(V) + (\kappa+b)V(X_s) + \frac{\kappa}{r+s}\phi(s).$$
- The same domination condition (3.2) leads also to
$$W(X_s,X_s) \leq 2\kappa V(X_s).$$

Putting the pieces together, we find the following inequality (denoting by C_0 and C_1 two deterministic positive constants)

$$\phi(t) \leq k_\epsilon t + \epsilon \left(C_0 + (b+\kappa)\phi(t) + \kappa \int_0^t \frac{\mathrm{d}s}{r+s}\phi(s) + C_1 t + \frac{4\kappa}{r}\phi(t) \right).$$

Now for ϵ small enough, we can have $\epsilon\left(b + \kappa(1+\frac{4}{r})\right) < 1$ thus, we have:

$$\phi(t) \leq C_0' + C_1' t + \int_0^t \frac{C_2' \mathrm{d}s}{r+s}\phi(s),$$

with C_i' another positive constants. If we note $u(t) := \int_0^t \frac{\mathrm{d}s}{r+s}\phi(s)$, solving the preceding inequality is equivalent to solving the inequality $\dot{u} \leq C + C_2' \frac{u}{r+t}$ which solution is $u(t) \leq C(r+t)^{C_2'}$. Finally, we have that there exists a positive deterministic constant β' such that $\phi(t) \leq \beta' t$ as required (because $C_2' < 1$). □

Proposition 7.1. *$\mathcal{P}_\beta(\mathbb{R}^d; V)$ is a weakly compact subset of $\mathcal{P}(\mathbb{R}^d; V)$.*

Proof. Consider a sequence $(\nu_n, n \geq 1)$ of $\mathcal{P}_\beta(\mathbb{R}^d; V)$, such that ν_n converges weakly to a measure $\nu \in \mathcal{P}(\mathbb{R}^d; V)$. We have to show that $\int V(x)\nu(\mathrm{d}x) \leq \beta$. We have for all $x \in \mathbb{R}^d$, for all $R > 0$, that $R \wedge V(x) \leq V(x)$. We know, by hypothesis, that for all n, $\int R \wedge V(x)\nu_n(\mathrm{d}x) \leq \beta$. Moreover

$$\int R \wedge V(x)\nu_n(\mathrm{d}x) \longrightarrow \int R \wedge V(x)\nu(\mathrm{d}x).$$

But $R \mapsto R \wedge V$ is a non-decreasing function, converging to V, so that by the monotone convergence theorem, we deduce that $\int V(x)\nu(\mathrm{d}x) \leq \beta$ and $\nu \in \mathcal{P}_\beta(\mathbb{R}^d; V)$. Hence $\mathcal{P}_\beta(\mathbb{R}^d; V)$ is closed. It is also easily seen that all $\nu \in \mathcal{P}_\beta(\mathbb{R}^d; V)$ is a tight measure a.s. (see the preceding proof which implies that μ_t is a.s. tight). Since the space $\mathcal{P}_\beta(\mathbb{R}^d; V)$ is Polish, this space is relatively compact as a consequence of Prokhorov theorem. It asserts that $\mathcal{P}_\beta(\mathbb{R}^d; V)$ is a compact set. □

Proposition 7.2. *Let $\beta > 0$ such that $\mu_t \in \mathcal{P}_\beta(\mathbb{R}^d; V)$ for all $t \geq 0$. For all $n \in \mathbb{N}$, we have that $\mathbb{E}_{x,r,\mu}(V^n(X_t))$ is bounded.*

Proof. We do the case $n = 1$ and drop the subscripts x, r, μ in the following. We will prove that the proposition is true for the Lyapunov function $\mathcal{E}_\mu(x)$ instead of V.

We apply the Itô formula to prove that:

$$\begin{aligned} \mathrm{d}\mathcal{E}_{\mu_s}(X_s) &= (\nabla \mathcal{E}_{\mu_s}(X_s), \mathrm{d}B_s) - |\nabla \mathcal{E}_{\mu_s}(X_s)|^2 \mathrm{d}s + \frac{1}{2}\Delta \mathcal{E}_{\mu_s}(X_s)\mathrm{d}s \\ &+ (W(X_s, X_s) - W * \mu_s(X_s)) \frac{\mathrm{d}s}{r+s}. \end{aligned}$$

The strict convexity of $V + W_1 * \mu$ (which is uniform in μ because of the curvature assumption (3.3)) and the boundedness assumptions on W_2 and its derivatives lead to:

$$\forall \alpha > 0, \exists K_\alpha = K(\alpha, \beta, V);\ \mathcal{E}_{\mu_t}(X_t) \leq \alpha |\nabla \mathcal{E}_{\mu_t}(X_t)|^2 + K_\alpha.$$

From the domination condition (3.2) on W and the growth condition on V, we get that there exists $a > 0$ such that $\Delta \mathcal{E}_{\mu_t}(X_t) \leq a(1 + \mathcal{E}_{\mu_t}(X_t))$. These bounds lead for all $t \geq s \geq 0$ to

$$\begin{aligned} \mathbb{E}\mathcal{E}_{\mu_t}(X_t) &\leq \mathbb{E}\mathcal{E}_{\mu_s}(X_s) + \frac{1}{2\alpha}\int_s^t \left(K_\alpha - \mathbb{E}\mathcal{E}_{\mu_u}(X_u)\right) \mathrm{d}u + \frac{a}{2}\int_s^t (1 + \mathbb{E}\mathcal{E}_{\mu_u}(X_u))\mathrm{d}u \\ &+ \kappa \int_s^t \mathbb{E}V(X_u)(r+u)^{-1}\mathrm{d}u. \end{aligned}$$

Now, we can choose α such that $1/\alpha - a = 2a$ and we recall that $V(X_t) = O(t)$. Therefore the preceding inequality becomes (with M depending only on V and β)

$$\mathbb{E}\mathcal{E}_{\mu_t}(X_t) \leq \mathbb{E}\mathcal{E}_{\mu_s}(X_s) - a\int_s^t \mathbb{E}\mathcal{E}_{\mu_u}(X_u)\mathrm{d}u + M(t-s)$$

We divide both sides by $t - s$ and let $s \to t$. Let $x(t) := \mathbb{E}\mathcal{E}_{\mu_t}(X_t)$. Solving the preceding inequality boil down to solve $\dot{x} \leq M - ax$. The solution is $x(t) \leq \left(x(0) + M\int_0^t e^{as}\mathrm{d}s\right)e^{-at}$ and we finally obtain the following:

$$\mathbb{E}\mathcal{E}_{\mu_t}(X_t) \leq \left(KV(x) + \frac{M}{a}(e^{at} - 1)\right)e^{-at}$$

and we are done. □

7.2. Back to the dynamical system: a global attractor for the flow. We have defined in Section 5 a smooth dynamical system, with the strong topology. But, as we want to study a probabilistic objet, actually the asymptotic behavior of $(\mu_t, t \geq 0)$, it is natural to work with the weak* topology. That is the reason why we consider the following set: we denote by $L(\mu_t)$ the ω-limit set of μ_t that is $L(\mu_t) = \left\{\bigcap_{t\geq 0} \overline{\{\mu_s; s \geq t\}}\right\} \subset \mathcal{P}(\mathbb{R}^d; V)$[3]. We remark that $L(\mu_t)$ is composed by all the accumulation points of μ_t. This set is well-defined because it is invariant for the flow Φ which restriction to $L(\mu_t)$ does not explode in a finite time.

Definition 7.3. *The dynamical system associated to V and W is the mapping $\Psi : \mathbb{R} \times L(\mu_t) \to L(\mu_t)$, $(t, \mu) \mapsto \Psi_t(\mu)$ given by $\Psi_t(\mu) = \Phi_t(\mu)$ (where Φ is the flow defined in Section 5).*

[3]We emphasize that $L(\mu_t)$ is the ω-limit set of (μ_t) and not $(\Phi_t(\mu))$. In particular, $L(\mu_t)$ is random and depends on $(X_t, t \geq 0)$.

As we will never consider the first flow Φ in the following, by an abuse of notations (we hope it will not be confusing), we will denote the flow Ψ by Φ.

Remark 7.4. *It is easily proved that the function Π is continuous with respect to the weak* topology (because of the domination assumption on W and Π is just a composition with the exponential map). By use of the preceding results, we can show that the new mapping Φ satisfies the flow property and is continuous with respect to the weak* topology. For a detailed proof, we refer the reader to* [**6**] *(lemma 3.3).*

Our aim is now to find a global attracting set for Φ. A natural idea would be to prove that $\widehat{\mathrm{Im}(\Pi)}$ is suitable, but the problem that this space is quite large and therefore not necessarily compact. As a consequence, we decide to reduce to the non empty set $\widehat{\mathrm{Im}(\Pi)} \cap L(\mu_t)$.

Theorem 7.5. *$\widehat{\mathrm{Im}(\Pi)} \cap L(\mu_t)$ is a.s. a global attracting set for Φ.*

Proof. We begin to notice that $\widehat{\mathrm{Im}(\Pi)} \cap L(\mu_t)$ is a.s. compact for the weak* topology, by definition of $L(\mu_t)$ (it is a closed subset of $\mathcal{P}_\beta(\mathbb{R}^d)$).
The next step is to prove that for all $\mu \in \mathcal{P}_\beta(\mathbb{R}^d; V)$, we have that the distance $\mathrm{dist}(\Phi_t(\mu), \widehat{\mathrm{Im}(\Pi)} \cap L(\mu_t))$ converges to 0. But we already know that

$$\lim_{t\to\infty} \mathrm{dist}(\Phi_t(\mu), \widehat{\mathrm{Im}(\Pi)}) = 0.$$

Now because of the proposition 7.7, which asserts that $(\mu_{h(t)})$ is an asymptotic pseudo-trajectory for the flow Φ and thus $L(\mu_t)$ is an attractor free set (see corollary 7.13), we have that

$$\lim_{t\to\infty} \mathrm{dist}(\Phi_t(\mu), L(\mu_t)) = 0$$

and finally we are done because for two sets A, B, we have that $\mathrm{dist}(\Phi_t(\mu), A\cap B) \leq \mathrm{dist}(\Phi_t(\mu), A) + \mathrm{dist}(\Phi_t(\mu), B)$.
It is clear, by the definition of the set, that $\widehat{\mathrm{Im}(\Pi)} \cap L(\mu_t)$ is positively invariant. Finally, this set is a.s. a global attracting set for the flow. □

Corollary 7.6. *$L(\mu_t)$ is a.s. a subset of $\widehat{\mathrm{Im}(\Pi)}$.*

Proof. $\widehat{\mathrm{Im}(\Pi)} \cap L(\mu_t)$ is a global attracting set for Φ and is compact for the weak* topology. But $\emptyset$ and $L(\mu_t)$ are the only attractors of Φ. This due to the fact that $\mu_{h(t)}$ is an asymptotic pseudo-trajectory for the flow, which implies that $L(\mu_t)$ is attractor free. Therefore, $\widehat{\mathrm{Im}(\Pi)} \cap L(\mu_t) = L(\mu_t)$. Consequently, we find that $L(\mu_t) \subset \widehat{\mathrm{Im}(\Pi)}$. □

7.3. Asymptotic behavior. We define a family of measures $\{\varepsilon_{t,t+s}; t \geq 0, s \geq 0\}$ by

$$\varepsilon_{t,t+s} := \int_t^{t+s} (\delta_{X_{h(u)}} - \Pi(\mu_{h(u)}))\mathrm{d}u. \tag{7.1}$$

This family will play an important role in this section: it will be essential for proving that $t \mapsto \mu_{h(t)}$ is an asymptotic pseudo-trajectory for Φ.

PROPOSITION 7.7. *For all $f \in \mathcal{C}^\infty(\mathbb{R}^d; V)$ and every $T > 0$ there exists a positive constant K (depending only on V, W and the initial point x) such that for all $\delta > 0$*

$$\mathbb{P}_{x,r,\mu}\left(\sup_{0\leq s\leq T} |\varepsilon_{t,t+s} f| > \delta\right) \leq K\delta^{-2}e^{-t}||f||_V.$$

The proof will need the uniform estimates on the family of semi-groups (P_t^μ) proved in the last section.

PROOF. Let $f \in \mathcal{C}^\infty(\mathbb{R}^d; V)$. We begin to rewrite

$$\varepsilon_{t,t+s} f = \int_{h(t)}^{h(t+s)} A_{\mu_u} Q_{\mu_u} f \frac{\mathrm{d}u}{r+u}.$$

We consider the $\mathcal{C}^2$-valued process $(t,x) \mapsto Q_{\mu_{h(t)}} f(x)$, which is of class $\mathcal{C}^2$ and a $\mathcal{C}^1$-semi-martingale. Indeed it is easy to see that $t \mapsto \mu_{h(t)}$ is *a.s.* a bounded variation process with values in $\mathcal{M}(\mathbb{R}^d; V)$. Since we already know from the last section that $\mu \mapsto Q_\mu f$ is also $\mathcal{C}^1$, the claim follows by composition.

Therefore we can apply the generalized Itô formula to $(t,x) \mapsto h(t)^{-1} Q_{\mu_{h(t)}} f(x)$ and decompose the measure $\varepsilon_{t,t+s}$ in four parts (and we will control each term separately):

$$\varepsilon_{t,t+s} f = \varepsilon^1_{t,t+s} f + \varepsilon^2_{t,t+s} f + \varepsilon^3_{t,t+s} f + \varepsilon^4_{t,t+s} f$$

with

$$\begin{aligned}
\varepsilon^1_{t,t+s} f &= -\frac{1}{h(t+s)} Q_{\mu_{h(t+s)}} f(X_{h(t+s)}) + \frac{1}{h(t)} Q_{\mu_{h(t)}} f(X_{h(t)}) \\
\varepsilon^2_{t,t+s} f &= -\int_{h(t)}^{h(t+s)} Q_{\mu_u} f(X_u) \frac{\mathrm{d}u}{(r+u)^2} \\
\varepsilon^3_{t,t+s} f &= \int_{h(t)}^{h(t+s)} \frac{\partial}{\partial u} Q_{\mu_u} f(X_u) \frac{\mathrm{d}u}{r+u} \\
\varepsilon^4_{t,t+s} f &= M^f_{h(t+s)} - M^f_{h(t)}
\end{aligned}$$

where M^f_t is the local martingale $M^f_t := \int_0^t \nabla Q_{\mu_u} f(X_u) \frac{\mathrm{d}B_u}{r+u}$.

We recall the estimate proved in the last section: $\forall \varepsilon > 0, x \in \mathbb{R}^d$

$$|Q_{\mu_{h(t)}} f(X_{h(t)})| \leq ||f||_V (\varepsilon V(X_{h(t)}) + C(\varepsilon)) \;\; \forall f \in L^\infty(\Pi(\mu)).$$

and

$$|\nabla Q_{\mu_{h(t)}} f(X_{h(t)})| \leq K||f||_V (\varepsilon V(X_{h(t)}) + C(\varepsilon)).$$

We also remind that $\int_0^t V(X_s)\mathrm{d}s = O(t)$. Now, we can control each part of $\varepsilon_{t,t+s}$ and find for all $\varepsilon > 0$:

$$
\begin{aligned}
|\varepsilon^1_{t,t+s}f| &\leq h(t)^{-1}(|Q_{\mu_{h(t+s)}}f| + |Q_{\mu_{h(t)}}f|)\\
&\leq h(t)^{-1}||f||_V(\varepsilon(V(X_{h(t+s)}) + V(X_{h(t)})) + C(\varepsilon)) \xrightarrow[t\to\infty]{a.s.} 0;
\end{aligned}
$$

and

$$
\begin{aligned}
|\varepsilon^2_{t,t+s}f| &\leq \int_{h(t)}^{h(t+s)} ||Q_{\mu_u}f||_V V(X_u)\frac{\mathrm{d}u}{(r+u)^2}\\
&\leq Kh(t)^{-2}||f||_V \int_{h(t)}^{h(t+s)} V(X_u)\mathrm{d}u \xrightarrow{a.s.} 0.
\end{aligned}
$$

For the third part of $\varepsilon_{t,t+s}$, we will use the Markov inequality and the bound on the differential of Q_μ given in the preceding section:

$$
\begin{aligned}
\mathbb{P}\left(\sup_{0\leq s\leq T}|\varepsilon^3_{t,t+s}f| \geq \delta\right) &\leq \delta^{-2}\mathbb{E}|\varepsilon^3_{t,t+T}f|^2\\
&\leq \delta^{-2}\int_{h(t)}^{h(t+T)} \mathbb{E}|(DQ_{\mu_u}\cdot\mu_u)(f)(X_u)|^2\frac{\mathrm{d}u}{r+u}\\
&\leq C\delta^{-2}||f||_V^2\int_{h(t)}^{h(t+T)} \mathbb{E}\left[(\varepsilon V(X_u) + K(\varepsilon))^2\, V(X_u)\right]\frac{\mathrm{d}u}{(r+u)^2}
\end{aligned}
$$

We now recall that we have proved that for all $\varepsilon > 0$, for all $n \in \mathbb{N}$ and $t > 0$, we obtain $\mathbb{E}[V^n(X_t)] = o(t^\varepsilon)$. We use it to find that there exists some constant K (uniform in μ) such that

$$
\mathbb{P}\left(\sup_{0\leq s\leq T}|\varepsilon^3_{t,t+s}f| \geq \delta\right) \leq \frac{1}{\delta^2}K||f||_V^2 h(t)^{-1/2}.
$$

Since the quadratic variation of $M^f_{h(t+s)} - M^f_{h(t)}$ is bounded by $K||f||_V^2(\varepsilon + h(t)^{-1})$, Burkholder-Davis-Gundy inequality implies directly

$$
\mathbb{P}_{x,r,\mu}\left(\sup_{s\in[0,T]}|\varepsilon^4_{t,t+s}f| \geq \delta\right) \leq \frac{1}{\delta^2}K||f||_V^2(\varepsilon + h(t)^{-1}) \tag{7.2}
$$

□

Corollary 7.8. *For all $T > 0$ and all $f \in \mathcal{C}^\infty(\mathbb{R}^d; V)$, the following holds a.s.*

$$
\lim_{t\to\infty}\sup_{0\leq s\leq T}|\varepsilon_{t,t+s}f| = 0.
$$

Proof. Let $T > 0$ and $f \in \mathcal{C}^\infty(\mathbb{R}^d; V)$. We just need to prove that

$$
\lim_{t\to\infty}\sup_{0\leq s\leq T}|\varepsilon^4_{t,t+s}f| = 0.
$$

We recall that the quadratic variation of $M^f_{h(t+s)} - M^f_{h(t)}$ is bounded by $K||f||_V^2(\varepsilon+h(t)^{-1})$. We will use the well-known Borel-Cantelli lemma. First, for all $\varepsilon > 0$, we have by the Doob inequality added to the Burkholder-Davis-Gundy inequality that

$$\begin{aligned}\mathbb{P}_{x,r,\mu}\left(\sup_{n\leq t<n+1}\sup_{s\in[0,T]}|\varepsilon^4_{t,t+s}f|\geq\delta\right) &\leq \frac{1}{\delta^2}K\sup_{n\leq t<n+1}||f||_V^2(\varepsilon+h(t)^{-1})\\ &\leq \frac{1}{\delta^2}K||f||_V^2(\varepsilon+h(n)^{-1}).\end{aligned}$$

As it is true for all $\varepsilon > 0$, we deduce from the preceding inequality that

$$\mathbb{P}_{x,r,\mu}\left(\sup_{n\leq t<n+1}\sup_{0\leq s\leq T}|\varepsilon^4_{t,t+s}f|\geq\delta\right)\leq\frac{1}{\delta^2}K||f||_V^2h(n)^{-1}.$$

As we know that the sum $\sum_n h(n)^{-1}$ converges, an easy application of the Borel-Cantelli lemma permits us to conclude that a.s.

$$\lim_{n\to\infty}\sup_{n\leq t<n+1}\sup_{0\leq s\leq T}|\varepsilon^4_{t,t+s}f|=0$$

and the result follows. □

Lemma 7.9. *If for all $T > 0$, all $f \in \mathcal{C}^\infty(\mathbb{R}^d; V)$, it holds*

$$\lim_{t\to\infty}\sup_{0\leq s\leq T}|\varepsilon_{t,t+s}f|=0\ \ a.s.,$$

then the time-changed process, given by the function $\mathbb{R}_+ \to \mathcal{P}(\mathbb{R}^d; V)$, $t \mapsto \mu_{h(t)}$ is a.s. an asymptotic pseudo-trajectory for Φ (for the weak topology of measures).*

Proof. We have, for all $t, s \geq 0$,

$$\mu_{h(t+s)}-\Phi_s(\mu_{h(t)})=\int_0^s\big(F(\mu_{h(t+u)})-F(\Phi_u(\mu_{h(t)}))\big)\,\mathrm{d}u+\varepsilon_{t,t+s}.$$

Now for t large enough, we have already obtained that $\mu_{h(t)} \in \mathcal{P}_\beta(\mathbb{R}^d; V)$. Moreover, for all $T > 0$ and $0 \leq s \leq T$, there exists $\beta(T) \geq \beta$ (increasing with T) such that $\Phi_s(\mu_{h(t)}) \in \mathcal{P}_{\beta(T)}(\mathbb{R}^d; V)$. Let $f \in \mathcal{C}^\infty(\mathbb{R}^d; V)$. As for all $\mu \in \mathcal{P}_{\beta(T)}(\mathbb{R}^d; V)$, there exists a constant $C_\beta(T) > 0$ such that $\Pi(\mu)(\mathrm{d}x) \leq C_\beta(T)e^{-2W*\mu(x)}\gamma(\mathrm{d}x)$, the nonnegativity of W implies directly

$$0\leq\int_{\mathbb{R}^d}f(x)\Pi(\mu)(\mathrm{d}x)\leq C_\beta(T)\int_{\mathbb{R}^d}f(x)\gamma(\mathrm{d}x).$$

As a consequence, there exists a positive constant $C(\beta, T)$ such that

$$|\mu_{h(t+s)}f-\Phi_s(\mu_{h(t)})f|\leq\int_0^s|\mu_{h(t+u)}f-\Phi_u(\mu_{h(t)})f|\mathrm{d}u+sC(\beta,T)+|\varepsilon_{t,t+s}f|.$$

Now for any $T > 0$, the Gronwall lemma applied on $[0, T]$ then leads to:

$$|\mu_{h(t+s)}f-\Phi_s(\mu_{h(t)})f|\leq C(\beta,T)Te^T\sup_{0\leq s\leq T}|\varepsilon_{t,t+s}f|.$$

We take the supremum (for $0 \le s \le T$) of each side of the inequality and we are done:

$$\sup_{0\le s\le T} |\mu_{h(t+s)}f - \Phi_s(\mu_{h(t)})f| \le C(\beta, T)Te^T \sup_{0\le s\le T} |\varepsilon_{t,t+s}f|.$$

$\square$

THEOREM 2.2 *Under $\mathbb{P}_{x,r,\mu}$, the function $t \mapsto \mu_{h(t)}$ is almost surely an asymptotic pseudo-trajectory for Φ (for the weak* topology).*

PROOF. It suffices to combine the proposition 7.7 with the previous assertion linking the asymptotic pseudo-trajectory of the dynamical system Φ and $\varepsilon_{t,t+s}$. $\square$

COROLLARY 7.10. *([**6**]) Suppose that, for all $y \in \mathbb{R}^d$, V and $W(\cdot, y)$ are $\mathcal{C}^k(\mathbb{R}^d)$, for $k \ge 2$. Then $\mathbb{P}_{x,r,\mu}$-almost surely, every limit point of $(\mu_t, t \ge 0)$ has a $\mathcal{C}^k(\mathbb{R}^d)$ density with respect to the Lebesgue measure.*

PROOF. We have already proved that $\widehat{\mathrm{Im}(\Pi)} \cap L$ contains every subset of $\mathcal{P}(\mathbb{R}^d; V)$ negatively invariant under Φ and is a compact set for the weak* topology. The result is now just a consequence of the preceding theorem. $\square$

COROLLARY 7.11. $\mathbb{P}_{x,r,\mu}(\overline{\lim}_{t\to+\infty} |X_t| = +\infty) = 1$.

PROOF. Let A be a open subset of $\mathbb{R}^d$ such that $\gamma(A) > 0$. Since the measure γ is diffusive, we have that for all $\nu \in \widehat{\mathrm{Im}(\Pi)} \cap L(\mu_t)$, there exist $m, M > 0$ (independent of μ, $m \le \frac{e^{-W*\mu}}{Z(\mu)} \le M$) such that

$$m\gamma \le \nu \le M\gamma.$$

Now, if we consider a sequence $(\nu_{t_n}, n \ge 0)$ of $\mathcal{P}(\mathcal{P}(\mathbb{R}^d; V))$, the limits of its convergent subsequences will belong to $\widehat{\mathrm{Im}(\Pi)} \cap L(\mu_t)$, because the limit set of $\{\mu_t, t \ge 0\}$ is *a.s.* an attractor free set of Φ. Thus, there exists a subsequence $(\nu_{t_{n_k}})$ of (ν_{t_n}) such that $\nu_{t_{n_k}}$ converges almost surely to $\nu \in \widehat{\mathrm{Im}(\Pi)} \cap L(\mu_t)$ for the weak* topology. For all smooth function φ compactly supported, we have that

$$\nu_{t_n}(\varphi) \xrightarrow{w} \nu(\varphi).$$

If we consider φ such that it equals 1 on A and 0 out of a set B containing A, we find that $\nu(\varphi) \ge \nu(A) > 0$. Thus

$$\nu(B) \ge \limsup \nu_t(\varphi) \ge \liminf \nu_t(\varphi) \ge \nu(A) \ge m\gamma(A).$$

Therefore, it implies that

$$\int_0^{t_n} \delta_{X_s}(A)\mathrm{d}s \sim t_n m\gamma(A)$$

which in turn gives $\int_0^\infty \delta_{X_s}(A)\mathrm{d}s = \infty$ a.s. It implies that for all constant $K > 0$, $\int_0^\infty \delta_{X_s}(\mathbb{R}^d \setminus \overline{B}_K)\mathrm{d}s = \infty$ a.s., where $\overline{B}_K$ is the closed ball of radius K and so

$$\mathbb{P}_{x,r,\mu}\left(\bigcap_K \left\{ \int_0^\infty \mathrm{d}s \mathbb{1}_{\{|X_s| \geq K\}} = \infty \right\}\right) = 1.$$

We conclude that $\mathbb{P}_{x,r,\mu}\left(\overline{\lim_{t\to+\infty}} |X_t| = +\infty\right) = 1.$ □

THEOREM 7.12. *([**5**]) If the limit set of an asymptotic pseudo-trajectory is relatively compact, then this limit set is an attractor-free set.*

COROLLARY 7.13. *The limit set of $\{\mu_t(r,\mu)\}_{t\geq 0}$ is $\mathbb{P}_{x,r,\mu}$-almost surely an attractor free set of Φ.*

PROOF. This is a consequence of the two preceding theorems. □

PROPOSITION 7.14. *([**6**]) Let $L \subset \mathcal{P}_\beta(\mathbb{R}^d; V)$ be an attractor free set for Φ and $A \subset \mathcal{P}_\beta(\mathbb{R}^d; V)$ an attractor for Φ. If $L \cap B(A) \neq \emptyset$, then $L \subset A$. (Here $B(A)$ is the basin of attraction of A.)*

PROOF. If $L \cap B(A) \neq \emptyset$, the invariance of L makes $L \cap A$ a nonempty attractor for $\Phi|A$. Thus $L \subset A$. □

LEMMA 7.15. *The set of the fixed points of Π, $\{\mu \in \mathcal{P}_\beta(\mathbb{R}^d); F(\mu) = 0\}$, is a nonempty compact (for the weak* topology) subset of $\mathcal{P}_\beta(\mathbb{R}^d; V)$.*

PROOF. For the weak* topology, Π maps continuously the compact convex set $\mathcal{P}_\beta(\mathbb{R}^d; V)$ into a compact subset of $\mathcal{P}(\mathbb{R}^d; V)$. The Leray-Schauder fixed point theorem then applies and shows that the set $\{\mu \in \mathcal{P}_\beta(\mathbb{R}^d; V); F(\mu) = 0\}$ is nonempty. □

THEOREM 7.16. *(Tromba [**36**]) Let $\mathcal{B}$ be a $\mathcal{C}^\infty$ Banach manifold, F a $\mathcal{C}^\infty$ vector field on $\mathcal{B}$ and $\mathcal{E} : \mathcal{B} \to \mathbb{R}$ a $\mathcal{C}^\infty$ function. Assume that:*

(1) *$D\mathcal{E}(\mu) = 0$ if and only if $F(\mu) = 0$;*
(2) *$F^{-1}(0)$ is compact;*
(3) *for each $\mu \in F^{-1}(0)$, $D\mathcal{E}(\mu)$ is a Fredholm operator.*

Then $\mathcal{E}(F^{-1}(0))$ has an empty interior.

PROPOSITION 7.17. *([**4**], proposition 6.4) Let Λ be a compact invariant set for a flow Φ on a metric space E. Assume that there exists a continuous function $\mathcal{V} : E \to \mathbb{R}$ such that:*

(1) *$\mathcal{V}(\Phi_t(x)) < \mathcal{V}(x)$ for $x \in E\backslash\Lambda$ and $t > 0$;*
(2) *$\mathcal{V}(\Phi_t(x)) = \mathcal{V}(x)$ for $x \in \Lambda$ and $t > 0$.*

If $\mathcal{V}$ has empty interior, then every attractor free set K for Φ is contained in Λ. Furthermore, $\mathcal{V}$ restricted to K is constant.

THEOREM 4.6 *Suppose that W is symmetric. Then the limit set of $(\mu_t, t \geq 0)$ is $\mathbb{P}_{x,r,\mu}$-a.s. a compact connected subset of the fixed points of Π.*

PROOF. We follow the lines of [**8**] and work only with probability measures having a density with respect to the measure γ. The proof of this result relies on the preceding results. We want to use the preceding proposition with the Lyapunov function $\mathcal{E}$ (that is the free energy composed with Π), which satisfies the required condition. The preceding lemma shows that $\left\{\mu \in \mathcal{P}_\beta(\mathbb{R}^d; V); \Pi(\mu) = \mu\right\}$ is a non empty compact subset of $\mathcal{P}_\beta(\mathbb{R}^d; V)$. We still know that $F^{-1}(0)$ is compact for the weak* topology. Therefore, the only thing we have to show, thanks to the Tromba theorem, is that $\mathcal{E}(F^{-1}(0))$ has an empty interior. Let $\mu \in F^{-1}(0)$ and prove that $DF(\mu)$ is a Fredholm operator. Let us show that the operator $W : \mathcal{M}(\mathbb{R}^d; V) \to \mathcal{C}^0(\mathbb{R}^d; V), \mu \mapsto W * \mu(x)$ is compact. We recall the lemma 5.1 which asserts that

$$||W * \mu||_V \leq 2\kappa ||\mu||_V.$$

Moreover, we get for all $u, v \in \mathbb{R}^d$:

$$|W * \mu(u) - W * \mu(v)| \leq ||W(u, \cdot) - W(v, \cdot)||_V ||\mu||_V.$$

As a consequence, the set $\{W * \mu; ||\mu||_V \leq 1\}$ is bounded and equicontinuous. By the theorem of Ascoli, we conclude that the preceding set is relatively compact in $\mathcal{C}^0(\mathbb{R}^d; V)$ and thus the operator W is compact. By definition, its restriction to $L^2(\gamma)$ is also compact. As $D\Pi$ is a compact operator, the same holds for $DF(\mu)$. Moreover, this operator is self-adjoint (see [**?**], proposition 2.9). As a consequence, it follows from the spectral theory of compact self-adjoint operators that DF has at most countably many real eigenvalues ; the set of nonzero eigenvalues is either finite or can be ordered as $|\lambda_1| > |\lambda_2| > \ldots > 0$ with $\lim_{n\to\infty} \lambda_n = 0$. Therefore, we apply the result of Tromba which asserts that $\mathcal{E}(F^{-1}(0))$ has an empty interior. The conclusion is just an application of the proposition 7.17. □

COROLLARY 4.7 *Assume that W is symmetric. If the fixed point set of Π contains only finitely many isolated points, then μ_t converges almost surely.*

PROOF. Straightforward. □

8. Some ideas for diffusions in a Riemannian manifold

For the sake of simplicity, we have supposed the Brownian motion to live in $\mathbb{R}^d$, but we can also work in a Riemannian manifold. We just need to be more precise with our assumptions, by taking care of the Ricci curvature, and explain where and how we have to adapt the proofs.

Let M be a d-dimensional, connected complete smooth Riemannian manifold, with boundary ∂M empty. Denote by $\mathrm{d}x$ the Riemannian volume element and by Ric the Ricci curvature of the manifold. Let q be the squared Riemannian distance function from a given fixed point $o \in M$.

We will remind the reader some well-known notions of Riemannian geometry (see e.g. [**15**]). Let $U_o = \{v \in TM; q(v) > 1\}$.

Definition 8.1. *The cut locus of $o \in M$, denoted by cut(o), is the closure of the set containing all points $p \in M$ such that p has at least two shortest (straight line) segments to o. It means that we have cut$(o) = \exp_o(\partial U_o)$.*

Proposition 8.2. *cut(o) has zero measure and there exists a set U such that $M =$ cut$(o) \cup U$ where $U = \exp_o(U_o)$ and cut(o) are disjoint.*

Remark 8.3. *If we consider the function $x \mapsto q(x)$, then $\nabla q(x)$ is well defined for all $x \in M \setminus \{\text{cut}(o)\}$.*

8.1. Assumptions. We need here to replace all the assumptions given in the second paragraph by the following:

(1) *(regularity and positivity)* $V \in \mathcal{C}^2(M)$ and $W \in \mathcal{C}^2(M \times M)$, $V \geq 1$ and $W \geq 0$;
(2) *(convexity)* V is a strictly uniformly convex function (the constant of convexity is denoted by C);
(3) *(growth)* $\exists c > 0$, $\delta > 1$ such that $\forall x \in M$ we have

$$(\nabla V(x), \nabla q(x)) \geq cq(x)^\delta, \tag{8.1}$$

and there exists $C > 0$ such that for all $x, y \in M$ we have

$$|\nabla V(x) - \nabla V(y)| \leq C(|\nabla q(x) - \nabla q(y)| \wedge 1)(V(x) + V(y))$$

(4) *(domination)* there exists $\kappa > 0$ such that for all $x, y \in M$, $W(x, y) \leq \kappa V(x)$ and

$$|\nabla_x W(x, y)| + |\nabla^2_{xx} W(x, y)| \leq \kappa(V(x) + V(y)); \tag{8.2}$$

(5) *(curvature)* we suppose that we can decompose $W = W_1 + W_2$, where W_2, $\nabla_x W_2$ and $\nabla^2_{xx} W_2$ are three bounded functions in the variable x, and that there exists K such that $K + C/2 > 0$ and for all $y \in M$,

$$\text{Ric}(Y, Y) + (\nabla_Y \nabla(\frac{1}{2}V + W_1(\cdot, y)), Y) \geq K(Y, Y); \tag{8.3}$$

Of course, $(\cdot, \cdot)$, ∇ and Δ stand respectively for the Riemannian inner product, the associated gradient and Laplace-Beltrami operator on M.

Remark 8.4. *We emphasize that in a Riemannian manifold, we need one more condition including the Ricci curvature.*

Remark 8.5. *The construction of a standard Brownian motion on a Riemannian manifold is classical and uses the Stratonovich differential, see e.g.* [**7**] *V34.*

8.2. The diffusion. In a Riemannian manifold, the self-attracting diffusion has to be written under the following form:

$$\mathrm{d}X_t = \sum_{k=0}^{N} F_k(X_t) \circ \mathrm{d}B_t^k - (\nabla V(X_t) + \nabla W * \mu_t(X_t))\,\mathrm{d}t$$

where the integral is Stratonovich and $\sum_{k=0}^{N} F_k(F_k f) = \Delta f$

The only result we need to improve is that the corollary of Röckner & Wang is still valid. But we know that we just have to find a function α such that

$$A_\mu q(x) \leq c - \alpha(q(x))$$

for some $c > 0$ and all $x \in M \setminus \{\mathrm{cut}(o)\}$. The function α defined in the preceding section satisfies also this equation in the Riemannian case. This shows that we can easily extend all our result to a Riemannian manifold.

We also remind Myers theorem (see *e.g.* [**15**] p.162):

THEOREM 8.6. *(Myers) If (M, g) is a complete Riemannian manifold such that $Ric \geq (n-1)r^{-2}g$ where $r > 0$, then $diam(M, g) \leq diam(S^n(r))$. In particular, M is compact.*

Thus if M is a Riemannian manifold whose Ricci curvature is bounded from below by a positive constant, then M is compact. We then refer to [**6**] for examples.

9. Conclusion

We have obtained that the asymptotic behavior of $(\mu_t, t \geq 0)$ can be related to the dynamical system $\Phi_t(\mu)$. In a preceding section, we saw on an example that μ_t converges to a fixed point of Φ_t (a fixed point of Π). This raises two natural questions: in a compact space, Benaïm *& al.* showed that Π can have a continuum of fixed points and, more important, that μ_t can circle around; what happens in non compact spaces? does μ_t avoid traps? We manage to answer to the first question. In our setting, we deeply believe that Π can have only finitely many isolated fixed points. We also want to emphasize that when W is not symmetric, then it can happen that there exists no Lyapunov function and that the limit set of (μ_t) may be a non trivial orbit. Suppose for instance that (in dimension two) $W(x, y) = (x, Ry)$ where R is a rotation matrix and V is a polynomial. Then, depending on R (and V), one expects

- either the unique invariant set for the flow is γ and thus μ_t converges a.s. to γ;
- or μ_t converges a.s. to a random measure, related to the critical points of the free energy;
- or the limit set of μ_t is a periodic orbit related to the measure γ.

This example will be studied in details in a forthcoming paper.

Furthermore, we believe that our assumptions on the potentials are not optimal and may be weakened:

- we point out that we can not work without the confinement potential V (it is essential for proving the ultracontractivity of the semi-group P_t^μ). But we should lead the same study with a function $W * \mu$ strictly convex (uniformly in μ)...

- we should find the same kind of results without controlling from above W by V.

We also believe that it should be possible to prove that we have an asymptotic pseudo-trajectory for the strong topology of measures and not only for the weak topology. But to this aim, we need to compute the Burkholder-Davies-Gundy inequality for local martingales in the Banach space $(\mathcal{M}(\mathbb{R}^d; V), || \cdot ||_V)$.

Nevertheless, it should be possible to prove the convergence result of the last section under little different assumptions (for instance, supposing that V and W are two convex functions, one of them being uniformly convex and W is symmetric), together with obtaining the rate of convergence in some cases.

Bibliography

[1] Ané C., Blachère S., Chafaï D., Fougères P., Gentil Y., Malrieu F., Roberto C. *&* Scheffer G. [2001], *Sur les inégalités de Sobolev logarithmiques*, Panoramas et Synthèses 10, SMF.

[2] Bakry D. [1994], L'hypercontractivité et son utilisation en théorie des semigroupes, *Lectures on Prob. Th. and Stat., Ecole de Prob. de St-Flour*, Springer, 1-114.

[3] Bakry D. *&* Emery M. [1985], Diffusions hypercontractives, *Sém. Prob. XIX*, Springer LNM 1123, 177-206.

[4] Benaïm M. [1999], Dynamics of stochastic approximation algorithms, *Sém. Prob. XXXIII, Lecture Notes in Math.* **1709**, 1-68, Springer.

[5] Benaïm M. *&* Hirsch M.W. [1996], Asymptotic pseudotrajectories and chain reccurent flows, *J. Dynam. Diff. Equ.* **8**, 141-176.

[6] Benaïm M., Ledoux M. *&* Raimond O. [2002], Self-interacting diffusions, *Prob. Theory Relat. Fields* **122**, 1-41.

[7] Benaïm M. *&* Raimond O. [2003], Self-interacting diffusions II: convergence in law, *An. Inst. H. Poincaré, Prob. & Stat.* **39**(6), 1043-1055.

[8] Benaïm M. *&* Raimond O. [2005], Self-interacting diffusions III: symmetric interactions, *Ann. Prob.* **33**(5), 1716-1759.

[9] Bismut J.M. [1981], A generalized formula of Itô and some other properties of solution flows, *Z. Wahrsch. Verw. Gebiete* **55**, 331-350.

[10] Conley C.C. [1978], Isolated invariant sets and the Morse index, *CBMS Regional conference series in mathematics* **38**, Am. Math. Soc., Providence.

[11] Cranston M. *&* Le Jan Y. [1995], Self-attracting diffusions: two cases studies, *Math. Ann.* **303**, 87-93.

[12] Cranston M. *&* Mountford T.S. [1996], The strong law of large numbers for a Brownian polymer, *Ann. Prob.* **2**(3), 1300-1323.

[13] Davies E.B *&* Simon B. [1984], Ultracontractivity and the heat kernel for Schrödinger operators and Dirichlet Laplacians, *J. Func. An.* **59**, 335-395.

[14] Durrett R.T. *&* Rogers L.C.G. [1992], Asymptotic behavior of Brownian polymers, *Prob. Th. Rel. Fields* **92**(3), 337-349.

[15] Gallot S., Hulin D. *&* Lafontaine J. [2004], *Riemannian Geometry*, 3rd edition, Springer "Universitext".

[16] Herrmann S. *&* Roynette B. [2003], Boundedness and convergence of some self-attracting diffusions, *Math. Ann.* **325**(1), 81-96.

[17] Hirsch M.W. *&* Smale S. [1974], *Differential Equations, Dynamical Systems, and Linear Algebra*, Academic Press.

[18] Holley R. *&* Stroock D. [1987], Logarithmic Sobolev Inequalities and Stochastic Ising Models, *J. Stat. Phys.* **46**, 1159-1194.

[19] Kallenberg 0. [2001], *Foundations of Modern Probability*, 2nd edition, Probability and its Applications, Springer.

[20] Kontoyiannis I. *&* Meyn S.P. [2003], Spectral Theory and Limit Theory for Geometrically Ergodic Markov Processes, *Ann. App. Prob.* **13**, 304-362.

[21] Kunita H. [1990], *Stochastic flows and stochastic differential equations*, Cambridge studies in advanced mathematics.

[22] Ledoux M. [2000], The geometry of Markov diffusion generators, *Ann. Fac. Sci. Toulouse* **IX**, 305-366.

[23] McCann R. [1997], A convexity principle for interacting gases, *Adv. Math.* **128**, 153-179.

[24] Meyn S.P. *&* Tweedie R.L. [1993], *Markov Chains and Stochastic Stability*, Springer-Verlag.

[25] Mountford T.S. *&* Tarrès P. [2007], An asymptotic result for Brownian polymers, *Ann. Inst. H. Poincaré, Prob. Stat.*, to appear.

[26] Norris J.R., Rogers L.C.G. *&* Williams D. [1987], Self-avoiding random-walk: a Brownian motion model with local time drift, *Prob. Th. Rel. Fields* **74**, 2, 271-287.

[27] Norton D.E. [1995], The fundamental theorem of dynamical systems, *Comment. Math. Univ. Carolinae* **36**, 3, 585-597.

[28] Pemantle R. [2007], A survey of random processes with reinforcement, *Prob. Surveys* **4**, 1-79.

[29] Raimond O. [1997], Self-attracting diffusions: case of constant interaction, *Prob. Theory Relat. Fields* **107**, 177-196.

[30] Revuz D. *&* Yor M. [1998], *Continuous Martingales and Brownian Motion*, 3rd edition, Springer.

[31] Röckner M. *&* Wang F.Y. [2003], Supercontractivity and Ultracontractivity for (Nonsymmetric) Diffusions Semigroups on Manifolds, *Forum Math.* **15**, 893-921.

[32] Rogers L.C.G. *&* Williams D. [2000], *Diffusions, Markov processes and Martingales*, 2nd edition, vol. 2 "Itô Calculus", Cambridge Univ. Press.

[33] Rothaus O. [1981], Diffusions on compact Riemannian manifolds and logarithmic Sobolev inequalities, *J. Func. An.* **42**, 102-109.

[34] Tóth B. [1999], Self-interacting random motions - a survey, in P.Révész and B. Tóth editors, *Random Walks, Bolyai Soc. Math. St.* **9**, 349-384.

[35] Tóth B. *&* Werner W. [1998], The true self-repelling motion, *Prob. Theory Rel. Fields* **111**, 375-452.

[36] Tromba A.J. [1977], The Morse-Sard-Brown theorem for functionals and the problem of Plateau, *Amer. J. Math.* **99**, 1251-1256.

[37] Villani C. [2003], *Topics in Optimal Transportation*, Graduate Studies in Mathematics, Vol. 58, *AMS*.

CHAPITRE 5

Exemple de diffusion renforcée sur $\mathbb{R}^2$

Dans ce court chapitre, nous illustrons l'étude générale des diffusions renforcées précédente par un exemple. En effet, nous considérons un potentiel d'interaction $W(x,y) = (x, Ry)$ où R est une matrice de rotation sur $\mathbb{R}^2$. A travers l'étude complète de cet exemple, nous exhibons les différents comportements asymptotiques de la mesure d'occupation normalisée μ_t. En particulier, nous verrons que l'attraction peut être suffisamment forte pour forcer μ_t à tourner et alors son ensemble limite est un cercle de mesures.

1. Introduction

In this short paper, we will illustrate the previous study of self-interacting diffusions living in $\mathbb{R}^d$ (Kurtzmann [**4**]) with some examples in the two-dimensional case. The preceding paper contains abstract results, and therefore we describe here a simple example and illustrate some of our previous results. For the moment, let us recall in few words the problem we are dealing with. Consider a confinement potential $V : \mathbb{R}^2 \to \mathbb{R}_+$ and an interaction potential $W : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}_+$. For any Borel bounded measure μ, we consider the 'convoled' function

$$W * \mu : \mathbb{R}^2 \to \mathbb{R}, \;\; W * \mu(x) := \int_{\mathbb{R}^2} W(x,y)\mu(\mathrm{d}y).$$

Our main object of interest is the self-interacting diffusion solution to

$$\text{(1.1)} \qquad \begin{cases} \mathrm{d}X_t = \mathrm{d}B_t - (\nabla V(X_t) + \nabla W * \mu_t(X_t))\,\mathrm{d}t \\ \mathrm{d}\mu_t = (\delta_{X_t} - \mu_t)\frac{\mathrm{d}t}{r+t} \\ X_0 = x, \;\; \mu_0 = \mu \end{cases}$$

where (B_t) is a two-dimensional Brownian motion, and (μ_t) the normalized occupation measure of the process X, with initial weight $r > 0$ and initial probability measure μ. That is the sequence $(\mu_t; t \geq 0)$ defined by

$$\text{(1.2)} \qquad \mu_t := \frac{r\mu + \int_0^t \delta_{X_s}\mathrm{d}s}{r+t}.$$

We will analyze a simple class of self-interacting diffusions: the Euclidian scalar product $W(x,y) = (x, Ry)$ where R is a rotation matrix in $\mathbb{R}^2$, of angle θ, and $V(x) = V(|x|)$ is a polynomial. For these particular potentials, we manage to describe precisely the asymptotic behavior of the normalized occupation measure of X. In the applications, we will suppose for simplicity that $V(x) = a|x|^2 + b|x|^4$. We will see that the asymptotic behavior of μ_t depends crucially on V, and more precisely on the sign of $1 + cos\theta \int_0^\infty \rho^2\gamma(\rho)\mathrm{d}\rho$, where γ corresponds to the Gibbs probability measure $\gamma(\mathrm{d}x) := \frac{1}{Z}e^{-2V(x)}\mathrm{d}x$ (and Z is the normalization constant). We also introduce the function $H : \mathbb{R}_+ \to \mathbb{R}_+$ defined by

$$H(\alpha) := \int_0^\infty \mathrm{d}\rho\gamma(\rho) \int_0^{2\pi} \mathrm{d}v e^{-\alpha\rho\cos v}.$$

Of course, we deal with a probabilistic object and therefore the question is whether it converges for the weak convergence of measures (denoted by (w)) or not. Our main result is the following:

THEOREM 1.1. *Consider the self-interacting diffusion on $\mathbb{R}^2$ associated to V and $W(x,y) = (x, Ry)$, where R is the rotation matrix $R = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$, with $0 \leq \theta < 2\pi$. Then one of the following holds:*

(1) *If V is such that $1 + \int_0^\infty \gamma(\rho)\rho^2\mathrm{d}\rho\cos\theta > 0$, then with probability one $\mu_t \xrightarrow{(w)} \gamma$;*

(2) *If V is such that $1 + \int_0^\infty \gamma(\rho)\rho^2\mathrm{d}\rho\cos\theta \leq 0$, then we have two different cases:*

a) if $\theta = \pi$ then there exists a random variable $v \in \mathbb{S}^1$ such that a.s. $\mu_t \xrightarrow{(w)} \mu_\infty^v$ with

$$\mu_\infty^v(\mathrm{d}x) = \frac{e^{\alpha_1(x,v)}}{Z_1}\gamma(\mathrm{d}x),$$

where Z_1 is the normalization constant and α_1 is the unique positive solution to $-\alpha + 2\frac{H'(\alpha)}{H(\alpha)} = 0$,

b) else $\theta \neq \pi$, and then the limit set of (μ_t) equals $\{\nu(\delta), 0 \leq \delta < 2\pi\}$ with probability one, where $\nu(\delta) = \frac{1}{e^{T_\theta}-1}\int_0^{T_\theta} \mathrm{d}s e^s \mu_\infty^{v(s\tan\theta+\delta),\theta}$ with $T_\theta = 2\pi(\tan\theta)^{-1}$ and $\mu_\infty^{v,\theta}$ is the unique positive solution to $-\alpha + 2\cos\theta\frac{H'(\alpha)}{H(\alpha)} = 0$.

This theorem means that the probability measure γ has the same role as the Lebesgue measure usually plays. Therefore, if we suppose that $\int_0^\infty \mathrm{d}\rho\gamma(\rho)\rho^2\cos(\theta) > -1$, then the empirical measure behaves like the "Brownian motion" (constructed with respect to the measure γ).

On the contrary, if $\theta = \pi$ and $\int_0^\infty \mathrm{d}\rho\gamma(\rho)\rho^2 \geq 1$, then there is enough attraction for the convergence of μ_t (for the topology of weak convergence). Actually, we manage to counter the effect of the Brownian motion and the empirical occupation measure converges almost surely to a probability measure, which is approximatively a Gaussian distribution.

If we now suppose that $\theta \neq \pi$ and that there is enough attraction, then the term induced by θ forces μ_t to circle around and the limit set of (μ_t) is a circle of measures $\{\nu(\delta), 0 \leq \delta < 2\pi\}$.

The paper is organized in the following way. We begin to recall the notation and former results concerning self-interacting diffusions in non-compact spaces. Section 3 is devoted to the study of the particular example $W(x,y) = (x, Ry)$ and is divided in two parts. First, we study in details the rich case $R = -Id$ and after, we go further with the general case, where R is a rotation matrix. This study is illustrated with some numerical simulations.

2. Notation and background

Let $(\cdot,\cdot)$ and $|\cdot|$ denote the Euclidian scalar product and respectively the Euclidian norm in $\mathbb{R}^2$.

2.1. Assumptions on the potentials. We recall the technical assumptions on the potentials V and W:

i) (*regularity*) $V \in \mathcal{C}^2(\mathbb{R}^d)$ and $W \in \mathcal{C}^2(\mathbb{R}^d \times \mathbb{R}^d)$;

ii) (*convexity*) V is a strictly uniformly convex function, *i.e.* there exists $K > 0$ such that for all $x, \xi \in \mathbb{R}^d$: $(\nabla^2 V(x)\xi, \xi) \geq K|\xi|^2$;

iii) (*growth*) there exist $R, c > 0$, $\delta > 1$ such that for all $|x| \geq R$, $(\nabla V(x), x) \geq c|x|^{2\delta}$ and there exists $C > 0$ such that for all $x, y \in \mathbb{R}^d$ we have

$$|\nabla V(x) - \nabla V(y)| \leq C(|x-y| \wedge 1)(V(x) + V(y));$$

iv) (*domination*) there exists $\kappa > 0$ such that for all $x, y \in \mathbb{R}^d$,

$$W(x,y) + |\nabla_x W(x,y)| + |\nabla^2_{xx} W(x,y)| \leq \kappa \left(V(x) + V(y)\right);$$

v) (*curvature*) we suppose that we can decompose $W(x,y) = W_1(x,y) + W_2(x,y)$, where W_2 and its two first derivatives with respect to x are three bounded functions in the variable x and there exists $M > 0$ such that for all $x, y, \xi \in \mathbb{R}^d$ we have

$$\left((\nabla^2 V(x) + \nabla^2_{xx} W_1(x,y))\xi, \xi\right) \geq M|\xi|^2.$$

2.2. Measure space. As usual, we denote by $\mathcal{M}(\mathbb{R}^d)$ the space of signed (bounded) Borel measures on $\mathbb{R}^d$ and by $\mathcal{P}(\mathbb{R}^d)$ its subspace of probability measures. We will need the following measure space:

$$\mathcal{M}(\mathbb{R}^d; V) := \{\mu \in \mathcal{M}(\mathbb{R}^d); \int_{\mathbb{R}^d} V(y)|\mu|(\mathrm{d}y) < \infty\}, \tag{2.1}$$

where $|\mu|$ is the variation of μ (that is $|\mu| := \mu^+ + \mu^-$ with (μ^+, μ^-) the Hahn-Jordan decomposition of μ). This space will enable us to always check the integrability of V (and therefore of W and its derivatives thanks to the domination condition with respect to the (random) measures to be considered. We endow this space with the following dual weighted supremum norm (or dual V-norm) defined by

$$||\mu||_V := \sup_{\varphi; |\varphi| \leq V} \left| \int \varphi \mathrm{d}\mu \right|, \quad \mu \in \mathcal{M}(\mathbb{R}^d; V). \tag{2.2}$$

For $\beta > 0$, we similarly define the spaces

$$\mathcal{M}_\beta(\mathbb{R}^d; V) := \{\mu \in \mathcal{M}(\mathbb{R}^d); \int_{\mathbb{R}^d} V(y)|\mu|(\mathrm{d}y) \leq \beta\}, \tag{2.3}$$

and $\mathcal{P}_\beta(\mathbb{R}^d; V) := \mathcal{M}_\beta(\mathbb{R}^d; V) \cap \mathcal{P}(\mathbb{R}^d; V)$.

2.3. Some definitions. Now we need to refresh the reader's memory with a list of important definitions of the classical theory of dynamical systems. First we recall that $\mathcal{P}(\mathbb{R}^d; V)$ is a metrizable space and we can choose the following metric: since $\mathcal{C}^0(\mathbb{R}^d; V)$ is separable, we exhibit a sequence $(f_k)_k$ dense in $\{f \in \mathcal{C}^0(\mathbb{R}^d; V)/||f||_V \leq 1\}$, and set for all $\mu, \nu \in \mathcal{P}(\mathbb{R}^d; V)$:

$$\mathrm{dist}(\mu, \nu) := \sum_{k=1}^{\infty} 2^{-k} |\mu(f_k) - \nu(f_k)|. \tag{2.4}$$

Suppose we are given a deterministic flow Φ (we will define it precisely later).

DEFINITION 2.1. *a) A subset A of $\mathcal{P}(\mathbb{R}^d; V)$ is positively invariant (resp. negatively invariant, resp. invariant) for Φ provided $\Phi_t(A) \subset A$ (respectively $A \subset \Phi_t(A)$, respectively $\Phi_t(A) = A$) for all $t \geq 0$.*

b) A subset A of $\mathcal{P}(\mathbb{R}^d; V)$ is an attracting set (respectively attractor) for Φ provided:

1. *A is nonempty, compact for the weak* topology and positively invariant, (respectively invariant) and*
2. *A has a neighborhood $\mathcal{N} \subset \mathcal{P}(\mathbb{R}^d; V)$ such that $\mathrm{dist}(\Phi_t(\mu), A) \to 0$ as $t \to +\infty$ uniformly in $\mu \in \mathcal{N}$.*

c) The basin of attraction of an attractor $K \subset A$ for $\Phi|A = (\Phi_t|A)_t$ is the positively invariant open set (in A) comprising all points x such that $\mathrm{dist}(\Phi_t(x), A) \to 0$ as $t \to +\infty$, it means that it is the set of measures in the space $\mathcal{P}(\mathbb{R}^d; V)$ such that all initial conditions chosen in this set dynamically evolve to A

$$B(K, \Phi|A) = \{\mu \in A; \lim_{t\to\infty} \mathrm{dist}(\Phi_t(\mu), K) = 0\}.$$

d) A global attracting set (respectively global attractor) is an attracting set (respectively attractor) whose basin is the whole space $\mathcal{P}(\mathbb{R}^d; V)$.

e) Let A be a positively invariant set for Φ. An attractor for $\Phi|A = (\Phi_t|A)_t$ (defined by taking the restriction of (Φ_t) to A) is proper if it is different from $\emptyset$ and A.

f) An attractor-free set is a nonempty compact invariant set A such that $\Phi|A$ has no proper attractor.

2.4. Former results. In this section, we recall all the former results we have still obtained.

Let us consider a confinement potential $V : \mathbb{R}^d \to \mathbb{R}^+$ and an interaction potential $W : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^+$. For any Borel bounded measure μ, we consider the 'convoled' function

$$W * \mu : \mathbb{R}^d \to \mathbb{R}, \;\; W * \mu(x) := \int_{\mathbb{R}^d} W(x,y)\mu(\mathrm{d}y).$$

Our main object of interest is the self-interacting diffusion solution to

$$\left\{ \begin{array}{l} \mathrm{d}X_t = \mathrm{d}B_t - (\nabla V(X_t) + \nabla W * \mu_t(X_t))\,\mathrm{d}t \\ \mathrm{d}\mu_t = (\delta_{X_t} - \mu_t)\frac{\mathrm{d}t}{r+t} \\ X_0 = x, \;\; \mu_0 = \mu \end{array} \right. \tag{2.5}$$

where (B_t) is a d-dimensional Brownian motion, and (μ_t) the normalized occupation measure of the process with initial weight $r > 0$ and initial probability measure μ. Its means that the sequence $(\mu_t; t \geq 0)$ is defined by

$$\mu_t := \frac{r\mu + \int_0^t \delta_{X_s} \mathrm{d}s}{r+t}. \tag{2.6}$$

We note by γ the probability measure defined by

$$\gamma(\mathrm{d}x) := e^{-2V(x)}\mathrm{d}x,$$

and for a probability measure μ, we define the new probability measure

$$\Pi(\mu)(\mathrm{d}x) := \frac{1}{Z} e^{-2W*\mu(x)}\gamma(\mathrm{d}x) \tag{2.7}$$

where Z is the normalization constant.

The question is the asymptotic study of the empirical law of the process X. To this aim, we introduce a deterministic dynamical system $\Phi : \mathbb{R} \times \mathcal{M}(\mathbb{R}^d; V) \to \mathcal{M}(\mathbb{R}^d; V)$, which flow is defined by

$$\Phi_0(\mu) = \mu, \;\; \frac{\mathrm{d}}{\mathrm{d}t}\Phi_t(\mu) = \Pi(\Phi_t(\mu)) - \Phi_t(\mu). \tag{2.8}$$

Let us define the deterministic time-change $h(t) := r(e^t - 1)$. The key idea is to prove that the time-changed process $(\mu_{h(t)})$ is an *asymptotic pseudo-trajectory* of the dynamical system (2.8) that is for all $T > 0$

$$\lim_{t \to +\infty} \sup_{0 \leq s \leq T} |\mu_{h(t+s)} - \Phi_s(\mu_{h(t)}| = 0.$$

PROPOSITION 2.1. *There exists $\beta > 0$ such that for all $t \geq 0$, we have $\mu_t \in \mathcal{P}_\beta(\mathbb{R}^d; V)$.*

THEOREM 2.2. *Under $\mathbb{P}_{x,r,\mu}$, the function $t \mapsto \mu_{h(t)}$ is almost surely an asymptotic pseudo-trajectory for Φ (for the weak* topology).*

THEOREM 2.3. *The limit set of $(\mu_t; t \geq 0)$ is $\mathbb{P}_{x,r,\mu}$-almost surely an attractor free set of Φ.*

The next proposition comes from Benaïm *& al* [**2**] Proposition 3.9:

PROPOSITION 2.2. *Let $L \subset \mathcal{P}_\beta(\mathbb{R}^2; V)$ be an attractor free set for Φ and $A \subset \mathcal{P}_\beta(\mathbb{R}^2; V)$ an attractor for Φ. If $L \cap B(A) \neq \emptyset$, then $L \subset A$.*

3. Example

We suppose in the following that the dimension $d = 2$. We consider the function $W(x,y) := (x, Ry)$, where R is a matrix in $\mathbb{R}^2$, and $V(x) = V(|x|) := a|x|^4 + b|x|^2 + 1$. W and V obviously satisfy the conditions of the previous section. We denote by $(1,0)^T$ the transpose vector of $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

LEMMA 3.1. *We define $p = (1,0)^T$. For all continuous $\varphi : \mathbb{R} \to \mathbb{R}$, for all $y \in \mathbb{S}^1$ we have*

$$\begin{aligned} \int_{\mathbb{R}^2} \left(\varphi((x,y)) - \varphi((x,p))\right) \gamma(\mathrm{d}x) &= 0, \\ \int_{\mathbb{R}^2} \varphi((x,y))(x - (x,y)y)\gamma(\mathrm{d}x) &= 0. \end{aligned}$$

PROOF. We adapt the proof of [**2**]. We denote by $O(2)$ the orthogonal group of $\mathbb{R}^2$. For all $y \in \mathbb{S}^1$, there exists $g \in O(2)$ such that $y = gp$. We recall that $V(x) = V(|x|)$ and we therefore find by changing the variable that

$$\int_{\mathbb{R}^2} \varphi((x,y))\gamma(\mathrm{d}x) = \int_{\mathbb{R}^2} \varphi((g^{-1}x,p))\gamma(\mathrm{d}x) = \int_{\mathbb{R}^2} \varphi((x,p))\gamma(\mathrm{d}x).$$

We have proved the first equality. For the second one, let us define $\phi(y) := \int_{\mathbb{R}^2} \varphi((x,y))(x - (x,y)y)\gamma(\mathrm{d}x)$. We clearly have that $(\phi(y), y) = 0$ and for all $g \in O(2)$, we obtain by the invariance of γ under g that $\phi(gp) = g\phi(p)$. For each $h \in \{h \in O(2); hp = p\}$, we have $\phi(p) = h\phi(p)$. As a consequence, $\phi(p) = 0$ and thus $\phi(y) = 0$. $\square$

For all probability measure $\mu \in \mathcal{P}(\mathbb{R}^2; V)$, let $\bar{\mu}$ denote the mean of μ that is $\bar{\mu} := \int_{\mathbb{R}^2} x\mu(\mathrm{d}x)$. If we define the probability measure

$$\bar{\Pi}(\bar{\mu})(\mathrm{d}x) := \frac{e^{-2(x,R\bar{\mu})}}{Z(\bar{\mu})}\gamma(\mathrm{d}x), \tag{3.1}$$

then $\bar{\Pi}(\bar{\mu}) = \Pi(\mu)$. Let $\bar{F}(\mu) := -\mu + \int_{\mathbb{R}^2} x\bar{\Pi}(\mu)(\mathrm{d}x)$. It is readily shown that $\bar{\Phi}_t(\mu)$ satisfies the ODE

$$\frac{\mathrm{d}}{\mathrm{d}t}\bar{\mu} = \bar{F}(\bar{\mu}), \;\; \bar{\mu}_0 = \bar{\mu}. \tag{3.2}$$

LEMMA 3.2. *Let $m = \rho v$ with $\rho \geq 0$ and $v \in \mathbb{S}^1$. Then we get*

$$\int_{\mathbb{R}^2} x\bar{\Pi}(m)(\mathrm{d}x) = -\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}\rho}\log\left(\int_{\mathbb{R}^2} e^{-2\rho(x,v)}\gamma(\mathrm{d}x)\right)Rv.$$

PROOF. We have $m = \rho v$ with $\rho = |m| \in [0, +\infty[$ and $v = \frac{m}{|m|} \in \mathbb{S}^1$. We wonder if a separation of variables is possible, that is finding two functions f, g such that $\int_{\mathbb{R}^2} x\bar{\Pi}(m)(\mathrm{d}x) = f(\rho)g(v)$.

Let $v = gp$ with $g \in O(2)$, $p = (1,0)^T$ and $\alpha = 2\rho$. Because $V(x) = V(|x|)$, a pedestrian computation gives

$$\begin{aligned}\int_{\mathbb{R}^2} x\bar{\Pi}(m)(\mathrm{d}x) &= \frac{1}{\int_{\mathbb{R}^2} e^{-2(x,\rho Rv)}\gamma(\mathrm{d}x)} \int_{\mathbb{R}^2} xe^{-2(x,\rho Rv)}\gamma(\mathrm{d}x) \\ &= \frac{1}{\int_{\mathbb{R}^2} e^{-\alpha(x,v)}\gamma(\mathrm{d}x)} \int_{\mathbb{R}^2} Rxe^{-\alpha(x,v)}\gamma(\mathrm{d}x) \\ &= \frac{1}{\int_{\mathbb{R}^2} e^{-\alpha(x,v)}\gamma(\mathrm{d}x)} \int_{\mathbb{R}^2} (x,v)e^{-\alpha(x,v)}\gamma(\mathrm{d}x)Rv.\end{aligned}$$

Using the second equality of the lemma 3.1, we obtain

$$\int_{\mathbb{R}^2} x\bar{\Pi}(m)(\mathrm{d}x) = -\frac{\mathrm{d}}{\mathrm{d}\alpha}\log\left(\int_{\mathbb{R}^2} e^{-\alpha(x,v)}\gamma(\mathrm{d}x)\right) Rv$$

□

Let $m = \rho v$ be the solution to the ODE $\dot{m} = \bar{F}(m)$. We decompose $m = \rho v$ with $\rho = |m|$ and $v \in \mathbb{S}^1$. Then we have by the preceding lemma:

$$\frac{\mathrm{d}}{\mathrm{d}t}v = 0.$$

Moreover, if we let $\alpha = 2\rho$, then the preceding result implies that α satisfies the one-dimensional ODE

$$\frac{\mathrm{d}}{\mathrm{d}t}\alpha = F(\alpha) = -\alpha + 2\frac{\mathrm{d}}{\mathrm{d}\alpha}\log\left(\int_{\mathbb{R}^2} e^{-\alpha(x,Rp)}\gamma(\mathrm{d}x)\right). \tag{3.3}$$

Let us define some useful functions expressed in the polar coordinates:

$$H(\alpha) := \int_0^\infty \mathrm{d}\rho\gamma(\rho)\int_0^{2\pi} \mathrm{d}v e^{-\alpha\rho\cos v} \tag{3.4}$$

and

$$\tilde{H}'(\alpha) := \int_0^\infty \mathrm{d}\rho\gamma(\rho)\rho^2\int_0^{2\pi} \mathrm{d}v \sin^2 v e^{-\alpha\rho\cos v}. \tag{3.5}$$

3.1. The case $R = -Id$. Suppose that $R = -Id$. In fact, it means that W is a symmetric function. We express the problem in polar coordinates and we get $F(\alpha) = -\alpha\left(1 - 2\frac{\tilde{H}'(\alpha)}{H(\alpha)}\right)$.

PROPOSITION 3.1. *If $\int_0^\infty \rho^2\gamma(\rho)\mathrm{d}\rho \le 1$, then 0 is the unique equilibrium of (3.3) and 0 is stable. The basin of attraction of 0 is $\mathbb{R}_+$.*

If $\int_0^\infty \rho^2\gamma(\rho)\mathrm{d}\rho > 1$, then 0 is linearly unstable and there is another equilibrium α_1 for the equation (3.3) which is stable. Moreover, the basin of attraction of α_1 is $\mathbb{R}_+^*$.

Proof. We remark that the function F is $\mathcal{C}^\infty$. More precisely, a computation yields to

$$\begin{aligned} F'(\alpha) &= -1+2\frac{H''(\alpha)}{H(\alpha)}-2\left(\frac{H'(\alpha)}{H(\alpha)}\right)^2; \\ F''(\alpha) &= 2\frac{H^{(3)}(\alpha)}{H(\alpha)}-6\frac{H''(\alpha)}{H(\alpha)}\frac{H'(\alpha)}{H(\alpha)}+4\left(\frac{H'(\alpha)}{H(\alpha)}\right)^3; \\ F^{(3)}(\alpha) &= 2\frac{H^{(4)}(\alpha)}{H(\alpha)}-8\frac{H^{(3)}(\alpha)}{H(\alpha)}\frac{H'(\alpha)}{H(\alpha)}+24\frac{H''(\alpha)}{H(\alpha)}\left(\frac{H'(\alpha)}{H(\alpha)}\right)^2-12\left(\frac{H'(\alpha)}{H(\alpha)}\right)^4. \end{aligned}$$

We let the reader write the explicit expression of F (and its derivative) in terms of γ and $\cos v$. We wonder for the sign of $F^{(3)}$. We remark that $F^{(3)}$ corresponds to (twice) the kurtosis of the projection on the axis x of a random variable X (expressed in polar coordinates) such that X has the law γ. We need here to be careful! Actually, one can believe that, for a density function such that the graph of its asymptotic distribution is below the graph of the standard Gaussian, the kurtosis is negative. But this is false in general! For the assertion to be true, one need to ask for an other condition: the graph of the symmetric part of the density function has to cut exactly twice the graph of the Gaussian variable having the same mean and variance (see [**3**], p.95). But this is our case because of the assumption on V. Therefore, we get that the kurtosis of our variable is negative, that is $F^{(3)}(\alpha)<0$ for all $\alpha>0$ and $F^{(3)}(0)=0$. This means that the function $F''(\alpha)$ is non-increasing: for all $\alpha\geq 0$, $F''(\alpha)\leq F''(0)=0$. Similarly, we get

$$F'(\alpha)\leq F'(0)=-1+\int_0^\infty \mathrm{d}\rho\gamma(\rho)\rho^2.$$

Therefore, if $F'(0)\leq 0$, then F is non-increasing function and as $F(0)=0$, the first part of the result is proved.

Else we have $F'(0)>0$. But we remember that F' is a non-increasing function and $\lim_{\alpha\to\infty}F'(\alpha)=-1$. Thus (the continuity of F' implies that) there exists $\alpha_0>0$ such that $F'(\alpha_0)=0$. Moreover, we have that $\lim_{\alpha\to\infty}F(\alpha)=-\infty$. As a consequence, there exists a positive solution to $F(\alpha)=0$. Finally, we can conclude that there exists a positive solution to $F(\alpha)=0$ if and only if $\int_0^\infty \mathrm{d}\rho\gamma(\rho)\rho^2>1$. In that specific case, the point 0 is unstable and there exists an other equilibrium, which is stable. $\square$

Remark 3.3. *1) The positivity of $\int_0^\infty \mathrm{d}\rho\gamma(\rho)\rho^2-1$ relies completely on the coefficients of V. Actually, for $V(x)=10^{-3}|x|^2+10^{-4}|x|^4$, we obtain that the preceding expression is positive, whereas it is negative for $V(x)=|x|^2+|x|^4$.*
2) We recognize that the function $t\mapsto\int_0^{2\pi}e^{-t\cos v}\mathrm{d}v$ is the Bessel function $I_0(t)$.

LEMMA 3.4. *([**2**] corollary 3.10) Let (E,d) be a metric space, $\bar{\Phi}: E\times\mathbb{R}\to E$ a flow on E and $G:\mathcal{P}_\beta(\mathbb{R}^d;V)\to E$ a continuous function. Assume that $G\circ\Phi_t=\bar{\Phi}_t\circ G$. Let L denote the limit set of $\{\mu_t\}$. Then for almost all $\omega\in\Omega$, $G(L)$ is an attractor free set of $\bar{\Phi}$.*

PROOF. Let $\mu_{h(t)}$ be an asymptotic pseudo-trajectory for the flow generated by Φ. The compactness of $\mathcal{P}_\beta(\mathbb{R}^d;V)$ and the continuity of G imply that $G(\mu_{h(t)})$ is almost-surely an asymptotic pseudo-trajectory for $\bar{\Phi}$. Therefore, its limit set is an attractor free set for $\bar{\Phi}$ by the theorem 2.3. By compactness of $\mathcal{P}_\beta(\mathbb{R}^d;V)$ and continuity of G, this limit set is the image under G of the limit set of $\mu_{h(t)}$. $\square$

We need the following simplified version of the corollary 8.10 of [**1**].

LEMMA 3.5. *Let $\bar{\Phi}$ denote a smooth flow on $\mathbb{R}^2$. Let $A\subset\mathbb{R}^2$ be a compact submanifold invariant by $\bar{\Phi}$. Let $D\bar{\Phi}\big|_A(x)$ denote the derivative at x of $\bar{\Phi}\big|_A$. Let $\bar{\xi}:\mathbb{R}_+\to\mathbb{R}^2$ be a continuous function. Assume that*

(1) *there exists $\lambda<0$ such that for all $T\geq 0$*
$$\limsup_{t\to\infty}\frac{1}{t}\log\left(\sup_{0\leq s\leq T}|\bar{\xi}_{t+s}-\bar{\Phi}_s(\bar{\xi}_t)|\right)\leq\lambda;$$
(2) *the limit set of $\bar{\xi}$ is included in A;*
(3) *there is a neighborhood U of A which is attracted exponentially fast by A, that is there exists $\sigma<0$ such that (for $d(x,A)=\inf\{|x-a|;a\in A\}$),*
$$\limsup_{t\to\infty}\frac{1}{t}\log\left(\sup_{x\in U}\frac{d(\bar{\Phi}_t,A)}{d(x,A)}\right)\leq\sigma;$$
(4) *for $\mathcal{E}(\bar{\Phi}|_A):=\lim_{t\to\infty}\frac{1}{t}\log\left(\inf_{x\in A}|D\bar{\Phi}|_A(x)^{-1}|^{-1}\right)$, we have*
$$\sup(\sigma,\lambda)<\min(0,\mathcal{E}(\bar{\Phi}|_A)).$$

Then there exist $r\geq 0$ and $x\in A$ such that
$$\limsup_{t\to\infty}\frac{1}{t}\log|\bar{\xi}_t-\bar{\Phi}_{t+r}(x)|\leq\sigma.$$

REMARK 3.6. *The quantity $\mathcal{E}(\bar{\Phi}|_A)$ is called the expansion rate of $\bar{\Phi}|_A$.*

THEOREM 3.7. *Consider the self-interacting diffusion on $\mathbb{R}^2$ associated to V and W with $R=-Id$. Then we have two different cases:*

(1) *If $\int_0^\infty \mathrm{d}\rho\gamma(\rho)\rho^2\leq 1$, then a.s. $\mu_t\xrightarrow{(w)}\gamma$;*
(2) *If $\int_0^\infty \mathrm{d}\rho\gamma(\rho)\rho^2>1$, then there exists a random variable $v\in\mathbb{S}^1$ such that a.s. $\mu_t\xrightarrow{(w)}\mu_\infty^v$ with*
$$\mu_\infty^v(\mathrm{d}x)=\frac{e^{\alpha_1(x,v)}}{Z_1}\gamma(\mathrm{d}x),$$

where Z_1 is the normalization constant and α_1 is the unique positive solution to the equation $F(\alpha) = -\alpha + 2\frac{H'(\alpha)}{H(\alpha)} = 0$.

PROOF. Let $G : \mathcal{P}_\beta(\mathbb{R}^2; V) \to \mathbb{R}^2$ be the mapping defined by $G(\mu) = \bar{\mu}$. By the lemma 3.4, the limit set of $\bar{\mu}_t$ is a.s. an attractor free set of $\bar{\Phi}$. When $\int_0^\infty \mathrm{d}\rho\gamma(\rho)\rho^2 \leq 1$, then 0 is a global attractor for the dynamical system generated by $\bar{\Phi}$. Therefore, each attractor free set of $\bar{\Phi}$ reduces to 0. As a consequence, $\bar{\mu}_t \xrightarrow{(w)} 0$ almost surely and $L \subset G^{-1}(0)$. The definitions of $\bar{\Pi}(\bar{\mu})$ and F imply that $G^{-1}(0)$ is an invariant set under the action of Φ and

$$\Phi\big|_{G^{-1}(0)}(\mu) = e^{-t}(\mu - \gamma) + \gamma.$$

This last formula comes from the fact that $\Pi(\Phi_t\big|_{G^{-1}(0)}(\mu)) = \gamma$. Therefore, γ is a global attractor for $\Phi\big|_{G^{-1}(0)}$. The proposition 2.2 then implies that each attractor free set reduces to γ. Consequently, the theorem 3.4 enables us to conclude that $L = \gamma$. We are done for the first part of the theorem.

Suppose now that 0 is unstable for $\bar{F}$, where we recall that $\bar{F}(m) = -m + \int_{\mathbb{R}^2} x\bar{\Pi}(m)(\mathrm{d}x)$. It holds for all $f \in \mathcal{C}^\infty(\mathbb{R}^2; V)$

$$\frac{\mathrm{d}}{\mathrm{d}t}\mu_{h(t)}f = -\mu_{h(t)}f + \Pi(\mu_{h(t)})f + \frac{\mathrm{d}}{\mathrm{d}s}\varepsilon_{t,t+s}\big|_{s=0}f.$$

If we consider the projection map $f(x) = P_i(x) = x_i$, then we get $\frac{\mathrm{d}}{\mathrm{d}t}\bar{\mu}_{h(t)} = \bar{F}(\bar{\mu}_{h(t)}) + \eta_t$ where η_t is a random vector in $\mathbb{R}^2$ and more precisely $\eta_t = \frac{\mathrm{d}}{\mathrm{d}s}\varepsilon_{t,t+s}\big|_{s=0}(P_1, P_2)^T$. As 0 is an unstable linear equilibrium for $\bar{F}$, we apply the result of Tarrès ([**9**], part 3) to prove that

$$\mathbb{P}\left(\lim_{t\to\infty} \bar{\mu}_{h(t)} = 0\right) = 0.$$

We recall that F is a Lipschitz function (for the strong topology of measure), thus $\bar{F}$ is also Lipschitz (and continuous). We have for all $t, s \geq 0$

$$\bar{\mu}_{h(t+s)} - \bar{\Phi}_s(\bar{\mu}_{h(t)}) = \int_0^s (\bar{F}(\bar{\mu}_{h(t+u)}) - \bar{F}(\bar{\Phi}_u(\bar{\mu}_{h(t)}))\mathrm{d}u + \eta_{t+s}.$$

But for t large enough, we have already obtained that $\mu_{h(t)} \in \mathcal{P}_\beta(\mathbb{R}^2; V)$. Moreover, for all $T > 0$ and $0 \leq s \leq T$, there exists $\beta(T) \geq \beta$ (increasing with T) such that $\Phi_s(\mu_{h(t)}) \in \mathcal{P}_\beta(\mathbb{R}^2; V)$. As the restriction of F to $\mathcal{P}_{\beta(T)}(\mathbb{R}^2; V)$ is Lipschitz, there exists $K(T)$ such that

$$|\bar{\mu}_{h(t+s)} - \bar{\Phi}_s(\bar{\mu}_{h(t)})| \leq K(T)\int_0^s |\bar{\mu}_{h(t+u)} - \bar{\Phi}_u(\bar{\mu}_{h(t)})|\mathrm{d}u + |\eta_{t+s}|.$$

Thus for any $T > 0$, the Gronwall lemma applied on $[0, T]$ then leads to

$$|\bar{\mu}_{h(t+s)} - \bar{\Phi}_s(\bar{\mu}_{h(t)})| \leq e^{K(T)s} \sup_{0\leq s\leq T} \max(|\varepsilon_{t,t+s}P_1|, |\varepsilon_{t,t+s}P_2|).$$

We take the supremum (for $0 \leq s \leq T$) of each side of the inequality and we obtain

$$\sup_{0\leq s\leq T} |\bar{\mu}_{h(t+s)} - \bar{\Phi}_s(\bar{\mu}_{h(t)})| \leq e^{K(T)T} \sup_{0\leq s\leq T} \max(|\varepsilon_{t,t+s}P_1|, |\varepsilon_{t,t+s}P_2|).$$

Therefore, we find that

$$\limsup_{t\to\infty} \frac{1}{t} \log\left(\sup_{0\leq s\leq T} |\bar{\mu}_{h(t+s)} - \bar{\Phi}_s(\bar{\mu}_{h(t)})| \leq -\frac{1}{2}\right).$$

We remind that

$$\partial_t \rho = -\rho - \frac{H'(\alpha)}{H(\alpha)} \tag{3.6}$$

and we denote by $\alpha_1(\gamma)$ the unique positive solution to $-\alpha + 2\frac{H'(\alpha)}{H(\alpha)} = 0$. We introduce the invariant set (for the flow $\bar{\Phi}$)

$$A := \{m = \rho v; \rho = \alpha_1(\gamma), v \in \mathbb{S}^1\}.$$

It just remains to use the lemma 3.5 for the flow $\bar{\Phi}$ induced by $\bar{F}$.

- The first assertion is satisfied for $\lambda = -1/2$.
- The limit set of $\bar{\mu}_{h(t)}$ being an attractor free set by the theorem 2.2, the ODE (3.6) satisfied by ρ implies that this limit set either reduces to $\{0\}$, or is included in A. But we know that $\mathbb{P}\left(\lim_{t\to\infty} \bar{\mu}_{h(t)} = 0\right) = 0$ and thus the limit set of $\bar{\mu}_{h(t)}$ is a subset of A with probability one. The second assertion follows.
- Moreover, α_1 is stable for the equation (3.6), thus $F'(\alpha_1) < 0$ and the set A attracts a neighborhood of itself at any exponential rate $F'(\alpha_1) < \sigma < 0$. We are done for the third assertion.
- It remains to prove the last point of the lemma 3.5. We recall that $\partial_t v = 0$. As a consequence, it is clear that $\bar{\Phi}_t\big|_A = Id\big|_A$ and $\mathcal{E}(\bar{\Phi}_t\big|_A) = 0$.

Finally, the lemma 3.5 implies that there exists $v \in \mathbb{S}^1$ such that

$$\limsup_{t\to\infty} \frac{1}{t} \log\left(|\bar{\mu}_{h(t)} - \alpha_1 v|\right) \leq \max(F'(\alpha_1), -1/2).$$

To conclude, we have on one side that the limit set of (μ_t) is an attractor free set of $\Phi\big|_{G^{-1}(\alpha_1 v)}$ and on the other side, the flow $\Phi\big|_{G^{-1}(\alpha_1 v)}$ admits μ_∞^v as a global attractor. This leads to the announced result: $L(\mu_t) = \mu_\infty^v$. $\square$

The case $R = Id$. We suppose here that $W(x,y) = (x,y)$. In particular, we note that W is a symmetric function. In that special case, we use the polar coordinates and we get $F(\alpha) = -\alpha\left(1 + 2\frac{\tilde{H}'(\alpha)}{H(\alpha)}\right)$. It is obvious that $F(\alpha) \leq 0$ for all $\alpha \geq 0$ and $F(\alpha) = 0$ if and only if $\alpha = 0$. Therefore, 0 is the unique equilibrium of (3.3) and 0 is stable. The basin of attraction of 0 is

$\mathbb{R}_+$. As consequence, we easily get that with probability 1, $\mu_t \xrightarrow[t\to\infty]{(w)} \gamma$. That is why this case is not very interesting in comparison with the previous one.

Here is an illustration of the phenomenon: we have chosen the initial point $x_0 = (1, -2)$, the initial weight $r = 0.4$ and the initial probability measure $\mu_0(\mathrm{d}x) = (2\pi)^{-1/2} e^{-\frac{(x-1)^2}{2}} \mathrm{d}x$. Now, let us show with a simulation what happens for $W(x, y) = 2(x, y)$ and $V(x) = |x|^4 + |x|^2 + 1$. $(T = 510)$

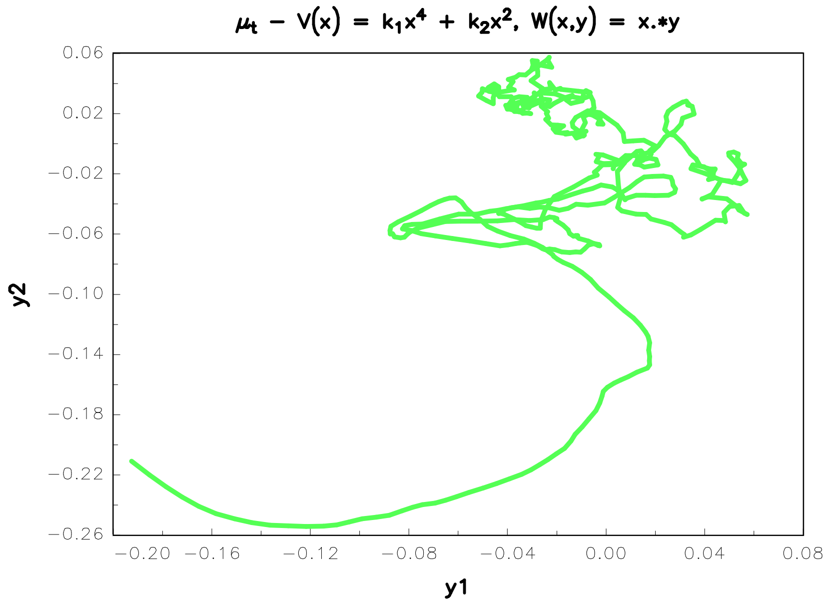

FIGURE 1. The convergence of μ_t toward γ.

3.2. The case "R is a rotation". We assume here that R is a rotation, that is $R \in O(2)$ and $R = R(\theta)$ is defined as $R = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}$, with $0 \le \theta < 2\pi$. We emphasize that (else if $\theta = 0, \pi$) W is *not* a symmetric function.

THEOREM 3.8. *Consider the self-interacting diffusion on $\mathbb{R}^2$ associated to V and $W(x, y) = (x, Ry)$. Then one of the following holds:*

(1) *If V is such that $\int_0^\infty \mathrm{d}\rho\gamma(\rho)\rho^2\cos(\theta) > -1$, then with probability one $\mu_t \xrightarrow{(w)} \gamma$;*

(2) *If V is such that $\int_0^\infty \mathrm{d}\rho\gamma(\rho)\rho^2\cos(\theta) \le -1$, then we get two different cases:*
a) if $\theta = \pi$ then there exists a random variable $v \in \mathbb{S}^1$ such that a.s. $\mu_t \xrightarrow{(w)} \mu_\infty^v$ with

$$\mu_\infty^v(\mathrm{d}x) = \frac{e^{\alpha_1(x,v)}}{Z_1}\gamma(\mathrm{d}x),$$

where Z_1 is the normalization constant and α_1 is the unique positive solution to $-\alpha + 2\frac{H'(\alpha)}{H(\alpha)} = 0$,

b) if $\theta \neq \pi$, then the ω-limit set of (μ_t) equals $\{\nu(\delta), 0 \le \delta < 2\pi\}$ with probability one, where

$$\nu(\delta) = \frac{1}{e^{T_\theta} - 1}\int_0^{T_\theta} \mathrm{d}s e^s \mu_\infty^{v,\theta},$$

with $T_\theta = 2\pi(\tan\theta)^{-1}$ and $\mu_\infty^{v,\theta}$ is the unique positive solution to $-\alpha + 2\cos\theta\frac{H'(\alpha)}{H(\alpha)} = 0$.

Proof. We recall the second equality of the lemma 3.1: for all $v \in \mathbb{S}^1$, $v = gp$ ($g \in O(2)$) it holds by the invariance of γ by Rg

$$\frac{\int_{\mathbb{R}^2} xe^{-\alpha(x,Rv)}\gamma(\mathrm{d}x)}{\int_{\mathbb{R}^2} e^{-\alpha(x,Rv)}\gamma(\mathrm{d}x)} = \frac{\int_{\mathbb{R}^2}(x,Rv)Rve^{-\alpha(x,Rv)}\gamma(\mathrm{d}x)}{\int_{\mathbb{R}^2} e^{-\alpha(x,Rv)}\gamma(\mathrm{d}x)} = -\frac{H'(\alpha)}{H(\alpha)}Rv.$$

Let $v = gp$ with $g \in O(2)$ and $m = \rho v = \alpha v/2$. We remind the equations

$$\begin{aligned}
\frac{\mathrm{d}m}{\mathrm{d}t} &= \bar{F}(m) = -m + \bar{\Pi}(m) = \rho\partial_t v + v\partial_t\rho;\\
\frac{\mathrm{d}\alpha}{\mathrm{d}t} &= -\alpha - 2\frac{H'(\alpha)}{H(\alpha)}(Rv,v);\\
\frac{\mathrm{d}v}{dt} &= -\frac{2}{\alpha}\frac{H'(\alpha)}{H(\alpha)}\left((Rv,v)v - Rv\right).
\end{aligned}$$

But by definition of R and $v = (v_1,v_2)^T = (\cos\sigma,\sin\sigma)^T$, we have $Rv = (v_1\cos\theta + v_2\sin\theta, -v_1\sin\theta + v_2\cos\theta)^T$ and therefore we obtain $(Rv,v) = \cos\theta$. A simple computation yields to the vector

$$(Rv,v)v - Rv = \begin{pmatrix} -\sin\theta\sin\sigma \\ \sin\theta\cos\sigma \end{pmatrix}.$$

We finally get after some easy calculations

$$\left\{\begin{array}{rcl} \frac{\mathrm{d}\alpha}{\mathrm{d}t} &=& -\alpha - \frac{2H'(\alpha)}{H(\alpha)}\cos\theta; \\ \frac{\mathrm{d}\sigma}{\mathrm{d}t} &=& \frac{2H'(\alpha)}{\alpha H(\alpha)}\sin\theta. \end{array}\right. \tag{3.7}$$

We recall (see the proof of the theorem 3.7) that $\frac{H'(\alpha)}{H(\alpha)} > 0$ for $\alpha > 0$. It just remains to use the proposition 3.1. We know that we have a bifurcation at $\cos\theta \int_0^\infty \gamma(\mathrm{d}\rho)\rho^2 = 1$. More precisely:

- If $\cos\theta \int_0^\infty \gamma(\mathrm{d}\rho)\rho^2 \geq 1$, we find a global attracting set for (3.7), that is the set $\{(\sigma,\alpha);\alpha = 0\}$ and therefore, adapting the proof of the previous theorem 3.7, we obtain a.s. $\mu_t \xrightarrow{(w)} \gamma$;
- If $\cos\theta \int_0^\infty \gamma(\mathrm{d}\rho)\rho^2 < 1$, then the set $\{(\sigma,\alpha);\alpha = \alpha_1(\cos\theta)\}$ is a global attracting set for (3.7). On this set, the dynamics is given by
$$\frac{\mathrm{d}\sigma}{\mathrm{d}t} = \frac{2H'(\alpha_1(\cos\theta))}{\alpha_1(\cos\theta)H(\alpha_1(\cos\theta))}\sin\theta = \tan\theta.$$

We mimic the proof of the theorem 3.7, by using the lemma 3.5, and show that there exists a constant $\lambda < 0$ and a random variable σ_0 such that with probability 1
$$\limsup_{t\to\infty}\frac{1}{t}\log\left(\left|\bar{\mu}_{h(t)} - \frac{\alpha_1(\cos\theta)}{2}v(t\tan\theta + \sigma_0)\right|\right) < \lambda.$$
For the moment, we know the dynamics on the set $\tilde{A} := \{(\sigma,\alpha);\alpha = \alpha_1(\cos\theta)\}$. Unfortunately, we need to more than that in order to finish the study of our dynamical system. That is the reason why we study the coupled system defined on $\mathcal{M}(\mathbb{R}^2;V)\times\mathbb{R}^2$ by

$$\left\{\begin{array}{l} \frac{\mathrm{d}m}{\mathrm{d}t} = \bar{F}(m) = -m + \bar{\Pi}(m);\\ \frac{\mathrm{d}\nu}{\mathrm{d}t} = -\nu + \bar{\Pi}(m). \end{array}\right. \tag{3.8}$$

The set $L(\mu_t)\times\tilde{A}$ (where $L(\mu_t)$ denotes the limit set of μ_t) is an attractor free set, by the theorem 2.2, for the flow induced by the dynamical system 3.8 restricted to $\mathcal{P}(\mathbb{R}^2;V)\times\mathbb{R}^2$. The dynamics on $L(\mu_t)\times\tilde{A}$ is given by

$$\left\{\begin{array}{l} \frac{\mathrm{d}\sigma}{\mathrm{d}t} = \tan\theta;\\ \frac{\mathrm{d}\nu}{\mathrm{d}t} = -\nu + f(\sigma) = -\nu + \mu_\infty^{v,\cos\theta}. \end{array}\right. \tag{3.9}$$

We remark that f is a 2π-periodic function. We can compute the explicit solution to (3.9), which is given by

$$\left\{\begin{array}{l} \nu_t = e^{-t}\left(\nu_0 + \int_0^t e^s f(s\tan\theta + \sigma_0)\mathrm{d}s\right);\\ \sigma_t = \sigma_0 + t\tan\theta. \end{array}\right. \tag{3.10}$$

For the rest of the proof, it remains to adapt the last part of the proof of [**2**] (theorem 4.11). Let us define $f_\sigma(s) := f(s+\sigma)$ and $T_\theta := 2\pi(\tan\theta)^{-1}$. We get

$$\begin{aligned} \nu_{t+T_\theta} &= e^{-(t+T_\theta)}\left(\nu_0 + \int_0^{t+T_\theta} e^s f_{\sigma_0}(s\tan\theta)\mathrm{d}s\right)\\ &= e^{-T_\theta}\left(\nu_t + \int_0^{T_\theta} e^s f_{\sigma_0}(s\tan\theta)\mathrm{d}s\right). \end{aligned}$$

Let $\nu^\theta(\sigma_0) = \frac{1}{e^{T_\theta}-1}\int_0^{T_\theta} e^s f_{\sigma_0}(s\tan\theta)\mathrm{d}s$. We get for all $t \in \mathbb{R}$:

$$\nu_{t+T_\theta} - \nu^\theta(\sigma_0) = e^{-T_\theta}(\nu_t - \nu^\theta(\sigma_0))$$

and thus for all $n \in \mathbb{N}$ we obtain $\nu_{-nT_\theta} - \nu^\theta(\sigma_0) = e^{-nT_\theta}(\nu_0 - \nu^\theta(\sigma_0))$. Suppose now that $(\nu_0, \frac{1}{2}\alpha_1(\cos\theta)v(\tan\sigma_0))$ belongs to the set $L(\mu_t) \times \tilde{A}$. As the set $L(\mu_t) \times \tilde{A}$ is compact (for the weak* topology) and invariant in $\mathcal{P}(\mathbb{R}^2; V) \times \mathbb{R}^2$, then ν_t is a probability measure for all $t \in \mathbb{R}$. Thus we have $\nu_0 = \nu^\theta(\sigma_0)$ and ν_t is T_θ-periodic. For all $0 \le t \le T_\theta$, we get after some easy calculations

$$\begin{aligned} \nu_t &= e^{-t}\left(\nu(\sigma_0) + \int_0^t e^s f_{\sigma_0}(s\tan\theta)\mathrm{d}s\right) \\ &= \nu(t\tan\theta + \sigma_0). \end{aligned}$$

To conclude, we have with probability one

$$L(\mu_t) = \{\nu(\sigma); 0 \le \sigma < 2\pi\}.$$

As a consequence, there exists some continuous function $\chi : \mathbb{R}_+ \to \mathbb{R}$ such that $\lim_{t\to\infty} \mathrm{dist}(\mu_t, \nu(\chi_t)) = 0$. As the application $G : \mathcal{P}(\mathbb{R}^2; V) \to \mathbb{R}^2$ is uniformly continuous, we get

$$\lim_{t\to\infty} \left|\bar{\mu}_t - \frac{\alpha_1(\cos\theta)}{2} v(\chi_t)\right| = 0.$$

But we recall that there exists $\lambda < 0$ such that

$$\limsup_{t\to\infty} \log\left|\bar{\mu}_{h(t)} - \frac{\alpha_1(\cos\theta)}{2} v(t\tan\theta + \sigma_0)\right| < \lambda.$$

As a consequence, we have the wanted result:

$$\lim_{t\to\infty} |v(\chi_t) - v(h^{-1}(t)\tan\theta + \sigma_0)| = 0.$$

□

Here is an illustration of the preceding theorem. We will see the three asymptotic behaviors of μ_t. For all the figures, we have chosen the initial point $x_0 = (1, -2)$, the initial weight $r = 0.4$ and the initial probability measure $\mu_0 = (2\pi)^{-1/2} e^{-\frac{(x-1)^2}{2}}$. The following illustration corresponds to the case $W(x,y) = 8(x, Ry)$, with $\theta = 3\pi/4$ and $V(x) = 0.002x^4 + 0.001x^2 + 1$. $(T = 542)$

The last figure corresponds to the case $W(x,y) = 15(x, Ry)$, with $\theta = \pi$ and $V(x) = 0.02|x|^4 + 0.2|x|^2 + 1$. $(T = 417)$

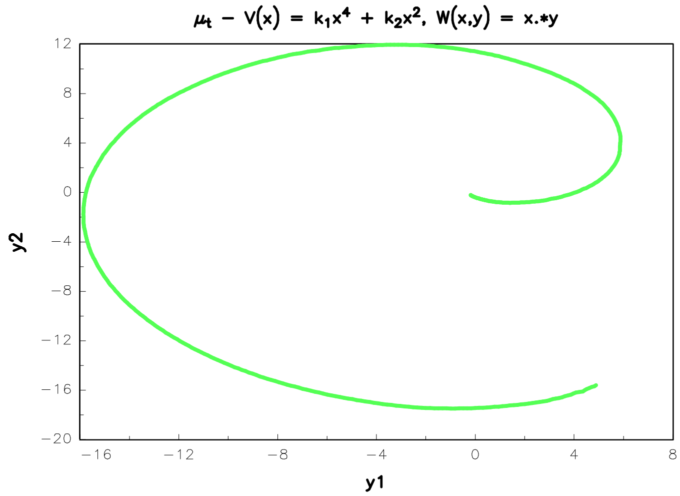

FIGURE 2. Oscillations: μ_t circles around.

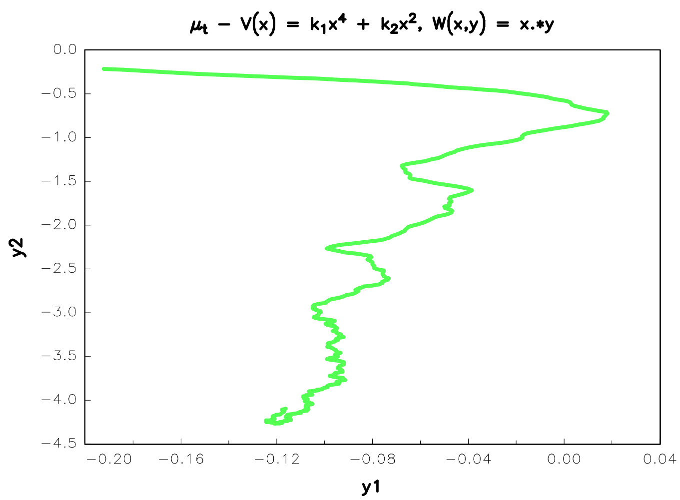

FIGURE 3. Convergence of μ_t toward the positive fixed point of F.

Bibliography

[1] BENAÏM M. [1999], Dynamics of stochastic approximation algorithms, *Sém. Prob. XXXIII, Lecture Notes in Math.* **1709**, 1-68, Springer.

[2] BENAÏM M., LEDOUX M. & RAIMOND O. [2002], Self-interacting diffusions, *Prob. Theory Relat. Fields* **122**, 1-41.

[3] KENDALL M. & STUART A. [1977], *The advanced theory of statistics. Vol. 1*, 4th ed., Macmillan Publishing Co., Inc., New York.

[4] KURTZMANN A. [2007], The ODE method for some self-interacting diffusions on non-compact spaces, preprint.

[5] PEMANTLE R. [2007], A survey of random processes with reinforcement, *Prob. Surveys* **4**, 1-79.

[6] REVUZ D. & YOR M. [1998], *Continuous Martingales and Brownian Motion*, 3rd edition, Springer.

[7] ROGERS L.C.G. & WILLIAMS D. [2000], *Diffusions, Markov processes and Martingales*, 2nd edition, vol. 2 "Itô Calculus", Cambridge Univ. Press.

[8] TARRÈS P. [2000], Pièges répulsifs, *Comptes Rendus Ac. Sc.* **330**, Série I, 125-130.

[9] TARRÈS P. [2001], *Traps of stochastic algorithms and vertex-reinforced random walks*, Doctoral dissertation, Ecole Normale Supérieure de Cachan.

ANNEXE A

Diffusions attirées par leur maximum

Le titre initial de cette note est “A note on some diffusions attracted by their maximum”. Elle est écrite en anglais.

We study the diffusion given by the stochastic differential equation

$$\begin{cases} \mathrm{d}X_t = \mathrm{d}B_t + V'\left(\sup_{s\le t} X_s - X_t\right)\mathrm{d}t \\ X_0 = 0 \end{cases} \tag{0.1}$$

where $(B_t, t \geq 0)$ is a standard Brownian motion and V a symmetric function of class $\mathcal{C}^2(\mathbb{R})$, such that $\int_0^\infty e^{-2V(x)}\mathrm{d}x < \infty$.

We begin to notice that the existence and uniqueness of the solution to the equation is straightforward if V' is a Lipschitz function (see e.g. Rogers & Williams, theorem 11.2).

Proposition 0.1. *There exists a unique global solution to the equation (0.1).*

Démonstration. As V' is a locally Lipschitz function, there exists a unique local solution to (0.1). It remains to prove that the process X does not explode in a finite time. Let $Z_t := \sup_{s\le t} X_s - X_t$ and use the usual notation $M_t := \sup_{s\le t} X_s$. We get

$$\mathrm{d}Z_t = -\mathrm{d}B_t - V'(Z_t)\mathrm{d}t + \mathrm{d}M_t.$$

We emphasize that the process Z is a continuous non-negative semi-martingale. The Tanaka formula then implies that $Z_t - Z_0 = \int_0^t sgn(Z_s)\mathrm{d}Z_s + L_t^0 = \int_0^t (1 - 2\mathbb{1}_{\{Z_s=0\}})\mathrm{d}Z_s + L_t^0$ where L_t^0 is the local time of Z in 0. We therefore get

$$\begin{aligned} L_t^0 &= 2\int_0^t \mathbb{1}_{\{Z_s=0\}}\mathrm{d}Z_s \\ &= -2\int_0^t \mathbb{1}_{\{Z_s=0\}}\mathrm{d}B_s - 2\int_0^t \mathbb{1}_{\{Z_s=0\}}V'(Z_s)\mathrm{d}s + 2\int_0^t \mathbb{1}_{\{Z_s=0\}}\mathrm{d}M_s. \end{aligned}$$

But the bracket of the local martingale $\int_0^t \mathbb{1}_{\{Z_s=0\}}\mathrm{d}B_s$ vanishes ($4\int_0^t \mathbb{1}_{\{Z_s=0\}}\mathrm{d}s = 0$) and as a consequence, this martingale vanishes a.s. Moreover, the function V is symmetric, thus $V'(0) = 0$. Finally, we have (by definition) that

$L_t^0 = 2\int_0^t \mathbb{1}_{\{Z_s=0\}}\mathrm{d}M_s = 2M_t$, that is

$$(L_t^0, t \geq 0) \stackrel{(d)}{=} (2M_t, t \geq 0).$$

But we know that the local time L_t^0 does not explode in a finite time. Therefore, the process M does not explode in a finite time and we conclude because $X_t \leq M_t$ for all $t \geq 0$. □

PROPOSITION 0.2. *We have the following identity*

$$(X_t; t \geq 0) \stackrel{(d)}{=} (L_t - |Y_t|; t \geq 0)$$

where Y satisfies $\mathrm{d}Y_t = \mathrm{d}B_t - V'(Y_t)\mathrm{d}t$ and L is its local time in 0.

DÉMONSTRATION. Let T be a positive constant and F a non negative functional on the space $\mathcal{C}^0([0,T],\mathbb{R})$. By use of the Girsanov theorem, we find :

$$\mathbb{E}[F(X_t; t \leq T)] = \mathbb{E}[F(B_t; t \leq T)\mathcal{E}(\int_0^{\cdot} V'(M_s - B_s)\mathrm{d}B_s)_T]$$

where $M_s = \sup_{u \leq s} B_u$ and $\mathcal{E}$ is the usual Doléans-Dade martingale.
We apply now the identity of Lévy

$$(M_t - B_t, M_t; t \geq 0) \stackrel{(d)}{=} (|B_t|, L_t; t \geq 0)$$

and with the help of Tanaka's formula for the local time, we remark that $B_t \stackrel{(d)}{=} L_t - |B_t|$ and $M_t - B_t \stackrel{(d)}{=} |B_t|$. Thus, noticing that $V'(0) = 0$, we can rewrite the preceding equality as

$$\begin{aligned}\mathbb{E}[F(X_t; t \leq T)] &= \mathbb{E}[F(L_t - |B_t|; t \leq T)\mathcal{E}\left(-\int_0^{\cdot} V'(|B_s|)\mathrm{d}(L_s - |B_s|)\right)_T] \\ &= \mathbb{E}[F(L_t - |B_t|; t \leq T)\mathcal{E}(-\int_0^{\cdot} V'(B_s)\mathrm{d}B_s)_T].\end{aligned}$$

This last result implies the announced result for $t \leq T$. We just have to let T increase to the infinity. The result follows by applying the Girsanov theorem :

$$\mathbb{E}[F(L_t - |B_t|; t \leq T)\mathcal{E}(-\int_0^{\cdot} V'(B_s)\mathrm{d}B_s)_T] = \mathbb{E}[F(L_t^Y - |Y_t|; t \leq T)],$$

where Y is the announced process. □

REMARK 0.3. *1) If $V(x) = c|x|^2$, where $c > 0$, then Y is an Ornstein-Uhlenbeck process.*
2) If $X_0 = x \in \mathbb{R}$, we just have to adapt the proof and replace the local time in 0 of the process Y by its local time in $-x$.

COROLLARY 0.4. *Suppose that $V(x) \geq 0$. Then if we define the probability measure $\gamma(\mathrm{d}x) = \frac{1}{Z}e^{-2V(x)}\mathrm{d}x$, we get*

$$\frac{X_t}{t} \xrightarrow[t\to\infty]{a.s.} \int_{\mathbb{R}} |V'(x)|\gamma(\mathrm{d}x).$$

In particular, if $V(x) = cx^2$, where c is a positive constant, then

$$\frac{X_t}{t} \xrightarrow[t\to\infty]{a.s.} \sqrt{\frac{c}{\pi}}.$$

DÉMONSTRATION. By the preceding proposition, it is sufficient to prove the corollary for $(L_t - |Y_t|; t \geq 0)$. By the formula of Tanaka, we have

$$\frac{L_t - |Y_t|}{t} = -\frac{1}{t}\int_0^t \operatorname{sgn}(Y_s)\mathrm{d}Y_s = -\frac{1}{t}\int_0^t \operatorname{sgn}(Y_s)\mathrm{d}W_s + \frac{1}{t}\int_0^t |V'(Y_s)|\mathrm{d}s.$$

The strong law of large numbers implies then that the local martingale $-\frac{1}{t}\int_0^t \operatorname{sgn}(Y_s)\mathrm{d}W_s$ converges to 0 when t increases to the infinity. Moreover, the ergodic lemma (see e.g. Kallenberg, p.464) implies, because the invariant measure of the process Y is the probability measure $\gamma(\mathrm{d}x) = \frac{1}{Z}e^{-2V(x)}\mathrm{d}x$ (where Z is the normalization constant), that

$$\frac{1}{t}\int_0^t |V'(Y_s)|\mathrm{d}s \xrightarrow{a.s.} \int_{\mathbb{R}} |V'(x)|\gamma(\mathrm{d}x).$$

It then implies that $\frac{X_t}{t} \xrightarrow{a.s.} \int_{\mathbb{R}} |V'(x)|\gamma(\mathrm{d}x)$. □

REMARK 0.5. *If $V(x) = cx^2$, where the constant c is negative, then the Ornstein-Uhlenbeck process considered is transient (see e.g. Borodin & Salminen).*

The question is now the asymptotic study of the process

$$\mathrm{d}X_t = \mathrm{d}B_t + c\left(\sup_{s\leq t} X_s + \inf_{s\leq t} X_s - 2X_t\right)\mathrm{d}t,$$

or more generally (say for a convex and symmetric function V)

$$\mathrm{d}X_t = \mathrm{d}B_t + V'\left(\sup_{s\leq t} X_s + \inf_{s\leq t} X_s - 2X_t\right)\mathrm{d}t.$$

ANNEXE B

Décomposition d'une fonction strictement convexe à l'infini

We will prove here the following lemma:

LEMMA 0.1. *Any function f, strictly uniformly convex out of a compact subset of $\mathbb{R}^d$, can be written as a bounded function added to a strictly uniformly convex function.*

PROOF. We will build the bounded function. Let $\psi : \mathbb{R}_+ \to \mathbb{R}$ be a bounded function C^∞, which characteristics will be described later.
Let $U(x) := f(x) + \psi(|x|^2)$. For all vector h of $\mathbb{R}^d$, we have:

$$(h\nabla^2 U(x), h) = (h\nabla^2 f(x), h) + 4|x \cdot h|^2 \psi''(|x|^2) + 2|h|^2 \psi'(|x|^2).$$

But we know, by definition of f, that there exist two positive constants C, ε and a "border" R such that

$$\begin{cases} (h\nabla^2 f(x), h) \geq C|h|^2 & \text{for } |x| \geq R \\ (h\nabla^2 f(x), h) \geq -\varepsilon|h|^2 & \text{otherwise} \end{cases}$$

We now have to choose ψ.
For $|x| < R$, we want ψ to be such that $\psi'' > 0$ and $\psi' > \varepsilon > 0$. Then we find

$$(h\nabla^2 U(x), h) \geq -\varepsilon|h|^2 + 2\varepsilon|h|^2 = \varepsilon|h|^2.$$

If on the contrary, $|x| \geq R$, we enforce on ψ that $\psi' \geq 0$, ψ'' is bounded and $x\psi(x) \geq -m$, where m is a positive constant such that $C - 4m > 0$. We have then the following inequality:

$$(h\nabla^2 U(x), h) \geq C|h|^2 + 4|x|^2|h|^2\psi''(|x|^2) = (C - 4m)|h|^2.$$

We conclude that U is strictly uniformly convex, *i.e.* for all $x, h \in \mathbb{R}^d$,

$$(h\nabla^2 U(x), h) \geq \min(\varepsilon, C - 4m)|h|^2$$

and by construction, ψ is bounded everywhere. □

REMARK 0.2.
- *For instance, we can choose $\psi(x) = -(R+\varepsilon)^2 e^{-\frac{1}{2}\left(\frac{x+\varepsilon}{R+\varepsilon}\right)^2}$.*
- *We can also notice that the converse is false (for example if $U(x) = x^2$ and $\psi(x) = 3\sin x$, the function $f(x) = U(x) + \psi(x)$ is not strictly uniformly convex out of a compact set.)*

- *With a convolution and a gluing, we can prove in addition that the support of the bounded function is compact (but we think that the proof given is much more readable). For a proof including the compactness of the support of ψ, we refer the reader to Ghomi (2002).*

Résumé

Keywords: self-interacting diffusion, random reinforced process, asymptotic pseudo-trajectories, stochastic approximation, dynamical system.

Summary: **First part**. The first chapter is concerned with some self-interacting diffusions $(X_t, t \geq 0)$ living on $\mathbb{R}^d$. These diffusions are solutions to stochastic differential equations:

$$\mathrm{d}X_t = \mathrm{d}B_t - g(t)\nabla V(X_t - \overline{\mu}_t)\mathrm{d}t,$$

where $\overline{\mu}_t$ is the empirical mean of the process X, V is an asymptotically strictly convex potential and g is a given function. We study the ergodic behavior of X and prove that it is strongly related to g. Actually, X and $\overline{\mu}_t$ have the same asymptotic behavior and we will give necessary and sufficient conditions (on g and V) for the almost sure convergence of X. In chapter 2, we finish the previous study. We have still studied the ergodic behavior of X and proved that it is strongly related to g. We go further and give necessary and sufficient conditions (for small g's) in order that X converges in law to X_∞ (which is related to the global minima of V).

Second part. We begin to situate our study in Chapter 3. Self-interacting diffusions are solutions to SDEs with a drift term depending on the process and its normalized occupation measure μ_t (via an interaction potential V and a confinement potential W):

$$\begin{cases} \mathrm{d}X_t = \mathrm{d}B_t - \left(\nabla V(X_t) + \frac{1}{t}\int_0^t \nabla_x W(X_t, X_s)\mathrm{d}s\right)\mathrm{d}t \\ \mathrm{d}\mu_t = (\delta_{X_t} - \mu_t)\frac{\mathrm{d}t}{r+t} \\ X_0 = x,\ \ \mu_0 = \mu. \end{cases}$$

We establish a relation between the asymptotic behavior of μ_t and the asymptotic behavior of a deterministic dynamical flow (defined on the space of the Borel probability measures). We extend previous results on $\mathbb{R}^d$ or more generally a smooth complete connected Riemannian manifold without boundary. We will also give some sufficient conditions for the convergence of μ_t. We then illustrate, in Chapter 5, the previous study of self-interacting diffusions living in $\mathbb{R}^d$ with some examples in the two-dimensional case. The preceding paper contains abstract results, and therefore we describe here a simple example and illustrate some of our previous results. We will show in particular that, depending on W, either the empirical measure behaves like the "Brownian motion" (constructed with respect to the measure $e^{V(x)}\mathrm{d}x$);

or the empirical occupation measure converges almost surely to a probability measure, which is approximatively a Gaussian distribution ; or there is enough attraction, and then the term induced by W forces μ_t to circle around and the limit set of (μ_t) is a circle of measures $\{\nu(\delta), 0 \leq \delta < 2\pi\}$.

Mots-clés : diffusion, processus renforcé, auto-interaction, pseudo-trajectoire asymptotique, approximation stochastique, système dynamique.

Résumé. Le but de cette thèse est d'étudier le comportement asymptotique de diffusions auto-interactives sur $\mathbb{R}^d$. Nous étudions deux familles de processus renforcés. La première est régie par l'équation différentielle stochastique

$$\mathrm{d}X_t = \mathrm{d}B_t - g(t)\nabla V(X_t - \overline{\mu}_t)\mathrm{d}t,$$

où $\overline{\mu}_t$ est la moyenne de la mesure empirique du processus X, V est potentiel strictement uniformément convexe en dehors d'un compact et g est une fonction donnée. Nous étudions alors le comportement asymptotique de X, en fonction de g. Selon la forme de g, on peut montrer que X converge presque-sûrement (chapitre 1) ou alors converge en loi (chapitre 2). Dans une seconde partie, nous nous intéressons à une famille plus complexe, correspondant aux diffusions renforcées par la mesure d'occupation. Il s'agit de processus satisfaisant l'équation

$$\begin{cases} \mathrm{d}X_t = \mathrm{d}B_t - \left(\nabla V(X_t) + \frac{1}{t}\int_0^t \nabla_x W(X_t, X_s)\mathrm{d}s\right)\mathrm{d}t \\ \mathrm{d}\mu_t = (\delta_{X_t} - \mu_t)\frac{\mathrm{d}t}{r+t} \\ X_0 = x, \;\; \mu_0 = \mu. \end{cases}$$

Nous établissons une relation entre le comportement asymptotique de μ_t et le comportement asymptotique dun système dynamique déterministe (défini sur l'espace des probabilités). Nous étendons alors de précédents résultats à $\mathbb{R}^d$. Nous donnons également des conditions suffisantes pour la convergence de μ_t. Enfin, nous illustrons, au chaptre 5, l'étude précédente de diffusions auto-interactives par quelques exemples en dimension deux.