UNIVERSITÉ DE TOURS

École Doctorale MIPTIS
Institut Denis Poisson

THÈSE présentée par :

Ninon FÉTIQUE

soutenue le : 27 juin 2019

pour obtenir le grade de : Docteur de l'université de Tours

Discipline/ Spécialité : Mathématiques

ÉTUDE DU COMPORTEMENT EN TEMPS LONG DE CERTAINS PROCESSUS DE MARKOV DÉTERMINISTES PAR MORCEAUX DE TYPE POSITION-VITESSE

THÈSE DIRIGÉE PAR :

MALRIEU Florent	Professeur, Université de Tours
GUÉRIN Hélène	Professeure, Université du Québec à Montréal

RAPPORTEURS :

DE SAPORTA Benoîte	Professeure, Université de Montpellier
THIEULLEN Michèle	Maîtresse de conférences (HDR), Sorbonne Université

JURY :

BRETON Jean-Christophe	Professeur, Université de Rennes 1
DE SAPORTA Benoîte	Professeure, Université de Montpellier
GOUÉRÉ Jean-Baptiste	Professeur, Université de Tours
GUÉRIN Hélène	Professeure, Université du Québec à Montréal
MALRIEU Florent	Professeur, Université de Tours
THIEULLEN Michèle	Maîtresse de conférences, Sorbonne Université
ZITT Pierre-André	Maître de conférences, Université Paris - Est - Marne-la-Vallée

Résumé

L'objet de cette thèse est l'étude de certains processus de Markov déterministes par morceaux (abrégé dans la suite par PDMPs pour l'acronyme anglais), et plus particulièrement de leur comportement en temps long. Pour cela, nous utilisons des méthodes de couplage. Dans un premier chapitre, nous nous intéressons à un PDMP appelé « billard stochastique », décrivant le mouvement d'une particule dans un convexe borné du plan. Nous donnons, dans des cas particuliers de convexes, une borne explicite sur la vitesse de convergence à l'équilibre du processus. Dans un deuxième chapitre, nous étudions l'ergodicité d'un PDMP que l'on a appelé « processus Zig-zag généralisé », pouvant décrire le mouvement linéaire par morceaux d'une bactérie attirée par un nutriment fixe dans son environnement. Enfin, dans un dernier chapitre, nous étudions un système de particules, dont chaque particule est un processus Zig-zag (cas particulier du processus étudié dans le deuxième chapitre), non plus attiré par un nutriment, mais par la moyenne spatiale du système de particules. Nous étudions la propagation du chaos de ce système de particules, ainsi que le comportement en temps long du processus limite.

Mots clés : Processus de Markov déterministes par morceaux ; Comportement en temps long ; Vitesse de convergence ; Méthode de couplage ; Système de particules en interaction ; Propagation du chaos.

Abstract

The purpose of this PhD thesis is the study of some Piecewise deterministic Markov processes (PDMPs), and in particular of their long-time behaviour. For that, we use coupling methods. In a first chapter, we are interested in a PDMP called "stochastic billiard", that describes the movement of a particle in a bounded convex set of the plan. In particular cases of convex sets, we give an explicit bound for the speed of convergence of the process. In a second chapter, we study a PDMP that we have named "generalised Zig-zag process", that can describe the piecewise linear movement of a bacteria which is attracted by a fixed nutriment in its environment. Finally, in a last chapter, we study a particle system in which each particle is a Zig-zag process (particular case of the process studied in the second chapter), attracted by the spatial mean of the particle system. We study the propagation of chaos of this particle system, and the long-time behaviour of the limit process.

Keywords : Piecewise deterministic Markov processes ; Long-time behaviour ; Speed of convergence ; Coupling method ; Interacting particle system ; Propagation of chaos.

Table des matières

Introduction

Dans cette thèse, nous nous intéressons à une classe particulière de processus de Markov, les processus de Markov déterministes par morceaux, que l'on abrège par PDMPs pour Piecewise Deterministic Markov Processes. Ces processus ont été introduits par Davis en 1984 dans [Davis, 1984] pour les distinguer des diffusions. Ce sont des processus aujourd'hui très étudiés, car ils permettent de modéliser de nombreux phénomènes, en biologie ou finance par exemple, mais également car ils ont un intérêt dans des problèmes d'échantillonnage.
Un PDMP est un processus de Markov qui évolue de manière déterministe pendant un temps aléatoire, dépendant de son état. Puis, au bout de ce temps, le processus "saute", c'est-à-dire qu'il change aléatoirement d'état. Dans cette thèse, nous nous intéresserons à une classe particulière de PDMPs, des PDMPs sous forme d'un couple position-vitesse. Ainsi, l'évolution déterministe du processus est linéaire, et à chaque saut, seule la vitesse du processus change. Plus particulièrement, notre intérêt sera le comportement en temps long de ces processus, et nous utiliserons pour cela des méthodes de couplage.

Ce manuscrit se divise en quatre chapitres. Dans une première partie, nous introduisons les notions et résultats utiles pour comprendre la suite du manuscrit, et pour replacer la thèse dans son contexte. Dans une deuxième partie nous nous intéressons à la vitesse de convergence d'un PDMP appelé billard stochastique, évoluant dans des convexes particuliers du plan. Nous construisons pour cela des couplages astucieux. Dans la troisième partie, nous étudions la convergence d'un processus que l'on a appelé "processus Zig-zag généralisé", pouvant modéliser le mouvement d'une bactérie attirée par un nutriment fixe dans son environnement. Pour cela, nous utilisons les méthodes de couplage de Meyn et Tweedie. Enfin, l'objet d'étude de la dernière partie est un système de PDMPs en interaction par leur moyenne. Nous nous intéressons à la propagation du chaos de ce système vers un processus non-linéaire, ainsi qu'au comportement en temps long de ce processus limite.

Les simulations présentées dans ce manuscrit ont été faites avec les logiciels Python et Scilab, et les illustrations avec Geogebra et TikZ.

Le Chapitre 2 de ce manuscrit est l'article [Fétique, 2019], accepté dans le journal Séminaire de Probabilités, et le Chapitre 3 est l'article [Fétique, 2017], soumis.

Chapitre 1

Introduction générale

Dans ce chapitre nous introduisons les notions et définitions nécessaires pour la suite du manuscrit. En particulier nous introduisons la notion de processus de Markov déterministes par morceaux, qui est l'objet central de cette thèse, et nous donnons des résultats qui permettent d'étudier leur comportement en temps long.

Nous commençons par introduire quelques notations utilisées tout au long de cette thèse.

Notations

- pour $k \in \mathbb{N}$, $\mathcal{C}^k(E)$ désigne l'ensemble des fonctions réelles sur $E \subset \mathbb{R}^d$ qui sont k-fois continument différentiables. On écrira $\mathcal{C}(E)$ au lieu de $\mathcal{C}^0(E)$;
- $\mathcal{C}_0(E)$ désigne l'ensemble des fonctions réelles continues sur E telles que $\lim_{\|x\| \to \infty} f(x) = 0$;
- $\mathcal{C}_b(E)$ désigne l'ensemble des fonctions réelles bornées sur E ;
- $\mathcal{M}(\mathbb{X})$ désigne l'ensemble des mesures de probabilité sur l'espace $\mathbb{X}$;
- on note $\mathcal{L}(X)$ la loi d'une variable aléatoire X, et on notera $X \sim \mu$ pour dire que la variable X suit la loi de probabilité μ. De plus, si deux variables X et Y sont égales en loi, on notera $X \overset{\mathcal{L}}{=} Y$;
- on note de la façon suivante les lois usuelles : $\mathcal{E}(\lambda)$ est la loi exponentielle de paramètre λ ; $Rad(p)$ est la loi de Rademacher de paramètre p ; $\mathcal{U}(A)$ est la loi uniforme sur un intervalle A de $\mathbb{R}$; $Beta(\alpha, \beta)$ est la loi Beta de paramètres α et β ;
- pour $x \in \mathbb{R}^d$, δ_x désigne la mesure de Dirac en x ;
- pour $A \subset \mathbb{R}^d$, $\mathbf{1}_A$ désigne la fonction indicatrice (ou fonction caractéristique) de A ;
- si $A \subset \mathbb{R}^d$, $\mathcal{B}(A)$ désigne l'ensemble des boréliens de A ;
- si $x, y \in \mathbb{R}^d$, on notera $\langle x, y \rangle$ le produit scalaire entre x et y, ou $x \cdot y$ pour

alléger les notations, quand cela reste lisible ;
- de manière générale, $\mathbb{E}_\nu[\cdot]$ désigne l'espérance sous la loi initiale ν. On écrit $\mathbb{E}_x[\cdot]$ lorsque la loi initiale est la masse de Dirac δ_x.

Dans toute l'introduction, on se place dans un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$.

1.1 Processus de Markov déterministes par morceaux

1.1.1 Processus de Markov

Dans cette première partie, nous introduisons les notions de base sur les processus de Markov, qui forment le cadre général de cette thèse. Pour plus de détails, on pourra consulter [Ethier et Kurtz, 1986].
Dans toute l'introduction, E désignera l'espace $\mathbb{R}^d$ ou une partie de $\mathbb{R}^d$, $d \in \mathbb{N}^*$.

On considère $(X_t)_{t\geq 0}$ un processus de Markov à valeurs dans E. Rappelons que $(X_t)_{t\geq 0}$ est un processus de Markov si pour tous $s, t \geq 0$ et pour toute fonction mesurable bornée $f : E \longrightarrow \mathbb{R}$ on a

$$\mathbb{E}\left[f(X_{t+s})|\mathcal{F}_t\right] = \mathbb{E}\left[f(X_{t+s})|X_t\right],$$

où $(\mathcal{F}_t)_{t\geq 0}$ désigne la filtration canonique associée à $(X_t)_{t\geq 0}$.
On se place dans le cadre d'un processus de Markov X homogène, c'est-à-dire que la loi de X_{t+s} sachant X_t ne dépend que de s.
Au processus X on associe le semi-groupe $(P_t)_{t\geq 0}$ agissant sur les fonctions mesurables bornées par

$$P_t f(x) = \mathbb{E}\left[f(X_t)|X_0 = x\right] = \mathbb{E}_x\left[f(X_t)\right].$$

Le semi-groupe caractérise la dynamique du processus puisqu'on a $P_t \mathbf{1}_A(x) = \mathbb{P}_x(X_t \in A)$. Ainsi, P donne accès aux lois fini-dimensionnelles de X, puis le théorème d'extension de Kolmogorov assure que P caractérise la loi du processus X, connaissant sa distribution initiale.
Comme son nom l'indique, le semi-groupe P associé au processus de Markov X est un semi-groupe de transition, c'est-à-dire qu'il vérifie les deux propriétés suivantes :

$$P_0 = Id \quad \text{et} \quad \forall s, t \geq 0,\ P_{s+t} = P_s \circ P_t.$$

Le semi-groupe P est dit de Feller si :

1. pour tout fonction $f \in \mathcal{C}_0(E)$, $P_t f \in \mathcal{C}_0(E)$ pour tout $t \geq 0$,
2. pour toute fonction continue $f : E \to \mathbb{R}$, on a $P_t f(x) \underset{t\to 0}{\longrightarrow} f(x)$ pour tout $x \in E$.

Si P est un semi-groupe de Feller, on définit alors son générateur infinitésimal L par

$$Lf = \lim_{t\to 0} \frac{P_t f - f}{t}$$

pour $f \in \mathcal{C}_0(E)$ telle que la limite dans $\mathcal{C}_0(E)$ existe ; on notera $\mathcal{D}(L)$ cet ensemble, appelé domaine de L.
Pour $f \in \mathcal{D}(L)$ on a $P_t f \in \mathcal{D}(L)$ et on a les égalités suivantes :

$$\partial_t P_t f = LP_t f = P_t Lf \quad \text{et} \quad P_t f = f + \int_0^t LP_s f \mathrm{d}s.$$

En particulier, le générateur vérifie la formule de Dynkin :

$$\mathbb{E}_x \left[f(X_t)\right] = f(x) + \mathbb{E}_x \left[\int_0^t Lf(X_s)\mathrm{d}s\right].$$

Ainsi, pour toute fonction $f \in \mathcal{D}(L)$, le processus $(M_t^f)_{t\geq 0}$ défini par

$$M_t^f = f(X_t) - f(X_0) - \int_0^t Lf(X_s)\mathrm{d}s$$

est une martingale par rapport à la filtration $(\mathcal{F}_t)_{t\geq 0}$ engendrée par le processus de Markov.
Le semi-groupe P, et donc la dynamique du processus de Markov X, sont entièrement caractérisés par le générateur L et son domaine. Le générateur a l'avantage d'être en général explicite, contrairement au semi-groupe dont on ne connaît la plupart du temps pas d'expression. Dans la suite, on utilisera donc principalement le générateur des processus de Markov auxquels on s'intéresse pour les décrire et les étudier.
On finit cette section par la notion de mesure invariante. On dit que $\pi \in \mathcal{M}(E)$ est une mesure invariante pour le processus X si pour tout $t \geq 0$, et toute fonction $f \in \mathcal{C}_0(E)$ on a

$$\int P_t f \mathrm{d}\pi = \int f \mathrm{d}\pi.$$

Cela signifie que si la loi initiale du processus X à un instant s est π, alors pour tout $t \geq s$, la loi de X_t est également donnée par π. De manière équivalente, π est une mesure invariante pour X si pour toute fonction $f \in \mathcal{D}(L)$ on a

$$\int Lf \mathrm{d}\pi = 0.$$

1.1.2 Processus de Markov déterministes par morceaux

1.1.2.1 Processus de Markov déterministes par morceaux à valeurs dans $\mathbb{R}^d$

Les processus de Markov déterministes par morceaux (PDMPs) sont un cas particulier de processus de Markov. Ils ont été introduits par Davis ([Davis, 1984]) pour les distinguer des processus diffusifs.
Nous commençons par considérer le cas des processus dont l'espace d'états est $\mathbb{R}^d$ tout entier. Le cas d'un espace d'états strictement inclus dans $\mathbb{R}^d$ nécessite que le processus saute lorsqu'il atteint la frontière de cet ensemble,

ce qui complique un peu les écritures. Ce cas sera évoqué dans la section 1.1.2.2. Pour plus de résultats sur les PDMPs, on pourra consulter par exemple [Davis, 1984, Davis, 1993, Jacobsen, 2006].
Le mouvement d'un PDMP $(X_t)_{t\geq 0}$ à valeurs dans $\mathbb{R}^d$ dépend de trois caractéristiques (nous donnons juste après des conditions assurant l'existence d'un PDMP ayant ces caractéristiques) :

- le taux de saut $\lambda : \mathbb{R}^d \longrightarrow \mathbb{R}^+$: quand $X_t = x$, le processus a un taux de saut de $\lambda(x)$, donc plus le taux de saut est élevé, plus le processus a de chances de sauter ;
- le noyau de transition $Q(x, \mathrm{d}y)$, qui décrit les sauts de X ;
- le champ de vecteurs $b : \mathbb{R}^+ \times \mathbb{R}^d \longrightarrow \mathbb{R}^d$ qui décrit l'évolution déterministe du processus en dehors des sauts : $\mathrm{d}X_t = b(t, X_t)\mathrm{d}t$.

Ici, nous nous intéresserons au cas où le champ de vecteurs b ne dépend pas du temps, i.e. $b : \mathbb{R}^d \longrightarrow \mathbb{R}^d$. Dans ce cas, le générateur d'un tel processus s'écrit :

$$Lf(x) = b(x) \cdot \nabla f(x) + \lambda(x) \int (f(y) - f(x))Q(x, \mathrm{d}y). \tag{1.1}$$

Exemple 1.1.1. Donnons un premier exemple simple de PDMP, qui est à la base de cette thèse, et qui a été étudié par exemple dans [Fontbona *et al.*, 2012], et sous une forme plus générale dans [Bierkens et Roberts, 2017] et [Fontbona *et al.*, 2016]. Le processus Zig-Zag $((X_t, V_t))_{t\geq 0}$ à valeurs dans $\mathbb{R} \times \{-1, +1\}$ est le processus dont le générateur L est donné par : pour $(x, v) \in \mathbb{R} \times \{-1, +1\}$,

$$Lf(x, v) = v\partial_x f(x, v) + (a\mathbf{1}_{xv<0} + b\mathbf{1}_{xv\geq 0}) (f(x, -v) - f(x, v)),$$

où a et b sont deux constantes strictement positives.
On pourra observer sur la Figure 1.1 une trajectoire du processus Zig-zag, avec $a = 1$ et $b = 1.1$.
D'après l'expression du générateur, la composante X du processus est continue et évolue suivant l'équation différentielle $\frac{\mathrm{d}X_t}{\mathrm{d}t} = V_t$, alors que la composante V est constante par morceaux, et est changée en son opposé avec un taux de saut égal à a ou b selon que X se rapproche ou s'éloigne de l'origine.
Dans le cas $a < b$, ce processus est utilisé pour modéliser l'évolution d'une bactérie attirée par un nutriment qui se trouverait en l'origine. La première composante X_t représente alors la position de la bactérie à l'instant t, et V_t sa vitesse.

Décrivons maintenant la construction d'un PDMP dont les caractéristiques λ, Q et b satisfont les hypothèses suivantes :

- le taux de saut λ est une fonction mesurable positive ;
- le noyau de transition Q est un noyau markovien ;
- le champ de vecteur b est globalement lipschitzien, de façon à ce qu'il existe un unique flot $\varphi(t, x)$ tel que $t \mapsto \varphi(t, x)$ est solution de l'équation différentielle suivante :

$$\begin{cases} \mathrm{d}x_t = b(x_t)\mathrm{d}t \quad \text{pour} \quad t > 0, \\ x_0 = x. \end{cases}$$

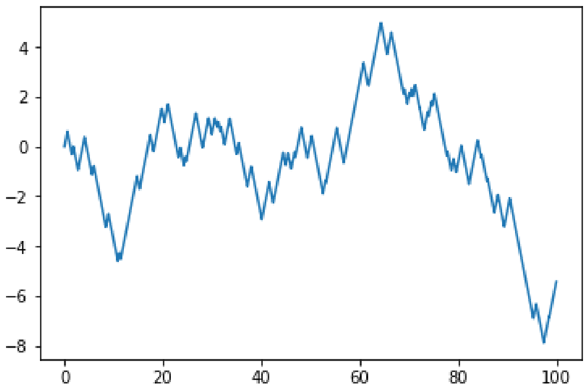


FIGURE 1.1 – Trajectoire du processus Zig-zag avec $a = 1$ et $b = 1.1$, avec conditions initiales $(x_0, v_0) = (0, 1)$

Considérons la fonction définie pour $t \geq 0$ par

$$F(t) = 1 - \exp\left(-\int_0^t \lambda(u)\mathrm{d}u\right).$$

La fonction F est la fonction de répartition d'une variable aléatoire T positive, dont la densité par rapport à la mesure de Lebesgue est

$$\forall t \geq 0, \quad f(t) = \lambda(t)\exp\left(-\int_0^t \lambda(u)\mathrm{d}u\right).$$

Dans un contexte général, on dit que λ est le taux de défaillance associé à la variable T (ou taux de saut dans le langage des PDMPs).
Si $E \sim \mathcal{E}(1)$, alors on vérifie facilement que la variable aléatoire définie par

$$\widetilde{T} = \inf\left\{t \geq 0 : \int_0^t \lambda(u)\mathrm{d}u \geq E\right\}$$

a un taux de défaillance λ. Remarquons que ce résultat peut être utilisé en particulier pour simuler une variable aléatoire de taux de défaillance λ.
On peut alors construire le PDMP de générateur L donné par (1.1). Pour cela, on commence par construire la chaîne de Markov $(\overline{X}_n, \tau_n)_{n\geq 1}$, où la suite $(\tau_n)_{n\geq 1}$ représente la suite des inter-sauts du PDMP que l'on cherche à construire, et la suite $(\overline{X}_n)_{n\geq 1}$ représente la position du processus X après chaque instant de saut. Cette chaîne de Markov $(\overline{X}_n, \tau_n)_{n\geq 1}$ est définie par

$$\begin{cases} \tau_{n+1} = \inf\left\{t \geq 0 : \int_0^t \lambda(\varphi(u, \overline{X}_n))\mathrm{d}u \geq E_{n+1}\right\} \\ \overline{X}_{n+1} = J(\varphi(\tau_{n+1}, X_n), U_{n+1}), \end{cases}$$

où $(E_i)_{i\geq 1}$ et $(U_i)_{i\geq 1}$ sont des suites de variables aléatoires indépendantes de lois respectives $\mathcal{E}(1)$ et $\mathcal{U}([0,1])$, et J est une fonction mesurable de $\mathbb{R}^d \times [0,1]$

dans $\mathbb{R}^d$ telle que $J(x,U) \sim Q(x,\mathrm{d}y)$ si $U \sim \mathcal{U}([0,1])$.
On définit alors la suite $(T_n)_{n\geq 0}$ des temps de saut du processus X de la façon suivante :

$$T_0 = 0 \quad \text{et} \quad T_{n+1} = T_n + \tau_{n+1}.$$

Finalement, on définit le processus $(X_t)_{t\geq 0}$ par

$$X_t = \varphi(t - T_n, \overline{X}_n) \text{ pour } t \in [T_n, T_{n+1}[.$$

Montrons maintenant que le processus $(X_t)_{t\geq 0}$ ainsi défini est bien un processus de Markov de générateur infinitésimal donné par (1.1).
Considérons le processus ponctuel N défini par $N_t = \sum_{n\geq 1} \mathbf{1}_{T_n \leq t}$: N_t représente le nombre de sauts du processus avant le temps t.
Pour $f : \mathbb{R}^d \longrightarrow \mathbb{R} \in \mathcal{C}_b^1$ on a :

$$\begin{aligned}
P_t f(x) &= \mathbb{E}_x\left[f(X_t)\mathbf{1}_{N_t=0}\right] + \mathbb{E}_x\left[f(X_t)\mathbf{1}_{N_t=1}\right] + \mathbb{E}_x\left[f(X_t)\mathbf{1}_{N_t\geq 2}\right] \\
&= f(\varphi(t,x))\mathbb{P}_x(T_1 > t) + \mathbb{E}_x\left[f(\varphi(t - T_1, X_{T_1}))\mathbf{1}_{T_1 \leq t < T_2}\right] \\
&\qquad + \mathbb{E}_x\left[f(X_t)\mathbf{1}_{N_t\geq 2}\right].
\end{aligned}$$

Par un développement de Taylor de f, le premier terme donne

$$\begin{aligned}
f(\varphi(t,x))\mathbb{P}_x(T_1 > t) &= \left(f(x) + tb(x)\cdot\nabla f(x) + o(t)\right)\left(1 - t\lambda(x) + o(t)\right) \\
&= f(x) + tb(x)\cdot\nabla f(x) - \lambda f(x) + o(t).
\end{aligned}$$

Sous $\mathbb{P}_x$, la loi du premier temps de saut a pour fonction de répartition la fonction de répartition associée au taux de défaillance $t \mapsto \lambda(\varphi(t,x))$. On notera G_x cette fonction, et g_x la densité associée. Le second terme donne donc

$$\begin{aligned}
&\mathbb{E}_x\left[f(\varphi(t - T_1, X_{T_1}))\mathbf{1}_{T_1 \leq t < T_2}\right] \\
&= \int_0^t \int_{\mathbb{R}^d} f(\varphi(t-s,y))(1 - G_y(t-s))Q(\varphi(s,x),\mathrm{d}y)g_x(s)\mathrm{d}s \\
&= t\lambda(x)\int_{\mathbb{R}^d} f(y)Q(x,\mathrm{d}y) + o(t)
\end{aligned}$$

en utilisant le fait que φ et Q sont continues, et que g_x est continue en 0 avec $g_x(0) = \lambda(x)$.
Enfin, on a $\mathbb{P}_x(N_t \geq 2) = O(t^2)$ d'après les résultats classiques sur les processus de Poisson.
Donc par les calculs précédents, et comme $Lf(x) = \lim_{t\to 0} \frac{P_t f(x) - f(x)}{t}$, on en déduit que le processus construit ci-dessus a pour pour générateur L défini par (1.1).

Enfin, nous mentionnons le fait qu'au lieu d'être décrits par leur générateur, les PDMPs peuvent être vus comme solutions d'équations différentielles stochastiques (EDS) conduites par des processus de Poisson (on pourra consulter [Gikhman et Skorohod, 1972, Ikeda et Watanabe, 2014] pour plus de détails).
Soit X un PDMP de caractéristiques λ, Q, et b, dont on rappelle que le générateur est alors donné par

$$Lf(x) = b(x)\cdot\nabla f(x) + \lambda(x)\int (f(y) - f(x))Q(x,dy).$$

Le processus X est alors solution de l'EDS suivante :

$$\mathrm{d}X_t = b(X_{t^-})\mathrm{d}t + \int_0^1 \int_{\mathbb{R}^d} \int_0^{+\infty} (\Psi(x,u) - x)\, \mathbf{1}_{z \le \lambda(X_{t^-})} \mathcal{N}(\mathrm{d}z, \mathrm{d}t, \mathrm{d}u),$$

où $\mathcal{N}$ est un processus de Poisson sur $\mathbb{R}^+ \times \mathbb{R}^+ \times [0,1]$ d'intensité $\mathrm{d}z\mathrm{d}t\mathrm{d}u$, et Ψ est une fonction mesurable de $\mathbb{R}^d \times [0,1]$ dans $\mathbb{R}^d$ telle que $\Psi(x,U) \sim Q(x,\mathrm{d}y)$ si $U \sim \mathcal{U}([0,1])$.

Exemple 1.1.2. Nous reprenons l'Exemple 1.1.1 du processus Zig-zag (X,V), dont on rappelle le générateur : pour $(x,v) \in \mathbb{R} \times \{-1,+1\}$,

$$Lf(x,v) = v\partial_x f(x,v) + (a\mathbf{1}_{xv<0} + b\mathbf{1}_{xv\ge 0})\,(f(x,-v) - f(x,v)).$$

Ce processus est alors solution du système stochastique suivant :

$$\begin{cases} \mathrm{d}X_t & = V_t \mathrm{d}t \\ \mathrm{d}V_t & = -2V_{t^-} \int \mathbf{1}_{z \le \lambda(X_t V_{t^-})} \mathcal{N}(\mathrm{d}z, \mathrm{d}t), \end{cases}$$

où $\lambda(xv) = a\mathbf{1}_{xv<0} + b\mathbf{1}_{xv\ge 0}$ et où $\mathcal{N}$ est un processus de Poisson d'intensité la mesure de Lebesgue sur $\mathbb{R}^+ \times \mathbb{R}^+$.

1.1.2.2 Processus de Markov déterministes par morceaux à valeurs dans un sous-ensemble borné de $\mathbb{R}^d$

Dans cette partie, on décrit la construction des PDMPs évoluant dans un sous-ensemble de $\mathbb{R}^d$.
Soit M un ouvert de $\mathbb{R}^d$. On note ∂M sa frontière et $\overline{M}$ sa fermeture. Comme pour les PDMPs évoluant dans $\mathbb{R}^d$ tout entier, un PDMP évoluant dans M dépend uniquement de trois caractéristiques, que nous réintroduisons par soucis de clarté dans les notations, et pour introduire les hypothèses nécessaires :

— le taux de saut $\lambda : \overline{M} \to \mathbb{R}^+$, que l'on suppose mesurable et satisfaisant

$$\forall x \in M,\ \exists \varepsilon > 0,\ \int_0^\varepsilon \lambda(\varphi(x,s))\mathrm{d}s < \infty;$$

— le noyau de saut markovien Q sur $(\overline{M}, \mathcal{B}(\overline{M}))$, satisfaisant la propriété

$$\forall x \in \overline{M},\ Q(x, M \setminus \{x\}) = 1.$$

— le champ de vecteurs $b : \mathbb{R}^d \to \mathbb{R}^d$ associé au flot $\varphi : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^d$, que l'on suppose continu et tel que pour tous $t \in \mathbb{R}$, $\varphi(\cdot,t)$ est un homéomorphisme satisfaisant la propriété de semi-groupe : $\varphi(\cdot,t+s) = \varphi(\varphi(\cdot,s),t)$ pour tout $s,t \in \mathbb{R}$;

Le fait de considérer les processus évoluant dans un sous-espace de $\mathbb{R}^d$ nous oblige à introduire le temps d'atteinte de la frontière de M partant d'un point $x \in M$:

$$t^*(x) = \inf\{t > 0 : \varphi(t,x) \in \partial M\},$$

avec la convention $\inf \varnothing = +\infty$.
On ne détaille pas ici la construction du PDMP $(X_t)_{t\geq 0}$ à valeurs dans M de caractéristiques λ, b et Q. La principale différence par rapport à un PDMP vivant dans $\mathbb{R}^d$ tout entier est que partant de l'état $x \in M$, le temps T_1 avant le prochain saut du PDMP est donné par

$$\mathbb{P}_x(T_1 > t) = \exp\left(-\int_0^t \lambda(\varphi(x,s))\mathrm{d}s\right)\mathbf{1}_{t\leq t^*(x)}.$$

Le générateur de X est alors donné par

$$\widetilde{L}f(x) = b(x)\cdot\nabla f(x) + \lambda(x)\int_M (f(y)-f(x))Q(x,\mathrm{d}y),$$

pour les fonctions f satisfaisant la condition de bord suivante :

$$f(x) = \int_M f(y)Q(x,\mathrm{d}y), \quad \forall x \in \partial M.$$

La forme du générateur est ici la même que dans le cas d'un PDMP à valeurs dans $\mathbb{R}^d$ tout entier, la différence se lit dans la condition de bord ci-dessus, qui traduit les sauts forcés sur le bord de M.

Exemple 1.1.3. Le processus appelé billard stochastique peut être décrit de la manière suivante : une particule se déplace à vitesse unitaire à l'intérieur d'un ensemble K jusqu'à ce qu'elle touche son bord, et est alors réfléchie de manière aléatoire, indépendamment de sa position et de sa vitesse précédente.
Décrivons plus précisément ce processus, afin d'en donner son générateur infinitésimal. Soit $K \subset \mathbb{R}^2$ un compact convexe à bord de classe $\mathcal{C}^1$.
Soit $e = (1,0)$ le premier vecteur de la base canonique de $\mathbb{R}^2$. Soit γ une loi sur la demi-sphère $\mathbb{S}_e = \{v \in \mathbb{S}^1 : e\cdot v \geq 0\}$. Soit de plus $(U_x, x \in \partial K)$ une famille de rotations de $\mathbb{S}^1$ telle que $U_x e = n_x$, où on note n_x le vecteur normal rentrant à ∂K au point x.
Soit $(\eta_n)_{n\geq 0}$ une suite de variables aléatoires i.i.d. sur $\mathbb{S}_e$ de loi γ.
Pour $(x_0, v_0) \in K\times\mathbb{S}^1$, le billard stochastique $((X_t,V_t))_{t\geq 0}$ évoluant dans $K\times\mathbb{S}^1$ peut être construit de la manière suivante :

- Si $x_0 \in K\setminus\partial K$, on définit $T_0 = \inf\{t > 0 : x_0 + tv_0 \notin K\}$. Pour $t \in [0, T_0)$ on pose alors $X_t = x_0 + tv_0$ et $V_t = V_0$. Si $x_0 \in \partial K$ on définit $T_0 = 0$.
- On pose $X_{T_0} = x_0 + T_0 v_0$, et $V_{T_0} = U_{X_{T_0}}\eta_0$.
- Soit $\tau_1 = \inf\{t > 0 : X_{T_0} + tV_{T_0} \notin K\}$ et $T_1 = T_0 + \tau_1$. On pose $X_t = X_{T_0} + tV_{T_0}$, $V_t = V_{T_0}$ pour $t \in [T_0, T_1)$, et $X_{T_1} = X_{T_0} + \tau_1 V_{T_0}$. Puis, on définit $V_{T_1} = U_{X_{T_1}}\eta_1$.
- Soit $\tau_2 = \inf\{t > 0 : X_{T_1} + tV_{T_1} \notin K\}$ et $T_2 = T_1 + \tau_2$. On pose $X_t = X_{T_1} + tV_{T_1}$, $V_t = V_{T_1}$ pour $t \in [T_1, T_2)$, et $X_{T_2} = X_{T_1} + \tau_2 V_{T_1}$. La vitesse à l'instant T_2 est alors définie par $V_{T_2} = U_{X_{T_2}}\eta_2$.
- Et ainsi de suite ...

Le générateur infinitésimal du processus $((X_t,V_t))_{t\geq 0}$ est alors donné par : pour $(x,v) \in K\times\mathbb{S}^1$,

$$Lf(x,v) = v\cdot\nabla_x f(x,v)$$

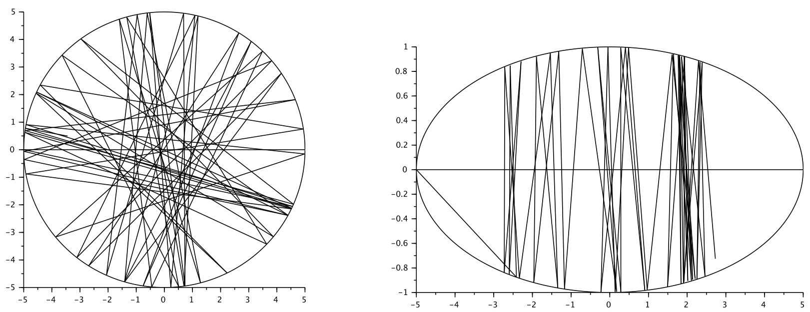


Figure 1.2 – Évolution du billard stochastique dans un cercle et dans une ellipse, avec pour loi de réflexion γ la loi uniforme sur l'ensemble $\{v \in \mathbb{S}^1 \ : \ \langle v, e\rangle \geq \cos\left(\frac{\pi}{8}\right)\} \subset \mathbb{S}_e$.

pour $f \in \mathcal{C}^{1,0}(K \times \mathbb{S}^1)$ vérifiant

$$f(x,v) = \int_{\{y \in \mathbb{S}^1 \ : \ y \cdot e \geq 0\}} f(x, U_x u)\gamma(u)\mathrm{d}u, \quad \forall\, x \in \partial K, \ v \cdot n_x \leq 0.$$

Ce processus est donc un PDMP vivant dans $K \times \mathbb{S}^1$. Il sera l'objet d'étude du Chapitre 2 de ce manuscrit, et nous nous intéresserons en particulier à sa vitesse de convergence à l'équilibre dans des ensembles K assez simples.
On pourra observer sur la Figure 1.2 à gauche une trajectoire du billard stochastique jusqu'au temps $t = 500$ dans le disque du plan centré en l'origine et de rayon 5, avec pour position initiale le point $x_0 = (5, 0)$ et pour vitesse initiale $v_0 = (-1, 0)$. La trajectoire a été générée avec pour loi de réflexion γ la loi uniforme sur l'ensemble $\{v \in \mathbb{S}^1 \ : \ \langle v, e\rangle \geq \cos\left(\frac{\pi}{8}\right)\} \subset \mathbb{S}_e$. Le dessin de droite représente quant à lui une trajectoire du billard stochastique dans l'ellipse d'équation $\left(\frac{x}{5}\right)^2 + y^2 = 1$, avec les mêmes conditions initiales et la même loi de réflexion, et jusqu'au temps $t = 100$.
Sur la Figure 1.3 on a représenté les positions au temps $t = 100$ de 100 billards stochastiques avec mêmes conditions initiales et même loi de réflexion que ci-dessus, à gauche dans le disque, et à droite dans l'ellipse.
On observe sur ces deux figures que le billard stochastique dans le cercle semble mieux explorer l'espace que celui dans l'ellipse. Il est en fait facile de se convaincre que, pour une loi de réflexion donnée, plus le domaine est « courbé », plus il sera difficile pour le billard stochastique d'atteindre les régions du domaine à fortes courbures. De plus, on peut également aisément conjecturer que plus la loi de réflexion charge une grande partie de la demi-sphère $\mathbb{S}_e$, plus le processus va explorer le domaine dans lequel il évolue. Nous retrouverons ces résultats dans le Chapitre 2.

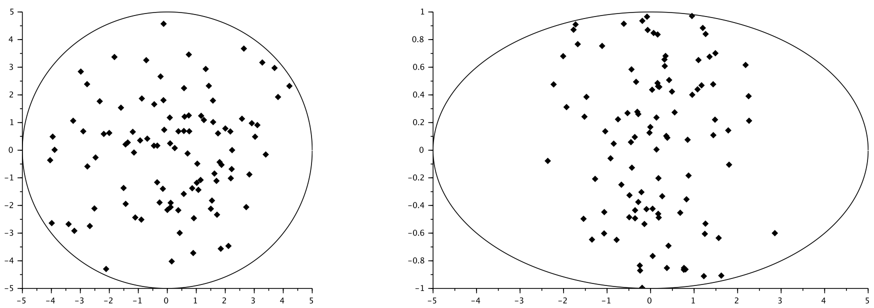


FIGURE 1.3 – Positions de 100 billards stochastiques au temps $t = 100$ dans un cercle et dans une ellipse, avec pour loi de réflexion γ la loi uniforme sur l'ensemble $\{v \in \mathbb{S}^1 \ : \ \langle v, e\rangle \geq \cos\left(\frac{\pi}{8}\right)\} \subset \mathbb{S}_e$.

1.2 Comportement en temps long

1.2.1 Couplage

Dans cette thèse, notre principal objectif est d'étudier la convergence de certains PDMPs vers leur mesure invariante, lorsque celle-ci existe et est unique, et d'en estimer la vitesse si c'est possible. Pour cela, on introduit la notion de distance en variation totale, qui est la distance que l'on considérera dans la suite.

Soient $\nu, \tilde{\nu} \in \mathcal{M}(E)$. La distance en variation totale entre ν et $\widetilde{\nu}$, notée $\|\nu - \widetilde{\nu}\|_{TV}$, est définie par

$$\|\nu - \widetilde{\nu}\|_{TV} = \sup_{A \in \mathcal{B}(\mathbb{R}^d)} \{|\nu(A) - \tilde{\nu}(A)|\} = \frac{1}{2} \sup_{\|\varphi\|_\infty \leq 1} \{|\nu(\varphi) - \tilde{\nu}(\varphi)|\}.$$

Il existe de nombreuses définitions équivalentes de la distance en variation totale (voir [Lindvall, 2002]). L'égalité ci-dessus entre les deux suprema est relativement facile à démontrer. Cependant, ces définitions ne sont pas très utiles en pratique pour notre usage. Dans la proposition suivante, nous allons donner d'autres écritures de la distance en variation totale entre deux mesures, qui seront plus adaptées pour nos études.
Un couplage de ν et $\widetilde{\nu}$ est une mesure de probabilité sur $E \times E$ dont la première (resp. deuxième) marginale est ν (resp. $\widetilde{\nu}$), ou de manière équivalente, un couple $(X, \widetilde{X})$ de variables aléatoires tel que $X \sim \nu$ et $\widetilde{X} \sim \widetilde{\nu}$.

Proposition 1.2.1. *Soient $\nu, \tilde{\nu} \in \mathcal{M}(E)$. Notons $f, \tilde{f}$ leurs densités respectives par rapport à une mesure μ. Alors on a :*

$$\|\nu - \tilde{\nu}\|_{TV} = \inf_{X \sim \nu, \tilde{X} \sim \tilde{\nu}} \mathbb{P}(X \neq \tilde{X}) = 1 - \int (f \wedge \tilde{f}) \mathrm{d}\mu = \frac{1}{2} \int |f - \tilde{f}| \mathrm{d}\mu. \quad (1.2)$$

Un couplage $(X, \tilde{X})$ qui réalise l'infimum dans (1.2) est appelé un couplage maximal (puisqu'il maximise la probabilité $\mathbb{P}(X = \tilde{X})$).

Démonstration. Remarquons d'abord que l'égalité

$$1 - \int (f \wedge \tilde{f})\mathrm{d}\mu = \frac{1}{2}\int |f - \tilde{f}|\mathrm{d}\mu$$

est facile à démontrer, puisque f et $\tilde{f}$ sont des densités, donc d'intégrale 1. Nous n'écrivons pas les détails de ce calcul.
Notons $p = \int (f \wedge \tilde{f})\mathrm{d}\mu$ et $A = \{f \leq \tilde{f}\}$. On a :

$$|\nu(A) - \tilde{\nu}(A)| = \int_A (\tilde{f} - f)\mathrm{d}\mu = 1 - p,$$

et on obtient ainsi directement l'inégalité suivante :

$$\|\nu - \tilde{\nu}\|_{TV} \geq 1 - p.$$

D'autre part, si $X \sim \nu$, $\tilde{X} \sim \tilde{\nu}$ et $B \in \mathcal{B}(\mathbb{R}^d)$ alors

$$\begin{aligned}
|\nu(B) - \tilde{\nu}(B)| &= |\mathbb{P}(X \in B) - \mathbb{P}(\tilde{X} \in B)| \\
&= |\mathbb{P}(X \in B, X \neq \tilde{X}) + \mathbb{P}(X \in B, X = \tilde{X}) \\
&\qquad - \mathbb{P}(\tilde{X} \in B, X \neq \tilde{X}) - \mathbb{P}(\tilde{X} \in B, X = \tilde{X})| \\
&= |\mathbb{P}(X \in B, X \neq \tilde{X}) - \mathbb{P}(\tilde{X} \in B, X \neq \tilde{X})| \\
&\leq \mathbb{P}(X \neq \tilde{X}).
\end{aligned}$$

Ainsi on a finalement

$$1 - p \leq \|\nu - \tilde{\nu}\|_{TV} \leq \inf_{X \sim \nu, \tilde{X} \sim \tilde{\nu}} \mathbb{P}(X \neq \tilde{X}).$$

Il suffit alors d'exhiber un couplage $(X, \tilde{X})$ de ν et $\tilde{\nu}$ tel que $\mathbb{P}(X \neq \tilde{X}) \leq 1 - p$ pour terminer la preuve. Pour cela, soit Y une variable aléatoire de loi de Bernoulli de paramètre p. Alors :

- si $Y = 1$ (avec probabilité p donc), on tire X selon la loi $\frac{f \wedge \tilde{f}}{p}\mu$ et on pose $\tilde{X} = X$;
- si $Y = 0$, on tire X selon la loi $\frac{f - f \wedge \tilde{f}}{1-p}\mu$ et $\tilde{X}$ selon la loi $\frac{\tilde{f} - f \wedge \tilde{f}}{1-p}\mu$.

Vérifions que X et $\tilde{X}$ ainsi construits forment bien un couplage de ν et $\tilde{\nu}$. Soit $B \in \mathcal{B}(\mathbb{R}^d)$, on a :

$$\begin{aligned}
\mathbb{P}(X \in B) &= \mathbb{P}(X \in B, Y = 1) + \mathbb{P}(X \in B, Y = 0) \\
&= p \int_B \frac{f \wedge \tilde{f}}{p}\mathrm{d}\mu + (1 - p)\int_B \frac{f - f \wedge \tilde{f}}{1 - p}\mathrm{d}\mu \\
&= \int_B f \mathrm{d}\mu
\end{aligned}$$

$$= \nu(B),$$

et de même, $\mathbb{P}(\tilde{X} \in B) = \tilde{\nu}(B)$, donc $(X, \tilde{X})$ est bien un couplage de ν et $\tilde{\nu}$. Enfin, puisque sur l'événement $Y = 1$ on a posé $X = \tilde{X}$, on a

$$\mathbb{P}(X \neq \tilde{X}) \leq \mathbb{P}(Y \neq 1) = 1 - p,$$

ce qui conclut la preuve. □

La convergence en distance en variation totale implique la convergence en loi, mais la réciproque n'est pas vraie en général. En effet, dans $\mathbb{R}$, alors que $(\delta_{1/t})_{t\geq 0}$ converge en loi vers δ_0, la convergence en variation totale n'a pas lieu puisqu'on a $\|\delta_{1/t} - \delta_0\|_{TV} = 1$ pour tout $t \geq 0$, car ces deux mesures sont à supports disjoints. Notons que dans un espace de probabilité fini ou dénombrable, l'équivalence entre la convergence en loi et la convergence en variation totale est par contre bien vérifiée.

Enfin, terminons cette partie en évoquant comment la première égalité dans (1.2) peut permettre d'obtenir des vitesses de convergence d'un processus de Markov vers sa mesure invariante.
Les méthodes de couplage reposent sur l'idée suivante : considérons deux processus de Markov $(X_t)_{t\geq 0}$ et $(\tilde{X}_t)_{t\geq 0}$ suivant la même dynamique. Pour obtenir une borne sur la distance en variation totale entre les lois de ces deux processus partant d'états initiaux différents, on essaie de les construire jusqu'à leur temps de couplage $T_c = \inf \left\{ t \geq 0 : X_t = \tilde{X}_t \right\}$. Alors, ces deux processus étant markovien et suivant la même dynamique, on peut les construire de façon à ce qu'ils restent égaux après le temps T_c. La première égalité de (1.2), donne alors immédiatement :

$$\|\mathcal{L}(X_t) - \mathcal{L}(\tilde{X}_t)\|_{TV} \leq \mathbb{P}\left(T_c > t\right).$$

Soit T^* une variable aléatoire stochastiquement plus grande que T_c, $T_c \leq_{st} T^*$, ce qui signifie que $\mathbb{P}\left(T_c \leq t\right) \geq \mathbb{P}\left(T^* \leq t\right)$ pour tout $t \geq 0$. Si T^* possède un moment exponentiel fini, l'inégalité de Markov donne alors, pour tout λ tel que la transformée de Laplace de T^* est bien définie :

$$\|\mathcal{L}(X_t) - \mathcal{L}(\tilde{X}_t)\|_{TV} \leq \mathbb{P}\left(T^* > t\right) \leq \mathrm{e}^{-\lambda t}\mathbb{E}\left[\mathrm{e}^{\lambda T^*}\right].$$

Ainsi, pour montrer qu'un processus de Markov, dont on sait qu'il possède une unique mesure de probabilité invariante, converge vers cette mesure, on peut procéder de la manière suivante : on considère deux processus de Markov de même générateur mais avec des conditions initiales différentes. Si l'on réussit à les construire de façon à ce que leur temps de couplage ait un moment exponentiel fini, alors on en déduit la convergence à vitesse exponentielle du processus vers sa mesure invariante. De plus, si l'on sait estimer les moments exponentiels du temps de couplage, alors cela fournit une borne pour la vitesse de convergence. Ainsi, l'intérêt des méthodes de couplage réside dans le choix ou la construction du meilleur couplage des lois que l'on considère, pour obtenir les meilleures bornes possibles sur la distance en variation totale considérée. Cependant, en règle générale il n'est pas forcément aisé de construire un couplage explicite dont on peut estimer le temps de couplage. Ainsi, nous introduisons dans la partie suivante des résultats généraux qui permettent de montrer la convergence à vitesse exponentielle d'un processus de Markov à l'équilibre.

1.2.2 Approche de Meyn et Tweedie

Dans cette partie, on donne des conditions pour obtenir la convergence à vitesse exponentielle d'un processus de Markov vers sa mesure invariante en distance en variation totale (les notions et résultats de cette section se trouvent dans [Meyn et Tweedie, 1993a, Meyn et Tweedie, 1993b, Down *et al.*, 1995]).
Un processus $(X_t)_{t\geq 0}$ est dit f-exponentiellement ergodique, où f est une fonction mesurable de E dans $[1,+\infty[$, s'il existe une mesure de probabilité π, une fonction $M : E \longrightarrow \mathbb{R}^+$ et $0 < \rho < 1$ tels que

$$\|\mathcal{L}(X_t|X_0 = x) - \pi\|_f \leq M(x)\rho^t, \quad \text{pour } t \geq 0;$$

où $\|\cdot\|_f$ est définie pour toute mesure signée μ par $\|\mu\|_f = \sup_{|g|\leq f} \left|\int g(y)\mu(\mathrm{d}y)\right|$.
Le processus X est dit exponentiellement ergodique s'il existe $f \geq 1$ telle que X est f-exponentiellement ergodique.
Afin de montrer l'ergocité exponentielle d'un processus de Markov $(X_t)_{t\geq 0}$, on peut utiliser le critère de Foster-Lyapunov qui consiste en l'exhibition d'une fonction de Lyapunov associée à un ensemble petit pour le processus étudié $(X_t)_{t\geq 0}$. Nous décrivons cette méthode ci-dessous.
Un ensemble $K \subset \mathbb{R}^d$ est dit petit pour le processus $(X_t)_{t\geq 0}$ s'il existe une mesure de probabilité ν sur $\mathbb{R}^+$ et une mesure positive non triviale μ sur $\mathbb{R}^d$ telles que, pour tout $x \in K$, $\int_0^\infty P_t(x,\cdot)\nu(\mathrm{d}t) \geq \mu(\cdot)$. La Proposition 1.2.3 ainsi que la discussion à la fin de cette section permettent de mieux comprendre cette notion.
Si $K \subset \mathbb{R}^d$ est un ensemble compact petit pour le processus $(X_t)_{t\geq 0}$, alors une fonction $H : \mathbb{R}^d \to \mathbb{R}$ est une fonction de Lyapunov associée à l'ensemble K pour le processus $(X_t)_{t\geq 0}$ si $H(x) \geq 1$ pour tout $x \in \mathbb{R}^d$, et s'il existe des constantes $\alpha > 0$ et $\beta \geq 0$ telles que pour tout $x \in \mathbb{R}^d$,

$$LH(x) \leq -\alpha H(x) + \beta \mathbf{1}_K(x). \tag{1.3}$$

Le résultat suivant donne alors l'ergocité exponentielle d'un processus sous l'existence d'une telle fonction.

Théorème 1.2.1. *[Théorème 5.2 de [Down et al., 1995]] Soit $(X_t)_{t\geq 0}$ un processus de Markov irréductible et apériodique (voir [Down et al., 1995] pour la définition précise de ces notions).*
S'il existe un ensemble petit $K \subset \mathbb{R}^d$ et une fonction de Lyapunov associée à l'ensemble K pour le processus $(X_t)_{t\geq 0}$, alors X est exponentiellement ergodique.

Exemple 1.2.2. (voir [Bierkens et Roberts, 2017, Fontbona *et al.*, 2016]) Revenons sur l'exemple du processus Zig-zag introduit dans l'Exemple 1.1.1. Nous rappelons qu'il s'agit du processus $((X_t, V_t))_{t\geq 0}$ de générateur donné par : pour $(x, v) \in \mathbb{R} \times \{-1, +1\}$,

$$Lf(x, v) = v\partial_x f(x, v) + (a\mathbf{1}_{xv<0} + b\mathbf{1}_{xv\geq 0}) (f(x, -v) - f(x, v)),$$

avec $0 < a < b$.
Alors il existe $x_1 > 0$ et $\alpha, \beta > 0$ tels que la fonction définie pour $|x|\geq x_1$ et $v \in \{-1, +1\}$ par $H(x, v) = \mathrm{e}^{\alpha|x|+\beta \mathrm{sgn}(xv)}$ est une fonction de Lyapunov pour le générateur L associée au compact $K = [-x_1, x_1] \times \{-1, +1\}$.
En effet, on choisit x_1 assez grand de façon à ce qu'on puisse construire H positive et de classe $\mathcal{C}^1$ en sa première variable sur $\mathbb{R}$. De plus, soient $\alpha, \beta > 0$ tels que $a\left(\mathrm{e}^{2\beta} - 1\right) < \alpha < b\left(1 - \mathrm{e}^{-2\beta}\right)$. On a alors, pour $x \geq x_1$,

$$LH(x, 1) = \left(\alpha - b\left(1 - \mathrm{e}^{-2\beta}\right)\right) H(x, 1)$$

et

$$LH(x, -1) = \left(-\alpha + a\left(\mathrm{e}^{2\beta} - 1\right)\right) H(x, -1).$$

De plus pour $x \leq -x_1$ on a

$$LH(x, 1) = \left(-\alpha + a\left(\mathrm{e}^{2\beta} - 1\right)\right) H(x, 1)$$

et

$$LH(x, -1) = \left(\alpha - b\left(1 - \mathrm{e}^{-2\beta}\right)\right) H(x, -1).$$

Donc par construction de α et β, il existe bien $\eta > 0$ tel que pour $|x|\geq x_1$ et $v \in \{-1, +1\}$, $LH(x, v) \leq -\eta H(x, v)$. Et pour $x \in [-x_1, x_1]$, $LH(x, v)$ est bornée. Ainsi H est bien une fonction de Lyapunov pour le générateur L.
De plus, pour tout $x_1 > 0$, on peut montrer que l'ensemble K est un ensemble petit pour le processus Zig-zag, qui est irréductible et apériodique. On en déduit donc l'exponentielle ergodicité de ce processus.
Enfin, on a même l'expression explicite de la mesure invariante du processus Zig-zag :

$$\mu(\mathrm{d}x, \mathrm{d}v) = \frac{b - a}{2}\mathrm{e}^{-(b-a)|x|} \otimes \frac{1}{2}\left(\delta_{-1} + \delta_{+1}\right)(\mathrm{d}v).$$

Pour le voir, il suffit de vérifier que pour toute fonction φ de classe $\mathcal{C}^{1,0}$ sur $\mathbb{R} \times \{-1, +1\}$ et à support compact, on a $\int L\varphi(x, v)\mu(\mathrm{d}x, \mathrm{d}v) = 0$.

Mentionnons maintenant un résultat similaire au Théorème 1.2.1, permettant également d'obtenir l'ergodicité exponentielle d'un processus, sans avoir besoin d'expliciter une fonction de Lyapunov, mais reposant sur l'étude du temps de retour dans un ensemble petit. Ce résultat se démontre par l'introduction d'une fonction qui va jouer le rôle de fonction de Lyapunov pour une chaîne associée au processus de Markov étudié.

Théorème 1.2.2. *(Théorème 6.2 de [Down et al., 1995]) Soit $(X_t)_{t\geq 0}$ un processus de Markov irréductible et apériodique. Supposons qu'il existe une fonction $f \geq 1$, un ensemble fermé $C \subset \mathbb{R}^d$ et des constantes $\delta, \eta > 0$, $M < \infty$ tels que*

$$\mathbb{E}_x\left[\int_0^{\tau_C(\delta)} \mathrm{e}^{\eta t} f(X_t)\mathrm{d}t\right] < \infty, \quad \textit{pour tout } x \notin C$$

et

$$\sup_{x\in C} \mathbb{E}_x\left[\int_0^{\tau_C(\delta)} \mathrm{e}^{\eta t} f(X_t)\mathrm{d}t\right] \leq M$$

où $\tau_C(\delta) = \inf\{t \geq \delta, X_t \in C\}$.
Si l'ensemble C est petit pour $(X_t)_{t\geq 0}$, alors le processus est exponentiellement ergodique.

Nous ne donnons pas ici de preuves pour les deux théorèmes précédents, mais nous donnons deux résultats permettant de se faire une idée de la façon dont fonctionnent les choses.
Dans un premier temps, remarquons le résultat suivant sur les ensembles petits :

Proposition 1.2.3. *Si l'espace d'état $\mathbb{R}^d$ du processus de Markov $(X_t)_{t\geq 0}$ est un ensemble petit, alors le processus $(X_t)_{t\geq 0}$ converge exponentiellement vite en variation totale.*

Démonstration. Nous donnons ici une preuve dans un cas particulier des ensembles petits : le cas où $\mathbb{R}^d$ est un « small set », c'est-à-dire qu'il existe $\varepsilon > 0$, $t_0 > 0$ et μ une mesure de probabilité sur $\mathbb{R}^d$ tels que pour tout $x \in \mathbb{R}^d$, $\mathbb{P}_x(X_{t_0} \in \cdot) \geq \varepsilon\mu(\cdot)$.
Sous cette hypothèse, en notant P le semi-groupe du processus que l'on considère, pour tout couple $(x, y) \in E^2$, il existe un couplage (X, X') des mesures $\delta_x P_{t_0}$ et $\delta_y P_{t_0}$ tel que

$$\mathbb{P}(X = X') \geq \varepsilon.$$

Cela signifie qu'en un temps t_0, on peut réussir à coller deux copies du processus partant de deux états x et y, avec une probabilité au moins ε. Si ce couplage échoue, on réessaie de coller nos processus sur l'intervalle $[t_0, 2t_0[$, et ainsi de suite. En itérant on peut ainsi construire un couplage (X_n, X'_n) des mesures $\delta_x P_{nt_0}$ et $\delta_y P_{nt_0}$ tel que

$$\mathbb{P}(X_n \neq X'_n) \leq (1-\varepsilon)^n.$$

On obtient ainsi la convergence exponentielle du processus de semi-groupe P en distance en variation totale. □

Voyons maintenant le rôle d'une fonction de Lyapunov à travers le résultat suivant.

Proposition 1.2.4. *Soit $(X_t)_{t\geq 0}$ un processus de Markov. Soit K un compact, et soit τ_K le temps d'atteinte de K : $\tau_K = \inf\{t \geq 0 : X_t \in K\}$.*
Si H est une fonction de Lyapunov associée à l'ensemble K, i.e. H satisfait (1.3), alors τ_K a des moments exponentiels. En particulier :

$$\mathbb{E}_x\left[e^{\alpha\tau_K}\right] \leq H(x).$$

Démonstration. L'idée de la preuve est la suivante : d'après l'inégalité (1.3) vérifiée par la fonction de Lyapunov, $H(X_t)$ décroit exponentiellement vite avec un taux α tant que X_t n'est pas dans K. Mais H étant minorée par 1, $H(X_t)$ ne peut pas décroitre sans cesse, donc X_t va atteindre l'ensemble K « rapidement ».
Démontrons maintenant ce résultat rigoureusement. Considérons le processus $(\widetilde{X})_{t\geq 0} = (t \wedge \tau_K, X_{t\wedge\tau_K})_{t\geq 0}$. Notons $\widetilde{P}$ son semi-groupe et $\widetilde{L}$ son générateur. Soit $f : \mathbb{R}^+ \times E \to \mathbb{R}$ une fonction mesurable bornée. D'après la propriété de Markov on a

$$\widetilde{P}_s f(t,x) = \widetilde{P}_{s-t} f(0,x),$$

donc il suffit de calculer le terme de droite pour connaître le semi-groupe $\widetilde{P}$.
Soient alors $x \in E$ et $s \geq 0$:

$$\begin{aligned}\widetilde{P}_s f(0,x) &= \mathbb{E}_{0,x}\left[f\left(s\wedge\tau_K, X_{s\wedge\tau_K}\right)\right]\\ &= \mathbb{E}_{0,x}\left[f\left(\tau_K, X_{\tau_K}\right)\mathbf{1}_{s>\tau_K}\right] + \mathbb{E}_{0,x}\left[f(s,X_s)\mathbf{1}_{s\leq\tau_K}\right].\end{aligned}$$

Donc

$$\begin{aligned}\widetilde{L}f(0,x) &= \lim_{s\to 0}\frac{\widetilde{P}_s f(0,x) - f(0,x)}{s}\\ &= \lim_{s\to 0}\frac{\mathbb{E}_{0,x}\left[f(s,X_s)\mathbf{1}_{s\leq\tau_K}\right] - f(0,x)}{s}.\end{aligned}$$

Puis en utilisant le développement de Taylor de f en sa première variable on obtient :

$$\begin{aligned}\widetilde{L}f(0,x) &= \lim_{s\to 0}\frac{\mathbb{E}_{0,x}\left[\left(f(0,X_s) + s\frac{\partial f}{\partial t}(0,X_s)\right)\mathbf{1}_{s<\tau_K}\right] - f(0,x)}{s}\\ &= \left(Lf_0(x) + \frac{\partial f}{\partial t}(0,x)\right)\mathbf{1}_{x\notin K},\end{aligned}$$

où on note $f_t : x \longmapsto f(t,x)$.
Ainsi, grâce à la propriété de Markov, l'expression générale du générateur $\widetilde{L}$ est :

$$\widetilde{L}f(t,x) = \left(\frac{\partial f}{\partial t}(t,x) + Lf_t(x)\right)\mathbf{1}_{x\notin K}.$$

Appliquant cette relation à $f(t,x) = e^{\alpha t}H(x)$ avec H la fonction de Lyapunov on en déduit :

$$\widetilde{L}f(t,x) = e^{\alpha t}\left(\alpha V(x) + LH(x)\right)\mathbf{1}_{x\notin K} \leq 0.$$

Alors, la formule de Dynkin appliquée au processus $\widetilde{X}$ donne

$$\mathbb{E}_x\left[f(\widetilde{X}_t)\right] = f(0,x) + \int_0^t \widetilde{P}_s\widetilde{L}f(0,x)\mathrm{d}s \leq f(0,x),$$

c'est-à-dire

$$\mathbb{E}_x\left[\mathrm{e}^{\alpha t\wedge\tau_K}H(X_{t\wedge\tau_K})\right] \leq H(x).$$

En utilisant le fait que H est minorée par 1, et en faisant tendre t vers l'infini grâce au théorème de convergence dominée on obtient finalement le résultat recherché :

$$\mathbb{E}_x\left[\mathrm{e}^{\alpha\tau_K}\right] \leq H(x).$$

□

Avec ces deux résultats, on peut alors se convaincre du Théorème 1.2.1. En effet, soit un processus de Markov vérifiant les hypothèses de ce théorème, on peut construire un couplage $\left((X_t,\tilde{X}_t)\right)_{t\geq 0}$ selon l'idée suivante :

— partant de lois initiales μ_0 et $\tilde{\mu}_0$, l'existence de la fonction de Lyapunov associée à l'ensemble K assure que les temps d'atteinte de l'ensemble K par les processus X et $\tilde{X}$ sont finis ;

— une fois que les deux processus sont à un même instant dans l'ensemble K, K étant petit, avec une certaine probabilité on arrive à coupler X et $\tilde{X}$. Cette probabilité dépend de la mesure μ dans la définition d'un ensemble petit, et le temps de couplage est lui donné par la loi ν ;

— si le couplage échoue, alors on recommence : on attend que les deux processus soient au même moment dans l'ensemble K, puis on essaie de les coupler.

De cette manière on a construit un couplage de $\left((X_t,\tilde{X}_t)\right)_{t\geq 0}$, dont on pourrait montrer que le temps de couplage admet des moments exponentiels. D'où l'ergodicité exponentielle obtenue dans le Théorème 1.2.1.

1.2.3 PDMDs et échantionnage

En dehors de l'utilité des PDMPs pour la modélisation de phénomènes biologiques ou autres, un des aspects qui explique l'attrait fort des PDMPs est leur intérêt dans les problèmes d'échantillonnage, les algorithmes MCMC.
Imaginons que l'on veuille simuler une loi de densité $\pi(x)$ sur $\mathbb{R}^d$ proportionnelle à $e^{-U(x)}$, où U est un certain potentiel. Les méthodes MCMC pour simuler une approximation de la loi π reposent sur l'introduction d'une chaîne de Markov ergodique et de loi invariante π. L'étude de la vitesse de convergence de cette chaîne de Markov donne alors une indication sur la vitesse de convergence de l'algorithme sous-jacent.
Alors que la plupart des méthodes MCMC sont basées sur des chaînes de Markov réversibles, il semblerait que des chaînes non réversibles donnent de meilleures vitesses de convergence. En effet, alors que les processus réversibles peuvent

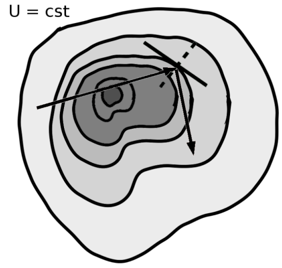


FIGURE 1.4 – Lors d'un saut du Bouncy Particle Sampler, la vitesse rebondit sur les lignes de niveau de U. Cette image provient du papier [Monmarché, 2016].

« revenir sur leurs pas », la non-réversibilité ne le permet pas, et l'exploration de l'espace est alors plus rapide, donnant une meilleure vitesse de convergence. Une classe de processus non-réversibles aujourd'hui beaucoup utilisés pour simuler la loi π est la classe des PDMPs sous forme d'un couple position-vitesse, et dont la loi limite a pour première marginale la loi π. En plus de leur apparente meilleure vitesse de convergence que d'autre processus utilisés pour les méthodes MCMC, ces PDMPs ont pour avantage la possibilité d'être simulés de manière exacte sous de bonnes conditions sur U (voir par exemple [Lemaire *et al.*, 2018]). De plus, considérant des PDMPs dont la deuxième composante est la vitesse de la première, la partie déterministe du processus ne requiert aucune approximation (on pourra voir [De Saporta *et al.*, 2015] pour des méthodes d'approximations de PDMPs dont l'évolution déterministe est générale, et qui requiert donc parfois une approximation pour être simulée).
De manière générale (dans tous les travaux cités ci-dessous sur ce sujet), si la loi cible est π comme ci-dessus, le taux de saut du processus $((X_t, V_t))_{t\geq 0}$ utilisé pour échantillonner π est choisi de la forme $\lambda(x, v) = c + (\langle v, \nabla_x U(x)\rangle)_+$, pour une constante $c \geq 0$, de façon à ce que la première marginale de la loi limite du processus soit π. La vitesse, quant à elle, a une évolution qui varie selon les modèles.
Commençons par citer les travaux [Bierkens et Roberts, 2017] et [Bierkens *et al.*, 2019], dans lesquels les auteurs s'intéressent au processus Zig-zag, c'est-à-dire avec une vitesse dans $\{-1, +1\}^d$, dans le cas particulier où le processus peut être vu comme le produit de d processus Zig-zag en dimension 1. Dans [Bierkens *et al.*, 2017], les résultats précédents sont généralisés puisque les auteurs considèrent un processus Zig-zag en dimension d sous forme très générale. Un autre PDMP utilisé pour l'échantillonnage est le « Bouncy Particle Sampler », étudié dans [Bouchard-Côté *et al.*, 2018, Deligiannidis *et al.*, 2019, Monmarché, 2016], à valeurs dans $\mathbb{R}^d \times \mathbb{R}^d$ ou $\mathbb{R}^d \times \mathbb{S}^1$. Pour ce processus, le mécanisme de saut est le suivant : lors d'un saut, la vitesse est réfléchie suivant la loi optique le long de la ligne de niveau de U que la position a atteinte (voir Figure 1.4). Enfin, dans [Wu et Robert, 2018], les auteurs introduisent le

« coordinate sampler », qui est une variante du processus Zig-zag : la vitesse du processus vit dans la base canonique de $\mathbb{R}^d$, et lors d'un saut, une composante de la vitesse est switchée. Contrairement au processus Zig-zag, entre deux sauts de vitesse, seule une composante de la position évolue, alors que les autres restent inactives.
Des comparaisons de l'efficacité des différents algorithmes sous-jacents à ces processus sont faites dans [Andrieu *et al.*, 2018, Wu et Robert, 2018].

1.3 Systèmes de particules en interaction

Dans le Chapitre 4 de ce manuscrit, nous nous intéressons à un système de particules en interaction. Nous introduisons donc dans cette partie la notion de propagation du chaos, centrale dans l'étude de systèmes de particules en interaction. Pour plus de détails, on pourra voir le chapitre « Topics in propagation of chaos » de Sznitman dans le livre [Burkholder *et al.*, 1991].
Soit E un espace métrique. Si $\mu \in \mathcal{M}(E)$ et $f \in \mathcal{C}_b(E)$, on définit $\langle \mu, f \rangle = \int f \mathrm{d}\mu$. On dit qu'une mesure $\mu_N \in \mathcal{M}(E^N)$ est symétrique si pour toute permutation $\sigma \in \mathfrak{S}_N$ et toutes fonctions $f_1, \cdots, f_N \in \mathcal{C}_b(E)$ on a $\langle \mu_N, f_1 \otimes \cdots \otimes f_N \rangle = \langle \mu_N, f_{\sigma(1)} \otimes \cdots \otimes f_{\sigma(N)} \rangle$.
Soit $(u_N)_{N \geq 1}$ une suite de mesures de probabilité symétriques sur E^N. Soit u une mesure de probabilité sur E. On dit que $(u_N)_{N \geq 1}$ est u-chaotique si pour tous $\phi_1, \cdots, \phi_k \in C_b(E)$ et $k \geq 1$ on a

$$\lim_{N \to \infty} \langle u_N, \phi_1 \otimes \cdots \otimes \phi_k \otimes 1 \cdots \otimes 1 \rangle = \prod_{i=1}^{k} \langle u, \phi_i \rangle. \tag{1.4}$$

La notion de u-chaotique signifie que les mesures empiriques des variables coordonnées de E^N, sous u_N, tendent à se concentrer autour de u. C'est donc une sorte de loi des grands nombres. Nous observons cela dans la proposition suivante :

Proposition 1.3.1 ([Burkholder *et al.*, 1991]).

1. *(u_N) est u-chaotique est équivalent au fait que $\bar{X}_N = \frac{1}{N} \sum_{i=1}^{N} \delta_{X_i}$ converge en loi vers u. C'est aussi équivalent à la condition (1.4) avec $k = 2$.*
2. *Si E est un espace polonais, la suite de mesures $(\bar{X}_N)_N$ est tendue si et seulement si les lois sur E de X_1 sous u_N sont tendues.*

En général, on dispose d'une loi symétrique P_N sur $C(\mathbb{R}^+, \mathbb{R}^d)^N$ par exemple, décrivant un système de particules en interaction. Les conditions initiales sont supposées être u_0-chaotiques, où u_0 est une mesure de probabilité sur $\mathbb{R}^d$. Dans le contexte de propagation du chaos, le but est alors de montrer que la suite $(P_N)_{N \geq 1}$ est P-chaotique, pour une certaine loi P sur $C(\mathbb{R}^+, \mathbb{R}^d)^N$. La loi P est alors la loi d'un processus solution d'une certaine EDS, limite du système de particules. Cette EDS est en général non-linéaire : l'interaction entre les particules se traduit par une dépendance du processus limite en sa propre loi.

Un processus non-linéaire, c'est-à-dire solution d'une EDS faisant intervenir sa loi, n'est en général pas facile à simuler. On peut alors utiliser un système de particules pour lequel il y aurait propagation du chaos vers ce processus non-linéaire pour le simuler. L'étude de la vitesse de la propagation du chaos permet alors d'estimer l'erreur commise lors de l'approximation du processus par le système de particules.
D'autres questions peuvent également se poser lorsque l'on étudie un système de particules en interaction. On peut par exemple étudier le comportement en temps long du système de particules, avec un nombre de particules N fixé. De plus, lorsqu'il y a propagation du chaos, on peut également s'intéresser au comportement en temps long du processus limite. Enfin, il peut être intéressant de voir si les passages à la limite quand le nombre de particules N tend vers l'infini, et en temps long, sont commutatifs.

Dans la suite de cette partie, nous donnons des exemples, plus ou moins détaillés, d'études de systèmes de particules en interaction et du processus non-linéaire associé, afin de faire ressortir les points importants et les difficultés dans ce contexte.

Exemple 1.3.2. Nous commençons par donner un exemple simple d'étude d'un système de particules en interaction, que l'on peut retrouver dans [Chevallier, 2017], pour illustrer la propagation du chaos, et la méthode classique utilisée pour la démontrer.
On considère un système de N particules décrites par leurs positions $X_t^{1,N}, \cdots, X_t^{N,N}$ pour tout temps t positif, satisfaisant, pour tout $i \in \{1, \cdots, N\}$

$$\mathrm{d}X_t^{i,N} = -\left(X_t^{i,N} - \frac{1}{N}\sum_{k=1}^{N} X_t^{k,N}\right)\mathrm{d}t + \mathrm{d}B_t^i, \tag{1.5}$$

où les $(B_t^i)_{t\geq 0}$ sont des mouvements Browniens indépendants. On suppose que les positions initiales $X_0^{1,N}, \cdots, X_0^{N,N}$ sont indépendantes et identiquement distribuées de loi μ_0, et indépendantes des B^i, $i \in \{1, \cdots, N\}$.
La première chose à observer est qu'il y a existence et unicité trajectorielle des solutions de l'EDS (1.5), puisqu'elle est linéaire.
Lorsque l'on fait tendre le nombre de particules N vers l'infini, on s'attend à un effet « loi des grands nombres », et donc à ce qu'à la limite, le terme $\frac{1}{N}\sum_{k=1}^{N} X_t^{k,N}$ soit remplacé par une espérance. On introduit donc l'EDS suivante :

$$\mathrm{d}\bar{X}_t = -\left(\bar{X}_t - \mathbb{E}[\bar{X}_t]\right)\mathrm{d}t + \mathrm{d}B_t, \tag{1.6}$$

où $(B_t)_{t\geq 0}$ est un mouvement brownien, et on considère la condition initiale $\bar{X}_0$ distribuée selon la loi μ_0.
En règle générale, le caractère bien posé de l'équation limite dans un problème de système de particules en interaction n'est pas trivial. En effet, cette équation fait intervenir la loi du processus solution de l'EDS.
Cependant, dans ce cas particulier, l'EDS se simplifie. En effet, supposons que

μ_0 possède un moment d'ordre 1, et notons alors $m_0 = \mathbb{E}[\bar{X}_0]$. L'espérance de $\bar{X}_t$ étant constante au cours du temps, l'EDS (1.6) est alors équivalente à l'EDS suivante

$$\mathrm{d}\bar{X}_t = -\left(\bar{X}_t - m_0\right)\mathrm{d}t + \mathrm{d}B_t. \tag{1.7}$$

Ainsi, l'existence et l'unicité de la solution de l'équation non-linéaire (1.6) sont immédiates. On peut même expliciter cette solution, qui est en fait un processus d'Ornstein-Uhlenbeck :

$$\bar{X}_t = \bar{X}_0\mathrm{e}^{-t} + m_0(1-\mathrm{e}^{-t}) + \int_0^t \mathrm{e}^{-(t-s)}\mathrm{d}B_s.$$

Voyons maintenant qu'il y a propagation du chaos : à la limite quand $N \to \infty$, les particules sont indépendantes et identiquement distribuées, de même loi que le processus non-linéaire solution de l'EDS (1.6). Pour cela nous allons, de manière classique, coupler les particules en interaction avec des copies i.i.d. du processus non-linéaire. On considère donc

— les processus $(X_t^{i,N})_{t\geq 0}$, $i = 1, \cdots, n$ solutions de

$$\begin{cases} \mathrm{d}X_t^{i,N} & = -\left(X_t^{i,N} - \frac{1}{N}\sum_{k=1}^N X_t^{k,N}\right)\mathrm{d}t + \mathrm{d}B_t^i, \\ X_0^{i,N} & = Y^i \end{cases}$$

— les processus $(\bar{X}_t^i)_{t\geq 0}$, $i \geq 1$ solutions de

$$\begin{cases} \mathrm{d}\bar{X}_t^i & = -\left(\bar{X}_t^i - \mathbb{E}[\bar{X}_t^i]\right)\mathrm{d}t + \mathrm{d}B_t^i, \\ X_0^{i,N} & = Y^i \end{cases}$$

où les conditions initiales Y^i sont i.i.d. de loi μ_0.
Sous l'hypothèse que μ_0 possède un moment d'ordre 2, il y a propagation du chaos en temps fini. Plus précisément, en notant $v_0 = Var(\bar{X}_0)$, on a, pour tout $i \in \{1, \cdots, N\}$ et $T > 0$,

$$\mathbb{E}\left[\sup_{0\leq t\leq T}|X_t^{i,N} - \bar{X}_t^i|\right] \leq \left(v_0 + \frac{1}{2}\right)\frac{T\mathrm{e}^{2T}}{\sqrt{N}}.$$

En effet, les conditions initiales et les mouvements Browniens étant les mêmes pour $X^{i,N}$ et $\bar{X}^i$ on a

$$X_t^{i,N} - \bar{X}_t^i = -\int_0^t \left((X_s^{i,N} - \bar{X}_s^i) - \left(\frac{1}{N}\sum_{k=1}^N X_s^{k,N} - \mathbb{E}\left[\bar{X}_s^i\right]\right)\right)\mathrm{d}s,$$

puis par inégalité triangulaire :

$$\left|X_t^{i,N} - \bar{X}_t^i\right| \\ \leq \int_0^t \left(|X_s^{i,N} - \bar{X}_s^i| + \left|\frac{1}{N}\sum_{k=1}^N (X_s^{k,N} - \bar{X}_s^k)\right| + \left|\frac{1}{N}\sum_{k=1}^N \bar{X}_s^k - \mathbb{E}\left[\bar{X}_s^i\right]\right|\right)\mathrm{d}s.$$

En prenant l'espérance et en notant $\alpha(t) = \mathbb{E}\left[\sup_{0\leq s\leq t}|X_s^{i,N} - \bar{X}_s^i|\right]$, on obtient

$$\begin{aligned}
\alpha(t) &\leq 2\int_0^t \alpha(s)\mathrm{d}s + \int_0^t \mathbb{E}\left[\left|\frac{1}{N}\sum_{k=1}^N \bar{X}_s^k - \mathbb{E}\left[\bar{X}_s^i\right]\right|\right]\mathrm{d}s \\
&\leq 2\int_0^t \alpha(s)\mathrm{d}s + \int_0^t \frac{1}{\sqrt{N}} Var(\bar{X}_s^1)\mathrm{d}s \\
&\leq 2\int_0^t \alpha(s)\mathrm{d}s + \left(v_0 + \frac{1}{2}\right)\frac{t}{\sqrt{N}},
\end{aligned}$$

où on a utilisé l'expression explicite de $\bar{X}^1$ pour voir que $Var(\bar{X}_s^1) \leq \left(v_0 + \frac{1}{2}\right)$.
Le lemme de Gronwall permet alors d'obtenir la borne annoncée pour $\alpha(t)$.
De ce résultat, on peut en déduire la convergence en loi de la trajectoire de $\left(X_t^{i,N}\right)_{0\leq t\leq T}$ vers celle de $\left(\bar{X}_t^i\right)_{0\leq t\leq T}$. De plus, on a également la convergence en probabilité de $\frac{1}{N}\sum_{k=1}^N \delta_{(X_t^{k,N})_{0\leq t\leq T}}$ vers la loi de $\left(\bar{X}_t^i\right)_{0\leq t\leq T}$.

Exemple 1.3.3. Citons maintenant une généralisation de l'exemple ci-dessus, et dans lequel on s'intéresse à la propagation du chaos, mais également au comportement en temps long du système de particules et du processus non-linéaire (voir [Malrieu, 2003]). Nous ne donnons ici que les résultats principaux, sans donner de démonstration.
Soient $(B_\cdot^i)_{i\in\mathbb{N}}$ une suite de mouvements Browniens indépendants et $(X_0^i)_{i\in\mathbb{N}}$ une suite de variables aléatoires indépendantes de loi u_0, et indépendante des B^i, $i \in \mathbb{N}$. Soit $(X_t^N)_{t\geq 0}$ le processus solution de :

$$\begin{cases} \mathrm{d}X_t^{i,N} = \sqrt{2}\mathrm{d}B_t^i - \frac{1}{N}\sum_{j=1}^N \nabla W(X_t^{i,N} - X_t^{j,N})\mathrm{d}t & \text{pour } i = 1, \cdots, N \\ X_0^{i,N} = X_0^i & \text{pour } i = 1, \cdots, N, \end{cases}$$

avec $W : \mathbb{R}^d \to \mathbb{R}$ symétrique (ie pour tout $x \in \mathbb{R}^d$, $W(x) = W(-x)$), uniformément convexe et tel que son gradient est localement Lipschitz avec croissance polynomiale.
Le processus non-linéaire associé est le processus $(\bar{X}_t)_{t\geq 0}$ solution de :

$$\begin{cases} \mathrm{d}\bar{X}_t = \sqrt{2}\mathrm{d}B_t - \nabla W * u_t(\bar{X}_t)\mathrm{d}t \\ \mathcal{L}(\bar{X}_t) = u_t(\mathrm{d}y), \quad \text{pour } t \geq 0, \end{cases} \tag{1.8}$$

où $*$ est l'opérateur de convolution $\nabla W * u(x) = \displaystyle\int \nabla W(x-y)u(\mathrm{d}y)$.

Pour montrer la propagation du chaos, on introduit de manière classique la famille $(\bar{X}_\cdot^i)_{i\in\mathbb{N}^*}$ de processus non-linéaire indépendants définis par :

$$\begin{cases} \mathrm{d}\bar{X}_t^i = \sqrt{2}\mathrm{d}B_t^i - \nabla W * u_t(\bar{X}_t^i)\mathrm{d}t \\ \mathcal{L}(\bar{X}_t^i) = u_t(\mathrm{d}y), \quad \text{pour } t \geq 0 \\ \bar{X}_0^i = X_0^i. \end{cases}$$

On peut alors montrer le résultat suivant : si $\mathbb{E}\left[|X_0|^2\right] < \infty$, alors il existe une constante C telle que pour tout $T \geq 0$ et tout $N \in \mathbb{N}$,

$$\sup_{0\leq t\leq T} \mathbb{E}\left[|X_t^{i,N} - \bar{X}_t^i|^2\right] \leq \frac{CT^2}{N}.$$

La propagation du chaos a donc bien lieu, mais n'est pas uniforme en temps.
Concernant le processus non-linéaire limite, solution de (1.8), on peut montrer qu'il converge à vitesse exponentielle vers l'équilibre.
Par contre, le système de particules lui ne converge pas vers une mesure de probabilité quand t tend vers l'infini. En effet, par symétrie de W, ∇W est impair, et donc la moyenne de la mesure empirique est égale à

$$\frac{1}{N}\sum_{i=1}^{N} X_t^{i,N} = \frac{\sqrt{2}}{N}\sum_{i=1}^{N} B_t^i + \frac{1}{N}\sum_{i=1}^{N} X_0^i.$$

Le premier terme est une variable gaussienne $\sqrt{2}\mathcal{N}(0, \frac{t}{N})$, et le deuxième une variable aléatoire indépendante de t. Ainsi cette quantité ne converge pas vers une mesure de probabilité quand t tend vers l'infini. Ceci suggère que la direction $(\mathbf{1}, \cdots, \mathbf{1})$ a une mauvaise influence sur le comportement en temps long du système de particules.
On est donc conduit à introduire le processus $(Y_t^N)_{t\geq 0}$ à valeurs dans $\mathbb{R}^{dN}$ défini par

$$Y_t^{i,N} = X_t^{i,N} - \frac{1}{N}\sum_{j=1}^{N} X_t^{j,N} \quad \text{pour } i = 1, \cdots, N.$$

On peut obtenir une expression explicite de la loi invariante $u_\infty^{(N)}$ de (Y_t^N), et une estimée de la vitesse de convergence à l'équilibre en terme d'entropie relative. Cela se démontre à l'aide d'une inégalité de Sobolev logarithmique sur la matrice Hessienne de U, où $U(x_1, \cdots, x_N) = \frac{1}{2N}\sum_{i,j} W(x_i - x_j)$.
Quant à la propagation du chaos pour ce système recentré, on peut montrer qu'elle a lieu uniformément en temps : il existe une constante C telle que pour tout $N \in \mathbb{N}^*$,

$$\sup_{t\geq 0} \mathbb{E}\left[|Y_t^{i,N} - \bar{X}_t^i|^2\right] \leq \frac{C}{N}.$$

Enfin, en utilisant cette propagation du chaos uniforme en temps et la convergence à l'équilibre du système de particules recentré, on aboutit à la convergence à l'équilibre à vitesse exponentielle du processus non-linéaire, en distance converge à $\mathcal{W}_2$.
Pour finir, mentionnons le fait que dans [Carrillo *et al.*, 2003], les auteurs montrent la convergence exponentielle à l'équilibre du processus non-linéaire par une approche analytique.

Cet exemple illustre les différentes questions qui peuvent se poser lors de l'étude d'un système de particules en interaction : la propagation du chaos,

uniforme ou non en temps, et l'existence du processus non-linéaire, ainsi que la convergence en temps long du système de particules et du processus non-linéaire associé.
Dans les deux exemples ci-dessus, on voit que l'interaction entre les particules peut conduire à une non-linéarité du processus limite par rapport à sa moyenne, ou à sa loi. Mais on peut également imaginer des systèmes de particules qui interagiraient par leur médiane par exemple.

Citons maintenant quelques travaux sur les systèmes de PDMPs en interaction, puisque c'est dans ce cadre que le Chapitre 4 se place.

Exemple 1.3.4. Dans [Thai, 2015] l'auteur considère N particules $X^{1,N}, \cdots, X^{N,N}$ qui évoluent dans $\mathbb{N}$, et dont le générateur du système de particules est

$$Lf(x) = \sum_{i=1}^{N} \left[\left(b_{x_i} + q^+(x_i, M^N)\right) \left(f(x+e_i) - f(x)\right) \right. \\ \left. + \left(d_{x_i} + q^-(x_i, M^N)\right) \left(f(x-e_i) - f(x)\right) \right], \qquad (1.9)$$

où les e_i sont les vecteurs de la base canonique, $M^N = M^N(x) = \frac{1}{N}\sum_{i=1}^N x_i$, et où $b_i > 0$, $d_i > 0$ pour tout i, et $d_0 = q^-(0,m) = 0$ pour tout m.
Sous de bonnes hypothèses sur b, d, q^+ et q^- (hypothèse de convexité et lipschitzianité), l'auteur montre plusieurs points :

- le système de particules converge à vitesse exponentielle en distance de Wasserstein W^1 ;
- il y a propagation du chaos uniforme en temps ;
- le processus non-linéaire, limite du système de particules quand $N \to +\infty$, converge à l'équilibre à vitesse exponentielle en distance de Wasserstein ;
- les limites $t \to +\infty$ et $N \to +\infty$ sont commutatives : la limite en temps du système de particules converge quand N tend vers l'infini vers la limite du processus non-linéaire.

Dans cet exemple, on obtient donc toutes les convergences que l'on peut espérer dans l'étude d'un système de particules.

Exemple 1.3.5. Dans [Monmarché, 2018], Monmarché s'intéresse à des systèmes de particules en interaction ainsi qu'au processus non-linéaire associé, en les étudiant à partir des équations intégro-différentielles associées. Ses résultats, très généraux, incluent en particulier le cas de PDMPs en interaction, l'interaction ayant lieu dans le mécanisme de saut. Ses objets d'intérêt sont d'abord l'existence et l'unicité de la solution d'une équation intégro-différentielle associée à un processus non-linéaire, ainsi que le comportement en temps long d'un tel processus, ou du système de particules associé. Pour cela, il se place dans un régime perturbatif, c'est-à-dire dans le cas où l'interaction est faible par rapport à l'effet mélangeant du processus sans interaction.
Pour obtenir ce comportement en temps long en distance en variation totale,

Monmarché utilise des méthodes de couplage, en ne travaillant pas sur les trajectoires des processus mais sur les lois. Cependant, alors que dans le cas de processus de Markov, si l'on réussit à coupler deux processus, on peut les laisser égaux ensuite grâce à l'absence de mémoire des processus markoviens, cela ne se passe pas si bien pour des processus non-linéaires. En effet, si deux processus en interaction avec leur loi sont égaux à un instant, leur non-linéarité fait qu'ils n'ont pas la même loi, donc pas la même dynamique, et sont donc susceptibles de se séparer après le temps de croisement. Néanmoins, tant qu'ils restent égaux, leurs lois se « rapprochent ». Ainsi, si l'interaction est assez faible par rapport à l'effet mélangeant des processus, la contraction due à ce mécanisme de base va l'emporter sur la probabilité qu'ils se séparent.
Citons en particulier une application de son travail : la convergence du processus Zig-zag en interaction en champ-moyen, c'est-à-dire le processus associé à l'équation intégro-différentielle suivante :

$$\partial_t m_t(x,v) + y\partial_x m_t(x,v) = \lambda_{m_t}(x,-v)m_t(x,v) - \lambda_{m_t}(x,v)m_t(x,v),$$

avec $\lambda_\nu = r\left(v(x-\theta x_\nu)\right)$, où $x_\nu = \int_{\mathbb{R}\times\{-1,+1\}} x\nu(\mathrm{d}x,\mathrm{d}v)$, $\theta \in (0,1)$ et $r : \mathbb{R} \to \mathbb{R}^+$ est une fonction Lipschitz telle que $\lim_{s\to-\infty} r(s) < \lim_{s\to+\infty} r(s)$ et $\inf_{s\in\mathbb{R}} r(s) > 0$.
Cette équation décrit un processus Zig-Zag en dimension 1, à valeurs dans $\mathbb{R}\times\{-1,+1\}$ donc, qui au lieu d'être attiré par l'origine, est attiré par une moyenne entre l'origine et sa propre moyenne en espace.
Monmarché montre que pour θ assez petit, c'est-à-dire lorsque l'attraction vers la moyenne du processus n'est pas trop forte, le processus converge à l'équilibre à vitesse exponentielle.

La littérature sur les systèmes de particules en interaction est très riche, il n'est donc pas possible d'en donner une description exhaustive. Nous finissons donc par citer quelques travaux, étudiant différents types de systèmes de particules en interaction, sans rentrer dans les détails.
Nous mentionnons d'abord deux travaux sur des PDMPs en interaction. Dans [Fournier et Löcherbach, 2016], les auteurs s'intéressent à un système de neurones en interaction, dont l'évolution du potentiel de membrane est décrite par des PDMPs en interaction à travers le taux de saut, mais également l'évolution déterministe entre les sauts. Dans [Graham et Robert, 2009] les auteurs étudient un système de PDMPs en interaction décrivant l'interaction de plusieurs classes de connexions dans un réseau, ainsi que le processus non-linéaire associé.
Dans [Andreis *et al.*, 2018], Andreis et ses co-auteurs s'intéressent à la propagation du chaos d'un système de particules en interaction par leur moyenne, pouvant modéliser des systèmes de neurones. Ces particules sont des diffusions avec sauts simultanés.
Finissons par citer l'article [Fontbona *et al.*, 2009], dans lequel les auteurs s'intéressent à des processus non-linéaires solutions d'EDS dirigées par des bruits blancs espace-temps. Leur but dans ce travail est de construire un système de diffusions en interaction facilement simulable, et approchant le système non-linéaire de départ.

1.4 Présentation des résultats principaux de la thèse

Dans cette section, nous résumons les trois chapitres qui constituent cette thèse, en décrivant dans chaque cas le travail qui est fait, et en donnant les résultats principaux.

1.4.1 Chapitre 2 : Explicit speed of convergence of the stochastic billiard in a convex set

Dans le Chapitre 2, nous étudions le comportement en temps long du billard stochastique $((X_t, V_t))_{t\geq 0}$ dans un convexe K de $\mathbb{R}^2$, décrit dans l'Exemple 1.1.3 dont nous gardons les notations, et nous nous intéressons plus particulièrement à sa vitesse de convergence à l'équilibre.
Rappelons que ce processus décrit le mouvement d'une particule dans un convexe de $\mathbb{R}^2$, qui avance à vitesse unitaire constante jusqu'au moment où elle touche le bord du convexe. À cet instant, la particule est alors réfléchie aléatoirement à l'intérieur du convexe, indépendamment de sa position et de sa vitesse précédente. Les instants de rebond de la particule sont alors notés T_n, $n \geq 0$, avec $T_0 = 0$ si la position initiale de la particule est sur le bord du convexe.

Pour $x \in \partial K$, ∂K désignant la frontière de K, il est équivalent de considérer la vitesse après un saut dans $\mathbb{S}_x = \{v \in \mathbb{S}^1 \ : \ \langle v, n_x\rangle \geq 0\}$ ou l'angle dans $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ entre ce vecteur vitesse et la normale rentrante n_x. Pour $n \geq 1$, on considère donc Θ_n la variable aléatoire à valeurs dans $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ telle que $r_{X_{T_n},\Theta_n}(n_{X_{T_n}}) \overset{\mathcal{L}}{=} V_{T_n}$, où pour $x \in \partial K$ et $\theta \in \mathbb{R}$, $r_{x,\theta}$ est la rotation de centre x et d'angle θ.
Nous faisons alors une hypothèse sur la loi γ, que l'on peut écrire de manière équivalente sous l'une des deux formes suivantes $(\mathcal{H})$ ou $(\mathcal{H}')$:

Hypothèse $(\mathcal{H})$:

La loi γ a une densité ρ par rapport à la mesure de Haar sur $\mathbb{S}_e$, qui vérifie : il existe $\mathcal{J} \subset \mathbb{S}_e$, symétrique par rapport à e, et $\rho_{\min} > 0$ tels que :

$$\rho(u) \geq \rho_{\min}, \quad \text{pour tout } u \in \mathcal{J}.$$

Hypothèse $(\mathcal{H}')$:

Les variables Θ_n, $n \geq 0$, ont une densité commune f par rapport à la mesure de Lebesgue sur $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ satisfaisant : il existe $f_{\min} > 0$ et $\theta^* \in (0, \pi)$ tels que :

$$f(\theta) \geq f_{\min}, \quad \text{pour tout } \theta \in \left[-\frac{\theta^*}{2}, \frac{\theta^*}{2}\right].$$

L'équivalence de ces hypothèses implique que

$$\rho_{\min} = f_{\min} \quad \text{et} \quad |\mathcal{J}| = \theta^*.$$

L'une ou l'autre des descriptions du processus seront utilisées en fonction de leur praticité dans les différentes études.

Le comportement en temps long de tels processus a beaucoup été étudié, et sous différentes hypothèses. Citons le travail d'Evans [Evans, 2001], dans lequel il s'intéresse à la convergence du billard stochastique évoluant dans une région bornée à bord $\mathcal{C}^1$ en dimension d, ainsi que dans des régions polygonales du plan, et avec une loi de réflexion uniforme. Dans [Dieker et Vempala, 2015], les auteurs étudient eux la trace du billard stochastique sur le bord d'un convexe à courbure minorée, et avec pour loi de réflexion la loi du cosinus. Dans ce papier, ils obtiennent une borne sur la vitesse de convergence de ce processus vers la mesure uniforme sur le bord du convexe. Enfin, mentionnons l'article [Comets *et al.*, 2009], dont le Chapitre 2 est en partie inspiré. Dans ce travail, les auteurs s'intéressent à la convergence du billard stochastique ainsi que sa chaîne incluse, dans un cadre très général, puisqu'ils considèrent des domaines de $\mathbb{R}^d$ à bord localement Lipschitz et presque partout C^1, et une loi de réflexion absolument continue par rapport à la mesure de Haar sur la sphère unité, et dont la densité est strictement positive sur la demi-sphère des vecteurs pointant vers l'intérieur du domaine. La convergence exponentielle de ces deux processus est obtenue en décrivant un couplage astucieux. Cependant, le domaine dans lequel vivent ces processus étant tellement général, les vitesses de convergence obtenues par leur méthode dans des domaines plus particuliers ne sont pas bonnes. Le but du Chapitre 2 est donc de construire des couplages adaptés pour obtenir des bornes sur les vitesses de convergence du billard stochastique et de sa chaîne incluse, dans un premier temps dans un disque, puis plus généralement dans un convexe du plan à courbure minorée et majorée.

Décrivons brièvement et de manière informelle les stratégies utilisées pour obtenir nos vitesses de convergence. Pour la trace du billard stochastique sur le bord du domaine considéré, nous estimons le nombre de pas nécessaires n_0 pour que la position de la chaîne charge plus de la moitié du domaine dans lequel elle évolue. Ainsi, considérant deux processus partant d'états initiaux différents, on peut les construire de façon à ce qu'après n_0 pas, ils aient une probabilité strictement positive d'être égaux. Après un nombre géométrique d'essais, on réussit alors à coupler nos deux processus, et l'estimation de ce temps donne une borne sur la vitesse de convergence de la chaîne incluse.

Pour le processus en temps continu, c'est plus compliqué puisque, considérant deux billards stochastiques partant d'états initiaux différents, on doit réussir à coupler leurs positions et vitesses au même instant, pour pouvoir ensuite les laisser égaux. Nous faisons alors cela en deux temps. D'abord on construit les deux processus jusqu'à ce qu'ils touchent le bord du domaine à un même instant, mais pas nécessairement au même endroit. Puis, partant du bord, on essaye de les faire toucher le bord au même endroit et au même instant, après deux rebonds. Avec une certaine probabilité cet essai échoue, et on recommence alors le couplage depuis le début, jusqu'à ce que la deuxième étape soit un succès. L'estimation de ce temps de couplage permet alors d'estimer la vitesse de convergence du billard stochastique.

1.4.1.1 Billard stochastique dans un disque

On s'intéresse dans un premier temps au cas particulier du billard stochastique évoluant dans un disque, ou autrement dit une boule de $\mathbb{R}^2$ de la forme $K = \mathcal{B}(0,r)$, pour un certain $r > 0$ fixé.
Dans ce cas, pour $n \geq 0$, le couple $(X_{T_n}, V_{T_n}) \in \partial\mathcal{B}(0,r) \times \mathbb{S}^1$ peut être représenté de manière équivalente par le couple $(\Phi_n, \Theta_n) \in [0, 2\pi) \times \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ défini comme suit :

— à une position x sur $\partial\mathcal{B}(0,r)$ correspond un unique angle $\phi \in [0, 2\pi)$. La variable Φ_n désigne donc cet unique angle associé à X_{T_n}, i.e. (r, Φ_n) sont les coordonnées polaires de X_{T_n}.

— à la vitesse V_{T_n} on associe la variable Θ_n introduite précédemment, et satisfaisant l'hypothèse $(\mathcal{H}')$.

Il est alors immédiat d'observer les relations suivantes, en rappelant que les variables τ_n, $n \geq 1$, sont définies par $\tau_n = T_n - T_{n-1}$:

Proposition 1.4.1. *Pour tout $n \geq 1$ on a :*

$$\tau_n = 2r\cos(\Theta_{n-1}) \quad \textit{et} \quad \Phi_n = \pi + 2\Theta_{n-1} + \Phi_{n-1}.$$

Ces relations entre les différentes variables décrivant le billard stochastique dans le disque permettent de faire des calculs explicites. Ainsi, en utilisant les hypothèses $(\mathcal{H})$ ou $(\mathcal{H}')$, on obtient des estimations des densités de la position X_{T_n}, des temps T_n, ou encore des couples (X_{T_n}, T_n).

On observe alors que si l'angle θ^* (qui décrit la longueur de l'intervalle que charge la loi γ) est strictement plus grand que $\frac{\pi}{2}$, on peut espérer coupler deux versions de la chaîne incluse en un pas. Pour θ^* plus petit, le nombre de rebonds nécessaires (n_0 dans le théorème suivant) pour avoir une probabilité strictement positive de coller deux billards stochastiques sur le bord du disque est alors plus élevé. On obtient finalement le résultat suivant.

Théorème 1.4.1. *Soit $(\Phi_n)_{n\geq 0}$ la chaîne de Markov associée au billard stochastique dans la boule $\mathcal{B}(0,r)$, satisfaisant l'hypothèse $(\mathcal{H}')$.*
Il existe une unique mesure de probabilité ν sur $[0, 2\pi)$ invariante pour la chaîne $(\Phi_n)_{n\geq 0}$, et on a :

1. *si $\theta^* > \frac{\pi}{2}$, pour tout $n \geq 0$,*

$$\|\mathbb{P}\left(\Phi_n \in \cdot\right) - \nu\|_{TV} \leq \left(1 - f_{\min}(2\theta^* - \pi)\right)^{n-1},$$

2. *si $\theta^* \leq \frac{\pi}{2}$, pour tout $n \geq 0$ et tout $\varepsilon \in (0, \theta^*)$,*

$$\|\mathbb{P}\left(\Phi_n \in \cdot\right) - \nu\|_{TV} \leq (1 - \alpha)^{\frac{n}{n_0} - 1},$$

où

$$n_0 = \left\lfloor \frac{\pi - 2\varepsilon}{2(\theta^* - \varepsilon)} \right\rfloor + 1 \quad \textit{et} \quad \alpha = \left(\frac{\varepsilon}{2}\right)^{n_0 - 1} {f_{\min}}^{n_0} \left(2n_0\theta^* - 2(n_0 - 1)\varepsilon - \pi\right).$$

Pour le processus en temps continu, utilisant encore les relations de la Proposition 1.4.1 et l'hypothèse $(\mathcal{H})$, on observe que si θ^* est strictement plus grand que $\frac{2\pi}{3}$, la probabilité de coupler les deuxièmes temps de rebonds de deux billards stochastiques est strictement positive. De plus, pour un tel θ^*, la probabilité de coupler deux billards stochastiques partant du bord du cercle en deux rebonds est également strictement positive. Ainsi, on peut construire un couplage de deux billards stochastiques comme décrit à la fin de la section 1.4.1, et dont on peut estimer le temps de couplage. On obtient alors le résultat suivant donnant la vitesse de convergence à l'équilibre du billard stochastique dans un disque.

Théorème 1.4.2. *Soit $((X_t, V_t))_{t\geq 0}$ le billard stochastique évoluant dans $\mathcal{B}(0, r)$ satisfaisant l'hypothèse $(\mathcal{H}')$ avec $\theta^* \in \left(\frac{2\pi}{3}, \pi\right)$.*
Il existe une unique mesure de probabilité invariante χ sur $\mathcal{B}(0, r) \times \mathbb{S}^1$ pour le processus $(X_t, V_t)_{t\geq 0}$.
De plus, soit $\eta \in \left(0, r\left(1 - 2\cos\left(\frac{\theta^*}{2}\right)\right)\right)$ et $\varepsilon \in \left(0, \frac{2\theta^* - \pi}{8}\right)$.
Il existe $\delta, h, \alpha > 0$ tels que pour tout $t \geq 0$ et tout $\lambda < \lambda_M$ on a

$$\|\mathbb{P}\left(X_t \in \cdot, V_t \in \cdot\right) - \chi\|_{TV} \leq C_\lambda \mathrm{e}^{-\lambda t},$$

où

$$\lambda_M = \min\left\{\frac{1}{4r}\log\left(\frac{1}{1-\delta h}\right); \frac{1}{4r}\log\left(\frac{-(1-\delta h) + \sqrt{(1-\delta h)^2 + 4\delta h(1-\alpha)}}{2\delta h(1-\alpha)}\right)\right\}.$$

et

$$C_\lambda = \frac{\alpha \delta h \mathrm{e}^{10\lambda r}}{1 - \mathrm{e}^{4\lambda r}(1-\delta h) - \mathrm{e}^{8\lambda r}\delta h(1-\alpha)}.$$

1.4.1.2 Billard stochastique dans un convexe du plan à courbure majorée et minorée

On considère ensuite un cas plus général : le billard stochastique évoluant dans un convexe à courbure majorée et minorée. On introduit donc l'hypothèse suivante :

Hypothèse $(\mathcal{K})$ **:**

> K est un convexe compact du plan à courbure majorée par $C < \infty$
> et minorée par $c > 0$.

Cela signifie qu'en tout point $x \in \partial K$, il existe une boule B_1 de rayon $\frac{1}{C}$ incluse dans K et une boule B_2 de rayon $\frac{1}{c}$ contenant K, telles que les tangentes à K, B_1 et B_2 au point x coïncident.
On introduit également D le diamètre de K :

$$D = \max\{\|x - y\|: x, y \in \partial K\}.$$

Dans ce cadre, puisque le convexe K n'est pas explicite, on ne peut pas obtenir de relations entre les temps et positions de rebond, comme on avait obtenu dans la Proposition 1.4.1 pour le disque. Il faut donc utiliser l'hypothèse $(\mathcal{K})$ et des arguments géométriques élémentaires pour estimer les densités des différentes variables, ou plutôt pour estimer la taille des intervalles que ces densités chargent.
Comme pour le disque, nous nous intéressons d'abord à la chaîne incluse $(X_{T_n})_{n\geq 0}$. Son noyau de transition est connu (on peut le trouver dans [Comets *et al.*, 2009] par exemple), et on peut alors en déduire une borne sur sa vitesse de convergence à l'équilibre :

Théorème 1.4.3. *Soit $K \subset \mathbb{R}^2$ satisfaisant l'hypothèse $(\mathcal{K})$, de diamètre D. Soit $(X_{T_n})_{n\geq 0}$ la chaîne de Markov associée au billard stochastique évoluant dans K, et vérifiant l'hypothèse $(\mathcal{H})$.*
Il existe une unique mesure de probabilité invariante ν sur ∂K pour $(X_{T_n})_{n\geq 0}$. De plus, rappelant que $\theta^* = |\mathcal{J}|$ dans l'hypothèse $(\mathcal{H})$, on a :

1. *si $\theta^* > \frac{C|\partial K|}{8}$, pour tout $n \geq 0$,*
$$\|\mathbb{P}\left(X_{T_n} \in \cdot\right) - \nu\|_{TV} \leq \left(1 - q_{\min}\left(\frac{8\theta^*}{C} - |\partial K|\right)\right)^{n-1};$$
2. *si $\theta^* \leq \frac{C|\partial K|}{8}$, pour tout $n \geq 0$ et tout $\varepsilon \in \left(0, \frac{2\theta^*}{C}\right)$,*
$$\|\mathbb{P}\left(X_{T_n} \in \cdot\right) - \nu\|_{TV} \leq (1-\alpha)^{\frac{n}{n_0}-1}$$
où
$$n_0 = \left\lfloor \frac{\frac{|\partial K|}{2} - 2\varepsilon}{\frac{4\theta^*}{C} - 2\varepsilon} \right\rfloor + 1$$
$$\text{et} \quad \alpha = \left(\frac{4\theta^*}{C}\right)^{n_0-1} {q_{\min}}^{n_0}\left(4\left(\frac{2n_0\theta^*}{C} - (n_0-1)\varepsilon\right) - |\partial K|\right)$$

avec
$$q_{\min} = \frac{c\rho_{\min}\cos\left(\frac{\theta^*}{2}\right)}{CD}.$$

Enfin, on s'intéresse à la vitesse de convergence du processus en temps continu. Dans ce cas, nous arrivons à estimer, en fonction de θ^*, le nombre de rebonds nécessaires pour pouvoir coupler les temps de rebonds de deux billards stochastiques. Cependant, pour ensuite coller les positions et temps de rebonds de nos deux processus après deux sauts, nous utilisons un couplage dans lequel le saut intermédiaire des deux processus peut avoir lieu partout sur le convexe. Ainsi, nous avons besoin de supposer que $\theta^* = \pi$, c'est-à-dire que la vitesse lors d'un rebond charge tout le convexe, pour que ce rebond intermédiaire soit possible.
On obtient donc finalement le théorème suivant donnant la convergence du processus en temps continu, sous l'hypothèse $\theta^* = \pi$.

Théorème 1.4.4. *Soit $K \subset \mathbb{R}^2$ satisfaisant l'hypothèse $(\mathcal{K})$, de diamètre D. Soit $((X_t, V_t))_{t\geq 0}$ le billard stochastique évoluant dans K et satisfaisant l'hypothèse $(\mathcal{H})$ avec $|\mathcal{J}| = \pi$.*
Il existe une unique mesure de probabilité χ sur $K \times \mathbb{S}^1$ invariante pour le processus $((X_t, V_t))_{t\geq 0}$.
De plus, il existe $\kappa, p > 0$ tels que pour tout $t \geq 0$ et tout $\lambda < \lambda_M$:

$$\|\mathbb{P}(X_t \in \cdot, V_t \in \cdot) - \chi\|_{TV} \leq C_\lambda e^{-\lambda t},$$

où

$$\lambda_M = \min\left\{\frac{1}{2D}\log\left(\frac{1}{1-p}\right); \frac{1}{2D}\log\left(\frac{-(1-p)+\sqrt{(1-p)^2+4p(1-\kappa)}}{2p(1-\kappa)}\right)\right\}$$

et

$$C_\lambda = \frac{p\kappa e^{5\lambda D}}{1 - e^{2\lambda D}(1-p) - e^{4\lambda D}p(1-\kappa)}.$$

Nous terminons le Chapitre 2 par une discussion expliquant comment généraliser ces résultats à la dimension $n \geq 2$.

1.4.2 Chapitre 3 : Long-time behaviour of generalized Zig-Zag process

Dans le Chapitre 3, on s'intéresse au comportement en temps long du PDMP $((X_t, V_t))_{t\geq 0}$ à valeurs dans $E = \mathbb{R}^d \times \mathcal{B}(1)$, où $\mathcal{B}(1) = \left\{v \in \mathbb{R}^d, |v| \leq 1\right\}$ est la boule euclidienne de rayon 1, dont le générateur infinitésimal L est donné par, pour $h \in \mathcal{D}(L)$ et $(x, v) \in \mathbb{R}^d \times \mathcal{B}(1)$:

$$Lh(x, v) = v \cdot \nabla_x h(x, v) + \lambda(x, v)\int_{\mathcal{B}(1)} \big(h(x, v') - h(x, v)\big)\, Q(x, v, dv'), \quad (1.10)$$

où $Q(x, v, \cdot)$ est un noyau de transition sur $\mathcal{B}(1)$.
On appelle ce processus le « processus Zig-Zag généralisé » puisqu'il est une généralisation du processus Zig-Zag étudié dans [Bierkens et Roberts, 2017], [Bierkens *et al.*, 2019] et [Fontbona *et al.*, 2012]. Un cas particulier du processus de générateur (1.10) a été étudié de manière analytique dans [Calvez *et al.*, 2015] : les auteurs s'intéressent dans ce papier au cas particulier de la dimension 1, et avec un noyau de saut Q uniforme sur $[-1, 1]$. Les résultats du Chapitre 3, donnant la convergence exponentielle du processus de générateur donné par (1.10) sous de bonnes hypothèses, sont une généralisation des articles cités ci-dessus puisqu'on considère ici un processus en dimension d dont la première composante est dans $\mathbb{R}^d$ et la seconde dans la boule unité.
Dans ce modèle, X_t représente la position d'une bactérie à l'instant t, et V_t sa vitesse. La forme du générateur indique que la composante X est continue et évolue suivant l'équation $\frac{dX_t}{dt} = V_t$, tandis que V est constante pendant un temps aléatoire, puis saute selon le noyau Q avec un taux de saut $\lambda(x, v)$ quand

$(X_t, V_t) = (x, v)$.
Dans cette étude, nous cherchons à modéliser le mouvement d'une bactérie dans $\mathbb{R}^d$, attirée par un nutriment positionné à l'origine. Nous traduisons cela par les hypothèses $(\mathcal{H}_3)$ et $(\mathcal{H}_4)$ décrites ci-dessous.
Les hypothèses que l'on fait sur notre modèle sont les suivantes :

$(\mathcal{H}_1)$: Il existe $\lambda_{\min} > 0$ tel que pour tout $(x, v) \in E$, $\lambda(x, v) \geq \lambda_{\min}$;
$(\mathcal{H}_2)$: La quantité $\lambda_{\max} = \sup_{\{(x,v)\in E\ :\ x\cdot v\leq 0\}} \lambda(x, v)$ est finie ;
$(\mathcal{H}_3)$: Il existe $p > 0$, $\theta_0, \bar{\theta} \in (0, 1]$ tels que pour tout $(x, v) \in E$ satisfaisant $\frac{x\cdot v}{|x|} > -\bar{\theta}$,

$$\int_{\{v'\in\mathcal{V}\ :\ \frac{x\cdot v'}{|x|}\leq -\theta_0\}} Q(x, v, \mathrm{d}v') \geq p;$$

$(\mathcal{H}_4)$: Il existe $\theta_* \in \left[0, (p\theta_0)^2 \frac{\lambda_{\min}}{\lambda_{\max}}\right)$, $\beta > \frac{1}{(p\theta_0)^2}$ et $\Delta > 0$ tels que

$$\inf_{\left\{\frac{x\cdot v}{|x|}\geq\theta_*, |x|\geq\Delta\right\}} \lambda(x, v) \geq \beta\lambda_{\max}.$$

Les hypothèses $(\mathcal{H}_1)$ et $(\mathcal{H}_2)$ signifient que le taux de saut λ est uniformément minoré, et qu'il est majoré lorsque la bactérie se dirige vers le nutriment situé à l'origine. L'existence de la borne $\lambda_{\min}$ permet d'assurer l'irréductibilité du processus.
Comme mentionné plus haut, les deux dernières hypothèses reflètent l'attraction de la bactérie vers l'origine. Plus précisément, l'hypothèse $(\mathcal{H}_3)$ traduit le fait que si la bactérie ne se dirige « pas assez » vers l'origine, lorsqu'un saut de vitesse a lieu, elle a une probabilité uniformément minorée de ramener la bactérie dans la direction du nutriment. De plus, dans l'hypothèse $(\mathcal{H}_4)$, on suppose une sorte de monotonie du taux de saut : on suppose que lorsque la bactérie est loin de l'origine, et se dirige dans une « trop mauvaise » direction, son taux de saut est toujours plus grand que lorsque la bactérie de dirige vers le nutriment.
Le résultat principal du Chapitre 3 est le suivant :

Théorème 1.4.5. *Soit $((X_t, V_t))_{t\geq 0}$ le PDMP sur $E = \mathbb{R}^d \times \mathcal{B}(1)$ de générateur infinitésimal donné par* (1.10).
Si λ et Q satisfont les hypothèses $(\mathcal{H}_1)$, $(\mathcal{H}_2)$, $(\mathcal{H}_3)$ et $(\mathcal{H}_4)$, alors le processus (X, V) est exponentiellement ergodique.

Pour montrer ce résultat, nous introduisons une fonction de Lyapunov pour le processus (X, V), pour ensuite pouvoir appliquer le Théorème (1.2.1).

Puis, la théorie des processus régénératifs nous permet d'obtenir l'existence de moments exponentiels de la mesure invariante du PDMP (X, V) :

Théorème 1.4.6. *Soit $((X_t, V_t))_{t\geq 0}$ le processus Zig-zag généralisé satisfaisant les hypothèses $(\mathcal{H}_1)$, $(\mathcal{H}_2)$, $(\mathcal{H}_3)$ et $(\mathcal{H}_4)$.*
Il existe $\eta > 0$ tel que pour tout $0 < \beta < \eta$ et tout $\gamma > 0$, :

$$\int_E \mathrm{e}^{\beta|x|+\gamma|v|} \pi(dx, dv) < \infty,$$

où π est l'unique mesure de probabilité invariante pour (X, V).

Enfin, dans une dernière partie, nous nous intéressons au processus Zig-zag généralisé dans la cas particulier de la dimension 1, cas dans lequel on montre des résultats similaires, avec une approche un peu différente. Cette étude particulière est faite sous des hypothèses différentes du cas de la dimension d, et lui est donc complémentaire.
On considère donc le PDMP, que l'on note encore $((X_t, V_t))_{t\geq 0}$, à valeurs dans $\mathbb{R} \times [-1, 1]$, de générateur infinitésimal donné par

$$Lh(x, v) = v\partial_x h(x, v) + \lambda(x, v) \int_{-1}^{1} \big(h(x, v') - h(x, v)\big)\, Q(x, v, \mathrm{d}v'), \qquad (1.11)$$

où $Q(x, v, \cdot)$ est un noyau de transition sur $[-1, 1]$.
Dans ce cas particulier de la dimension 1, nous faisons les hypothèses suivantes sur le modèle :

- $(\mathcal{A}_1)$: $Q(x, v, \mathrm{d}v') = q(v, v')(\nu(\mathrm{d}v') + \mu(\mathrm{d}v'))$ avec ν une mesure discrète et μ une restriction de la mesure de Lebesgue. On note $\mathcal{V}$ le support de Q et on suppose qu'il existe $q_{\min} > 0$ tel que $q(v, v') \geq q_{\min}$ pour tout $v, v' \in \mathcal{V}$;
- $(\mathcal{A}_2)$: Le processus est symétrique : $\mathcal{V}$ est symétrique par rapport à 0 et $\lambda(x, v) = \lambda(-x, -v)$ pour tout $(x, v) \in \mathbb{R} \times \mathcal{V}$;
- $(\mathcal{A}_3)$: Il existe $\lambda_{\min} > 0$ tel que pour tout $(x, v) \in \mathbb{R} \times \mathcal{V}$, $\lambda_{\min} \leq \lambda(x, v)$, et la quantité $\sup_{x\geq 0, v\leq 0} \lambda(x, v)$ est finie ;
- $(\mathcal{A}_4)$:

$$\exists I_* \subset \left[0, \inf_{x\geq 0, v\in(0,1]} \frac{\lambda(x, v)}{v}\right) \text{ an interval } I_*, \exists 0 < J_* < 1,$$
$$\forall \alpha \in I_*, \forall v' \in \mathcal{V}, \int_{-1}^{1} G(\alpha, v) Q(v', \mathrm{d}v) \leq J_*,$$

où

$$G(\alpha, v) = \frac{\sup\limits_{x\geq 0} \lambda(x, v)}{\sup\limits_{x\geq 0} \lambda(x, v) - \alpha v} \mathbf{1}_{v<0} + \frac{\inf\limits_{x\geq 0} \lambda(x, v)}{\inf\limits_{x\geq 0} \lambda(x, v) - \alpha v} \mathbf{1}_{v\geq 0}$$

pour $\alpha \geq 0$ et $v \in \mathcal{V}$.

L'hypothèse $(\mathcal{A}_2)$ qui suppose la symétrie du processus n'est pas indispensable pour obtenir les résultats qui suivent. Elle permet simplement de réduire les calculs nécessaires dans les preuves. L'hypothèse $(\mathcal{A}_4)$ est celle qui traduit l'attraction de la bactérie vers l'origine. Cette hypothèse apparaît naturellement dans les calculs, mais on n'y lit pas directement sa signification. Nous donnerons dans le Chapitre 3 des conditions plus fortes qui permettent de mieux la comprendre.
Une étude poussée du temps de retour en 0 de notre PDMP (X, V) et l'application du Théorème 1.2.2 permettent alors d'obtenir le résultat suivant :

Théorème 1.4.7. *Soit* $((X_t, V_t))_{t\geq 0}$ *le PDMP sur* $\mathbb{R} \times [-1,1]$ *de générateur infinitésimal donné par* (1.11). *Sous les hypothèses* $(\mathcal{A}_1)$, $(\mathcal{A}_2)$, $(\mathcal{A}_3)$ *et* $(\mathcal{A}_4)$, *le processus est exponentiellement ergodique.*

Puis, de la même manière qu'en dimension d, la théorie des processus régératifs permet d'obtenir l'existence de moments exponentiels de la mesure invariante π du processus :

Théorème 1.4.8. *Soit* $\gamma > 0$ *tel que* $\frac{\lambda_{\min}}{\lambda_{\min}-\gamma} J_* < 1$.
Pour tout $0 < \alpha < \gamma$ *et tout* $\beta > 0$, *on a :*

$$\int_{\mathbb{R}\times[-1,1]} \mathrm{e}^{\alpha|x|+\beta|v|}\pi(dx, dv) < \infty.$$

1.4.3 Chapitre 4 : Zig-zag processes in interaction

Enfin, dans le dernier chapitre, on s'intéresse à un système de processus Zig-zag uni-dimensionnels en interaction par leur moyenne, et au processus non-linéaire associé. Le générateur du système de particules que l'on étudie est donné par, pour $(x, v) = ((x_1, \cdots, x_N), (v_1, \cdots, v_N)) \in \mathbb{R}^N \times \{-1, +1\}^N$,

$$\mathcal{L}_N f(x, v) = \sum_{i=1}^{N} v_i \partial_{x_i} f(x, v) + \sum_{i=1}^{N} \lambda\left((x_i - \overline{x})v_i\right)\left(f(x, v^{(i)}) - f(x, v)\right), \quad (1.12)$$

où $v_j^{(i)} = \begin{cases} v_j & \text{si } j \neq i \\ -v_i & \text{sinon.} \end{cases}$ et $\overline{x} = \frac{1}{N}\sum_{k=1}^{N} x_i$, et où λ est une fonction positive sur $\mathbb{R}$.
L'écriture du taux de saut sous la forme $\lambda\left((x_i - \overline{x})v_i\right)$ impose une symétrie par rapport à la moyenne spatiale des particules. Ainsi, sous des conditions classiques sur le taux de saut λ, les N particules sont attirées par leur position moyenne.
Autrement dit on s'intéresse aux couples $\left((X_t^{i,N}, V_t^{i,N})\right)_{t\geq 0}$ pour $i \in \{1, \cdots, N\}$, solutions du système stochastique suivant :

$$\begin{cases} \mathrm{d}X_t^{i,N} &= V_t^{i,N}\mathrm{d}t \\ \mathrm{d}V_t^{i,N} &= -2V_{t^-}^{i,N} \int \mathbf{1}_{z\leq\lambda((X_t^{i,N} - \frac{1}{N}\sum_{j=1}^{N} X_t^{j,N})V_{t^-}^{i,N})}\mathcal{N}^i(\mathrm{d}z, \mathrm{d}t) \end{cases} \quad (1.13)$$

où les $\mathcal{N}^i$, $i = 1, \cdots, N$, sont N processus de Poisson indépendants sur $(\mathbb{R}^+)^2$, d'intensité la mesure de Lebesgue.
Le processus non-linéaire associé, attendu comme étant la limite du système de particules, est alors le processus $((X_t, V_t))_{t\geq 0}$ solution (si elle existe) du système non-linéaire suivant :

$$\begin{cases} \mathrm{d}X_t &= V_t\mathrm{d}t \\ \mathrm{d}V_t &= -2V_{t^-} \int \mathbf{1}_{z\leq\lambda((X_t - \mathbb{E}[X_t])V_{t^-})}\mathcal{N}(\mathrm{d}z, \mathrm{d}t), \end{cases} \quad (1.14)$$

où $\mathcal{N}$ est un processus de Poisson sur $(\mathbb{R}^+)^2$ d'intensité la mesure de Lebesgue.

Nous faisons une première hypothèse sur le taux de saut λ, pour assurer l'existence et unicité du processus non-linéaire, ainsi que la propagation du chaos. Il s'agit d'une hypothèse classique permettant de montrer des résultats d'existence et d'unicité des solutions d'une EDS standard.

Hypothèse $(\mathcal{H}_\lambda)$ **:**

On suppose que λ est une fonction positive lipschitzienne sur $\mathbb{R}$, et on note K sa constante de Lipschitz :

$$\forall x, y \in \mathbb{R}, \quad |\lambda(x) - \lambda(y)| \leq K |x - y| .$$

De plus, nous avons besoin d'une hypothèse sur les conditions initiales du processus non-linéaire pour assurer sa bonne définition (cette hypothèse est forte, et nous espérons pouvoir l'affaiblir par la suite).

Hypothèse $(\mathcal{H}_\alpha)$ **:**

La variable aléatoire X_0 à valeurs dans $\mathbb{R}$ satisfait l'hypothèse $(\mathcal{H}_\alpha)$ s'il existe $\alpha > 0$ tel que $\mathbb{E}\left[\mathrm{e}^{\alpha|X_0|}\right] < \infty$.

On peut alors montrer l'existence et l'unicité des solutions du système (1.14) sous ces deux hypothèses, ainsi que la propagation du chaos, en temps fini. On appelle alors processus Zig-zag non-linéaire la solution de (1.14). Ces résultats se montrent de manière classique : on commence par s'intéresser au système (1.14) où l'on remplace $\mathbb{E}[X_t]$ par une fonction $t \mapsto m_t$ déterministe, et on montre la continuité de la solution en les conditions initiales et en cette fonction m. Puis on utilise une méthode d'itérations de Picard pour en déduire l'existence et l'unicité des solutions du système non-linéaire (1.14). La propagation du chaos se démontre ensuite en couplant N particules en interaction avec N copies indépendantes du processus Zig-zag non-linéaire, partant des mêmes conditions initiales i.i.d. Les résultats sont alors les suivants.

Théorème 1.4.9. *Si X_0 satisfait la condition $(\mathcal{H}_\alpha)$, et sous l'hypothèse $(\mathcal{H}_\lambda)$, il y a existence et unicité trajectorielle d'une solution du système (1.14) avec conditions initiales (X_0, V_0), V_0 étant une variable à valeurs dans $\{-1, +1\}$. De plus, soient $((X_t, V_t))_{t\geq 0}$ et $\left((\tilde{X}_t, \tilde{V}_t)\right)_{t\geq 0}$ deux solutions de (1.14) avec conditions initiales respectives (X_0, V_0) et $(\tilde{X}_0, \tilde{V}_0)$ satisfaisant $(\mathcal{H}_\alpha)$. Alors, pour tout $T \geq 0$, il existe des constantes $C_T, C'_T > 0$ telles que :*

$$\begin{aligned}\mathbb{E}&\left[\|X - \tilde{X}\|_T + \|V - \tilde{V}\|_T\right]\\ &\leq \mathbb{E}\left[\left(|X_0 - \tilde{X}_0| + |V_0 - \tilde{V}_0|\right)\mathrm{e}^{\alpha|X_0|}\right]\left(1 + C_T \exp\left(C'_T \mathbb{E}\left[\mathrm{e}^{\alpha|X_0|}\right]\right)\mathbb{E}\left[\mathrm{e}^{\alpha|X_0|}\right]\right).\end{aligned}$$

Théorème 1.4.10. *Considérons $\left((\overline{X}_t^{i,N}, \overline{V}_t^{i,N})\right)_{t\geq 0}$, $i \in \{1, \cdots, N\}$, N copies indépendantes du processus non-linéaire, dirigés par les mêmes processus de Poisson que les $\left((X_t^{i,N}, V_t^{i,N})\right)_{t\geq 0}$. On suppose que pour $i \in \{1, \cdots, N\}$, les*

processus $\left((X_t^{i,N}, {}_t^{i,N})\right)_{t\geq 0}$ *et* $\left((\overline{X}_t^{i,N}, \overline{V}_t^{i,N})\right)_{t\geq 0}$ *ont les mêmes conditions initiales* $\left(X_0^{i,N}, V_0^{i,N}\right)$.
Supposons de plus que :

- *les variables aléatoires* $\left(\left(X_0^{i,N}, V_0^{i,N}\right)\right)_{1\leq i\leq N}$ *sont indépendantes et identiquement distribuées ;*
- *les variables aléatoires* $\left(X_0^{i,N}\right)_{1\leq i\leq N}$ *satisfont* (H_α).

Alors, pour tout $t \geq 0$, *il existe une constante* C_t, *indépendante de* N, *telle que pour tout* $i \in \{1, \cdots, N\}$

$$\mathbb{E}\left[\left\|X^{i,N} - \overline{X}^{i,N}\right\|_t + \left\|V^{i,N} - \overline{V}^{i,N}\right\|_t\right] \leq \frac{C_t}{\sqrt{N}}.$$

Ces deux résultats sont démontrés sous l'Hypothèse $(\mathcal{H}_\lambda)$ qui suppose que le taux de saut est Lipschitz, hypothèse classique dans ce genre de preuves. Cependant, on peut penser, bien que nous ne l'ayons pas encore démontré, que l'existence et unicité du processus non-linéaire ainsi que la propagation du chaos restent vraies avec un taux de saut de la forme $\lambda(z) = a\mathbf{1}_{z<0} + b\mathbf{1}_{z\geq 0}$.

Une fois ces résultats prouvés, notre but est d'étudier le comportement en temps long du processus Zig-zag non-linéaire, sous de bonnes hypothèses sur le taux de saut (hypothèses traduisant l'attraction vers la moyenne). Cependant, alors que dans le cas diffusif, l'espérance de processus non-linéaires similaires reste constante au cours du temps (voir [Malrieu, 2003] par exemple), le comportement de la position moyenne du processus Zig-zag non-linéaire n'est pas trivial. Ainsi, pour se placer dans un cadre plus simple, nous introduisons un nouveau système de particules, dans lequel on a recentré les positions par rapport au système (1.13). Nous introduisons également le processus non-linéaire associé à ce nouveau système de particules.
Posant, pour $i \in \{1, \cdots, N\}$, $Y_t^{i,N} = X_t^{i,N} - \sum_{k=1}^N X_t^{k,N}$, on s'intéresse donc au processus $\left(Y^{i,N}, V^{i,N}\right)$ solution du système

$$\begin{cases} \mathrm{d}Y_t^{i,N} &= \left(V_t^{i,N} - \frac{1}{N}\sum_{k=1}^N V_t^{k,N}\right)\mathrm{d}t \\ \mathrm{d}V_t^{i,N} &= -2V_{t^-}^{i,N} \int \mathbf{1}_{z\leq\lambda(Y_t^{i,N} V_{t^-}^{i,N}))}\mathcal{N}^i(\mathrm{d}z, \mathrm{d}t). \end{cases} \tag{1.15}$$

L'EDS non-linéaire associée est alors la suivante :

$$\begin{cases} \mathrm{d}Y_t &= (V_t - \mathbb{E}[V_t])\mathrm{d}t \\ \mathrm{d}V_t &= -2\int V_{t^-}\mathbb{1}_{z\leq\lambda(Y_t V_{t^-})}\mathcal{N}(\mathrm{d}z, \mathrm{d}t). \end{cases} \tag{1.16}$$

Les preuves d'existence et d'unicité et de propagation du chaos se font alors de la même manière que dans le cas non-recentré, et on a donc le résultat suivant.

Théorème 1.4.11. *Supposons que* $(\mathcal{H}_\lambda)$ *est satisfaite et que* Y_0 *vérifie* $(\mathcal{H}_\alpha)$. *Alors il y a existence et unicité trajectorielle pour le système stochastique* (1.16) *de condition initiale* $(Y_0, V_0) \in \mathbb{R} \times \{-1, 1\}$.
De plus, il y a propagation du chaos du système de particules (1.15) *vers la solution du système* (1.16), *à vitesse* $\sqrt{N}$.

On appelle processus Zig-zag non-linéaire centré l'unique solution du système non-linéaire (1.16). L'objet d'étude est alors maintenant le comportement en temps long de ce processus.
Nous faisons alors l'hypothèse suivante, classique pour démontrer l'ergodicité d'un processus Zig-zag (voir [Bierkens et Roberts, 2017] et [Fontbona *et al.*, 2012] par exemple) :

Hypothèse $(\mathcal{A})$ **:**

Il existe $y_0 > 0$ et $\lambda_{\min} > 0$ tels que $\lambda(y) \geq \lambda_{\min}$ pour tout $y \geq y_0$ et tels que

$$\overline{\lambda}_0 := \inf_{y \geq y_0} \lambda(y) > \sup_{y \leq -y_0} \lambda(y) =: \underline{\lambda}_0.$$

Nous discutons alors du résultat suivant, qui est une conjecture pour le moment.

Conjecture. *Supposons que les hypothèses* $(\mathcal{H}_\lambda)$ *et* $(\mathcal{A})$ *sont satisfaites.*
Soit (Y, V) *le processus Zig-zag non-linéaire centré, c'est-à-dire la solution de* (1.16)*, avec* $\mathbb{E}[Y_0] = 0$*. Alors* (Y, V) *possède une unique mesure de probabilité invariante* ν_0 *définie par :*

$$\nu_0(\mathrm{d}z, \mathrm{d}w) = \frac{1}{C_0} \mathrm{e}^{-F_0(z)} \mathrm{d}z \otimes \frac{1}{2} \left(\delta_1 + \delta_{-1}\right)(\mathrm{d}w)$$

où

$$F_0(z) = \int_0^z \left(\lambda(u) - \lambda(-u)\right) \mathrm{d}u \quad \text{et} \quad C_0 = \int_{\mathbb{R}} \exp(-F_0(z)) \mathrm{d}z.$$

De plus, il converge en temps long vers cette mesure de probabilité.

Ce travail étant encore un travail en cours, nous donnons à la fin du Chapitre 4 des pistes pour démontrer cette conjecture, et nous discutons des perspectives de travail sur ce sujet.

Chapitre 2

Explicit speed of convergence of the stochastic billiard in a convex set

This Chapter is the reproduction of the paper [Fétique, 2019], accepted in the journal Séminaire de Probabilités..

In this paper, we are interested in the speed of convergence of the stochastic billiard evolving in a convex set K. This process can be described as follows : a particle moves at unit speed inside the set K until it hits the boundary, and is randomly reflected, independently of its position and previous velocity. We focus on convex sets in $\mathbb{R}^2$ with a curvature bounded from above and below. We give an explicit coupling for both the continuous-time process and the embedded Markov chain of hitting points on the boundary, which leads to an explicit speed of convergence to equilibrium.

2.1 Introduction

In this paper, our goal is to give explicit bounds on the speed of convergence of a process, called "stochastic billiard", towards its invariant measure, under some assumptions that we will detail further. This process can be informally described as follows : a particle moves at unit speed inside a domain until it hits the boundary. At this time, the particle is reflected inside the domain according to a random distribution on the unit sphere, independently on its position and previous velocity.
The stochastic billiard is a generalisation of shake-and-bake algorithm (see [Boender *et al.*, 1991]), in which the reflection law is the cosine law. In that case, it has been proved that the Markov chain of hitting points on the boundary has a uniform stationary distribution. In [Boender *et al.*, 1991], the shake-and-bake

algorithm is introduced for generating uniform points on the boundary of bounded polyhedra. More generally, stochastic billiards can be used for sampling from a bounded set or the boundary of such a set, through the Markov Chain Monte Carlo algorithms. In that sense, it is therefore important to have an idea of the speed of convergence of the process towards its invariant distribution.

Stochastic billiards have been studied a lot, under different assumptions on the domain in which it lives and on the reflection law. Let us mention some of these works. In [Evans, 2001], Evans considers the stochastic billiard with uniform reflection law in a bounded d-dimensional region with C^1 boundary, and also in polygonal regions in the plane. He proves first the exponentially fast total variation convergence of the Markov chain, and moreover the uniform total variation Césaro convergence for the continuous-time process. In [Dieker et Vempala, 2015], the authors only consider the stochastic billiard Markov chain, in a bounded convex set with curvature bounded from above and with a cosine distribution for the reflection law. They give a bound for the speed of convergence of this chain towards its invariant measure, that is the uniform distribution on the boundary of the set, in order to get a bound for the number of steps of the Markov chain required to sample approximatively the uniform distribution. Finally, let us mention the work of Comets, Popov, Schütz and Vachkovskaia [Comets *et al.*, 2009], in which some ideas have been picked and used in the present paper. They study the convergence of the stochastic billiard and its associated Markov chain in a bounded domain in $\mathbb{R}^d$ with a boundary locally Lipschitz and almost everywhere C^1. They consider the case of a reflection law which is absolutely continuous with respect to the Haar measure on the unit sphere of $\mathbb{R}^d$, and supported on the whole half-sphere that points into the domain. They show the exponential ergodicity of the Markov chain and the continuous-time process and also their Gaussian fluctuations. The particular case of the cosine reflection law is discussed. Even if they do not give speeds of convergence, their proofs could lead to explicit speeds if we write them in particular cases (as for the stochastic billiard in a disc of $\mathbb{R}^2$ for instance). However, as we will mention in Section 2.2.3, the speed of convergence obtained in particular cases will not be relevant, since their proof is adapted to their very general framework, and not for more particular and simple domains.

The goal of this paper is to give explicit bounds on the speed of convergence of the stochastic billiard and its embedded Markov chain towards their invariant measures. For that purpose, we are going to give an explicit coupling of which we can estimate the coupling time.
In a first part, we study the particular case of the billiard in a disc. In that case, everything is quite simple since all the quantities can be explicitly expressed.
Then, in a second part, we extend the results for the case of the stochastic billiard in a compact convex set of $\mathbb{R}^2$ with curvature bounded from above and below. In that case, we can no more do explicit computations on the quantities describing the process, since we do not know exactly the geometry of the convex set. However, thanks to the assumptions on the curvature, we are able to estimate the needed quantities.

In both cases, the disc and the convex set, we suppose that the reflection law has a density function which is bounded from below by a strictly positive constant on a part of the sphere. The speed of convergence will obviously depend on it. However, for the convergence of the stochastic billiard process in a convex set, we will need to suppose that the reflection law is supported on the whole half sphere that points inside the domain.
At the end of this paper, we briefly discuss the extension of the results to higher dimensions.

Notations

We introduce some notations used in the paper :

— for $A \subset \mathbb{R}$, $\mathbf{1}_A$ denotes the indicator function of the set A, that is $\mathbf{1}_A(x)$ is equal to 1 if $x \in A$ and 0 otherwise ;
— for $x \in \mathbb{R}$, $\lfloor x \rfloor$ denotes the floor of the real x ;
— for $x, y \in \mathbb{R}^2$, we note by $\|x\|$ the euclidean norm of x and we write $\langle x, y \rangle$ for the scalar product of x and y ;
— for $A \subset \mathbb{R}^2$, ∂A denotes the boundary of the set A ;
— $\mathcal{B}_r$ denotes the closed ball of $\mathbb{R}^2$ centred at the origin with radius r, i.e. $\mathcal{B}_r = \{x \in \mathbb{R}^2 : \|x\| \leq r\}$, and $\mathbb{S}^1$ denotes the unit sphere of $\mathbb{R}^2$, i.e. $\mathbb{S}^1 = \{x \in \mathbb{R}^2 : \|x\| = 1\}$;
— for $\mathcal{I} \subset \mathbb{R}$, $|\mathcal{I}|$ denotes the Lebesgue measure of the set $\mathcal{I}$;
— for $K \subset \mathbb{R}^2$ a compact convex set, we consider the 1-dimensional Hausdorff measure in $\mathbb{R}^2$ restricted to ∂K. Therefore, if $A \subset \partial K$, $|A|$ denotes this Hausdorff measure of A ;
— for $A \subset \mathbb{R}^2$, if $x \in \partial A$, we write n_x the unitary normal vector of ∂A at x pointing to the interior of A and we define $\mathbb{S}_x$ the set of vectors that point to the interior of A : $\mathbb{S}_x = \{v \in \mathbb{S}^1 : \langle v, n_x \rangle \geq 0\}$;
— if two random variables X and Y are equal in law we write $X \overset{\mathcal{L}}{=} Y$, and we write $X \sim \mu$ to say that the random variable X has μ for law, or simply $\mathcal{L}(X)$ to nominate the law of X ;
— we denote by $\mathcal{G}(p)$ the geometric law with parameter p.

2.2 Coupling for the stochastic billiard

2.2.1 Generalities on coupling

In order to describe the way we will prove the exponential convergences and obtain bounds on the speeds of convergence, we first need to introduce some notions.
Let ν and $\widetilde{\nu}$ be two probability measures on a measurable space E. We say that a probability measure on $E \times E$ is a coupling of ν and $\widetilde{\nu}$ if its marginals are ν and $\widetilde{\nu}$. Denoting by $\Gamma(\nu, \widetilde{\nu})$ the set of all the couplings of ν and $\widetilde{\nu}$, we say that two random variables Y and $\widetilde{Y}$ satisfy $(Y, \widetilde{Y}) \in \Gamma(\nu, \widetilde{\nu})$ if ν and $\widetilde{\nu}$ are

the respective laws of Y and $\widetilde{Y}$. The total variation distance between these two probability measures is then defined by

$$\|\nu - \widetilde{\nu}\|_{TV} = \inf_{(Y,\widetilde{Y})\in\Gamma(\nu,\widetilde{\nu})} \mathbb{P}(Y \neq \widetilde{Y}).$$

For other equivalent definitions of the total variation distance and its properties, see for instance [Lindvall, 2002].
Let $(Y_t)_{t\geq 0}$ and $(\tilde{Y}_t)_{t\geq 0}$ be two Markov processes. A coupling $((Y_t, \tilde{Y}_t))_{t\geq 0}$ is called a coalescent coupling if there exists an almost surely finite random time T, such that $Y_{T+s} = \tilde{Y}_{T+s}$ for all $s \geq 0$. In that case, $T_c = \inf\left\{t \geq 0 : Y_t = \tilde{Y}_t\right\}$ is called the coupling time of Y and $\tilde{Y}$, and from the definition of the total variation distance, it immediately follows that

$$\|\mathcal{L}(Y_t) - \mathcal{L}(\tilde{Y}_t)\|_{TV} \leq \mathbb{P}\left(T_c > t\right).$$

Therefore, let T^* be a random variable stochastically bigger than T_c, $T_c \leq_{st} T^*$, which means that $\mathbb{P}\left(T_c \leq t\right) \geq \mathbb{P}\left(T^* \leq t\right)$ for all $t \geq 0$. If T^* has a finite exponential moment, Markov's inequality gives then, for any λ such that the Laplace transform of T^* is well defined :

$$\|\mathcal{L}(Y_t) - \mathcal{L}(\tilde{Y}_t)\|_{TV} \leq \mathbb{P}\left(T^* > t\right) \leq \mathrm{e}^{-\lambda t}\mathbb{E}\left[\mathrm{e}^{\lambda T^*}\right].$$

Thus, if we manage to stochastically bound the coupling time of two stochastic billiards by a random time whose Laplace transform can be estimated, we get an exponential bound for the speed of convergence of the stochastic billiard towards its invariant measure.

Let us now speak about maximal coupling, a result that we will use a lot in the proofs of our main results. Let us consider μ and ν two probability distributions on $\mathbb{R}^d$ with respective density functions f and g with respect to the Lebesgue measure. Let us suppose that there exists a constant $c > 0$ and an interval I such that for all $x \in I$, $f(x) \geq c$ and $g(x) \geq c$. Then, there exists a coupling (X, Y) (called a maximal coupling) of μ and ν such that $\mathbb{P}\left(X = Y\right) \geq c|I|$. For more details, see for instance Section 4 of Chapter 1 of [Thorisson, 2000].

We end this part with a definition that we will use throughout this paper.

Definition 2.2.1. Let $K \subset \mathbb{R}^2$ be a compact convex set.
We say that a pair of random variables (X, T) living in $\partial K \times \mathbb{R}^+$ is α-continuous on the set $A \times B \subset \partial K \times \mathbb{R}^+$ if for any measurable $A_1 \subset A$, $B_1 \subset B$:

$$\mathbb{P}\left(X \in A_1, T \in B_1\right) \geq \alpha|A_1||B_1|.$$

We can also adapt this definition for a single random variable.

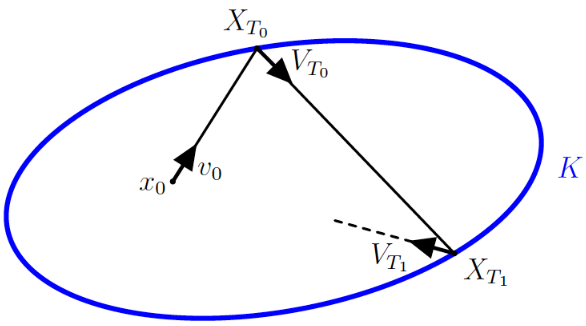


FIGURE 2.1 – A trajectory of the stochastic billiard in a set K, starting in the interior of K

2.2.2 Description of the process

Let us now give a precise description of the stochastic billiard $((X_t, V_t))_{t\geq 0}$ in a set K.
We assume that $K \subset \mathbb{R}^2$ is a compact convex set with a boundary at least C^1.
Let $e = (1, 0)$ be the first coordinate vector of the canonical basis of $\mathbb{R}^2$. We consider a law γ on the half-sphere $\mathbb{S}_e = \{v \in \mathbb{S}^1 : e \cdot v \geq 0\}$. Let moreover $(U_x, x \in \partial K)$ be a family of rotations of $\mathbb{S}^1$ such that $U_x e = n_x$, where we recall that n_x is the normal vector of ∂K at x pointing to the interior of K.
Let $(\eta_n)_{n\geq 0}$ be a sequence of i.i.d. random variables on $\mathbb{S}_e$ with law γ.
Given $(x_0, v_0) \in K \times \mathbb{S}^1$, we consider the process $((X_t, V_t))_{t\geq 0}$ living in $K \times \mathbb{S}^1$ constructed as follows (see Figure 2.1) :

- If $x_0 \in K\backslash\partial K$, let $(X_0, V_0) = (x_0, v_0)$ and let $T_0 = \inf\{t > 0 : x_0 + tv_0 \notin K\}$. For $t \in [0, T_0)$ let then $X_t = x_0 + tv_0$ and $V_t = v_0$. Else, i.e. if $x_0 \in \partial K$, let $T_0 = 0$.
- Let $X_{T_0} = x_0 + T_0 v_0$, and $V_{T_0} = U_{X_0}\eta_0$.
- Let $\tau_1 = \inf\{t > 0 : X_{T_0} + tV_{T_0} \notin K\}$ and define $T_1 = \tau_1 + T_0$. We put $X_t = X_{T_0} + tV_{T_0}$, $V_t = V_{T_0}$ for $t \in [T_0, T_1)$, and $X_{T_1} = X_{T_0} + \tau_1 V_{T_0}$. Then, let $V_{T_1} = U_{X_{T_1}}\eta_1$.
- Let $\tau_2 = \inf\{t > 0 : X_{T_1} + tV_{T_1} \notin K\}$ and define $T_2 = T_1 + \tau_2$. We put $X_t = X_{T_1} + tV_{T_1}$, $V_t = V_{T_1}$ for $t \in [T_1, T_2)$, and $X_{T_2} = X_{T_1} + \tau_2 V_{T_1}$. Then, let $V_{T_2} = U_{X_{T_2}}\eta_2$.
- And we start again ...

As mentioned in the introduction $(X_{T_n})_{n\geq 0}$ is a Markov chain living in ∂K and the process $((X_t, V_t))_{t\geq 0}$ is a Markov process living in $K \times \mathbb{S}^1$.

For $x \in \partial K$, it is equivalent to consider the new speed in $\mathbb{S}_x$ or to consider the angle in $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ between this vector speed and the normal vector n_x.

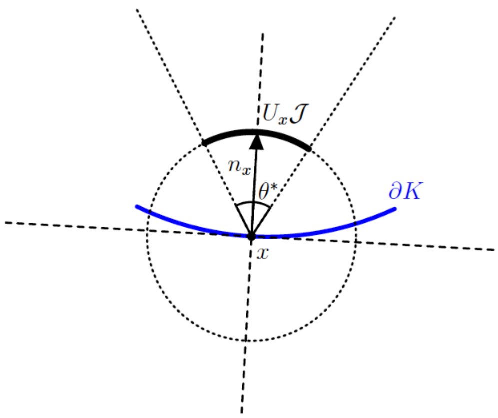


FIGURE 2.2 – Illustration of Assumptions $(\mathcal{H})$ and $(\mathcal{A})$

For $n \geq 1$, we thus denote by Θ_n the random variable in $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ such that $r_{X_{T_n},\Theta_n}(n_{X_{T_n}}) \overset{\mathcal{L}}{=} V_{T_n}$, where for $x \in \partial K$ and $\theta \in \mathbb{R}$, $r_{x,\theta}$ denotes the rotation with center x and angle θ.
We make the following assumption on γ (see Figure 2.2) :

Assumption $(\mathcal{H})$ **:**

The law γ has a density function ρ with respect to the Haar measure on $\mathbb{S}_e$, which satisfies : there exist $\mathcal{J}$ an open subset of $\mathbb{S}_e$, containing e and symmetric with respect to e, and $\rho_{\min} > 0$ such that :

$$\rho(u) \geq \rho_{\min}, \quad \text{for all } u \in \mathcal{J}.$$

This assumption is equivalent to the following one on the variables $(\Theta_n)_{n\geq 0}$:

Assumption $(\mathcal{H}')$ **:**

The variables Θ_n, $n \geq 0$, have a density function f with respect to the Lebesgue measure on $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ satisfying : there exist $f_{\min} > 0$ and $\theta^* \in (0, \pi)$ such that :

$$f(\theta) \geq f_{\min}, \quad \text{for all } \theta \in \left[-\frac{\theta^*}{2}, \frac{\theta^*}{2}\right].$$

In fact, since these two assumptions are equivalent, we have

$$\rho_{\min} = f_{\min} \quad \text{and} \quad |\mathcal{J}| = \theta^*.$$

In the sequel, we will use both descriptions of the speed vector depending on which is the most suitable.

2.2.3 A coupling for the stochastic billiard

Let us now informally describe the idea of the couplings used to explicit the speeds of convergence of our processes to equilibrium. They will be explained explicitly in Sections 2.3 and 2.4.

To get a bound on the speed of convergence of the Markov chain recording the location of hitting points on the boundary of the stochastic billiard, the strategy is the following. We consider two stochastic billiard Markov chains with different initial conditions. We estimate the number of steps that they have to do before they have a strictly positive probability to arrive on the same place at a same step. In particular, it is sufficient to know the number of steps needed before the position of each chain charges the half of the boundary of the set on which they evolve. Then, their coupling time is stochastically smaller than a geometric time whose Laplace transform is known.
The case of the continuous-time process is a bit more complicated. To couple two stochastic billiards, it is not sufficient to make them cross in the interior of the set where they live. Indeed, if they cross with a different speed, then they will not be equal after. So the strategy is to make them arrive at the same place on the boundary of the set at the same time, and then they can always keep the same velocity and stay equal. We will do this in two steps. First, we will make the two processes hit the boundary at the same time, but not necessarily at the same point. This will take some random time, that we will be able to quantify. And secondly, with some strictly positive probability, after two bounces, the two processes will have hit the boundary at the same point at the same time. We repeat the scheme until the first success. This leads us to a stochastic upper bound for the coupling time of two stochastic billiards.

Obviously, the way that we couple our processes is only one way to do that, and there are as many as we want. Let us for instance describe the coupling constructed in [Comets *et al.*, 2009]. Consider two stochastic billiard processes evolving in the set K with different initial conditions. Their first step is to make the processes hit the boundary in the neighbourhood of a good $x_1 \in \partial K$. This can be done after n_0 bounces, where n_0 is the minimum number of bounces needed to connect any two points of the boundary of K. Once the two processes have succeeded, they are in the neighbourhood of x_1, but at different times. Then, the strategy used by the authors of [Comets *et al.*, 2009] is to make the two processes do round trips between the neighbourhood of x_1 and the neighbourhood of another good $y_1 \in \partial K$. Thereby, if the point y_1 is well chosen, the time difference between the two processes decreases gradually, while the positions of the processes stay the same after one round trip. However, the number of round trips needed to compensate for the possibly big difference of times could be very high. This particular coupling is therefore well adapted for sets whose boundary can be quite "chaotic", but not for convex sets with smooth boundary as we consider in this paper.

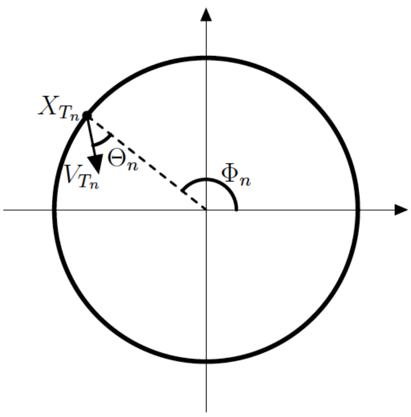


FIGURE 2.3 – Definition of the variables Φ_n and Θ_n in bijection with the variables X_{T_n} and V_{T_n}

2.3 Stochastic billiard in the disc

In this section, we consider the particular case where K is a ball : $K = \mathcal{B}_r$, for some fixed $r > 0$.
In that case, for each $n \geq 0$, the couple $(X_{T_n}, V_{T_n}) \in \partial\mathcal{B}_r \times \mathbb{S}^1$ can be represented by a couple $(\Phi_n, \Theta_n) \in [0, 2\pi) \times \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ as follows (see Figure 2.3) :

- to a position x on $\partial\mathcal{B}_r$ corresponds a unique angle $\phi \in [0, 2\pi)$. The variable Φ_n nominates this unique angle associated to X_{T_n}, i.e. (r, Φ_n) are the polar coordinates of X_{T_n}.
- at each speed V_{T_n} we associate the variable Θ_n introduced in Section 2.2.2, satisfying Assumption $(\mathcal{H}')$.

Remark that for all $n \geq 0$, the random variable Θ_n is independent of Φ_k for all $k \in \{0, \cdots, n\}$. We also recall that the variables Θ_n, $n \geq 0$, are all independent.

In the sequel, we do not care about the congruence modulo 2π : it is implicit that when we write Φ, we consider its representative in $[0, 2\pi)$.

Let us state the following proposition that links the different random variables together.

Proposition 2.3.1. *For all $n \geq 1$ we have :*

$$\tau_n = 2r\cos(\Theta_{n-1}) \quad \textit{and} \quad \Phi_n = \pi + 2\Theta_{n-1} + \Phi_{n-1} \tag{2.1}$$

Proof. The relationships are immediate with geometric considerations. □

2.3.1 The embedded Markov chain

In this section, the goal is to obtain a control of the speed of convergence of the stochastic billiard Markov chain on the circle. For this purpose, we study the distribution of the position of the Markov chain at each step.

Let $\Phi_0 = \phi_0 \in [0, 2\pi)$.

Proposition 2.3.2. *Let $(\Phi_n)_{n\geq 0}$ be the stochastic billiard Markov chain evolving on $\partial\mathcal{B}_r$, satisfying assumption $(\mathcal{H}')$. Let us denote by f_{Φ_n} the density function of Φ_n, for $n \geq 1$.*
We have

$$f_{\Phi_1}(u) \geq \frac{f_{\min}}{2}, \quad \forall u \in \mathcal{I}_1 = [\pi - \theta^* + \phi_0, \pi + \theta^* + \phi_0] .$$

Moreover, for all $n \geq 2$, for all $\eta_2, \cdots, \eta_n$ such that $\eta_2 \in (0, 2\theta^*)$, and for $k \in \{2, \cdots, n-1\}$, $\eta_{k+1} \in \left(0, \min\left\{2\theta^*; k\theta^* - \sum_{\ell=2}^{k} \eta_\ell\right\}\right)$, we have

$$f_{\Phi_n}(u) \geq \left(\frac{f_{\min}}{2}\right)^n \prod_{k=2}^{n} \eta_k,$$
$$\forall u \in \mathcal{I}_n = \left[n(\pi - \theta^*) + \phi_0 + \sum_{k=2}^{n} \eta_k, n(\pi + \theta^*) + \phi_0 - \sum_{k=2}^{n} \eta_k\right].$$

Proof. Since the Markov chain is rotationally symmetric, we do the computations with $\phi_0 = 0$.

— Case $n = 1$:
We have, thanks to (2.1), $\Phi_1 = \pi + 2\Theta_0 + \phi_0 = \pi + 2\Theta_0$. Thus, for any measurable bounded function g, we get :

$$\mathbb{E}\left[g(\Phi_1)\right] = \mathbb{E}\left[g\left(\pi + 2\Theta_0\right)\right] = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} g\left(\pi + 2x\right) f(x)\mathrm{d}x$$
$$\geq f_{\min} \int_{-\frac{\theta^*}{2}}^{\frac{\theta^*}{2}} g\left(\pi + 2x\right)\mathrm{d}x = \frac{f_{\min}}{2} \int_{\pi-\theta^*}^{\pi+\theta^*} g(u)\mathrm{d}u.$$

We deduce :

$$f_{\Phi_1}(u) \geq \frac{f_{\min}}{2}, \quad \forall u \in [\pi - \theta^*, \pi + \theta^*] .$$

— Induction : let us suppose that for some $n \geq 1$, $f_{\Phi_n}(u) \geq c_n$ for all $u \in [a_n, b_n]$. Then, using the relationship (2.1) and the independence between Θ_n and Φ_n we have, for any measurable bounded function g :

$$\mathbb{E}\left[g(\Phi_{n+1})\right] = \mathbb{E}\left[g(\pi + 2\Theta_n + \Phi_n)\right]$$
$$\geq f_{\min} c_n \int_{-\frac{\theta^*}{2}}^{\frac{\theta^*}{2}} \int_{a_n}^{b_n} g(\pi + 2\theta + x)\mathrm{d}x\mathrm{d}\theta.$$

Using the substitution $u = \pi + 2\theta + x$ in the integral with respect to x and Fubini's theorem, we have :

$$\mathbb{E}\left[g(\Phi_{n+1})\right] \geq f_{\min} c_n \int_{\pi-\theta^*+a_n}^{\pi+\theta^*+b_n} \left(\int_{-\frac{\theta^*}{2}}^{\frac{\theta^*}{2}} \mathbf{1}_{\frac{1}{2}(u-\pi-b_n)\leq\theta\leq\frac{1}{2}(u-\pi-a_n)} \mathrm{d}\theta \right) g(u)\mathrm{d}u,$$

and we deduce the following lower bound of the density function $f_{\Phi_{n+1}}$ of Φ_{n+1} :

$$f_{\Phi_{n+1}}(u) \geq f_{\min} c_n \left| \left[-\frac{\theta^*}{2}, \frac{\theta^*}{2} \right] \cap \left[\frac{1}{2}\left(u - \pi - b_n\right), \frac{1}{2}\left(u - \pi - a_n\right) \right] \right|,$$

for all $u \in [\pi - \theta^* + a_n, \pi + \theta^* + b_n]$.
When u is equal to one extremal point of this interval, this lower bound is equal to 0. However, let $\eta_{n+1} \in \left(0, \min\left\{2\theta^*; \frac{1}{2}(b_n - a_n)\right\}\right)$. Then the intersection $\left[-\frac{\theta^*}{2}, \frac{\theta^*}{2}\right] \cap \left[\frac{1}{2}\left(u - \pi - b_n\right), \frac{1}{2}\left(u - \pi - a_n\right)\right]$ is non-empty, and we have, for
$u \in [\pi - \theta^* + a_n + \eta_{n+1}, \pi + \theta^* + b_n - \eta_{n+1}]$:

$$f_{\Phi_{n+1}}(u) \geq f_{\min} c_n \min\left\{\theta^*; \frac{\eta_{n+1}}{2}\right\} = f_{\min} c_n \frac{\eta_{n+1}}{2}.$$

The result follows immediately. □

By choosing a constant sequence for the η_k, $k \geq 2$ in the Proposition 2.3.2, we immediately deduce :

Corollaire 2.3.3. *For all $n \geq 2$, for all $\varepsilon \in (0, \theta^*)$, we have*

$$f_{\Phi_n}(u) \geq \left(\frac{f_{\min}}{2}\right)^n \varepsilon^{n-1},$$

$\forall u \in \mathcal{J}_n = [n(\pi - \theta^*) + \phi_0 + (n-1)\varepsilon, n(\pi + \theta^*) + \phi_0 - (n-1)\varepsilon]$.

Let $(\mathcal{J}_n)_{n\geq 2}$ defined as in Corollary 2.3.3. We put $\mathcal{J}_1 = \mathcal{I}_1$ with $\mathcal{I}_1$ defined in Proposition 2.3.2.

Theorem 2.3.4. *Let $(\Phi_n)_{n\geq 0}$ be the stochastic billiard Markov chain on the circle $\partial\mathcal{B}_r$, satisfying Assumption $(\mathcal{H}')$.*
There exists a unique invariant probability measure ν on $[0, 2\pi)$ for the Markov chain $(\Phi_n)_{n\geq 0}$, and we have :

1. *if $\theta^* > \frac{\pi}{2}$, for all $n \geq 0$,*

$$\|\mathbb{P}\left(\Phi_n \in \cdot\right) - \nu\|_{TV} \leq \left(1 - f_{\min}(2\theta^* - \pi)\right)^{n-1},$$

2. *if* $\theta^* \leq \frac{\pi}{2}$, *for all* $n \geq 0$ *and all* $\varepsilon \in (0, \theta^*)$,

$$\|\mathbb{P}(\Phi_n \in \cdot) - \nu\|_{TV} \leq (1-\alpha)^{\frac{n}{n_0}-1},$$

where

$$n_0 = \left\lfloor \frac{\pi - 2\varepsilon}{2(\theta^* - \varepsilon)} \right\rfloor + 1 \text{ and } \alpha = \left(\frac{\varepsilon}{2}\right)^{n_0-1} {f_{\min}}^{n_0} \left(2n_0\theta^* - 2(n_0-1)\varepsilon - \pi\right).$$

Proof. The existence of the invariant measure is immediate thanks to the compactness of $\partial\mathcal{B}_r$ (see [Ethier et Kurtz, 1986]). The following proof leads to its uniqueness and the speed of convergence.
Let $((\Phi_n, \Theta_n))_{n\geq 0}$ and $((\tilde{\Phi}_n, \tilde{\Theta}_n))_{n\geq 0}$ be two versions of the process described above, with initial positions ϕ_0 and $\tilde{\phi}_0$ on $\partial\mathcal{B}_r$.
In order to couple Φ_n and $\tilde{\Phi}_n$ at some time n, it is sufficient to show that the intervals $\mathcal{J}_n$ and $\tilde{\mathcal{J}}_n$ corresponding to Corollary 2.3.3 have a non empty intersection. Since these intervals are included in $[0, 2\pi)$, a sufficient condition to have $\mathcal{J}_n \cap \tilde{\mathcal{J}}_n \neq \varnothing$ is that the length of these two intervals is strictly bigger than π.

Let $\varepsilon \in (0, \theta^*)$. We have

$$|\mathcal{J}_1| = |\tilde{\mathcal{J}}_1| = 2\theta^*,$$

and for $n \geq 2$,

$$|\mathcal{J}_n| = |\tilde{\mathcal{J}}_n| = 2n\theta^* - 2(n-1)\varepsilon.$$

Therefore the length of $\mathcal{J}_n$ is a strictly increasing function of n (which is intuitively clear).

— Case 1 : $\theta^* > \frac{\pi}{2}$. In that case we have $|\mathcal{J}_1| = |\tilde{\mathcal{J}}_1| > \pi$. Therefore we can construct a coupling $\left(\Phi_1, \tilde{\Phi}_1\right)$ (see the reminder on maximal coupling in Section 2.2.1) such that we have, using Proposition 2.3.2 :

$$\begin{aligned}\mathbb{P}\left(\Phi_1 = \tilde{\Phi}_1\right) &\geq \frac{f_{\min}}{2}\left|\mathcal{J}_1 \cap \tilde{\mathcal{J}}_1\right| \\ &\geq \frac{f_{\min}}{2} 2(2\theta^* - \pi) \\ &= f_{\min}(2\theta^* - \pi).\end{aligned}$$

— Case 2 : $\theta^* \leq \frac{\pi}{2}$. Here we need more jumps before having a positive probability to couple Φ_n and $\tilde{\Phi}_n$. Let us thus define

$$n_0 = \min\{n \geq 2 : 2n\theta^* - 2(n-1)\varepsilon > \pi\} = \left\lfloor \frac{\pi - 2\varepsilon}{2(\theta^* - \varepsilon)} \right\rfloor + 1.$$

Using the lower bound of the density function of Φ_{n_0} obtained in Corollary 2.3.3, we deduce that we can construct a coupling $\left(\Phi_{n_0}, \tilde{\Phi}_{n_0}\right)$ such that :

$$\mathbb{P}\left(\Phi_{n_0} = \tilde{\Phi}_{n_0}\right) \geq \left(\frac{f_{\min}}{2}\right)^{n_0} \varepsilon^{n_0-1} \left|\mathcal{J}_{n_0} \cap \tilde{\mathcal{J}}_{n_0}\right|$$

$$\geq \left(\frac{f_{\min}}{2}\right)^{n_0} \varepsilon^{n_0-1} 2\left(2n_0\theta^* - 2(n_0-1)\varepsilon - \pi\right)$$
$$= \left(\frac{\varepsilon}{2}\right)^{n_0-1} (f_{\min})^{n_0} \left(2n_0\theta^* - 2(n_0-1)\varepsilon - \pi\right).$$

To treat both cases together, let us define

$$m_0 = \mathbf{1}_{\theta^* > \frac{\pi}{2}} + \left(\left\lfloor \frac{\pi - 2\varepsilon}{2(\theta^* - \varepsilon)} \right\rfloor + 1\right) \mathbf{1}_{\theta^* \leq \frac{\pi}{2}}.$$

and

$$\alpha = f_{\min}(2\theta^* - \pi)\mathbf{1}_{\theta^* > \frac{\pi}{2}} + \left(\frac{\varepsilon}{2}\right)^{m_0-1} (f_{\min})^{m_0} \left(2m_0\theta^* - 2(m_0-1)\varepsilon - \pi\right) \mathbf{1}_{\theta^* \leq \frac{\pi}{2}}.$$

Since our processes $(\Phi_n)_{n\geq 0}$ and $(\tilde{\Phi}_n)_{n\geq 0}$ are Markovian processes, once they are equal, we can let them equal afterwards. And then we get :

$$\begin{aligned}
\|\mathbb{P}(\Phi_n \in \cdot) - \nu\|_{TV} &\leq \mathbb{P}\left(\Phi_n \neq \tilde{\Phi}_n\right) \\
&\leq \mathbb{P}\left(\Phi_{\lfloor \frac{n}{m_0} \rfloor m_0} \neq \tilde{\Phi}_{\lfloor \frac{n}{m_0} \rfloor m_0}\right) \\
&\leq (1-\alpha)^{\lfloor \frac{n}{m_0} \rfloor} \\
&\leq (1-\alpha)^{\frac{n}{m_0} - 1}.
\end{aligned}$$

□

2.3.2 The continuous-time process

We assume here that the constant θ^* introduced in Assumption $(\mathcal{H}')$ satisfies

$$\theta^* \in \left(\frac{2\pi}{3}, \pi\right).$$

This condition on θ^* is essential in the proof of Theorem 2.3.7 to couple our processes with "two jumps". However, if $\theta^* \in \left(0, \frac{2\pi}{3}\right]$ we can adapt our method (see Remark 2.3.9).

Notation : We define a sequence $(S_n)_{n\geq 0}$ of random times by

$$S_0 = 0 \quad \text{and for } n \geq 1, \;\; S_n = T_n - T_0.$$

By this way, we avoid the presence of the time T_0 in the computations, and the law of the random time S_n is the law of the instant of the n^{th} bounce, when starting on the boundary of $\mathcal{B}_r$. If $T_0 = 0$, that is if the process starts on the boundary of $\mathcal{B}_r$, then $(S_n)_{n\geq 0} = (T_n)_{n\geq 0}$.
We observe that thanks to the rotational symmetry of the process, for all $m, n, p \geq 0$ we have the following equality in law : $S_{n+m} - S_m \overset{\mathcal{L}}{=} S_n \overset{\mathcal{L}}{=} T_{n+p} - T_p$.

Proposition 2.3.5. *Let $((X_t, V_t))_{t\geq 0}$ be the stochastic billiard process in the ball $\mathcal{B}_r$ satisfying Assumption $(\mathcal{H}')$ with $\theta^* \in \left(\frac{2\pi}{3}, \pi\right)$.*
We denote by f_{S_2} the density function of S_2. Let $\eta \in \left(0, 2r\left(1-\cos\left(\frac{\theta^*}{2}\right)\right)\right)$. We have

$$f_{S_2}(x) \geq \delta \quad \text{for all } x \in [4r\cos\left(\frac{\theta^*}{2}\right) + \eta, 4r - \eta],$$

where

$$\delta = \frac{2f_{\min}^2}{r\sin\left(\frac{\theta^*}{2}\right)} \min\left\{\frac{\theta^*}{2} - \arccos\left(\cos\left(\frac{\theta^*}{2}\right) + \frac{\eta}{2r}\right); \arccos\left(1 - \frac{\eta}{2r}\right)\right\}. \quad (2.2)$$

Proof. If the density function f is supported on $\left[-\frac{\theta^*}{2}, \frac{\theta^*}{2}\right]$, it is immediate to observe that $4r\cos\left(\frac{\theta^*}{2}\right) \leq S_2 \leq 4r$. But let us be more precise.
Let $g : \mathbb{R} \to \mathbb{R}$ be a bounded measurable function. Thanks to (2.1) we have $S_2 = 2r\left(\cos(\Theta_0) + \cos(\Theta_1)\right)$ with Θ_0, Θ_1 two independent random variables with density function f. Therefore, using Assumption $(\mathcal{H}')$ we have :

$$\begin{aligned}
\mathbb{E}\left[g(S_2)\right] &= \mathbb{E}\left[g\left(2r\left(\cos(\Theta_0) + \cos(\Theta_1)\right)\right)\right] \\
&\geq f_{\min}^2 \int_{-\frac{\theta^*}{2}}^{\frac{\theta^*}{2}} \int_{-\frac{\theta^*}{2}}^{\frac{\theta^*}{2}} g\left(2r\left(\cos(u) + \cos(v)\right)\right) \mathrm{d}u\mathrm{d}v \\
&= 4f_{\min}^2 \int_0^{\frac{\theta^*}{2}} \int_0^{\frac{\theta^*}{2}} g\left(2r\left(\cos(u) + \cos(v)\right)\right) \mathrm{d}u\mathrm{d}v.
\end{aligned}$$

The substitution $x = 2r\left(\cos(u) + \cos(v)\right)$ in the integral with respect to u gives then :

$$\mathbb{E}\left[g(S_2)\right] \geq 4f_{\min}^2 \int_0^{\frac{\theta^*}{2}} \int_{2r\left(\cos\left(\frac{\theta^*}{2}\right) + \cos(v)\right)}^{2r(1+\cos(v))} g(x) \frac{1}{2r\sin\left(\arccos\left(\frac{x}{2r} - \cos(v)\right)\right)} \mathrm{d}x\mathrm{d}v.$$

Fubini's theorem leads to

$$\mathbb{E}\left[g(S_2)\right] \geq \frac{2f_{\min}^2}{r} \int_{4r\cos\left(\frac{\theta^*}{2}\right)}^{4r} \left(\int_0^{\frac{\theta^*}{2}} \frac{1}{\sqrt{1 - \left(\frac{x}{2r} - \cos(v)\right)^2}} \mathbf{1}_{\frac{x}{2r} - 1 < \cos(v) < \frac{x}{2r} - \cos\left(\frac{\theta^*}{2}\right)} \mathrm{d}v\right) g(x)\mathrm{d}x.$$

We then deduce a lower-bound for the density function of S_2 :

$$f_{S_2}(x) \geq \frac{2f_{\min}^2}{r} \int_0^{\frac{\theta^*}{2}} \frac{1}{\sqrt{1 - \left(\frac{x}{2r} - \cos(v)\right)^2}} \mathbf{1}_{\frac{x}{2r} - 1 < \cos(v) < \frac{x}{2r} - \cos\left(\frac{\theta^*}{2}\right)} \mathrm{d}v \mathbf{1}_{x \in \left(4r\cos\left(\frac{\theta^*}{2}\right), 4r\right)}.$$

Let $x \in \left(4r\cos\left(\frac{\theta^*}{2}\right), 4r\right)$. Cutting the interval $\left(4r\cos\left(\frac{\theta^*}{2}\right), 4r\right)$ at point $2r\left(1+\cos\left(\frac{\theta^*}{2}\right)\right)$, we get :

$$\begin{aligned}
f_{S_2}(x) \geq \frac{2f_{\min}^2}{r}&\mathbf{1}_{x\in\left(4r\cos\left(\frac{\theta^*}{2}\right),2r\left(1+\cos\left(\frac{\theta^*}{2}\right)\right)\right]} \\
&\int_0^{\frac{\theta^*}{2}} \frac{1}{\sqrt{1-\left(\frac{x}{2r}-\cos(v)\right)^2}}\mathbf{1}_{\frac{x}{2r}-1<\cos(v)<\frac{x}{2r}-\cos\left(\frac{\theta^*}{2}\right)}\mathrm{d}v \\
+ \frac{2f_{\min}^2}{r}&\mathbf{1}_{x\in\left[2r\left(1+\cos\left(\frac{\theta^*}{2}\right)\right),4r\right)} \\
&\int_0^{\frac{\theta^*}{2}} \frac{1}{\sqrt{1-\left(\frac{x}{2r}-\cos(v)\right)^2}}\mathbf{1}_{\frac{x}{2r}-1<\cos(v)<\frac{x}{2r}-\cos\left(\frac{\theta^*}{2}\right)}\mathrm{d}v \\
= \frac{2f_{\min}^2}{r}&\mathbf{1}_{x\in\left(4r\cos\left(\frac{\theta^*}{2}\right),2r\left(1+\cos\left(\frac{\theta^*}{2}\right)\right)\right]} \\
&\int_{\arccos\left(\frac{x}{2r}-\cos\left(\frac{\theta^*}{2}\right)\right)}^{\frac{\theta^*}{2}} \frac{1}{\sqrt{1-\left(\frac{x}{2r}-\cos(v)\right)^2}}\mathrm{d}v \\
+ \frac{2f_{\min}^2}{r}&\int_0^{\arccos\left(\frac{x}{2r}-1\right)} \frac{1}{\sqrt{1-\left(\frac{x}{2r}-\cos(v)\right)^2}}\mathrm{d}v\mathbf{1}_{x\in\left[2r\left(1+\cos\left(\frac{\theta^*}{2}\right)\right),4r\right)}.
\end{aligned}$$

Then, for $v \in \left(\arccos\left(\frac{x}{2r}-\cos\left(\frac{\theta^*}{2}\right)\right), \frac{\theta^*}{2}\right)$ we have $\cos(v) \leq \frac{x}{2r}-\cos\left(\frac{\theta^*}{2}\right)$, and for $v \in \left(0, \arccos\left(\frac{x}{2r}-1\right)\right)$ we have $\cos(v) \leq 1$. We thus have :

$$\begin{aligned}
f_{S_2}(x) \geq \frac{2f_{\min}^2}{r\sin\left(\frac{\theta^*}{2}\right)}&\left(\frac{\theta^*}{2}-\arccos\left(\frac{x}{2r}-\cos\left(\frac{\theta^*}{2}\right)\right)\right)\mathbf{1}_{x\in\left(4r\cos\left(\frac{\theta^*}{2}\right),2r\left(1+\cos\left(\frac{\theta^*}{2}\right)\right)\right]} \\
&+ \frac{2f_{\min}^2}{r}\frac{\arccos\left(\frac{x}{2r}-1\right)}{\sqrt{\frac{x}{r}\left(1-\frac{x}{4r}\right)}}\mathbf{1}_{x\in\left[2r\left(1+\cos\left(\frac{\theta^*}{2}\right)\right),4r\right)}.
\end{aligned}$$

We can observe that the lower bound of f_{S_2} is strictly positive for $x \in \left(4r\cos\left(\frac{\theta^*}{2}\right), 4r\right)$, but is equal to 0 when x is one of the extremal points of this interval. Let us therefore introduce $\eta \in \left(0, 2r\left(1-\cos\left(\frac{\theta^*}{2}\right)\right)\right)$. We have :

— for $x \in [4r\cos\left(\frac{\theta^*}{2}\right)+\eta, 2r\left(1+\cos\left(\frac{\theta^*}{2}\right)\right)]$ we have

$$\begin{aligned}
&\frac{2f_{\min}^2}{r\sin\left(\frac{\theta^*}{2}\right)}\left(\frac{\theta^*}{2}-\arccos\left(\frac{x}{2r}-\cos\left(\frac{\theta^*}{2}\right)\right)\right) \\
&\qquad\geq \frac{2f_{\min}^2}{r\sin\left(\frac{\theta^*}{2}\right)}\left(\frac{\theta^*}{2}-\arccos\left(\frac{4r\cos\left(\frac{\theta^*}{2}\right)+\eta}{2r}-\cos\left(\frac{\theta^*}{2}\right)\right)\right) \\
&\qquad= \frac{2f_{\min}^2}{r\sin\left(\frac{\theta^*}{2}\right)}\left(\frac{\theta^*}{2}-\arccos\left(\cos\left(\frac{\theta^*}{2}\right)+\frac{\eta}{2r}\right)\right)
\end{aligned}$$

— for $x \in \left[2r\left(1+\cos\left(\frac{\theta^*}{2}\right)\right), 4r-\eta\right]$ we have

$$\begin{aligned}
\frac{2f_{\min}^2}{r}\frac{\arccos\left(\frac{x}{2r}-1\right)}{\sqrt{\frac{x}{r}\left(1-\frac{x}{4r}\right)}} &\geq \frac{2f_{\min}^2}{r}\frac{\arccos\left(\frac{4r-\eta}{2r}-1\right)}{\sqrt{\frac{2r\left(1+\cos\left(\frac{\theta^*}{2}\right)\right)}{r}\left(1-\frac{2r\left(1+\cos\left(\frac{\theta^*}{2}\right)\right)}{4r}\right)}} \\
&= \frac{2f_{\min}^2}{r}\frac{\arccos\left(1-\frac{\eta}{2r}\right)}{\sqrt{\left(1+\cos\left(\frac{\theta^*}{2}\right)\right)\left(1-\cos\left(\frac{\theta^*}{2}\right)\right)}} \\
&= \frac{2f_{\min}^2}{r\sin\left(\frac{\theta^*}{2}\right)}\arccos\left(1-\frac{\eta}{2r}\right).
\end{aligned}$$

The result follows then immediately.

□

Notation : For $x \in \partial\mathcal{B}_r$, we denote by φ_x the unique angle in $[0, 2\pi)$ describing the position of x on $\partial\mathcal{B}_r$. Moreover, we write Φ_n^x for the position of the Markov chain after n steps, and that started at position x on $\partial\mathcal{B}_r$.
Let us remark that thanks to the rotational symmetry of the process in the disc, we do not have to take care of the starting position on ∂K when we look at the inter-jump times.

Proposition 2.3.6. *Let $((X_t, V_t))_{t\geq 0}$ be the stochastic billiard process in $\mathcal{B}_r$ satisfying Assumption $(\mathcal{H}')$ with $\theta^* \in \left(\frac{2\pi}{3}, \pi\right)$.*
For all $\varepsilon \in \left(0, \frac{\theta^*}{4}\right)$, the pair (Φ_2^x, S_2) is $\frac{f_{\min}^2}{2r\sin\left(\frac{\theta^*}{4}\right)}$-continuous on $(\varphi_x - \theta^* + 4\varepsilon, \varphi_x + \theta^* - 4\varepsilon) \times \left(4r\cos\left(\frac{\theta^*}{4}\right), 4r\cos\left(\frac{\theta^*}{4}-\varepsilon\right)\right)$ for all $x \in \partial\mathcal{B}_r$.

Proof. By symmetry of the process, it is sufficient to prove the lemma for $x \in \partial\mathcal{B}_r$ such that $\varphi_x = 0$, what we do.
Let $\varepsilon \in \left(0, \frac{\theta^*}{4}\right)$, $A \subset (-\theta^*+4\varepsilon, \theta^*-4\varepsilon)$ and $(r_1, r_2) \subset \left(2r\cos\left(\frac{\theta^*}{4}\right), 2r\cos\left(\frac{\theta^*}{4}-\varepsilon\right)\right)$.
Let us recall that $\Phi_2^0 = 2\Theta_0 + 2\Theta_1$ and $S_2 = 2r(\cos(\Theta_0)+\cos(\Theta_1))$, where Θ_0, Θ_1 are independent variables with density function f. We thus have :

$$\begin{aligned}
&\mathbb{P}\left(\Phi_2^0 \in A, S_2 \in (r_1, r_2)\right) \\
&\quad = \mathbb{P}\left(2\Theta_0 + 2\Theta_1 \in A, 2r(\cos(\Theta_0)+\cos(\Theta_1)) \in (r_1, r_2)\right) \\
&\quad = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \mathbf{1}_{2u+2v\in A}\mathbf{1}_{\cos(u)+\cos(v)\in\left(\frac{r_1}{2r},\frac{r_2}{2r}\right)} f(u)f(v)\mathrm{d}u\mathrm{d}v \\
&\quad \geq f_{\min}^2\int_{-\frac{\theta^*}{2}}^{\frac{\theta^*}{2}}\int_{-\frac{\theta^*}{2}}^{\frac{\theta^*}{2}} \mathbf{1}_{\frac{u+v}{2}\in\frac{A}{4}}\mathbf{1}_{\cos\left(\frac{u+v}{2}\right)\cos\left(\frac{u-v}{2}\right)\in\left(\frac{r_1}{4r},\frac{r_2}{4r}\right)}\mathrm{d}u\mathrm{d}v.
\end{aligned}$$

Let us consider

$$g \;:\; (u,v) \in \left[-\frac{\theta^*}{2}, \frac{\theta^*}{2}\right]^2 \longmapsto \left(\frac{u+v}{2}, \frac{u-v}{2}\right).$$

We have
$$\left[-\frac{\theta^*}{4},\frac{\theta^*}{4}\right]^2 \subset g\left(\left[-\frac{\theta^*}{2},\frac{\theta^*}{2}\right]^2\right),$$
and
$$|\det \mathrm{Jac}_g| = \frac{1}{2}.$$
With this substitution, and using Fubini's theorem, we get :
$$\begin{aligned}
&\mathbb{P}\left(\Phi_2^0 \in A, S_2 \in (r_1,r_2)\right)\\
&\geq f_{\min}^2 \int_{-\frac{\theta^*}{4}}^{\frac{\theta^*}{4}}\int_{-\frac{\theta^*}{4}}^{\frac{\theta^*}{4}} \mathbf{1}_{x\in\frac{A}{4}}\mathbf{1}_{\cos(x)\cos(y)\in\left(\frac{r_1}{4r},\frac{r_2}{4r}\right)} 2\mathrm{d}x\mathrm{d}y\\
&= 4f_{\min}^2 \int_{-\frac{\theta^*}{4}}^{\frac{\theta^*}{4}}\int_{0}^{\frac{\theta^*}{4}} \mathbf{1}_{\cos(x)\cos(y)\in\left(\frac{r_1}{4r},\frac{r_2}{4r}\right)}\mathrm{d}y\mathbf{1}_{x\in\frac{A}{4}}\mathrm{d}x.
\end{aligned}$$
We now do the substitution $z=\cos(x)\cos(y)$ in the integral with respect to $\mathrm{d}y$:
$$\begin{aligned}
&\mathbb{P}\left(\Phi_2^0 \in A, S_2 \in (r_1,r_2)\right)\\
&\geq 4f_{\min}^2 \int_{-\frac{\theta^*}{4}}^{\frac{\theta^*}{4}}\int_{\cos\left(\frac{\theta^*}{4}\right)\cos(x)}^{\cos(x)} \mathbf{1}_{z\in\left(\frac{r_1}{4r},\frac{r_2}{4r}\right)}\frac{1}{\sqrt{\cos^2(x)-z^2}}\mathrm{d}z\mathbf{1}_{x\in\frac{A}{4}}\mathrm{d}x\\
&\geq 4f_{\min}^2 \int_{-\frac{\theta^*}{4}}^{\frac{\theta^*}{4}}\int_{\cos\left(\frac{\theta^*}{4}\right)\cos(x)}^{\cos(x)} \mathbf{1}_{z\in\left(\frac{r_1}{4r},\frac{r_2}{4r}\right)}\frac{1}{\sin\left(\frac{\theta^*}{4}\right)}\mathrm{d}z\mathbf{1}_{x\in\frac{A}{4}}\mathrm{d}x\\
&\geq \frac{4f_{\min}^2}{\sin\left(\frac{\theta^*}{4}\right)} \int_{-\frac{\theta^*}{4}+\varepsilon}^{\frac{\theta^*}{4}-\varepsilon}\int_{\cos\left(\frac{\theta^*}{4}\right)}^{\cos\left(\frac{\theta^*}{4}-\varepsilon\right)} \mathbf{1}_{z\in\left(\frac{r_1}{4r},\frac{r_2}{4r}\right)}\mathrm{d}z\mathbf{1}_{x\in\frac{A}{4}}\mathrm{d}x\\
&= \frac{f_{\min}^2}{4r\sin\left(\frac{\theta^*}{4}\right)}(r_2-r_1)\left|A\right|,
\end{aligned}$$
where we have used for the last equality the fact that $A\subset[-\theta^*+4\varepsilon,\theta^*-4\varepsilon)$ and $(r_1,r_2)\subset\left(4r\cos\left(\frac{\theta^*}{4}\right),4r\cos\left(\frac{\theta^*}{4}-\varepsilon\right)\right)$.
This ends the proof. □

Let us fix $\eta\in\left(0,r\left(1-2\cos\left(\frac{\theta^*}{2}\right)\right)\right)$ and $\varepsilon\in\left(0,\frac{2\theta^*-\pi}{8}\right)$ (the condition $\theta^*>\frac{2\pi}{3}$ ensures that we can take such η and ε).
Let us define
$$h = 4r\left(1-\cos\left(\frac{\theta^*}{2}\right)\right)-2\eta-2r = 2r\left(1-2\cos\left(\frac{\theta^*}{2}\right)\right)-2\eta>0 \qquad (2.3)$$
and
$$\begin{aligned}
\alpha &= \frac{f_{\min}^2}{2r\sin\left(\frac{\theta^*}{4}\right)}(4\theta^*-2\pi-16\varepsilon)4r\left(\cos\left(\frac{\theta^*}{4}-\varepsilon\right)-\cos\left(\frac{\theta^*}{4}\right)\right)\\
&= \frac{2f_{\min}^2}{\sin\left(\frac{\theta^*}{4}\right)}(4\theta^*-2\pi-16\varepsilon)\left(\cos\left(\frac{\theta^*}{4}-\varepsilon\right)-\cos\left(\frac{\theta^*}{4}\right)\right) \qquad (2.4)
\end{aligned}$$

Theorem 2.3.7. *Let $((X_t, V_t))_{t\geq 0}$ be the stochastic billiard process in $\mathcal{B}_r$ satisfying Assumption $(\mathcal{H}')$ with $\theta^* \in \left(\frac{2\pi}{3}, \pi\right)$.*
There exists a unique invariant probability measure χ on $\mathcal{B}_r \times \mathbb{S}^1$ for the process $((X_t, V_t))_{t\geq 0}$.
Moreover let $\eta \in \left(0, r\left(1 - 2\cos\left(\frac{\theta^*}{2}\right)\right)\right)$ and $\varepsilon \in \left(0, \frac{2\theta^* - \pi}{8}\right)$. For all $t \geq 0$ and all $\lambda < \lambda_M$ we have

$$\|\mathbb{P}(X_t \in \cdot, V_t \in \cdot) - \chi\|_{TV} \leq C_\lambda \mathrm{e}^{-\lambda t},$$

where

$$\lambda_M = \min\left\{\frac{1}{4r}\log\left(\frac{1}{1-\delta h}\right); \frac{1}{4r}\log\left(\frac{-(1-\delta h) + \sqrt{(1-\delta h)^2 + 4\delta h(1-\alpha)}}{2\delta h(1-\alpha)}\right)\right\}. \tag{2.5}$$

and

$$C_\lambda = \frac{\alpha\delta h \mathrm{e}^{10\lambda r}}{1 - \mathrm{e}^{4\lambda r}(1-\delta h) - \mathrm{e}^{8\lambda r}\delta h(1-\alpha)},$$

with δ, h and α respectively given by (2.2), (2.3) and (2.4).

Remark 2.3.8. *The following proof of this theorem is largely inspired by the proof of Theorem 2.2 in [Comets et al., 2009].*

Proof. The existence of the invariant probability measure comes from the compactness of the space $\mathcal{B}_r \times \mathbb{S}^1$. The following proof shows its uniqueness and gives the speed of convergence of the stochastic billiard to equilibrium.
Let $((X_t, V_t))_{t\geq 0}$ and $((\tilde{X}_t, \tilde{V}_t))_{t\geq 0}$ be two versions of the stochastic billiard with $(X_0, V_0) = (x_0, v_0) \in \mathcal{B}_r \times \mathbb{S}^1$ and $(\tilde{X}_0, \tilde{V}_0) = (\tilde{x}_0, \tilde{v}_0) \in \mathcal{B}_r \times \mathbb{S}^1$. We are going to construct these two processes until they become equal.
We recall the definition of T_0 and $\tilde{T}_0$:

$$T_0 = \inf\{t \geq 0 : x_0 + tv_0 \notin K\}, \quad \text{and} \quad \tilde{T}_0 = \inf\{t \geq 0 : \tilde{x}_0 + t\tilde{v}_0 \notin K\}.$$

We are going to couple (X_t, V_t) and $(\tilde{X}_t, \tilde{V}_t)$ in two steps : we first couple the times, so that the two processes hit $\partial\mathcal{B}_r$ at a same time, and then we couple both position and time.

Step 1. Proposition 2.3.5 ensures that S_2 and $\tilde{S}_2$ are both δ-continuous on $\left[4r\cos\left(\frac{\theta^*}{2}\right) + \eta, 4r - \eta\right]$. Therefore, the variables T_2 and $\tilde{T}_2$ are δ-continuous on $\left[T_0 + 4r\cos\left(\frac{\theta^*}{2}\right) + \eta, T_0 + 4r - \eta\right] \cap \left[\tilde{T}_0 + 4r\cos\left(\frac{\theta^*}{2}\right) + \eta, \tilde{T}_0 + 4r - \eta\right]$, with $\left|\left[T_0 + 4r\cos\left(\frac{\theta^*}{2}\right) + \eta, T_0 + 4r - \eta\right] \cap \left[\tilde{T}_0 + 4r\cos\left(\frac{\theta^*}{2}\right) + \eta, \tilde{T}_0 + 4r - \eta\right]\right| \geq h$ since $|T_0 - \tilde{T}_0| \leq 2r$. Note that the condition $\theta^* > \frac{2\pi}{3}$ has been introduced to ensure that this intersection is non-empty.
Thus, there exists a coupling of T_2 and $\tilde{T}_2$ such that

$$\mathbb{P}(E_1) \geq \delta h,$$

where

$$E_1 = \left\{T_2 = \tilde{T}_2\right\}.$$

On the event E_1 we define $T_c^1 = T_2$.
On the event E_1^c, we can suppose, by symmetry that $T_2 \leq \tilde{T}_2$. In order to try again to couple the hitting times, we need to begin at times whose difference is smaller than $2r$. Let us thus define

$$m_1 = \min\left\{n > 2 : T_n > \tilde{T}_2\right\} \quad \text{and} \quad \tilde{m}_1 = 2.$$

We then have, by construction of m_1 and $\tilde{m}_1$, $\left|T_{m_1} - \tilde{T}_{\tilde{m}_1}\right| \leq 2r$. Therefore, as previously, there exists a coupling of T_{m_1+2} and $\tilde{T}_{\tilde{m}_1+2}$ such that

$$\mathbb{P}\left(E_2|E_1^c\right) \geq \delta h,$$

where

$$E_2 = \left\{T_{m_1+2} = \tilde{T}_{\tilde{m}_1+2}\right\}.$$

On the event $E_1^c \cap E_2$ we define $T_c^1 = T_{m_1+2}$.
We then repeat the same procedure. We thus construct two sequences of stopping times $(m_k)_{k\geq 1}$, $(\tilde{m}_k)_{k\geq 1}$ and a sequence of events $(E_k)_{k\geq 1}$, with

$$E_k = \left\{T_{m_{k-1}+2} = \tilde{T}_{\tilde{m}_{k-1}+2}\right\},$$

satisfying

$$\mathbb{P}\left(E_k|E_1^c \cap \dots \cap E_{k-1}^c\right) \geq \delta h.$$

On the event $E_1^c \cap \dots \cap E_{k-1}^c \cap E_k$ we define $T_c^1 = T_{m_k+2}$. By construction, T_c^1 is the coupling time of the hitting times of the boundary.
Since the inter-jump times are directly linked to the speeds of the process, and thus to its positions, we can construct both stochastic billiards (X_t, V_t) and $(\tilde{X}_t, \tilde{V}_t)$ until time T_c^1, and along with T_c^1.
We observe that by this construction of T_c^1, we have

$$T_c^1 \leq_{st} T_0 + \sum_{l=1}^{G^1} S^{1,l} \tag{2.6}$$

with $G^1 \sim \mathcal{G}(\delta h)$ and $S^{1,l}$, $l \geq 1$, independent random times with distribution f_{S_2}, and independent of G^1.

Step 2. Let us now suppose that T_c^1, $((X_t, V_t))_{0\leq t\leq T_1^c}$ and $((\tilde{X}_t, \tilde{V}_t))_{0\leq t\leq T_1^c}$ are constructed as described above, .
We define $y = X_{T_c^1}$ and $\tilde{y} = \tilde{X}_{T_c^1}$, which are by construction of T_c^1 on $\partial\mathcal{B}_r$. We also define $N_c^1 = \min\{n > 0 : X_{T_n} = y\}$, which is deterministic conditionally to T_c^1.
Proposition 2.3.6 ensures that the couples $\left(X_{T_{N_c^1+2}}, T_{N_c^1+2} - T_{N_c^1}\right)$ and $\left(\tilde{X}_{T_{N_c^1+2}}, \tilde{T}_{N_c^1+2} - \tilde{T}_{N_c^1}\right)$ are both $\frac{f_{\min}^2}{2r\sin\left(\frac{\theta^*}{4}\right)}$-continuous on the set $((\varphi_y - \theta^* + 4\varepsilon, \varphi_y + \theta^* - 4\varepsilon) \cap (\varphi_{\tilde{y}} - \theta^* + 4\varepsilon, \varphi_{\tilde{y}} + \theta^* - 4\varepsilon))$
$\times \left(4r\cos\left(\frac{\theta^*}{4}\right), 4r\cos\left(\frac{\theta^*}{4} - \varepsilon\right)\right)$, with

$|(\varphi_y - \theta^* + 4\varepsilon, \varphi_y + \theta^* - 4\varepsilon) \cap (\varphi_{\tilde{y}} - \theta^* + 4\varepsilon, \varphi_{\tilde{y}} + \theta^* - 4\varepsilon)| \geq 4\theta^* - 2\pi - 16\varepsilon$ (let us mention that these intervals are seen in $[0, 2\pi]/(2\pi\mathbb{Z})$ since they are intervals of angles). Note that the condition $\theta^* > \frac{2\pi}{3}$ implies in particular that the previous intersection in non-empty.
Therefore we can construct a coupling such that

$$\mathbb{P}\left(F|E_1^c \cap \cdots \cap E_{N_c^1-1}^c \cap E_{N_c^1}, T_c^1\right) \geq \alpha,$$

where

$$F = \left\{X_{T_{N_c^1+2}} = \tilde{X}_{T_{N_c^1+2}} \text{ and } T_{N_c^1+2} = \tilde{T}_{N_c^1+2}\right\}.$$

On the event F we define $T_c = T_{N_c^1+2}$, and we construct $((X_t, V_t))_{T_c^1 \leq t \leq T_c}$ and $((\tilde{X}_t, \tilde{V}_t))_{T_c^1 \leq t \leq T_c}$ along with the coupling of $(X_{T_{N_c^1+2}}, T_{N_c^1+2})$ and $(\tilde{X}_{T_{N_c^1+2}}, \tilde{T}_{N_c^1+2})$. If F does not occur, we can not directly try to couple both position and time since the two processes have not necessarily hit $\partial\mathcal{B}_r$ at the same time. We thus have to couple first the hitting times, as we have done in step 1.

Let us suppose that on $\left(E_1^c \cap \cdots \cap E_{N_c^1-1}^c \cap E_{N_c^1}\right) \cap F^c$, we have $T_{N_c^1+2} \leq \tilde{T}_{N_c^1+2}$ (the other case can be treated in the same way thanks to the symmetry of the problem). Let us define

$$\ell = \min\left\{n > N_c^1 + 2 : T_n > \tilde{T}_{N_c^1+2}\right\} \text{ and } \tilde{\ell} = N_c^1 + 2$$

We clearly have $\left|T_\ell - \tilde{T}_{\tilde{\ell}}\right| \leq 2r$. Therefore, we can start again : we try to couple the times at which the two processes hit the boundary, and then to couple the positions and times together.

Finally, the probability that we succeed to couple the positions and times in "one step" (Step 1 and Step 2) is :

$$\begin{aligned}
&\mathbb{P}\left(\left(\bigcup_{k\geq 1}\left(E_1^c \cap \cdots \cap E_{k-1}^c \cap E_k\right)\right) \cap F\right) \\
&= \mathbb{P}\left(F \,\middle|\, \bigcup_{k\geq 1}\left(E_1^c \cap \cdots \cap E_{k-1}^c \cap E_k\right)\right) \mathbb{P}\left(\bigcup_{k\geq 1}\left(E_1^c \cap \cdots \cap E_{k-1}^c \cap E_k\right)\right) \\
&= \mathbb{P}\left(F \,\middle|\, \bigcup_{k\geq 1}\left(E_1^c \cap \cdots \cap E_{k-1}^c \cap E_k\right)\right) \\
&\geq \alpha.
\end{aligned}$$

Therefore, the coupling time $\hat{T}$ of the couples position-time satisfies :

$$\hat{T} \leq_{st} T_0 + \sum_{k=1}^{G}\left(\left(\sum_{l=1}^{G^k} S^{k,l}\right) + S^k\right)$$

where $G \sim \mathcal{G}(\alpha)$, $G^1, G^2, \cdots \sim \mathcal{G}(\delta h)$ are independent geometric variables, and $\left(S^{k,l}\right)_{k,l\geq 1}$, $\left(S^k\right)_{k\geq 1}$ are independent random variables, independent from the

geometric variables, with distribution f_{S_2}.
Let $\lambda \in (0, \lambda_M)$, with λ_M defined in equation (2.5). Since all the random variables $S^{k,l}$ and S^k, $k, l \geq 1$, are almost surely smaller than two times the diameter of the ball $\mathcal{B}_r$, and since T_0 is almost surely smaller than this diameter, we have :

$$
\begin{aligned}
\mathbb{P}\left(\hat{T} > t\right) &\leq \mathrm{e}^{-\lambda t}\mathbb{E}\left[\mathrm{e}^{\lambda \hat{T}}\right] \\
&\leq \mathrm{e}^{\lambda(T_0 - t)}\mathbb{E}\left[\exp\left(\lambda \sum_{k=1}^{G}\left(\left(\sum_{l=1}^{G^k} S^{k,l}\right) + S^k\right)\right)\right] \\
&\leq \mathrm{e}^{\lambda(2r - t)}\mathbb{E}\left[\prod_{k=1}^{G}\left(\left(\prod_{l=1}^{G^k}\exp(\lambda 4r)\right)\exp(\lambda 4r)\right)\right] \\
&= \mathrm{e}^{\lambda(2r - t)}\mathbb{E}\left[\prod_{k=1}^{G}\mathbb{E}\left[\mathrm{e}^{4\lambda r(G^k + 1)}\right]\right].
\end{aligned}
$$

Now, using the expression of generating function of a geometric random variable we get :

$$
\begin{aligned}
\mathbb{P}\left(\hat{T} > t\right) &\leq \mathrm{e}^{\lambda(2r - t)}\mathbb{E}\left[\prod_{k=1}^{G}\left(\sum_{l=1}^{\infty}\mathrm{e}^{4\lambda r(l+1)}\delta h(1 - \delta h)^{l-1}\right)\right] \\
&= \mathrm{e}^{\lambda(2r - t)}\mathbb{E}\left[\left(\frac{\mathrm{e}^{8\lambda r}\delta h}{1 - \mathrm{e}^{4\lambda r}(1 - \delta h)}\right)^{G}\right] \\
&= \mathrm{e}^{\lambda(2r - t)}\frac{\alpha \mathrm{e}^{8\lambda r}\delta h}{1 - \mathrm{e}^{4\lambda r}(1 - \delta h)}\frac{1}{1 - \frac{\mathrm{e}^{8\lambda r}\delta h(1-\alpha)}{1 - \mathrm{e}^{4\lambda r}(1 - \delta h)}} \\
&= \mathrm{e}^{-\lambda t}\frac{\alpha \mathrm{e}^{10\lambda r}\delta h}{1 - \mathrm{e}^{4\lambda r}(1 - \delta h) - \mathrm{e}^{8\lambda r}\delta h(1 - \alpha)}.
\end{aligned}
$$

This calculations are valid for $\lambda > 0$ such that the generating functions are well defined, that is for $\lambda > 0$ satisfying

$$
\mathrm{e}^{4\lambda r}(1 - \delta h) < 1 \quad \text{and} \quad \frac{\mathrm{e}^{8\lambda r}\delta h(1 - \alpha)}{1 - \mathrm{e}^{4\lambda r}(1 - \delta h)} < 1.
$$

The first condition is equivalent to $\lambda < \frac{1}{4r}\log\left(\frac{1}{1-\delta h}\right)$.
The second condition is equivalent to $\delta h(1-\alpha)s^2 + (1-\delta h)s - 1 < 0$ with $s = \mathrm{e}^{4\lambda r}$. It gives $s_1 < s < s_2$ with $s_1 = \frac{-(1-\delta h)-\sqrt{\Delta}}{2\delta h(1-\alpha)} < 0$ and $s_2 = \frac{-(1-\delta h)+\sqrt{\Delta}}{2\delta h(1-\alpha)} > 1$ where $\Delta = (1-\delta h)^2 + 4\delta h(1-\alpha) > 0$. And finally we get $\lambda < \frac{1}{4r}\log(s_2)$.
Therefore, the estimation for $\mathbb{P}\left(\hat{T} > t\right)$ is indeed valid for all $\lambda \in (0, \lambda_M)$. The conclusion of the theorem follows immediately.

□

Remark 2.3.9. *If $\theta^* \in \left(0, \frac{2\pi}{3}\right]$, Step 1 of the proof of Theorem 2.3.7 fails : the intervals on which the random variables S_2 and $\tilde{S}_2$ are continuous can have an empty intersection. Similarly, in Step 2, the intersection of the intervals on which the couples $\left(X_{T_{N_c^1+2}}, T_{N_c^1+2} - T_{N_c^1}\right)$ and $\left(\tilde{X}_{T_{N_c^1+2}}, \tilde{T}_{N_c^1+2} - \tilde{T}_{N_c^1}\right)$ are continuous can be empty depending on the value of θ^*.*
However, instead of trying to couple the times or both positions and times in two jumps, we just need more jumps to do that. Therefore, the method and the results are similar in the case $\theta^* \leq \frac{2\pi}{3}$, the only difference is that the computations and notations will be much more awful.

2.4 Stochastic billiard in a convex set with bounded curvature

We make the following assumption on the set K in which the stochastic billiard evolves :

Assumption $(\mathcal{K})$ **:**

K is a compact convex set with curvature bounded from above by $C < \infty$ and bounded from below by $c > 0$.

This means that for each $x \in \partial K$, there is a ball B_1 with radius $\frac{1}{C}$ included in K and a ball B_2 containing K, so that the tangent planes of K, B_1 and B_2 at x coincide (see Figure 2.4). In fact, for $x \in \partial K$, the ball B_1 is the ball with radius $\frac{1}{C}$ and with center the unique point at distance $\frac{1}{C}$ from x in the direction of n_x. And B_2 is the one with the center at distance $\frac{1}{c}$ from x in the direction of n_x.

In this section, we consider the stochastic billiard in such a convex K.
Let us observe that the case of the disc is a particular case. Moreover, Assumption $(\mathcal{K})$ excludes in particular the case of the polygons : because of the upper bound C on the curvature, the boundary of K can not have "corners", and because of the lower bound c, the boundary can not have straight lines.

In the following, D will denote the diameter of K, that is

$$D = \max\{\|x - y\| \colon x, y \in \partial K\}.$$

2.4.1 The embedded Markov chain

Notation : We define $l_{x,y} = \frac{y-x}{\|x-y\|} = -l_{y,x}$ and we denote by $\varphi_{x,y}$ the angle between $l_{x,y}$ and the normal n_x to ∂K at the point x (see Figure 2.5).

The following property, proved by Comets and al. in [Comets *et al.*, 2009], gives the dynamics of the Markov chain $(X_{T_n})_{n\geq 0}$ defined in Section 2.2.2

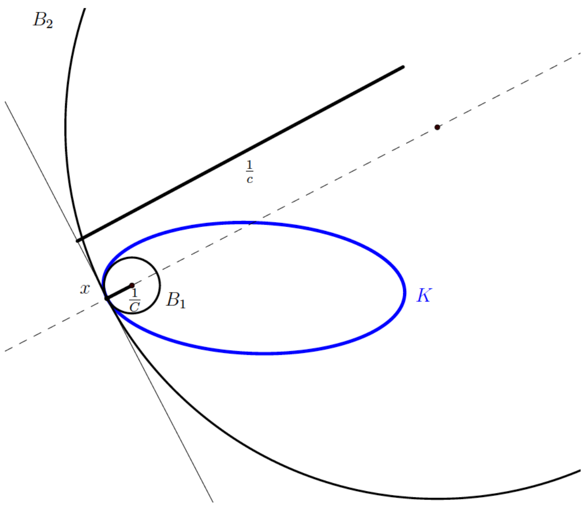


FIGURE 2.4 – Illustration of Assumption $(\mathcal{K})$

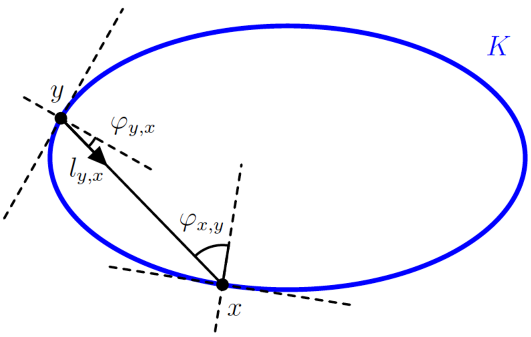


FIGURE 2.5 – Definition of the quantities $\varphi_{x,y}$ and $l_{y,x}$ for $x, y \in \partial K$

Proposition 2.4.1. *The transition kernel of the chain $(X_{T_n})_{n\geq 0}$ is given by :*

$$\mathbb{P}\left(X_{T_{n+1}} \in A \,|X_{T_n} = x\right) = \int_A Q(x,y)\mathrm{d}y$$

where

$$Q(x,y) = \frac{\rho(U_x^{-1}l_{x,y})\cos(\varphi_{y,x})}{\|x-y\|}.$$

This proposition is one of the main ingredients to obtain the exponentially-fast convergence of the stochastic billiard Markov chain towards its invariant probability measure.

Theorem 2.4.2. *Let $K \subset \mathbb{R}^2$ satisfying Assumption $(\mathcal{K})$ with diameter D. Let $(X_{T_n})_{n\geq 0}$ be the stochastic billiard Markov chain on ∂K verifying Assumption $(\mathcal{H})$.*
There exists a unique invariant measure ν on ∂K for $(X_{T_n})_{n\geq 0}$.
Moreover, recalling that $\theta^* = |\mathcal{J}|$ in Assumption $(\mathcal{H})$, we have :

1. *if $\theta^* > \frac{C|\partial K|}{8}$, for all $n \geq 0$,*

$$\|\mathbb{P}\left(X_{T_n} \in \cdot\right) - \nu\|_{TV} \leq \left(1 - q_{\min}\left(\frac{8\theta^*}{C} - |\partial K|\right)\right)^{n-1};$$

2. *if $\theta^* \leq \frac{C|\partial K|}{8}$, for all $n \geq 0$ and all $\varepsilon \in \left(0, \frac{2\theta^*}{C}\right)$,*

$$\|\mathbb{P}\left(X_{T_n} \in \cdot\right) - \nu\|_{TV} \leq (1-\alpha)^{\frac{n}{n_0}-1}$$

where

$$n_0 = \left\lfloor \frac{\frac{|\partial K|}{2} - 2\varepsilon}{\frac{4\theta^*}{C} - 2\varepsilon} \right\rfloor + 1$$
$$\text{and} \quad \alpha = (\frac{4\theta^*}{C})^{n_0 - 1} {q_{\min}}^{n_0}\left(4\left(\frac{2n_0\theta^*}{C} - (n_0-1)\varepsilon\right) - |\partial K|\right)$$

with

$$q_{\min} = \frac{c\rho_{\min}\cos\left(\frac{\theta^*}{2}\right)}{CD}.$$

Proof. Once more, the existence of the invariant measure is immediate since the state space ∂K of the Markov chain is compact. The following shows its uniqueness and gives the speed of convergence of $(X_{T_n})_{n\geq 0}$ towards ν.
Let $(X_{T_n})_{n\geq 0}$ and $(\tilde{X}_{T_n})_{n\geq 0}$ be two versions of the Markov chain with initial conditions x_0 and $\tilde{x}_0$ on ∂K. In order to have a strictly positive probability to couple X_{T_n} and $\tilde{X}_{T_n}$ at time n, it is sufficient that their density functions are bounded from below on an interval of length strictly bigger than $\frac{|\partial K|}{2}$. Let us therefore study the length of set on which the density function $f_{X_{T_n}}$ of X_{T_n} is bounded from below by a strictly positive constant.

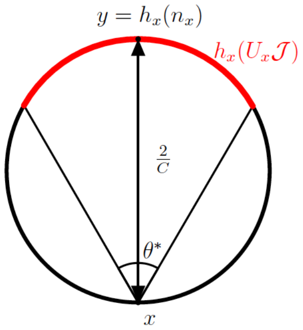


FIGURE 2.6 – Worst scenario for the length of $h_x(U_x\mathcal{J})$

Let $x \in \partial K$. For $v \in \mathbb{S}_x$, we denote by $h_x(v)$ the unique point on ∂K seen from x in the direction of v. We firstly get a lower bound on $|h_x(U_x\mathcal{J})|$, the length of the subset of ∂K seen from x with a strictly positive density.
It is easy to observe, with a drawing for instance, the following facts :
— $|h_x(U_x\mathcal{J})|$ increases when $\|x - h_x(n_x)\|$ increases,
— $|h_x(U_x\mathcal{J})|$ decreases when the curvature at $h_x(n_x)$ increases,
— $|h_x(U_x\mathcal{J})|$ decreases when $|\varphi_{h_x(n_x),x}|$ increases.
Therefore, $|h_x(U_x\mathcal{J})|$ is minimal when $\|x-h_x(n_x)\|$ is minimal, when the curvature at $h_x(n_x)$ is maximal, and then equal to C, and finally when $\varphi_{h_x(n_x),x} = 0$. Moreover, the minimal value of $\|x - h_x(n_x)\|$ is $\frac{2}{C}$ since C is the upper bound for the curvature of ∂K. The configuration that makes the quantity $|h_x(U_x\mathcal{J})|$ minimal is thus the case where x and $h_x(n_x)$ define a diameter on a circle of diameter $\frac{2}{C}$ (see Figure 2.6). We immediately deduce a lower bound for $|h_x(U_x\mathcal{J})|$:

$$|h_x(U_x\mathcal{J})| \geq 2\theta^* \times \frac{2}{C} = \frac{4\theta^*}{C}.$$

This means that the density function $f_{X_{T_1}}$ of X_{T_1} is strictly positive on a subset of ∂K of length at least $\frac{4\theta^*}{C}$.
Let now $\varepsilon \in \left(0, \frac{2\theta^*}{C}\right)$. As it has been done in Section 2.3 for the disc, we can deduce that for all $n \geq 2$, the density function $f_{X_{T_n}}$ is strictly positive on a set of length at least $2n\theta^*\frac{2}{C} - 2(n-1)\varepsilon = \frac{4n\theta^*}{C} - 2(n-1)\varepsilon$.
Let us define, for $x \in \partial K$ and $n \geq 1$, $\mathcal{J}_x^n$ the set of points of ∂K that can be reached from x in n bounces by picking for each bounce a velocity in $\mathcal{J}$.
We now separate the cases where we can couple in one jump, and where we need more jumps.
— Case 1 : $\theta^* > \frac{C|\partial K|}{8}$. In that case we have, for all $x \in \partial K$, $|\mathcal{J}_x^1| \geq \frac{4\theta^*}{C} > \frac{|\partial K|}{2}$, and we can thus construct a coupling $(X_{T_1}, \tilde{X}_{T_1})$ such that :

$$\begin{aligned}\mathbb{P}\left(X_{T_1} = \tilde{X}_{T_1}\right) &\geq q_{\min}\left|\mathcal{J}_{x_0}^1 \cap \tilde{\mathcal{J}}_{\tilde{x}_0}^1\right| \\ &\geq q_{\min} \times 2\left(\frac{4\theta^*}{C} - \frac{|\partial K|}{2}\right)\end{aligned}$$

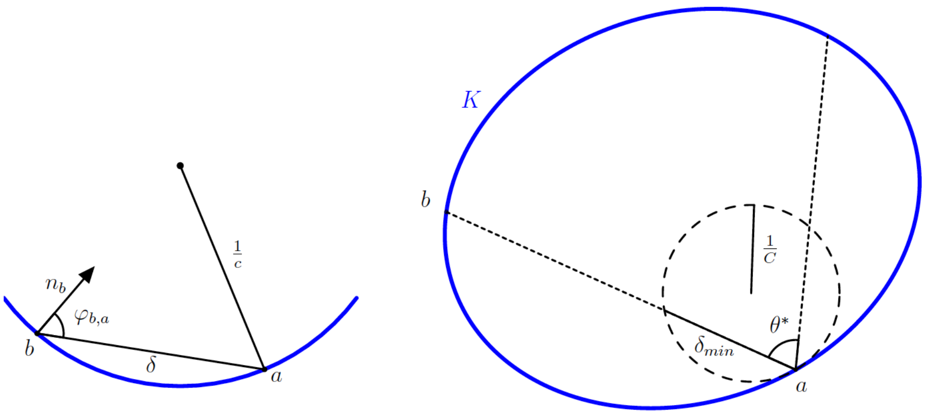


FIGURE 2.7 – Illustration for the calculation of a lower bound for $\cos(\varphi_{b,a})$ with $a \in \partial K$ and $b \in h_a(U_a\mathcal{J})$

$$= q_{\min}\left(\frac{8\theta^*}{C} - |\partial K|\right),$$

where $q_{\min}$ is a uniform lower bound of $Q(a,b)$ with $a \in \partial K$ and $b \in h_a(U_a\mathcal{J})$, i.e.

$$q_{\min} \leq \min_{a\in\partial K, b\in h_a(U_a\mathcal{J})} Q(a,b).$$

Let us thus give an explicit expression for $q_{\min}$. Let $a \in \partial K$ and $b \in h_a(U_a\mathcal{J})$. We have

$$Q(a,b) \geq \frac{\rho_{\min}\cos(\varphi_{b,a})}{D}.$$

We could have $\cos(\varphi_{b,a}) = 0$ if a and b were on a straight part of ∂K, which is not possible since the curvature of K is bounded from below by c. Thus, the quantity $\cos(\varphi_{b,a})$ is minimal when a and b are on a part of a disc with curvature c. In that case, $\cos(\varphi_{b,a}) = \frac{\delta c}{2}$, where δ is the distance between a and b (see the first picture of Figure 2.7). Since $b \in h_a(U_a\mathcal{J})$, we have $\delta \geq \delta_{\min} := \frac{2\cos\left(\frac{\theta^*}{2}\right)}{C}$ (see the second picture of Figure 2.7). Finally we get

$$Q(a,b) \geq \frac{c\rho_{\min}\cos\left(\frac{\theta^*}{2}\right)}{CD} =: q_{\min}.$$

— Case 2 : $\theta^* \leq \frac{C|\partial K|}{8}$. In that case, we need more than one jump to couple the two Markov chains. Therefore, defining

$$n_0 = \min\left\{n \geq 2 : \frac{4n\theta^*}{C} - 2(n-1)\varepsilon > \frac{\partial K}{2}\right\} = \left\lfloor \frac{\frac{|\partial K|}{2} - 2\varepsilon}{\frac{4\theta^*}{C} - 2\varepsilon} \right\rfloor + 1,$$

we get that the intersection $\mathcal{J}_{x_0}^{n_0} \cap \tilde{\mathcal{J}}_{\tilde{x}_0}^{n_0}$ is non-empty, and then we can construct $X_{T_{n_0}}$ and $\tilde{X}_{T_{n_0}}$ such that the probability $\mathbb{P}\left(X_{T_{n_0}} = \tilde{X}_{T_{n_0}}\right)$ is strictly positive. It remains to estimate a lower bound of this probability. First, we have

$$\begin{aligned}\left|\mathcal{J}_{x_0}^{n_0} \cap \tilde{\mathcal{J}}_{\tilde{x}_0}^{n_0}\right| &\geq 2\left(\frac{4n_0\theta^*}{C} - 2(n_0-1)\varepsilon - \frac{|\partial K|}{2}\right) \\ &= 4\left(\frac{2n_0\theta^*}{C} - (n_0-1)\varepsilon\right) - |\partial K|.\end{aligned}$$

Moreover, let $x \in \{x_0, \tilde{x}_0\}$ and $y \in \mathcal{J}_{x_0}^{n_0} \cap \tilde{\mathcal{J}}_{\tilde{x}_0}^{n_0}$. We have :

$$\begin{aligned}&Q^{n_0}(x,y) \\ &\quad \geq \int_{h_x(U_x\mathcal{J})}\int_{h_{z_1}(U_{z_1}\mathcal{J})}\cdots\int_{h_{z_{n-2}}(U_{z_{n-2}}\mathcal{J})} \\ &\qquad\qquad Q(x,z_1)Q(z_1,z_2)\cdots Q(z_{n_0-1},y)\mathrm{d}z_1\mathrm{d}z_2\cdots\mathrm{d}z_{n_0-1} \\ &\quad \geq (\frac{4\theta^*}{C})^{n_0-1}{q_{\min}}^{n_0}.\end{aligned}$$

We thus deduce :

$$\begin{aligned}\mathbb{P}\left(X_{T_{n_0}} = \tilde{X}_{T_{n_0}}\right) &\geq (\frac{4\theta^*}{C})^{n_0-1}{q_{\min}}^{n_0}\left|\mathcal{J}_{x_0}^{n_0} \cap \tilde{\mathcal{J}}_{\tilde{x}_0}^{n_0}\right| \\ &\geq (\frac{4\theta^*}{C})^{n_0-1}{q_{\min}}^{n_0}\left(4\left(\frac{2n_0\theta^*}{C} - (n_0-1)\varepsilon\right) - |\partial K|\right).\end{aligned}$$

We can now conclude, including the two cases : let us define

$$m_0 = \mathbf{1}_{\theta^* > \frac{C|\partial K|}{8}} + \left(\left\lfloor \frac{\frac{|\partial K|}{2} - 2\varepsilon}{\frac{4\theta^*}{C} - 2\varepsilon} \right\rfloor + 1\right)\mathbf{1}_{\theta^* \leq \frac{C|\partial K|}{8}}$$

and

$$\begin{aligned}\alpha = q_{\min}&\left(\frac{8\theta^*}{C} - |\partial K|\right)\mathbf{1}_{\theta^* > \frac{C|\partial K|}{8}} \\ &+ (\frac{4\theta^*}{C})^{m_0-1}{q_{\min}}^{m_0}\left(4\left(\frac{2m_0\theta^*}{C} - (m_0-1)\varepsilon\right) - |\partial K|\right)\mathbf{1}_{\theta^* \leq \frac{C|\partial K|}{8}}.\end{aligned}$$

We have proved that we can construct a coupling $\left(X_{T_{m_0}}, \tilde{X}_{T_{m_0}}\right)$ such that $\mathbb{P}\left(X_{T_{m_0}} = \tilde{X}_{T_{m_0}}\right) \geq \alpha$, and then we get

$$\|\mathbb{P}\left(X_{T_n} \in \cdot\right) - \nu\|_{TV} \leq (1-\alpha)^{\frac{n}{m_0}-1}.$$

□

2.4.2 The continuous-time process

In this section, we will work in the case $|\mathcal{J}|= \theta^* = \pi$.

Notations : We still use the following notation, already introduced in the case of the disc :

$$S_0 = 0 \quad \text{and} \quad \text{for } n \geq 1, \quad S_n = T_n - T_0.$$

Moreover, for $x \in \partial K$, we write S_n^x an $X_{S_n}^x$ respectively for the n^{th} hitting time of ∂K and the position of the Markov chain after n steps, and that started at position $x \in \partial K$.
Thereby, we have the following equalities : $\mathcal{L}(X_{S_n}^x) = \mathcal{L}(X_{T_n}|X_{T_0} = x)$ and $\mathcal{L}(S_n^x) = \mathcal{L}(T_n - T_0|X_{T_0} = x)$.

Proposition 2.4.3. *Let $K \subset \mathbb{R}^2$ satisfying Assumption $(\mathcal{K})$. Let $((X_t, V_t))_{t\geq 0}$ the stochastic billiard process evolving in K and verifying Assumption $(\mathcal{H})$ with $|\mathcal{J}|= \pi$.*
For all $x \in \partial K$, the random time S_1^x, which is the first hitting time of ∂K when starting at point x, is $c\rho_{\min}$-continuous on $\left[0, \frac{2}{C}\right]$.

Proof. Let $x \in \partial K$. Let us recall that the curvature of K is bounded from above by C, which means that for each $x \in \partial K$, there is a ball B_1 with radius $\frac{1}{C}$ included in K so that the tangent planes of K and B_1 at x coincide. Therefore, starting from x, the maximal time to go on another point of ∂K is bigger than $\frac{2}{C}$ (the diameter of the ball B_1).
That is why we are going to prove the continuity of S_1^x on the interval $\left[0, \frac{2}{C}\right]$.
Let thus $0 \leq r \leq R \leq \frac{2}{C}$.
Let Θ be a random variable living in $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ such that the velocity vector $(\cos(\Theta), \sin(\Theta))$ follows the law γ.
The time S_1^x being completely determined by the velocity V_{T_0} and thus by its angle with respect to n_x, it is clear that there exist $-\frac{\pi}{2} \leq \theta_1 \leq \theta_2 \leq \theta_3 \leq \theta_4 \leq \frac{\pi}{2}$ such that we have :

$$\mathbb{P}\left(S_1^x \in [r, R]\right) = \mathbb{P}\left(\Theta \in [\theta_1, \theta_2] \cup [\theta_3, \theta_4]\right).$$

Then, thanks to Assumption $(\mathcal{H})$ on the law γ, and since we assume here that $|\mathcal{J}|= \pi$, the density function of Θ is bounded from below by $\rho_{\min}$ on $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$.
It gives :

$$\mathbb{P}\left(S_1^x \in [r, R]\right) \geq \rho_{\min}\left(\theta_2 - \theta_1 + \theta_4 - \theta_3\right).$$

Moreover, since the curvature is bounded from below by c, there exists a ball B_2 with radius $\frac{1}{c}$ containing K so that the tangent planes of K and B_2 at x coincide. And it is easy to see that the differences $\theta_2 - \theta_1$ and $\theta_4 - \theta_2$ are larger than the difference $\alpha_2 - \alpha_1$ where α_1 and α_2 are the angles corresponding to the distances r and R starting from x and to arrive on the ball B_2.
The time of hitting the boundary of B_2 is equal to $d \in \left[0, \frac{2}{C}\right]$ if the angle between n_x and the velocity is equal to $\arccos\left(\frac{cd}{2}\right)$. We thus deduce :

$$\mathbb{P}\left(S_1^x \in [r, R]\right) \geq 2\rho_{\min}\left(\arccos\left(\frac{cr}{2}\right) - \arccos\left(\frac{cR}{2}\right)\right)$$

$$\geq 2\rho_{\min}\left|\frac{cr}{2}-\frac{cR}{2}\right|$$
$$= \rho_{\min} c\,(R-r)\,,$$

where we have used the mean value theorem for the second inequality.

□

Let us introduce some constants that will appear in the following results.
Let $\beta > 0$ and $\delta > 0$ such that $\frac{|\partial K|}{3} - \max\{2\delta;\beta+\delta\} > 0$.
Let $\varepsilon \in \left(0, \min\{\beta; \frac{2}{C}\}\right)$ such that $h > 0$ where

$$h = \frac{\delta}{D}\left(\frac{\beta c}{2}\right)^2 - \varepsilon M, \tag{2.7}$$

with

$$M = 2\left(\frac{1}{\frac{1}{C}-\varepsilon} + \frac{1}{\beta-\varepsilon} + C\right). \tag{2.8}$$

Let us remark that M is non decreasing with ε, so that it is possible to take ε small enough to have $h > 0$.

Proposition 2.4.4. *Let $K \subset \mathbb{R}^2$ satisfying Assumption $(\mathcal{K})$ with diameter D. Let $((X_t, V_t))_{t\geq 0}$ the stochastic billiard process evolving in K and verifying Assumption $(\mathcal{H})$ with $|\mathcal{J}| = \pi$.*
Let $x, \tilde{x} \in \partial K$ with $x \neq \tilde{x}$.
There exist $R_1 > 0$, $R_2 > 0$ and $J^* \subset \partial K$, with $|J^*| < h\varepsilon$, such that the couples $(X^x_{S_2}, S^x_2)$ and $(\tilde{X}^{\tilde{x}}_{\tilde{S}_2}, \tilde{S}^{\tilde{x}}_2)$ are both η-continuous on $J^* \times (R_1, R_2)$, with

$$\eta = \frac{1}{2}\left(\frac{c\rho_{\min}}{2D}\right)^2\left(\frac{1}{C}-\varepsilon\right)(\beta-\varepsilon)\,.$$

Moreover we have $R_2 - R_1 \geq 2\,(h\varepsilon - |J^*|)$.

Remark 2.4.5. *The following proof is largely inspired by the proof of Lemma 5.1 in [Comets et al., 2009].*

Proof. Let $x, \tilde{x} \in \partial K$, $x \neq \tilde{x}$. Let us denote by $\Delta_{x\tilde{x}}$ the bisector of the segment defined by the two points x and $\tilde{x}$. The intersection $\Delta_{x\tilde{x}} \cap \partial K$ contains two points, let us thus define $\bar{y}$ the one which achieves the larger distance towards x and $\tilde{x}$ (we consider this point of intersection since we need in the sequel to have a lower bound on $\|x-\bar{y}\|$ and $\|\tilde{x}-\bar{y}\|$).
Let $t \in I \mapsto g(t)$ be a parametrization of ∂K with $g(0) = \bar{y}$, such that $\|g'(t)\| = 1$ for all $t \in I$. Consequently, the length of an arc satisfies length$(g_{|[s,t]}) = ||g(t) - g(s)|| = |t-s|$. We can thus write $I = [0, |\partial K|]$, and $g(0) = g(|\partial K|)$. Note that the parametrization g is C^2 thanks to Assumption $(\mathcal{K})$.
In the sequel, for $z \in \partial K$, we write s_z (or t_z) for the unique $s \in I$ such that $g(s) = z$. And for $A \subset \partial K$, we define $I_A = \{t \in I : g(t) \in A\}$.
Let us define, for $s, t \in I$ and $w \in \{x, \tilde{x}\}$:

$$\varphi_w(s,t) = \|w - g(s)\| + \|g(s) - g(t)\|.$$

Lemma 2.4.6. *There exists an interval $I^*_{\beta,\delta} \subset I$, satisfying $|I^*_{\beta,\delta}| < h\varepsilon$, such that for $w \in \{x, \tilde{x}\}$:*

$$|\partial_s \varphi_w(s,t)| \geq h, \quad \textit{for } s \in B^\varepsilon_{\bar{y}} \textit{ and } t \in I^*_{\beta,\delta},$$

where $B^\varepsilon_{\bar{y}} = \{s \in I; |s - s_{\bar{y}}| \leq \varepsilon\}$.

We admit this lemma for the moment and prove it after the end of the current proof.
Let us suppose for instance that $\partial_s \varphi_w(s,t)$ is positive for $s \in B^\varepsilon_{\bar{y}}$ and $t \in I^*_{\beta,\delta}$, for $w = x$ and $w = \tilde{x}$. If one or both of $\partial_s \varphi_x(s,t)$ and $\partial_s \varphi_{\tilde{x}}(s,t)$ are negative, we just need to consider $|\varphi_x|$ or $|\varphi_{\tilde{x}}|$, and everything works similarly.
We thus have, by the lemma :

$$\partial_s \varphi_w(s,t) \geq h, \quad \text{for } s \in B^\varepsilon_{\bar{y}} \text{ and } t \in I^*_{\beta,\delta}.$$

Let us now define :

$$r_1 = \sup_{t \in I^*_{\beta,\delta}} \inf_{s \in B^\varepsilon_{\bar{y}}} \varphi_x(s,t) \quad \text{and} \quad r_2 = \inf_{t \in I^*_{\beta,\delta}} \sup_{s \in B^\varepsilon_{\bar{y}}} \varphi_x(s,t)$$

and

$$\tilde{r_1} = \sup_{t \in I^*_{\beta,\delta}} \inf_{s \in B^\varepsilon_{\bar{y}}} \varphi_{\tilde{x}}(s,t) \quad \text{and} \quad \tilde{r_2} = \inf_{t \in I^*_{\beta,\delta}} \sup_{s \in B^\varepsilon_{\bar{y}}} \varphi_{\tilde{x}}(s,t).$$

Since $s \mapsto \varphi_x(s,t)$ and $s \mapsto \varphi_{\tilde{x}}(s,t)$ are strictly increasing on $B^\varepsilon_{\bar{y}}$ for all $t \in I^*_{\beta,\delta}$, we deduce that, considering $B^\varepsilon_{\bar{y}}$ as the interval (s_1, s_2),

$$r_1 = \sup_{t \in I^*_{\beta,\delta}} \varphi_x(s_1,t), \quad r_2 = \inf_{t \in I^*_{\beta,\delta}} \varphi_x(s_2,t)$$

$$\tilde{r_1} = \sup_{t \in I^*_{\beta,\delta}} \varphi_{\tilde{x}}(s_1,t), \quad \tilde{r_2} = \inf_{t \in I^*_{\beta,\delta}} \varphi_{\tilde{x}}(s_2,t).$$

Lemma 2.4.7. *We have $r_1 < r_2$ and $\tilde{r_1} < \tilde{r_2}$.*
Moreover, there exist R_1, R_2 with $0 \leq R_1 < R_2$ satisfying $R_2 - R_1 \geq 2(h\varepsilon - |I^*_{\beta,\delta}|)$, such that $(r_1, r_2) \cap (\tilde{r_1}, \tilde{r_2}) = (R_1, R_2)$.

We admit this result to continue the proof, and will give a demonstration later.
We can now prove that the pairs $\left(X^x_{S_2}, S^x_2\right)$ and $\left(\tilde{X}^{\tilde{x}}_{\tilde{S}_2}, \tilde{S}^{\tilde{x}}_2\right)$ are both η-continuous on $I^*_{\beta,\delta} \times (R_1, R_2)$ with some $\eta > 0$ that we are going to define after the computations.
We first prove that $\left(X^x_{S_2}, S^x_2\right)$ is η-continuous on $I^*_{\beta,\delta} \times (r_1, r_2)$. By the same way we can prove that $\left(\tilde{X}^{\tilde{x}}_{\tilde{S}_2}, \tilde{S}^{\tilde{x}}_2\right)$ is η-continuous on $I^*_{\beta,\delta} \times (\tilde{r}_1, \tilde{r}_2)$. These two facts imply immediately the continuity with (R_1, R_2) since the interval (R_1, R_2) is included in (r_1, r_2) and $(\tilde{r}_1, \tilde{r}_2)$.
Let $(u_1, u_2) \subset (r_1, r_2)$ and $A \subset I^*_{\beta,\delta}$. We have :

$$\mathbb{P}\left(X^x_{S_2} \in A, S^x_2 \in (u_1, u_2)\right) \geq \int_{I_A} \int_{B^\varepsilon_{\bar{y}}} Q(x, g(s)) Q(g(s), g(t)) \mathbf{1}_{\varphi_x(s,t) \in (u_1,u_2)} \mathrm{d}s\mathrm{d}t.$$

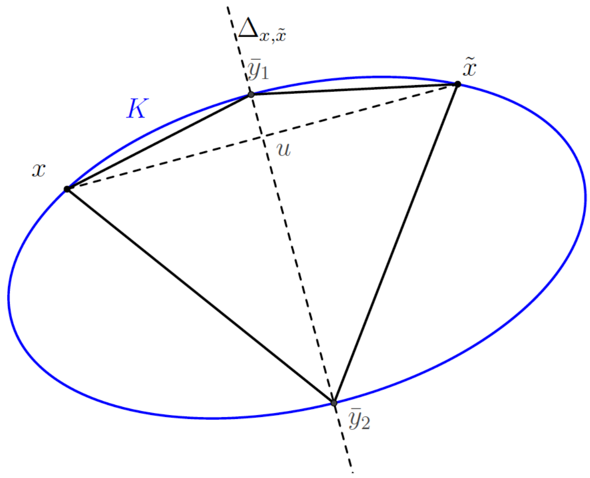


FIGURE 2.8 – Upper bound for the distance $\|w - \bar{y}\|$, $w \in \{x, \tilde{x}\}$

Let $s \in B^{\varepsilon}_{\bar{y}}$ and $t \in I^*_{\beta,\delta}$. We now give a lower bound of $Q(x, g(s))$ and $Q(g(s), g(t))$.
Proposition 2.4.1 gives :

$$\begin{aligned} Q(x, g(s)) &= \frac{\rho(U_x^{-1} l_{x,g(s)}) \cos\left(\varphi_{g(s),x}\right)}{\|x - g(s)\|} \\ &\geq \frac{c\rho_{\min}}{2D}\left(\frac{1}{C} - \varepsilon\right), \end{aligned}$$

where we have used the same method as in he proof of Theorem 2.4.2 (with Figure 2.7) to get that $\cos\left(\varphi_{g(s),x}\right) \geq \frac{\|x-g(s)\|c}{2}$, and then the fact that $\|x - g(s)\| \geq \frac{1}{C} - \varepsilon$. Let us prove this latter. With the notations of Figure ??, by Pythagore's theorem we have, for $\bar{y} \in \{\bar{y}_1, \bar{y}_2\}$, $\|x - \bar{y}\|^2 = \left(\frac{\|x-\tilde{x}\|}{2}\right)^2 + \|u - \bar{y}\|^2$. Moreover, since the curvature of K is bounded by C, it follows that $\|\bar{y}_1 - \bar{y}_2\| \geq \frac{2}{C}$, and then $\max\{\|u - \bar{y}_1\|; \|u - \bar{y}_2\|\} \geq \frac{1}{C}$. We deduce : $\max\{\|x - \bar{y}_1\|; \|x - \bar{y}_2\|\} \geq \frac{1}{C}$. Therefore, by the definition of $\bar{y}$, we have $\|x - \bar{y}\| \geq \frac{1}{C}$. Thus, the reverse triangle inequality gives, for $s \in B^{\varepsilon}_{\bar{y}}$, $\|x - g(s)\| \geq \frac{1}{C} - \varepsilon$.
By the same way, since $\|g(t) - g(s)\| \geq \beta - \varepsilon$, we have :

$$Q\left(g(s), g(t)\right) \geq \frac{c\rho_{\min}}{2D}\left(\beta - \varepsilon\right).$$

Therefore we get :

$$\mathbb{P}\left(X^x_{S_2} \in A, S^x_2 \in (u_1, u_2)\right) \geq a \int_{I_A} \int_{B^{\varepsilon}_{\bar{y}}} \mathbf{1}_{\varphi_x(s,t) \in (u_1,u_2)} \mathrm{d}s\mathrm{d}t,$$

with

$$a = \left(\frac{c\rho_{\min}}{2D}\right)^2 \left(\frac{1}{C} - \varepsilon\right)\left(\beta - \varepsilon\right). \tag{2.9}$$

Let us define, for $t \in I^*_{\beta,\delta}$:

$$M_{x,t}(u_1,u_2) := \left\{ s \in B^{\varepsilon}_{\tilde{y}} : \varphi_x(s,t) \in (u_1,u_2) \right\}.$$

Using the fact that $s \mapsto \varphi_x(s,t)$ is strictly increasing on $B^{\varepsilon}_{\tilde{y}}$ for $t \in I^*_{\beta,\delta}$ we get ($\varphi_w^{-1}(s,t)$ stands for the inverse function of $s \mapsto \varphi_x(s,t)$) :

$$\begin{aligned} |M_{x,t}(u_1,u_2)| &= \left| \left\{ s \in B^{\varepsilon}_{\tilde{y}} : s \in \left(\varphi_x^{-1}(u_1,t), \varphi_x^{-1}(u_2,t) \right) \right\} \right| \\ &= \left| (s_1,s_2) \cap \left(\varphi_x^{-1}(u_1,t), \varphi_x^{-1}(u_2,t) \right) \right|. \end{aligned}$$

By definition of r_1 and r_2, and since $(u_1,u_2) \subset (r_1,r_2)$ we have :

$$\varphi_x(s_1,t) \leq r_1 \leq u_1 \quad \text{and} \quad \varphi_x(s_2,t) \geq r_2 \geq u_2,$$

and since $s \mapsto \varphi_x(s,t)$ is strictly increasing :

$$s_1 \leq \varphi_x^{-1}(u_1,t) \quad \text{and} \quad s_2 \geq \varphi_x^{-1}(u_2,t).$$

Therefore we deduce :

$$\begin{aligned} |M_{x,t}(u_1,u_2)| &= \left| \left(\varphi_x^{-1}(u_1,t), \varphi_x^{-1}(u_2,t) \right) \right| \\ &= \left| \varphi_x^{-1}\left((u_1,u_2), t \right) \right| \\ &\geq \frac{1}{2}(u_2-u_1). \end{aligned}$$

For the last inequality we have used the following property. Let $\psi : \mathbb{R} \mapsto \mathbb{R}$ a function. If for all $x \in [a_1,a_2]$ we have $c_1 < \psi'(x) < c_2$ with $0 < c_1 < c_2 < \infty$, then for any interval $I \subset [\psi(a_1), \psi(a_2)]$, we have $c_2^{-1}|I| \leq |\psi^{-1}(I)| \leq c_1^{-1}|I|$. In our case, the Cauchy-Schwarz inequality gives $\partial_s \varphi_x(s,t) \leq 2$ (see Equation (2.10) for the expression of $\partial_s \varphi_x(s,t)$).
Finally we get, with a given by (2.9) :

$$\begin{aligned} \mathbb{P}\left(X^x_{S_2} \in A, S^x_2 \in (u_1,u_2) \right) &\geq a \int_A \frac{1}{2}(u_2-u_1) \mathrm{d}z \\ &= \frac{a}{2}(u_2-u_1)|A|, \end{aligned}$$

which proves that $\left(X^x_{S_2}, S^x_2 \right)$ is $\frac{a}{2}$-continuous on $I^*_{\beta,\delta} \times (\tilde{r}_1, \tilde{r}_2)$.
Thanks to the remarks before, the proof is completed with $\eta = \frac{a}{2}$ and $J = I^*_{\beta,\delta}$. □

Let us now give the proofs of Lemma 2.4.6 and 2.4.7 that we have admitted so far.

Proof of Lemma 2.4.6. We use the notations introduced at the end of the proof of Proposition 2.4.4.
We have, for $s,t \in I$:

$$\partial_s \varphi_w(s,t) = \left\langle \frac{g(s)-w}{\|g(s)-w\|} + \frac{g(s)-g(t)}{\|g(s)-g(t)\|}, g'(s) \right\rangle. \tag{2.10}$$

By the definition of g, we note that $g'(s)$ is a director vector of the tangent line of ∂K at point $g(s)$.
It is easy to verify that for $w \in \{x, \tilde{x}\}$, there exists a unique $t \in I \setminus \{s_{\bar{y}}\}$ such that

$$\partial_s \varphi_w(s_{\bar{y}}, t) = 0. \tag{2.11}$$

For $w = x$ (resp. $w = \tilde{x}$), we denote by t_{z_x} (resp. $t_{z_{\tilde{x}}}$) this unique element of I. With our notations we thus have $g(t_{z_x}) = z_x$ and $g(t_{z_{\tilde{x}}}) = z_{\tilde{x}}$.
Let $w \in \{x, \tilde{x}\}$. We have :

$$\begin{aligned}
\partial_t \partial_s \varphi_w(s,t) &= \partial_t \left(\left\langle \frac{g(s)-g(t)}{\|g(s)-g(t)\|}, g'(s) \right\rangle \right) \\
&= \frac{1}{\|g(t)-g(s)\|} \left(-\langle g'(t), g'(s) \rangle + \right. \\
&\qquad \left. \left\langle \frac{g(t)-g(s)}{\|g(t)-g(s)\|}, g'(t) \right\rangle \left\langle \frac{g(t)-g(s)}{\|g(t)-g(s)\|}, g'(s) \right\rangle \right).
\end{aligned}$$

Let us look at the term in parenthesis. Let us denote by $[\![u,v]\!]$ the oriented angle between the vectors $u, v \in \mathbb{R}^2$. We have :

$$\begin{aligned}
&-\langle g'(t), g'(s) \rangle + \left\langle \frac{g(t)-g(s)}{\|g(t)-g(s)\|}, g'(t) \right\rangle \left\langle \frac{g(t)-g(s)}{\|g(t)-g(s)\|}, g'(s) \right\rangle \\
&= -\cos\left([\![g'(t), g'(s)]\!]\right) + \cos\left([\![g(t)-g(s), g'(t)]\!]\right) \cos\left([\![g(t)-g(s), g'(s)]\!]\right) \\
&= -\cos\left([\![g'(t), g'(s)]\!]\right) + \frac{1}{2}\cos\left([\![g(t)-g(s), g'(t)]\!] - [\![g(t)-g(s), g'(s)]\!]\right) \\
&\qquad + \frac{1}{2}\cos\left([\![g(t)-g(s), g'(t)]\!] + [\![g(t)-g(s), g'(s)]\!]\right) \\
&= -\cos\left([\![g'(t), g'(s)]\!]\right) + \frac{1}{2}\cos\left([\![g'(s), g'(t)]\!]\right) \\
&\qquad + \frac{1}{2}\cos\left([\![g(t)-g(s), g'(t)]\!] + [\![g(t)-g(s), g'(s)]\!]\right) \\
&= -\frac{1}{2}\cos\left([\![g'(t), g'(s)]\!]\right) + \frac{1}{2}\cos\left([\![g(t)-g(s), g'(t)]\!] + [\![g(t)-g(s), g'(s)]\!]\right) \\
&= -\sin\left(\frac{1}{2}\left([\![g(t)-g(s), g'(t)]\!] + [\![g(t)-g(s), g'(s)]\!] + [\![g'(t), g'(s)]\!]\right)\right) \times \\
&\qquad \sin\left(\frac{1}{2}\left([\![g(t)-g(s), g'(t)]\!] + [\![g(t)-g(s), g'(s)]\!] - [\![g'(t), g'(s)]\!]\right)\right) \\
&= -\sin\left([\![g(t)-g(s), g'(s)]\!]\right) \sin\left([\![g(t)-g(s), g'(t)]\!]\right).
\end{aligned}$$

Therefore we get

$$\partial_t \partial_s \varphi_w(s,t) = -\frac{1}{\|g(t)-g(s)\|} \sin\left([\![g(t)-g(s), g'(s)]\!]\right) \sin\left([\![g(t)-g(s), g'(t)]\!]\right),$$

and then

$$|\partial_t \partial_s \varphi_w(s,t)| = \frac{1}{\|g(t)-g(s)\|} \left|\sin\left([\![g(t)-g(s), g'(s)]\!]\right) \sin\left([\![g(t)-g(s), g'(t)]\!]\right)\right|$$

$$= \frac{1}{\|g(t)-g(s)\|} \left|\cos\left(\varphi_{g(s),g(t)}\right)\cos\left(\varphi_{g(t),g(s)}\right)\right|$$

Let $t \in I$ such that $|t - s_{\bar{y}}| \geq \beta$ (we recall that β is introduced at the beginning of the section). Using once more Figure 2.7, we get, as we have done in the proof of Theorem 2.4.2 :

$$\begin{aligned} |\partial_t \partial_s \varphi_w(s,t)| &\geq \frac{1}{\|g(t)-g(s)\|}\left(\frac{\beta c}{2}\right)^2 \\ &\geq \frac{1}{D}\left(\frac{\beta c}{2}\right)^2. \end{aligned} \tag{2.12}$$

Using Equations (2.11) and Equation (2.12), the mean value theorem gives : for $t \in I$ such that $|t - s_{\bar{y}}| \geq \beta$ and $|t - t_{z_w}| \geq \delta$ (δ is introduced at the beginning of the section),

$$|\partial_s \varphi_w(s_{\bar{y}},t)| = |\partial_s \varphi_w(s_{\bar{y}},t) - \partial_s \varphi_w(s_{\bar{y}},t_{z_w})| \geq \frac{1}{D}\left(\frac{\beta c}{2}\right)^2 |t-t_{z_w}| \geq \frac{\delta}{D}\left(\frac{\beta c}{2}\right)^2. \tag{2.13}$$

We want now such an inequality for $s \in I$ near from $s_{\bar{y}}$. We thus compute :

$$\begin{aligned} &\partial_s^2 \varphi_w(s,t) \\ &= \frac{1}{\|w-g(s)\|} + \frac{1}{\|g(s)-g(t)\|} + \left\langle \frac{g(s)-w}{\|g(s)-w\|} + \frac{g(s)-g(t)}{\|g(s)-g(t)\|}, g''(s)\right\rangle \\ &- \frac{1}{\|w-g(s)\|}\left\langle \frac{w-g(s)}{\|w-g(s)\|}, g'(s)\right\rangle^2 - \frac{1}{\|g(s)-g(t)\|}\left\langle \frac{g(s)-g(t)}{\|g(s)-g(t)\|}, g'(s)\right\rangle^2. \end{aligned}$$

We immediately deduce, using the Cauchy-Schwarz inequality, and the fact that $\|g'(s)\| = 1$ for all $s \in I$:

$$\begin{aligned} &|\partial_s^2 \varphi_w(s,t)| \\ &\leq \frac{1}{\|w-g(s)\|} + \frac{1}{\|g(s)-g(t)\|} + 2\|g''(s)\| + \frac{1}{\|w-g(s)\|} + \frac{1}{\|g(s)-g(t)\|} \\ &\leq 2\left(\frac{1}{\|w-g(s)\|} + \frac{1}{\|g(s)-g(t)\|} + C\right), \end{aligned}$$

where we recall that C is the upper bound on the curvature of K.
Let now $t \in I$ such that $|t - s_{\bar{y}}| \geq \beta$ and $|t - t_{z_w}| \geq \delta$, and let $s \in I$ such that $|s - s_{\bar{y}}| \leq \varepsilon$. With such s and t we have $|t - s| \geq \beta - \varepsilon$. Moreover, we have already seen in proof of Proposition 2.4.4 (with Figure ??) that $\|w - g(s)\| \geq \frac{1}{C} - \varepsilon$ for $s \in B_{\bar{y}}^{\varepsilon}$. Therefore, for such s and t :

$$|\partial_s^2 \varphi_w(s,t)| \leq 2\left(\frac{1}{\frac{1}{C}-\varepsilon} + \frac{1}{\beta-\varepsilon} + C\right) = M > 0. \tag{2.14}$$

Using once again the mean value theorem with Equations (2.13) and (2.14), we deduce that for all $t \in I$ such that $|t - s_{\bar{y}}| \geq \beta$ and $|t - t_{z_w}| \geq \delta$, and for all $s \in I$

such that $|s - s_{\bar{y}}| \leq \varepsilon$:

$$|\partial_s \varphi_w(s,t)| \geq \frac{\delta}{D}\left(\frac{\beta c}{2}\right)^2 - \varepsilon M = h > 0.$$

Let us now take $I^*_{\beta,\delta} \subset I \setminus \{s_{\bar{y}}, t_{z_x}, t_{z_{\tilde{x}}}\}$ an interval of length strictly smaller than $h\varepsilon$ (this condition on the length of $I^*_{\beta,\delta}$ is not necessary for the lemma, but for the continuation of the proof of the proposition), and such that for all $t \in I^*_{\beta,\delta}$, $|t - t_{z_x}| \geq \delta$, $|t - t_{z_{\tilde{x}}}| \geq \delta$ and $|t - s_{\bar{y}}| \geq \beta$. In order to ensure that $I^*_{\beta,\delta}$ is not empty, we take β and δ such that $\frac{|\partial K|}{3} - \max\{2\delta; \beta+\delta\} > 0$. Indeed, it is necessary that one of the intervals "$(t_{z_x}, t_{z_{\tilde{x}}})$", "$(t_{z_x}, s_{\bar{y}})$" and "$(s_{\bar{y}}, t_{z_{\tilde{x}}})$" at which we removes a length β or δ on the good extremity, is not empty. And since the larger of these intervals has a length at least $\frac{\partial K}{3}$, we obtain the good condition on β and δ.
We thus just proved that $|\partial_s \varphi_w(s,t)| \geq h$ for $s \in B^{\varepsilon}_{\bar{y}}$ and $t \in I^*_{\beta,\delta}$, which is the result of the lemma. □

Proof of Lemma 2.4.7. Let us first prove that $r_1 < r_2$. We do it only for r_1 and r_2 since it is the same for $\tilde{r_1}$ and $\tilde{r_2}$. We have :

$$\begin{aligned} r_2 - r_1 &= \inf_{t\in I^*_{\beta,\delta}} \varphi_x(s_2,t) - \sup_{t\in I^*_{\beta,\delta}} \varphi_x(s_1,t) \\ &= \inf_{t\in I^*_{\beta,\delta}} \varphi_x(s_2,t) - \inf_{t\in I^*_{\beta,\delta}} \varphi_x(s_1,t) - \left(\sup_{t\in I^*_{\beta,\delta}} \varphi_x(s_1,t) - \inf_{t\in I^*_{\beta,\delta}} \varphi_x(s_1,t)\right) \\ &\geq h(s_2 - s_1) - \left(\sup_{t\in I^*_{\beta,\delta}} |\partial_t \varphi_x(s_1,t)|\right) \left|I^*_{\beta,\delta}\right| \\ &\geq 2h\varepsilon - \left|I^*_{\beta,\delta}\right|, \end{aligned}$$

and this quantity is strictly positive since $|I^*_{\beta,\delta}| < h\varepsilon$ by construction.
For the first inequality, we have used the mean value theorem twice, and for the last inequality, we have used the fact that
$\sup_{t\in I^*_{\beta,\delta}} |\partial_t \varphi_x(s_1,t)| = \sup_{t\in I^*_{\beta,\delta}} \left|\left\langle \frac{g(t)-g(s_1)}{\|g(t)-g(s_1)\|}, g'(t)\right\rangle\right| \leq 1$ thanks to the Cauchy-Schwarz inequality.
Let us now prove that the intersection $(r_1, r_2) \cap (\tilde{r}_1, \tilde{r}_2)$ is not empty.
Let $t \in I^*_{\beta,\delta}$, we have :

$$\begin{aligned} r_2 - \varphi_x(s_{\bar{y}},t) &= \inf_{t\in I^*_{\beta,\delta}} \varphi_x(s_2,t) - \varphi_x(s_{\bar{y}},t) \\ &= \inf_{t\in I^*_{\beta,\delta}} \varphi_x(s_2,t) - \inf_{t\in I^*_{\beta,\delta}} \varphi_x(s_{\bar{y}},t) - \left(\varphi_x(s_{\bar{y}},t) - \inf_{t\in I^*_{\beta,\delta}} \varphi_x(s_{\bar{y}},t)\right) \\ &\geq h(s_2 - s_{\bar{y}}) - |I^*_{\beta,\delta}| \\ &= h\varepsilon - |I^*_{\beta,\delta}| \\ &> 0, \end{aligned}$$

once again thanks to the mean value theorem. Similarly we have

$$\begin{aligned}\varphi_x(s_{\bar{y}},t)-r_1&=\varphi_x(s_{\bar{y}},t)-\sup_{t\in I^*_{\beta,\delta}}\varphi_x(s_1,t)\\&=\varphi_x(s_{\bar{y}},t)-\sup_{t\in I^*_{\beta,\delta}}\varphi_x(s_{\bar{y}},t)-\left(\sup_{t\in I^*_{\beta,\delta}}\varphi_x(s_1,t)-\sup_{t\in I^*_{\beta,\delta}}\varphi_x(s_{\bar{y}},t)\right)\\&\geq -|I^*_{\beta,\delta}|+h(s_{\bar{y}}-s_1)\\&=h\varepsilon-|I^*_{\beta,\delta}|\\&>0.\end{aligned}$$

Moreover, since $\bar{y}\in\Delta_{x,\tilde{x}}$, we have $\varphi_x(s_{\bar{y}},t)=\varphi_{\tilde{x}}(s_{\bar{y}},t)$, and we thus can prove the same inequalities with $\tilde{r}_1$ and $\tilde{r}_2$ instead of r_1 and r_2.
Finally we thus get that the interval $(R_1,R_2)=(r_1,r_2)\cap(\tilde{r_1},\tilde{r_2})$ is well defined and

$$R_2-R_1\geq 2\left(h\varepsilon-|I^*_{\beta,\delta}|\right).$$

□

Remark 2.4.8. *The fact that $|\mathcal{J}|=\pi$ is here to ensure that the process can go from x and $\tilde{x}$ to $\bar{y}$ in the proof of Proposition 2.4.4. If $|\mathcal{J}|<\pi$, since x and $\tilde{x}$ are unspecified and $\bar{y}$ can therefore be everywhere on ∂K, nothing ensures that this path is available.*

We can now state the following theorem on the speed of convergence of the stochastic billiard in the convex set K.

Theorem 2.4.9. *Let $K\subset\mathbb{R}^2$ satisfying Assumption $(\mathcal{K})$ with diameter D. Let $((X_t,V_t))_{t\geq 0}$ the stochastic billiard process evolving in K and verifying Assumption $(\mathcal{H})$ with $|\mathcal{J}|=\pi$.*
There exists a unique invariant probability measure χ on $K\times\mathbb{S}$ for the process $((X_t,V_t))_{t\geq 0}$.
Moreover, let us define n_0 and p by (2.15) and (2.16) with $\zeta\in\left(0,\frac{2}{C}\right)$. Let us consider η, $I^*_{\beta,\delta}$, R_1,R_2 as in Proposition 2.4.4 and Lemma 2.4.6, and let us define κ by (2.17).
For all $t\geq 0$ and all $\lambda<\lambda_M$:

$$\|\mathbb{P}\left(X_t\in\cdot,V_t\in\cdot\right)-\chi\|_{TV}\leq C_\lambda \mathrm{e}^{-\lambda t},$$

where

$$\lambda_M=\min\left\{\frac{1}{2D}\log\left(\frac{1}{1-p}\right);\frac{1}{2D}\log\left(\frac{-(1-p)+\sqrt{(1-p)^2+4p(1-\kappa)}}{2p(1-\kappa)}\right)\right\}$$

and

$$C_\lambda=\frac{p\kappa\mathrm{e}^{5\lambda D}}{1-\mathrm{e}^{2\lambda D}(1-p)-\mathrm{e}^{4\lambda D}p(1-\kappa)}.$$

Proof. As previously, the existence of an invariant probability measure for the stochastic billiard process comes from the compactness of $K\times\mathbb{S}^1$. The following proof ensures its uniqueness and gives an explicit speed of convergence.
Let $((X_t,V_t))_{t\geq 0}$ and $((\tilde{X}_t,\tilde{V}_t))_{t\geq 0}$ be two versions of the stochastic billiard with $(X_0,V_0)=(x_0,v_0)\in K\times\mathbb{S}^1$ and $(\tilde{X}_0,\tilde{V}_0)=(\tilde{x}_0,\tilde{v}_0)\in K\times\mathbb{S}^1$.
We define (or recall the definition for T_0 and $\tilde{T}_0$) :

$$T_0=\inf\{t\geq 0, x_0+tv_0\notin K\},\qquad w=x_0+T_0v_0\in\partial K,$$

and

$$\tilde{T}_0=\inf\{t\geq 0, \tilde{x}_0+t\tilde{v}_0\notin K\},\qquad \tilde{w}=\tilde{x}_0+\tilde{T}_0\tilde{v}_0\in\partial K.$$

Step 1. From Proposition 2.4.3, we deduce that for all $x\in\partial K$ and all $\zeta\in\left(0,\frac{1}{C}\right)$, S_n^x is $(c\rho_{\min})^n\zeta^{n-1}$-continuous on the interval
$\Gamma_n=\left[(n-1)\zeta,\frac{nC}{2}-(n-1)\zeta\right]$.
Let thus $\zeta\in\left(0,\frac{1}{C}\right)$ and let us define

$$n_0=\min\left\{n\geq 1:|\Gamma_n|>D\right\}=\left\lfloor\frac{D-2\zeta}{2\left(\frac{1}{C}-1\right)}\right\rfloor+1. \tag{2.15}$$

The variables T_{n_0} and $\tilde{T}_{n_0}$ are both $(c\rho_{\min})^{n_0}\zeta^{n_0-1}$-continuous on
$\left[T_0+(n_0-1)\zeta,T_0+\frac{nC}{2}-(n_0-1)\zeta\right]\cap\left[\tilde{T}_0+(n_0-1)\zeta,\tilde{T}_0+\frac{nC}{2}-(n_0-1)\zeta\right]$.
Since $\left|T_0-\tilde{T}_0\right|\leq D$, this intersection is non-empty and its length is larger that $\frac{2n_0}{C}-2(n_0-1)\zeta-D$.
Let us define

$$p=(c\rho_{\min})^{n_0}\zeta^{n_0-1}\left(\frac{2n_0}{C}-2(n_0-1)\zeta-D\right). \tag{2.16}$$

Using the fact that the for all $x\in\partial K$, $S_{n_0}^x\leq n_0D$ almost surely, we deduce that we can construct a coupling such that the coupling-time T_c^1 of T_{n_0} and $\tilde{T}_{n_0}$ satisfies :

$$T_c^1\leq_{st}T_0+n_0DG^1$$

with $G^1\sim\mathcal{G}(p)$.

Step 2. Let us now suppose that T_c^1, $((X_t,V_t))_{0\leq t\leq T_1^c}$ and $((\tilde{X}_t,\tilde{V}_t))_{0\leq t\leq T_1^c}$ are constructed as described above, .
We define $y=X_{T_c^1}$ and $\tilde{y}=\tilde{X}_{T_c^1}$, which are by construction of T_c^1 on ∂K. We also define $N_c^1=\min\left\{n>0:X_{T_n}=y\right\}$, which is deterministic conditionally to T_c^1.
By the Proposition 2.4.4, we can construct a coupling of $(X_{S_2}^y,S_2^y)$ and $(\tilde{X}_{\tilde{S}_2}^{\tilde{y}},\tilde{S}_2^y)$ such that

$$\mathbb{P}\left(X_{S_2}^y=\tilde{X}_{\tilde{S}_2}^{\tilde{y}}\ \text{ and }\ S_2^y=\tilde{S}_2^y\right)\geq\eta|I_{\beta,\delta}^*|(R_2-R_1).$$

Therefore we can construct $((X_t,V_t))$ and $((\tilde{X}_t,\tilde{X}_t))$ until time $T_{N_c^1+2}$ such that

$$\mathbb{P}\left(X_{T_{N_c^1+2}}=\tilde{X}_{T_{N_c^1+2}}\ \text{ and }\ T_{N_c^1+2}=\tilde{T}_{N_c^1+2}\right)\geq\eta|I_{\beta,\delta}^*|(R_2-R_1).$$

Defining

$$\kappa = \eta |I^*_{\beta,\delta}|(R_2 - R_1), \tag{2.17}$$

we get that the entire coupling-time of the two processes satisfies :

$$\hat{T} \leq_{st} T_0 + \sum_{l=1}^{G} \left(n_0 D G^l + n_0 D \right) = T_0 + \sum_{l=1}^{G} \left(n_0 D (G^l + 1) \right)$$

where G as a geometric distribution with parameter κ and the $(G^l)_{l\geq 1}$ are independent geometric random variables with parameter p, and independent of G.
Finally, we get

$$\mathbb{P}\left(\hat{T} > t\right) \leq \mathrm{e}^{-\lambda t} \frac{p\kappa \mathrm{e}^{5\lambda D}}{1 - \mathrm{e}^{2\lambda D}(1-p) - \mathrm{e}^{4\lambda D} p (1-\kappa)},$$

for all $\lambda \in (0, \lambda_M)$.

□

2.5 Discussion

All the results presented in this paper are in dimension 2. However, the ideas developed here can be adapted to higher dimensions. Let us briefly explain it.

Stochastic billiard in a ball of $\mathbb{R}^d$

Let us first look at the stochastic billiard (X, V) in a ball $\mathcal{B} \subset \mathbb{R}^d$ with $d \geq 2$. As we have done in Section 2.3, we can represent the Markov chain $(X_{T_n}, V_{T_n})_{n\geq 0}$ by another Markov chain. Indeed, for $n \geq 1$, the position $X_{T_n} \in \partial\mathcal{B}$ can be uniquely represented by its hyperspherical coordinates : a $(d-1)$-tuple $(\Phi_n^1, \cdots, \Phi_n^{d-1})$ with $\Phi_n^1, \cdots, \Phi_n^{d-2} \in [0, \pi)$ and $\Phi_n^{d-1} \in [0, 2\pi)$.
Similarly, for $n \geq 1$, the vector speed $V_{T_n} \in \left\{ v \in \mathbb{S}^{d-1} : v \cdot n_{X_{T_n}} \geq 0 \right\}$ can be represented by its hyperspherical coordinates.
Thereby, we can give relations between the different random variables as in Proposition 2.3.1, and in theory, we can do explicit computations to get lower bounds on the needed density function. Then the same coupling method in two steps can be applied. Nevertheless, it could be difficult to manage the computations in practice when the dimension increases.

Stochastic billiard in a convex set $K \subset \mathbb{R}^d$

To get bounds on the speed of convergence of the stochastic billiard (X, V) in a convex set $K \subset \mathbb{R}^d$, $d \geq 2$, satisfying Assumption $(\mathcal{K})$, we can apply exactly the same method as in Section 2.4. The main difficulty could be the proof of the equivalent of Proposition 2.4.4. But it can easily be adapted, and we refer to the proof of Lemma 5.1 in [Comets *et al.*, 2009], where the authors lead the proof in dimension $d \geq 3$.

Chapitre 3

Long-time behaviour of generalized Zig-Zag process

This Chapter is the reproduction of the paper [Fétique, 2019].

We study the long-time behaviour of a class of piecewise-deterministic Markov processes which are an extension of some recent works. These d-dimensional processes, $d \geq 1$, can especially be used to model the motion of a bacterium in presence of a chemo-attractant, with parameters depending both on the position and the velocity of the bacterium. Using the method of Meyn and Tweedie ([Meyn et Tweedie, 1993a, Meyn et Tweedie, 1993b]), we show that under some good assumptions on the parameters of the model, such a process converges exponentially fast towards its invariant measure. We also establish the existence of exponential moments of the invariant measure using results on semi-regenerative processes. The one-dimensional case is studied separately since complementary results can be obtained in that particular case.

3.1 Introduction

Piecewise-deterministic Markov processes (PDMPs) have been introduced by Davis ([Davis, 1984]) to distinguish these particular processes from diffusive processes. They are the subject of much current work in various domains, since they are a simple alternative to diffusions to model stochastic systems (see [Azaïs *et al.*, 2014] for an overview of recent results on PDMPs). In this paper, we study a PDMP that comes from biological modeling for the motion of flagellated bacteria which are in presence of a chemo-attractant. The motion of such a bacterium has been described as run-and-tumble, which means that the bacterium alternates sequences of linear runs with periods of random reorientation (tumbling). The tumble-periods being typically much shorter than the run-periods, we can suppose them to be instantaneous. Moreover, the presence

of a chemo-attractant in the environment of the bacterium influences the rate at which the bacterium changes its direction, and also the new direction it takes (see [Othmer *et al.*, 1988] for more details on the model).
More precisely, we consider the PDMP $((X_t, V_t))_{t\geq 0}$ with values in $E = \mathbb{R}^d \times \mathcal{B}(1)$, where $\mathcal{B}(1) = \left\{v \in \mathbb{R}^d, |v| \leq 1\right\}$ is the Euclidian ball of radius 1, with infinitesimal generator given by, for h in the domain of L (see [Davis, 1984] for a definition of the domain of the infinitesimal generator) :

$$Lh(x,v) = v \cdot \nabla_x h(x,v) + \lambda(x,v) \int_{\mathcal{B}(1)} \left(h(x,v') - h(x,v)\right) Q(x,v,\mathrm{d}v'), \quad (3.1)$$

where $Q(x,v,\cdot)$ is a probability kernel on $\mathcal{B}(1)$. We call this process "generalized Zig-Zag process" since it is a generalization of the Zig-Zag process studied in [Bierkens et Roberts, 2017] and [Bierkens *et al.*, 2019], in the sense that we do not any more have a velocity with values in $\{-1,1\}^d$, but in $\mathcal{B}(1)$.
In our model, X_t represents the position of a bacterium at instant t, and V_t its velocity. The form of the generator indicates that the first component X is continuous and evolves according to $\frac{\mathrm{d}X_t}{\mathrm{d}t} = V_t$, whereas V is constant during a random time, and jumps according to the kernel Q at rate $\lambda(x,v)$ when $(X_t, V_t) = (x,v)$. The fact that the motion of the bacterium is biased by the presence of a chemo-attractant will be taken into account in the assumptions we will make further on the jump rate λ and the velocity kernel Q.
In this paper, we are interested in the long-time behaviour of the process driven by (3.1) under some good assumptions.

Many special cases of this process, driven by (3.1), have already been studied in different ways and under different assumptions. Let first mention some works on the process in dimension 1 with a modeling point of view : in [Fontbona *et al.*, 2012], Fontbona, Guérin and Malrieu have shown the exponential ergodicity of the process with a jump rate equal to $a\mathbf{1}_{xv\leq 0} + b\mathbf{1}_{xv>0}$ with $b > a > 0$, and the velocity taking its values in $\{-1,+1\}$. For this, they give an exact description of the excursions of the process away from the origin and give an explicit construction of a coalescent coupling for both velocity and position. In [Fontbona *et al.*, 2016] and [Bierkens et Roberts, 2017], the previous result has been extended by considering a more general jump rate, depending on the position and the velocity of the bacterium. Calvez, Raoul and Schmeiser have shown in [Calvez *et al.*, 2015] by an analytical method the exponential ergodicity of the process driven by (3.1) in the particular case where the kernel Q is the uniform kernel on $[-1,1]$, and under similar assumptions to the ones introduced here (see Section 3.5).
Furthermore, there exist also results for this process in high dimension. We can for instance cite [Bierkens *et al.*, 2019], in which Bierkens, Fearnhead and Roberts study the Zig-Zag process, that is the process with values in $\mathbb{R}^d \times \{-1,+1\}^d$. They prove its ergodicity in the case where it can be seen as a product of independent one-dimensional Zig-Zag processes. Then, Bierkens, Roberts and Zitt generalize these results in [Bierkens *et al.*, 2017], not considering

only the case of a product of one-dimensional Zig-Zag process. In [Bouchard-Côté *et al.*, 2018], [Deligiannidis *et al.*, 2019] and [Monmarché, 2016], the authors are interested in the ergodicity of the bouncy particle sampler, with values in $\mathbb{R}^d \times \mathbb{R}^d$ or $\mathbb{R}^d \times \mathcal{S}^{d-1}$, where $\mathcal{S}^{d-1}$ is the unit sphere of $\mathbb{R}^d$. This PDMP is a particular case of the process driven by (3.1) : for instance in [Monmarché, 2016], the jump rate is given by $\lambda(x, v) = (v \cdot \nabla_x U(x))_+$, where U is a potential, and at each jump, the velocity is reflected according to optical laws on the level set of U it has reached. Recently, a new model has been studied : in [Wu et Robert, 2018], Robert and Wu introduce the coordinate sample, which is a variant of the Zig-Zag process, since the velocity does not live in $\{-1, +1\}^d$ but in $\{e_i, 1 \leq i \leq d\}$ the canonical base of $\mathbb{R}^d$. For this process, only one component of the position evolves between the jumps. The authors show in the paper the exponential ergodicity of this process under some conditions.
Finally, let us mention that the study of this kind of processes has an interest not only for biological modelling, but also for simulating a target distribution. In fact, almost all of the recent study in dimension d do this for the sampling. Let us refer to [Bierkens *et al.*, 2019], [Bierkens et Roberts, 2017], [Bierkens *et al.*, 2017], [Bouchard-Côté *et al.*, 2018], [Deligiannidis *et al.*, 2019], [Durmus *et al.*, 2018], [Monmarché, 2016] and [Wu et Robert, 2018], where the authors want to sample from a distribution with a density proportional to e^{-U}, where U is a potential on $\mathbb{R}^d$. For a jump rate λ and a jump kernel Q well chosen (with respect to U), the PDMP converges towards the targeted distribution. The estimation of the speed of convergence to equilibrium of these processes gives then information on the efficiency of the corresponding algorithms to sample from the target distribution. Comparisions of the efficiency of the different samplers are done in [Andrieu *et al.*, 2018] and [Wu et Robert, 2018] for instance. An interest of considering PDMPs to catch a distribution is the irreversibility of PDMPs. Indeed, while many Markov chain Monte Carlo (MCMC) methods rely on realisations from a discrete reversible ergodic Markov chain, it has been observed that non-reversibility often implies favourable convergence properties (see for instance [Hwang *et al.*, 2005, Lelièvre *et al.*, 2013]). Moreover, PDMPs have the advantage to be easy to sample, and can even in some cases being simulated without discretization error.

Framework

Let us now introduce the framework of the paper. Denoting by $x \cdot v$ the scalar product of $x \in \mathbb{R}^d$ and $v \in \mathbb{R}^d$, and $|x|$ the Euclidean norm of the vector x, the assumptions made on the model are the following :

$(\mathcal{H}_1)$: There exists $\lambda_{\min} > 0$ such that for all $(x, v) \in E$, $\lambda(x, v) \geq \lambda_{\min}$;

$(\mathcal{H}_2)$: The quantity $\lambda_{\max} = \sup_{\{(x,v)\in E \,:\, x\cdot v\leq 0\}} \lambda(x, v)$ is finite ;

$(\mathcal{H}_3)$: There exists $p > 0, \theta_0, \bar{\theta} \in (0, 1]$ such that for all $(x, v) \in E$ satisfying $\frac{x\cdot v}{|x|} > -\bar{\theta}$,

$$\int_{\{v'\in\mathcal{V} \,:\, \frac{x\cdot v'}{|x|}\leq -\theta_0\}} Q(x, v, \mathrm{d}v') \geq p;$$

$(\mathcal{H}_4)$: There exists $\theta_* \in \left[0, (p\theta_0)^2 \frac{\lambda_{\min}}{\lambda_{\max}}\right)$, $\beta > \frac{1}{(p\theta_0)^2}$ and $\Delta > 0$ such that

$$\inf_{\left\{\frac{x \cdot v}{|x|} \geq \theta_*, |x| \geq \Delta\right\}} \lambda(x, v) \geq \beta \lambda_{\max}.$$

Assumptions $(\mathcal{H}_1)$ and $(\mathcal{H}_2)$ mean that the jump rate is uniformly bounded by below, and that it is bounded from above when the bacterium is moving towards 0, where the chemo-attractant is assumed to be. The existence of the lower bound $\lambda_{\min}$ is here to ensure the irreducibility of the process. We refer to Section 1.3 of [Bierkens *et al.*, 2017] for illustrations of the difficulties if the switching rate λ can be zero.
Assumptions $(\mathcal{H}_3)$ and $(\mathcal{H}_4)$ reflect the attraction of the bacterium to the origin. Indeed, in Assumption $(\mathcal{H}_3)$, we suppose that if the bacterium does not "enough" go towards the origin, when a jump happens, it has a chance to go towards it. Moreover, in Assumption $(\mathcal{H}_4)$, we assume a kind of monotonicity of the jump rate. Roughly speaking, we suppose that when the bacterium is far from the origin, and goes in a too bad direction, its jump rate is strictly bigger than $\lambda_{\max}$, which is the maximum of the jump rate when the bacterium is coming closer the origin.
Assumptions $(\mathcal{H}_1), (\mathcal{H}_2)$ and $(\mathcal{H}_4)$ appear in the works [Bierkens et Roberts, 2017, Fontbona *et al.*, 2012, Fontbona *et al.*, 2016]. However, in these papers, there is no assumption equivalent to $(\mathcal{H}_3)$. But it is in fact normal : the Zig-Zag process has a velocity in $\{-1, 1\}$, and thus, as soon as the bacterium goes in a bad direction, the jump makes it go towards the origin, and the assumption is in fact satisfied.
All of these assumptions seem to be reasonable for the modelling of the motion of bacteria as described above.

Remark 3.1.1. *We only consider the case of the speed in the ball unit ball of $\mathbb{R}^d$, but the results can easily be adapted if the speed lies in a compact set K of $\mathbb{R}^d$.*
Moreover, although we could hope to obtain similar results with an unbounded velocity (with probably other assumptions on the jump rate), we do not deal with this case since it is not really relevant from a modelling point of view.

Under these assumptions, we can show the exponential ergodicity (see Section 3.2 for the definition) of the generalized Zig-zag process (X, V).

Theorem 3.1.2. *Let $(X_t, V_t)_{t \geq 0}$ be a PDMP on $E = \mathbb{R}^d \times \mathcal{B}(1)$ with infinitesimal generator given by (3.1).*
If λ and Q satisfy Assumptions $(\mathcal{H}_1)$, $(\mathcal{H}_2)$, $(\mathcal{H}_3)$ and $(\mathcal{H}_4)$, then the process is exponentially ergodic.

In Figure 3.1, we can observe the convergence of the empirical law of $(X_t)_{t\geq 0}$ in the case $\lambda(x, v) = \mathbf{1}_{xv<0} + 2\mathbf{1}_{xv \geq 0}$ and $v \in \{-1, +1\}$, that is $Q(x, v, \mathrm{d}v') = \delta_{-v}(v')\mathrm{d}v'$, and in Figure 3.2, we observe it with the same jump rate and initial conditions, but with $Q(x, v, \mathrm{d}v') = \frac{1}{2}\mathbf{1}_{[-1,1]}(v')\mathrm{d}v'$.

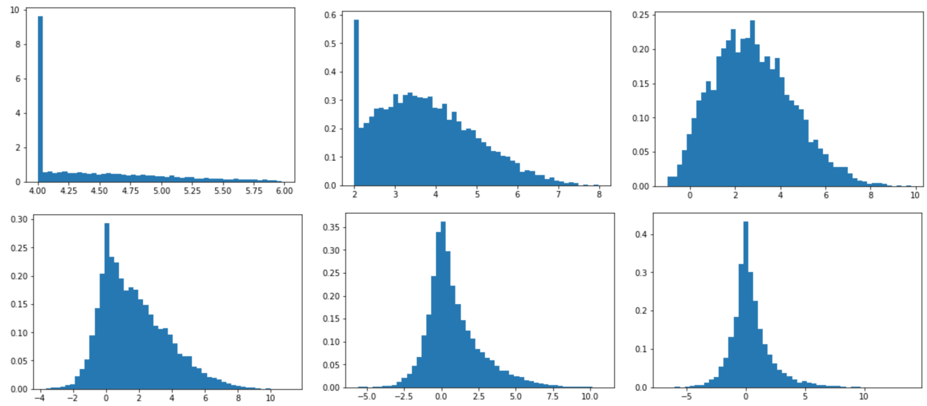


FIGURE 3.1 – Empirical distribution of X_t starting at $(5, -1)$ for $t \in \{1, 3, 6, 10, 16, 22\}$ with $\lambda(x, v) = \mathbf{1}_{xv<0} + 2\mathbf{1}_{xv\geq 0}$ and $v \in \{-1, +1\}$.

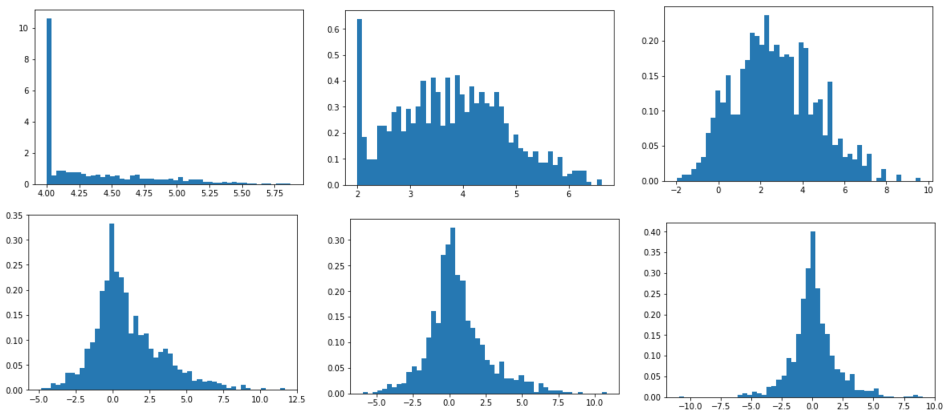


FIGURE 3.2 – Empirical distribution of X_t starting at $(5, -1)$ for $t \in \{1, 3, 10, 30, 40, 100\}$ with $\lambda(x, v) = \mathbf{1}_{xv<0} + 2\mathbf{1}_{xv\geq 0}$ and $Q(x, v, \mathrm{d}v') = \frac{1}{2}\mathbf{1}_{[-1,1]}(v')\mathrm{d}v'$.

The interest of this convergence result in relation to those already existing is that it concerns a very general class of PDMPs.
The approach we will carry out in Section 3.3 to prove Theorem 3.1.2 is the classical method of Meyn and Tweedie ([Down *et al.*, 1995, Meyn et Tweedie, 1993a, Meyn et Tweedie, 1993b]), by showing the existence of a Lyapunov function for the process, and the existence of petite sets. In Section 3.2, we thus first briefly recall the generalities about ergodicity of Markov processes.
Then, we will study in Section 3.4 the existence of exponential moments for the invariant measure in Theorem 3.4.2. For this purpose, we will use a convergence result on semi-regenerative processes, since our PDMP is such a process.
Finally, we will go back to the particular case of the dimension 1 in Section 3.5, in which we study our process with a different approach than in Section 3.3. In particular, in this section, we will assume that the jump kernel does not depend on the position of the bacterium. Thus, only the jump rate will favour the return towards the origin, and therefore the ergodicity of the process. We will establish in Theorems 3.5.3 and 3.5.9 the exponential ergodicity of the process and the existence of exponential moments for its invariant probability measure. The results obtained in this section can be linked with the previous works [Calvez *et al.*, 2015], [Fontbona *et al.*, 2012] and [Fontbona *et al.*, 2016], since they are in fact a generalization of these latter.
We mention at the end of the paper that the two studies applied in dimension 1 are complementary, since one can be applied at some process whereas the other can not, and conversely.

3.2 Preliminaries

3.2.1 About ergodicity

In this paper, we study the convergence of our process with the total variation distance. Let us recall its definition. Let ν and $\widetilde{\nu}$ be two probability measures on a measurable space E. We say that a probability measure on $E \times E$ is a coupling of ν and $\widetilde{\nu}$ if its marginals are ν and $\widetilde{\nu}$. Denoting by $\Gamma(\nu, \widetilde{\nu})$ the set of all the couplings of ν and $\widetilde{\nu}$, we say that two random variables X and $\widetilde{X}$ satisfy $(X, \widetilde{X}) \in \Gamma(\nu, \widetilde{\nu})$ if ν and $\widetilde{\nu}$ are the respective laws of X and $\widetilde{X}$. The total variation distance between these two probability measures is then defined by

$$\|\nu - \widetilde{\nu}\|_{TV} = \inf_{(X,\widetilde{X}) \in \Gamma(\nu,\widetilde{\nu})} \mathbb{P}(X \neq \widetilde{X}). \tag{3.2}$$

For other definitions of the total variation distance and its properties, see for instance [Lindvall, 2002].

Let us now introduce some usefull results to study the convergence of a Markov process towards its invariant measure (see [Down *et al.*, 1995], [Meyn et Tweedie, 1993a], [Meyn et Tweedie, 1993b]).

Let $(Y_t)_{t\geq 0}$ be a Markov process on the state space E, and denote by $\mathtt{P}$ its semi-group and $\mathtt{L}$ its infinitesimal generator. We say that the Markov process Y is exponentially ergodic if there exists a probability measure π, a function $M : E \longrightarrow \mathbb{R}_+$ and a constant $0 < \rho < 1$ such that

$$\|\mathtt{P}_t(x,\cdot) - \pi\|_{TV} \leq M(x)\rho^t, \text{ for all } t \geq 0, \tag{3.3}$$

where $\mathtt{P}_t(x,\cdot) = \mathbb{P}_x(Y_t \in \cdot)$.
A set K is said to be petite for the process Y if there exists a probability measure ν on $\mathbb{R}_+$ and a nontrivial measure μ on E such that, for all $x \in K$,

$$\int_0^\infty \mathtt{P}_t(x,\cdot)\nu(\mathrm{d}t) \geq \mu(\cdot). \tag{3.4}$$

This notion is weaker than the notion of small sets : K is said to be a small set for the process Y if there exists $t_0 > 0$ and a non trivial measure μ on E such that, for all $x \in K$,

$$\mathtt{P}_{t_0}(x,\cdot) \geq \mu(\cdot). \tag{3.5}$$

Let $K \subset E$ be a compact set, and let $H : E \longrightarrow \mathbb{R}$ be a function. We say that H is a Lyapunov function (associated to the set K) for the process Y if $H(x) \geq 1$ for all $x \in E$ and if there exists some constants $\alpha > 0$ and $\beta \geq 0$ such that for all $x \in E$,

$$\mathtt{L}H(x) \leq -\alpha H(x) + \beta \mathbf{1}_K(x). \tag{3.6}$$

Finally, we recall that the Markov process Y is called φ-irreducible if there exists a σ-finite measure φ such that for all measurable set A with $\varphi(A) > 0$ we have, for all $x \in E$, $\mathbb{E}_x\left[\int_0^\infty \mathbf{1}_{Y_t\in A}\mathrm{d}t\right] > 0$. In that case, there exists a maximal irreducibility measure ψ such that for any other irreducibility measure ν, ν is absolutely continuous with respect to ψ. We then write $\mathcal{B}^+(E) = \{A \subset E \text{ measurable}, \psi(A) > 0\}$. The process Y is said to be aperiodic if for some small set $C \in \mathcal{B}^+(E)$, there exists $T > 0$ such that for all $t \geq T$ and all $x \in C$ we have $\mathtt{P}_t(x, C) > 0$.

We can now recall the main result we will use in Section 3.3.2 to prove Theorem 3.1.2.

Theorem 3.2.1. *([Down et al., 1995]) Let Y be an irreducible and aperiodic Markov process. If there exists a petite set K for the process Y and a Lyapunov function associated to this set K (i.e. (3.4) and (3.6) hold) , then the process Y is exponentially ergodic.*

3.2.2 Description of the process

Let us now describe the dynamics of the process.
The variables $0 = T_0, T_1, T_2, \ldots$ refer to the successive jumping times of the process, and for $n \geq 1$ we denote by τ_n the inter-jump time : $\tau_n = T_n - T_{n-1}$.
In order to make the paper easier to read, we note V_i the velocity on the time

interval $[T_i, T_{i+1})$, instead of V_{T_i}.
Finally, we denote by $\overline{F}((\cdot, x, v)$ the survival function of T_1 with initial data $(X_0, V_0) = (x, v)$: $\overline{F}(\cdot, x, v) = \mathbb{P}_{x,v}(T_1 > \cdot)$.
The time T_1 satisfies, when $(X_0, V_0) = (x, v)$:

$$T_1 = \inf\left\{t \geq 0, \int_0^t \lambda(X_s, V_s)\mathrm{d}s \geq E\right\} = \inf\left\{t \geq 0, \int_0^t \lambda(x + vs, v)\mathrm{d}s \geq E\right\}$$

where E is an exponential variable with unit mean, because the process is deterministic between jump times. We then get :

$$\begin{aligned}
\overline{F}(t, x, v) &= \mathbb{P}_{x,v}(T_1 > t) \\
&= \mathbb{P}_{x,v}\left(\int_0^t \lambda(x + vs, v)\mathrm{d}s \leq E\right) \\
&= \exp\left(-\int_0^t \lambda(x + vs, v)\mathrm{d}s\right).
\end{aligned}$$

The conditional distribution of the inter-jump times is then given by, for all $n \geq 0$:

$$\mathbb{P}\left(\tau_{n+1} \geq t \,|X_{T_n}, V_n\right) = \exp\left(-\int_0^t \lambda(X_{T_n} + V_n s, V_n)\mathrm{d}s\right).$$

3.3 Main result

In this section we prove our main result, Theorem 3.1.2, using Theorem 3.2.1. We first need to find a Lyapunov function for the process.

3.3.1 A Lyapunov function

Let us introduce some constants which will appear in the definition of our Lyapunov function. Thanks to Assumptions $(\mathcal{H}_1)$, $(\mathcal{H}_2)$, $(\mathcal{H}_3)$ and $(\mathcal{H}_4)$, the following quantities are well defined :

$$\begin{aligned}
\theta_1 &\in \left(\max\left\{\frac{\theta_* \lambda_{\max}}{p\theta_0\lambda_{\min}}; \frac{1}{p\theta_0\beta}\right\} \,,\, \min\left\{p\theta_0, \bar{\theta}\right\}\right); \\
\alpha &\in \left(0 \,,\, \min\left\{\frac{p\theta_0\lambda_{\min}}{\theta_*} - \frac{\lambda_{\max}}{\theta_1}; p\theta_0\beta\lambda_{\max} - \frac{\lambda_{\max}}{\theta_1}\right\}\right); \\
a &\in \left(1 + \frac{\lambda_{\max}}{\alpha\theta_1} \,,\, \min\left\{\frac{p\theta_0\lambda_{\min}}{\alpha\theta_*}; \frac{p\theta_0\beta\lambda_{\max}}{\alpha}\right\}\right).
\end{aligned}$$

We then consider the function H defined on $E = \mathbb{R}^d \times \mathcal{B}(1)$ by

$$H(x, v) = \mathrm{e}^{\alpha|x|}\left(a + \varphi\left(\frac{x \cdot v}{|x|}\right)\right),$$

where φ is a non-decreasing function of class C^1 on $[-1,1]$, with $\varphi(\theta) = \theta$ if $\theta \in [-1, -\theta_1]$, and $\varphi(\theta) = 0$ if $\theta \in [0,1]$. We introduce m the supremum norm of the derivative of φ, that is :

$$m = \sup_{\theta \in [-1,1]} |\varphi^{'}(\theta)|.$$

Then, we have the following result :

Proposition 3.3.1. *There exist some constants $R > 0$, $\eta > 0$ such that for all $(x,v) \in (\mathbb{R}^d \backslash \mathcal{B}(R)) \times \mathcal{B}(1)$,*

$$LH(x,v) \leq -\eta H(x,v).$$

The function H is then a Lyapunov function associated to the set $\mathcal{B}(R) \times \mathcal{B}(1)$ for the process (X,V) driven by the generator L.

Remark 3.3.2. *This function H can be compared to the one of [Bierkens et Roberts, 2017, Fontbona* et al.*, 2016] in dimension 1. It is also qualitatively equivalent to the ones of [Bierkens* et al.*, 2017, Durmus* et al.*, 2018, Deligiannidis* et al.*, 2019], but here it is written in terms of the jump rate λ, while in the papers cited before, they give a Lyapunov function depending on their potential U.*

Proof. We can check that H is in the domain of the generator L. For $(x,v) \in E$ we have :

$$LH(x,v) = \mathrm{e}^{\alpha|x|}\left(A_1 + A_2 + A_3\right) \tag{3.7}$$

with

$$A_1 = \alpha \frac{x \cdot v}{|x|}\left(a + \varphi\left(\frac{x \cdot v}{|x|}\right)\right)$$
$$A_2 = \varphi^{'}\left(\frac{x \cdot v}{|x|}\right)\frac{1}{|x|}\left(|v|^2 - \left(\frac{x \cdot v}{|x|}\right)^2\right)$$
$$A_3 = \lambda(x,v)\int_{\mathcal{B}(1)}\left(\varphi\left(\frac{x \cdot v'}{|x|}\right) - \varphi\left(\frac{x \cdot v}{|x|}\right)\right)Q(x,v,\mathrm{d}v').$$

Let first remark that since the derivative of φ is bounded by m, we always have

$$A_2 \leq \frac{m}{|x|}.$$

Moreover, thanks to Assumption $(\mathcal{H}_3)$ and the definition of the function φ, if x and v are such that $\frac{x \cdot v}{|x|} > -\bar{\theta}$, we can bound A_3 as follows :

$$A_3 \leq \lambda(x,v)\left[\int_{\{v' \in \mathcal{B}(1)/\frac{x \cdot v'}{|x|} \leq -\theta_0\}} \varphi\left(\frac{x \cdot v'}{|x|}\right)Q(x,v,\mathrm{d}v') - \varphi\left(\frac{x \cdot v}{|x|}\right)\right]$$

$$\leq \lambda(x,v)\left[-\theta_0 \int_{\{v'\in\mathcal{B}(1)/\frac{x\cdot v'}{|x|}\leq -\theta_0\}} Q(x,v,\mathrm{d}v') - \varphi\left(\frac{x\cdot v}{|x|}\right)\right]$$
$$\leq -\lambda(x,v)\left[p\theta_0 + \varphi\left(\frac{x\cdot v}{|x|}\right)\right].$$

We can now study Equation (3.7) depending on the different values taken by $\frac{x\cdot v}{|x|}$.

- If $\frac{x\cdot v}{|x|} \in [-1,-\theta_1]$:

$$LH(x,v) \leq \mathrm{e}^{\alpha|x|}\left[-\alpha\theta_1(a-1) + \frac{m}{|x|} + \lambda_{\max}\right].$$

- If $\frac{x\cdot v}{|x|} \in (-\theta_1, 0)$:

$$LH(x,v) \leq \mathrm{e}^{\alpha|x|}\left[\frac{m}{|x|} - \lambda(x,v)\left(p\theta_0 - \theta_1\right)\right]$$
$$\leq \mathrm{e}^{\alpha|x|}\left[\frac{m}{|x|} - \lambda_{\min}\left(p\theta_0 - \theta_1\right)\right].$$

- If $\frac{x\cdot v}{|x|} \in [0, \theta_*]$:

$$LH(x,v) \leq \mathrm{e}^{\alpha|x|}\left[\alpha\theta_* a + \frac{m}{|x|} - \lambda_{\min} p\theta_0\right].$$

- If $\frac{x\cdot v}{|x|} \in (\theta_*, 1]$, and if $|x| \geq \Delta$, with Δ defined in Assumption $(\mathcal{H}_4)$:

$$LH(x,v) \leq \mathrm{e}^{\alpha|x|}\left[\alpha a + \frac{m}{|x|} - p\theta_0\beta\lambda_{\max}\right].$$

Let us now take $R > r$ where

$$r = \max\left\{\frac{m}{\alpha\theta_1(a-1)-\lambda_{\max}}; \frac{m}{\lambda_{\min}\left(p\theta_0-\theta_1\right)}; \frac{m}{p\theta_0\lambda_{\min}-\alpha\theta_* a}; \Delta; \frac{m}{p\theta_0\beta\lambda_{\max}-\alpha a}\right\}.$$

Thanks to the assumptions made on each parameter, the constant r is well defined and finite. Therefore, the previous calculations give the existence of a constant $\eta > 0$ such that for all $(x,v) \in (\mathbb{R}^d\backslash\mathcal{B}(R)) \times \mathcal{B}(1)$:

$$LH(x,v) \leq -\eta \mathrm{e}^{\alpha|x|}.$$

Finally, the function $(y,w) \longmapsto a + \varphi\left(\frac{y\cdot w}{|y|}\right)$ being bounded from above by a, we get :

$$LH(x,v) \leq -\eta a H(x,v),$$

which ends the proof. □

3.3.2 Proof of Theorem 3.1.2

Since we have already found a Lyapunov function, we still have to review three points in order to use Theorem 3.2.1 : the irreducibility and the aperiodicity of the process, and the existence of petite sets. These are the object of the following proposition.

Proposition 3.3.3. *All compact sets of the form $K \times \mathcal{B}(1)$, with K a compact set of $\mathbb{R}^d$, are petite for the process (X, V).*
Moreover, the process is irreducible and aperiodic.

Proof. We do not detail the proof since we just have follow the ideas of the proof of Lemma 2 in [Deligiannidis *et al.*, 2019]. It is indeed enough to show that for all $M > 0$, for all $(x, v) \in \mathcal{B}(M) \times \mathcal{B}(1)$, for all positive bounded function $\varphi : E \to \mathbb{R}$, there exist a constant $C > 0$ and a compact set $A \subset \mathbb{R}^d \times \mathcal{B}(1)$, both independent of M and φ, such that :

$$\int_0^\infty \mathrm{e}^{-t}\mathbb{E}_{(x,v)}\left[\varphi(X_t, V_t)\right] \mathrm{d}t \geq C \int\int_A \varphi(y, w)\mathrm{d}y\mathrm{d}w.$$

to deduce the first result of the proposition. Let us mention that the proof is based only on Assumptions $(\mathcal{H}_1)$, $(\mathcal{H}_2)$ and $(\mathcal{H}_3)$.
Then, the irreducibility of the process follows easily, and we refer once more to [Deligiannidis *et al.*, 2019], Lemma 3, for the aperiodicity. □

Remark 3.3.4. *Let us mention different proofs of this result in some particular cases.*
In dimension 1, a proof has been given in [Bierkens et Roberts, 2017] for the process with a velocity equal to -1 or $+1$, with the same main ideas that in [Deligiannidis et al.*, 2019] and [Monmarché, 2016].*
For the process in dimension 2, if the jump kernel has a density with respect to the Lebesgue measure which is bounded from below by a strictly positive constant, and if the jump rate is bounded, the previous proposition can be proved by geometric considerations, as in the proof of Lemma 4.5 in [Champagnat et Villemonais, 2017].

Proof of Theorem 3.1.2. By Proposition 3.3.1 and 3.3.3, all conditions of 3.2.1 are satisfied, so that the stated result follows. □

3.4 Exponential moments for the invariant measure

In this section, we want to study the existence of exponential moments for the unique invariant probability measure of the process (X, V), that we denote by π, whose existence follows from Theorem 3.1.2.

Let first say that if the lengths of the consecutive excursions away from the origin were independent and identically distributed, the process (X, V) would

be a regenerative process, and we could then apply standard results on regenerative processes (see [Asmussen, 2003] or [Cocozza-Thivent, 2018] for instance) to collect some information on the invariant measure of the process. This is for instance what is done in [Fontbona *et al.*, 2012], since the simple Zig-zag process is a regenerative process when we look at it at the successive hitting-times of the origin. However, the excursions away from the origin of the generalized Zig-zag process do not satisfy this assumptions of independence and identical distribution because of the dependence in speed between two consecutive excursions.
Nevertheless, the generalized Zig-zag process is a semi-regenerative process (see [Asmussen, 2003] or [Cocozza-Thivent, 2018] for the definition). Indeed, let $Z_0 = 0$ and $(Z_n)_{n\geq 1}$ be the successive times at which the process X enters the ball $\mathcal{B}(R)$ (with R defined in Proposition 3.3.1). The PDMP (X, V) is a semi-regenerative process associated to the renewal process $((X_{Z_n}, V_{Z_n}), Z_n)_{n\geq 0}$.

Before stating the result on regenerative process that we will use in this part, let us first briefly speak about Harris-recurrence and non-arithmetic process (see for instance [Alsmeyer, 1997, Asmussen, 2003, Cocozza-Thivent, 2018]). The Markov chain Y with state space F is called Harris-recurrent if there exists a measurable set $A \subset F$, $c > 0$, $m \geq 1$, and a distribution φ such that

1. for all $y \in F$, $\mathbb{P}_y(\tau_A < \infty) = 1$, where $\tau_A = \inf\{n \geq 0, Y_n \in A\}$;
2. for all $y \in A$, $\mathbb{P}_y(Y_m \in \cdot) \geq c\varphi(\cdot)$.

Then, we recall that the process $(Y_n, T_n)_{n\geq 0}$ is non-arithmetic if the laws of the variables $T_n - T_{n-1}$, $n \geq 1$ have a part which is absolutely continuous with respect to the Lebesgue measure (see [Cocozza-Thivent, 2018] for the definition of an arithmetic process).

We have the following result concerning the convergence of semi-regenerative process, that we can found in [Alsmeyer, 1997, Asmussen, 2003] or [Cocozza-Thivent, 2018] :

Theorem 3.4.1. *Let $(\Psi_t)_{t\geq 0}$ be a semi-regenerative process associated to the renewal process $(Y, T) = (Y_n, T_n)_{n\geq 0}$.*
We suppose that (Y, T) is non-arithmetic and that Y is Harris-recurrent, and let ν be an invariant measure for this chain.
Let $f : F \longrightarrow \mathbb{R}$ be a measurable positive function such that $(z, t) \longmapsto \mathbb{E}[f(\Psi_t)|Y_0 = z]$ is bounded on $F \times [0, t]$ for all t. We define $g : F \times \mathbb{R}_+ \longrightarrow \mathbb{R}$ by $g(z, t) = \mathbb{E}[f(\Psi_t)\mathbf{1}_{t<T_1}|Y_0 = z]$.
We suppose :

1. *for ν-almost all $z \in F$, $t \mapsto g(z, t)$ is almost everywhere continuous with respect to the Lebesgue measure ;*
2. *there exists $\Delta > 0$ such that*

$$\int_F \sum_{n\in\mathbb{N}} \sup_{n\Delta\leq t<(n+1)\Delta} |g(z, t)|\nu(dz) < \infty.$$

Then for ν-almost all $y \in F$:

$$\mathbb{E}\left[f(\Psi_t)|Y_0 = y\right] \underset{t\to\infty}{\longrightarrow} \frac{\int_F \mathbb{E}\left[\int_0^{T_1} f(\Psi_s)ds|Y_0 = z\right]\nu(dz)}{\int_F \mathbb{E}\left[T_1|Y_0 = z\right]\nu(dz)}. \tag{3.8}$$

Consequently, the previous theorem can be applied to our process, and it implies the existence of exponential moments for the invariant measure π of the PDMP (X, V). More precisely we have the following result :

Theorem 3.4.2. *Let $(X_t, V_t)_{t\geq 0}$ be the generalized Zig-zag process satisfying Assumptions $(\mathcal{H}_1)$, $(\mathcal{H}_2)$, $(\mathcal{H}_3)$ and $(\mathcal{H}_4)$.*
Let η be as in Proposition 3.3.1. Then for all $0 < \beta < \eta$ and all $\gamma > 0$, we have :

$$\int_E \mathrm{e}^{\beta|x|+\gamma|v|}\pi(dx, dv) < \infty,$$

where we recall that π is the unique invariant probability measure of the process (X, V).

Remark 3.4.3. *This result ensures that at the equilibrium, the process is concentrated around the origin. From a modelling point of view, it means that in long time, the bacterium evolves around the chemo-attractant and does not go too far from it.*

This theorem is a consequence of Proposition 3.3.1, which implies the existence of exponential moments for the hitting time of the compact set $\mathcal{B}(R)$ associated to the Lyapunov function. Let us precise this fact :

Proposition 3.4.4. *Let us note $\tau_{\mathcal{B}(R)} = \inf\{t \geq 0, X_t \in \mathcal{B}(R)\}$ the hitting time of $\mathcal{B}(R)$. For all $(x, v) \in E$,*

$$\mathbb{E}_{x,v}\left[\mathrm{e}^{\eta\tau_{\mathcal{B}(R)}}\right] \leq H(x, v),$$

where η and H are defined in Proposition 3.3.1.

Proof. In order to make the proof easier to read, we note τ instead of $\tau_{\mathcal{B}(R)}$.
For $(x, v) \in E$, Dynkin's formula gives :

$$\begin{aligned}
\mathbb{E}_{x,v}\left[\mathrm{e}^{\eta(t\wedge\tau)}\right] &\leq \mathbb{E}_{x,v}\left[H\left(X_{t\wedge\tau}, V_{t\wedge\tau}\right)\mathrm{e}^{\eta(t\wedge\tau)}\right] \\
&= H(x, v) + \mathbb{E}_{x,v}\left[\int_0^{t\wedge\tau} (\eta + L)H\left(X_{s\wedge\tau}, V_{s\wedge\tau}\right)\mathrm{e}^{\eta(s\wedge\tau)}\mathrm{d}s\right] \\
&\leq H(x, v),
\end{aligned}$$

the last inequality resulting from Proposition 3.3.1.
Then, when t goes to infinity, the monotone convergence theorem gives :

$$\mathbb{E}_{x,v}\left[\mathrm{e}^{\eta\tau}\right] \leq H(x, v).$$

□

Proof of Theorem 3.4.2. Let first remark that the chain $(X_{Z_n}, V_{Z_n})_{n\geq 0}$ is Harris-recurrent. Indeed, let $A = \mathcal{S}_R^{d-1}\times\mathcal{B}(1)$ be the state space of $(X_{Z_n}, V_{Z_n})_{n\geq 0}$ ($\mathcal{S}_R^{d-1}$ denotes the sphere of $\mathbb{R}^d$ with radius R). We obviously have, for all $(x,v)\in A$, $\mathbb{P}(\tau_A<\infty)=1$. The second point in the definition of the Harris-recurrence can be proved with φ the Lebesgue measure on $\mathbb{R}^d$, using in particular the fact that Z_m is almost-surely finite for each $m\geq 1$ thanks to Proposition 3.4.4.
Moreover, the chain $((X_{Z_n}, V_{Z_n}), Z_n)_{n\geq 0}$ is non-arithmetic because the law of the times $Z_{n+1}-Z_n$ has a part which is absolutely continuous with respect to the Lebesgue measure.
Let us now introduce $f : (x,v)\in\mathbb{R}^d\times\mathcal{B}(1)\longmapsto \mathrm{e}^{\beta|x|+\gamma|v|}$ for $0\leq\beta<\eta$ and $\gamma\geq 0$.
We first observe that f is a positive measurable function, and that

$$((x,v),t)\longrightarrow \mathbb{E}\left[f(X_t,V_t)|(X_{Z_n},V_{Z_n})=(x,v)\right]$$

is bounded on $\left(\mathcal{S}_R^{d-1}\times\mathcal{B}(1)\right)\times[0,t]$ for all t.
Let ν be the unique invariant probability measure of the chain $(X_{Z_n}, V_{Z_n})_{n\geq 0}$, which exists since the chain is positive Harris-recurrent.
Let us define the function g on $\left(\mathcal{S}_R^{d-1}\times\mathcal{B}(1)\right)\times\mathbb{R}_+$ by

$$g\left((x,v),t\right)=\mathbb{E}_{x,v}\left[\mathrm{e}^{\beta|X_t|+\gamma|V_t|}\mathbf{1}_{t<Z_1}\right],$$

and let us check if g satisfies the assumption of Theorem 3.4.1 :

1. For ν-almost all $(x,v)\in\mathcal{S}_R^{d-1}\times\mathcal{B}(1)$, the function $t\longmapsto g((x,v),t)$ is almost everywhere continuous with respect to the Lebesgue measure since it is right-continuous and has thus an at most countable set of discontinuities.
2. Let $\Delta>0$. According to Markov inequality we have

$$\mathbb{P}_{x,v}\left(Z_1>t\right)\leq \mathrm{e}^{-\eta t}\mathbb{E}_{x,v}\left[\mathrm{e}^{\eta Z_1}\right].$$

Using Proposition 3.4.4 and the fact that the Lyapunov function H is uniformly bounded on $\mathcal{S}_R^{d-1}\times\mathcal{B}(1)$, we get the existence of a finite constant M such that for all $(x,v)\in\mathcal{S}_R^{d-1}\times\mathcal{B}(1)$,

$$\mathbb{P}_{x,v}\left(Z_1>t\right)\leq M\mathrm{e}^{-\eta t}.$$

Moreover, let remark that if $(X_0,V_0)\in\mathcal{S}_R^{d-1}\times\mathcal{B}(1)$, since $|X_0|=R$ and the velocity of the process lives in $\mathcal{B}(1)$, then for all $t\geq 0$, $|X_t|\leq R+t$. Therefore we have

$$\begin{aligned}&\int_{\mathcal{S}^{d-1}\times\mathcal{B}(1)}\sum_{n\in\mathbb{N}}\sup_{n\Delta\leq t<(n+1)\Delta}|g\left((x,v),t\right)|\nu(\mathrm{d}x,\mathrm{d}v)\\ &=\int_{\mathcal{S}^{d-1}\times\mathcal{B}(1)}\sum_{n\in\mathbb{N}}\sup_{n\Delta\leq t<(n+1)\Delta}\mathbb{E}_{x,v}\left[\mathrm{e}^{\beta|X_t|+\gamma|V_t|}\mathbf{1}_{t<Z_1}\right]\nu(\mathrm{d}x,\mathrm{d}v)\end{aligned}$$

$$\le \int_{\mathcal{S}^{d-1}\times\mathcal{B}(1)} \sum_{n\in\mathbb{N}} \sup_{n\Delta\le t<(n+1)\Delta} \mathrm{e}^{\beta(R+t)+\gamma}\mathbb{P}_{x,v}\left(Z_1>t\right)\nu(\mathrm{d}x,\mathrm{d}v)$$
$$\le \int_{\mathcal{S}^{d-1}\times\mathcal{B}(1)} \sum_{n\in\mathbb{N}} \sup_{n\Delta\le t<(n+1)\Delta} \mathrm{e}^{\beta(R+t)+\gamma}M\mathrm{e}^{-\eta t}\nu(\mathrm{d}x,\mathrm{d}v)$$
$$\le M\mathrm{e}^{\beta R+\gamma}\sum_{n\in\mathbb{N}}\mathrm{e}^{\beta(n+1)\Delta-\eta n\Delta}$$
$$= M\mathrm{e}^{\beta(R+\Delta)+\gamma}\sum_{n\in\mathbb{N}}\left(\mathrm{e}^{\Delta(\beta-\eta)}\right)^n.$$

This quantity is finite since $\beta<\eta$.

The function g satisfying all the required assumptions, we can apply Theorem 3.4.1 : for ν-almost all $(x_0,v_0)\in\mathcal{S}_R^{d-1}\times\mathcal{B}(1)$ we have

$$\mathbb{E}_{x_0,v_0}\left[\mathrm{e}^{\beta|X_t|+\gamma|V_t|}\right]\underset{t\to\infty}{\longrightarrow}\frac{\int_{\mathcal{S}_R^{d-1}\times\mathcal{B}(1)}\mathbb{E}_{x,v}\left[\int_0^{Z_1}\mathrm{e}^{\beta|X_s|+\gamma|V_s|}\mathrm{d}s\right]\nu(\mathrm{d}x,\mathrm{d}v)}{\int_{\mathcal{S}_R^{d-1}\times\mathcal{B}(1)}\mathbb{E}_{x,v}\left[Z_1\right]\nu(\mathrm{d}x,\mathrm{d}v)}.$$

The quantity on the right is clearly finite thanks to the previous calculations. Moreover, the ergodic theorem gives :

$$\mathbb{E}_{x_0,v_0}\left[\mathrm{e}^{\beta|X_t|+\gamma|V_t|}\right]\underset{t\to\infty}{\longrightarrow}\int_{\mathbb{R}^d\times\mathcal{B}(1)}\mathrm{e}^{\beta|x|+\gamma|v|}\pi(\mathrm{d}x,\mathrm{d}v).$$

We then deduce that

$$\int_{\mathbb{R}^d\times\mathcal{B}(1)}\mathrm{e}^{\beta|x|+\gamma|v|}\pi(\mathrm{d}x,\mathrm{d}v)=\frac{\int_{\mathcal{S}_R^{d-1}\times\mathcal{B}(1)}\mathbb{E}_{x,v}\left[\int_0^{Z_1}\mathrm{e}^{\beta|X_s|+\gamma|V_s|}\mathrm{d}s\right]\nu(\mathrm{d}x,\mathrm{d}v)}{\int_{\mathcal{S}_R^{d-1}\times\mathcal{B}(1)}\mathbb{E}_{x,v}\left[Z_1\right]\nu(\mathrm{d}x,\mathrm{d}v)},$$

and we have proved the theorem. □

3.5 The particular case of dimension 1

In this section, we study PDMPs in dimension 1, but in a different way, and under some different assumptions (see Section 3.5.4 for a comparison between the two approaches in dimension 1). This section is thus complementary to the previous study applied at the one-dimension.

We consider here the PDMP $((X_t,V_t))_{t\ge0}$ with values in $\mathbb{R}\times[-1,1]$ with infinitesimal generator given by

$$Lh(x,v)=v\partial_x h(x,v)+\lambda(x,v)\int_{-1}^{1}\left(h(x,v')-h(x,v)\right)Q(x,v,\mathrm{d}v'),\qquad(3.9)$$

where $Q(x,v,\cdot)$ is a probability kernel on $[-1,1]$. This process is the one-dimensional version of the process driven by (3.1).

The framework of this section is the following :

$(\mathcal{A}_1)$: $Q(x, v, \mathrm{d}v') = q(v, v')(\nu(\mathrm{d}v') + \mu(\mathrm{d}v'))$ with ν a discrete measure and μ a restriction of the Lebesgue measure. We denote by $\mathcal{V}$ the support of Q and suppose that there exists a constant $q_{\min} > 0$ such that $q(v, v') \geq q_{\min}$ for all $v, v' \in \mathcal{V}$;

$(\mathcal{A}_2)$: The process is symmetric : $\mathcal{V}$ is symmetric with respect to 0 and $\lambda(x, v) = \lambda(-x, -v)$ for all $(x, v) \in \mathbb{R} \times \mathcal{V}$;

$(\mathcal{A}_3)$: There exists $0 < \lambda_{\min}$ such that for all $(x, v) \in \mathbb{R} \times \mathcal{V}$, $\lambda_{\min} \leq \lambda(x, v)$, and the quantity $\sup_{x \geq 0, v \leq 0} \lambda(x, v)$ is finite.

The importance of the existence of q_{min} in Assumption $(\mathcal{A}_1)$ will not explicitly appear in the following since we are not going to give all the proofs in detail, but we note that it is useful in Proposition 3.5.8 and Theorem 3.5.9.
Assumption $(\mathcal{A}_2)$ implies that the process is spatially symmetric with respect to the origin, which will allow us to reduce the number of computations. Nevertheless, the following results are still available without this symmetry.
The fact that we suppose the kernel Q to be independent of the position of the bacterium implies that the chain composed by the velocities at the jump times is a Markov chain with kernel Q. In particular, contrary to what is assumed in higher dimension, the attraction of the bacterium by the origin is not favored by the jump kernel since it does not take care of the position of the bacterium with respect to the origin.
We make thus an additional assumption $(\mathcal{A}_4)$ that takes into account this attraction to the origin, because of the presence of a chemo-attractant there. This assumption is the one that makes the process ergodic.

$$(\mathcal{A}_4) : \exists I_* \subset \left[0, \inf_{x \geq 0, v \in (0,1]} \frac{\lambda(x, v)}{v}\right) \text{ an interval }, \exists 0 < J_* < 1,$$
$$\forall \alpha \in I_*, \forall v' \in \mathcal{V}, \int_{-1}^{1} G(\alpha, v) Q(v', \mathrm{d}v) \leq J_*,$$

where

$$G(\alpha, v) = \frac{\sup_{x \geq 0} \lambda(x, v)}{\sup_{x \geq 0} \lambda(x, v) - \alpha v} \mathbf{1}_{v<0} + \frac{\inf_{x \geq 0} \lambda(x, v)}{\inf_{x \geq 0} \lambda(x, v) - \alpha v} \mathbf{1}_{v \geq 0} \tag{3.10}$$

for $\alpha \geq 0$ and $v \in \mathcal{V}$.

Before stating the theorem that gives the exponential ergodicity of the process, we make some remarks on Assumption $(\mathcal{A}_4)$.

Remark 3.5.1. *Let us give a sufficient condition to satisfy Assumption $(\mathcal{A}_4)$ when the jump kernel Q does not depend on the previous velocity.*
Let us write $J_{v'}(\alpha) = \int_{-1}^{1} G(\alpha, v) Q(v', \mathrm{d}v)$.

For all $v' \in [-1,1]$, the function $J_{v'}$ is C^1 on $[0, \inf_{x\geq 0, v\in]0,1]} \frac{\lambda(x,v)}{v})$, is convex and

$$J_{v'}(0) = 1 \ , \ J'_{v'}(0) = \int_{-1}^{1} \left(\frac{v}{\sup_{x\geq 0}\lambda(x,v)} \mathbf{1}_{v<0} + \frac{v}{\inf_{x\geq 0}\lambda(x,v)} \mathbf{1}_{v\geq 0} \right) Q(v', \mathrm{d}v)$$

$$\textit{and} \quad \lim_{\alpha \to \inf_{x\geq 0, v\in]0,1]} \frac{\lambda(x,v)}{v}} J_{v'}(\alpha) = +\infty.$$

If we assume that $J'_{v'}(0) < 0$, we get that there exists a unique $\hat{\alpha} \in (0, \inf_{x\geq 0, v\in]0,1]} \frac{\lambda(x,v)}{v})$ such that $J(\hat{\alpha}) = 1$, and then there exists an interval $\hat{I} \subset (0, \alpha)$ such that for all $\alpha \in \hat{I}$, $J_{v'}(\alpha)$ is uniformly bounded by a constant strictly smaller than 1 on the interval $\hat{I}$.
Consequently, for all $v' \in [-1,1]$,

$$\exists \hat{I} \subset (0, \inf_{x\geq 0, v\in(0,1]} \frac{\lambda(x,v)}{v}), \exists 0 < \hat{K} < 1, \forall \alpha \in \hat{I}, \int_{-1}^{1} G(\alpha, v) Q(v', \mathrm{d}v) \leq \hat{K}.$$

In Assumption $(\mathcal{A}_4)$, we suppose in addition that the interval $\hat{I}$ and the constant $\hat{K}$ are uniform in v'. In particular, if $Q(v', \mathrm{d}v) = Q(\mathrm{d}v)$, the assumption $J'(0) < 0$, that is

$$\int_{-1}^{1} \left(\frac{v}{\sup_{x\geq 0}\lambda(x,v)} \mathbf{1}_{v<0} + \frac{v}{\inf_{x\geq 0}\lambda(x,v)} \mathbf{1}_{v\geq 0} \right) Q(\mathrm{d}v) < 0$$

is sufficient for Assumption $(\mathcal{A}_4)$ to be satisfied.
This assumption is the equivalent of Assumption $(H3)$ made by Calvez, Raoul and Schmeiser in [Calvez et al., 2015]. The probability study carried out in this section covers thus the framework of [Calvez et al., 2015].

Remark 3.5.2. *Let see that in the case where the kernel Q does not depend on the previous velocity and is symmetric in the sense that $q(v) = q(-v)$ for all $v \in \mathcal{V}$, Assumption $(\mathcal{A}_4)$ holds in the particular case where $\inf_{x\geq 0}\lambda(x,v) > \sup_{x\leq 0}\lambda(x,v)$ for all $v \in [0,1]$, i.e. when the velocity tends to jump even more when the bacterium goes away from 0 than when it moves towards the origin. This case is the one considered for instance in [Bierkens et Roberts, 2017], [Fontbona et al., 2012] and [Fontbona et al., 2016], with a velocity taking its values in $\{-1,+1\}$*
Let us prove this fact. Under this assumption on the jump rate, with the same notations as in the previous remark, and using the symmetries of the process we have :

$$J'_{v'}(0) = \int_{-1}^{1} \left(\frac{v}{\sup_{x\geq 0}\lambda(x,v)} \mathbf{1}_{v<0} + \frac{v}{\inf_{x\geq 0}\lambda(x,v)} \mathbf{1}_{v\geq 0} \right) q(v)\mathrm{d}v$$

$$= \int_{-1}^{0} \frac{v}{\sup\limits_{x\geq 0}\lambda(x,v)} q(v)\mathrm{d}v + \int_{0}^{1} \frac{v}{\inf\limits_{x\geq 0}\lambda(x,v)} q(v)\mathrm{d}v$$
$$= -\int_{0}^{1} \frac{v}{\sup\limits_{x\geq 0}\lambda(x,-v)} q(v)\mathrm{d}v + \int_{0}^{1} \frac{v}{\inf\limits_{x\geq 0}\lambda(x,v)} q(v)\mathrm{d}v$$
$$= -\int_{0}^{1} \frac{v}{\sup\limits_{x\leq 0}\lambda(x,v)} q(v)\mathrm{d}v + \int_{0}^{1} \frac{v}{\inf\limits_{x\geq 0}\lambda(x,v)} q(v)\mathrm{d}v$$
$$< 0.$$

And the end of Remark 3.5.1 ensures that Assumption $(\mathcal{A}_4)$ is satisfied.

Let us now give the main theorem of this section.

Theorem 3.5.3. *Let $(X_t, V_t)_{t\geq 0}$ be the PDMP on $\mathbb{R}\times[-1,1]$ whose infinitesimal generator is given by*

$$Lh(x,v) = v\partial_x h(x,v) + \lambda(x,v)\int_{-1}^{1}\big(h(x,v') - h(x,v)\big)\,Q(x,v,\mathrm{d}v').$$

Under Assumptions $(\mathcal{A}_1)$, $(\mathcal{A}_2)$, $(\mathcal{A}_3)$ and $(\mathcal{A}_4)$, the process is exponentially ergodic.

3.5.1 The hitting time of the origin

As mentioned before, we are going to estimate the exponential moments of the hitting times of compact sets in order to prove Theorem 3.5.3. We introduce two new notations.
We denote by Z the first hitting time of 0, i.e.

$$Z = \inf\{t > 0, X_t = 0\}, \tag{3.11}$$

and S is the index of the first jump-time at which the position of the process has changed its sign : $S = \inf\{n \geq 1, X_{T_n}X_0 \leq 0\}$.
Our goal in this section is to get information on the exponential moments of Z. For this purpose, we will first study the random variable S. Then, the inequality $Z \leq \sum_{i=1}^{S}\tau_i$ a.s. will allow us to come back to Z.

Proposition 3.5.4. *For all $(x_0, v_0) \in \mathbb{R}\times\mathcal{V}$ and all α such that $|\alpha|\in I_*$, we have, for all $n \geq 1$,*

$$\mathbb{P}_{x_0,v_0}(S > n) \leq \kappa_{\alpha,x_0,v_0} {J_*}^{n-1},$$

with

$$\kappa_{\alpha,x_0,v_0} = \mathrm{e}^{|\alpha x_0|} G(|\alpha|, v_0),$$

where G and J_* are defined by Assumption $(\mathcal{A}_4)$.

Proof. In order to make the proof easier to read, we distinguish the cases where x_0 and v_0 are positive or negative. We first look at the case $x_0 > 0$ and $v_0 \in \mathcal{V} \cap [0,1]$, and the other cases are similar because of the symmetry of the process.

Let $x_0 > 0$ and $v_0 \in \mathcal{V} \cap [0,1]$.
We have $S = \inf\{n \geq 1, X_{T_n} \leq 0\}$, and $X_{T_n} = X_0 + \sum_{i=0}^{n-1} V_i \tau_{i+1} = x_0 + v_0\tau_1 + \sum_{i=1}^{n-1} V_i \tau_{i+1}$.
Let $\alpha \in [0, \inf_{x \geq 0, v \in (0,1]} \frac{\lambda(x,v)}{v})$.
Since x_0 is positive, on the event $\{S > n\}$, X_{T_n} is also positive, and the sign of α implies :

$$\begin{aligned}
\mathbb{P}_{x_0,v_0}(S > n) &\leq \mathbb{E}_{x_0,v_0}\left[\mathrm{e}^{\alpha X_{T_n}} \mathbf{1}_{S>n}\right] \\
&= \mathrm{e}^{\alpha x_0} \mathbb{E}_{x_0,v_0}\left[\mathrm{e}^{\alpha\left(v_0\tau_1 + \sum_{i=1}^{n-1} V_i\tau_{i+1}\right)} \mathbf{1}_{S>n}\right] \\
&= \mathrm{e}^{\alpha x_0} \mathbb{E}_{x_0,v_0}\left[\mathrm{e}^{\alpha v_0 \tau_1} \prod_{i=1}^{n-1} \mathrm{e}^{V_i\tau_{i+1}} \mathbf{1}_{S>n}\right] \\
&= \mathrm{e}^{\alpha x_0} \mathbb{E}_{x_0,v_0}\left[\mathbb{E}_{x_0,v_0}\left[\mathrm{e}^{\alpha v_0 \tau_1} \prod_{i=1}^{n-1} \mathrm{e}^{V_i\tau_{i+1}} \mathbf{1}_{S>n} | X_{T_1}, V_1, \ldots, X_{T_{n-1}}, V_{n-1}\right]\right].
\end{aligned}$$

Since the inter-jump times $\tau_1, \ldots, \tau_n$ are independent conditionally to the couples $(X_0, V_0), (X_{T_1}, V_1), \ldots, (X_{T_{n-1}}, V_{n-1})$, and since $\{S > n\} = \{X_{T_1} > 0, \cdots, X_{T_n} > 0\}$ we have :

$$\begin{aligned}
&\mathbb{P}_{x_0,v_0}(S > n) \\
&\leq \mathrm{e}^{\alpha x_0} \mathbb{E}_{x_0,v_0}\left[\mathbb{E}_{x_0,v_0}\left[\mathrm{e}^{\alpha v_0\tau_1}\right] \prod_{i=1}^{n-1} \mathbb{E}_{x_0,v_0}\left[\mathrm{e}^{\alpha V_i \tau_{i+1}} | X_{T_i}, V_i\right] \mathbf{1}_{X_{T_1}>0,\cdots,X_{T_n}>0}\right].
\end{aligned} \tag{3.12}$$

Moreover, we know that

$$\mathbb{P}\left(\tau_{i+1} \geq t | X_{T_i}, V_i\right) = \overline{F}(t, X_{T_i}, V_i),$$

which gives, for $1 \leq i \leq n-1$:

$$\begin{aligned}
&\mathbb{E}\left[\mathrm{e}^{\alpha V_i \tau_{i+1}} | X_{T_i}, V_i\right] \\
&= \int_0^{+\infty} \mathrm{e}^{\alpha V_i t} \lambda(X_{T_i} + V_i t, V_i) \mathrm{e}^{-\int_0^t \lambda(X_{T_i} + V_i s, V_i)\mathrm{d}s} \mathrm{d}t \\
&= \int_0^{+\infty} \frac{\lambda(X_{T_i} + V_i t, V_i)}{\lambda(X_{T_i} + V_i t, V_i) - \alpha V_i} \left(\lambda(X_{T_i} + V_i t, V_i) - \alpha V_i\right) \\
&\qquad\qquad \mathrm{e}^{-\int_0^t \left(\lambda(X_{T_i} + V_i t, V_i) - \alpha V_i\right)\mathrm{d}s} \mathrm{d}t.
\end{aligned}$$

Furthermore, for $\alpha \in \mathbb{R}$ and $v \in [-1,1]$ the function $\lambda \in \mathbb{R}_+ \longmapsto \frac{\lambda}{\lambda-\alpha v}$ is increasing if $\alpha v \leq 0$ and decreasing otherwise. Therefore, if $V_i \geq 0$, since on $\{S > n\}$ the position of the bacterium is positive, we get :

$$\begin{aligned}
&\mathbb{E}\left[\mathrm{e}^{\alpha V_i \tau_{i+1}} | X_{T_i}, V_i\right] \mathbf{1}_{X_{T_1}>0,\cdots,X_{T_n}>0} \\
&\leq \frac{\inf\limits_{x\geq 0} \lambda(x,V_i)}{\inf\limits_{x\geq 0}\lambda(x,V_i) - \alpha V_i} \int_0^{+\infty} \left(\lambda(X_{T_i} + V_i t, V_i) - \alpha V_i\right) \mathrm{e}^{-\int_0^t \left(\lambda(X_{T_i}+V_i t, V_i) - \alpha V_i\right)\mathrm{d}s} \mathrm{d}t \\
&= \frac{\inf\limits_{x\geq 0} \lambda(x,V_i)}{\inf\limits_{x\geq 0}\lambda(x,V_i) - \alpha V_i}.
\end{aligned}$$

In the case $V_i < 0$ we get in the same way :

$$\mathbb{E}\left[\mathrm{e}^{\alpha V_i \tau_{i+1}} | X_{T_i}, V_i\right] \mathbf{1}_{X_{T_1}>0,\cdots,X_{T_n}>0} \leq \frac{\sup\limits_{x\geq 0} \lambda(x,V_i)}{\sup\limits_{x\geq 0}\lambda(x,V_i) - \alpha V_i}.$$

We have thus obtained the following inequality :

$$\mathbb{E}\left[\mathrm{e}^{\alpha V_i \tau_{i+1}} | X_{T_i}, V_i, S\right] \mathbf{1}_{X_{T_1}>0,\cdots,X_{T_n}>0} \leq G(\alpha, V_i),$$

where G is given by (3.10).
Getting back to (3.12), we get :

$$\mathbb{P}_{x_0,v_0}\left(S > n\right) \leq \mathrm{e}^{\alpha x_0} G(\alpha, v_0) \mathbb{E}_{x_0,v_0}\left[\prod_{i=1}^{n-1} G(\alpha, V_i)\right].$$

Using now the fact that $(V_i)_{i\geq 0}$ is a Markov chain with kernel Q, we have :

$$\begin{aligned}
&\mathbb{E}_{x_0,v_0}\left[\prod_{i=1}^{n-1} G(\alpha, V_i)\right] \\
&\qquad \leq \int_{-1}^{1} G(\alpha, v_1) \times \ldots \times \int_{-1}^{1} G(\alpha, v_{n-1}) Q(v_{n-2}, \mathrm{d}v_{n-1}) \ldots Q(v_0, \mathrm{d}v_1).
\end{aligned}$$

We deduce from Assumption $(\mathcal{A}_4)$, that if $\alpha \in I_*$ we have

$$\mathbb{E}_{x_0,v_0}\left[\prod_{i=1}^{n-1} G(\alpha, V_i)\right] \leq {J_*}^{n-1}.$$

And finally :

$$\mathbb{P}_{x_0,v_0}\left(S > n\right) \leq \mathrm{e}^{\alpha x_0} G(\alpha, v_0) {J_*}^{n-1}, \tag{3.13}$$

for all $\alpha \in I_*$, $x_0 > 0$ and $v_0 \in \mathcal{V} \cap [0,1]$.

For $x_0 > 0$ and $v_0 \in \mathcal{V} \cap [-1, 0)$, the calculations are exactly the same. The case $x_0 \leq 0$ can also be made in the same way. The symmetry of the process and Assumption $(\mathcal{A}_4)$ ensure that, for all α such that $-\alpha \in I_*$, for all $x_0 \leq 0$ and all $v_0 \in \mathcal{V}$,

$$\mathbb{P}_{x_0,v_0}(S > n) \leq \mathrm{e}^{\alpha x_0} G(-\alpha, v_0) {J_*}^{n-1}. \tag{3.14}$$

Finally, we have proved in (3.13) and (3.14) that for all $(x_0, v_0) \in \mathbb{R} \times \mathcal{V}$ and all α such that $|\alpha| \in I_*$,

$$\mathbb{P}_{x_0,v_0}(S > n) \leq \mathrm{e}^{|\alpha x_0|} G(|\alpha|, v_0) {J_*}^{n-1}, \tag{3.15}$$

which is the result of the proposition.

□

Proposition 3.5.5. *For all $(x_0, v_0) \in \mathbb{R} \times \mathcal{V}$, all $0 < \rho < \lambda_{\min}(1 - J_*)$ and all α such that $|\alpha|$ we have :*

$$\mathbb{E}_{x_0,v_0}\left[\mathrm{e}^{\rho Z}\right] \leq K_{\alpha,x_0,v_0} \frac{1}{1 - \frac{\lambda_{\min}}{\lambda_{\min} - \rho} J_*},$$

where

$$K_{\alpha,x_0,v_0} = \kappa_{\alpha,x_0,v_0} \frac{1 - J_*}{{J_*}^2} = \mathrm{e}^{|\alpha x_0|} G\left(|\alpha|, v_0\right) \frac{1 - J_*}{{J_*}^2}.$$

Moreover, K_{α,x_0,v_0} is uniformly bounded from above for $(x_0, v_0) \in [-x_1, x_1] \times \mathcal{V}$ for all $x_1 > 0$.

Proof. As mentioned before, we are going to use the inequality $Z \leq \sum_{i=1}^{S} \tau_i$ a.s. to get some information on the hitting time of the origin.
Let us notice the variables $\tau_i, i \geq 1$ are stochastically smaller than i.i.d. exponential variables $E_i, i \geq 1$ with mean $\frac{1}{\lambda_{\min}}$. Indeed, the survival function of the variable τ_{i+1}, conditionally to X_{T_i} and V_i, is
$\overline{F}(\cdot, X_{T_i}, V_i) = \exp\left(\int_0^t \lambda(X_{T_i} + V_i s, V_i)\mathrm{d}s\right)$, and the jump rate λ is uniformly bounded from below by $\lambda_{\min}$. The variables $(E_i)_{i\geq 1}$ can be taken as independent because of the independence of the variables $(\tau_i)_{i\geq 1}$ conditionally to $(X_{T_i}, V_i)_{i\geq 0}$, and can also be taken as independent of the variables $(\tau_i)_{i\geq 1}$.
Thanks to this comment and the previous proposition, we get, for $\rho > 0$ and α such that $|\alpha| \in I_*$:

$$\begin{aligned}
\mathbb{E}\left[\mathrm{e}^{\rho Z}\right] &\leq \mathbb{E}_{x_0,v_0}\left[\mathrm{e}^{\rho \sum_{i=1}^{S} \tau_i}\right] \\
&\leq \mathbb{E}_{x_0,v_0}\left[\mathrm{e}^{\rho \sum_{i=1}^{S} \tilde{E}_i}\right] \\
&= \mathbb{E}_{x_0,v_0}\left[\mathbb{E}\left[\mathrm{e}^{\rho \sum_{i=1}^{S} \tilde{E}_i} | S\right]\right] \\
&= \mathbb{E}_{x_0,v_0}\left[\prod_{i=1}^{S} \mathbb{E}\left[\mathrm{e}^{\rho \tilde{E}_i}\right]\right]
\end{aligned}$$

$$
\begin{aligned}
&= \mathbb{E}_{x_0,v_0}\left[\left(\frac{\lambda_{min}}{\lambda_{min}-\rho}\right)^S\right] \\
&= \sum_{n=0}^{+\infty}\left(\frac{\lambda_{min}}{\lambda_{min}-\rho}\right)^n \mathbb{P}_{x_0,v_0}(S=n) \\
&= \sum_{n=0}^{+\infty}\left(\frac{\lambda_{min}}{\lambda_{min}-\rho}\right)^n (\mathbb{P}(S_{x_0,v_0}>n-1)-\mathbb{P}(S_{x_0,v_0}>n)) \\
&\leq \mathrm{e}^{|\alpha x_0|}G(|\alpha|,v_0)\frac{1-J_*}{J_*^2}\sum_{n=0}^{+\infty}\left(\frac{\lambda_{min}}{\lambda_{min}-\rho}J_*\right)^n.
\end{aligned}
$$

Finally, for all $\rho>0$ such that $\frac{\lambda_{min}}{\lambda_{min}-\rho}J_*<1$, we get :

$$
\mathbb{E}\left[\mathrm{e}^{\rho Z}\right] \leq \mathrm{e}^{|\alpha x_0|}G(|\alpha|,v_0)\frac{1-J_*}{J_*^2}\frac{1}{1-\frac{\lambda_{min}}{\lambda_{min}-\rho}J_*}<\infty, \tag{3.16}
$$

which ends the proof of the proposition. □

3.5.2 Exponential ergodicity of the process

In this section, we are going to give a proof of the exponential ergodicity of our process based on Proposition 3.5.5. For this purpose, we recall another result on the exponential ergodicity of a Markov process.

Theorem 3.5.6. *(Theorem 6.2 in [Down et al., 1995]) Let Y be an irreducible and aperiodic Markov process. Suppose that there exists a function $f\geq 1$, a closed measurable set $C\in E$ and some constants $\delta,\eta>0$, $M<\infty$ such that*

$$
\mathbb{E}_x\left[\int_0^{\tau_C(\delta)}\mathrm{e}^{\eta t}f(Y_t)\mathrm{d}t\right]<\infty, \quad \textit{for all } x\notin C
$$

and

$$
\sup_{x\in C}\mathbb{E}_x\left[\int_0^{\tau_C(\delta)}\mathrm{e}^{\eta t}f(Y_t)\mathrm{d}t\right]\leq M
$$

where $\tau_C(\delta)=\inf\{t\geq\delta, Y_t\in C\}$.
If the set C is petite for Y, then the process is exponentially ergodic.

Remark 3.5.7. *The proof of this theorem relies on the introduction of the function*
$H_0(x):=1+\mathbb{E}_x\left[\int_0^{\tau_C(\delta)}\mathrm{e}^{\eta t}f(Y_t)\mathrm{d}t\right]$, which plays the role of a Lyapunov function for a Markov chain linked to the process Y.

To apply this theorem, we need then, as previously, to find a petite set. The following proposition states that all compact sets are petite for the process (X,V), and is just a consequence of Proposition 3.3.3.

Proposition 3.5.8. *For all $x_1 > 0$, the set $C = [-x_1, x_1] \times \mathcal{V}$ is petite for the process (X, V).*

We can now prove Theorem 3.5.3 :

Proof of Theorem 3.5.3. Let $x_1 > 0$. Proposition 3.5.8 ensures that the closed set $C = [-x_1, x_1] \times \mathcal{V}$ is petite for the process (X, V).
Moreover, let us consider the quantity $\mathbb{E}_{x,v}\left[e^{\eta\tau_C(\delta)}\right]$ with $\eta, \delta > 0$. It satisfies the assumptions of Theorem 3.5.6 (we have taken $f \equiv 1$) for $0 < \eta < \lambda_{\min}(1 - J_*)$ thanks to Proposition 3.5.5.
Finally, let *Leb* be the Lebesgue measure on $\mathbb{R}$. The process (X, V) is $Leb \otimes Q$-irreducible, and aperiodic. The proof of these facts can be handled as mentioned previously in the general case. The conclusion of Theorem 3.5.6 gives then the exponential ergodicity of (X, V). □

3.5.3 Exponential moments of the invariant measure

As in the general case, we can show that the invariant measure, that we still denote by π, has exponential moments.

Theorem 3.5.9. *Let $\gamma > 0$ such that $\frac{\lambda_{\min}}{\lambda_{\min}-\gamma} J_* < 1$.*
For all $0 < \alpha < \gamma$ and all $\beta > 0$, we have :

$$\int_{\mathbb{R}\times[-1,1]} e^{\alpha|x|+\beta|v|}\pi(dx, dv) < \infty.$$

Proof. The proof relies, as in Section 3.4, on Theorem 3.4.1. We see the process (X, V) as a semi-regenerative process between the successive hitting times of 0. Then, using the upper bound of the exponential moments of the hitting time of 0 obtained in Proposition 3.5.5, we can verify that the assumptions of Theorem 3.4.1 are satisfied, and the result follows. □

3.5.4 Comparison between the two studies

Let us first recall that the proof of the exponential ergodicity carried out in any dimension is obviously relevant in dimension 1, whereas the converse is not possible, or at least not directly. Indeed, the process in dimension 1 is quite simple since it goes towards the origin or not in terms of the sign of the scalar product $X_t \cdot V_t$. In higher dimension, if we want to estimate the hitting time of a compact set, let say a ball, we can not see if the process is evolving towards the ball just in terms of the sign of $X_t \cdot V_t$, and the computations do not proceed as well as in dimension 1.
Let now see that the two approaches carried out in dimension 1 and in higher dimension cover different types of PDMPs.

— Let look at the process studied in [Fontbona *et al.*, 2012], [Fontbona *et al.*, 2016] and [Bierkens et Roberts, 2017], whose generator has the following form :

$$Lf(x,v) = v\partial_x f(x,v) + \lambda(x,v)\left(f(x,-v) - f(x,v)\right),$$

for $(x,v) \in \mathbb{R} \times \{-1,+1\}$.
This process does not satisfy Assumption $(\mathcal{A}_4)$, because for $v' = -1$ we have

$$\int_{-1}^{1} G(\tilde{\alpha},v)Q(-1,\mathrm{d}v) = G(\tilde{\alpha},1) = \frac{\inf\limits_{x\geq 0}\lambda(x,1)}{\inf\limits_{x\geq 0}\lambda(x,1) - \alpha} > 1,$$

whereas it is still ergodic. Indeed, its ergodicity has been prooved in [Bierkens et Roberts, 2017] or [Fontbona *et al.*, 2016] for instance, with the introduction of a Lyapunov function, and this can also be deduce from Theorem 3.1.2, if the jump-rate λ satisfies the good assumptions.

— In the d-dimensional case, Assumption $(\mathcal{H}_4)$ assume a kind of monotony of the jump rate, which is not necessary in Assumption $(\mathcal{A}_4)$ in the one-dimensional case. We can thus construct a particular jump rate which, associated to the kernel Q, satisfies Assumption $(\mathcal{A}_4)$, but does not verify Assumption $(\mathcal{H}_4)$. Let for instance consider the case

$$Lf(x,v) = v\partial_x f(x,v) + (2\mathbf{1}_{\mathrm{sgn}(x)v\geq -\frac{1}{2}} + \frac{1}{2}\mathbf{1}_{\mathrm{sgn}(x)v< -\frac{1}{2}})\frac{1}{2}\int_{-1}^{1}\left(f(x,v') - f(x,v)\right)\mathrm{d}v',$$

for $(x,v) \in \mathbb{R} \times [-1,1]$.
This process obviously does not satisfies Assumption $(\mathcal{H}_4)$ since for all $\theta_* \in [0,1]$, and all $\Delta > 0$,
$\inf_{\{\mathrm{sgn}(x)v\geq\theta_*,|x|\geq\Delta\}}\lambda(x,v) = 2 = \sup_{\{\mathrm{sgn}(x)v\leq 0\}}\lambda(x,v)$.
Nevertheless, Assumption $(\mathcal{A}_4)$ is verified for this process. Indeed, let $\alpha \in \left(0, \inf\limits_{x\geq 0, v\in(0,1]}\frac{\lambda(x,v)}{v}\right)$, we have :

$$\begin{aligned}
&\frac{1}{2}\int_{-1}^{1} G(\alpha,v)\mathrm{d}v \\
&= \frac{1}{2}\int_{-1}^{-\frac{1}{2}} \frac{\sup\limits_{x\geq 0}\lambda(x,v)}{\sup\limits_{x\geq 0}\lambda(x,v) - \alpha v}\mathrm{d}v + \frac{1}{2}\int_{-\frac{1}{2}}^{-0} \frac{\sup\limits_{x\geq 0}\lambda(x,v)}{\sup\limits_{x\geq 0}\lambda(x,v) - \alpha v}\mathrm{d}v \\
&\qquad + \frac{1}{2}\int_{0}^{1} \frac{\inf\limits_{x\geq 0}\lambda(x,v)}{\inf\limits_{x\geq 0}\lambda(x,v) - \alpha v}\mathrm{d}v \\
&= \frac{1}{2}\int_{-1}^{-\frac{1}{2}} \frac{\frac{1}{2}}{\frac{1}{2} - \alpha v}\mathrm{d}v + \frac{1}{2}\int_{-\frac{1}{2}}^{1} \frac{2}{2 - \alpha v}\mathrm{d}v
\end{aligned}$$

$$= \frac{1}{4\alpha} \log\left(\frac{\frac{1}{2}+\alpha}{\frac{1}{2}+\frac{\alpha}{2}}\right) + \frac{1}{\alpha}\log\left(\frac{2+\frac{\alpha}{2}}{2-\alpha}\right).$$

A curve sketching shows that there exists an interval $I_* \subset \left[0, \inf_{x\geq 0, v\in(0,1]} \frac{\lambda(x,v)}{v}\right)$ and a constant $J_* \in (0,1)$ such that for all $\alpha \in I_*$, $\frac{1}{2}\int_{-1}^{1} G(\alpha, v)\mathrm{d}v \leq J_*$, i.e. Assumption $(\mathcal{A}_4)$ is satisfied.

Chapitre 4

Zig-zag processes in interaction

In this chapter, we are interested in a system of one-dimensional Zig-zag processes in mean field type interaction. Our aim is to study first the propagation of chaos and the existence of the non-linear limit process, and then to investigate its long-time behaviour. For the latter point, we are led to work with a centred particle system and its associated centred non-linear process.
Since this work is in progress, with my PhD advisors, some of our conjectures are not yet proved, but we will illustrate the expected results with some simulations.

4.1 Introduction

We consider a system of N couples $\big((X^{1,N},V^{1,N}),\cdots,(X^{N,N},V^{N,N})\big)$ in $\mathbb{R}\times\{-1,+1\}$ where each couple is a one-dimensional Zig-zag process attracted by the mean position of all the particles. We suppose that the jump-rate of each particle is of the form $\lambda\left(\left(X^{i,N}-\frac{1}{N}\sum_{k=1}^N X^{k,N}\right)V^{i,N}\right)$, which suppose a sort of symmetry around the mean position of the system. The particle system has thus the following generator :
for $f\in\mathcal{C}^{1,0}(\mathbb{R}^N\times\{-1,+1\}^N,\mathbb{R})$ and $(x,v)=((x_1,\cdots,x_N),(v_1,\cdots,v_N))\in\mathbb{R}^N\times\{-1,+1\}^N$,

$$\mathcal{L}_N f(x,v)=\sum_{i=1}^N v_i\partial_{x_i}f(x,v)+\sum_{i=1}^N\lambda\left((x_i-\overline{x})v_i\right)\Big(f(x,v^{(i)})-f(x,v)\Big),\quad (4.1)$$

where $v_j^{(i)}=\begin{cases} v_j & \text{if } j\neq i\\ -v_i & \text{else.}\end{cases}$ and $\overline{x}=\frac{1}{N}\sum_{k=1}^N x_i$, and where λ is a non-negative function on $\mathbb{R}$, with assumptions that will be detailed thereafter.
In other words, for $i\in\{1,\cdots,N\}$, the couple $(X^{i,N},V^{i,N})$ is solution of the following stochastic system :

$$\begin{cases} \mathrm{d}X_t^{i,N}=V_t^{i,N}\mathrm{d}t\\ \mathrm{d}V_t^{i,N}=-2V_{t^-}^{i,N}\int\mathbf{1}_{z\leq\lambda((X_t^{i,N}-\frac{1}{N}\sum_{j=1}^N X_t^{j,N})V_{t^-}^{i,N}))}\mathcal{N}^i(\mathrm{d}z,\mathrm{d}t)\end{cases}\quad (4.2)$$

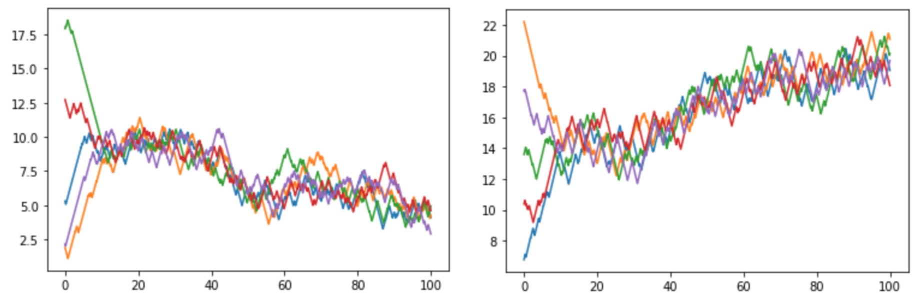


FIGURE 4.1 – Trajectories of five particles following (4.2), left with $\lambda(z) = \arctan(z) + 3$, and right with $\lambda(z) = 1 + z\mathbf{1}_{z>0}$.

where the $\mathcal{N}^i$, $i = 1, \cdots, N$, are N independent Poisson measures on $\mathbb{R}_+^2$ with Lebesgue intensity measure.
We can observe on Figure 4.1 two trajectories of five particles following (4.2), with $\lambda(z) = \arctan(z) + 3$, and $\lambda(z) = 1 + z\mathbf{1}_{z>0}$. The initial positions are independent variables with law $\mathcal{E}(\frac{1}{10})$, and the initial speeds are i.i.d. Rademacher variables with parameter $\frac{1}{2}$.
Since under good assumptions on the jump-rate λ, a Zig-zag process can model the movement of a bacterium which is attracted by a fixed nutriment (see [Bierkens et Roberts, 2017], [Fontbona *et al.*, 2012], [Fontbona *et al.*, 2016], [Calvez *et al.*, 2015] and [Fétique, 2017]) the particle system with infinitesimal generator (4.1) can model the movement of a group of N bacteria that are attracted by their mean position. Under such assumptions on λ, this particle system of PDMPs in interaction models therefore a population of collaborative bacteria.
Our goal is to investigate the behaviour of one particle interacting with the others, when the number N of particles tends to infinity. We will see that when the particles start with i.i.d. initial conditions, the time-evolution of such a bacterium can be approximated by a non-linear process. We say that propagation of chaos holds. In our case, since the interaction between the particles solutions of (4.2) holds by their mean position $\frac{1}{N}\sum_{k=1}^N X^{k,N}$, at the limit when N tends to infinity, the non-linear process interacts with its law, and particularly with its mean position. Therefore, the propagation of chaos holds towards a process, that we will call the non-linear Zig-zag process, solution of the following non-linear system :

$$\begin{cases} \mathrm{d}X_t = V_t \mathrm{d}t \\ \mathrm{d}V_t = -2V_{t^-} \int \mathbf{1}_{z \le \lambda((X_t - \mathbb{E}[X_t])V_{t^-})} \mathcal{N}(\mathrm{d}z, \mathrm{d}t), \end{cases} \tag{4.3}$$

where $\mathcal{N}$ is a Poisson measure on $\mathbb{R}_+^2$ with Lebesgue intensity measure.
Since such a system is non-linear, due to the dependence of the evolution of X_t on its mean $\mathbb{E}[X_t]$, the existence and uniqueness of the solution of (4.3) is non-trivial. We thus first study this point, using a classical method of Picard iterations. Then, the propagation of chaos can be proved with the same kind of calculations, by coupling N particles starting at i.i.d. initial conditions with N

independent copies of the non-linear Zig-zag process driven by the same Poisson processes as the particles, and with the same initial conditions.
Once these two points, existence and uniqueness of the non-linear process and propagation of chaos, have been proved, our goal is to investigate the long-time behaviour of the non-linear Zig-zag process. However, its behaviour depends on its mean position, whose evolution is not trivial. We are therefore led to introduce a new particle system and its associated non-linear process, that will be easier to study. We thus consider the couples $\left((Y_t^{i,N}, V_t^{i,N})\right)_{t\geq 0}$, $i \in \{1, \cdots, N\}$, where $Y_t^{i,N} = X_t^{i,N} - \frac{1}{N}\sum_{k=1}^N X_t^{k,N}$, and the associated non-linear process $((Y_t, V_t))_{t\geq 0}$, that we will name the non-linear centred Zig-zag process. The evolution of $V^{i,N}$ now only depends on $Y^{i,N}$ and $V^{i,N}$, and the interaction lies in the evolution of $Y^{i,N}$ since $\mathrm{d}Y_t^{i,N} = \left(V_t^{i,N} - \frac{1}{N}\sum_{k=1}^N V_t^{k,N}\right)\mathrm{d}t$. Therefore, the non-linearity of the non-linear centred Zig-zag process now holds with $\mathbb{E}[V_t]$, that seems easier to study than $\mathbb{E}[X_t]$. Our goal is to prove that, under assumptions on the jump-rate λ, the non-linear centred Zig-zag process has a unique invariant probability measure, and that it converges towards it in time. These facts still are conjectures, and we will discuss about it.

Mean-field type interacting particles and propagation of chaos where firstly introduced by Kac in [Kac, 1956], and we can also cite the seminal works of McKean [McKean, 1967] and Snzitman [Burkholder *et al.*, 1991] on this topic. The literature on interacting particle system is huge, and we can not cite all of the works on this topic, so we only give few examples that study different kinds of interacting particle systems. We mention the work of Malrieu [Malrieu, 2003], in which the author studies a mean-field diffusive particle system. He studies the propagation of chaos, that holds uniformly in time, towards a non-linear process, in which the non-linearity holds with the law of the process. He also investigates the exponentially fast convergence of the non-linear process to equilibrium. In [Fontbona *et al.*, 2009], the authors are interested in non-linear processes that are solutions of non-linear stochastic equations driven by space-time white noises. Their goal is to construct a diffusive interacting particle system, that is easy to simulate, and that approximates the non-linear process. We can also cite the work [Andreis *et al.*, 2018], in which Adreis and her co-authors investigate the propagation of chaos of a particle system that can model a system of neurons interacting by their mean. The particles are diffusions with simultaneous jumps.
Finally, let us mention some works on PDMPs in interaction, which is the context of this chapter. In [Thai, 2015], the authors is interested in a system of birth and death processes in mean-field type interaction in discrete space. Under good assumptions on the parameters on which the model depends, she shows the exponentially fast convergence of the particle system in Wasserstein distance, the uniform propagation of chaos, and the convergence of the non-linear limit process. The paper [Graham et Robert, 2009] deals with a system of PDMPs in interaction that models the interaction of several classes of per-

manent connections in a network. They analyse the existence and uniqueness of the associated non-linear process, and the propagation of chaos. The computations that we make for the corresponding results in our study are very close to theirs. Then, they make a link between the existence of invariant measures for the non-linear process and the existence of solutions of a fixed-point equation. Let us also mention the work [Fournier et Löcherbach, 2016]. They study a stochastic system of interacting neurons, modeled by a system of PDMPs in interaction. The interaction holds in the jump process, and also in the deterministic behaviour. They prove the propagation of chaos towards a non-linear process, and investigate the long-time behaviour of this latter. Finally, let us speak about the recent work [Monmarché, 2018], which is related to ours. In his paper, Monmarché is interested in the long-time behaviour of mean-field type interacting particle systems and non-linear processes, and he studies them with their associated mean-field integro-differential equations. The framework of his work is the following : he considers processes, including PDMPs, in the perturbative regime, that is the case where the interactions have a negligible effect with respect to the mixing effect of the process without interactions, and the interactions are supposed to hold in the jump mechanism of the processes. He proves convergence in total variation distance by working on the laws of the processes, whereas we work in this Chapter on the trajectories. Let us finally describe one example detailed in his paper : the study of a Zig-zag process in mean-field type interaction. He considers the following non-linear integro-differential equation on $\mathbb{R} \times \{-1, +1\}$:

$$\partial_t m_t(x, v) + y\partial_x m_t(x, v) = \lambda_{m_t}(x, -v)m_t(x, v) - \lambda_{m_t}(x, v)m_t(x, v),$$

with $\lambda_\nu = r\left(v(x - \theta x_\nu)\right)$, where $x_\nu = \int_{\mathbb{R}\times\{-1,+1\}} x\nu(\mathrm{d}x, \mathrm{d}v)$, $\theta \in (0, 1)$ and $r : \mathbb{R} \to \mathbb{R}^+$ is a Lipschitz function such that $\lim_{s\to-\infty} r(s) < \lim_{s\to+\infty} r(s)$ and $\inf_{s\in\mathbb{R}} r(s) > 0$. This equation describes a one-dimensional Zig-zag process which is attracted by an average of the origin and the barycentre of its law. Monmarché shows that for θ small enough, that is when the attraction to the origin is not too strong, the process converges exponentially fast to equilibrium. Therefore, this example is near to our study of the non-linear Zig-zag process, but the difference, that makes the behaviour of the process completely different, is that we do not consider any attraction to the origin.

This chapter is organized as follows : in Section 4.2, we investigate the existence and uniqueness of the non-linear Zig-zag process, and the propagation of chaos of the particle system towards this process. Then, in Section 4.3, we introduce the centred particle system and non-linear process, in order to investigate the long-time behaviour of the latter. Finally, we discuss the prospects to improve and complete this work in Section 4.4.

4.2 Study of the interacting particles system

In this section, we first investigate the existence and uniqueness of the non-linear process solution of (4.3), process that we name the non-linear Zig-zag process. Then, we prove the propagation of chaos of the particle system (4.2). For this purpose, we make the following assumption on the jump rate λ, which is a classical assumption to prove the existence and uniqueness of solutions of stochastic differential equations :

Assumption $(\mathcal{H}_\lambda)$ **:**

The jump-rate λ is a non-negative Lipschitz continuous function on $\mathbb{R}$, and we denote by K its Lipschitz constant :

$$\forall x, y \in \mathbb{R}, \quad |\lambda(x) - \lambda(y)| \leq K |x - y| .$$

4.2.1 Pathwise existence and uniqueness of the non-linear system

For $f : \mathbb{R}^+ \to \mathbb{R}^2$ and $T > 0$, we define

$$\|f\|_T = \sup_{0 \leq s \leq T} \|f(s)\|.$$

In order to prove the existence and uniqueness of a solution of the system (4.3), which is not trivial due to the dependence of a solution on its own law, we use a standard method of fix point.

Proposition 4.2.1. *We assume $(\mathcal{H}_\lambda)$. For a (deterministic) measurable function $m : \mathbb{R}_+ \to \mathbb{R}$ and a couple (X_0, V_0) of random variables on $\mathbb{R} \times \{-1, +1\}$, let $\phi(X_0, V_0, m) = (X_t, V_t)_{t \geq 0}$ be the solution of the linear stochastic differential equation*

$$\begin{cases} \mathrm{d}X_t &= V_t \mathrm{d}t \\ \mathrm{d}V_t &= -2V_{t^-} \int \mathbf{1}_{z \leq \lambda((X_t - m_t)V_{t-})} \mathcal{N}(\mathrm{d}z, \mathrm{d}t) \end{cases} \tag{4.4}$$

starting at (X_0, V_0), where $\mathcal{N}$ is a Poisson measure with Lebesgue intensity measure. Then, for any such functions $m, \tilde{m}$ and couples of random variables (X_0, V_0), $(\tilde{X}_0, \tilde{V}_0)$, any $T > 0$, we have

$$\mathbb{E}_0 \left[\left\| \phi(X_0, V_0, m) - \phi\left(\tilde{X}_0, \tilde{V}_0, \tilde{m}\right) \right\|_T \right] \leq C_0(T) \mathrm{e}^{\int_0^T h(s) \mathrm{d}s}, \tag{4.5}$$

with

$$C_0(T) = |X_0 - \tilde{X}_0| + |V_0 - \tilde{V}_0| + 2K \int_0^T \|m - \tilde{m}\|_s \mathrm{d}s$$
$$h(T) = 1 + 2K(|X_0| + T + |m_T| + 1).$$

where $\mathbb{E}_0[\cdot]$ denotes the conditional expectation with respect to $(X_0, V_0, \tilde{X}_0, \tilde{V}_0)$.

Proof. The linear Markov process $\phi(X_0, V_0, m)$ is well defined (see for instance [Gikhman et Skorohod, 1972] for details on stochastic differential equations driven by Poisson measures) and can be constructed by introducing the successive jump times of the process as follows. Let $(E_i)_{i\geq 1}$ i.i.d. exponential random variables, also independent of (X_0, V_0).

For $(X_0, V_0) = (x_0, v_0)$, let $T_1 = \inf\left\{t \geq 0 : \int_0^t \lambda((x_0 + v_0 s - m_s)v_0) > E_1\right\}$ be the first jump time. For $t \in [0, T_1)$, we set $V_t = v_0$ and $X_t = x_0 + v_0 t$, and for $t = T_1$, $V_{T_1} = -v_0$ and $X_{T_1} = x_0 + v_0 T_1$. Then we introduce the successive inter-jump times τ_n, and the jump times T_n, $n \geq 1$, by $\tau_1 = T_1$ and for $n \geq 1$,

$$\tau_{n+1} = \inf\left\{t \geq 0 : \int_0^t \lambda((X_{T_n} + V_{T_n} s - m_{s+T_n})V_{T_n}) > E_{n+1}\right\},$$

with $\inf \varnothing = +\infty$, and $T_{n+1} = T_n + \tau_{n+1}$.
Then we have for $t \in [T_n, T_{n+1})$, we set $V_t = V_{T_n}$ and $X_t = X_{T_n} + V_{T_n} t$, and for $t = T_{n+1}$, $V_{T_{n+1}} = -V_{T_n}$ and $X_t = X_{T_n} + V_{T_n} T_{n+1}$.

We denote by $(X_t, V_t) = \phi\left(X_0, V_0, m\right)(t)$ and $\left(\tilde{X}_t, \tilde{V}_t\right) = \phi\left(\tilde{X}_0, \tilde{V}_0, \tilde{m}\right)(t)$. To simplify the read of the proof, we introduce the notations

$$\lambda_{s^-} := \lambda\left((X_s - m_s)V_{s^-}\right), \quad \lambda_s := \lambda\left((X_s - m_s)V_s\right)$$
$$\tilde{\lambda}_{s^-} := \lambda\left((\tilde{X}_s - \tilde{m}_s)\tilde{V}_{s^-}\right), \quad \tilde{\lambda}_s := \lambda\left((\tilde{X}_s - \tilde{m}_s)\tilde{V}_s\right).$$

By integrating equation (4.4) for (X_t, V_t) and $(\tilde{X}_t, \tilde{V}_t)$ we have :

$$\begin{cases} X_t - \tilde{X}_t & = X_0 - \tilde{X}_0 + \int_0^t \left(V_s - \tilde{V}_s\right) \mathrm{d}s \\ V_t - \tilde{V}_t & = V_0 - \tilde{V}_0 - 2\int_0^t \int \left(V_{s^-} \mathbf{1}_{z \leq \lambda_{s^-}} - \tilde{V}_{s^-} \mathbf{1}_{z \leq \tilde{\lambda}_{s^-}}\right) \mathcal{N}(\mathrm{d}z, \mathrm{d}s). \end{cases}$$

Itô's formula for the absolute value gives then :

$$\begin{aligned} &\left|V_t - \tilde{V}_t\right| = \left|V_0 - \tilde{V}_0\right| \\ &+ \int_0^t \int \left(\left|V_s - \tilde{V}_s - 2V_{s^-}\mathbf{1}_{z \leq \lambda_{s^-}} + 2\tilde{V}_{s^-}\mathbf{1}_{z \leq \tilde{\lambda}_{s^-}}\right| - \left|V_s - \tilde{V}_s\right|\right) \mathcal{N}(\mathrm{d}z, \mathrm{d}s) \\ &= \left|V_0 - \tilde{V}_0\right| + \int_0^t \int \left(|V_s + \tilde{V}_s| - |V_s - \tilde{V}_s|\right) \\ &\qquad \mathbf{1}_{\min\{\lambda_{s^-};\tilde{\lambda}_{s^-}\} \leq z \leq \max\{\lambda_{s^-};\tilde{\lambda}_{s^-}\}} \mathcal{N}(\mathrm{d}z, \mathrm{d}s) \end{aligned}$$

so that, taking conditional expectation we get :

$$\mathbb{E}_0\left[\left\|V - \tilde{V}\right\|_t\right] = \left|V_0 - \tilde{V}_0\right| + \int_0^t \int \mathbb{E}_0\left[\left(|V_s + \tilde{V}_s| - |V_s - \tilde{V}_s|\right) \mathbf{1}_{\min\{\lambda_s;\tilde{\lambda}_s\} \leq z \leq \max\{\lambda_s;\tilde{\lambda}_s\}}\right] \mathrm{d}z\mathrm{d}s$$

$$\leq \left|V_0 - \tilde{V}_0\right| + 2\int_0^t \mathbb{E}_0\left[\left|\lambda_s - \tilde{\lambda}_s\right|\right] \mathrm{d}s.$$

Since λ is K-Lipschitz-continuous, for all $t \geq 0$, $|V_t| = 1$ and $|X_t| \leq |X_0| + t$, we deduce

$$
\begin{aligned}
&\mathbb{E}_0\left[\left\|V-\tilde{V}\right\|_t\right] \\
&\leq \left|V_0-\tilde{V}_0\right| + 2K\int_0^t \mathbb{E}_0\left[\left|(X_s-m_s)V_s-(\tilde{X}_s-\tilde{m}_s)\tilde{V}_s\right|\right]\mathrm{d}s \\
&\leq \left|V_0-\tilde{V}_0\right| + 2K\int_0^t \mathbb{E}_0\left[\left|(X_s-m_s)(V_s-\tilde{V}_s)+\tilde{V}_s(X_s-\tilde{X}_s-(m_s-\tilde{m}_s))\right|\right]\mathrm{d}s \\
&\leq \left|V_0-\tilde{V}_0\right| + 2K\int_0^t (|X_0|+s+|m_s|)\mathbb{E}_0\left[\left|V_s-\tilde{V}_s\right|\right]\mathrm{d}s \\
&\qquad + 2K\int_0^t \mathbb{E}_0\left[\left|X_s-\tilde{X}_s\right|\right]\mathrm{d}s + 2K\int_0^t |m_s-\tilde{m}_s|\mathrm{d}s \\
&\leq \left|V_0-\tilde{V}_0\right| + 2K\int_0^t (|X_0|+s+|m_s|)\mathbb{E}_0\left[\left\|V-\tilde{V}\right\|_s\right]\mathrm{d}s \\
&\qquad + 2K\int_0^t \mathbb{E}_0\left[\left\|X-\tilde{X}\right\|_s\right]\mathrm{d}s + 2K\int_0^t \|m-\tilde{m}\|_s\mathrm{d}s.
\end{aligned} \tag{4.6}
$$

Moreover we have

$$
\mathbb{E}_0\left[\left\|X-\tilde{X}\right\|_t\right] \leq \left|X_0-\tilde{X}_0\right| + \int_0^t \mathbb{E}_0\left[\left\|V-\tilde{V}\right\|_s\right]\mathrm{d}s.
$$

The two equations above give :

$$
\begin{aligned}
&\mathbb{E}_0\left[\|X-\tilde{X}\|_t+\|V-\tilde{V}\|_t\right] \\
&\quad\leq |X_0-\tilde{X}_0|+|V_0-\tilde{V}_0|+2K\int_0^t \|m-\tilde{m}\|_s\,\mathrm{d}s \\
&\qquad + \int_0^t (1+2K(|X_0|+s+|m_s|+1))\left(\mathbb{E}_0\left[\|X-\tilde{X}\|_s+\|V-\tilde{V}\|_s\right]\right)\mathrm{d}s.
\end{aligned}
$$

We thus have, with the notations introduced in the proposition :

$$
\mathbb{E}_0\left[\|X-\tilde{X}\|_t+\|V-\tilde{V}\|_t\right] \leq C_0(t) + \int_0^t h(s)\mathbb{E}_0\left[\|X-\tilde{X}\|_s+\|V-\tilde{V}\|_s\right]\mathrm{d}s.
$$

Since $t\mapsto C_0(t)$ is non-decreasing, Gronwall's lemma gives :

$$
\mathbb{E}_0\left[\|X-\tilde{X}\|_t+\|V-\tilde{V}\|_t\right] \leq C_0(t)\mathrm{e}^{\int_0^t h(s)\mathrm{d}s},
$$

and the result is proved. □

Let us now introduce a condition on the initial data.

Assumption $(\mathcal{H}_\alpha)$ **:**

There exists $\alpha>0$ such that $\mathbb{E}\left[\mathrm{e}^{\alpha|X_0|}\right]<\infty$.

We can now state the result that establishes the existence and uniqueness of a solution of (4.3), namely the non-linear Zig-zag process.

Theorem 4.2.2. *Let us suppose that Assumption* $(\mathcal{H}_\lambda)$ *holds. If* X_0 *satisfies Assumption* $(\mathcal{H}_\alpha)$*, there is pathwise existence and uniqueness of a solution of* (4.3) *starting at* (X_0, V_0)*.*
Moreover, let (X, V) *and* $(\tilde{X}, \tilde{V})$ *be solutions of* (4.3) *with respective initial conditions* (X_0, V_0) *and* $(\tilde{X}_0, \tilde{V}_0)$*. Suppose that* X_0 *and* $\tilde{X}_0$ *satisfy Assumption* $(\mathcal{H}_\alpha)$*. Then for any* $T \geq 0$*, there exists some constants* $C_T, C'_T > 0$ *such that :*

$$\begin{aligned}\mathbb{E}&\left[\|X-\tilde{X}\|_T+\|V-\tilde{V}\|_T\right]\\ &\leq \mathbb{E}\left[\left(|X_0-\tilde{X}_0|+|V_0-\tilde{V}_0|\right)\mathrm{e}^{\alpha|X_0|}\right]\left(1+C_T\exp\left(C'_T\mathbb{E}\left[\mathrm{e}^{\alpha|X_0|}\right]\right)\mathbb{E}\left[\mathrm{e}^{\alpha|X_0|}\right]\right).\end{aligned} \tag{4.7}$$

Remark 4.2.3. *The last inequality of the theorem describes the continuity of the solution of* (4.3) *with respect to the initial condition.*

Proof. We use a classical Picard iteration and thus construct a sequence $(X^n, V^n)_{n\geq 0}$ by induction as follows :

— $(X^0, V^0) = (X_0, V_0)$ and $m^0 = 0$,

— for $n \geq 1$, let (X^n, V^n) be the solution of (4.4) with $(X_0^n, V_0^n) = (X_0, V_0)$ and $m = m^{n-1}_{\cdot} := t \mapsto \mathbb{E}\left[X_t^{n-1}\right]$.

Applying (4.5) we have, for $t \geq 0$:

$$\mathbb{E}_0\left[\|X^{n+1}-X^n\|_t+\|V^{n+1}-V^n\|_t\right] \leq 2K\mathrm{e}^{\int_0^t h(s)\mathrm{d}s}\int_0^t \|m^n - m^{n-1}\|_s \mathrm{d}s.$$

Since $h(s) = 1+2K(|X_0|+s+|m_s^n|+1)$ and, by definition of m^n, $|m_s^n| \leq \mathbb{E}\left[|X_0|\right]+s$, we deduce

$$\mathrm{e}^{\int_0^t h(s)\mathrm{d}s} \leq A_t \mathrm{e}^{2Kt|X_0|}, \tag{4.8}$$

with $A_t := \mathrm{e}^{(1+2K(1+\mathbb{E}[|X_0|]))t+2Kt^2}$.
Let define $t_0 = \frac{\alpha}{2K}$, where α is the constant introduced in Assumption $(\mathcal{H}_\alpha)$. We note that $\delta := 2KA_{t_0}\mathbb{E}\left[\mathrm{e}^{\alpha|X_0|}\right]$ is finite, and we have, for $t \leq t_0$,

$$\begin{aligned}\mathbb{E}\left[\|X^{n+1}-X^n\|_t+\|V^{n+1}-V^n\|_t\right] &= \mathbb{E}\left[\mathbb{E}_0\left[\|X^{n+1}-X^n\|_t+\|V^{n+1}-V^n\|_t\right]\right]\\ &\leq \delta\int_0^t \|m^n-m^{n-1}\|_s\mathrm{d}s.\end{aligned} \tag{4.9}$$

On the other hand, we have

$$\begin{aligned}\|m^n-m^{n-1}\|_t &\leq \mathbb{E}\left[\|X^n-X^{n-1}\|_t\right]\\ &\leq \delta\int_0^t \|m^{n-1}-m^{n-2}\|_s\mathrm{d}s\\ &\leq \delta^2\int_0^t\int_0^{s_1}\|m^{n-2}-m^{n-3}\|_{s_2}\mathrm{d}s_2\mathrm{d}s_1\end{aligned}$$

$$\begin{aligned}
&\leq \cdots \\
&\leq \delta^n \int_0^t \int_0^{s_1} \cdots \int_0^{s_{n-1}} \|m^1 - m^0\|_{s_{n-1}} \mathrm{d}s_{n-1} \cdots \mathrm{d}s_2 \mathrm{d}s_1 \\
&\leq \frac{(\delta t)^{n-1}}{(n-1)!} \|m^1 - m^0\|_t. \qquad (4.10)
\end{aligned}$$

By plugging this relation in (4.9), we get :

$$\begin{aligned}
\mathbb{E}\left[\|X^{n+1} - X^n\|_{t_0} + \|V^{n+1} - V^n\|_{t_0}\right] &\leq \delta \int_0^{t_0} \frac{(\delta s)^{n-1}}{(n-1)!} \|m^1 - m^0\|_s \mathrm{d}s \\
&\leq \frac{(\delta t_0)^n}{n!} \|m^1 - m^0\|_{t_0}.
\end{aligned}$$

We deduce that the sequence of processes $((X_t^n, V_t^n)_{0 \leq t \leq t_0})_{n \geq 0}$ converges uniformly to some process $((X_t, V_t))_{0 \leq t \leq t_0}$, and we have :

$$\begin{aligned}
\mathbb{E}\left[\|X^n - X\|_{t_0} + \|V^n - V\|_{t_0}\right] &\leq \mathbb{E}\left[\sum_{k=n}^{+\infty} \left(\|X^k - X^{k+1}\|_{t_0} + \|V^k - V^{k+1}\|_{t_0}\right)\right] \\
&\leq \sum_{k=n}^{+\infty} \frac{(\delta t_0)^k}{k!} \|m^1 - m^0\|_{t_0} \\
&= \frac{(\delta t_0)^n}{n!} \mathrm{e}^{\delta t_0} \|m^1 - m^0\|_{t_0}.
\end{aligned}$$

This relation implies that the sequence of functions $(t \mapsto \mathbb{E}[X_t^n])_{n \geq 0}$ converges uniformly to $m := t \mapsto \mathbb{E}[X_t]$ on $[0, t_0]$ (we already knew that this sequence converges thanks to (4.10), but we explicit now its limit).
In order to verify that the limit (X, V) satisfies the system (4.3), we write, by definition of (X^{n+1}, V^{n+1}), for $t \leq t_0$:

$$\begin{cases}
X_t^{n+1} &= X_0 + \displaystyle\int_0^t V_s^{n+1} \mathrm{d}s \\
V_t^{n+1} &= V_0 - 2 \displaystyle\int_0^t V_{s^-}^{n+1} \int \mathbf{1}_{z \leq \lambda((X_s^{n+1} - \mathbb{E}[X_s^n]) V_{s^-}^{n+1}))} \mathcal{N}(\mathrm{d}z, \mathrm{d}s).
\end{cases}$$

The uniform convergences, the continuity of λ and the continuity properties of the Poisson process (see [Kingman, 1992] for instance) show that the limit process (X, V) indeed satisfies (4.3) on $[0, t_0]$. We have then the existence of a solution of (4.3) on $[0, t_0]$.
Let us now show the uniqueness on this interval. Let $(\tilde{X}, \tilde{V})$ be another solution of (4.3) with initial condition (X_0, V_0). We note $\tilde{m}_t = \mathbb{E}\left[\tilde{X}_t\right]$. Using again (4.5) we get, as previously, for $t \leq t_0$ with $t_0 = \frac{\alpha}{2K}$:

$$\|m - \tilde{m}\|_t \leq 2K \mathrm{e}^{\int_0^t h(s) \mathrm{d}s} \int_0^t \|m - \tilde{m}\|_s \mathrm{d}s \leq \delta \int_0^t \|m - \tilde{m}\|_s \mathrm{d}s,$$

with $\delta = 2KA_{t_0} \mathbb{E}\left[\mathrm{e}^{\alpha |X_0|}\right]$.
We deduce by Gronwall's lemma that $m = \tilde{m}$ on $[0, t_0]$, and then $(X, V) =$

$(\tilde{X}, \tilde{V})$ on $[0, t_0]$, hence the uniqueness of the solution.

To extend the existence and uniqueness on $\mathbb{R}_+$, we first do the same on the interval $[t_0, 2t_0]$ with initial condition (X_{t_0}, V_{t_0}), by translating the Poisson process by t_0 on the second coordinate. Let us notice that X_{t_0} satisfy condition $(\mathcal{H}_\alpha)$ since we have $|X_{t_0}|\leq |X_0|+t_0$. Then, we conclude by induction.

It remains to prove the continuity of the solution on the initial condition. Let thus (X, V) and $(\tilde{X}, \tilde{V})$ be the solutions of (4.3) with initial conditions (X_0, V_0) and $(\tilde{X}_0, \tilde{V}_0)$. One more time, using (4.5) gives :

$$
\begin{aligned}
&\|m - \tilde{m}\|_t \\
&\leq \mathbb{E}\left[\|X - \tilde{X}\|_t\right] \\
&\leq \mathbb{E}\left[\left(|X_0 - \tilde{X}_0|+|V_0 - \tilde{V}_0|+2K\int_0^t \|m - \tilde{m}\|_s \mathrm{d}s\right) \mathrm{e}^{\int_0^t h(s)\mathrm{d}s}\right] \\
&= \mathbb{E}\left[\left(|X_0 - \tilde{X}_0|+|V_0 - \tilde{V}_0|\right) \mathrm{e}^{\int_0^t h(s)\mathrm{d}s}\right] + 2K\mathbb{E}\left[\mathrm{e}^{\int_0^t h(s)\mathrm{d}s}\right] \int_0^t \|m - \tilde{m}\|_s \mathrm{d}s.
\end{aligned}
$$

With Gronwall's lemma we then get, for $t \leq t_0$:

$$
\begin{aligned}
\|m - \tilde{m}\|_t &\leq \mathbb{E}\left[\left(|X_0 - \tilde{X}_0|+|V_0 - \tilde{V}_0|\right) \mathrm{e}^{\int_0^t h(s)\mathrm{d}s}\right] \exp\left(2Kt\mathbb{E}\left[\mathrm{e}^{\int_0^t h(s)\mathrm{d}s}\right]\right) \\
&\leq A_{t_0}\mathbb{E}\left[\left(|X_0 - \tilde{X}_0|+|V_0 - \tilde{V}_0|\right) \mathrm{e}^{\alpha|X_0|}\right] \mathrm{e}^{\delta t_0},
\end{aligned}
$$

where A_t is defined by (4.8), and $\delta := 2KA_{t_0}\mathbb{E}\left[\mathrm{e}^{\alpha|X_0|}\right]$. Using now Relation (4.5) we get :

$$
\begin{aligned}
&\mathbb{E}\left[\|X - \tilde{X}\|_{t_0}+\|V - \tilde{V}\|_{t_0}\right] \\
&\quad\leq \mathbb{E}\left[\left(|X_0 - \tilde{X}_0|+|V_0 - \tilde{V}_0|\right.\right. \\
&\qquad\qquad\left.\left.+2Kt_0A_{t_0}\mathbb{E}\left[\left(|X_0 - \tilde{X}_0|+|V_0 - \tilde{V}_0|\right) \mathrm{e}^{\alpha|X_0|}\right] \mathrm{e}^{\delta t_0}\right) \mathrm{e}^{\int_0^{t_0} h(s)\mathrm{d}s}\right] \\
&\quad\leq A_{t_0}\mathbb{E}\left[\left(|X_0 - \tilde{X}_0|+|V_0 - \tilde{V}_0|\right) \mathrm{e}^{\alpha|X_0|}\right] \left(1 + 2Kt_0A_{t_0}\mathrm{e}^{\delta t_0}\mathbb{E}\left[\mathrm{e}^{\alpha|X_0|}\right]\right).
\end{aligned}
$$

Thus, we have obtained the conclusion of the theorem on $[0, t_0]$. Then we can get it on the interval $[0, nt_0]$ for any $n \geq 1$, the constants depending on n, by the same procedure. And the proof is complete. □

Remark 4.2.4. *Assumption $(\mathcal{H}_\alpha)$ is very strong and is due to the use of the rough estimation $|X_t|\leq |X_0|+t$. But we could hope that the result remains true for $X_0 \in L^1(\mathbb{R})$.*

Remark 4.2.5. *We obtained the existence and uniqueness of a solution of the system (4.3) for jump-rates that are Lipschitz. We can hope that this remains true for a jump-rate of the form $\lambda(z) = a\mathbf{1}_{z<0} + b\mathbf{1}_{z\geq 0}$, but our method does not work for discontinuous jump-rates.*

4.2.2 Propagation of chaos

We investigate now the convergence of the particle system towards the non-linear Zig-zag process, that is the solution of (4.3), when the number of particles goes to infinity. In other words, we are going to prove that if N particles are interacting, and start from i.i.d. initial conditions, when N tends to infinity, one of these particles behaves like the non-linear Zig-zag process. The method used here is classic : we couple N particles solutions of the particle system (4.2) with N independent copies of the non-linear Zig-zag process.

Let $\left((X^{i,N}, V^{i,N})\right)_{1\leq i\leq N}$ be the solutions of the particle system (4.2), with i.i.d. initial conditions $\left((X_0^{i,N}, V_0^{i,N})\right)$, $1 \leq i \leq N$, independent of the Poisson processes $\mathcal{N}^i$, $1 \leq i \leq N$.
Moreover, let $\left(\left(\overline{X}^i, \overline{V}^i\right)\right)_{1\leq i\leq N}$ be N independent copies of the non-linear Zig-zag process with initial conditions $(X_0^{i,N}, V_0^{i,N})$, $1 \leq i \leq N$: for $i \in \{1, \cdots, N\}$, $\left(\overline{X}^i, \overline{V}^i\right)$ is the solution of the non-linear system

$$\begin{cases} \mathrm{d}\overline{X}^i = \overline{V}^i \mathrm{d}t \\ \mathrm{d}\overline{V}^i = -2\overline{V}^i_{t^-} \int \mathbf{1}_{z\leq\lambda((\overline{X}^i_t - \overline{m}^i(t))\overline{V}_{t^-}))} \mathcal{N}^i(\mathrm{d}z, \mathrm{d}t) \\ \left(\overline{X}^i_0, \overline{V}^i_0\right) = \left(X_0^{i,N}, V_0^{i,N}\right), \end{cases} \tag{4.11}$$

where the Poisson measure $\mathcal{N}^i$ is the one that drives the EDS of $V^{i,N}$ and where $\overline{m}^i(t) = \mathbb{E}\left[\overline{X}^i_t\right]$, for any $i \in \{1, \cdots, N\}$.

Theorem 4.2.6. *Let us suppose that Assumption $(\mathcal{H}_\lambda)$ holds. Let $(X_0^{i,N}, V_0^{i,N})$, $1 \leq i \leq N$, be i.i.d. random vectors in $L^1(\mathbb{R} \times \{-1, +1\})$. Let us suppose that the random variables $\left(X_0^{i,N}\right)_{1\leq i\leq N}$ satisfy Assumption $(\mathcal{H}_\alpha)$.*
Then, for every $t \geq 0$, there exists a constant C_t not depending on N, such that for all $i \in \{1, \cdots, N\}$

$$\mathbb{E}\left[\left\|X^{i,N} - \overline{X}^i\right\|_t + \left\|V^{i,N} - \overline{V}^{i,N}\right\|_t\right] \leq \frac{C_t}{\sqrt{N}},$$

where the processes $\left((X^{i,N}, V^{i,N})\right)_{1\leq i\leq N}$ and $\left(\left(\overline{X}^i, \overline{V}^i\right)\right)_{1\leq i\leq N}$ are defined above.

Proof. Let us define $m^N_\cdot := t \mapsto \frac{1}{N}\sum_{k=1}^N X_t^{k,N}$ the empirical mean-position of the particle system.
Since m^N is random, we can not directly apply (4.5), which has been obtained for deterministic functions m and $\tilde{m}$. We thus return to Relation (4.6) , which gives :

$$\begin{aligned}
&\mathbb{E}_0\left[\|V^{1,N}-\overline{V}^1\|_t\right] \\
&\quad\leq 2K\int_0^t \mathbb{E}_0\left[\left|\left(X_s^{1,N}-m_s^N\right)\left(V_s^{1,N}-\bar{V}_s^1\right)\right.\right. \\
&\qquad\qquad\qquad\qquad\left.\left.+\ \bar{V}_s^1\left(X_s^{1,N}-\bar{X}_s^1-(m_s^N-\bar{m}_s^1)\right)\right|\right]\mathrm{d}s \\
&\quad\leq 2K\int_0^t\left(\mathbb{E}_0\left[\left(|X_0^{1,N}|+s+|m_s^N|\right)\|V^{1,N}-\bar{V}^1\|_s\right]\right. \\
&\qquad\qquad\qquad\qquad\left.+\mathbb{E}_0\left[\|X^{1,N}-\bar{X}^1\|_s\right]+\mathbb{E}_0\left[|m_s^N-\bar{m}_s^1|\right]\right)\mathrm{d}s,
\end{aligned}$$

where $\mathbb{E}_0\left[.\right]$ stands for the conditional expectation with respect to the initial conditions $\left((X_0^{k,N},V_0^{k,N})\right)_{1\leq k\leq N}$.
Thus, as previously, and using the definition of m^N we have :

$$\begin{aligned}
&\mathbb{E}_0\left[\|V^{1,N}-\overline{V}^1\|_t+\|X^{1,N}-\overline{X}^1\|_t\right] \\
&\leq 2K\int_0^t\mathbb{E}_0\left[|m_s^N-\overline{m}_s^1|\right]\mathrm{d}s \\
&\quad+\int_0^t\mathbb{E}_0\left[\left(1+2K+2K\left(|X_0^{1,N}|+s+\frac{1}{N}\sum_{k=1}^N|X_s^{k,N}|\right)\right)\right. \\
&\qquad\qquad\qquad\qquad\left.\left(\|V^{1,N}-\overline{V}^1\|_s+\|X^{1,N}-\overline{X}^1\|_s\right)\right]\mathrm{d}s \\
&\leq 2K\int_0^t\mathbb{E}_0\left[|m_s^N-\overline{m}_s^1|\right]\mathrm{d}s \\
&\quad+\int_0^t\left(1+2K\left(1+|X_0^{1,N}|+2s+\frac{1}{N}\sum_{k=1}^N|X_0^{k,N}|\right)\right) \\
&\qquad\qquad\qquad\qquad\mathbb{E}_0\left[\left(\|V^{1,N}-\overline{V}^1\|_s+\|X^{1,N}-\overline{X}^1\|_s\right)\right]\mathrm{d}s.
\end{aligned}\tag{4.12}$$

Let us now look at the term $\mathbb{E}_0\left[|m_s^N-\overline{m}_s^1|\right]$. We have, by exchangeability property of the particles :

$$\begin{aligned}
\mathbb{E}_0\left[|m_s^N-\overline{m}_s^1|\right]&=\mathbb{E}_0\left[\left|\frac{1}{N}\sum_{k=1}^N X_s^{k,N}-\mathbb{E}\left[\overline{X}_s^1\right]\right|\right] \\
&\leq\mathbb{E}_0\left[\frac{1}{N}\sum_{k=1}^N\left|X_s^{k,N}-\overline{X}_s^k\right|\right]+\mathbb{E}_0\left[\left|\frac{1}{N}\sum_{k=1}^N\overline{X}_s^k-\mathbb{E}\left[\overline{X}_s^1\right]\right|\right] \\
&\leq\mathbb{E}_0\left[\left|X_s^{1,N}-\overline{X}_s^1\right|\right]+\mathbb{E}_0\left[\left|\frac{1}{N}\sum_{k=1}^N\overline{X}_s^k-\mathbb{E}\left[\overline{X}_s^1\right]\right|\right].
\end{aligned}$$

Defining

$$F^N(s)=\mathbb{E}_0\left[\left|\frac{1}{N}\sum_{k=1}^N\overline{X}_s^k-\mathbb{E}\left[\overline{X}_s^1\right]\right|\right],$$

$$g(s) = 1 + 2K\left(2 + |X_0^{1,N}| + 2s + \frac{1}{N}\sum_{k=1}^{N}|X_0^{k,N}|\right),$$

and going back to (4.12), we get :

$$\begin{aligned}
&\mathbb{E}_0\left[\|V^{1,N} - \overline{V}^1\|_t + \|X^{1,N} - \overline{X}^1\|_t\right] \\
&\leq 2K\int_0^t F^N(s)\mathrm{d}s + \int_0^t \left(1 + 2K\left(1 + |X_0^{1,N}| + 2s + \frac{1}{N}\sum_{k=1}^{N}|X_0^{k,N}|\right) + 2K\right) \\
&\qquad\qquad \mathbb{E}_0\left[\|V^{1,N} - \overline{V}^1\|_s + \|X^{1,N} - \overline{X}^1\|_s\right]\mathrm{d}s \\
&= 2K\int_0^t F^N(s)\mathrm{d}s + \int_0^t g(s)\mathbb{E}_0\left[\|V^{1,N} - \overline{V}^1\|_s + \|X^{1,N} - \overline{X}^1\|_s\right]\mathrm{d}s.
\end{aligned}$$

Applying Gronwall's lemma we finally have :

$$\mathbb{E}_0\left[\|V^{1,N} - \overline{V}^1\|_t + \|X^{1,N} - \overline{X}^1\|_t\right] \leq 2K\mathrm{e}^{\int_0^t g(s)\mathrm{d}s}\int_0^t F^N(s)\mathrm{d}s.$$

Taking the expectation, and using the Hölder inequality we get :

$$\begin{aligned}
\mathbb{E}\left[\|V^{1,N} - \overline{V}^1\|_t + \|X^{1,N} - \overline{X}^1\|_t\right] &\leq 2K\mathbb{E}\left[\mathrm{e}^{\int_0^t g(s)\mathrm{d}s}\int_0^t F^N(s)\mathrm{d}s\right] \\
&\leq 2K\mathbb{E}\left[\mathrm{e}^{2\int_0^t g(s)\mathrm{d}s}\right]^{\frac{1}{2}}\mathbb{E}\left[\left(\int_0^t F^N(s)\mathrm{d}s\right)^2\right]^{\frac{1}{2}}.
\end{aligned}$$

Let us now see that $\mathbb{E}\left[\left(\int_0^t F^N(s)\mathrm{d}s\right)^2\right]^{\frac{1}{2}}$ tends to 0 when N tends to infinity, with speed $\sqrt{N}$. Using the fact that the $\overline{X}^k$, for $k \in \{1, \cdots, N\}$ are independent and identically distributed we have $\mathbb{E}\left[\frac{1}{N}\sum_{k=1}^N \overline{X}_s^k\right] = \mathbb{E}\left[\overline{X}_s^1\right]$, and Hölder inequality gives then

$$\begin{aligned}
F^N(s) &= \mathbb{E}_0\left[\left|\frac{1}{N}\sum_{k=1}^{N}\overline{X}_s^k - \mathbb{E}\left[\overline{X}_s^1\right]\right|\right] \\
&\leq \sqrt{\mathrm{Var}_0\left(\frac{1}{N}\sum_{k=1}^{N}\overline{X}_s^k\right)} \\
&= \frac{1}{\sqrt{N}}\sqrt{\mathrm{Var}_0\left(\overline{X}_s^1\right)},
\end{aligned}$$

where Var_0 is the conditional variance with respect to $\left((X_0^{k,N}, V_0^{k,N})\right)_{1\leq k\leq N}$. Therefore we get :

$$\mathbb{E}\left[\left(\int_0^t F^N(s)\mathrm{d}s\right)^2\right] = \mathbb{E}\left[\left(\int_0^t \mathbb{E}_0\left[\left|\frac{1}{N}\sum_{k=1}^{N}\overline{X}_s^k - \mathbb{E}\left[\overline{X}_s^1\right]\right|\right]\mathrm{d}s\right)^2\right]$$

$$\leq \frac{1}{N}\mathbb{E}\left[\left(\int_0^t \sqrt{\mathrm{Var}_0\left(\overline{X}_s^1\right)}\mathrm{d}s\right)^2\right].$$

Since $\overline{X}_s^1 \leq X_0^{1,N} + s$, and since $X_0^{1,N}$ satisfies Assumption (H_α), the quantity $\mathrm{Var}_0\left(\overline{X}_s^1\right)$ is bounded from above by a constant which is independent on N. Therefore, there exists a constant C_t, dependent on t by not on N, such that

$$\mathbb{E}\left[\left(\int_0^t F^N(s)\mathrm{d}s\right)^2\right] \leq \frac{C_t}{N}.$$

Moreover, using that the initial conditions are i.i.d., and using Hölder and Jensen inequalities, we have :

$$\begin{aligned}
&\mathbb{E}\left[\mathrm{e}^{2\int_0^t g(s)\mathrm{d}s}\right] \\
&\quad \leq \mathbb{E}\left[\mathrm{e}^{t+2Kt\left(2+|X_0^{1,N}|+\frac{1}{N}\sum_{k=1}^N |X_0^{k,N}|\right)+2Kt^2}\right] \\
&\quad \leq \mathrm{e}^{(1+4K)t+2Kt^2}\mathbb{E}\left[\mathrm{e}^{4Kt|X_0^{1,N}|}\right]^{\frac{1}{2}}\mathbb{E}\left[\mathrm{e}^{\frac{4Kt}{N}\sum_{k=1}^N |X_0^{k,N}|}\right]^{\frac{1}{2}} \\
&\quad = \mathrm{e}^{(1+4K)t+2Kt^2}\mathbb{E}\left[\mathrm{e}^{4Kt|X_0^{1,N}|}\right]^{\frac{1}{2}}\left(\prod_{k=1}^N \mathbb{E}\left[\mathrm{e}^{\frac{4Kt}{N}|X_0^{k,N}|}\right]\right)^{\frac{1}{2}} \\
&\quad = \mathrm{e}^{(1+4K)t+2Kt^2}\mathbb{E}\left[\mathrm{e}^{4Kt|X_0^{1,N}|}\right]^{\frac{1}{2}}\left(\mathbb{E}\left[\mathrm{e}^{\frac{4Kt}{N}|X_0^{1,N}|}\right]^N\right)^{\frac{1}{2}} \\
&\quad \leq \mathrm{e}^{(1+4K)t+2Kt^2}\mathbb{E}\left[\mathrm{e}^{4Kt|X_0^{1,N}|}\right]^{\frac{1}{2}}\mathbb{E}\left[\mathrm{e}^{4Kt|X_0^{1,N}|}\right]^{\frac{1}{2}} \\
&\quad = \mathrm{e}^{(1+4K)t+2Kt^2}\mathbb{E}\left[\mathrm{e}^{4Kt|X_0^{1,N}|}\right].
\end{aligned}$$

Therefore, for $t \leq t_0 := \frac{\alpha}{4K}$, the quantity $\mathbb{E}\left[\mathrm{e}^{2\int_0^t g(s)\mathrm{d}s}\right]$ is finite.
We deduce that for all $t \leq t_0$, there exists a constant, that we still denote by C_t, depending on t but not on N, such that

$$\mathbb{E}\left[\|V^{1,N} - \overline{V}^1\|_t + \|X^{1,N} - \overline{X}^1\|_t\right] \leq \frac{C_t}{\sqrt{N}}.$$

The extension to any arbitrary t can be established as it is done in the proof of Theorem 4.2.2. □

Remark 4.2.7. *The propagation of chaos is here not uniform in time as it can be obtained in other cases (see [Malrieu, 2003, Thai, 2015] for instance). Once again this is in particular due to the rough estimation* $|X_t| \leq |X_0| + t$.

Remark 4.2.8. *As for the existence and uniqueness of a solution of the system (4.3), we think that propagation of chaos holds also for a jump-rate of the form* $\lambda(z) = a\mathbf{1}_{z<0} + b\mathbf{1}_{z\geq 0}$, *but our proof does not work in that case.*

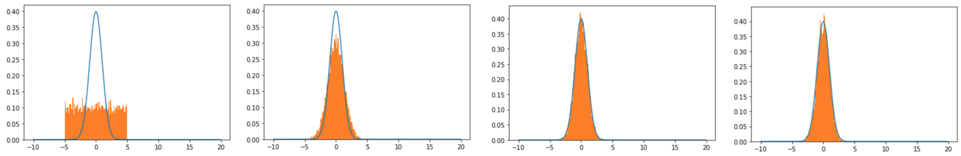


FIGURE 4.2 – Empirical distribution of X_t solution of (4.3) for $t \in \{0, 5, 10, 15\}$ with $\lambda(z) = 1 + z\mathbf{1}_{z>0}$. Simulation made from the evolution of 10000 particles solution of (4.2), with i.i.d. initial positions of law $\mathcal{U}([-5, 5])$ and i.i.d. initial speeds of law $Rad(0.5)$. The blue curve is $f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$.

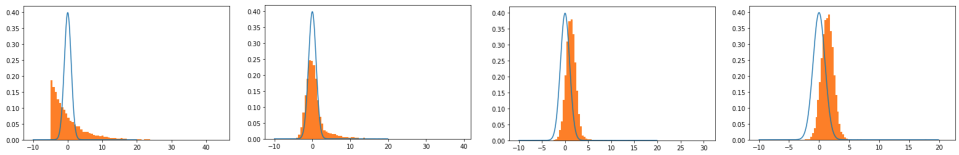


FIGURE 4.3 – Empirical distribution of X_t solution of (4.3) for $t \in \{0, 5, 10, 15\}$ with $\lambda(z) = 1 + z\mathbf{1}_{z>0}$. Simulation made from the evolution of 10000 particles solution of (4.2), with i.i.d. initial positions of law $\mathcal{E}(\frac{1}{5}) - 5$ and i.i.d. initial speeds of law $Rad(0.5)$. The blue curve is $f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$.

4.3 Long-time behaviour of the non-linear Zig-zag process

Our initial motivation is to investigate the long-time behaviour of the non-linear Zig-zag process.

On Figures 4.2, 4.3, 4.4, 4.5, 4.6 and 4.7 we have represented the empirical distribution of the first component X of the couple solution of (4.3), for different jump-rates and different initial conditions. This has been done from the simulation of 10000 particles solution of (4.3). We compare the empirical distribution of X to the density function proportional to $\exp\left(-\int_0^x (\lambda(y, +1) - \lambda(y, -1))\, dy\right)$. From these simulations, we can observe and conjecture the following facts : the law of X converges with time towards a shift of the law with density proportional to $\exp\left(-\int_0^x (\lambda(y, +1) - \lambda(y, -1))\, dy\right)$. This latter is in fact the first component of the invariant distribution of the one-dimensional Zig-zag process with jump-rate λ. The constant that gives the shift of this distribution seems to depend on the law of both the initial position and speed, and not only on their mean.

The behaviour of the non-linear Zig-zag process $((X_t, V_t))_{t\geq 0}$ depends heavily on the behaviour of its mean position $\mathbb{E}[X_t]$, which is not trivial to describe. Indeed, even if the simulations lead us to the conjecture that $\mathbb{E}[X_t]$ tends to a limit depending on the initial conditions, we did not manage to prove it for the moment. On the contrary, the mean of the non-linear diffusive process studied in [Malrieu, 2003] is trivial since it stays equal to zero with time.

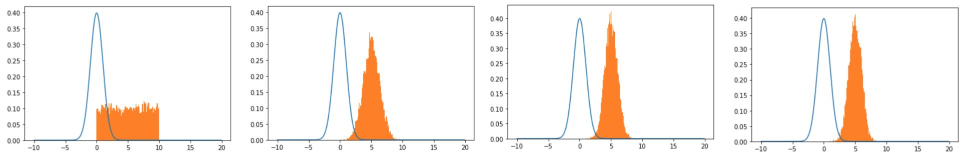


FIGURE 4.4 – Empirical distribution of X_t solution of (4.3) for $t \in \{0, 5, 10, 15\}$ with $\lambda(z) = 1 + z\mathbf{1}_{z>0}$. Simulation made from the evolution of 10000 particles solution of (4.2), with i.i.d. initial positions of law $\mathcal{U}([0, 10])$ and i.i.d. initial speeds of law $Rad(0.5)$. The blue curve is $f(x) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-x^2/2}$.

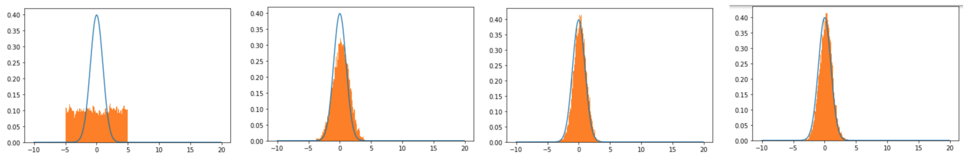


FIGURE 4.5 – Empirical distribution of X_t solution of (4.3) for $t \in \{0, 5, 10, 15\}$ with $\lambda(z) = 1 + z\mathbf{1}_{z>0}$. Simulation made from the evolution of 10000 particles solution of (4.2), with i.i.d. initial positions of law $\mathcal{U}([-5, 5])$ and initial speeds equal to 1. The blue curve is $f(x) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-x^2/2}$.

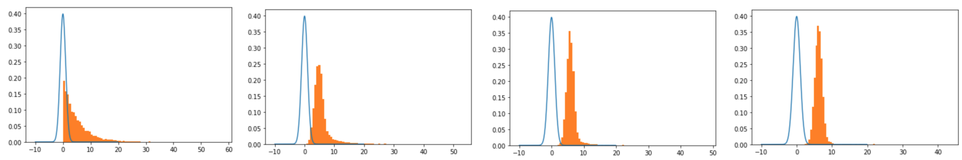


FIGURE 4.6 – Empirical distribution of X_t solution of (4.3) for $t \in \{0, 5, 10, 15\}$ with $\lambda(z) = 1 + z\mathbf{1}_{z>0}$. Simulation made from the evolution of 10000 particles solution of (4.2), with i.i.d. initial positions of law $\mathcal{E}(\frac{1}{5})$ and i.i.d. initial speeds of law $Rad(0.5)$. The blue curve is $f(x) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-x^2/2}$.

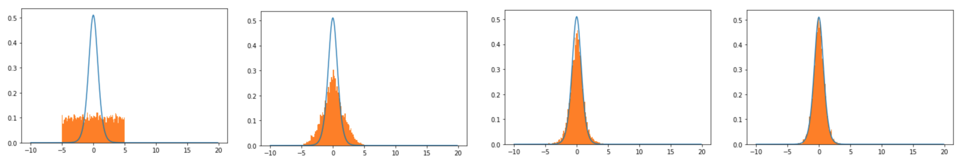


FIGURE 4.7 – Empirical distribution of X_t solution of (4.3) for $t \in \{0, 5, 10, 15\}$ with $\lambda(z) = \arctan(z) + 3$. Simulation made from the evolution of 10000 particles solution of (4.2), with i.i.d. initial positions of law $\mathcal{U}([-5, 5])$ and i.i.d. initial speeds of law $Rad(0.5)$. The blue curve is the density function proportional to $f(x) = \sqrt{1 + x^2} \mathrm{e}^{-2x \arctan(x)}$.

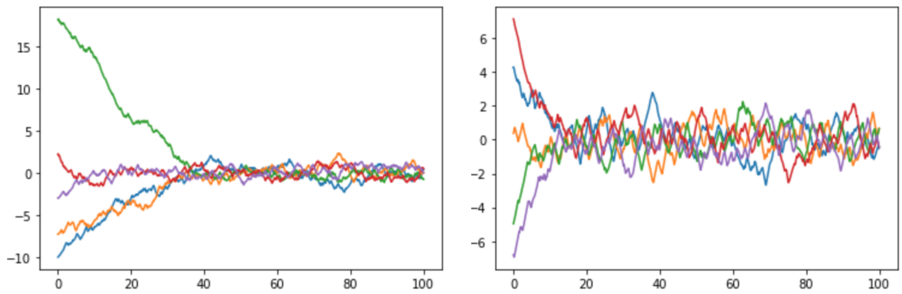


FIGURE 4.8 – Trajectories of five particles following (4.3), left with $\lambda(z) = \arctan(z) + 3$, and right with $\lambda(z) = 1 + z\mathbf{1}_{z>0}$.

An idea to simplify the problem is to work with a centred non-linear process and its associated centred particle system.

We thus consider new processes $\left(Y^{i,N}, V^{i,N}\right)$, $i \in \{1, \cdots, N\}$, with

$$Y_t^{i,N} = X_t^{i,N} - \frac{1}{N}\sum_{k=1}^{N} X_t^{k,N},$$

where the processes $\left(X^{i,N}, V^{i,N}\right)$, $i \in \{1, \cdots, N\}$, are the solutions of the system (4.2). The process $\left(Y^{i,N}, V^{i,N}\right)$ is therefore solution of the following system :

$$\begin{cases} \mathrm{d}Y_t^{i,N} &= \left(V_t^{i,N} - \frac{1}{N}\sum_{k=1}^{N} V_t^{k,N}\right)\mathrm{d}t \\ \mathrm{d}V_t^{i,N} &= -2V_{t^-}^{i,N}\int \mathbf{1}_{z\leq\lambda(Y_t^{i,N}V_{t^-}^{i,N}))}\mathcal{N}^i(\mathrm{d}z, \mathrm{d}t), \end{cases}$$

with $Y_0^{i,N} = X_0^{i,N} - \frac{1}{N}\sum_{k=1}^{N} X_0^{k,N}$.

We can observe on Figure 4.8 two trajectories of five particles following (4.3), with $\lambda(z) = \arctan(z) + 3$, and $\lambda(z) = 1 + z\mathbf{1}_{z>0}$. The initial positions are independent random variables with law $\mathcal{E}(\frac{1}{10}) - 10$, and the initial speeds are i.i.d. Rademacher variables with parameter $\frac{1}{2}$.
The following non-linear stochastic system is the expected limit of the centred particle system, when the number of particles increases :

$$\begin{cases} \mathrm{d}Y_t = (V_t - \mathbb{E}[V_t])\mathrm{d}t \\ \\ \mathrm{d}V_t = -2\int V_{t^-}\mathbf{1}_{z\leq\lambda(Y_tV_{t^-})}\mathcal{N}(\mathrm{d}z, \mathrm{d}t), \end{cases}$$

where $\mathcal{N}$ is a Poisson measure on $\mathbb{R}^+ \times \mathbb{R}^+$ with intensity the Lebesgue measure, and where the initial position Y_0 satisfies $\mathbb{E}[Y_0] = 0$ (since the initial position of $\left(Y^{i,N}, V^{i,N}\right)$ is centred by construction).
We name a solution (we will prove its existence and uniqueness thereafter) of this non-linear stochastic system the centred non-linear Zig-zag process.

As mentioned above, by construction, the initial positions of the centred particle system and non-linear process are centred. However, as long as it is possible, we will give results for the processes with general initial positions, not necessarily centred. We will thus in the sequel specify when the initial position is centred or not.

4.3.1 The centred non-linear Zig-zag process and propagation of chaos, for any initial position

In this section, we investigate the existence and uniqueness of a solution of the centred non-linear system, and the propagation of chaos of the centred particle system towards it, for any initial positions, that is not necessarily centred. We thus consider, for $i \in \{1, \cdots, N\}$, the couple $(Y^{i,N}, V^{i,N})$ solution of :

$$\begin{cases} \mathrm{d}Y_t^{i,N} = \left(V_t^{i,N} - \frac{1}{N}\sum_{k=1}^N V_t^{k,N}\right)\mathrm{d}t \\ \mathrm{d}V_t^{i,N} = -2V_{t-}^{i,N}\int \mathbf{1}_{z\le\lambda(Y_t^{i,N}V_{t-}^{i,N}))}\mathcal{N}^i(\mathrm{d}z,\mathrm{d}t) \\ (Y_0^{i,N}, V_0^{i,N}) \in L^1(\mathbb{R}\times\{-1,+1\}), \end{cases} \tag{4.13}$$

where the $\mathcal{N}^i$, $i = 1, \cdots, N$, are N independent Poisson measures on $\mathbb{R}_+^2$ with Lebesgue intensity measure, and independent of the initial conditions $\left((Y_0^{i,N}, V_0^{i,N})\right)_{1\le i\le N}$.
The expected limit EDS is the following :

$$\begin{cases} \mathrm{d}Y_t = (V_t - \mathbb{E}[V_t])\mathrm{d}t \\ \mathrm{d}V_t = -2\int V_{t-}\mathbf{1}_{z\le\lambda(Y_tV_{t-})}\mathcal{N}(\mathrm{d}z,\mathrm{d}t) \\ (Y_0, V_0) \in L^1(\mathbb{R}\times\{-1,+1\}), \end{cases} \tag{4.14}$$

where $\mathcal{N}$ is a Poisson measure on $\mathbb{R}^+\times\mathbb{R}^+$ with intensity the Lebesgue measure, independent of (Y_0, V_0).
To prove the existence and uniqueness of a solution of (4.14), process that we call the centred non-linear Zig-zag process, we use exactly the same method as for the non-centred process.

Proposition 4.3.1. *We assume $(\mathcal{H}_\lambda)$. For a (deterministic) measurable function $p : \mathbb{R}_+ \to [-1, 1]$ and a couple (Y_0, V_0) of random variables in $L^1(\mathbb{R}\times\{-1,+1\})$, we define $\Phi(Y_0, V_0, p) = ((Y_t, V_t))_{t\ge0}$ the solution of the linear stochastic differential equation*

$$\begin{cases} Y_t &= Y_0 + \int_0^t (V_s - p_s)\mathrm{d}s \\ V_t &= V_0 - 2\int_0^t V_{t-}\int \mathbf{1}_{z\le\lambda(Y_tV_{t-})}\mathcal{N}(\mathrm{d}z,\mathrm{d}t), \end{cases}$$

where $\mathcal{N}$ is a Poisson measure with Lebesgue intensity measure.
Then, for any such functions $p, \tilde{p}$ and couples of random variables (Y_0, V_0), $(\tilde{Y}_0, \tilde{V}_0)$, for any $T > 0$, we have

$$\mathbb{E}_0\left[\left\|\Phi(Y_0, V_0, p) - \Phi\left(\tilde{Y}_0, \tilde{V}_0, \tilde{p}\right)\right\|_T\right] \le C_0(T)\mathrm{e}^{\int_0^T h(s)\mathrm{d}s},$$

with

$$\begin{aligned} C_0(T) &= \left|Y_0 - \tilde{Y}_0\right| + \left|V_0 - \tilde{V}_0\right| + \int_0^T \|p - \tilde{p}\|_s \mathrm{d}s \\ h(T) &= 1 + 2K(|Y_0| + 2T + 1), \end{aligned}$$

where $\mathbb{E}_0[\cdot]$ *denotes the conditional expectation with respect to* $(Y_0, V_0, \tilde{Y}_0, \tilde{V}_0)$.

Proof. Since λ is Lipschitz continuous, the function Φ is well defined (see for instance [Gikhman et Skorohod, 1972]). We denote by $(Y_t, V_t) = \Phi(Y_0, V_0, p)(t)$ and $\left(\tilde{Y}_t, \tilde{V}_t\right) = \Phi\left(\tilde{Y}_0, \tilde{V}_0, \tilde{p}\right)(t)$.

We easily notice that

$$\mathbb{E}_0\left[\left\|Y - \tilde{Y}\right\|_t\right] \leq \left|Y_0 - \tilde{Y}_0\right| + \int_0^t \mathbb{E}_0\left[\left\|V - \tilde{V}\right\|_s\right]\mathrm{d}s + \int_0^t \|p - \tilde{p}\|_s \mathrm{d}s.$$

Let us introduce some notations to make the proof easier to read :

$$\begin{aligned} \lambda_{s^-} &:= \lambda\left(Y_s V_{s^-}\right), \quad \lambda_s := \lambda\left(Y_s V_s\right) \\ \tilde{\lambda}_{s^-} &:= \lambda\left(\tilde{Y}_s \tilde{V}_{s^-}\right), \quad \tilde{\lambda}_s := \lambda\left(\tilde{Y}_s \tilde{V}_s\right). \end{aligned}$$

We have

$$\begin{aligned} \left|V_t - \tilde{V}_t\right| &= \left|V_0 - \tilde{V}_0\right| \\ &\quad + \int_0^t \int \left(\left|V_s + \tilde{V}_s\right| - \left|V_s - \tilde{V}_s\right|\right) \mathbf{1}_{\min\{\lambda_{s^-}, \tilde{\lambda}_{s^-}\} \leq z \leq \max\{\tilde{\lambda}_{s^-}, \tilde{\lambda}_{s^-}\}} \mathcal{N}(\mathrm{d}z, \mathrm{d}s) \\ &\leq \left|V_0 - \tilde{V}_0\right| + 2\int_0^t \int \mathbf{1}_{\min\{\tilde{\lambda}_{s^-}, \tilde{\lambda}_{s^-}\} \leq z \leq \max\{\tilde{\lambda}_{s^-}, \tilde{\lambda}_{s^-}\}} \mathcal{N}(\mathrm{d}z, \mathrm{d}s). \end{aligned}$$

Then, since λ is a K-Lipschitz continuous function and for all $t \geq 0$, $|Y_t| \leq |Y_0| + 2t$,

$$\begin{aligned} &\mathbb{E}_0\left[\left\|V - \tilde{V}\right\|_t\right] \\ &\leq \left|V_0 - \tilde{V}_0\right| + 2\int_0^t \mathbb{E}_0\left[\left|\lambda_s - \tilde{\lambda}_s\right|\right]\mathrm{d}s \\ &\leq \left|V_0 - \tilde{V}_0\right| + 2K\int_0^t \mathbb{E}_0\left[\left|Y_s V_s - \tilde{Y}_s \tilde{V}_s\right|\right]\mathrm{d}s \\ &= \left|V_0 - \tilde{V}_0\right| + 2K\int_0^t \mathbb{E}_0\left[\left|Y_s(V_s - \tilde{V}_s) + (Y_t - \tilde{Y}_s)\tilde{V}_s\right|\right]\mathrm{d}s \\ &\leq \left|V_0 - \tilde{V}_0\right| + 2K\int_0^t (|Y_0| + 2s)\mathbb{E}_0\left[\left\|V - \tilde{V}\right\|_s\right]\mathrm{d}s + 2K\int_0^t \mathbb{E}_0\left[\left\|Y - \tilde{Y}\right\|_s\right]\mathrm{d}s \end{aligned}$$

Using Gronwall's Lemma, we deduce

$$\mathbb{E}_0\left[\left\|Y-\tilde{Y}\right\|_t+\left\|V-\tilde{V}\right\|_t\right]$$
$$\leq\left(\left|Y_0-\tilde{Y}_0\right|+\left|V_0-\tilde{V}_0\right|+\int_0^t\|p-\tilde{p}\|_s\mathrm{d}s\right)\mathrm{e}^{\int_0^t(1+2K(|Y_0|+2s+1))\mathrm{d}s}.$$

□

We observe in the previous proof that the centred process and the centred particle system can be studied in the same way than the initial non-linear process and the initial particle system. Consequently, we have similar results of existence and uniqueness for the centred non-linear equation (4.14) and the propagation of chaos by using the same methods than in Section 4.2.

Proposition 4.3.2. *If we assume that $(\mathcal{H}_\lambda)$ holds and Y_0 satisfies $(\mathcal{H}_\alpha)$, there is pathwise existence and uniqueness of the solution of (4.14).*

Moreover, let us consider $\left((Y^{i,N},V^{i,N})\right)_{1\leq i\leq N}$ the particle system solving (4.13). If the random vectors $\left(\left(Y_0^{i,N},V_0^{i,N}\right)\right)$, $i\in\{1,\cdots,N\}$ are independent and identically distributed, and if for $i\in\{1,\cdots,N\}$, $Y_0^{i,N}$ satisfies Assumption $(\mathcal{H}_\alpha)$, then there is propagation of chaos of the centred particle system towards the centred non-linear Zig-zag process, with speed $\sqrt{N}$.

4.3.2 Invariant distribution of the non-linear process, for any initial position

In this section, our goal is to get information on the invariant probability measures of the centred non-linear Zig-zag process solution of (4.14), that is for an initial position not necessarily centred.

Let $p\in[-1,1]$. We consider the following system

$$\begin{cases}\mathrm{d}Z_t=(W_t-p)\mathrm{d}t\\ \mathrm{d}W_t=-2\int W_{t^-}\mathbf{1}_{z\leq\lambda(Z_tW_{t^-})}\mathcal{N}(\mathrm{d}z,\mathrm{d}t),\end{cases}\tag{4.15}$$

where $\mathcal{N}$ is a Poisson measure on $\mathbb{R}^+\times\mathbb{R}^+$ with intensity the Lebesgue measure. This system looks like the centred non-linear system (4.14), but instead of the term $\mathbb{E}[V_t]$ for the evolution of Y_t, we have a constant $p\in[-1,1]$.
A process (Z,W) solution of (4.15) is a Markov process with infinitesimal generator $\mathcal{L}^p$ given by

$$\mathcal{L}^pf(z,w)=(w-p)\partial_zf(z,w)+\lambda(zw)\left(f(z,-w)-f(z,w)\right).$$

We almost recognize the infinitesimal generator of a classical one-dimension Zig-zag process attracted to the origin (under good assumption on the jump-rate λ), but here the derivative of the first component Y is not W but $W-p$.

Up to now, we made no assumption on the jump rate that models an attraction between the particles describing a group of collaborative bacteria. We make it now, in order to enable the ergodicity of our processes.

Assumption $(\mathcal{A})$ **:** There exist $y_0 > 0$ and $\lambda_{\min} > 0$ such that $\lambda(y) \geq \lambda_{\min}$ for all $y \geq y_0$ and such that

$$\overline{\lambda}_0 := \inf_{y \geq y_0} \lambda(y) > \sup_{y \leq -y_0} \lambda(y) =: \underline{\lambda}_0.$$

Assumption $(\mathcal{A})$ is classic in the study of the Zig-zag process, since it is a sufficient condition to prove its exponential ergodicity (see [Bierkens et Roberts, 2017, Fontbona *et al.*, 2012, Fontbona *et al.*, 2016]).

Proposition 4.3.3. *Let us assume Assumption* $(\mathcal{A})$ *and let* $p \in (-1, 1)$ *such that* $|p| < \frac{\overline{\lambda}_0 - \underline{\lambda}_0}{\overline{\lambda}_0 + \underline{\lambda}_0}$.
The process (Z, W) *solution of* (4.15) *has a unique invariant probability measure* ν_p *on* $\mathbb{R} \times \{-1, +1\}$*, which is given by :*

$$\nu_p(\mathrm{d}z, \mathrm{d}w) = \frac{1}{C_p} \mathrm{e}^{-F_p(z)} \mathrm{d}z \otimes \left(\frac{1+p}{2} \delta_1 + \frac{1-p}{2} \delta_{-1} \right) (\mathrm{d}w) \tag{4.16}$$

where

$$F_p(z) = \int_0^z \left(\frac{1}{1-p} \lambda(u) - \frac{1}{1+p} \lambda(-u) \right) \mathrm{d}u \quad \text{and} \quad C_p = \int_{\mathbb{R}} \exp(-F_p(z)) \mathrm{d}z.$$

Proof. Let us first remark that C_p is well defined when $|p| < \frac{\overline{\lambda}_0 - \underline{\lambda}_0}{\overline{\lambda}_0 + \underline{\lambda}_0}$. Indeed, we notice that, thanks to Assumption $(\mathcal{A})$,

$$\begin{aligned}
&\int_{y_0}^{\infty} \exp\left[- \int_0^z \left(\frac{1}{1-p} \lambda(u) - \frac{1}{1+p} \lambda(-u) \right) \mathrm{d}u \right] \mathrm{d}z \\
&= \exp\left[- \int_0^{y_0} \left(\frac{1}{1-p} \lambda(u) - \frac{1}{1+p} \lambda(-u) \right) \mathrm{d}u \right] \\
&\qquad\qquad \int_{y_0}^{\infty} \exp\left[- \int_{y_0}^z \left(\frac{1}{1-p} \lambda(u) - \frac{1}{1+p} \lambda(-u) \right) \mathrm{d}u \right] \mathrm{d}z \\
&\leq \exp\left[- \int_0^{y_0} \left(\frac{1}{1-p} \lambda(u) - \frac{1}{1+p} \lambda(-u) \right) \mathrm{d}u \right] \\
&\qquad\qquad \int_{y_0}^{\infty} \exp\left[- \int_0^z \left(\frac{1}{1-p} \overline{\lambda}_0 - \frac{1}{1+p} \underline{\lambda}_0 \right) \mathrm{d}u \right] \mathrm{d}z \\
&\leq \exp\left[- \int_0^{y_0} \left(\frac{1}{1-p} \lambda(u) - \frac{1}{1+p} \lambda(-u) \right) \mathrm{d}u \right] \\
&\qquad\qquad \int_{y_0}^{\infty} \exp\left[- \left(\frac{1}{1-p} \overline{\lambda}_0 - \frac{1}{1+p} \underline{\lambda}_0 \right) z \right] \mathrm{d}z,
\end{aligned}$$

which is finite for $-p < \frac{\overline{\lambda}_0-\underline{\lambda}_0}{\underline{\lambda}_0+\overline{\lambda}_0}$.
Similarly, since $\forall z \in \mathbb{R}$, $p \in (-1,1)$, $F_p(-z) = F_{-p}(z)$, the quantity

$$\int_{-\infty}^{-y_0} \exp\left[-\int_0^z \left(\frac{1}{1-p}\lambda(u) - \frac{1}{1+p}\lambda(-u)\right) \mathrm{d}u\right] \mathrm{d}z$$

is finite for $p < \frac{\overline{\lambda}_0-\underline{\lambda}_0}{\underline{\lambda}_0+\overline{\lambda}_0}$ and then C_p is well defined for $|p| < \frac{\overline{\lambda}_0-\underline{\lambda}_0}{\underline{\lambda}_0+\overline{\lambda}_0}$.

Let us now prove that ν_p is invariant for (Z, W). Let f be an arbitrary function on $\mathbb{R} \times \{-1,+1\}$ of class $\mathcal{C}^{1,0}$ with compact support. We have, for $z \in \mathbb{R}$:

$$\mathcal{L}f(z,1) = (1-p)\partial_z f(z,1) + \lambda(z)\left(f(z,-1) - f(z,1)\right)$$

and

$$\mathcal{L}f(z,-1) = -(1+p)\partial_z f(z,-1) + \lambda(-z)\left(f(z,1) - f(z,-1)\right)$$

so that

$$\begin{aligned}
&\frac{1+p}{2}\mathcal{L}f(z,1) + \frac{1-p}{2}\mathcal{L}f(z,-1) \\
&= \frac{1-p^2}{2}\left(\partial_z f(z,1) - \partial_z f(z,-1)\right) \\
&\qquad\qquad - \left(f(z,1) - f(z,-1)\right)\left(\frac{1+p}{2}\lambda(z) - \frac{1-p}{2}\lambda(-z)\right).
\end{aligned}$$

We can then easily verify that for any $\mathcal{C}^{1,0}$-functions f with compact support on $\mathbb{R} \times \{-1,+1\}$ we have

$$\begin{aligned}
&\int_{\mathbb{R}\times\{-1,+1\}} \mathcal{L}f(z,w)\nu_p(\mathrm{d}z,\mathrm{d}w) \\
&= \int_{\mathbb{R}} \frac{1}{C_p}\left(\frac{1+p}{2}\mathcal{L}f(z,1) + \frac{1-p}{2}\mathcal{L}f(z,-1)\right) \\
&\qquad\qquad \exp\left(-\int_0^z \left(\frac{1}{1-p}\lambda(u) - \frac{1}{1+p}\lambda(-u)\right)\mathrm{d}u\right)\mathrm{d}z \\
&= 0.
\end{aligned}$$

Therefore, ν_p is an invariant probability measure of the process (Z, W) solution of (4.15).
Moreover, this is its unique invariant probability measure since (Z, W) is positive Harris recurrent (this can be classically proved with Theorem 4.2 of [Meyn et Tweedie, 1993b]). $\square$

Remark 4.3.4. *We observe that if $(Z, W) \sim \nu_p$, we have $\mathbb{E}[W] = p$.*

Theorem 4.3.5. *Let us suppose that Assumptions $(\mathcal{H}_\lambda)$ and $(\mathcal{A})$ hold.
Let (Y, V) be solution of (4.14). Then, for all $p \in [-1,1]$ with $|p| < \frac{\overline{\lambda}_0-\underline{\lambda}_0}{\overline{\lambda}_0+\underline{\lambda}_0}$, the measure ν_p given by (4.16) is an invariant probability measure for (Y, V).*

Proof. Let $p \in [-1,1]$ with $|p| < \frac{\overline{\lambda}_0 - \underline{\lambda}_0}{\overline{\lambda}_0 + \underline{\lambda}_0}$.
Let $(Y_0, V_0) \sim \nu_p$ (we thus necessarily have $p = \mathbb{E}[V_0]$). Let us consider (Z, W) the solution of the stochastic system (4.15) with $(Z_0, W_0) = (Y_0, V_0)$. By Proposition 4.3.3, since ν_p is invariant for (4.15), we deduce that $p = \mathbb{E}[W_0] = \mathbb{E}[W_t]$ for all $t \geq 0$. We can thus rewrite the stochastic system (4.3) :

$$\begin{cases} \mathrm{d}Z_t = (W_t - p)\mathrm{d}t = (W_t - \mathbb{E}[W_t])\mathrm{d}t \\ \mathrm{d}W_t = -2\int W_{t^-}\mathbf{1}_{z \leq \lambda(Z_t W_{t^-})}\mathcal{N}(\mathrm{d}z, \mathrm{d}t). \end{cases}$$

Therefore, (Z, W) is solution of the non-linear system (4.3), and thus by uniqueness equal to (Y, V). Since ν_p is invariant for the process (Z, W), this proves that ν_p is invariant for (Y, V), that is for (4.3). □

Remark 4.3.6. *We can show a little bit more : let $\pi(\mathrm{d}y, \mathrm{d}w)$ be an invariant probability measure for the non-linear process (Y, V) solution of (4.3) such that $\int\int w\pi(\mathrm{d}y, \mathrm{d}w) = p$ with $|p| < \frac{\overline{\lambda}_0 - \underline{\lambda}_0}{\overline{\lambda}_0 + \underline{\lambda}_0}$. Then $\pi = \nu_p$.*
Indeed, since π is invariant, $\mathbb{E}[V_t] = p$ for all $t \geq 0$. Therefore, (Y, V) is solution of (4.15). Moreover, since (Y, V) is stationary by assumption, we deduce by Proposition 4.3.3 that $\pi = \nu_p$.

4.3.3 Conjectures on the long-time behaviour of the centred non-linear Zig-zag process with centred initial position

We now consider the centred non-linear Zig-zag process with centred initial position, that is the process (Y, V) solution of the following non-linear stochastic system :

$$\begin{cases} Y_t = Y_0 + \int_0^t (V_s - \mathbb{E}[V_s])\mathrm{d}s \\ V_t = V_0 - 2\int_0^t \int V_{s^-}\mathbf{1}_{z \leq \lambda(Y_s V_{s^-})}\mathcal{N}(\mathrm{d}z, \mathrm{d}s) \\ (Y_0, V_0) \in L^1(\mathbb{R} \times \{-1, +1\}) \text{ with } \mathbb{E}[Y_0] = 0, \end{cases} \tag{4.17}$$

where $\mathcal{N}$ is a Poisson measure on $\mathbb{R}^+ \times \mathbb{R}^+$ with intensity the Lebesgue measure. Our goal is to understand the long-time behaviour of this process. Let first do the following remark.

Remark 4.3.7. *Let us assume Assumption $(\mathcal{H}_\lambda)$. Let us consider (Y, V) the solution of (4.17), and let us suppose that Assumption $(\mathcal{A})$ holds.*
We can observe that among the probability measures ν_p defined by (4.16), with $p \in [-1,1]$ such that $|p| < \frac{\overline{\lambda}_0 - \underline{\lambda}_0}{\overline{\lambda}_0 + \underline{\lambda}_0}$, only ν_0 is invariant for (Y, V).
Indeed, we have $\mathbb{E}[Y_0] = 0$, and thus $\mathbb{E}[Y_t] = 0$ for all $t \geq 0$. Therefore, an invariant probability measure π for (Y, V) necessarily satisfies $\int y\pi(\mathrm{d}y, \mathrm{d}v) = 0$.
Let us see that only ν_0 satisfies this condition.
We note that $\int_{\mathbb{R}} z\mathrm{e}^{-F_0(z)}\mathrm{d}z = 0$, so that ν_0 indeed satisfies it. Moreover, since $C_p < \infty$, $\int_{\mathbb{R}} z\frac{\exp(-F_p(z))}{C_p}\mathrm{d}z = 0$ if and only if $\int_{\mathbb{R}} z\exp(-F_p(z))\mathrm{d}z = 0$. It is easy

to check that $p \mapsto \int_{\mathbb{R}} z\mathrm{e}^{-F_p(z)}\mathrm{d}z$ *is strictly decreasing for* $p \in \left(\frac{\underline{\lambda}_0-\overline{\lambda}_0}{\overline{\lambda}_0+\underline{\lambda}_0}, \frac{\overline{\lambda}_0-\underline{\lambda}_0}{\overline{\lambda}_0+\underline{\lambda}_0}\right)$. *We thus deduce that for* p *in this interval,* $\int_{\mathbb{R}} z\mathrm{e}^{-F_p(z)}\mathrm{d}z = 0$ *if and only if* $p = 0$.

Therefore, by this remark, we have reduced the possibilities for the invariant measures for the centred non-linear Zig-zag process with centred initial position. In fact, we conjecture the following result :

Conjecture $(\mathcal{C}_1)$ **:**

If Assumptions $(\mathcal{H}_\lambda)$ and $(\mathcal{A})$ are satisfied, the measure ν_0 given by (4.16) is the unique invariant probability measure of (Y, V).

To prove it, it is sufficient to prove that there exists $t_0 > 0$, $\tilde{y} > 0$ and $0 < \eta < \frac{\lambda(\tilde{y})-\lambda(-\tilde{y})}{\lambda(\tilde{y})+\lambda(-\tilde{y})}$ such that

$$|\mathbb{E}[V_t]| \leq \eta \quad \text{for all} \quad t \geq t_0. \tag{4.18}$$

Indeed, let us suppose that (4.18) holds. We have already seen that ν_0 is an invariant distribution for the process (Y, V) in Theorem 4.3.5. Moreover, let π be an invariant probability measure of (Y, V). By Remark 4.3.6, and since (4.18) holds, it ensures that $\pi = \nu_p$ for some p. And finally, Remark 4.3.7 gives that $\pi = \nu_0$ is the unique invariant distribution.
Thereby, a first goal is to prove that (4.18) holds.

Then, if we prove Conjecture $(\mathcal{C}_1)$, the next step is to prove the convergence of (Y, V) towards ν_0 :

Conjecture $(\mathcal{C}_2)$ **:**

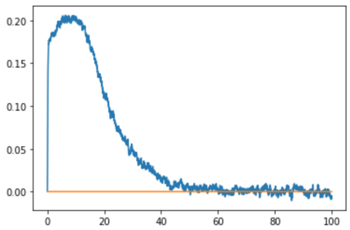
Under Assumptions $(\mathcal{H}_\lambda)$ and $(\mathcal{A})$, the distribution of (Y_t, V_t) converges towards ν_0 when t tends to infinity.

To prove it, it is sufficient to find a Lyapunov function for the infinitesimal generator of the process (Y, V). Indeed, the existence of the Lyapunov function and the uniqueness of the invariant measure ν_0 give the convergence of the process towards ν_0 by an argument of compactness. And in fact, a positive function g of class $\mathcal{C}^1$ in its first variable with

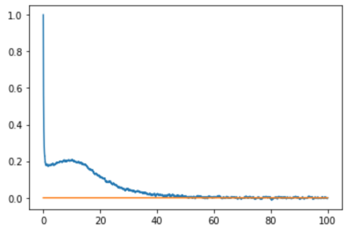
$$g(y, v) = \exp(\alpha|y| + \beta \mathrm{sgn}(yv)) \quad \text{for} \quad |y| \geq y_1$$

for some $y_1 > 0$, will indeed be a Lyapunov function for (Y, V) if relation (4.18) holds.

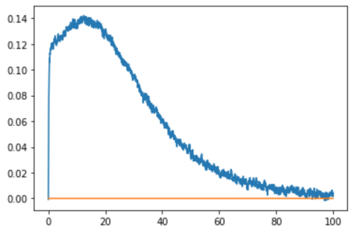
From the discussion above, we deduce that it would be sufficient to get relation (4.18) in order to conclude to the convergence of the centred non-linear Zig-zag process towards the probability measure ν_0. In fact, relation (4.18) seems to be true, since we observe on the simulations the convergence of p_t towards zero when t tends to infinity.
On Figure 4.9, we can observe four evolutions of p_t, for different jump-rates and initial conditions. In all cases, we observe the convergence of p_t towards zero, but not necessarily directly. Moreover, we observe that the speed of convergence of p_t depends heavily on the law of the initial positions. All these observations remain to be proved. The main difficulty to prove this is that our process is a couple, and we can not study the speed without taking into account the position.



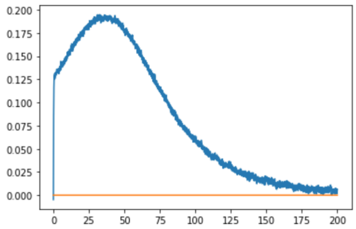
$\lambda(z) = 1 + z\mathbf{1}_{z>0}$,
initial law $\left(\mathcal{E}(\frac{1}{10}) - 10\right) \otimes Rad(\frac{1}{2})$



$\lambda(z) = 1 + z\mathbf{1}_{z>0}$,
initial law $\left(\mathcal{E}(\frac{1}{10}) - 10\right) \otimes \delta_1$



$\lambda(z) = \arctan(z) + 3$,
initial law $\left(\mathcal{E}(\frac{1}{10}) - 10\right) \otimes Rad(\frac{1}{2})$



$\lambda(z) = \arctan(z) + 3$,
initial law $\left(\mathcal{E}(\frac{1}{20}) - 20\right) \otimes Rad(\frac{1}{2})$

FIGURE 4.9 – Time evolution of $p_t = \mathbb{E}[V_t]$ with different jump-rates and initial conditions.

4.4 Prospects

We finally discuss of the prospects of this work, in addition of Conjectures $(\mathcal{C}_1)$ and $(\mathcal{C}_2)$.

Firstly, we have obtained the propagation of chaos for the initial particle system and the centred particle system, but not uniformly in time. We could thus try to have it, even if it does not seem to be simple with our computations. Moreover, the propagation of chaos and existence and uniqueness of the non-linear processes have been proved under Assumptions $(\mathcal{H}_\lambda)$ and $(\mathcal{H}_\alpha)$, but we hope to do better for both assumptions. Indeed, we guess that the results stay true for jump-rate of the form $\lambda(z) = a\mathbf{1}_{z<0} + b\mathbf{1}_{z\geq 0}$, and for an initial position in $L^1(\mathbb{R})$.
Then, if we indeed obtain the ergodicity of the centred non-linear Zig-zag process, we hope to be able to use it to get information on the initial non-linear Zig-zag process, that we also think to be ergodic as we have seen on Figures 4.2, 4.4, 4.5, 4.6 and 4.7.

Bibliographie

[Alsmeyer, 1997] ALSMEYER, G. (1997). The Markov renewal theorem and related results. *Markov Process. Related Fields*, 3(1):103–127.

[Andreis *et al.*, 2018] ANDREIS, L., DAI PRA, P. et FISCHER, M. (2018). McKean–Vlasov limit for interacting systems with simultaneous jumps. *Stoch. Anal. Appl.*, 36(6):960–995.

[Andrieu *et al.*, 2018] ANDRIEU, C., DURMUS, A., NÜSKEN, N. et ROUSSEL, J. (2018). Hypocoercivity of Piecewise Deterministic Markov Process-Monte Carlo. *arXiv e-prints*, page arXiv :1808.08592.

[Asmussen, 2003] ASMUSSEN, S. (2003). *Applied Probability and Queues.* Applications of mathematics : stochastic modelling and applied probability. Springer.

[Azaïs *et al.*, 2014] AZAÏS, R., BARDET, J.-B., GENADOT, A., KRELL, N. et ZITT, P.-A. (2014). Piecewise deterministic Markov process - recent results. *ESAIM : Proceedings*, 44:276–290. Clermont-Ferrand, France, 29-31 August 2012.

[Bierkens *et al.*, 2019] BIERKENS, J., FEARNHEAD, P. et ROBERTS, G. (2019). The zig-zag process and super-efficient sampling for Bayesian analysis of big data. *Ann. Statist.*, 47(3):1288–1320.

[Bierkens et Roberts, 2017] BIERKENS, J. et ROBERTS, G. (2017). A piecewise deterministic scaling limit of lifted Metropolis–Hastings in the Curie–Weiss model. *Ann. Appl. Probab.*, 27(2):846–882.

[Bierkens *et al.*, 2017] BIERKENS, J., ROBERTS, G. et ZITT, P.-A. (2017). Ergodicity of the zigzag process. *ArXiv e-prints.*

[Boender *et al.*, 1991] BOENDER, C., BOENDER, C., CARON, R., MCDONALD, J., RINNOOY KAN, A., RINNOOY KAN, A., ROMEIJN, H., SMITH, R., SMITH, R., TELGEN, J. et VORST, A. (1991). Shake-and-bake algorithms for generating uniform points on the boundary of bounded polyhedra. *Operations research*, 39(6):945–954.

[Bouchard-Côté *et al.*, 2018] BOUCHARD-CÔTÉ, A., VOLLMER, S. J. et DOUCET, A. (2018). The Bouncy Particle Sampler : a nonreversible rejection-free Markov chain Monte Carlo method. *J. Amer. Statist. Assoc.*, 113(522):855–867.

[Burkholder *et al.*, 1991] BURKHOLDER, D. L., PARDOUX, E. et SZNITMAN, A. (1991). *École d'Été de Probabilités de Saint-Flour XIX—1989*, volume 1464 de *Lecture Notes in Mathematics*. Springer-Verlag, Berlin. Papers from the school held in Saint-Flour, August 16–September 2, 1989, Edited by P. L. Hennequin.

[Calvez *et al.*, 2015] CALVEZ, V., RAOUL, G. et SCHMEISER, C. (2015). Confinement by biased velocity jumps : aggregation of *escherichia coli*. *Kinet. Relat. Models*, 8(4):651–666.

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

[Champagnat et Villemonais, 2017] CHAMPAGNAT, N. et VILLEMONAIS, D. (2017). Exponential convergence to quasi-stationary distribution for absorbed one-dimensional diffusions with killing. *ALEA Lat. Am. J. Probab. Math. Stat.*, 14(1):177–199.

[Chevallier, 2017] CHEVALLIER, J. (2017). Approximation par champ-moyen : le couplage à la Sznitman pour les nuls. *hal-01433292*.

[Cocozza-Thivent, 2018] COCOZZA-THIVENT, C. (2018). Processus de renouvellement markovien, Processus de Markov déterministes par morceaux. *hal-01418366*.

[Comets *et al.*, 2009] COMETS, F., POPOV, S., SCHÜTZ, G. M. et VACHKOVSKAIA, M. (2009). Billiards in a general domain with random reflections. *Arch. Ration. Mech. Anal.*, 191(3):497–537.

[Davis, 1984] DAVIS, M. H. A. (1984). Piecewise-deterministic Markov processes : a general class of nondiffusion stochastic models. *J. Roy. Statist. Soc. Ser. B*, 46(3):353–388. With discussion.

[Davis, 1993] DAVIS, M. H. A. (1993). *Markov models and optimization*, volume 49 de *Monographs on Statistics and Applied Probability*. Chapman & Hall, London.

[De Saporta *et al.*, 2015] DE SAPORTA, B., DUFOUR, F. et ZHANG, H. (2015). *Numerical Methods for Simulation and Optimization of Piecewise Deterministic Markov Processes*. Mathematics and statistics series. Wiley-ISTE.

[Deligiannidis *et al.*, 2019] DELIGIANNIDIS, G., BOUCHARD-CÔTÉ, A. et DOUCET, A. (2019). Exponential ergodicity of the bouncy particle sampler. *Ann. Statist.*, 47(3):1268–1287.

[Dieker et Vempala, 2015] DIEKER, A. B. et VEMPALA, S. S. (2015). Stochastic billiards for sampling from the boundary of a convex set. *Math. Oper. Res.*, 40(4):888–901.

[Down *et al.*, 1995] DOWN, D., MEYN, S. P. et TWEEDIE, R. L. (1995). Exponential and uniform ergodicity of Markov processes. *The Annals of Probability*, 23(4):1671–1691.

[Durmus *et al.*, 2018] Durmus, A., Guillin, A. et Monmarché, P. (2018). Geometric ergodicity of the bouncy particle sampler. *ArXiv e-prints*.

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

[Evans, 2001] Evans, S. N. (2001). Stochastic billiards on general tables. *Ann. Appl. Probab.*, 11(2):419–437.

[Fétique, 2017] Fétique, N. (2017). Long-time behaviour of generalised Zig-Zag process. *arXiv e-prints*, page arXiv :1710.01087.

[Fétique, 2019] Fétique, N. (2019). Explicit speed of convergence of the stochastic billiard in a convex set. *arXiv e-prints*, page arXiv :1901.02742.

[Fontbona *et al.*, 2012] Fontbona, J., Guérin, H. et Malrieu, F. (2012). Quantitative estimates for the long-time behavior of an ergodic variant of the telegraph process. *Adv. in Appl. Probab.*, 44(4):977–994.

[Fontbona *et al.*, 2016] Fontbona, J., Guérin, H. et Malrieu, F. (2016). Long time behavior of telegraph processes under convex potentials. *Stochastic Processes and their Applications*, 126(10):3077–3101.

[Fontbona *et al.*, 2009] Fontbona, J., Guérin, H. et Méléard, S. (2009). Measurability of optimal transportation and convergence rate for landau type interacting particle systems. *Probability Theory and Related Fields*, 143(3): 329–351.

[Fournier et Löcherbach, 2016] Fournier, N. et Löcherbach, E. (2016). On a toy model of interacting neurons. *Ann. Inst. Henri Poincaré Probab. Stat.*, 52(4):1844–1876.

[Gikhman et Skorohod, 1972] Gikhman, u. I. et Skorohod, A. V. (1972). *Stochastic differential equations*. Springer-Verlag, New York-Heidelberg. Translated from the Russian by Kenneth Wickwire, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 72.

[Graham et Robert, 2009] Graham, C. et Robert, P. (2009). Interacting multi-class transmissions in large stochastic networks. *Ann. Appl. Probab.*, 19(6):2334–2361.

[Hwang *et al.*, 2005] Hwang, C.-R., Hwang-Ma, S.-Y. et Sheu, S.-J. (2005). Accelerating diffusions. *Ann. Appl. Probab.*, 15(2):1433–1444.

[Ikeda et Watanabe, 2014] Ikeda, N. et Watanabe, S. (2014). *Stochastic Differential Equations and Diffusion Processes*. North-Holland Mathematical Library. Elsevier Science.

[Jacobsen, 2006] Jacobsen, M. (2006). *Point process theory and applications*. Probability and its Applications. Birkhäuser Boston, Inc., Boston, MA. Marked point and piecewise deterministic processes.

[Kac, 1956] Kac, M. (1956). Foundations of kinetic theory. pages 171–197.

[Kingman, 1992] KINGMAN, J. (1992). *Poisson Processes.* Oxford Studies in Probability. Clarendon Press.

[Lelièvre *et al.*, 2013] LELIÈVRE, T., NIER, F. et PAVLIOTIS, G. A. (2013). Optimal non-reversible linear drift for the convergence to equilibrium of a diffusion. *J. Stat. Phys.*, 152(2):237–274.

[Lemaire *et al.*, 2018] LEMAIRE, V., THIEULLEN, M. et THOMAS, N. (2018). Exact simulation of the jump times of a class of piecewise deterministic Markov processes. *J. Sci. Comput.*, 75(3):1776–1807.

[Lindvall, 2002] LINDVALL, T. (2002). *Lectures on the coupling method.* Dover Books on Mathematics Series. Dover Publications, Incorporated.

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

[McKean, 1967] MCKEAN, H. (1967). Propagation of chaos for a class of non-linear PDE's. *Lecture series in differential equations 7*, pages 41–57.

[Meyn et Tweedie, 1993a] MEYN, S. P. et TWEEDIE, R. L. (1993a). *Markov chains and stochastic stability.* Springer-Verlarg, London.

[Meyn et Tweedie, 1993b] MEYN, S. P. et TWEEDIE, R. L. (1993b). Stability of markovian processes iii : Foster-Lyapunov criteria for continuous-time processes. *Advances in Applied Probability*, 25(3):518–548.

[Monmarché, 2016] MONMARCHÉ, P. (2016). Piecewise deterministic simulated annealing. *ALEA Lat. Am. J. Probab. Math. Stat.*, 13(1):357–398.

[Monmarché, 2018] MONMARCHÉ, P. (2018). Elementary coupling approach for non-linear perturbation of Markov processes with mean-field jump mechanims and related problems. *arXiv e-prints*, page arXiv :1809.10953.

[Othmer *et al.*, 1988] OTHMER, H. G., DUNBAR, S. R. et ALT, W. (1988). Models of dispersal in biological systems. *Journal of Mathematical Biology*, 26:263–298.

[Thai, 2015] THAI, M.-N. (2015). Birth and Death process in mean field type interaction. *ArXiv e-prints.*

[Thorisson, 2000] THORISSON, H. (2000). *Coupling, stationarity, and regeneration.* New York ; London : Springer.

[Wu et Robert, 2018] WU, C. et ROBERT, C. P. (2018). The Coordinate Sampler : a non-reversible Gibbs-like MCMC sampler. *ArXiv e-prints.*

Résumé :

L'objet de cette thèse est l'étude de certains processus de Markov déterministes par morceaux (PDMPs), et plus particulièrement de leur comportement en temps long. Pour cela, nous utilisons des méthodes de couplage. Dans un premier chapitre, nous nous intéressons à un PDMP appelé "billard stochastique", décrivant le mouvement d'une particule dans un convexe borné du plan. Nous donnons, dans des cas particuliers de convexes, une borne explicite sur la vitesse de convergence à l'équilibre du processus. Dans un deuxième chapitre, nous étudions l'ergodicité d'un PDMP que l'on a appelé "processus Zig-zag généralisé", pouvant décrire le mouvement linéaire par morceaux d'une bactérie attirée par un nutriment fixe dans son environnement. Enfin, dans un dernier chapitre, nous étudions un système de particules, dont chaque particule est un processus Zig-zag (cas particulier du processus étudié dans le deuxième chapitre), non plus attiré par un nutriment, mais par la moyenne spatiale du système de particules. Nous étudions la propagation du chaos de ce système de particules, ainsi que le comportement en temps long du processus limite.

Mots clés :

Processus de Markov déterministes par morceaux ; Comportement en temps long ; Vitesse de convergence ; Méthode de couplage ; Système de particules en interaction ; Propagation du chaos.

Abstract :

The purpose of this PhD thesis is the study of some Piecewise deterministic Markov processes (PDMPs), and in particular of their long-time behaviour. For that, we use coupling methods. In a first chapter, we are interested in a PDMP called "stochastic billiard", that describes the movement of a particle in a bounded convex set of the plan. In particular cases of convex sets, we give an explicit bound for the speed of convergence of the process. In a second chapter, we study a PDMP that we have named "generalised Zig-zag process", that can describe the piecewise linear movement of a bacteria which is attracted by a fixed nutriment in its environment. Finally, in a last chapter, we study a particle system in which each particle is a Zig-zag process (particular case of the process studied in the second chapter), attracted by the spatial mean of the particle system. We study the propagation of chaos of this particle system, and the long-time behaviour of the limit process.

Keywords :

Piecewise deterministic Markov processes ; Long-time behaviour ; Speed of convergence ; Coupling method ; Interacting particle system ; Propagation of chaos.