A new way of study for meta-stability for Hamiltonian Systems

Michail Papadopoulos

July 11, 2023

Abstract

Mark Freidlin and Alexander D. Wentzell, through their study of general applications of stochastic mathematics on dynamical systems, have proven that by perturbing a Hamiltonian system with use of a white noise, we can approximate this systems dynamics with an appropriate diffusion on a respective graph, defined by the stability points of the Hamiltonian.The main purpose of the present diploma thesis is the proof of a result for the stochasticaly perturbed system of the harmonic oscillator, already covered by the classical Freidlin-Wentzell theory for Hamiltonian systems, but using a different approach, inspired by [**1**]. Namely, we will show the tightness of the distribution of the Hamiltonians of the perturbed solutions and later on we prove that any potential limit of these distributions can have a unique characterization through its respective behavior in terms of the martingale problem.

Abstract

Οι Mark Freidlin και Alexander D. Wentzell, μελετώντας γενικότερα εφαρμογές των στοχαστικών μαθηματικών σε δυναμικά συστήματα , έχουν αποδείξει ότι εάν διαταράξουμε ένα Χαμιλτονιανό σύστημα με χρήση λευκού θορύβου (white noise), μπορούμε να προσεγγίσουμε αυτή τη δυναμική για πολύ μικρές διαταραχές με μια διάχυση σε έναν κατάλληλο γράφο, ο οποίος προσδιορίζεται από τα σημεία ισορροπίας της Χαμιλτονιανής.Σκοπός της παρούσας διπλωματικής είναι η απόδειξη ενός αποτελέσματος αναφορικά με το στοχαστικά διαταραγμένο σύστημα του αρμονικού ταλαντωτή, το οποίο ήδη καλύπτεται μεν απ'τη κλασική θεωρία Freidlin-Wentzell για Χαμιλτονιανά συστήματα, αλλά με χρήση διαφορετικής προσέγγισης, η οποία έχει εμπνευστεί από το [**1**]. Ειδικότερα, θα αποδείξουμε τη σφιχτότητα των κατανομών της Χαμιλτονιανής για τις λύσεις του διαταραγμένου προβλήματος και ύστερα προβαίνουμε στο χαρακτηρισμό του ορίου μέσω της αντίστοιχης συμπεριφοράς των διαδικασιών ως προς το πρόβλημα Martingale.

Contents

Introduction

In a relatively recent paper [**1**] there was a new method introduced for the study of metastability. Namely, what was carried out was essentially studying the metastability of Markov chains by defining the so-called resolvent problem and by studying the limits of this problem's solutions- by means of the martingale problem- it was possible, eventually, to derive necessary and sufficient conditions for the metastable behaviour of a Markov chain with generator $\mathcal{L}$.

In this thesis, we will strive towards applying a similar reasoning to studying the metastability of continuous time processes and, in particular, in the case of Hamiltonian dynamical systems, that is systems that are defined by a Hamiltonian function $H(x)$ (usually-but not always- expressing the total amount of energy in the system). To that end, we will focus on the example of the harmonic oscillator to illustrate a first approach on such a method. Below follows the description of this problem's setting:

Let $\tilde{W}^1, \tilde{W}^2$ be two uncorrelated, standard Brownian motion. If q, p stand for the position and the momentum, respectively, of the harmonic oscillator, then the Hamiltonian for the harmonic oscillator will be $H(q,p) = \frac{q^2+p^2}{2}$ and so we consider the following system of (stochastic) differential equations:

$$\begin{cases} d\tilde{q}_t^\varepsilon = & \frac{\partial H}{\partial p}(\tilde{q}_t^\varepsilon, \tilde{p}_t^\varepsilon)dt + \varepsilon d\tilde{W}_t^1 = & \tilde{p}_t^\varepsilon dt + \varepsilon d\tilde{W}_t^1 \\ d\tilde{p}_t^\varepsilon = & -\frac{\partial H}{\partial q}(\tilde{q}_t^\varepsilon, \tilde{p}_t^\varepsilon)dt + \varepsilon d\tilde{W}_t^2 = & -\tilde{q}_t^\varepsilon dt + \varepsilon d\tilde{W}_t^2 \end{cases}.$$

This can also be summarised in vector notation as $d\tilde{x}_t^\varepsilon = b(\tilde{x}_t^\varepsilon)dt + \varepsilon d\tilde{W}_t$, where $\tilde{x}_t^\varepsilon = (\tilde{q}_t^\varepsilon, \tilde{p}_t^\varepsilon)$,

$b(\tilde{x}_t^\varepsilon) = (\tilde{p}_t^\varepsilon, -\tilde{q}_t^\varepsilon)$ and $\tilde{W}_t = (\tilde{W}_t^1, \tilde{W}_t^2)$. Our goal is to determine the behavior of the resulting process as $\varepsilon \to 0$.

Regarding the physics behind this system, we know that its unperturbed version would be moving across a circle in the phase space. However, due to the fact that we've introduced the stochastic perturbation, we will have a movement across different circles as time goes by. The way in which this happens is precisely what interests us, and so we need to rescale time in order to keep track of these changes, since the movement between circles (level sets in general) is much more slow than that across a particular circle. Hence, we focus on the process $x_t^\varepsilon = \tilde{x}^\varepsilon_{\frac{t}{\varepsilon^2}}$, with the Brownian motions introduced in the beginning undergoing the same time change, i.e. we work with $W^i = \tilde{W}^i_{\frac{t}{\varepsilon^2}}, i = 1, 2$. Hence, the resulting dynamical system- on which we will work from this point and on- is:

$$\begin{cases} dq_t^\varepsilon = & \frac{1}{\varepsilon^2} \cdot p_t^\varepsilon dt + W_t^1 \\ d\tilde{p}_t^\varepsilon = & -\frac{1}{\varepsilon^2} \cdot q_t^\varepsilon dt + W_t^2 \end{cases}.$$

What we will attempt is to see the behaviour of the dynamical system as $\varepsilon \to 0$. Before giving a description of the classical Freidlin Wentzell (FW) theory and our own method, we will dedicate the following two chapters to review the necessary tools in probability theory and stochastic differential equations, in order to proceed to the solution of the problem with both theories.

Chapter 1

Preliminaries on Probability Theory

We present a necessary introduction to the basic probabilistic tools that will be useful in the chapters to come.

Definition 1.0.1. Let $X \neq \emptyset$. A σ-algebra $\mathcal{F}$ over X is a collection of subsets of X (i.e. $\mathcal{F} \subset \mathcal{P}(X)$ where $\mathcal{P}(X)$ is the powerset of X) such that the following properties hold:

- $\emptyset,\ X \in \mathcal{F}$
- If $A \in \mathcal{F} \Rightarrow\ X \backslash A \in \mathcal{F}$
- For every sequence of sets $\{A_n\}_{n\in\mathbb{N}} \subset \mathcal{F}$ we have $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$

The notion of a σ-algebra is essential to the following:

Definition 1.0.2. Let $X \neq \emptyset$ and let $\mathcal{F}$ be a σ-algebra over X. A measure is a set function $\mu : \mathcal{F} \to [0, \infty]$ such that:

- $\mu(\emptyset) = 0$

- For every sequence of disjoint sets $\{A_n\}_{n\in\mathbb{N}} \subset \mathcal{F}$:

$$\mu(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n)$$

In addition, if $\mu(X) = 1$ then μ is called a probability measure. The triad $(X, \mathcal{F}, \mu)$ is called a measure space. In the case of a probability measure, the triad $(X, \mathcal{F}, \mu)$ (usually we will use the symbol Ω instead of X for the space) will be called a probability space.

For what will follow we shall not develop in full detail the basic properties and definitions pertaining measures and Lebesgue integration, but rather we shall take all this results as granted (refering to [**2**], [**3**]).

The set upon which everything is being built - σ-algebra, measure etc.- is what which we refer to as sample space Ω in probability. The σ-algebra is comprised of all possible events in a random experiment. If we define $X : \Omega \to \mathbb{R}$ then X is random variable if and only if it is a measurable function (i.e. inverse images of Borel subsets of $\mathbb{R}$ are events/ elements of $\mathcal{F}$). What is known to us as the expectation of a random variable is simply the Lebesgue integral of that random variable in terms of the probability measure:

$$\mathbb{E}[X] = \int_{\Omega} X(\omega) d\mathbb{P}(\omega)$$

We also define the variance of a random variable as:

$$\mathbb{V}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2]$$

whereas for X,Y random variables their covariance is defined as:

$$cov(X,Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])]$$

Definition 1.0.3 (π system,Dynkin System)**.** (i)A class $\mathcal{D}$ of subsets of Ω is called a Dynkin system if the following properties hold:

- $\Omega \in \mathcal{D}$
- If $A \subset B,\ A, B \in \mathcal{D}$ then $B \backslash A \in \mathcal{D}$
- If $\{A_n\}_{n \in \mathbb{N}}$ is an increasing family of sets in $\mathcal{D}$, then $\cup_{n=1}^{\infty} A_n \in \mathcal{D}$

(ii) A class $\mathcal{G}$ of subsets of Ω is called a π system if it's closed under intersections, i.e. if $A, B \in \mathcal{G}$ then $A \cap B \in \mathcal{G}$.

Theorem 1.0.1 (Dynkin's π-λ theorem)**.** *Let $\mathcal{C} \subset \mathcal{P}(\Omega)$ be a family which is closed under finite intersections (in other words, a π system). Then the dynkin system generated by $\mathcal{C}$ coincides with the σ algebra generated by $\mathcal{C}$, or $\delta(\mathcal{C}) = \sigma(\mathcal{C})$. In particular, a dynkin system that is a π system at the same moment is a σ algebra.*

Several important and very useful theorems regarding Lebesgue integration and convergence are the following:

Theorem 1.0.2 (Lebesgue's monotone convergence theorem)**.** *Let $\{X_n\}$ be a sequence of non-negative random variables with $X_n \uparrow$. Let $X = \lim_{n\to\infty} X_n$ (the pointwise limit). Then*

$$\lim_{n\to\infty} \int X_n d\mathbb{P} = \int X d\mathbb{P}$$

or, using the expectation notation:

$$\lim_{n\to\infty} \mathbb{E}[X_n] = E[X]$$

Lemma 1.0.3 (Fatou)**.** *Let X_n be a sequence of non-negative random variables. Then*

$$\int \liminf_{n\to\infty} X_n d\mathbb{P} \leq \liminf_{n\to\infty} \int X_n d\mathbb{P}$$

Theorem 1.0.4 (Lebesgue's dominated convergence theorem)**.** *Let X_n be a sequence of real-valued random variables such that* $\lim X_n = X$ *almost surely, i.e.* $\mathbb{P}[X_n \to X] = 1$ *and let there be a non-negative random variable Y such that $|X_n| \leq Y$ and $\mathbb{E}[|Y|] < \infty$. Then it is also $\mathbb{E}[|X|] < \infty$ and*

$$\lim_{n\to\infty} \mathbb{E}[X_n] = \mathbb{E}[X]$$

There will be instances where we have multiple integrations and, in those cases, we will require changing the order of integration. We will include the basic result for this process, the Fubini theorem.

Definition 1.0.4. We call a measurable rectangle in $X \times Y$ any set of the form $A \times B$ where A is a Borel set of X and B a Borel set of Y. The product $\sigma-$algebra is the $\sigma-$algebra generated by such rectangles, namely (denote the Borel sets of X, Y as $\mathcal{B}(X), \mathcal{B}(Y)$ respectively):

$$\mathcal{A} \otimes \mathcal{B} = \sigma(\{A \times B : A \in \mathcal{B}(X), B \in \mathcal{B}(Y)\})$$

It can be proven that if μ is a measure on $(X, \mathcal{B}(X))$ and ν a measure on $(Y, \mathcal{B}(Y))$, then there exists a unique measure m on the product space such that $m(A \times B) = \mu(A)\nu(B)$. This unique measure is called the product measure of μ, ν and is usually denoted as $\mu \otimes \nu$.

Theorem 1.0.5 (Fubini)**.** *Let $(X, \mathcal{B}(X), \mu)$, $(Y, \mathcal{B}(Y), \nu)$ be two spaces of $\sigma-$finite measure (this means there exists a sequence of sets such that their union is the entire space and each member of this sequence is of finite measure). Let $(X \times Y, \mathcal{B}(X) \otimes \mathcal{B}(Y), \mu \otimes \nu)$ be the product space and $f : X \times Y \to [-\infty, \infty]$ a measurable function. If $\int |f(x, y)| d(\mu \otimes \nu)(x, y) < \infty$, then*

$$\int f(x,y)d(\mu \otimes \nu)(x,y) = \int \left(\int f(x,y)d\nu(y) \right) d\mu(x) = \int \left(\int f(x,y)d\mu(x) \right) d\nu(y)$$

One of the main concepts we require in this diploma thesis is that of convergence of probability measures. We will not require any other probability-theoretic concepts of convergence in this text.

Definition 1.0.5. Let $\{\mathbb{P}^\varepsilon\}_{\varepsilon>0}$ be a family of probability measures on some space endowed with some topology. Then $\mathbb{P}^\varepsilon \xrightarrow{\varepsilon \to 0} \mathbb{P}$ if for any Borel set $\mathbb{P}^\varepsilon[A] \to \mathbb{P}[A]$ for any Borel set A. We will also say that if $\{X^\varepsilon\}_{\varepsilon>0}$ is a family of random variables, then X^ε converges weakly or in distribution to X (symb. $X^\varepsilon \Rightarrow X$ or $X^\varepsilon \xrightarrow{d} X$) if $\mathbb{P}^{X^\varepsilon} \to \mathbb{P}^X$

Additionally, we present a theorem particular to finite measure spaces and probability spaces

Corollary 1.0.5.1 (Lebesgue's Bounded Convergence theorem)**.** *Let X_n be a sequence of random variables,* $\lim_{n\to\infty} X_n = X$ *and $|X_n| \leq M < \infty$ where M is a constant. Then $\mathbb{E}[|X|] < \infty$ and*

$$\lim_{n\to\infty} \mathbb{E}[X_n] = \mathbb{E}[X]$$

Theorem 1.0.6. *Let $\{X^\varepsilon\}$ be a family of random variables with $X^\varepsilon : \Omega^\varepsilon \to (S, \mathcal{B}(S))$ ($\mathcal{B}(S)$: the Borel sets of S), then $X^\varepsilon \Rightarrow X$ iff :*

$$\mathbb{E}[f(X^\varepsilon)] \to \mathbb{E}[f(X)]$$

for any $f \in C_b(S)$ bounded and continuous function on S.

While this is an important characterization for weak convergence, there are times (as we'll see later on in this text) when this condition maybe prove impractical. We can, however, take advantage of another description of weak convergence:

Definition 1.0.6. Let $A \in \mathcal{S} = \mathcal{B}(S)$, for some metric space S. Then A is called a $\mathbb{P}-$continuity set if $\mathbb{P}[\partial A] = 0$.

Theorem 1.0.7 (Portmanteau's theorem)**.** *Let $\{\mathbb{P}_n\}_{n\in\mathbb{N}}$ be a sequence of probability measures.The following statements are equivalent:*

1. $\mathbb{P}_n \to \mathbb{P}$
2. $\int f(x)\mathbb{P}_n(dx) \to \int f(x)\mathbb{P}(dx), \forall f$ *bounded and continuous*
3. $\limsup_{n\to\infty} \mathbb{P}_n[F] \leq \mathbb{P}[F],\ \forall F$ *closed*
4. $\liminf_{n\to\infty} \mathbb{P}_n[G] \geq \mathbb{P}[G,\ \forall G$ *open*$]$
5. $\lim_{n\to\infty} \mathbb{P}_n[A] = \mathbb{P}[A],\ \forall A\ \mathbb{P}$ *-continuity sets*

The last condition in particular will prove very useful for our approach later on.

Theorem 1.0.8. *Let X^ε, X random variables of a metric spaces S. Let $h : S \to S'$ be a $S/S'-$measurable mapping and let D_h be the set of this function's discontinuity points. If $X^\varepsilon \to X$ and $\mathbb{P}[X \in D_h] = 0$ then $h(X^\varepsilon) \Rightarrow h(X)$, i.e. if $\mathbb{P}^\varepsilon \Rightarrow \mathbb{P}$ and $\mathbb{P}[D_h]$ then $\mathbb{P}^\varepsilon h^{-1} \Rightarrow \mathbb{P}h^{-1}$.*

Another integral concept for our study is that of tightness:

Definition 1.0.7. A family of probability measures Π on $(S, \mathcal{S})$ is called tight if for every $\varepsilon > 0$ there exists a $K \subset S$ compact such that $\mathbb{P}[K] > 1 - \varepsilon,\ \forall\ \mathbb{P} \in \Pi$

Definition 1.0.8. A family of probability measures Π on $(S, \mathcal{S})$ is called relatively compact if any sequence of its elements contains a weakly convergent subsequence. The limiting probability measures might be different for different subsequences and lie outside Π.

Definition 1.0.9. Let **P** be a probability measure on the metric space $(S, \mathcal{S})$. We define the Prokhorov distance $\pi(\mathbb{P}, \mathbb{Q})$ between two measures $\mathbb{P}, \mathbb{Q} \in \mathbf{P}$ to be the infimum of those positive ε for which:

$$\mathbb{P}[A] \leq \mathbb{Q}[A^\varepsilon] + \varepsilon,\ \mathbb{Q}[A] \leq \mathbb{P}[A^\varepsilon] + \varepsilon,\ \ \forall A \in \mathcal{S}$$

where A^ε is the cover of A with balls centered at points of A and radius ε.

Theorem 1.0.9. *Let $(S, \mathcal{S})$ be a complete separable metric space. Then weak convergence of probability measures is equivalent to convergence under the Prokhorov metric π, $(\mathbf{P}, \pi)$ is complete and separable, and $\Pi \subset \mathbf{P}$ is relatively compact iff it's $\pi-$closure is $\pi-$compact.*

We now proceed to one of the most important theorems of probability theory, which characterizes compactness for probability measures, namely the Prokhorov theorem.

Theorem 1.0.10 (Prokhorov)**.** *If a family of probability measures Π in $(S, \mathcal{S})$ is tight, then it is relatively compact. If, additionally, the space $(S, \mathcal{S})$ is complete and separable, and Π is relatively compact, then Π is tight.*

Since our application is referred to trajectories of functions, we will have to focus partially on the space $C[0,1]$. We know that under the metric defined by the supremum norm $||f||_\infty = \{|f(x)| : x \in [0,1]\}$, the space $C[0,1]$ is complete and separable. Furthermore, we have another way of examining tightness of probability measures on $C[0,1]$ apart from the definition, namely by means of the modulus of continuity:

Definition 1.0.10. Let $x \in [0,1] \to \mathbb{R}$. We define its modulus of continuity as:

$$w_x(\delta) = \sup_{|t-s|\leq\delta} \{|x(t) - x(s)|\}, \quad \delta \in (0,1]$$

It's clear that for any continuous function x it should be $w_x(\delta) \to 0$ as $\delta \to 0$. In fact, this is a characterization of continuous function on $[0,1]$. In any other case, the quantity $j_x = \lim_{\delta\to 0} w_x(\delta)$ is the value of the largest jump of x. For the following we define π_{t_0} the projection of a function on the time t_0, i.e. $\pi_{t_0}(x) = x(t_0)$

Theorem 1.0.11. *Let $\mathbb{P}^\varepsilon, \mathbb{P}$ be probability measures on $C[0,1]$. If $P^\varepsilon \pi^{-1}_{t_1,\dots,t_k} \to \mathbb{P}\pi^{-1}_{t_1,\dots,t_k}$ and additionally*

$$\lim_{\delta\to 0} \limsup_{\varepsilon\to 0} \mathbb{P}^\varepsilon[x : w_x(\delta) \geq l] = 0, \quad \forall l > 0$$

then $\mathbb{P}^\varepsilon \Rightarrow \mathbb{P}$.

We note that the condition we imposed on the modulus of continuity manages to yield the tightness of the family $\{\mathbb{P}^\varepsilon\}_{\varepsilon>0}$. Furthermore, if we manage to show that

$$\lim_{\delta\to 0} \mathbb{E}^{\mathbb{P}^\varepsilon}[w_x(\delta)] = 0$$

then this is sufficient a condition to satisfy last theorem's condition on the modulus of continuity. One needs only apply the Markov inequality, which we will present below, along with the equally important Jensen inequality.

Lemma 1.0.12 (Markov)**.** *Let X be a random variable. Then:*

$$\mathbb{P}[|X| \geq \varepsilon] \leq \frac{\mathbb{E}[|X|]}{\varepsilon}, \quad \forall \varepsilon > 0$$

Lemma 1.0.13 (Jensen)**.** *Let X be a random variable and f a convex function. Then:*

$$f(\mathbb{E}[(X)]) \leq \mathbb{E}[f(X)]$$

Conditional expectation

Definition 1.0.11. Let $X : \Omega \to \mathbb{R}$ be a random variable on $(\Omega, \mathcal{F}, \mathbb{P})$ with $\mathbb{E}[X] < \infty$. We define its conditional expectation w.r.t. a σ-algebra $\mathcal{G}$ to be a random variable $Y : \Omega \to \mathbb{R}$ such that:

1. Y is $\mathcal{G}-$measurable
2. $\int_A X d\mathbb{P} = \int_A Y d\mathbb{P}$

Essentially, the conditional expectation tells us the expected value of a random variable if an additional amount of information is known to us.

One may prove that under the conditions we've mentioned, the conditional expectation exists and is a.s. unique, i.e. for any two r.v.'s Y, Y' satisfying the definition of the conditional expectation we have $\mathbb{P}[Y = Y'] = 1$. In addendum, the following properties hold:

- (Linearity) $\mathbb{E}[aX + bY|\mathcal{G}] = a\mathbb{E}[X|\mathcal{G}] + b\mathbb{E}[Y|\mathcal{G}],\ \ \forall a, b \in \mathbb{R}$ and any r.v.'s X, Y.
- If $X \geq 0$, then $\mathbb{E}[X|\mathcal{G}] \geq 0$. Consequently, if $X \leq Y$ then $\mathbb{E}[X|\mathcal{G}] \leq \mathbb{E}[Y|\mathcal{G}]$.
- $\mathbb{E}[\mathbb{E}[X|\mathcal{G}]] = \mathbb{E}[X]$
- If X is $\mathcal{G}-$measurable, then $\mathbb{E}[X|\mathcal{G}] = X$. Furthermore, for any random variable Y, we have $\mathbb{E}[XY|\mathcal{G}] = X\mathbb{E}[Y|\mathcal{G}]$.
- If X is independent to $\mathcal{G}$, then $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X]$
- (Tower property) If $\mathcal{G}_1, \mathcal{G}_2$ σ-algebras with $\mathcal{G}_1 \subset \mathcal{G}_2$, then:

$$\mathbb{E}[\mathbb{E}[X|\mathcal{G}_2]|\mathcal{G}_1] = \mathbb{E}[\mathbb{E}[X|\mathcal{G}_1]|\mathcal{G}_2] = \mathbb{E}[\mathcal{G}_1]$$

Chapter 2

Elements of Stochastic Calculus

Having laid out the probabilistic essentials, we carry on with the description of the necessary material from the theory of Stochastic Calculus and Stochastic Differential equations.

Stochastic Processes-Filtration-Martingales

Definition 2.0.1. Let $(\Omega, \mathcal{F}, \mathbb{P}$ be a probability space. An $S-$valued stochastic process on this probability space is a measurable mapping $X : \Omega \times \mathbb{T} \to S$.

We refer to $\mathbb{T}$ as the time axis (usually it is a subinterval of $[0, \infty)$ or a subset of $\mathbb{N}_0$, i.e. the set of non-negative integes) and S is the so called state space.

For our purposes, $S = \mathbb{R}$ or $\mathbb{R}^d$ or some subset of either of these sets.

Moving on, we can "quantify" the information we have up to a certain point about a stochastic process by means of a filtration:

Definition 2.0.2. On a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, we define a filtration to be a collection $\{\mathcal{F}_t\}_{t\geq 0}$ of σ-algebras within $\mathcal{F}$ that is increasing, i.e. if $s < t \Rightarrow \mathcal{F}_s \subset \mathcal{F}_t$.

We observe that if $\{X_t\}$ is a stochastic process, one may define a filtration based on that process -or the filtration generated by this process- through defining $\mathcal{F}_t$ as the smallest σ-algebra for which the random variables $X_s,\ 0 \leq s \leq t$ are measurable.

Definition 2.0.3. A random variable τ is called a stopping time w.r.t. the filtration $\{\mathcal{F}_t\}_{t\geq 0}$ (or simply a $\mathcal{F}_t-$stopping time) if for any $t \geq 0$:

$$\{\tau \leq t\} \in \mathcal{F}_t$$

Simply put, stopping times are random variables for which we can determine if they have surpassed a time t or not when we have the information at time t at our disposal.

The above help us define a very useful type of stochastic processes: the martingales.

Definition 2.0.4. Let $\{\mathcal{F}_t\}_{t\geq 0}$ be a filtration and $\{M_t\}_{t\geq 0}$ a stochastic process on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We will say that M_t is a $\mathcal{F})_t-$martingale if the following hold:

1. $\mathbb{E}[|M_t|] < \infty,\ \forall t \geq 0$

2. $\mathbb{E}[M_t|\mathcal{F}_s] = M_s,\ \forall s, t \geq 0,\ s < t$

Intuitively, martingales represent processes that have the following property: if we have some information on the process at time s, that our estimation of what the process will be is precisely what we have at that given moment s. This can be translated in the field of games as a fair game, where if we have a certain amount of winnings, for example, at time s, then at any time $t > s$ neither we nor our opponent has an expected advantage.

Althought a very convenient property, martingality is substantially difficult to ascertain. This is the reason why usually most results in stochastic calculus are formulated for a broader class of process, the so-called local martingales:

Definition 2.0.5. A stochastic process M_t is called a local martingale w.r.t. the filtration $\{\mathcal{F}_t\}_{t\geq 0}$ if there is a sequence of stopping times $\{\tau_n\}_{n\in\mathbb{N}}$ with $\tau_n \to \infty$ so that the process $X_t^{\tau_n} = X_{t\wedge\tau_n}$ is an $\mathcal{F}_{t\wedge\tau_n}-$martingale.

Proposition 1. *A bounded local martingale $(X_t)_{t\geq 0}$ is a martingale.*

A fundamental result for the study of (local) martingales is the following:

Theorem 2.0.1 (Optional stopping theorem)**.** *Let X be a continuous martingale. If T is a bounded stopping time, then*

$$\mathbb{E}[X_T] = \mathbb{E}[X_0]$$

Brownian motion and Stochastic Integration

We carry on with a rather popular and frequently used tool of stochastic calculus, namely the Brownian motion:

Definition 2.0.6. We define a standard Brownian motion to be a $\mathbb{R}$-valued stochastic process $\{W_t\}_{t\geq 0}$ such that the following properties are satisfied:

1. $W_0 = 0$,
2. $W_t \sim \mathcal{N}(0, t)\ ,\ \forall t \geq 0$,
3. its increments are independent, i.e. for any $0 \leq t_1 < t_2 < \cdots < t_k$, the random variables $W_{t_1}, W_{t_2} - W_{t_1}, \ldots, W_{t_k} - W_{t_{k-1}}$ are independent,
4. its paths are a.s.- continuous.

One may easily generalize the above definition to extend it to a multidimensional Brownian motion:

Definition 2.0.7. A standard d-dimensional Brownian motion is a $\mathbb{R}^d$-valued stochastic process $\{W_t\}_{t\geq 0}$ such that:

1. $W_0 = \mathbf{0} \in \mathbb{R}^d$,

2. $W_t \sim \mathcal{N}(\mathbf{0}, t\mathcal{I}_d), \ \forall t \geq 0$, where $\mathcal{I}_d$ is the d-dimensional identity matrix,

3. its increments are independent,

4. its paths are a.s.-continuous

In other words, one may define a $d-$dimensional standard Brownian motion as a $d-$dimensional vector comprised by independent 1-d standard Brownian motions.

A prevalent alternative for the study of the Brownian motion is by means of the so-called Wiener measure, namely a probability measure defined on $C[0,\infty)$ (or $C([0,\infty);\mathbb{R}^d)$ respectively) under which the aforementioned properties hold. During this diploma thesis, we will be using these approaches interchangeably.

The properties above are, in fact, characteristic for a standard Brownian motion. That is, any stochastic process with independent, normally distributed increments and a.s.-continuous trajectories is a standard Brownian motion (provided, of course, that it starts from 0).

We also note that one may define the Brownian motion to start from a given point (other than zero) or even have a particular initial distribution.Indeed, if we consider W_t to be a standard Brownian motion and X a random variable with distribution μ, then $B_t = W_t + X$ is a Brownian motion with initial distribution μ. In particular, a Brownian motion starting from a point $x \in \mathbb{R}$ (or $\mathbb{R}^d$) can be defined in the same why by considering μ to be the Dirac measure δ_x.

We outline, now, some basic properties of a standard Brownian motion. For the following, let W_t be a standard Brownian motion

1. The process $B_t = -W_t$ is also a standard Brownian motion.

2. (Re-scaling) The process $B_t = \frac{1}{c} W_{c^2 t}, \ c \neq 0$, is a standard Brownian motion.

3. The process $B_t = t W_{\frac{1}{t}}$ for $t > 0$ and $B_0 = 0$ is a standard Brownian motion.

4. The quantities $W_t, W_t^2 - t$ and $e^{\lambda W_t - \frac{\lambda^2}{2} t}, \ \forall \lambda \geq 0$ are martingales (with respect to the filtration defined by W_t)

5. (Reflection principle) Let T be an a.s. finite stopping time. Then the process:

$$B_t = \begin{cases} W_t & \text{if } t \leq T \\ W_T - (W_t - W_T) & \text{if } t > T \end{cases}$$

is a standard Brownian motion.

Definition 2.0.8. Let $\{X_t\}_{t \geq 0}$ be a real-valued stochastic process. We define its quadratic variation as the process:

$$[X]_t = \lim_{||P|| \to 0} \sum_{i=1}^{n} (X_{t_i} - X_{t_{i-1}})^2$$

where the limit is taken on the partitions P of the interval $[0,t]$ and $||P||$ standards for the mesh of that partition. In the case where this limit exists, the convergence is in probability.

There is an analogous quantity with respect to a pair of stochastic processes:

Definition 2.0.9. Let $\{X_t\}_{t\geq 0}, \{Y_t\}_{t\geq 0}$ be two real-valued stochastic processes. We define their quadratic covariation as the process:

$$[X, Y]_t = \lim_{||P||\to 0} \sum_{i=1}^{n} (X_{t_i} - X_{t_{i-1}})(Y_{t_i} - Y_{t_{i-1}})$$

where the limit is meant in the same way as in the previous definition.

One may easily verify that $[X, X]_t = [X]_t$. Additionally, we can use a polarization identity to write:

$$[X, Y]_t = \frac{1}{2}\left([X + Y]_t - [X]_t - [Y]_t\right).$$

In addition, the above definitions can be generalized for an d-dimensional stochastic process by substituting the square with the square of the Euclidean norm and the ordinary product with the inner product in $\mathbb{R}^d$ respectively.

It can be proven that for a standard Brownian motion W_t we have $[W]_t = t$. In fact, according to Levy's theorem, that property is characteristic of the Brownian motion, i.e. any continuous martingale with quadratic variation at the time interval $[0, t]$ equal to the time t is none other than a Brownian motion.

Ito calculus

We will now illustrate the construction of the Ito integral and how this concept helps us develop the so-called Ito calculus. This construction works in a similar way to that of the Lebesgue integral, without being trivialized to the notion of the latter. We will restrict ourselves to the Ito integration, as the Stratonovich integral will not be utilized in this thesis.

Definition 2.0.10. A stochastic process h_t is called elementary/simple if it is piece-wise constant, so that there exist stopping times $0 = t_0 < t_1 < \cdots < t_n = T$ and a set of $\mathcal{F}_{t_i}-$measurable functions e_i such that

$$h_t = \sum_{k=0}^{n-1} e_k \mathbb{1}_{[t_k, t_{k+1}]}$$

Definition 2.0.11. For a simple stochastic process we define its stochastic integral or Ito integral in terms of the Brownian motion as:

$$\int_0^T h_u dW_u = \sum_{k=0}^{n-1} e_k (W_{t_{k+1}} - W_{t_k})$$

It can be proved, now, that a more general process X_t can be approximated by a sequence of simple processes, hence if $h_t^{(n)}$ is such a sequence, we define its Ito integral as

$$\int_0^T X_t dW_t = \lim_{n\to\infty} \int_0^T h_t^{(n)} dW_t$$

Usually we assume that our processes are square integrable in any interval $[0, T]$, i.e. $\mathbb{E}[\int_0^T X_t^2 dt] < \infty$ (we use the symbol $X \in L^2[0, T]$).

Theorem 2.0.2 (Ito's isometry)**.** *For any $X \in L^2[0, T]$ we have*

$$\mathbb{E}\left[\left(\int_0^T X_t dW_t\right)^2\right] = \mathbb{E}\left[\int_0^T X_t^2 dt\right]$$

Theorem 2.0.3. *If W_t is a standard Brownian motion, then the stochastic process $Y_t = \int_0^t X_u dW_u$ is a martingale for any $X_t \in L^2[0, T]$*

We move on to the application of the Ito integral to create a wide class of stochastic processes:

Definition 2.0.12. An n-dimensional Ito process X_t is a process that can be written as:

$$X_t = X_0 + \int_0^t a_s ds + \int_0^t b_s dW_s$$

where a_s is an n-dimensional $\{\mathcal{F}_t\}_{t\in[0,T]}-$adapted process, W_t is a standard m-dimensional Brownian motion, b_s is a $n \times m-$dimensional $\{\mathcal{F}_t\}_{t\in[0,T]}-$adapted process

(We remind that a process A_t is $\{\mathcal{F}_t\}-$adapted if A_t as a random variable is $\mathcal{F}_t-$measurable for any $t \in [0, T]$).

In order for us to do calculus on Ito processes, we need to define an analogue of the Fundamental theorem of calculus. That is achieved by the following:

Theorem 2.0.4 (Ito's formula, 1-d, Brownian Motion)**.** *Let W_t be a standard Brownian motion and $f \in C^2$. Then for any $t \leq T$:*

$$f(W_t) = f(0) + \int_0^t f'(W_s) dW_s + \frac{1}{2}\int_0^t f''(W_s) ds$$

Theorem 2.0.5 (Ito's formula, 1-d, General Ito process)**.** *Let X_t be an Ito process with*

$$dX_t = \mu_t dt + \sigma_t dW_t$$

.

If $f(t, x) : [0, T] \times \mathbb{R} \to \mathbb{R}$ is a $C^{1,2}$ function, then for $Z_t = f(t, X_t)$ we have :

$$dZ_t = \left(\frac{\partial f}{\partial t}(t, X_t) + \frac{\partial f}{\partial x}(t, X_t)\mu_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(t, X_t)\sigma_t^2\right) dt + \frac{\partial f}{\partial x}(t, X_t)\sigma_t dW_t$$

Theorem 2.0.6 (Ito's formula, multidimensional, General Ito process)**.** *Let X_t be an n-dimensional Ito process with*

$$dX_t = \mu_t dt + \sigma_t dW_t$$

, where μ_t is an n-dimensional vector, W_t is an m-dimensional Brownian motion and σ_t an $n \times m-$dimensional matrix. Then for $f(t,x) : [0,T] \times \mathbb{R}^n \to \mathbb{R}$ being a $C^{1,2}$ function and $Z_t = f(t, X_t)$ we have :

$$dZ_t = \frac{\partial f}{\partial t}(t, X_t)dt + (\nabla_X f(t, X_t))^T dX_t + (d[X]_t)^T H_X f(t, X_t) d[X]_t =$$

$$\left(\frac{\partial f}{\partial t}(t, X_t) + \nabla_X f(t, X_t)\mu_t + \frac{1}{2}Tr(\sigma_t^T (H_X f)\sigma_t)(t, X_t)\sigma_t^2\right) dt + \nabla_X f(t, X_t)\sigma_t dW_t$$

where we notate $\nabla_X f$ as f's gradient in terms of the variables that have to do with X,H_X is the Hessian matrix in terms of the same variables and Tr stands for the trace of a matrix.

A crucial step while we finalize our proof later on will be using a change of measure argument, by means of the Cameron-Martin and Girsanov theorems.

Theorem 2.0.7 (Girsanov)**.** *Let θ_s s.t. $\mathbb{E}^{\mathbb{P}}[exp\{\int_0^t \theta_s^2 ds\}] < \infty$ and let W_t be a standard Brownian motion under the probability measure $\mathbb{P}$. If $Z(t) = exp\{\int_0^t \theta_s dW_s - \frac{1}{2}\int_0^t \theta_s^2 ds\}$ and we define a new probability measure with the Radon-Nikodym derivative $\frac{d\mathbb{Q}}{d\mathbb{P}} = Z(t)$, then under the new measure $\mathbb{Q}$ the stochastic process $\hat{W}_t = W_t - \int_0^t \theta_s ds$ is a standard Brownian motion.*

Corollary 2.0.7.1 (Cameron-Martin Theorem)**.** *Let W_t be a standard Brownian motion under the probability measure $\mathbb{P}$. If $Z(t) = exp\{\theta W_s - \frac{\theta^2 t}{2}\}$ and we define a new probability measure with the Radon-Nikodym derivative $\frac{d\mathbb{Q}}{d\mathbb{P}} = Z(t)$, then under the new measure $\mathbb{Q}$ the stochastic process $\hat{W}_t = W_t - \theta t$ is a standard Brownian motion.*

An equation of the form $dX_t = b(X_t)dt + \sigma(X_t)dW_t$ is called a Stochastic Differential Equation (S.D.E.). Under certain conditions, we can prove that such a problem has a unique solution (we will soon provide an explanation as to what that means) and moreover that the solution process is a Markov process.

Definition 2.0.13. A Markov process on X is a stochastic process to which we correspond a set of probability measures $\mathbb{P}^\eta$, $\eta \in X$ on $X^{[0,\infty)]}$ such that:

1. $\mathbb{P}^\eta[\zeta \in X^{[0,\infty)]} : \zeta_0 = \eta] = 1$
2. The mapping $\eta \mapsto \mathbb{P}^\eta[A]$ from X to $[0,1]$ is measurable for any $A \in \mathcal{F}$
3. $\mathbb{P}^\eta[\eta_{s+\cdot} \in A|\mathcal{F}_s] = \mathbb{P}^{\eta_s}[[A]\ a.s.\ (\mathbb{P}^\eta)$ for every $\eta \in X$ and $A \in \mathcal{F}$

In the above, we denoted as $X^{[0,\infty)]}$ the space of functions from $[0,\infty)$ to X, $\mathcal{F}$ as the Borel σ-algebra of this space and $\mathcal{F}_t$ the σ-algebra generated by the projections $\pi_s,\ \forall s \leq t$.

Definition 2.0.14. A stochastic process is called a strong Markov process if it is a Markov process and additionally the following property holds: for any a.s.-finite stopping time τ we have:

$$\mathbb{P}^\eta[\eta_{\tau+\cdot} \in A|\mathcal{F}_\tau] = \mathbb{P}^{\eta_\tau}[A]\ a.s.$$

for any $\eta \in X$ and any $A \in \mathcal{F}$. That is, a strong Markov process is such that for any stopping time τ we can shift the process by this time τ creating a new Markov process starting at η_τ which is independent of the "past", i.e. the σ-algebra $\mathcal{F}_\tau$

We close this section with the basic concepts of SDEs in terms of existence and uniqueness of solution. To that end, we will consider an d-dimensional SDE $dX_t = b(X_t)dt + \sigma(X_t)dW_t$ with $X_0 = x$.

Definition 2.0.15. A continuous adapted stochastic process X_t is called a pathwise solution if it satisfies

$$X_0 = x \text{ and } X_t = x_0 + \int_0^t b(X_s)ds + \int_0^t \sigma(X_s) \cdot dW_s$$

We will refer to a problem having unique pathwise solution as a problem for which if X_t, X_t' are two processes satisfying the initial condition and the SDE, then there exists a set N s.t. $\mathbb{P}[N] = 0$ and outside of N we have $X_t = X_t',\ \forall t \geq 0$

A typical resulting for the proof of existence and uniqueness of solution is the following:

Theorem 2.0.8. *Let the coefficients b, σ be Lipschitz continuous, i.e. there exists constants L_b, L_σ s.t. $|b(x) - b(y)| \leq L_b|x - y|$ and $|\sigma(x) - \sigma(y)| \leq L_\sigma|x - y|$. Furthermore, we assume a condition of linear growth on b, σ, i.e we assume that there exists a constant c_1 such that:*

$$|\sigma(x)| + |b(x)| \leq c_1(1 + |x|)$$

Then, for any initial condition x, the problem $dX_t = b(X_t)dt + \sigma(X_t)dW_t,\ X_0 = x$ has a pathwise solution and it's pathwise unique.

Additionally, we have the notion of strong and weak solution of an SDE, as well as the respective type of uniqueness. From this point on, we will name the problem $dX_t = b(X_t)dt + \sigma(X_t)dW_t,\ X_0 = x$ as $(*)$

Definition 2.0.16. A strong solution to the problem $(*)$ exists if given a Brownian motion W_t there exists a process X_t so that the equations of $(*)$ are satisfied and X_t is adapted to W's filtration. A weak solution exists if there is a pair of processes (X_t, W_t) such that W_t is a Brownian motion and X_t satisfies $(*)$. We say that $(*)'$s solution is weakly unique if for any weak solutions $(X_t, W_t),\ (X_t', W_t')$ we have that the joint distributions of the pairs $(X, W),\ (X', W')$ are the same. In this case, we also say that the solution is unique in law.

Theorem 2.0.9. *Let b, σ be Lipschitz. Then the problem $(*)$ has a strong solution and we also have weak uniqueness.*

Theorem 2.0.10. *Suppose the matrix σ has an inverse that is bounded. Suppose σ and b are bounded and measurable. If $(*)$ has a strong solution and the solution to $(*)$ is weakly unique, then pathwise uniqueness holds for $(*)$.*

Chapter 3

Martingale problem: General Theory

In this chapter we will cover the theory regarding the so-called martingale problem. Prior to that, it is necessary to introduce the notion of the infinitesimal generator for an Ito process. Once we've covered that essential part of stochastic processes, we will define the martingale problem and study the issues of existence and uniqueness of solution.

Since-based on what we mentioned in the previous chapter- any Ito diffusion is essentially a Markov process, we will develop the respective theory around Markov processes and a special subset of these processes, we will define below. In the following, we denote $S(t)$ a collection of operators such that, if X_t is an X-valued Markov process and $f \in C(X)$, then $S(t)f(\eta) = \mathbb{E}_\eta[f(X_t)]$ (here the subscript of η indicates the initial point for this Markov process).

Definition 3.0.1 (Feller process)**.** A Markov stochastic process X_t is called a Feller process $S(t)f \in C(X)$ for every $t \geq 0$ and $f \in C(X)$.

Proposition 2. *Let X_t be a Feller process. Then the collection of operators $\{S(t)\}_{t\geq 0}$ on $C(X)$ satisfies the following properties :*

1. $S(0) = I$ *the identity operator on* $C(X)$
2. $t \mapsto S(t)f$ *from* $[0, \infty)$ *to* $C(X)$ *is right-continuous for every* $f \in C(X)$.
3. $S(t+s)f = S(t)S(s)f \; \forall f \in C(X)$ *and* $t, s \geq 0$
4. $S(t)1 = 1 \; \forall t \geq 0$
5. $S(t)f \geq 0$ *for all nonnegative* $f \in C(X)$

Definition 3.0.2. A family of linear operators on $C(X)$ that satisfies the properties $1 - 5$ of the above proposition is called a Markov Semigroup.

The following converse of Proposition 1 indicates that, in fact, we need only construct the semigroup of operators in order to fully determine the Markov process.

Theorem 3.0.1. *Suppose $\{S(t)\}_{t\geq 0}$ is a Markov semigroup on $C(X)$. Then there exists a unique Markov process such that:*

$$S(t)f(\eta) = \mathbb{E}^{\eta}[f(X_t)]$$

Let $\mathcal{P}$ be the space of probability measures on X, with the (metrizable) topology of weak convergence. We will, primarily, consider X to be a compact space and so $\mathcal{P}$ will also be compact. If we consider a $\mu \in \mathcal{P}$, then the Markov process with initial distribution μ is a stochastic process with distribution $\mathbb{P}^{\mu} = \int \mathbb{P}^{\eta}\mu(d\eta)$. In that case, we have

$$\mathbb{E}^{\mu}[f(X_t)] = \int S(t)fd\mu$$

Definition 3.0.3. Let $\{S(t)\}_{t\geq 0}$ be a Markov semigroup on $C(X)$. Given $\mu \in \mathcal{P},\ \mu S(t) \in \mathcal{P}$ is defined by the relation:

$$\int fd[\mu S(t)] = \int S(t)fd\mu$$

for all $f \in C(X)$ (this new probability measure is interpreted as the distribution of X_t at time t with initial distribution μ).

Definition 3.0.4. A probability measure $\mu \in \mathcal{P}$ is called invariant for a Markov semigroup $\{S(t)\}_{t\geq 0}$ if $\mu S(t) = \mu,\ \forall t \geq 0$. We will denoted the class of invariant measures with $\mathcal{I}(X)$.

Proposition 3.

1. *$\mu \in \mathcal{I}(X)$ iff $\int S(t)fd\mu = \int fd\mu\, f$*
2. *$\mathcal{I}(X)$ is a compact convex subset of $\mathcal{P}$*
3. *Let $\mathcal{I}_e$ be the set of extreme points of $\mathcal{I}(X)$. Then its closed convex hull is equal to $\mathcal{I}(X)$.*
4. *If $\nu = \lim_{t\to\infty} \mu S(t)$ exists for some $\mu \in \mathcal{P}$, then $\nu \in \mathcal{I}$.*
5. *If $\nu = \lim_{n\to\infty} T_n^{-1} \int_0^{T_n} \mu S(t)dt$ exists for some $\in \mu \in \mathcal{P}$ and some sequence $\{T_n\}$ with $T_n \to \infty$, then $\nu \in \mathcal{I}(X)$*
6. *$\mathcal{I}$ is non-empty.*

Definition 3.0.5. We will a Markov process ergodic if:

1. $\mathcal{I}(X)$ is a singleton and
2. $\lim_{t\to\infty} \mu S(t) = \nu$ for all $\mu \in \mathcal{P}$

Semigroups and their generators

Definition 3.0.6. A (usually unbounded) linear operator L on $C(X)$ with domain $\mathcal{D}(L)$ is said to be a Markov pregenerator if it satisfies the following:

1. $1 \in \mathcal{D}(L)$ and $L1 = 0$
2. $\mathcal{D}(L)$ is dense in $C(X)$
3. If $f \in \mathcal{D}(L)$, $\lambda \geq 0$ and $f - \lambda Lf = g$ then

$$\min_{\zeta \in X} f(\zeta) \geq \min_{\zeta \in X} g(\zeta)$$

.

For this last property, we conclude that $||f|| \leq ||g||$

Proposition 4. *Suppose that the linear operator L on $C(X)$ satisfies the following property: if $f \in \mathcal{D}(L)$ and $f(\eta) = \min_{\zeta \in X} f(\zeta)$, then $Lf(\eta) \geq 0$. Then L satisfies the third property of the previous definition.*

Definition 3.0.7. A linear operator L on $C(X)$ is said to be closed if its graph is a closed subset of $C(X) \times C(X)$. A linear operator $\bar{L}$ is called the closure of L if it is the smallest closed extension of L.

Although not every linear operator has to have a closure (this problem arises especially in the case of multivalued operator), this pathology does not arise in the case of Markov pregenerators.

Proposition 5. *Let L be a Markov pregenerator. Then there exists the closure of L, i.e. $\exists \bar{L}$ and it is a Markov pregenerator as well.*

Proposition 6. *Let L be a closed Markov pregenerator. Then the range of the operator $I - \lambda L$ is a closed subset of $C(X)$ for any $\lambda > 0$.*

We now combine the above to define the notion of a Markov generator:

Definition 3.0.8. A Markov generator L is a closed Markov pregenerator which satisfies $R(I - \lambda L) = C(X)$ for all sufficiently small $\lambda > 0$.

Additionally, there are some relatively easily verifiable conditions for a pregenerator to be a Markov generator, as can be seen by the following:

Proposition 7.

1. *A bounded (everywhere defined) Markov pregenerator is a Markov generator*
2. *A Markov pregenerator that satisfies $R(I - \lambda L) = C(X)$ for all $\lambda \geq 0$.*

We finish this preamble on generators with the famous theorem from functional analysis, the so-called Hille-Yosida, which in our case will also provide us with the expression most commonly used as the definition of the infinitesimal generator of a Markov process.

Theorem 3.0.2 (Hille-Yosida)**.** *There is a one-to-one correspondence between Markov generators on $C(X)$ and Markov semigroups on $C(X)$. The correspondence is given by:*

1. $$\mathcal{D}(L) = \{f \in C(X) : \lim_{t\to 0} \frac{S(t)f - f}{t} \textit{ exists}\}$$

   $$Lf = \lim_{t\to 0} \frac{S(t)f - f}{t}, \ f \in \mathcal{D}(L)$$

2. $S(t)f = \lim_{n\to\infty} \left(I - \frac{t}{n}L\right)^{-n} f, \ \forall f \in C(X), \ \forall t \geq 0$

   Furthermore,

3. *if $f \in \mathcal{D}(L)$ it follows that $S(t)f \in \mathcal{D}(L)$ and $\frac{d}{dt}S(t)f = LS(t)f = S(t)Lf$ and finally*

4. *for any $g \in C(X)$ and $\lambda \geq 0$ the solution to $f - \lambda Lf = g$ is given by:*

   $$f = \int_0^\infty e^{-t} S(\lambda t) g dt$$

Hence, we see that for a Markov process X_t with initial point x_0 we have that its infinitesimal generator is:

$$Lf = \lim_{t\to 0} \frac{\mathbb{E}[f(X_t)] - f(x_0)}{t}$$

The Martingale Problem

Existence

We initialize our discussion of Martingale problems by examining the issue of existence. Later on we will see how uniqueness can be tackled. Let $\mathcal{L}$ be an elliptic operator in the form:

$$\mathcal{L}f(x) = \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}\partial_{ij}f(x) + \sum_{i=1}^{d} b_i(x)\partial_i f(x), \quad f \in C^2$$

where ∂_i denotes the partial derivative in terms of the variable x_i and $\partial_{i,j}$ the second order partial derivative with respect to x_i, x_j. We assume throughout that $a_{i,j}, b_i$ are bounded and measurable functions. Since the coefficient of $\partial_{i,j}f(x)$ is $\frac{a_{ij}(x)+a_{ji}(x)}{2}$, we can assume without loss of generality that a_{ij} is symmetric, i.e. $a_{ij} = a_{ji}$. We define:

$$\mathcal{N}(\Lambda_1, \Lambda_2) = \{\mathcal{L} : \sup_{i\leq d} ||b_i||_\infty \leq \Lambda_2 \text{ and } \Lambda_1|y|^2 \leq \sum_{i,j=1}^{d} y_i y_j a_{ij}(x) \leq \Lambda_1^{-1}|y|^2, \ \forall x, y \in \mathbb{R}^d\}$$

If $\mathcal{L} \in \mathcal{N}(A, B)$ for some $A > 0$, then we say that $\mathcal{L}$ is uniformly elliptic.

Definition 3.0.9. We say that a probability measure is a solution to the martingale problem for $\mathcal{L}$ started at x if

$$\mathbb{P}[X_0 = x] = 1$$

and the process

$$f(X_t) - f(x) - \int_0^t \mathcal{L}f(X_s)ds$$

is a local martingale under $\mathbb{P}$ whenever f is in $C^2(\mathbb{R}^d)$.

Theorem 3.0.3. *Suppose that a_{ij}, b_i are bounded and continuous and $x \in \mathbb{R}^d$. Then there exists a solution to the martingale problem for $\mathcal{L}$ started at x.*

In the case where the operator $\mathcal{L}$ is uniformly elliptic, it suffices to examine only the part yielding from the diffusion, i.e. the operator

$$\mathcal{L}' = \frac{1}{2}\sum_{i,j=1}^{d} a_{ij}\partial_{ij}f(x)$$

(one can prove this result by merely applying Girsanov's theorem).

Theorem 3.0.4. *Suppose that $\mathcal{L} \in \mathcal{N}(\Lambda_1, \lambda_2)$. If there exists a solution to the martingale problem for the operator $\mathcal{L}'$ (as previously defined) started at x, then there exists a solution to the martingale problem for $\mathcal{L}$ started at x.*

Existence can be proven even if we lack the condition of continuity on a_{ij}, as long as we have that $\mathcal{L}'$s uniform elliptic.

Theorem 3.0.5. *Suppose that $\mathcal{L} \in \mathcal{N}(\Lambda_1, \Lambda_2)$ where a_{ij}, b_i are measurable (not necessarily continuous). If $x \in \mathbb{R}^d$, there exists a solution to the martingale problem for $\mathcal{L}$ started at x.*

On the subject of uniqueness, there is an intimate connection between the uniqueness of an SDE and the respective martingale problem.

Theorem 3.0.6. *Suppose $a = \sigma\sigma^T$. Then weak uniqueness for the problem: $\{dX_t = b(X_t)dt + \sigma(X_t)dW_t,\ X_0 = x\}, (*)$ holds if and only if the solution for the martingale problem for L started at x is unique. Weak existence for $(*)$ holds if and only if there exists a solution to the martingale problem for L started at x.*

We illustrated that the existence problem can be treated for the exclusive case of non-existent drift (by Theorem 3.04.). As we'll see in the next result, this remains true in the case of uniqueness, under the same conditions. The proof is yet another application of the Girsanov Theorem.

Theorem 3.0.7. *Suppose $\mathcal{L} \in \mathcal{N}(\Lambda_1, \Lambda_2)$. If we have uniqueness for the martingale problem for $\mathcal{L}'$ started at x, then the solution is unique for the martingale problem for $\mathcal{L}$ started at x.*

On a more practical note, to prove uniqueness it turns out that it is sufficient to look at quantities which are essentially λ-potentials (that is, λ-resolvents). It will be convenient to introduce the notation

$$\mathcal{M}(\mathcal{L}, x) = \{\mathbb{P} : \mathbb{P} \text{ is a solution to the martingale problem for } \mathcal{L} \text{ started at } x\}.$$

Theorem 3.0.8. *Suppose that for all $x \in \mathbb{R}^d$, $\lambda > 0$ and $f \in C^2(\mathbb{R}^d)$ we have:*

$$\mathbb{E}_1\left[\int_0^\infty e^{-\lambda t} f(X_t)dt\right] = \mathbb{E}_2\left[\int_0^\infty e^{-\lambda t} f(X_t)dt\right]$$

whenever $\mathbb{P}_1, \mathbb{P}_2 \in \mathcal{M}(\mathcal{L}, x)$. Then for any $x \in \mathbb{R}^d$ the solution to the martingale problem for $\mathcal{L}$ started at x is unique.

Additionally, it is useful to mention at this point that we need only examine uniqueness locally, since we can prove the following ”piecing-together” lemma:

Lemma 3.0.9. *Let $\mathcal{L}_1, \mathcal{L}_2$ be two elliptic operators with bounded coefficients and let $S = \inf\{t : |X_t - x| \geq r\}$. Let, also, $\mathbb{P}_1, \mathbb{P}_2$ be two solutions for the martingale problem for $\mathcal{L}_1, \mathcal{L}_2$ respectively started at x. Let $\mathbb{Q}_2$ be the conditional probability measure of $\mathbb{P}_{2,S}$ on $\mathcal{F}_S$, where $\mathbb{P}_{2,S}[A] = \mathbb{P}_2[A \circ \theta_S]$ (θ_t denotes the time-shifting operator by an amount of t). Define $\bar{\mathbb{P}}$ as*

$$\bar{\mathbb{P}}[B \circ \theta_S \cap A] = \mathbb{E}_{\mathbb{P}_1}[\mathbb{Q}_2[B]; A],\ A \in \mathcal{F}_S,\ B \in \mathcal{F}_\infty$$

If the coefficients of $\mathcal{L}_1, \mathcal{L}_2$ coincide on $B(x, r)$, then $\bar{\mathbb{P}}$ is a solution to the martingale problem. Here, $\bar{\mathbb{P}}$ expresses the process behaving according to $\mathbb{P}_1$ up to the time S and later on according to $\mathbb{P}_2$.

This result is crucial to proving the following:

Theorem 3.0.10. *Suppose $\mathcal{L} \in \mathcal{N}(\Lambda_1, \Lambda_2)$. Suppose for each $x \in \mathbb{R}^d$ there exist $r_x > 0$ and $\mathcal{K}(x) \in \mathcal{N}(\Lambda_1, \Lambda_2)$ such that the coefficients of $\mathcal{K}(x)$ agree with those of $\mathcal{L}$ in $B(x, r_x)$ and the solution to the martingale problem for $\mathcal{K}(x)$ is unique for every starting point. Then the martingale problem for $\mathcal{L}$ started at any point has a unique solution.*

From a thorough examination of the respective theory, one may find in ***cite Bass*** the following theorem:

Theorem 3.0.11. *Suppose d = 2, $\mathcal{L} \in \mathcal{N}(\Lambda, 0)$. Then the martingale problem for $\mathcal{L}$ started at any x is unique.*

Note that there are no assumptions on the coefficients regarding continuity.

Chapter 4

The Freidlin-Wentzell ("Classical") approach

We already have a way of tackling problems of metastability regarding Hamiltonian systems, namely the so-called Freidlin-Wentzell theory. Broadly speaking, what Freidlin and Wentzell showed for Hamiltonian systems is that, under certain conditions on the Hamiltonian, as $\varepsilon \to 0$ we can identify the shifting between level sets with a diffusion on a particular graph. In the case of the harmonic oscillator which we examine on this thesis, however, things are much simpler, since the graph we're referring to is merely the set $[0, \infty]$.

Let $\dot{x}_t = b(x_t)$ be an n-dimensional Hamiltonian system, where $x_t = (q_1, \dots, q_n, p_1, \dots, p_n)$, $H(x)$ is a sufficiently smooth function with $\lim_{|x|\to\infty} H(x) = \infty$ and

$$b(x_t) = (\frac{\partial H}{\partial p_1}(x_t), \dots, \frac{\partial H}{\partial p_n}(x_t), -\frac{\partial H}{\partial q_1}(x_t), \dots, -\frac{\partial H}{\partial q_n}(x_t))$$

.

What we're interested in is a stochastic perturbation of this problem by means of a white noise, i.e.:

$$\dot{\tilde{x}}_t^\varepsilon = b(\tilde{x}_t^\varepsilon) + \varepsilon \dot{\tilde{W}}_t$$

where $\tilde{W}_t$ is standard 2n-dimensional Brownian motion.

It is known from the theory of analytic mechanics that since the Hamiltonian does not depend on time explicitly, it is a constant of motion. This means that, for the unperturbed system, a typical trajectory will be of the form $H(x_t) = H_0 = H(x_0)$. However, since we introduce a perturbation by means of the white noise, there will be some motion across these level curves. Comparatively, the motion across those curves is much slower than the motion along one curve of this type. This justifies a change of time scale of the type $t \mapsto \frac{t}{\varepsilon^2}$, so we have a new, "fast-forwarded" system to describe the physics of the system, which is:

$$\dot{x}_t^\varepsilon = \frac{1}{\varepsilon^2} b(x_t) + \dot{W}_t$$

or, with by using notation with differentials:

$$dx_t^\varepsilon = \frac{1}{\varepsilon^2} b(x_t^\varepsilon) dt + dW_t$$

We further assume that the initial condition is of the form $X_0 = x_0$, i.e. we do not assume a more complex initial distribution for this problem.

The generator of the process x_t^ε is, in that case, the operator:

$$L^\varepsilon = \frac{1}{\varepsilon^2} b(x_t^\varepsilon) \cdot \nabla + \frac{1}{2}\Delta$$

and the Lebesgue measure is invariant for this process. Although, theoretically, we could introduce another perturbation, hence having the second term of L^ε be equal to some elliptic differential operator L_0, we remain on this case for the sake of simplicity.

If, now, we apply Ito's formula on $H(x_t^\varepsilon)$ we have:

$$dH(x_t^\varepsilon) = \nabla H(x_t^\varepsilon) \frac{1}{\varepsilon^2} b(x_t^\varepsilon) dt + \nabla H(x_t^\varepsilon) \cdot dW_t + \frac{1}{2} \Delta H(x_t^\varepsilon) dt$$

One may calculate quite easily that $\nabla H(x_t^\varepsilon) \cdot b(x_t^\varepsilon) = 0$, hence we write:

$$H(x_t^\varepsilon) = H_0 + \int_0^t \nabla H(x_s^\varepsilon) \cdot dW_s + \frac{1}{2} \int_0^t \Delta H(x_s^\varepsilon) ds$$

The idea, now, is that we have a large amount of rotations before $H(x_t^\varepsilon)$ changes significantly, and so the deterministic integral in the above equation can be approximated with $\int_0^t B(H(x_s^\varepsilon)) ds$, where:

$$B(H) = \frac{\oint \frac{\frac{1}{2}\Delta H(x)}{|b(x)|} dl}{\oint \frac{1}{|b(x)|} dl}$$

where the integral is calculated over the level set where $H(x) = H$. As for the stochastic integral, we have that:

$$\int_0^t \nabla H(x_t^\varepsilon) \cdot dW_t = w(\int_0^t |\nabla H(x_s^\varepsilon)|^2 ds)$$

where w is a standard 1-dimensional Brownian motion. In this case, we can approximate it with the quantity $\int_0^t A(H(x_s^\varepsilon)) ds$ where:

$$A(H) = \frac{\oint \frac{|\nabla H(x)|^2}{|b(x)|} dl}{\oint \frac{1}{|b(x)|} dl}, \quad B(H) = \frac{\oint \frac{\frac{1}{2}\Delta H(x)}{|b(x)|} dl}{\oint \frac{1}{|b(x)|} dl}$$

with the integration taken in the same way as in $B(H)$. This means that the "slow" process can be aproximated by a process with the generator:

$$\mathcal{L} f(H) = \frac{1}{2} A(H) f''(H) + B(H) f'(H)$$

This gives a description of the "slow" process in the space obtained by identifying all points on the same trajectory of the unperturbed system, but only while the process x_t^ε is moving in a region covered by closed trajectories. However, this process can move from one region covered by closed trajectories to another, and such regions are separated by the components of level curves $\{x : H(x) = H\}$ that are not closed trajectories. The folowing image illustrates an example of such a Hamiltonian system

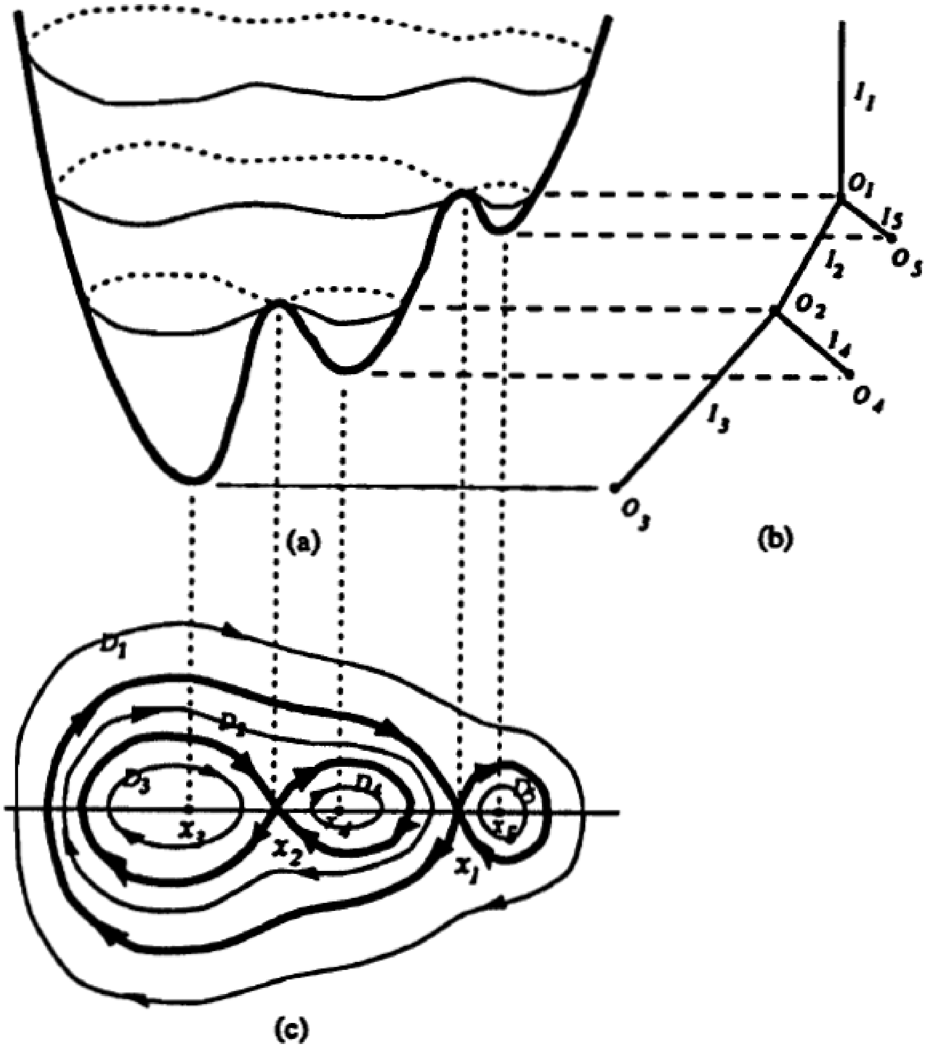

Figure 4.1: Example of a Hamiltonian system with multiple equilibrium points and its representation through a graph (image was found in [**13**])

The procedure that we have followed here is the following: all the points of a given level curve we identify with a single point, hence we get a graph with several segments, corresponding I_1 to the trajectories of the domain D_1 outside the $\infty-$shaped curve, I_2 to the trajectories in D_2 between the outer and inner $\infty-$shaped curves, I_3 and I_4 to the trajectories inside the two loops of the inner $\infty-$shaped curve (domains D_3 and D_4) and I_5 to those inside the right loop of the outer $\infty-$shaped curve (domain D_5).

The ends of the segments are vertices O_1, O_2 corresponding to the $\infty-$shaped curves, O_3, O_4, O_5 corresponding to the extrema x_3, x_4, x_5 (in our example, it just so happens that the curves corresponding to O_1, O_2 each contain a critical point of H, namely a saddle point). Let us complement the graph by assigning another vertex O_∞ to the point at infinity. We will denote the resulting graph as Γ.

We consider $Y(x)$ the mapping assigning each point $x \in \mathbb{R}^2$ to the corresponding point of the graph. We will denote the function H carried over to the graph under this mapping also by H (note that $H(0_\infty) = \infty$). The function H can be taken as a local coordinate on this graph.

Couples (i, H) where is the number of the segment I_i define global coordinates on the graph. Several such couples may correspond to a single vertex.

In the case where $x \in \mathbb{R}^2$ the graph has the structure of a tree. However, in a different manifold the appearance of loops is also possible.

We note that $i(x)$, the number of the segment of the graph containing $Y(x)$, is preserved in the case of the unperturbed system, that is $i(X_t^x) = i(x), \forall t$. This implies that i is a -discrete- first integral (i.e. a conserved qunatity) for this system. In systems where the Hamiltonian possesses more than one critical points, it is possible that we have $H(x) = H(y)$ for x, y s.t. $i(x) \neq i(y)$. In that case, the Hamiltonian system has two independent first integrals, namely $H(x)$ and $i(x)$. As a result, the process $H(x_t^\varepsilon)$ does not converge to a Markov process as $\varepsilon \to 0$ in the case of several critical points. If there are multiple first integrals, we must broaden our phase space and introduce all first integrals as coordinates in it.

In the present case, we have the two coordinates (i, H).

The couple $(i(x_t^\varepsilon), H(x_t^\varepsilon) = Y(x_t^\varepsilon)$ is a stochastic process on the graph Γ and it is reasonable to expect that it will converge to some diffusion process Y_t on the graph Γ as $\varepsilon \to 0$. What interests us is what can be said, in general, about this resulting diffusion Y_t.

First of all, on any segment I_i we can associate a respective generator acting on functions on this segment:

$$\mathcal{L}_i f(H) = \frac{1}{2} A_i(H) f''(H) + B_i(H) f'(H)$$

where A_i, B_i are computed similarly to the formulae we introduced earlier and the respective contour integrals are computed upon the level curve $\{x : H(x) = H\}$ lying in the region D_i corresponding to I_i. These generators define the behaviour of Y_t on the graph as long as the process remains within a particular segment. But what happens when the process exits a segment?

As it turns out, in order to determine the behaviour of such a diffusion exiting the interior of a segment, certain boundary conditions are required, but only for the ends of the segment that are accessible from the inside. Criteria of accessibility of an end from the inside and also of reaching the insider from an end have been given. One of them is the following: if the integral

$$\int e^{\{-\int \frac{2B(H)}{A(H)} dH\}} dH$$

diverges at the end of H_k, then H_k is not accessible from the inside. A simpler formulation of these criteria, and also the boundary conditions, can be attained by representing $\mathcal{L}f$ as a generalized second derivative $\frac{d}{du}\frac{d}{dv}f$ with respect to two increasing functions $u(H), v(H)$. For instance, condition for the integral above diverging at H_k can be rephrased as the unboundedness of $u(H)$ at H_k (in fact, the above integral can serve as a choice for $u(H)$). On the other hand, the end H_k is accessible from the inside and the insidie is accessible from H_k if and only if $u(H), v(H)$ are bounded at H_k. Yet another useful condition for inacessibility is the following: if the integral $\int v(H) dH$ diverges at H_k, then the end H_k is not accessible from the inside.

Representing the differential operator in the form of a generalized second derivative is especially convenient in our case, because the operators $\mathcal{L}_i$ degenerate at the ends of the segment I_i: its coefficients $A_i(H), B_i(H)$ given by the above formulae have finite limits, but $() \to 0$ at these ends (in the case of nondegenerate critical points, at an inverse logarithmic rate at an end corresponding to a level curve that contains a saddle point, and linearly at an end corresponding to an extremum).

The above idea, owed to Feller, can be carried over to diffusions on graphs; if some segments I_i, meet at a vertex O_k (which we write as $I_i\ O_k$) and O_k is accessible from the inside of at least one segment, then some "interior boundary" conditions (or "gluing" conditions) have to be prescribed at O_k. (If the vertex $O_k$is inaccessible from any of the segments, no condition has to be given.) If for all segments$I_i\ O_k$ the end of I_i corresponding to O_k is accessible from the inside of I_i, and the inside of I_i can be reached from this end, then the general interior boundary condition can be written in the form

$$\alpha_k Lf(O_k) = \sum_{i:I_i \sim O_k} (\pm\beta_k i) \frac{df}{du_i}(O_k)$$

where $\mathcal{L}f(O_k)$ is the common limit at O_k of the functions $\mathcal{L}_i f$ defined above for all segments $I_i \sim O_k$; u_i is the function on I_i, used in the representation $\mathcal{L}_i = (d/dvi)(d/dui)$; $\alpha_k \geq 0$, $\beta_{ki} \geq 0$, and the β_{ik} is taken with $+$ if the function u_i has its minimum at O_k, and with — if it has its maximum there; and $\alpha_k + \sum_{i:I_i \sim O_k} \beta_{ki} > 0$ (otherwise the general interior boundary condition is reduced to $0 = 0$). The coefficients α_k is not zero if and only if the process spends a positive amount of time at the point O_k.

For a more concise formulation of the results, we introduce the following notation:

- D_i denotes the set of all points in $x \in \mathbb{R}^2$ such that $Y(x)$ belongs to the interior of the segment I_i
- $C_k = \{x : Y(x) = O_k\}$
- $C_{ki} = C_k \cap \partial D_i$

  For H being one of the values of the function $H(x)$
- $C(H) = \{x : H(x) = H\}$

  For H being one of the values of the function $H(x)$ on $\bar{D}_i$
- $C_i(H) = \{x \in \bar{D}_i : H(x) = H\}$

  For two numbers $H_1 < H_2$:
- $D_i(H_1, H_2) = D_i(H_2, H_1) = \{x \in D_i : H_1 < H(x) < H_2\}$

  For a vertex O_k and a small number $\delta > 0$
- $D_k(\pm\delta)$ is the connected component of the set $\{H(O_k) - \delta < x < H(O_k) + \delta\}$ containing C_k
- $D(\pm\delta) = \cup_k D_k(\pm\delta)$

  For a vertex O_k, $I_i \sim O_k$ and a small $\delta > 0$
- $C_{ki}(\delta) = \{x \in D_i : H(x) = H(O_k) \pm \delta\}$ (the sets $C_{ki}(\pm\delta)$ are the connected components of the boundary of $D_k(\pm\delta)$).

If D with some subscripts and the like describe a region in $\mathbb{R}^2$, then τ^ε with the same subscripts and the like denotes the first exit time of the process x_t^ε from that region; for example $\tau_k^\varepsilon(\pm\delta) = \inf\{t \geq 0 : x_t^\varepsilon \notin D_k(\pm\delta)\}$

The pictures of the domains$D_k(\pm\delta)$ and $D_i(H(O_k)\pm\delta,{}_1)$ are different for a vertex corresponding to an extremum point x_k and for one corresponding to a level curve containing a saddle point x_k, as is illustrated below.

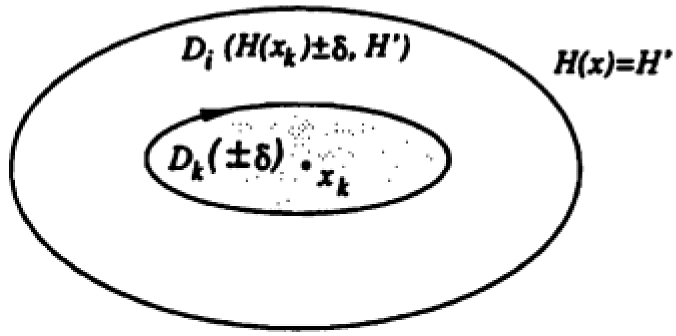

Figure 4.2: Extremum case

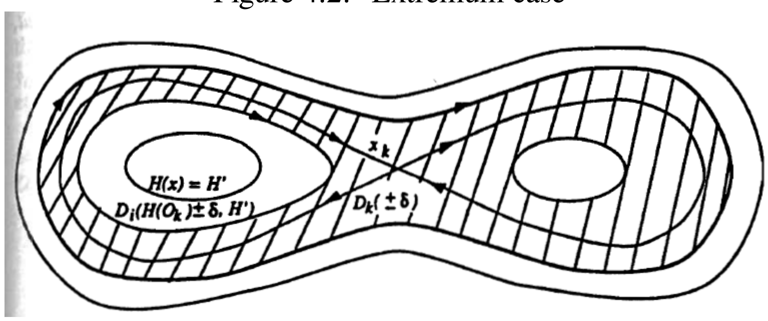

Figure 4.3: Saddle point case

According to the notation we have just introduced, we write:

$$A_i(H)=\frac{\oint_{C_i(H)}\frac{(|\nabla H(x)|)^2}{|b(x)|}dl}{\oint_{C_i(H)}\frac{1}{|b(x)|}dl}$$

$$B_i(H)=\frac{\oint_{C_i(H)}\frac{\frac{1}{2}\Delta H(x)}{|b(x)|}dl}{\oint_{C_i(H)}\frac{1}{|b(x)|}dl}$$

Integrals of this form can be computed by using the following:

Lemma 4.0.1. *Let f be a continuously differentiable function on the closed region $\bar{D}_i(H_1,H_2)$. Then for $H,H_0\in[H_1,H_2]$,*

$$\oint_{C_i(H)}f(x)|\nabla H(x)|dl=\int_{C_i(H_0)}f(x)|\nabla H(x)|dl\pm\iint_{D_i[H_0,H]}(\nabla f(x)\cdot\nabla H(x)+f(x)\Delta H(x))dx$$

where the sign $+$ or $-$ is taken whether $\nabla H(x)$ at $C_i(H)$ is pointing outside the region $D_i(H_0,H)$ or inside it, respectively. Additionally, we have:

$$\frac{d}{dH}\left(\oint_{C_i(H)}f(x)|\nabla H(x)|dl\right)=\oint_{C_i(H)}\left[\frac{\nabla f(x)\cdot\nabla H(x)}{|\nabla H(x)|}+f(x)\frac{\Delta H(x)}{|\nabla H(x)|}\right]dl$$

The integral on the denominator of A_i, B_i can be treated by this lemma by setting $f(x) = \frac{1}{|\nabla H(x)|^2}$, while the numerators will result by setting $f(x) = 1$ and $f(x) = \frac{\Delta H(x)}{|\nabla H(x)|^2}$.

By this lemma, we can conclude that $A_i(H),\ B_i(H)$ are continuously differentiable in the interior points of I_i up to $k-1$ times, if H is k times continuously differentiable.

Let us, now, consider the behaviour of H as it approaches an end of the interval $H(I_i)$, restricting ourselves only to the case where we have only non-degenerate critical points (i.e. critical points with non-degenerate Hessian matrix for the Hamiltonian).

As H approaches an end of the interval $H(I_i)$ corresponding to a non- degenerate extremum x_k of the Hamiltonian, the integral $\oint_{C_i(H)} \frac{1}{|\nabla H(x)|} dl$ which is equal to the period of the orbit $C_i(H)$, converges to:

$$T_k = \frac{2\pi}{\sqrt{\frac{\partial^2 H}{\partial q^2}(x)\frac{\partial^2 H}{\partial p^2}(x) - \left(\frac{\partial^2 H}{\partial q \partial p}(x)\right)^2}} > 0$$

$$\oint_{C_i(H)} |\nabla H(x)| dl \sim \text{constant} \cdot (H - H(O_k)) \to 0$$

and

$$\oint_{C_i(H)} \left(\frac{\Delta H)}{|\nabla H|}\right) dl \to T_k \cdot \Delta H(x_k).$$

So $A_i(H) \to 0$ as $H \to H(O_k) = H(x_k)$, $B_i(H) \to \frac{1}{2}\Delta H(x_k)$ (positive if x_k is a minimum, and negative in the case of a maximum).

In the case where $H \to H(O_k)$ where O_k corresponds to a level curve containing a non-degenerate saddle point, we have $\oint_{C_i(H)} \frac{1}{|\nabla H(x)|} dl \sim \text{constant} \cdot |\ln(H - H(0_k))| \to \infty$ and $\oint_{C_i(H)} |\nabla H(x)| dl \to \int_{C_{ki}} |\nabla H(x)| dl > 0$. In this case, $A_i(H) \to 0$ (at an inverse logarithmic rate) and again $B_i(H) \to \frac{1}{2}\Delta H(x_k)$ (which can be positive, negative or zero).

One can obtain some accurate estimates of the derivatives $A_i'(H), B_i'(H)$ as H approaches the ends of $H(I_i)$ corresponding to critical points of the Hamiltonian, but we do not need such, for, by the previous lemma, one can show that:

$$|A_i'(H)|,\ |B_i'(H)| \leq |H - H(O_k)|^{-A_0}$$

for sufficiently small $|H - H(O_k)|$, where A_0 is a positive constant.

Another useful application of that lemma is having an easy choice as to what the u_i, v_i's will be in the representation of $\mathcal{L}_i = \frac{d}{du_i}\frac{d}{dv_i}$, namely the can be chosen as:

$$v_i'(H) = \oint_{C_i(H)} \frac{1}{|\nabla H(x)|} dl$$

$$u_i'(H) = 2\left(\oint_{C_i(H)} |\nabla H(x)| dl\right)^{-1}$$

Indeed, we have that:

$$\frac{d}{dv_i}\frac{df}{du_i} = \frac{1}{u_i'(H)v_i'(H)} \cdot f''(H) + \frac{1}{v_i'(H)}\left(\frac{1}{u_i'(H)}\right)' f'(H)$$

and the first coefficient is none other than $A_i(H)$, while the second one is $B_i(H)$, noting that the derivative of $\oint_{C_i(H)} |\nabla H(x)|dl$ is equal to $\oint_{C_i(H)} \frac{\nabla H(x)}{|\nabla H(x)|}dl$. The function $v_i(H)$ can be taken equal to $\pm$ the area enclosed by $C_i(H)$.

Let us consider the vertices O_k of the graph from the point of view of their accessibility. The functions $v_i(H)$ are bounded at all vertices O_k except at O_∞. Now,

$$\lim_{H\to H(O_k)} u_i'(H) = 2\left(\oint_{C_{ki}} |\nabla H(x)|\right)^{-1}$$

(if the limit is finite, the right-hand side coincides with the one-sided derivative of u_i at O_k). This limit is finite for a vertex O_k corresponding to a separatrix containing a saddle point, and the function u_i is bounded at the end corresponding to O_k, the point O_k is accessible. Since $u_i'(H)$ has a finite positive limit at the end $H(O_k)$ corresponding to O_k, we can rewrite the interior boundary condition in the form

$$\alpha_k \mathcal{L}_i f(O_k) = \sum_{i:I_i\sim O_k} (\pm\beta_{ki]}) f_i'(H(O_k))$$

where f_i' denotes the derivative w.r.t. the local coordinate H on the $i-$th segment, the coefficients β_{ki} (that are different from those in the respective previous formula) are taken with $+$ if $H \geq H(O_k)$ on I_i and with $-$ if $H \leq H(O_k)$ on I_i.

As for a vertex O_k corresponding to an extremum x_k, we have $\oint_{C_i(H)} |\nabla H(x)|dl \sim$ constant$\cdot$ $(H - H(x_k))^2,\ u_i(H) \sim -$constant $\cdot\ (H - H(x_k))^{-1}$ as $H \to H(x_k)$, so the end O_k corresponding to an extremum x_k is inaccessible.

As for the vertex O_∞, we have that $v_i(H) = 2(u_i'(H))^{-1}$, the integral $\int_{H_0}^{\infty} v_i(H)du_i(H) = \int_{H_0}^{\infty} v_i(H)u_i'(H)dH = \int_{H_0}^{\infty} 2dH$ diverges and hence this vertex is inaccessible.

Theorem 4.0.2. *Let Γ be a graph consisting of closed segments $I_1,\ldots,I_N$ and vertices $O_1,\ldots,O_M$. Let a coordinate be defined in the interior of each segment I_i; let $u_i(y), v_i(y)$,for every segment I_i, be two functions on its interior that increase (strictly) as the coordinate increases; and let u_i be continuous. Suppose that the vertices are divided into two classes: interior vertices, for which $\lim_{y\to O_k} u_i(), \lim_{y\to O_k} v_i()$ are finite for all segments I_i meeting at O_k [notation: $I_i \sim O_k$) and exterior vertices, such that only one segment I_i enters O_k, and $\int(c + v_i(y))du_i(y)$ diverges at the end O_k for some constant For each interior vertex O_k, let b_{ki} be nonnegative constants defined for i such that $I_i \sim O_k$; $\sum_{i:I_i\sim O_k} b_{ki} > 0$. Consider the set $D(A) \subset \Gamma$ consisting of all functions f such that f has a continuous generalized derivative $\frac{d}{dv_i}\frac{d}{du_i}f$ in the interior of each segment I_i; finite limits $\lim_{y\to O_k} \frac{d}{dv_i}\frac{d}{du_i}f$ exist at every vertex O_k, and they do not depend on the segment $I_i \sim O_k$; for each interior vertex Ok,*

$$\sum_{i:I_i\sim O_k}{}' \beta_{ki} \lim_{y\to O_k} \frac{df}{du_i}(y) - \sum_{i:I_i\sim O_k}{}'' \beta_{ki} \lim_{y\to O_k} \frac{df}{du_i}(y) = 0$$

where the sum $\sum'$ contains all i such that the coordinate on the $i-$th segment has a minimum at O_k, and $\sum''$ those for which it has a maximum. Define the operator A with domain of definition $D(A)$ by$Af(y) = \frac{d}{dv_i}\frac{d}{du_i}f(y)$ in the interior of every segment I_i, and at the vertices, as the limit of this expression. Then there exists a strong Markov process $(y_t, \mathbb{P}_y)$ on Γ with continuous trajectories whose infinitesimal operator is A. If we take the space$C[0,\infty)$ of all continuous *functions on $[0,\infty)$ with values in Γ as the sample space for this process, with y_t being the value of a function of this space at the point* t*, such a process is unique. If O_k is an exterior vertex, and $\neq O_k$, then with $\mathbb{P}_y$-probability 1 the process never reaches O_k.*

We know present the general theorem of FW theory that concerns Hamiltonian systems:

Theorem 4.0.3. *Let the Hamiltonian $H(x), x \in \mathbb{R}^2$, be four times continuously differentiable with bounded second derivatives, $H(x) \geq A_1|x|^2$, $|\nabla H(x)| \geq A_2|x|$, $\Delta H(x) \geq A_3$ for sufficiently large $|x|$, where A_1, A_2, A_3 are positive constants. Let $H(x)$ have a finite number of critical points $x_1, \ldots, x_N$, at which the matrix of second derivatives is non-degenerate. Let every level curve C_k contain only one critical point x_k. Let $(X_t^\varepsilon, \mathbb{P}_x^\varepsilon)$ be the diffusion process on $\mathbb{R}^2$ corresponding to the differential operator $\mathcal{L}^\varepsilon f(x) = \frac{1}{2}\Delta f(x) + \frac{1}{\varepsilon^2} b(x) \cdot \nabla f(x)$ (recall that $b(x) = (\frac{\partial H}{\partial p}, -\frac{\partial H}{\partial q})$). Then the distribution of the process $Y(X_t^\varepsilon)$ in the space of continuous functions on $[0,\infty)$ with values in $Y(\mathbb{R}^2)(\subset \Gamma)$ with respect to $\mathbb{P}_x^\varepsilon$ converges weakly to the probability measure $\mathbb{P}_{Y(x)}$, where $(y_t, \mathbb{P}_y)$ is the process on the graph whose existence is stated in Theorem 4.0.3, corresponding to the functions u_i, v_i,- defined by formulas), and to the coefficients b_{ki} given by*

$$b_{ki} = \oint_{C_{ki}} |\nabla H(x)| dl$$

The Harmonic Oscillator

We will outline at this point what we expect based upon this theory and, ultimately, what we seek to prove with our own approach. First of all, for the harmonic oscillator we have a 2-dimensional problem, i.e. there are only two variables for this system, the position q and the momentum p, while the Hamiltonian is equal to $H(q,p) = \frac{q^2}{2} + \frac{p^2}{2}$. Furthermore, it is an elementary calculation to show that the only critical point of this system is $(0,0)$, meaning that the resulting graph- related to the aforementioned theory- will be simply the interval $[0,\infty]$ (we close the end point at infinity, since we previously discussed including a point O_∞ to represent the point at infinity in our graph).Hence, by the above equations we get:

$$L^\varepsilon = \frac{1}{\varepsilon^2}(p_t^\varepsilon, -q_t^\varepsilon) \cdot \nabla + \frac{1}{2}\Delta$$

As we've seen in the general setup, for the process $z_t^\varepsilon = H(x_t^\varepsilon)$ its generators $\mathcal{L}^\varepsilon$ have the behaviour the behaviour $\mathcal{L}^\varepsilon \xrightarrow{\varepsilon \to 0} \mathcal{L}_{FW}$, where:

$$\mathcal{L}_{FW} = \frac{1}{2} A(H) \frac{d^2}{dH^2} + B(H) \frac{d}{dH}$$

and

$$A(H) = \frac{\oint \frac{|\nabla H(x)|^2}{|b(x)|} dl}{\oint \frac{1}{|b(x)|} dl}, \quad B(H) = \frac{\oint \frac{\frac{1}{2}\Delta H(x)}{|b(x)|} dl}{\oint \frac{1}{|b(x)|} dl}$$

where the contour integrals are taken upon a level set of the Hamiltonian, i.e. a set of the following form $\{x = (q,p) \in \mathbb{R}^2 : H(x) = H\}$ for some constant H. In particular, for the harmonic oscillator we have that

$$H(x) = H \Leftrightarrow q^2 + p^2 = 2H,$$

hence

$$A(H) = \frac{\oint \frac{|(q,p)|^2}{|(p,-q)|dl}}{\oint \frac{1}{|(p,-q)|dl}} = \frac{\oint \sqrt{2H}dl}{\oint \frac{1}{\sqrt{2H}}dl} = 2H\frac{\cancel{\oint dl}}{\cancel{\oint dl}} = 2H$$

and

$$B(H) = \frac{\oint \frac{1}{\cancel{2}}\frac{\cancel{2}}{|b(x)|}dl}{\oint \frac{1}{|b(x)|}dl} = 1$$

This yields that the generator of the limit process will be $\mathcal{L}_{FW} = H\frac{d^2}{dH^2} + \frac{d}{dH}$. In other words, what we expect to prove in the next section is that as ε becomes very small, the Hamiltonian will behave as a diffusion process on $[0, \infty]$ with its generator being $\mathcal{L}_{FW}$.

Since this end up yielding a one dimensional diffusion, we will dedicate the next chapter to examine more closely the properties of such processes. Before proceeding to this, however, we will illustrate further examples and generalizations of the Freidlin Wentzell theory.

Further Examples

Example of non-harmonic potential

We considered the rather simple case of the harmonic oscillator, so far, where the calculations and the study of the resulting diffusion on $[0, \infty)$ was easy, since the "fast motion" has no effect on the Hamiltonian (we saw that the drift component was independent of ε). However, this is not the case for other Hamiltonian systems. Let us consider, for example, the Hamiltonian $H(q, p) = \frac{p^2}{2} + V(q)$ where $V(q)$ is the potential function. In the example we will examine, we will consider $V(q) = \frac{q^4}{2}$. As before, we write the system of differential equations perturbed by a white noise, and so we get:

$$\begin{cases} d\tilde{q}_t^\varepsilon = & \frac{\partial H}{\partial p}(\tilde{q}_t^\varepsilon, \tilde{p}_t^\varepsilon)dt + \varepsilon d\tilde{W}_t^1 = & \tilde{p}_t^\varepsilon dt + \varepsilon d\tilde{W}_t^1 \\ \\ d\tilde{p}_t^\varepsilon = & -\frac{\partial H}{\partial q}(\tilde{q}_t^\varepsilon, \tilde{p}_t^\varepsilon)dt + \varepsilon d\tilde{W}_t^2 = & -2(\tilde{q}_t^\varepsilon)^3 dt + \varepsilon d\tilde{W}_t^2 \end{cases}$$

or, by carrying out the time re-scaling as before, we have:

$$\begin{cases} dq_t^\varepsilon = & \frac{1}{\varepsilon^2}p_t^\varepsilon dt + dW_t^1 \\ \\ dp_t^\varepsilon = & -\frac{2}{\varepsilon}(q_t^\varepsilon)^3 dt + dW_t^2 \end{cases}.$$

Furthermore, the resulting process for the Hamiltonian is $z_t^\varepsilon = H(x_t^\varepsilon)$ (if we set once again $x_t^\varepsilon = (q_t^\varepsilon, p_t^\varepsilon)$), so by application of the Ito Formula we get:

$$dz_t^\varepsilon = \nabla H(x_t^\varepsilon)dx_t^\varepsilon + \frac{1}{2}\Delta H(x_t\varepsilon)d[x^\varepsilon]_t =$$

$$= \cancelto{0}{(2(q_t^\varepsilon)^3, p_t^\varepsilon) \cdot (\frac{1}{\varepsilon^2}p_t^\varepsilon, -\frac{2}{\varepsilon^2}(q_t^\varepsilon)^3)dt} + (2(q_t^\varepsilon)^3, p_t^\varepsilon) \cdot dW_t + \frac{1}{2}(6(q_t^\varepsilon)^2 + 1)dt \Rightarrow$$

$$\Rightarrow dz_t^\varepsilon = \frac{1}{2}[6(q_t^\varepsilon)^2 + 1]dt + (2(q_t^\varepsilon)^3, p_t^\varepsilon) \cdot dW_t$$

Our goal is to showcase that the drift and diffusion coefficients can be averaged towards the coefficients $A(H), B(H)$ illustratred above, as $\varepsilon \to 0$. We first prove that our system moves fast along the trajectories of the Hamiltonian and more slowly across them. More specifically, we will show that if $\tilde{x}_t^\varepsilon = (\tilde{q}_t^\varepsilon, \tilde{p}_t^\varepsilon)$ and x_t is the unperturbed solution to the system, then for any $\delta > 0$:

$$\mathbb{P}_x^\varepsilon[\max_{0\le t\le T} |\tilde{x}_t^\varepsilon - x_t(x)| \ge \delta] \le 3\left(\frac{e^{2LT}-1}{2L}\right)^2 \frac{\varepsilon^4}{\delta^4}$$

and, furthermore, that

$$\mathbb{E}_x^\varepsilon[|\tilde{x}_t^\varepsilon - x_t(x)|] \le \left(\frac{e^{2LT}-1}{2L}\right)^{\frac{1}{2}} \varepsilon.$$

Indeed, for this last estimate, we start by applying Ito's Formula on the quantity $|\tilde{x}_t^\varepsilon - x_t|^2$, so we have:

$$|\tilde{x}_t^\varepsilon - x_t(x)|^2 = \int_0^t 2(\tilde{x}_t^\varepsilon - x_t(x)) \cdot (b(\tilde{x}_t^\varepsilon) - b(x_t(x))dt + \int_0^t 2\varepsilon(\tilde{x}_t^\varepsilon - x_t(x)) \cdot d\tilde{W}_t + \int_0^t \varepsilon^2 dt$$

where $b(q,p) = (p, -2q^3)$ is the right-hand side of the Hamiltonian system. Clearly, this function is locally Lipschitz (that's all that interests us, since we will restrict ourselves on a trajectory), so if we consider L>0 being such a Lipschitz constant, then we have:

$$\mathbb{E}_x^\varepsilon[|\tilde{x}_t^\varepsilon - x_t(x)|^2] \le \varepsilon^2 t + 2\int_0^t L\mathbb{E}_x^\varepsilon[|\tilde{x}_s^\varepsilon - x_s(x)|^2]ds$$

Hence, by applying the Gronwall inequality, we have the estimate:

$$\mathbb{E}_x^\varepsilon[|\tilde{x}_t^\varepsilon - x_t(x)|^2] \le \frac{e^{2Lt}-1}{2L}\varepsilon^2$$

Consequently, by application of the Jensen inequality, we have:

$$\mathbb{E}_x^\varepsilon[|\tilde{x}_t^\varepsilon - x_t(x)|] \le (\mathbb{E}_x^\varepsilon[|\tilde{x}_t^\varepsilon - x_t(x)|^2])^{\frac{1}{2}} \le \left(\frac{e^{2Lt}-1}{2L}\right)^{\frac{1}{2}} \varepsilon$$

which is precisely what we opted for. Now for the probabilistic estimate, we apply the Ito Formula on $|\tilde{x}_t^\varepsilon - x_t(x)|^4$ and so, in a similar fashion as before, we have:

$$|\tilde{x}_t^\varepsilon - x_t(x)|^4 \le \int_0^t 4L|\tilde{x}_s^\varepsilon - x_s(x)|^4 ds + \int_0^t 4\varepsilon|\tilde{x}_s^\varepsilon - x_s(x)|^2(\tilde{x}_s^\varepsilon - x_s(x))d\tilde{W}_s + \int_0^t 6\varepsilon^2|\tilde{x}_s^\varepsilon - x_s(x)|^2 ds$$

By taking expectations, we obtain in a similar way that:

$$\mathbb{E}_x^\varepsilon[|\tilde{x}_t^\varepsilon - x_t(x)|^4] \leq 3\left(\frac{e^{2Lt}-1}{2L}\right)^2 \varepsilon^4$$

If we, now, consider the stopping time $\tau = \min\{t : |\tilde{x}_t^\varepsilon - x_t(x)| \geq \delta\} \wedge T$, then

$$\delta^4 \mathbb{P}_x^\varepsilon[\max_{0\leq t\leq T} |\tilde{x}_t^\varepsilon - x_t(x)|^4 \geq \delta] \leq \mathbb{E}_x^\varepsilon[|\tilde{x}_\tau^\varepsilon - x_\tau(x)|^4] \leq 3\left(\frac{e^{2LT}-1}{2L}\right)^2 \varepsilon^4 \Rightarrow$$

$$\Rightarrow \mathbb{P}_x^\varepsilon[\max_{0\leq t\leq T} |\tilde{x}_t^\varepsilon - x_t(x)|^4 \geq \delta] \leq 3\left(\frac{e^{2LT}-1}{2L}\right)^2 \frac{\varepsilon^4}{\delta^4}$$

The above estimates serve to show that for the slow motion, as $\varepsilon \to 0$ the process's distribution will convergence to the invariant measure along the trajectory. What remains to be shown is what this invariant measure will be. In order to find the invariant measure, we first need to consider the Fokker-Planck equation (or otherwise known as Forward Kolmogorov Equation):

$$\frac{\partial P}{\partial t}(x,t) = L^* P(x,t)$$

where L^* is the adjoint operator of the generator of the stochastic process. A stationary solution to the FK-equation, i.e. a candidate for the invariant measure, is the solution to the equation

$$L^* P = 0$$

In our case, we have $L^* = \frac{1}{2}\Delta - b(x)\cdot\nabla$, where again $b(x) = (p, -2q^3)$. Firstly, we observe that any constant solves the stationary FK equation $L^*P = 0$, so the uniform distribution is invariant.

Let us, now, consider the domain D_h which is the domain bounded between the level sets $C(H)$ and $C(H+h)$ of the Hamiltonian. This domain is invariant for the averaged process: if $x \in D_h$, then the averaged process will remain on D_h with probability 1. Let us call $\tilde{x}_t$ that process. Liouville's principle in analytical mechanics states that the volume of the phase space is invariant for $\tilde{x}_t$, so essentially we have that for any continuous function f defined on $\mathbb{R}^2$ it's:

$$\int_{D_h} f(x)dA = \int_{D_h} \mathbb{E}_x[f(\tilde{x}_t)]dA, \ \forall x \in D_h, \ \forall t > 0.$$

where dA is the element of area in $\mathbb{R}^2$.Let ds be an infinitesimal of the curve $C(H)$. We observe that the distance between a point on the curve $C(H)$ and $C(H+h)$ is of the form

$$\frac{h}{\nabla H(x)} + o(h) \text{ as } h \to 0.$$

Hence, by utilizing that f is a continuous function, we have that:

$$\oint_{C(H)} h\frac{f(x)}{|\nabla H(x)|}ds = \oint_{C(H)} h\frac{\mathbb{E}_x[f(\tilde{x}_t)]}{\nabla H(x)}ds + o(h)$$

hence if we divide with h and let $h \to 0$, while we also note that $|\nabla H(x)| = |b(x)|$, we get:

$$\oint_{C(H)} \frac{f(x)}{|b(x)|} ds = \oint_{C(H)} \frac{\mathbb{E}_x[f(\tilde{x}_t)]}{|b(x)|} ds$$

Hence

$$\oint_{C(H)} \frac{1}{|b(x)|} \frac{\mathbb{E}_x[f(\tilde{x}_t)] - f(x)}{t} ds = 0$$

and by taking $t \to 0$:

$$\oint_{C(H)} \frac{1}{|b(x)|} L[f] ds = 0 \Leftrightarrow \oint_{C(H)} L^* \left(\frac{1}{|b(x)|} \right) f ds = 0$$

for any continuous function f. Hence, the invariant distribution μ is proportional to $\frac{1}{|b(x)|}$ on a trajectory of the Hamiltonian. This, now, finalizes our argument, since by all of the above we have

$$\frac{1}{2}(6(q_t^\varepsilon)^2 + 1) \xrightarrow{\varepsilon \to 0} \mathbb{E}^\mu \left[\frac{1 + 6q^2}{2} \right] = B(H) = \frac{1}{\oint_{C(H)} \frac{1}{|b(x)|} ds} \oint_{q^4 + p^2 = 2H} \frac{1 + 6q^2}{2(|b(x)|)} ds =$$

$$= \frac{1}{\oint_{q^4 + p^2 = 2H} \frac{1}{\sqrt{p^2 + 4q^6}} ds} \oint_{q^4 + p^2 = 2H} \frac{1 + 6q^2}{2\sqrt{p^2 + 4q^6}} ds$$

and similarly the diffusion coefficient is averaged to

$$A(H) = \mathbb{E}^\mu [p^2 + (2q^3)^2] = \frac{1}{\oint_{C(H)} \frac{1}{|b(x)|} ds} \oint_{q^4 + p^2 = 2H} \frac{p^2 + 4q^6}{|b(x)|} ds =$$

$$\frac{1}{\oint_{q^4 + p^2 = 2H} \frac{1}{\sqrt{p^2 + 4q^6}} ds} \oint_{q^4 + p^2 = 2H} \sqrt{p^2 + 4q^6} ds$$

This concludes our analysis of the metastability for another, non-harmonic example of a Hamiltonian system which can be studied as a diffusion on the line $[0, \infty)$.

Example of a system with multiple stationary points

Until now, we have limited our explicit analysis on potentials with a single critical point ((0,0)) and so the resulting "graph" was but the line of non-negative real numbers. Hence, we will now outline an example of a Hamiltonian system the structure of which yields the metastability on a non-trivial graph.

Let us consider the Hamiltonian $H(q, p) = \frac{p^2}{2} + \frac{1}{2} q^2 (q^2 - 1)$. The unperturbed equations of this system are

$$\begin{cases} \dot{q}_t = & p_t \\ \dot{p}_t = & -2q_t^3 + q_t = -q_t(2q_t^2 - 1) \end{cases}.$$

There are three stationary points for this system, namely $(0,0)$, $(\frac{1}{\sqrt{2}},0)$ and $(-\frac{1}{\sqrt{2}},0)$. Among the three, the first one can be proved to be a saddle point, while the other two achieve a global minimum for the Hamiltonian. The following graphs allow us to get a clear view of the Hamiltonian's behaviour:

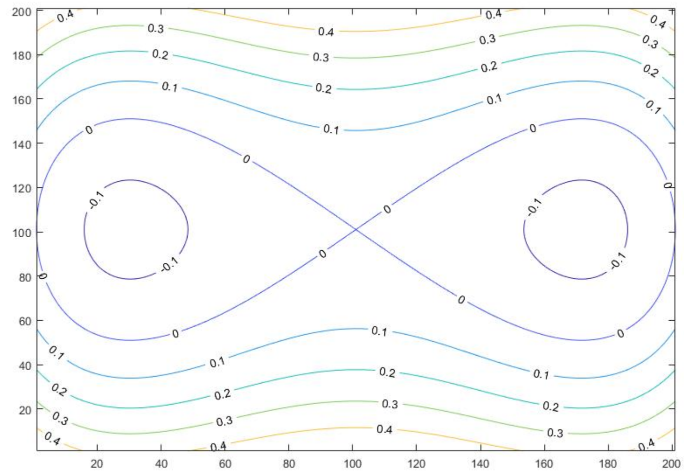

Figure 4.4: Contour plot for the example's Hamiltonian

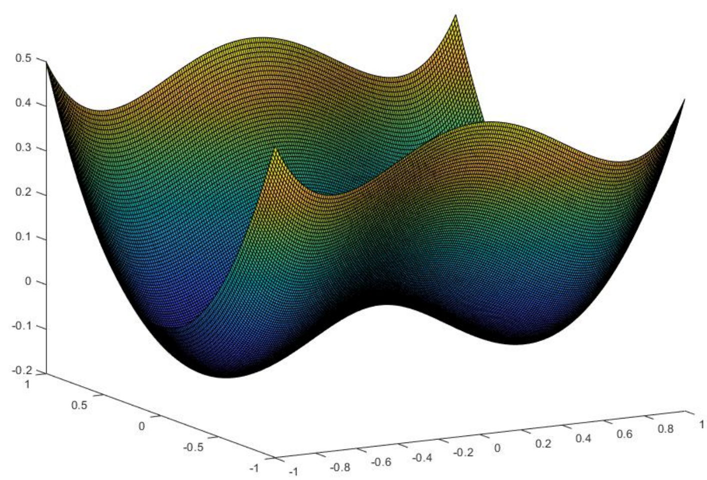

Figure 4.5: Surface plot of the example's Hamiltonian

The assessment of the equilibrium points is also visually verifiable by the above plots, with the saddle point resulting to "8"-shaped curves in the contour plot. If we identify these points with O_1, O_2, O_3 respectively and consider the connecting "edges" as I_1, I_2, I_3 (in the sense of the equivalent graph) then we get the following shape for the resulting graph (once again, O_∞ denotes the point to infinity):

We can compute the functions v_i, u_i as discussed in the relevant FW-theory, so as to examine the accessibility of each node (i.e. critical point for the Hamiltonian system). According to

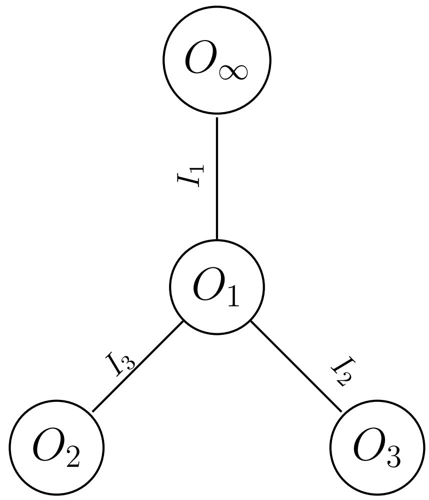

Figure 4.6: Equivalent graph to the Hamiltonian System

([**13**]) the function v_i on each edge of the graph is equal to the area enclosed by the curve $C_i(H)$, so we have

$$v_i(H) = \int_{p^2+q^2(q^2-1)\leq 2H} dqdp$$

For the equation $q^4 - q^2 - 2H = 0$ we have $\Delta = 1 + 8H$, so $(q)^2 = \frac{1\pm\sqrt{1+8H}}{2}$. If we consider the area of the edge I_1, i.e. $H \geq 0$ then there are only to real roots resulting from the above equation, namely $q_{1,2} = \pm\sqrt{\frac{1+\sqrt{1+8H}}{2}}$ (since $H \geq 0$, and so $1 \leq \sqrt{1+8H}$), hence

$$v_1(H) = \int_{-\sqrt{\frac{1+\sqrt{1+8H}}{2}}}^{\sqrt{\frac{1+\sqrt{1+8H}}{2}}} \int_{-\sqrt{2H-q^2(q^2-1)}}^{\sqrt{2H-q^2(q^2-1)}} dpdq = \int_{-\sqrt{\frac{1+\sqrt{1+8H}}{2}}}^{\sqrt{\frac{1+\sqrt{1+8H}}{2}}} 2\sqrt{2H - q^2(q^2-1)}dq$$

If we set $H = 0$, then we have

$$v_1(0) = \int_{-1}^{1} 2\sqrt{q^2(1-q^2)}dq = \int_{0}^{1} 4q\sqrt{1-q^2}dq = \left[-4\frac{(1-q^2)^{\frac{3}{2}}}{3}\right]_{q=0}^{1} = \frac{4}{3}$$

.

If, on the other hand, we consider $H \in [-\frac{1}{8}, 0]$ then the aforementioned equation has exactly four real solutions, namely $q_{1,2} = \sqrt{\frac{1\pm\sqrt{1+8H}}{2}}$ and $q_{3,4} = -\sqrt{\frac{1\pm\sqrt{1+8H}}{2}}$. Observe, also, that for symmetry reasons we will have $u_2(H) = u_3(H)$ and $v_2(H) = v_3(H)$. Hence:

$$v_2(H) = v_3(H) = \int_{\sqrt{\frac{1-\sqrt{1+8H}}{2}}}^{\sqrt{\frac{1+\sqrt{1+8H}}{2}}} \int_{-\sqrt{2H-q^2(q^2-1)}}^{\sqrt{2H-q^2(q^2-1)}} dpdq =$$

$$= \int_{\sqrt{\frac{1-\sqrt{1+8H}}{2}}}^{\sqrt{\frac{1+\sqrt{1+8H}}{2}}} 2\sqrt{2H + q^2 - q^4}dq$$

If we consider the case $H = 0$ (i.e. we examine the accessibility of the saddle point 0 from the edges I_2, I_3) we have

$$v_3(0) = v_3(0) = \int_0^1 2\sqrt{q^2 - q^4}dq = \int_0^1 2q\sqrt{1-q^2}dq = \left[-2\frac{(1-q^2)^{\frac{3}{2}}}{3}\right]_{q=0}^1 = \frac{2}{3} = \frac{1}{2}v_1(0)$$

while $v_2(-\frac{1}{8}) = v_3(-\frac{1}{8}) = 0$ (which was to be expected, since the curve is shrinked into a single point, hence there is no enclosed area).

Furthermore, we have that for $H \geq 0$:

$$u_1'(H) = \frac{2}{\int_{p^2+q^2(q^2-1)=2H} |\nabla H(x)| dl} = \frac{1}{\int_{p^2+q^2(q^2-1)=2H, p\geq 0} \sqrt{p^2 + (2q^3-q)^2} dl} =$$

$$= \frac{1}{\int_{-\sqrt{\frac{1+\sqrt{1+8H}}{2}}}^{\sqrt{\frac{1+\sqrt{1+8H}}{2}}} \sqrt{(2q^3-q)^2 + 2H - q^4 + q^2}\sqrt{1 + \frac{(q-2q^3)^2}{2H+q^2-q^4}}dq} =$$

$$= \frac{1}{\int_{-\sqrt{\frac{1+\sqrt{1+8H}}{2}}}^{\sqrt{\frac{1+\sqrt{1+8H}}{2}}} \frac{(2q^3-q)^2+2H-q^4+q^2}{\sqrt{2H+q^2-q^4}}dq}$$

If we consider $H = 0$, then

$$u_1'(0) = \frac{1}{\int_{-1}^1 \frac{(2q^3-q)^2-q^4+q^2}{\sqrt{q^2-q^4}}dq} = \frac{1}{2\int_0^1 \frac{(2q^3-q)^2-q^4+q^2}{\sqrt{q^2-q^4}}dq} = \frac{1}{2\int_0^1 \frac{4q^6-4q^4+q^2-q^4+q^2}{q\sqrt{1-q^2}}dq} =$$

$$= \frac{1}{2\int_0^1 \frac{4q^5-5q^3+2q}{\sqrt{1-q^2}}dq} = \frac{1}{2\int_0^1 \frac{q(4q^4-5q^2+2)}{\sqrt{1-q^2}}dq}$$

For the integral on the denominator, we have

$$2\int_0^1 \frac{q(4q^4-5q^2+2)}{\sqrt{1-q^2}}dq = \left[-2\sqrt{1-q^2}(4q^4-5q^2+2)\right]_{q=0}^1 + 2\int_0^1 \sqrt{1-q^2}(16q^3-10q)dq =$$

$$= 4 + 2\int_0^1 q\sqrt{1-q^2}(16q^2-10)dq = 4 + 2\left[-\frac{(1-q^2)^{\frac{3}{2}}}{3}(16q^2-10)\right]_{q=0}^1 + 2\int_0^1 (1-q^2)^{\frac{3}{2}}\frac{32}{3}qdq =$$

$$= 4 - \frac{20}{3} + \frac{64}{3}\left[-\frac{(1-q^2)^{\frac{5}{2}}}{5}\right]_{q=0}^1 =$$

$$= 4 - \frac{20}{3} + \frac{64}{15} = \frac{24}{15} = \frac{8}{5},$$

hence

$$u_1'(0) = \frac{5}{8}$$

Since this number and $v_1(0)$ are finite, we have that $u_1(H)$ is finite as H approaches zero, hence the critical (saddle) point $H = 0$ is accessible by the edge I_1.

Similarly, we have for $-\frac{1}{8} \leq H \leq 0$:

$$u_2'(H) = u_3'(H) = \frac{1}{\int_{p^2+q^2(q^2-1)=2H, p\geq 0} \sqrt{p^2 + (2q^3 - q)^2} dl} =$$

$$= \frac{1}{\int_{\sqrt{\frac{1-\sqrt{1+8H}}{2}}}^{\sqrt{\frac{1+\sqrt{1+8H}}{2}}} \frac{(2q^3-q)^2+2H-q^4+q^2}{\sqrt{2H+q^2-q^4}} dq}$$

For the point $H = 0$ we have that the denominator integral -by means of similar calculations to the ones above- will yield $\frac{4}{5}$, while for the critical point $H = -\frac{1}{8}$ we have that the respective integral is zero and so u' is unbounded, hence the two minima of the Hamiltonian for $H = -\frac{1}{8}$ are inaccessible.

Lastly, for the point at infinity, we can clearly see that $v_1(H) \to \infty$ as $H \to \infty$, hence the point O_∞ is -naturally- inaccessible as well, thus concluding or inquiry on the dynamics of the system on the resulting graph.

Chapter 5

One-Dimensional Diffusions

In this section we will concern ourselves with studying the problem $dX_t = b(X_t)dt + \sigma(X_t)dW_t$ relating it to a diffusion in a particular interval.

We start by assuming that the $(1D)$ the coefficients b, σ are continuous functions and $a(x) = \sigma^2(x) > 0$. We wish to answer to the following questions: under the condition $(1D)$, does a solution exist for the SDE? Is it unique? Does it explode?

The idea to tackle this issue is the following: we define a function ϕ so that if X_t is a solution to the SDE on $[0, \xi)$ then $Y_t = \phi(X_t)$ is a local martingale in $[0, \xi)$. Our Y_t has $[Y]_t = \int_0^t h(Y_s)ds$ for some function h, so we construct Y_t to be a solution to the Martingale problem for the coefficients $(0, h)$ by time changing the Brownian motion. Then, we define $X_t = \phi^{-1}(Y_t)$ and check that X_t is a solution to the initial SDE.

To begin to carry out this plan, we suppose that X_t is a solution to the martingale problem with coefficients (a, b) (we will denote this problem as MP(b, a)). If $f \in C^2$, then by Ito's formula:

$$f(X_t) - f(X_0) = \int_0^t f'(X_s)ds + \int_0^t \frac{1}{2} f''(X_s)d[X]_s = \text{local martingale} + \int_0^t \mathcal{L}f(X_s)ds$$

with $\mathcal{L}$ being the infinitesimal generator of X_t. It is clear from here to see that $f(X_t)$ is a local martingale iff $\mathcal{L}f = 0$. Setting $\mathcal{L}f = 0$ and remembering that from $(1D)$ we have $a(x) > 0$, we get the equation:

$$(f')' = \frac{-2b}{a} f'$$

Solving this equation, we find that:

$$f'(y) = Bexp\{\int_0^y \frac{-2b(z)}{a(z)}dz\}$$

and so we get:

$$f(x) = A + B\int_0^x Bexp\{\int_0^y \frac{-2b(z)}{a(z)}dz\}dy$$

Any of these functions can be referred to as the **natural scale**. We will select $A = 0,\ B = 1$ to get

$$\phi(x) = \int_0^x exp\{\int_0^y \frac{-2b(z)}{a(z)}dz\}dy$$

In some cases it can be useful to consider the lower endpoint of the integral other than 0. In addition, under the conditions we've assumed we have that $\phi \in C^2$, so this justifies the use of Ito's formula.

Since $Y_t = \phi(X_t)$ is by its construction a local martingale, we can time change it to a Brownian motion. We illustrate this in the following:

Theorem 5.0.1. *Let Y_t be a local martingale. Then we can change the time variable in such a way so that the resulting process is a Brownian motion, i.e. we can find a $Z_u = Y_{t(u)}$ which is a Brownian motion.*

Proof. It is known that a process Z_u is a Brownian motion iff Z_u and $Z_u^2 - u$ are local martingales. Hence, we will work towards the direction for our proof.

Let γ be the time changing mapping:

$$\gamma(u) = \inf\{t : [Y]_t > u\}$$

and let $Z_u = Y_{\gamma(u)}$. We have, first of all, that $\gamma([Y]_t) = t$ and so $Y_t = Z_{[Y]_t}$. We will show, now, that $Z_u, Z_u^2 - u$ are $\mathcal{F}_{\gamma(u)}$-local martingales (where $\mathcal{F}_t$ is Y's filtration). Let $T_n = \inf\{t : |X_t| > n\}$. From the optional stopping theorem we have that if $u < v$ then:

$$\mathbb{E}[Y_{\gamma(v)\wedge T_n}|\mathcal{F}_{\gamma(u)}] = Y_{\gamma(u)\wedge T_n}$$

Letting $n \to \infty$, we observe by Doob's maximal inequality, the fact that $Y^2_{\gamma(v)\wedge T_n} - [Y]_{\gamma(v)\wedge T_n}$ is a martingale and the definition of $\gamma(v)$ that:

$$\mathbb{E}[\sup_n Y_{\gamma(v)\wedge T_n}] \leq 4 \sup_n \mathbb{E}[X_{\gamma(v)\wedge T_n}] = 4 \sup_n \mathbb{E}[[Y]_{\gamma(v)\wedge T_n}] \leq 4v$$

From the last result and the Dominated Convergence Theorem, we have that as $n \to \infty$ it is $Y_{\gamma(t\wedge T_n} \to Y_{\gamma(t)}$ in L^2 for $t = u, v$. Since the conditional expectation is a contraction on L^2, we have that $\mathbb{E}[Y_{\gamma(v)\wedge T_n}|\mathcal{F}_{\gamma(u)}] \to \mathbb{E}[Y_{\gamma(v)}|\mathcal{F}_{\gamma(u)}]$ in L^2 and so we've shown that Z_u is a local martingale.

We continue by showing that $Z_u - u$ is a $\mathcal{F}_{\gamma(u)}$ local martingale. Again, by the optional stopping time theorem we have in a similar fashion the following:

$$\mathbb{E}\left[Y^2_{\gamma(v)\wedge T_n} - [Y]_{\gamma(v)\wedge T_n}|\mathcal{F}_{\gamma(u)}\right] = Y^2_{\gamma(u)} - [Y]_{\gamma(u)}$$

Using Young's inequality, we have:

$$\mathbb{E}[\sup_n\{Y^2_{\gamma(v)\wedge T_n} - [Y]_{\gamma(v)\wedge T_n}\}^2] \leq 2\mathbb{E}[\sup_n Y^4_{\gamma(v)\wedge T_n}] + 2\mathbb{E}[[Y]^2_{\gamma(v)}] \leq C\mathbb{E}[Y^2_{\gamma(v)}] \leq Cv^2$$

By applying again the dominated convergence theorem, we conclude the argument in the same way as before. Hence, according to the above, the process Z_u is a Brownian motion.

□

To see, now, more concretely the time change function, we note that:

$$[Y]_t = \int_0^t \phi'(X_s)^2 d[X]_s = \int_0^t \phi'(X_s)^2 a(X_s) ds = \int_0^t h(Y_s) ds$$

where $h(y) = \{\phi'(\phi^{-1}(y))\}^2 a(\phi^{-1}(y)) > 0$ is continuous. So, if we let $\tau_t = \inf\{s : [Y]_s > t\}$, then $W_t = Y_{\tau_t}$ is a Brownian time run for an amount of time $[Y]_\xi$.

To construct solutions of $MP(b,a)$ we will reverse the calculations above: we will use a time change of Brownian motion to construct a Y_t which solves $MP(O,h)$ and then let$X_t = \phi^{-1}(Y_t)$. It is possible to generalize our set-up to allow the coefficients b and u to be defined on an open interval (α, β) with $\alpha < 0 < \beta$. Since $\phi'(x) > 0$ for all x, the image of (α, β) under ϕ is an open interval (l, r) with $-\infty \leq l < 0 < r \leq \infty$.

Letting W_t be a Brownian motion, $\zeta = \inf\{t : W_t \notin (l, r)\}$, $g = \frac{1}{h}$,

$$\sigma_t = \int_0^t g(W_s) ds \text{ for } t < \zeta \;\text{ and }\; \gamma_s = \inf\{t : \sigma_t > s \text{ or } t \geq \zeta\}$$

It can be shown that $Y_s = W_{\gamma_s}$ is a solution to the martingale problem $MP(0,h)$ for $s < \xi = \sigma_\zeta$.

For the definition of X_t, we will consider ψ the inverse of ϕ and let $X_t = \psi(Y_t)$. To check that X_t solve $MP(b,a)$ until exiting the interval (α, β) at time ξ, we differentiate $\phi(\psi(x)) = x$ to get:

$$\psi'(x) = \frac{1}{\phi'(\psi(x))}$$

$$\psi''(x) = \frac{-1}{(\phi'(\psi(x))^2} \phi''(\psi(x))\psi'(x)$$

Using the fact that $\phi''(y) = \frac{-2b(y)}{a(y)} \phi'(y)$, we have:

$$\psi''(x) = \frac{1}{(\phi'(\psi(x)))^2} \frac{2b(\psi(x))}{a(\psi(x))}$$

These calculations show that $\psi \in C^2$ so using the above along with Ito's formula, we get that:

$$\psi(Y_t) - \psi(Y_0) = \int_0^t \psi'(Y_s) dY_s + \frac{1}{2} \int_0^t \psi''(Y_s) h(Y_s) ds$$

Note for the second term that from the formulae for ψ'', h (recall that $\psi = \phi^{-1}$) we have $\frac{1}{2}\psi''(y)h(y) = b(\psi(y))$. For the first term, we observe that given that Y_t solve $MP(0,h)$, there is a Brownian motion B_s such that $dY_s = \sqrt{h(Y_s)} dB_s$. Since $\psi'(y)\sqrt{h(y)} = \sigma(\psi(y))$, letting $X_t = \psi(Y_t)$ we have for $t < \xi$ that:

$$X_t - X_0 = \int_0^t \sigma(X_s) dB_s + \int_0^t b(X_s) ds$$

which indicates that the process X_t we constructed is indeed a solution to the $MP(b,a)$. We then get the following:

Theorem 5.0.2. *Consider $(C, \mathcal{C})$ and let $X_t(\omega) = \omega$. Let $\alpha < \beta$ and $\tau_{(\alpha,\beta)} = \inf\{t : X_t \notin (\alpha, \beta)\}$. Under (1D), uniqueness in distribution holds for $MP(b,a)$ on $[0, \tau_{(\alpha,\beta)})\cdot$*

Feller's test

Feller's test is a result which allows us to discern the absence of explosions in the unique solution $MP(b,a)$ defined on the interval $[0, \tau_{(\alpha,\beta)})$, i.e. if $\tau_{(\alpha,\beta)} = \infty$ a.s. We define $T_y = \inf\{t : X_t = y\}$ for $y \in (\alpha, \beta)$ and $T_\alpha = \lim_{y \downarrow \alpha} T_y$, $T_\beta = \lim_{y \uparrow \beta} T_y$. In stating this result, we make without loss of generality the assumption that $0 \in (\alpha, \beta)$. If that is not the case, one may simply select a $\gamma \in (\alpha, \beta)$ and shift the entire system by $-\gamma$.

Proposition 8. *Let $\phi(x)$ be the natural scale defined earlier and $m(x) = \frac{1}{\phi'(x)a(x)}$. Then:*
(a) $\mathbb{P}_x[T_\beta < T_0]$ is positive for some (all) $x \in (0, \beta)$ if and only if

$$\int_0^\beta dx m(x)(\phi(\beta) - \phi(x)) < \infty$$

(b) $\mathbb{P}_x[T_\alpha < T_0]$ is positive for some (all) $x \in (\alpha, 0)$ iff

$$\int_\alpha^0 dx m(x)(\phi(x) - \phi(\alpha)) < \infty$$

(c) If both integrals are finite, then $\mathbb{P}_x[\tau_{(\alpha,\beta)=\infty}] = 1$ for all $x \in (\alpha, \beta)$.

Remarks: If $\phi(\beta) = \infty$ then the first integrand is ∞ and so the integral as well is ∞. Similarly, the second integral becomes ∞ when $\phi(\alpha) = -\infty$. The statement of (a) means the following are equivalent:

- $\mathbb{P}_x[T_\beta < T_0] > 0$ for some $x \in (0, \beta)$
- $\int_0^\beta dx m(x)(\phi(\beta) - \phi(x)) < \infty$
- $\mathbb{P}_x[T_\beta < T_0] > 0$ for all $x \in (0, \beta)$.

Example: Let $\sigma(x) = 1$, $b(x) = \frac{(1+|x|)^\delta}{2}$ where $\delta > 0$. When $\delta \leq 1$, the coefficients are Lipschitz continuous, so there is no explosion. We will use Feller's test to that end and we will show that we have explosion of the solution $\delta > 1$. When $y < 0$, $\phi'(y) \geq 1$, so $\phi(-\infty) = -\infty$ and so the second integral in Feller's test is ∞. To evaluate the first integral, we have for $y > 0$ that :

$$\phi'(y) = exp\left(\int_0^y -(1+z)^\delta dz\right) = exp\left(\frac{-(1+y)^{\delta+1}+1}{1+\delta}\right)$$

so we have that $\phi(\infty) < \infty$. For Feller's test, note that $a(y) = 1$ and $m(y) = \frac{1}{\phi'(y)}$ so there is no explosion if and only if:

$$\infty = \int_0^\infty dv exp\left(\frac{(1+v)^{1+\delta}-1}{(1+\delta)}\right)\int_v^\infty exp\left(\frac{-(1+u)^{1+\delta}+1}{1+\delta}\right)du =$$

$$= \int_0^\infty dv \int_u^\infty exp\left(\frac{(1+v)^{1+\delta}-(1+u)^{1+\delta}}{1+\delta}\right) = \int_1^\infty dy \int_y^\infty dx exp\left(\frac{y^{\delta+1}-x^{\delta+1}}{1+\delta}\right)$$

where in the last step we used the substitution $y = 1 + v$ and $x = 1 + u$ to get rid of the 1's. We can, now, estimate that if $\delta > 0$ then:

$$y^{\delta} \int_{y}^{\infty} dxexp\left(\frac{y^{\delta+1} - x^{\delta+1}}{1+\delta}\right) \xrightarrow{y\to\infty} 1$$

Indeed, by using the substitution $x = y + zy^{-\delta}$ we have:

$$\int_{y}^{\infty} dxexp\left(\frac{y^{1+\delta} - x^{1+\delta}}{1+\delta}\right) = y^{-\delta} \int_{0}^{\infty} dz\, exp\left(-\int_{y}^{y+zy^{-\delta}} w^{\delta} dw\right)$$

Since $z \leq \int_{y}^{y+zy^{-\delta}} w^{\delta} dw \leq zy^{-\delta}(y + zy^{-\delta})^{\delta} \to z$ as $y \to \infty$. The estimate we provided is a result of the Dominated Convergence Theorem, hence we have the asymptotics that the above integral can be approximated by $\int_{0}^{\infty} \frac{1}{y^{\delta}} dy$ which diverges whenever $\delta \leq 1$ and converges for $\delta > 1$.

Recurrence and Transience

The natural scale, as was used for the construction of one dimensional diffusions, can also be utilized for the study of recurrence and transience. Let X_t be a solution to the problem $MP(b, a)$ and suppose the aforementioned condition (1D) is true. Let, also, the natural scale by defined by:

$$\phi(x) = \int_{0}^{x} exp\left(\int_{0}^{y} \frac{-2b(z)}{a(z)} dz\right) dy$$

Denote $T_y = \inf\{t > 0 : X_t = y\}$ and let $\tau = T_a \wedge T_b$. We start by showing:

Lemma 5.0.3. *If* $a < x < b$ *then* $\mathbb{P}_x[\tau < \infty] = 1$

Proof. We have seen that $Y_t = \phi(X_t)$ is a solution to $MP(0, h)$ where h is as defined earlier. Thus, as we've already stated, Y can be constructed as a time change of a Brownian motion, i.e. $Y_s = W_{\gamma(s)}$ where $\gamma(s) = \inf\{t : \sigma_t > s\}$ and $\sigma_t = \int_{0}^{t} \frac{1}{h(W_s)} ds$. We know, now, that the Brownian motion exits the interval $(\phi(a), \phi(b))$ with probability 1, it is evident that Y_t will exit $(\phi(a), \phi(b))$ with probability 1 and so X_t exits (a, b) with probability 1. □

We are now in position to study the recurrence and transience of one-dimensional diffusions. We know that the quantity $\phi(X_{t\wedge\tau})$ is a uniformly bounded martingale, so by the optional stopping time theorem we get:

$$\phi(x) = \mathbb{E}_x[\phi(X_\tau)] = \phi(a)\mathbb{P}_x[T_a < T_b] + \phi(b)[1 - \mathbb{P}_x[T_a < T_b]]$$

and solving this equation in terms of the probability, we have:

$$\mathbb{P}_x[T_a < T_b] = \frac{\phi(b) - \phi(x)}{\phi(b) - \phi(a)}$$

and

$$\mathbb{P}_x[T_a > T_b] = \frac{\phi(x) - \phi(a)}{\phi(b) - \phi(a)}$$

Letting $\phi(\infty) = \lim_{b\to\infty} \phi(b)$ and $\phi(-\infty) = \lim_{a\to-\infty} \phi(a)$ (these limits exist in the case where ϕ is strictly increasing), we have the following:

Theorem 5.0.4. *Suppose $a < x < b$. Then $\mathbb{P}_x[T_a < \infty] = 1$ iff $\phi(\infty) = \infty$ and $\mathbb{P}_x[T_b < \infty] = 1$ if and only if $\phi(-\infty) = -\infty$.*

In one dimension, we say that X is recurrent if $\mathbb{P}_x[T_y < \infty] = 1$ for all y. From the last theorem, we get:

Corollary 5.0.4.1. *X is recurrent iff $\phi(\mathbb{R}) = \mathbb{R}$.*

This result does not come as a surprise. We've shown that the process $Y_t = \phi(X_t)$ can be considered as a Brownian motion run for a random amount of time. So, if $\phi(\mathbb{R}) \neq \mathbb{R}$ then this random amount of time must be finite.

Green's functions

Suppose X_t is a solution to $MP(b, a)$ where a, b satisfy (1D). Let $a < b$ be real numbers and $D = (a, b)$, $\tau = \inf\{t : X_t \notin (a, b)\}$. In this section, we seek to show that if g is bounded and measurable then:

$$\mathbb{E}_x\left[\int_0^\tau g(X_s)\right] = \int G_D(x, y)g(y)dy$$

and further on to give a formula for this Green's function G_D. We initialize this study with a lemma:

Lemma 5.0.5. $\sup_{x\in(a,b)} \mathbb{E}_x[\tau] < \infty$

Theorem 5.0.6. *Suppose g is bounded. If there is a function v such that:*

1. *$v \in C^2$, $\mathcal{L}v = -g$ in (a, b) ($\mathcal{L}$ denotes the generator respective to the martingale problem at hand)*
2. *v is continuous at a, b with $v(a) = v(b) = 0$ then*

$$v(x) = \mathbb{E}_x\left[\int_0^\tau g(X_s)ds\right]$$

Proof. Let $M_t = v(X_t) + \int_0^t g(X_s)ds$. Then, for $t < \tau$ we have that:

$$v(X_t) - v(X_0) = \text{local martingale} - \int_0^t g(X_s)ds$$

so M_t is a local martingale on $[0, \tau)$. If v, g are bounded then for $t < \tau$:

$$|M_t| \leq \tau||g||_\infty + ||v||_\infty$$

which implies that the r.h. side is integrable, hence

$$M_\tau = \lim_{t \uparrow \tau} M_t = \int_0^\tau g(W_t)dt$$

So

$$v(x) = \mathbb{E}[X_0] = \mathbb{E}_x[M_\tau] = \mathbb{E}_x[\int_0^\tau g(W_t)dt]$$

$\square$

Now, on solving the theorem: it is necessary to find which function will play the part of that v. It is useful to consider the quantity $m(x) = \frac{1}{\phi'(x)a(x)}$, where $\phi(x)$ is the natural scale, and note that:

$$\frac{1}{2m(x)}\frac{d}{dx}\left(\frac{1}{\phi'(x)}\frac{df}{dx}\right) = \frac{a(x)}{2}\frac{d^2f}{dx^2} + \frac{a(x)}{2}\left(\frac{-\phi''(x)}{\phi(x)}\right)\frac{df}{dx} = \mathcal{L}f(x)$$

We will refer to $m(x)$ as the density of the speed measure (or in brevity, simply the speed measure). To solve the equation $\mathcal{L}f = -g$, we use the above relation to get

$$\frac{d}{dx}\left(\frac{1}{s(x)}\frac{dv}{dx}\right) = -2m(x)g(x)$$

Integrating once, we have:

$$\frac{1}{s(y)}\frac{dv}{dy} = \beta - 2\int_a^y dzm(z)g(z)$$

Multiplying by $s(y)$ on each side, integrating y from a to x and recalling that $v(a) = 0$ and $s = \phi'$ we have:

$$v(x) = \beta(\phi(x) - \phi(a)) - 2\int_a^x dys(y)\int_a^y dzm(z)g(z)$$

In order to have $v(b) = 0$, it must be:

$$\beta = \frac{2}{\phi(b) - \phi(a)}\int_a^b dys(y)\int_a^y dzm(z)g(z)$$

Plugging this formula in, and writting $u(x) = \frac{\phi(x)-\phi(a)}{\phi(b)-\phi(a)}$, we receive:

$$v(x) = 2u(x)\int_a^b dys(y)\int_a^y dzm(z)g(z) - 2\int_a^x dys(y)\int_a^y dzm(z)g(z)$$

Breaking the first integral $\int_a^b$ in the subintervals $[a, x]$, $[x, b]$ we get

$$v(x) = 2(u(x) - 1)\int_a^x dys(y)\int_a^y dzm(z)g(z) + 2u(x)\int_x^b dys(y)\int_a^y dzm(z)g(z)$$

Recalling that $s(y) = \phi'(y)$ we have

$$u(x)\int_x^b dys(y) = \frac{\phi(x) - \phi(a)}{\phi(b) - \phi(a)}(\phi(b) - \phi(x)) = (1 - u(x))\int_a^x dys(y)$$

Multiplying the last identity by $2\int_a^x dzm(z)g(z)$, we have:

$$2u(x)\int_x^b dys(y)\int_a^x dzm(z)g(z) = 2(1-u(x))\int_a^x dys(y)\int_a^x dzm(z)g(z)$$

Using this in the above, we can write:

$$v(x) = 2(1-u(x))\int_a^x dys(y)\int_y^x dzm(z)g(z) + 2u(x)\int_x^b dys(y)\int_x^y m(z)g(z)$$

Using Fubini's theorem results to:

$$v(x) = 2(1-u(x))\int_a^x dzm(z)g(z)\int_a^z dys(y) + 2u(x)\int_x^b dzm(z)g(z)\int_z^b dys(y)$$

If we define

$$G_D(x,z) = \begin{cases} 2\frac{\phi(x)-\phi(a)}{\phi(b)-\phi(a)}(\phi(b)-\phi(z))m(z) & \text{when } z \geq x \\ 2\frac{\phi(b)-\phi(x)}{\phi(b)-\phi(a)}(\phi(z)-\phi(a))m(z) & \text{when } z \leq x \end{cases}$$

then we have

$$v(x) = \int_a^b G_D(x,z)dz$$

Boundary Behaviour

So far, we have worked in cases in which the solution until the time it explodes. We will now present a framework through which it is possible in some times to "extend the life" of the process.

Example: Suppose $b(x) = 0$ and $\sigma(x)$ positive and continuous on $[0,\infty)$. To define a solution to $dX_t = \sigma(X_t)dW_t$ with a "reflecting boundary at 0", we extend σ to $\mathbb{R}$ by setting $\sigma(-x) = \sigma(x)$, and let Y_t be the solution to $dY_t = \sigma(Y_t)dW_t$, so that $X_t = |Y_t|$.

To use the standard procedure we developed earlier in this section, we start with X_t a solution to $MP(b,a)$ and let ϕ be the natural scale. If $Y_t = \phi(X_t)$ solves $MP(0,h)$ where

$$h(y) = \{\phi'(\phi^{-1}(y))\}^2 a(\phi^{-1}(y))$$

To see if we can start the process Y_t at 0, we use the same extension on h to the entire real line, namely we set $h(-y) = h(y)$ and let Z_t be the solution to $MP(0,h)$ on $\mathbb{R}$. Letting $\tilde{m}(y) = \frac{1}{h(|y|)}$ be the speed measure of Z_t, which is on its natural scale, we can say that:

$$\frac{1}{2}\mathbb{E}_0[\tau_{(-\varepsilon,\varepsilon)}] = \int_0^\varepsilon (\varepsilon - y)\tilde{m}(y)dy$$

Changing variables $y = \phi(x),\ dy = \phi'(x)dx,\ \varepsilon = \phi(\delta)$:

$$\frac{1}{2}\mathbb{E}_0[\tau_{(-\varepsilon,\varepsilon)}] = \int_0^\delta (\phi(\delta) - \phi(x))m(x)\frac{1}{\phi'(x)a(x)}dx$$

Now, recalling that the speed measure for the process is $m(x) = \frac{1}{\phi'(x)a(x)}$,changing notation $\phi'(z) = s(z)$ and using Fubini's theorem, we get:

$$\frac{1}{2}\mathbb{E}_0[\tau_{(-\varepsilon,\varepsilon)}] = \int_0^\delta \left(\int_x^\delta s(z)dz\right) m(x)dx = \int_0^\delta \int_0^z m(x)dxs(z)dz$$

Introducing as M the antiderivative of m to bring out the analogy with the condition of Feller's test, we have:

$$\mathbb{E}_0[\tau_{(-\varepsilon,\varepsilon)}] = 2\int_0^\delta (M(z) - M(0))s(z)dz$$

To see that this means the process cannot escape from 0, we require the following:

Lemma 5.0.7. *If* $\mathbb{P}_0[\tau_{(-\varepsilon,\varepsilon)} < \infty] > 0$ *then* $\mathbb{E}_0[\tau_{(-\varepsilon,\varepsilon)}] < \infty$

Consider, now, a diffusion on $(0, r)$ where $r \leq \infty$, let $q \in (0, r)$ and let :

$$I = \int_0^q (\phi(z) - \phi(0))m(z)dz$$
$$J = \int_0^q (M(z) - M(0))s(z)dz$$

Feller's test implies that when $I < \infty$ we can get IN to the boundary point, while the analysis above shows that when $J < \infty$ we can get OUT from the boundary point. As a result, we have the four possible combinations, which were named by Feller as follows:

I	J	name
$< \infty$	$< \infty$	regular
$< \infty$	$= \infty$	absorbing
$= \infty$	$< \infty$	entrance
$= \infty$	$= \infty$	natural

The second case is called absorbing since it is possoble to get in to the boundary point but impossible to escape it. The third is called an entrance because we cannot get to 0, however we can start the process from there. In the last case, we can neither enter nor exit 0, hence it is reasonable to exclude 0 from the state space. We will continue with providing some examples that illustrate these cases in a practical setting.

Example (Feller's branching diffusion): Let $dX_t = \beta X_t dt + \sigma\sqrt{X_t}dW_t$. Of course we want to suppose $\sigma > 0$, but we will additionally assume that $\beta > 0$, since the calculations are somewhat different in the cases where $\beta = 0$ or $\beta < 0$. Using the formula for the natural scale, we have:

$$\phi(x) = \int_0^x exp\left(\int_0^y \frac{-2\beta z}{\sigma^2 z}dz\right)dy = \int_0^x exp\left(\frac{-2\beta y}{\sigma^2}\right)dy = \frac{\sigma^2}{2\beta}\left(1 - exp\left(\frac{-2\beta x}{\sigma^2}\right)\right)$$

which maps $[0, \infty)$ to $[0, \frac{\sigma^2}{2\beta})$. The speed measure is

$$m(x) = \frac{1}{\phi'(x)a(x)} = \frac{e^{\frac{2\beta x}{\sigma^2}}}{\sigma^2 x}$$

To investigate the boundary 0, we note that

$$I = \int_0^1 m(x)(\phi(x) - \phi(0))dx = \int_0^1 \frac{1}{2\beta}(e^{\frac{2\beta x}{\sigma^2}} - 1)dx < \infty$$

since the integrant converges to $\frac{1}{\sigma^2}$ as $x \to 0$. To calculate J, we note that $m(x) \sim \frac{1}{\sigma^2 x}$ as $x \to 0$ so $M(0) = -\infty$ and so $J = \infty$. The combination $I < \infty$ and $J = \infty$ shows that the process may enter 0 but never get out, hence 0 is an absorbing point for the process.

As for the boundary at ∞, we note that:

$$\int_1^\infty m(x)(\phi(\infty) - \phi(x))dx = \int_1^\infty \frac{1}{2\beta x}dx = \infty$$

As for J, we note that when $\beta \geq 0$, $m(x) \geq \frac{1}{\sigma^2 x}$ so $M(\infty) = \infty$ and consequently $J = \infty$. In other words, we have the combination $I = \infty$ and $J = \infty$, so the point at infinity is a natural boundary for this process.

Example(Bessel process): Consider

$$dX_t = \frac{\gamma}{2X_t}dt + dW_t$$

Here $\gamma > -1$ is the index of the Bessel process. To explain the restriction for γ, we note that the radial part of a d-dimensional Brownian motion is a Bessel process with $\gamma = d - 1$. The natural scale for this process is

$$\phi(x) = \int_1^x exp\left(-\int_1^y \frac{\gamma}{z}dz\right)dy = \int_1^x y^{-\gamma}dy =$$

$$= \begin{cases} lnx & \text{if } \gamma = 1 \\ \frac{x^{1-\gamma}-1}{1-\gamma} & \text{if } \gamma \neq 1 \end{cases}$$

From the last computation we see that if $\gamma \geq 1$ then $\phi(0) = -\infty$ and $I = \infty$.

To handle $-1 < \gamma < 1$ we observe that the speed measure

$$m(z) = \frac{1}{\phi'(z)a(z)} = z^\gamma$$

So taking $q = 1$ in the definition of I:

$$I = \int_0^1 \frac{z^{1-\gamma}}{1-\gamma}z^\gamma dz < \infty$$

To compute J we observe that for any $\gamma > -1$, $M(z) = \frac{z^{\gamma+1}}{\gamma+1}$ and

$$J = \int_0^1 \frac{z^{\gamma+1}}{\gamma+1}z^{-\gamma}dz < \infty$$

Combining the above, we have that 0 is an entrance boundary if $\gamma \in [1, \infty)$ and a regular boundary if $\gamma \in (-1, 1)$.

Example(Power noise): Consider

$$dX_t = X_t^\delta dW_t$$

on $(0, \infty)$. The natural scale is $\phi(x) = x$ and the speed measure $m(x) = x^{-2\delta}$ so

$$I = \int_0^1 x^{1-2\delta} dx = \begin{cases} < \infty & \delta < 1 \\ = \infty & \delta \geq 1 \end{cases}$$

When $\delta \geq \frac{1}{2}$, $M(0) = -\infty$ and hence $J = \infty$. When $\delta < \frac{1}{2}$

$$J = \int_0^1 \frac{z^{1-2\delta}}{1-2\delta} dz < \infty$$

Combining the above conclusions, we have that the boundary point 0 is: (1) natural if $\delta \in [1, \infty)$, (2) absorbing if $\delta \in [\frac{1}{2}, 1)$ and (3) regular if $\delta \in (0, \frac{1}{2})$

Chapter 6

The Martingale Approach

6.1 The Harmonic Oscillator

Part I: Tightness

Now that we have outlined what is to be expected from our approach, we start our proof with showing the tightness of distributions that we study. Firstly, we state that it suffices to work on any time interval $[0, T]$ and to that end, without loss of generality, it suffices to work on the interval $[0, 1]$. For the remainder of this text, we will use the notation for the problem of the harmonic oscillator that we introduced in the last chapter We remind that $w_x(\delta) = \sup\{|x(t) - x(s)| : |t - s| \leq \delta\}, \delta \in [0, 1]$ is the modulus of continuity of a function x.

Theorem 6.1.1. *For the process z_t^ε, its family of distributions is tight,*

Proof. If we apply Ito's formula for the process z_t^ε, then:

$$dz_t^\varepsilon = dH(x_t^\varepsilon) = \nabla H(x_t^\varepsilon)dt + \frac{1}{2} \cdot 2d[x^\varepsilon]_t = (q_t^\varepsilon, p_t^\varepsilon)[\frac{1}{\varepsilon^2}(p_t^\varepsilon, -q_t^\varepsilon)dt + dW_t] + dt = dt + x_t^\varepsilon \cdot dW_t$$

where $W_t = (W_t^1, W_t^2)^T$. As a result, if we select $t, s \in [0, 1]$ with $|t - s| \leq \delta$, then:

$$z_t^\varepsilon - z_s^\varepsilon = t - s + \int_s^t x_u^\varepsilon \cdot dW_u$$

Then:

$$|z_t^\varepsilon - z_s^\varepsilon|^2 \leq 2(t-s)^2 + 2(\int_s^t x_u^\varepsilon \cdot dW_u)^2 \Rightarrow w_{z^\varepsilon}(\delta)^2 \leq 2\delta^2 + 2(\int_s^t x_u^\varepsilon \cdot dW_u)^2 \Rightarrow$$

$$\Rightarrow \mathbb{E}[w_{z^\varepsilon}(\delta)^2] \leq 2\delta^2 + 2\mathbb{E}[\int_s^t (x_u^\varepsilon)^2 du] = 2\delta^2 + 2\mathbb{E}[\int_s^t (q_u^\varepsilon)^2 + (p_u^\varepsilon)du] =$$

$$= 2\delta^2 + 4\mathbb{E}[\int_s^t z_u^\varepsilon du]$$

where by taking expectations we utilized the Ito isometry .However, based on the Ito formula we extracted above, we have that

$$\mathbb{E}[z_t^\varepsilon] = z_0 + t$$

and so:

$$\mathbb{E}[w_{z^\varepsilon}(\delta)^2] \leq 2\delta^2 + 4(z_0^\varepsilon(t-s) + \frac{(t-s)^2}{2}) \leq 2\delta^2 + 4(z_0^\varepsilon\delta + \frac{\delta^2}{2})$$

Consequently,

$$\lim_{\delta\to 0}\limsup_{\varepsilon\to 0}\mathbb{E}[w_{z^\varepsilon}(\delta)^2] = 0 \Rightarrow \lim_{\delta\to 0}\limsup_{\varepsilon\to 0}\mathbb{E}[w_{z^\varepsilon}(\delta)] = 0$$

Since we have this result for the modulus of continuity, it follows from Theorem 1.0.7. that the family of distributions of z_t^ε is tight. □

Having proven this result, we are certain by Prokhorov's theorem that the family of the distributions of z_t^ε is relatively compact, i.e. there exists at least a convergent subsequence to some probability measure on $C[0,1]$. This shows us that there convergent subsequences, but still there could be a multitude of possible limits. How could we venture towards showing that ultimately there is only one single limit with the desired generator?

Part II: Characterization of limits by means of the martingale problem

We've just finished proving tightness for the distributions of z_t^ε. This, as already stated, yields due to the Prokhorov theorem their sequential compactness, i.e. there exists at least a convergent subsequence. Our goal, now, will be to prove that whichever subsequence may be convergent must have a particular limit, hence the entire sequence will be converging to the process expected by the FW theory. The way we will do this is by examining the respective Martingale problem. Let $\{\mathcal{F}_s^\varepsilon\}_{s\geq 0}$ be x^ε's filtration and let g be a continuous function. Then for $\mathbb{P}$ to be the probability measure of the limit process, we want to show that:

$$g(z_t) - \int_0^t \mathcal{L}g(z_s)ds$$

is a martingale, that is

$$\mathbb{E}^{\mathbb{P}}[g(z_t) - \int_s^t \mathcal{L}g(z_u)du|\mathcal{G}_s] = g(z_s)$$

or, equivalently:

$$\int_A d\mathbb{P}g(z_t) - \int_0^t Lg(z_u)du = \int_A g(z_s)d\mathbb{P}$$

for any $A \in \mathcal{G}_s$, with $\{\mathcal{G}_s\}_s$ being the filtration of the limit process.

Theorem 6.1.2. *For any process z whose distribution is a limit of a subsequence of the distributions of z_t^ε, the generator of this process is $\mathcal{L} = \mathcal{L}_{FW}$*

Proof. It suffices to show that for any set of the form $A_k = \{(z_{t_1}, \ldots, z_{t_k} \in B\}$ where $0 \leq t_1 < \cdots < t_k \leq T$ and $B \subset \mathbb{R}^k$. First off, we have that the expression $g(w_t) - g(w_0) - \int_0^t \mathcal{L}^\varepsilon g(w_u)du$ is a $\mathbb{P}^\varepsilon-$martingale, so for a set A_k of the above form we have:

$$\int_{A_k} g(z_t) - \int_0^t \mathcal{L}^\varepsilon g(z_u) d\mathbb{P}^\varepsilon = \int_{A_k} g(z_s) - \int_0^s \mathcal{L}^\varepsilon g(z_u) d\mathbb{P}^\varepsilon$$

We have, additionally, that:

$$\mathcal{L}^\varepsilon g(w_t) = g'(w_t) + \frac{1}{2}|x_t|^2 g''(w_t) = g'(w_t) + w_t g''(w_t) = \mathcal{L}_{FW} g(w_t)$$

Note that we have that $\mathcal{L}^\varepsilon$ is in fact independent from ε. Furthermore, if $M_t^\varepsilon = g(w_t) - \int_0^t \mathcal{L}_{FW} g(w_u) du$ then :

$$\int_{A_k} M_t d\mathbb{P}^\varepsilon = \int_{A_k} M_s d\mathbb{P}^\varepsilon$$

To conclude the proof, we must send $\varepsilon \to 0$ and "substitute" in this way the measures $\mathbb{P}^\varepsilon$ with $\mathbb{P}$. Then, due to $\mathcal{L}$ being uniformly elliptic, we will have uniqueness for the solution to the martingale problem and thus have characterized the limit process. However, we must make sure that the continuity sets of the limit measure $\mathbb{P}$ include the Borel sets induced by the limit process z_t.

Let $\mathcal{A}_k = \{A \in \mathcal{G}_{t_k} : \mathbb{P}[\partial A] = 0\}$. We can clearly see the following:

1. $$\mathbb{P}[\partial \Omega] = \mathbb{P}[\bar{\Omega} \backslash \Omega] = 0$$

2. If $A \subset B \in \mathcal{A}_k$, then:

$$\partial(B \backslash A) \subset \partial B \cup \partial A \Rightarrow \mathbb{P}[\partial(B \backslash A)] \leq \mathbb{P}[\partial A \cup \partial B] \leq \mathbb{P}[\partial A] + \mathbb{P}[\partial B] = 0$$

3. Let $\{A_n\}_{n \in \mathbb{N}}$ increasing with $\mathbb{P}^\varepsilon[\partial A_n] = 0, \ \forall n \in \mathbb{N}$. This means that $\mathbb{P}[\bar{A}_n] = \mathbb{P}[int(A_n)]$. Then:

$$\mathbb{P}n[\partial(\cup_{n=1}^\infty A_n)] = \mathbb{P}^\varepsilon[\cup_{n=1}^\infty \bar{A}_n] - \mathbb{P}[int(\cup_{n=1}^\infty A_n)] \leq \mathbb{P}^\varepsilon[\cup_{n=1}^\infty \bar{A}_n] - \mathbb{P}[\cup_{n=1}^\infty int(A_n)] = 0$$

   since

$$\mathbb{P}[\cup_{n=1}^\infty \bar{A}_n] = \lim_{n \to \infty} \mathbb{P}[\bar{A}_n] = \lim_{n \to \infty} \mathbb{P}^\varepsilon[int(A_n)] = \mathbb{P}[\cup_{n=1}^\infty int(A_n)]$$

   Hence, $\mathbb{P}[\partial(\cup_{n=1}^\infty A_n)] = 0$

From the above, we have ensured that the class $\mathcal{A}_k$ we have defined is a Dynkin system. Furthermore, it is clear that that the borel sets of the form $(a_1, b_1) \times \cdots \times (a_k, b_k) \times \ldots$ belong in $\mathcal{A}_k$, since the $\mathbb{P}^\varepsilon$ can be transformed into Wiener measures. Indeed, as we've mentioned before we have that $\int_0^t x_u^\varepsilon \cdot dW_u = w(\int_0^t (x_u^\varepsilon)^2 du)$, where w is a standard Brownian motion. Then, by applying Girsanov's theorem, the process $z_t^\varepsilon - t$ is a Brownian motion under a new measure $\mathbb{Q}^\varepsilon$. In addendum, for a cylinder set of the form we mentioned, i.e. $A_k(a_1, b_1) \times \cdots \times (a_k, b_k) \times \mathbb{R} \times \ldots$, we have:

$$\mathbb{P}[\partial A_k] = \int_{\{a_1,b_1\}\times\cdots\times(a_k,b_k)} \frac{1}{\sqrt{2^k\pi^k t_1(t_2-t_1)\dots(t_k-t_{k-1})}} exp\{\frac{-x_1^2}{2t_1}+\cdots+\frac{-x_k^2}{2(t_k-t_{k-1})}\}dx_1\dots dx_k +$$

$$\cdots + \int_{(a_1,b_1)\times\cdots\times\{a_k,b_k\}} \frac{1}{\sqrt{2^k\pi^k t_1(t_2-t_1)\dots(t_k-t_{k-1})}} exp\{\frac{-x_1^2}{2t_1}+\cdots+\frac{-x_k^2}{2(t_k-t_{k-1})}\}dx_1\dots dx_k = 0$$

since we transfer our calculation to an k-dimensional Gaussian measure, which is absolutely continuous to the Lebesgue k-dimensional measure, hence the set we've introduced above must have measure 0.

Furthermore, the aforementioned cylinder sets are closed under intersections, and so by the $\pi - \lambda$ theorem, we have the desired property about the continuity sets of $\mathcal{G}_{t_k}$. Hence, we 've showed that the $\mathcal{L}_{FW}$ solves the martingale problem for the harmonic oscillator.

Regarding the uniqueness of the solution, we said in the relevant theory that we need only examine the diffusion part of the generator, showing that it is uniformly elliptic. In our case, we have for the process x_t^ε that the diffusion part of the generator is

$$L^{\varepsilon\prime} = \frac{1}{2}\Delta$$

If we represent this operator by means of a matrix, we have $a(x) = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$, hence we get that $\sum_{i,j=1}^2 a_{ij}(x)y_iy_j = \frac{|y|^2}{2}$, that is $L^{\varepsilon\prime} \in \mathcal{N}(\frac{1}{2}, 0)$ (with the inequalities being converted into equality in this case). Consequently, we have the desire uniqueness to the martingale problem, thus the limit process's generator is equal to $\mathcal{L}_{FW}$.

□

Study of the boundaries

Now that we have concluded our proof, we can examine the nature of the boundary points for the resulting diffusion for the system of the harmonic oscillator. Based on the framework we introduced for 1-d diffusions, it is $a(h) = 2h$ and $b(h) = 1$ for the limit system of the harmonic oscillator. The natural scale for this problem is:

$$\phi(h) = \int_1^h exp\left(\int_1^x \frac{-2b(y)}{a(y)}dy\right)dx = \int_1^h exp\left(\int_1^x \frac{-1}{y}dy\right)dx = \int_1^h exp(-lnx)dx = \int_1^h \frac{1}{x}dx = ln(h)$$

It is evident that $\phi(0) = -\infty$, hence - referring to the notation in the relevant chapter- we have $I = \infty$. Furthermore, the speed measure is $m(h) = \frac{1}{\phi'(h)a(h)} = \frac{1}{\frac{1}{h}2h} = \frac{1}{2} \Rightarrow M(h) = \frac{h}{2}$. Hence, we have:

$$J = \int_0^1 (M(z) - \cancelto{0}{M(0)})\phi'(z)dz = \int_0^1 \frac{z}{2}\cdot\frac{1}{z}dz = \frac{1}{2} < \infty$$

Consequently, based on Feller's classification of boundary points for 1-d diffusions, the point 0 is an entrance point for the process. This is a reasonable outcome, since we can start a process at 0, but if the system starts from a state of positive energy, it is not going to loose this amount of energy completely.

We then proceed to the point at infinity. Again, we have from the fact $\phi(\infty) = ln(\infty) = \infty$ that $I = \infty$. Additionally $M(\infty) = \infty \Rightarrow J = \infty$, hence the point at infinity is a natural boundary for the system. Again a logical finding, since it is practically impossible to start, reach or leave the level of infinite energy, so the fact that this point is a natural boundary was something to be expected.

Chapter 7

Conclusions

We have seen, through the elementary problem of the harmonic oscillator, that there is an alternative to describing the metastability of a Hamiltonian system. The Harmonic oscillator, however, is a rather ideal problem in this setting, while in other Hamiltonian systems proving tightness and deriving the FW through respective calculations is a more laborious procedure. Furthermore, this diploma thesis leaves open the issue of Hamiltonian systems with multiple extrema with the approach currently presented. As we've seen in the classical Freidlin-Wentzell theory, such systems can be studied as we presented in each section separately and upon the extrema the accessibility must be examined separately. The modification that this procedure yields is, essentially, a set of restrictions for the functions inserted into the operators $\mathcal{L}_i$ at the points of the extrema.

To sum it all up, the method presented here seems as a viable alternative to the clasical theory, leading to a more immediate derivation of metastability, exploiting the good properties of martingale problems (which have been studied extensively), rather than relying on the extensive and intricate classical theory.

Bibliography

[1] Landim, Claudio, D. Marcondes, and I. Seo. ”A resolvent approach to metastability.” arXiv preprint arXiv:2102.00998 (2021).

[2] Koumoulis, G., Negrepontis S., ”Θεωρία Μέτρου”, Συμμετρία Publication, 2005

[3] Cheliotis, D. (2015). ”Ένα δεύτερο μάθημα στις πιθανότητες” [Undergraduate textbook]. Kallipos, Open Academic Editions.

[4] Durrett, Rick. Probability: theory and examples. Vol. 49. Cambridge university press, 2019.

[5] Dudley, Richard M. Real analysis and probability. CRC Press, 2018.

[6] Stroock, Daniel W. Probability theory: an analytic view. Cambridge university press, 2010.

[7] Billingsley, Patrick. Convergence of probability measures. John Wiley Sons, 2013.

[8] Kallenberg, Olav, and Olav Kallenberg. Foundations of modern probability. Vol. 2. New York: springer, 1997.

[9] Haugh, Martin. ”Introduction to stochastic calculus.” Financial Engineering: Continuous-Time Models (2010).

[10] Lawler, Gregory F. ”Stochastic calculus: An introduction with applications.” American Mathematical Society (2010).

[11] Norris, James R. Markov chains. No. 2. Cambridge university press, 1998.

[12] Williams, David. Probability with martingales. Cambridge university press, 1991.

[13] Freĭdlin, Mark Iosifovich, and Alexander D. Wentzell. Random perturbations of Hamiltonian systems. Vol. 523. American Mathematical Soc., 1994.

[14] Freidlin, Mark, and Matthias Weber. ”On random perturbations of Hamiltonian systems with many degrees of freedom.” Stochastic processes and their applications 94.2 (2001): 199-239.

[15] Durrett, Richard. Stochastic calculus: a practical introduction. Vol. 6. CRC press, 1996.

[16] Fischer, Markus, and Giovanna Nappo. ”On the moments of the modulus of continuity of Itô processes.” Stochastic Analysis and Applications 28.1 (2009): 103-122.