Asymptotic Analysis for Markovian models in non-equilibrium Statistical Mechanics

by

Michela Ottobre

Department of Mathematics
Imperial College London
London SW7 2AZ
United Kingdom

Submitted to Imperial College London
for the degree of
Doctor of Philosophy

2011
Declaration

I the undersigned hereby declare that the work presented in this thesis is my own. When material from other authors has been used, these have been duly acknowledged. This thesis has not previously been presented for this or any other PhD examinations.

Michela Ottobre
Abstract

This thesis is mainly concerned with the problem of exponential convergence to equilibrium for open classical systems. We consider a model of a small Hamiltonian system coupled to a heat reservoir, which is described by the Generalized Langevin Equation (GLE) and we focus on a class of Markovian approximations to the GLE. The generator of these Markovian dynamics is an hypoelliptic non-selfadjoint operator. We look at the problem of exponential convergence to equilibrium by using and comparing three different approaches: classic ergodic theory, hypocoercivity theory and semiclassical analysis (singular space theory). In particular, we describe a technique to easily determine the spectrum of quadratic hypoelliptic operators (which are in general non-selfadjoint) and hence obtain the exact rate of convergence to equilibrium.
To my grandmother
Acknowledgments

Thanks to many many people and things,
thanks to my supervisor for inspiration, thoughtful guidance and for seconding my natural
incline while guiding me into scouting for new ideas and theories;
and to Paolo Buttá, for having introduced me to research and for having taught me to always
understand what is true and what is not before looking at the technical aspects;
and to Boguslaw Zegarlinski, for ideas, lovely and stimulating discussions and for
genuinely caring about these early stages of my career;
and for the exact same reason I thank many other people and the list is long, so we will leave
it to the reader, who surely knows what I am talking about; among others, my special thanks
surely go to D. Crisan, M. Hairer, A. Iserles, A. Jacquier, G. Moore, A. O. Parry, A. Stuart
and F. Theil;
and thanks to the Imperial College, for being a wonderful research environment;
and to the reading groups;
and to the other PhD students, all of them, for laughs, thoughts and memories;
and to London, the streets, the Opera and the lights;
and to Calvino, Montale and Moravia.
And to my parents;
and to my friends.
And, above all, thanks to my sister,
and to my grandmother, for seeing me the way she does.
Contents

Abstract 3

Acknowledgments 5

1 Introduction: open classical systems 15
   1.1 The Generalized Langevin Equation as a model in non-equilibrium statistical mechanics 15
   1.2 Markovian approximations of the GLE 19

2 Exponential Decay to equilibrium: Ergodicity, Hypocoercivity and Singular Space Theory 24
   2.1 Hypoellipticity 25
   2.2 Ergodic Theory for hypoelliptic diffusions 26
   2.3 Hypocoercivity 31
   2.4 Semiclassical Approach 37
      2.4.1 Basics of Semiclassical Analysis 37
      2.4.2 Singular Space theory and Spectral theory for quadratic hypoelliptic operators 39
   2.5 Comparison between hypocoercivity theory and singular space theory 45

3 Asymptotic Analysis for the Generalized Langevin Equation 46
   3.1 Statement of Main Results 48
      3.1.1 Notation 54
   3.2 Convergence to Equilibrium 55
      3.2.1 Hypocoercivity 56
      3.2.2 Proof of Theorem 3.1.3 57
Overview

The body of knowledge that goes under the name of Statistical Mechanics was initiated by Boltzmann, Gibbs, Maxwell and many others during the second half of the 1800’s. While equilibrium phenomena are completely understood, non-equilibrium Statistical Mechanics is still an open research field, encompassing a great variety of phenomena. Typically, the systems of interest in Statistical mechanics are many-particle systems where each particle moves according to deterministic laws; hence in principle the classic laws of motion can be used to describe the dynamics. Though this is in practice not doable and probability swoops in due to the large number of particles involved. Hence the state of the system, instead of being described by a point in state space, is described by a probability distribution.

Given a system in equilibrium, we can drive it away from its stationary state by either coupling it with one or more large Hamiltonian systems or by using non-Hamiltonian forces. The latter approach is used mainly to observe non-equilibrium steady states: if we apply non-Hamiltonian forces, energy is not conserved and the system, which is not constrained to a compact energy level, heats up; in order to keep it far from equilibrium while absorbing the excess of heat, it is then thermostatted ([71]). In the Hamiltonian approach, which is often referred to as the open systems theory, the system we are interested in is coupled to one or more heat reservoirs. Multiple reservoirs are used to study non equilibrium steady states; coupling with a single reservoir is used to study return to equilibrium, which is the main subject of the present thesis. Before getting into the matter of decay to equilibrium and giving an overview of the content of the following chapters, just one comment about irreversibility. The first result about relaxation to equilibrium, namely the Boltzmann $H$–theorem, had to face strong objections in the measure in which it raised an apparent contradiction between the macroscopic irreversibility and the microscopic reversibility of isolated many-particle systems; indeed, according to Poincaré’s recurrence theorem, the system should come back arbitrarily
close to its initial state (far from equilibrium) infinitely many times. In other words, if the microscopic dynamics of each particle is reversible, why is it not so for the macroscopic evolution? The paradox is only apparent, indeed the system does return to its initial configuration infinitely many times; though, the time of return increases with the number of particles, so that for the typical size of a system considered in Statistical Mechanics it can be calculated to be more than the age of universe! A more technically detailed answer to this question can be found in Lebowitz's review paper [46].

Chapter 1 shall provide more details on how to model a heat reservoir. For the time being, let us introduce the Generalized Langevin Equation (GLE):

\[
\ddot{q}(t) = -\partial_q V(q) - \int_0^t ds \gamma(t-s)\dot{q}(s) + F(t),
\]

(0.0.1)

which is a popular model for a particle immersed in a heat bath and has proven to be a very efficient tool in molecular dynamics. In (0.0.1), \( q(t) \) represents the position of the distinguished particle (here \( q(t) \in \mathbb{R} \) just for simplicity, the equation can be rewritten in \( \mathbb{R}^n \)), \( V = V(q) \) is a potential, \( \gamma(t) \) is a smooth kernel and \( F(t) \) is a mean zero stationary Gaussian process. Noise and memory kernel are related through the following fluctuation dissipation principle

\[
E(F(t)F(s)) = \beta^{-1} \gamma(t-s),
\]

(0.0.2)

i.e., the correlation function of the noise is proportional to the memory kernel through a constant \( \beta \) (inverse temperature of the bath). The GLE is a stochastic integro-differential equation and, for an arbitrary choice of the kernel \( \gamma(t) \), it is in general non-Markovian and thus less amenable to analysis. This is what brings us to consider Markovian approximations of the GLE. The general form of a Markovian system of ODEs which approximates the dynamics (0.0.1) reads as follows [39]:

\[
\begin{align*}
\dot{q} &= p \, dt \\
\dot{p} &= -\partial_q V(q) \, dt + g \cdot u \, dt \\
\dot{u} &= (-pg - A u) \, dt + C\,dW(t),
\end{align*}
\]

(0.0.3a)

(0.0.3b)

(0.0.3c)

where \((q,p) \in \mathbb{R}^2, \ u \) and \( g \) are column vectors of \( \mathbb{R}^d \), \( \cdot \) denotes Euclidean scalar product, \( W(t) = (W_1(t), \ldots, W_d(t)) \) is a \( d \)-dimensional Brownian motion and \( A \) and \( C \) are constant coefficients \( d \times d \) matrices, related through the fluctuation dissipation principle, which in the
The present case becomes
\[ A + A^T = CC^T. \]  

In (0.0.3) the variables \((u_1, \ldots, u_d)\) describe the heat bath. The simplest form of the above system is obtained when considering \(A\) to be a diagonal matrix, \(A = \text{diag}\{\alpha_1, \ldots, \alpha_d\}\) in which case the heat bath equations decouple and (0.0.3) turns into

\[
\begin{align*}
dq &= q(t)dt \\
\frac{dp}{dt} &= -\partial_q V(q(t))dt + \sum_{j=1}^{d} \lambda_j p_j(t)dt \\
\frac{du_j(t)}{dt} &= -\lambda_j p_j(t)dt - \alpha_j u_j(t)dt + \sqrt{2\alpha_j \beta^{-1}}dW_j, 
\end{align*}
\]

for \(j = 1, \ldots, d, \alpha_j > 0\) and \(\lambda_j > 0\). The notation for \(q\) and \(p\) in (0.0.3) and (0.0.5) should include a subscript \(d\), i.e. \(q_d, p_d\), as the solution will depend on the number of heat bath variables \(u_j\); we drop the subscript for notational convenience. More details on how to obtain the above system (0.0.5) will be given in Section 1.2. Under some assumptions on the matrix \(A\) and on the vector \(g\), the generator \(L\) of (0.0.3) is a hypoelliptic non-self adjoint operator; the spectral theory for operators of this type lacks of general results.

This thesis is mainly concerned with the problem of exponential convergence to equilibrium for degenerate Markovian systems; more precisely, we will focus our attention on Markovian systems with hypoelliptic and non-selfadjoint generator. In system (0.0.3), the degeneracy (hypoellipticity) is a consequence of the noise acting on the heat bath variables only. The idea that we wish to convey is that the problem of exponential convergence to equilibrium for this class of systems can be regarded by the standpoint of three different (yet closely related) theories: from the perspective of classic ergodic theory, adapted to the context of hypoelliptic diffusions ([51], [52]); using the functional analytic approach provided by the theory of hypocoercivity ([76]) and finally by the point of view of semiclassical and spectral analysis, in particular with the major help of the singular space theory ([62]). Basic facts about these three approaches and their links with the notion of hypoellipticity are presented in Chapter 2. We define a process to be ergodic when it admits a unique invariant measure. A process is geometrically ergodic when convergence to the equilibrium measure is exponentially fast. Classic ergodic theory provides a natural setting to study this phenomenon, through the combination of two key ingredients: the first one is the existence of a Lyapunov function, i.e. a function \(G(x)\) defined on the phase space, with compact level sets and satisfying
appropriate bounds involving the generator (see (2.2.2) and (2.2.5)). Once such a Lyapunov function is found, an invariant measure can be explicitly constructed, hence the existence of the Lyapunov function implies the existence of the equilibrium measure. The second ingredient is the so called aperiodicity and irreducibility of the process: for any open set and any \( x \) in the state space, \( P_t(x, A) > 0 \) for some \( t > 0 \), where \( P_t(\cdot, \cdot) \) is the transition probability of the Markov process. If the process is irreducible and aperiodic then the invariant measure is unique. Notice that this condition is the one containing the intuition about the notion of ergodicity: roughly speaking, the system explores the whole phase space. The ideal setting to study irreducibility and aperiodicity is provided by control theory, which is classically formulated for deterministic systems: the link between deterministic and stochastic control is then provided by Stroock-Varadhan support theorem (Appendix A.1). Once the unique invariant measure is constructed, the speed of convergence to the stationary state depends on the bounds satisfied by the Lyapunov function \( G(x) \). In the situations that we will examine, we shall always obtain exponentially fast convergence to equilibrium.

The theory of hypocoercivity, instead, is a purely functional analytic theory. It refers to dissipative Markovian evolutions whose generator can be written in the form

\[
\mathcal{L} = B + \sum_{i=1}^{n} A_i^* A_i,
\]

where \( A_i^* \) denotes the adjoint of \( A_i \) in the space \( L^2_\rho := \{ f \in L^2 : \int f \, d\mu < \infty \} \) and \( \mu \) is the invariant measure of the system, with density \( \rho \) with respect to the Lebesgue measure, i.e. \( d\mu(x) = \rho(x) \, dx \) (to be precise, here and throughout the thesis we denote by \( L \) the generator and \( \mathcal{L} = -L \)). The operator \( B \) is antisymmetric in \( L^2_\rho \), whereas \( \sum_{i=1}^{n} A_i^* A_i \) is clearly symmetric; hence the dynamics is nicely decomposed into a conservative (deterministic) part, described by \( B \) and a stochastic (dissipative) component, described by \( A \). Appropriate bounds on the successive commutators between \( A \) and \( B \) together with a Poincaré inequality guarantee hypocoercivity, that is, exponential convergence to equilibrium: there exist \( \kappa, \lambda > 0 \) such that

\[
\| e^{-Tt} h \|_{\tilde{H}} \leq \kappa e^{-\lambda t} \| h \|_{\tilde{H}} \quad \forall h \in \tilde{H} \text{ and } t \geq 0,
\]

where \( \tilde{H} \) is a Hilbert space which we shall be more specific about in Section 2.3, typically the space \( H^1 \) weighted by the invariant measure (modulo constants). Because the definition of hypocoercivity is invariant under a change of equivalent norm, the bulk of the theory is the construction of an appropriate auxiliary scalar product, \( (\cdot, \cdot) \), equivalent to the product
associated to the $\tilde{H}$ norm in which the operator is easily seen to be coercive (see Proposition 2.3.1). Both in Section 3.3 and in Chapter 6, we make use of Lyapunov functions which are designed in the spirit of [27] (and according to the same logic that leads to considering the auxiliary scalar product $((\cdot, \cdot))$) in order to determine the short time behaviour of the semigroup generated by $L$. Notice however that the bounds of Section 3.3 are in $L^2_\rho$ and they concern only first order derivatives, those in Chapter 6 are pointwise estimates on derivatives of any order. However we reckon that the proof of Section 3.3 can be extended to $n-th$ order derivatives, as well. The bounds we obtain are sharp and can be obtained also by using Malliavin calculus techniques ([56]). The operator $L$ can be recast in Hörmander sum of square’s form and theorems about hypocoercivity make use of successive commutators, as Hörmander’s theory of hypoellipticity does; nevertheless, in the study of regularity of solutions, we typically need to consider all the possible successive commutators between $A_i$ and $A_j$ and those between $A_i$ and $B$ whereas as far as convergence to equilibrium is concerned, only the commutators of the form $[A_i, B], [[A_i, B], B]$ etc are needed. A priori it makes no sense to compare the two concepts, hypoellipticity and hypocoercivity, simply because, even if they refer to analogous situations, they are concerned with solving different problems. Nonetheless, the two techniques are close in various respects (see Section 2.5).

Last, the singular space theory is based on semiclassical analysis and, in the form in which we shall use it, it refers to hypoelliptic quadratic evolutions in the flat $L^2$ i.e. to evolution equations whose generator is an hypoelliptic operator with quadratic Weyl symbol. The main results that we will use have been obtained in [62] by the third author. In particular, the singular space theory not only allows to establish exponential convergence to the stationary state, but it also proves to be a very strong tool to determine the whole spectrum of quadratic hypoelliptic operators, hence in particular the exact rate of convergence. We recall that these operators are in general non self-adjoint. To a quadratic operator $L$ we can associate in a unique way a matrix, the so called Hamilton map of the operator. The surprising outcome of this theory is that the spectrum of such operators is discrete and it consists of integer combinations of the eigenvalues of the Hamilton map which have positive imaginary part. Hence the calculation of the spectrum of the operator is reduced to the calculation of the spectrum of a matrix.

This thesis is organized as follows: Chapter 1 is an introductory chapter, which contains background material on the GLE, its derivation and its Markovian approximations. Chapter 2, after some preliminaries on hypoellipticity, presents the theory of geometric ergodicity, the
theory of hypocoercivity and the singular space theory. In particular, Section 2.5 contains a comment about the relation between these theories and the notion of hypoellipticity. As an example to which all the above mentioned approaches do apply, in Chapter 3 we study several properties of the approximating system (0.0.5); in particular, after proving the well-posedness of the semigroup associated to this dynamics (Proposition 3.2.1), we obtain exponential convergence to equilibrium both by using hypocoercivity theory and via geometric ergodicity. Also, we prove an homogenization result (Theorem 3.1.8) by using a functional central limit theorem. This requires the study of the Poisson problem associated with the generator, and therefore the compactness of the resolvent of $L$ (Proposition 3.1.9). The last result of Chapter 3 is a white noise limit (Theorem 3.1.10). Chapter 3 is the content of [60].

In Chapter 4 we consider the more general system (0.0.3), for which one could in principle carry out the same analysis done in Chapter 3. Instead, the main result of the section is that, under the sole assumption of hypoellipticity, system (0.0.3) is ergodic. This is done by using the theory of geometric ergodicity; our point is that the assumption of hypoellipticity gives precise instructions on how to construct the Lyapunov function needed to get uniqueness of the stationary state. Also, a classic result gives the controllability of the dynamics, as soon as the evolution is hypoelliptic. We think that the same kind of argument can be reproduced for a general Ornstein-Uhlenbeck process. In this chapter we also discuss the problem of the approximation itself which is, to the best of our knowledge, still an open problem. It is easy to show that when the kernel $\gamma(t)$ in (0.0.1) is a sum of exponentials of the type

$$\gamma(t) = \sum_{i=1}^{d} \lambda_i^2 e^{-\alpha_i t}, \quad \alpha_i > 0,$$

then the GLE is equivalent to system (0.0.5). Hence the natural idea is, roughly, to approximate an arbitrary function $\gamma(t)$ through sums of exponentials and in this way obtain a Markovian approximation of the non Markovian GLE. It turns out that the approximation of functions through sums of the type (0.0.6) is a quite hard task in approximation theory and in Section 4.2 we explain why. However, the following Section 4.3 makes sense, in an appropriate Hilbert setting, of the infinite dimensional approximation by proving the well-posedness of system (0.0.5) when we add infinitely many heat bath variables, i.e. when $d \to \infty$.

In Chapter 5 we show that the singular space theory can be employed to calculate the $L^2$ spectrum of the degenerate Ornstein-Uhlenbeck process, namely of the generator of the
process
\[ dX = -BX dt + \Sigma dW, \quad (0.0.7) \]

where \( B \) and \( \Sigma \) are \( n \times n \) matrices and \( \Sigma \Sigma^T = Q \) is degenerate (\( \det Q = 0 \)). The spectrum of the generator of (0.0.7), \( \sigma(L) \), was calculated in [53], where the authors show that \( \sigma(L) \) is given by integer combinations of the eigenvalues of \( B \) and hence it depends only on the drift. In Section 5.2 we show the equivalence between the result obtained in [53] and the one obtained via the calculation of the Hamilton map (which is a \( 2n \times 2n \) matrix as opposed to \( B \), which is a \( n \times n \) matrix). As an example on which calculations are easy and explicit, in Section 5.3 we apply the singular space theory to system (0.0.5).

Combining a classic semigroup approach ([5]) with the ideas contained in [27] regarding the construction of a Lyapunov functional, we prove in Chapter 6 pointwise estimates on the short (and long) time behaviour of hypoelliptic Markov semigroups in Hörmander’s sum of squares form. The work contained in this chapter has been done in collaboration with B. Zegarlinski. Finally, Appendix A.1 contains a statement of Stroock-Varadhan support theorem, Appendix A.2 is devoted to the Martingale central Limit theorem, which we use in Chapter 3, and Appendix A.3 gathers some useful definitions and results which we often refer to.

Throughout the the following chapters, we shall mainly consider \( q(t) \in \mathbb{R} \), (both for the GLE and for its approximations). This is in no way restrictive and comments on how to pass to the case \( q(t) \in \mathbb{R}^n \) will be made when needed. Also, the notation used in each chapter (or section) is local to that chapter (section), unless otherwise stated.
Chapter 1

Introduction: open classical systems

1.1 The Generalized Langevin Equation as a model in nonequilibrium statistical mechanics

Equation (0.0.1) is in general a non-Markovian stochastic integro-differential equation. To our knowledge, Langevin-type equations were first derived in their Markovian form \[ \ddot{q}(t) = -\partial_q V(q) - \gamma \dot{q}(t) + f(t), \] (1.1.1)

where $\gamma$ here is a constant and, by the fluctuation dissipation theorem, $f(t)$ is white noise. Equation (1.1.1) is just Newton’s equation of motion plus dumping and noise and it was introduced as a stochastic model for chemical reactions, in which a particle, held by intermolecular forces, undergoes the reaction when activated by random molecular collisions. In this framework, the term $-\gamma \dot{q}$ expresses the rate at which the reaction slows down due to such random interactions. This model assumes the collisions to occur instantaneously, hence it is not valid in physical situations in which such approximation cannot be made. In these cases a better description is given by the GLE:

\[ \ddot{q} = -\partial_q V(q) - \int_0^t \gamma(t-s) \dot{q}(s) \, ds + F(t). \] (1.1.2)

The GLE together with the fluctuation–dissipation theorem (0.0.2) appears in various applications such as surface diffusion [1] and polymer dynamics [73]. It also serves as one of the standard models of nonequilibrium statistical mechanics, describing the dynamics of a "small" Hamiltonian system (the distinguished particle) coupled to one or more heat baths which are
modelled as linear wave equations [68]:
\[
\partial_t^2 \varphi(t, x) = \partial_x^2 \varphi(t, x), \quad \varphi : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}.
\] (1.1.3)

(We consider $\varphi \in \mathbb{R}$ for notational simplicity, but everything holds for $\varphi \in \mathbb{R}^n$, like we said for (1.1.2)). Before sketching the derivation of the GLE, let us recall that the simplest case of the Bochner-Minlos theorem reads as follows: there is a one to one correspondence between probability measures on $\mathbb{R}$ and the class of continuous and positive definite functions on $\mathbb{R}$ such that $C(0) = 1$. Such a correspondence is given by
\[
C(t) = \int e^{ist} \mu(ds).
\]

Notice that this theorem applies to the covariance function of a stationary Gaussian process. Also, it still holds when $C$ is a function on $\mathcal{S}$, the space of Schwartz functions, and $\mu$ is a measure defined on the dual space $\mathcal{S}'$. Back to the Hamiltonian formulation of the wave equation, introducing the auxiliary variable $\pi(t, x) = \partial_t \varphi(t, x)$, we can rewrite (1.1.3) as follows:
\[
\begin{align*}
\partial_t \varphi(t, x) &= \pi(t, x) \\
\partial_t \pi(t, x) &= \partial_x^2 \varphi(t, x).
\end{align*}
\]

Denote $\phi = (\varphi, \pi)$ and let $\mathcal{H}_B$ be the completion of $H^1 \otimes L^2$ with norm induced by the inner product
\[
(\varphi_1, \pi_1) : (\varphi_2, \pi_2) = (\phi_1, \phi_2) := \int (\partial_x \varphi_1(x) \partial_x \varphi_2(x) + \pi_1(x) \pi_2(x)) dx.
\]

In this class of models the coupling between the distinguished particle and the heat bath is taken to be linear and is governed by a coupling function $\eta(x)$. The full Hamiltonian of the "particle+heat bath" is then
\[
H(q, p, \varphi, \pi) = H_P(q, p) + H_B(\varphi, \pi) + q \int \eta(x) \partial_x \varphi(x) dx,
\] (1.1.4)

where $H_P(q, p)$ denotes the Hamiltonian of the distinguished particle whose position and momentum are denoted by $q$ and $p$, respectively; $H_B(\varphi, \pi)$ is the Hamiltonian of the heat bath (wave equation) where $\varphi$ and $\pi$ are the canonically conjugate field variables:
\[
H_P(q, p) = \frac{p^2}{2} + V(q)
\]
and
\[
H_B(\varphi, \pi) = \frac{1}{2} \int_{\mathbb{R}} (\partial_x \varphi(x))^2 + (\pi(x))^2.
\]
The linear coupling in (1.1.4) is motivated by the dipole approximation from classical electrodynamics. Introducing $\alpha(x)$, defined in Fourier space by $\hat{\alpha}(k) = (-i\hat{\eta}(k)/k, 0)$ (so that $\partial_x \alpha = \eta$), the coupling term in (1.1.4) can be rewritten as $q\langle \phi, \alpha \rangle$. Let also $L$ be the operator

$$L = \begin{pmatrix} 0 & 1 \\ \partial_x^2 & 0 \end{pmatrix},$$

which in Fourier space reads

$$L = \begin{pmatrix} 0 & 1 \\ -k^2 & 0 \end{pmatrix}.$$  

Then

$$e^{tL} = \begin{pmatrix} \cos(kt) & \frac{\sin(kt)}{k} \\ -k \sin(kt) & \cos(kt) \end{pmatrix}.$$  

If we set $\gamma(t) := \langle e^{tL} \alpha, \alpha \rangle$, we have

$$\gamma(t) = \int |\hat{\eta}(k)|^2 e^{ikt} \, dk,$$  

(1.1.5)

where $\hat{\eta}(k)$ denotes the Fourier transform of $\eta(x)$ [34, 68]. Hence $\gamma(t)$ is the covariance function of a Gaussian process and $|\hat{\eta}(k)|^2$ is called the spectral density of the Gaussian process that has $\gamma(t)$ as covariance function. We now write down the equations of motion and integrate out the heat bath variables. The equations of motion of the system with Hamiltonian (1.1.4) are

$$\dot{q}_t = p_t$$  

(1.1.6)

$$\dot{p}_t = -\partial_q V(q) - \langle \phi, \alpha \rangle$$  

(1.1.7)

$$\dot{\phi}(k) = L(\dot{\phi}_t(k) + q_t \hat{\alpha}(k));$$  

(1.1.8)

integrating (1.1.8) and substituting back into (1.1.7) we get

$$\dot{q}_t = p_t$$  

(1.1.9)

$$\dot{p}_t = -\partial_q V(q) - \int_0^t \dot{\gamma}(t-s) q_s \, ds - \langle \phi_0, e^{-tL} \alpha \rangle.$$  

(1.1.10)
If $\phi_0$ is distributed according to a Gibbs measure \( \mu_\beta \) (with inverse temperature $\beta = 1$), then
\[
F(t) := \langle \phi_0, e^{-tL_\alpha} \rangle
\]
is still a Gaussian process with covariance function
\[
E(F(t)F(s)) = \int \langle \phi_0, e^{-tL_\alpha} \rangle \langle \phi_0, e^{-sL_\alpha} \rangle = \gamma(t-s),
\]
where the last equality follows by Gaussianity. Looking back at (1.1.9), (1.1.10) and (1.1.11), we have recovered the GLE together with the fluctuation dissipation principle. Also, we would like to explicitly notice that the fluctuation dissipation principle is only a consequence of assuming that the bath was initially in Gibbs equilibrium.

The GLE has also attracted attention in recent years in the context of mode reduction and coarse-graining for high dimensional dynamical systems [19]. One of the models that has been studied extensively within the framework of mode elimination is the Kac-Zwanzig model [80] and its variants [40, 39]. In these models the system “particle + bath” is described as a mechanical system in which a distinguished particle interacts with $n$ heat bath molecules of mass $\{m_j\}_{1 \leq j \leq n}$, through linear springs with random stiffness parameter $\{k_j\}_{1 \leq j \leq n}$; the Hamiltonian of the system is then:
\[
H(q_n, p_n, Q_1, ..., Q_n, P_1, ..., P_n) = V(q_n) + \frac{p_n^2}{2} + \frac{1}{2} \sum_{i=1}^n \frac{P_i^2}{m_i} + \frac{1}{2} \sum_{i=1}^n k_i (Q_i - q_n)^2,
\]
where $(q_n, p_n)$ and $(Q_1, ..., Q_n, P_1, ..., P_n)$ are the positions and momenta of the tagged particle and of the heat bath molecules, respectively (the notation $(q_n, p_n)$ is to stress that the position and momentum of the particle depend on the number of molecules it is coupled to). We can then write down the equations of motion of the system with Hamiltonian (1.1.12):
\[
\begin{align*}
\dot{q}_n &= p_n \\
\dot{p}_n &= -\partial_q V(q_n) + \sum_{i=1}^n k_i (Q_i - q_n) \\
\dot{Q}_i &= P_i/m_i \\
\dot{P}_i &= -k_i (Q_i - q_n) \quad 1 \leq i \leq n.
\end{align*}
\]
The initial conditions for the distinguished particle are assumed to be deterministic, namely $q_n(0) = q_0$ and $p_n(0) = p_0$; those for the heat bath are randomly drawn from a Gibbs distribution at inverse temperature $\beta$ and conditioned on $(q_0, p_0)$. Integrating out the heat
bath variables we obtain a closed equation for \( q_n \), of the form (1.1.2). In the thermodynamic limit as \( n \to \infty \) we recover the GLE. Under the assumption that at time \( t = 0 \) the heat bath is in equilibrium at inverse temperature \( \beta \), we obtain the fluctuation dissipation relation (0.0.2) as well. The form of the memory kernel \( \gamma(t) \) depends on the choice of the distribution of the spring constants of the harmonic oscillators in the heat bath [19].

The GLE (1.1.2) is a stochastic integrodifferential equation which is equivalent to the original infinite dimensional Hamiltonian system with random initial conditions. The non-Markovianity of the stochastic dynamics (1.1.2) renders the analysis of this dynamical system very difficult. This problem was studied in detail by Jaksic and Pillet in a series of papers [34, 36, 35]. In these works, existence and uniqueness of solutions as well as the ergodic properties of (1.1.2) were studied in detail. In particular, it was shown that the process \( \{q, p = \dot{q}\} \) is mixing with respect to the measure

\[
\nu_\beta(dqdp) = \frac{1}{Z_\beta} e^{-\beta H_P(q,p)} dqdp,
\]

where \( Z_\beta \) is the normalization constant. To our knowledge, no information concerning the rate of convergence to equilibrium for the non-Markovian dynamics (1.1.2) is known for general classes of memory kernels. Ergodic theory for a quite general class of non-Markovian processes has been developed recently, see [20] and the references therein.

One of the problems encountered when dealing with a non Markovian equation is that there is no general formalism to derive a (generalized) Fokker-Plank equation in closed form, that is, an equation for the time evolution of the probability density. Attempts in this direction have shown that even in simple cases the calculation is very convoluted, it does not always lead to a closed analytical expression and, when it does, the resulting equation is extremely involved (see for example, [2]). These considerations, together with the ones made so far, motivate the attempt to approximate the full dynamics (1.1.2) by a Markovian one.

### 1.2 Markovian approximations of the GLE

As noticed in [39], the problem can be recast as follows: we want to approximate the non Markovian process (1.1.2) with the Markovian dynamics given by the system of ODEs:
where we recall that \((q, p) \in \mathbb{R}^2\), \(u\) and \(g\) are column vectors of \(\mathbb{R}^d\), \(\cdot\) denotes Euclidean scalar product, \(W(t) = (W_1(t), \ldots, W_d(t))\) is a \(d\)-dimensional Brownian motion, \(V(q)\) is a potential and \(A\) and \(C\) are constant coefficients \(d \times d\) matrices, related through the fluctuation dissipation principle:

\[
A + A^T = CC^T. \tag{1.2.2}
\]

We also recall that the noise in (1.1.2) is Gaussian, stationary and mean zero. Because the memory kernel and the noise in (1.1.2) are related through the fluctuation-dissipation relation, the rough idea is that we might try to either approximate the noise and hence obtain the corresponding memory kernel or, the other way around, we could approximate the the correlation function and read off the noise. The latter is the approach that we shall follow in this section. As a motivation, we would like to notice that for some specific choices of the kernel, equation (1.1.2) is equivalent to a finite dimensional Markovian system in an extended state space (see [13]). If, for example, we choose \(\gamma(t) = \lambda^2 e^{-|t|}\), then (1.1.2) becomes

\[
\begin{align*}
\dot{q} &= p \\
\dot{p} &= -\partial_q V(q) - \lambda^2 \int_0^t e^{-(t-s)p(s)} ds + F(t);
\end{align*}
\tag{1.2.3}
\]

the fluctuation dissipation theorem (with \(\beta = 1\)) yields

\[
E(F(t + s)F(t)) = \lambda^2 e^{-|s|}. \tag{1.2.4}
\]

Since we are requiring \(F(t)\) to be stationary and Gaussian, (1.2.4) implies that \(F(t)\) is the Ornstein-Uhlenbeck process.\(^2\) If we write \(F(t) = \lambda v(t)\), with \(v(t)\) satisfying the equation \(\dot{v} = -v + \sqrt{2\beta^{-1}} \dot{W}\), and we define the new process

\[
z(t) = -\lambda \int_0^t e^{-(t-s)p(s)} ds + v(t), \tag{1.2.5}
\]

then (1.2.3) becomes

\[
\begin{align*}
\dot{q} &= p \\
\dot{p} &= -\partial_q V + \lambda z \\
\dot{z} &= -\lambda p - z + \sqrt{2} \dot{W},
\end{align*}
\]

\(^2\)The only stationary Gaussian process with autocorrelation function \(e^{-t}\) is the O-U process.
which is precisely system (0.0.5) with $d = 1$ and $\alpha_1 = 1$. The "Markovianization" of (1.1.2)
was first done by Mori ([55]) by first approximating the Laplace transform of the memory
kernel $\gamma(t)$, $\tilde{\gamma}(\xi)$, by a rational function (if and when this is possible) and then imposing
the fluctuation relation, which gives the matrices $A$ and $C$ as well as the vector $g$. If $\gamma(t)$
itself is a sum of exponentials, $\gamma_d(t) = \sum_{i=1}^d \lambda_i^2 e^{-\alpha_i t}$, then $\tilde{\gamma}_d = \sum_{i=1}^d \lambda_i^2 / (\xi_i + \alpha_i)$, so
the procedure indicated by Mori is clearly successful and it corresponds to the case in which
$A = \text{diag}\{\alpha_1, \ldots, \alpha_d\}$ and $g = (\lambda_1, \ldots, \lambda_d)^T$. Another typical situation is when the Laplace
transform of $\gamma$ has a continued fraction representation
$$
\tilde{\gamma}(\xi) = \frac{\epsilon_2}{\xi + \theta_1 + \frac{\epsilon_3}{\xi + \theta_2 + \frac{\epsilon_4}{\xi + \theta_3 + \cdots}}}, \quad \theta_i > 0.
$$
In this case the approximation is done by truncating the fraction at step $d$ and then reading
off the corresponding Markovian system of $(d + 2)$ SDEs. The matrix $A$ is then tridiagonal,
$$
A = \begin{pmatrix}
\theta_1 & -\epsilon_2 \\
\epsilon_2 & \theta_2 & -\epsilon_3 \\
& \epsilon_3 & \theta_3 & \ddots \\
& & & \ddots & \ddots \\
& & & & \theta_d
\end{pmatrix}
$$
and $g = (\epsilon_1, 0, \ldots, 0)^T$. A class of memory kernels for which more detailed information on the
long time asymptotics of the GLE (3.0.1) can be obtained was considered by Eckmann, Hairer,
Pillet and Rey-Bellet in a series of papers [69, 16, 17, 14]. It was observed in these works that
when the memory kernel $\gamma(t)$ has a rational spectral density, then the GLE is equivalent to
a finite dimensional Markovian system. This system is obtained by adding a finite number of
additional degrees of freedom which account for the memory in the system (along the lines
of what we did when we introduced the auxiliary variable $z(t)$ in (1.2.5)). These auxiliary
variables satisfy linear stochastic differential equations. As an example we mention the case
where $\tilde{\eta}(k)$ in (1.1.5) can be written as
$$
|\tilde{\eta}(k)|^2 = \frac{1}{|p(k)|^2},
$$
where $p(k) = \sum_{m=1}^d c_m(-ik)^m$ is a polynomial with real coefficients and roots in the upper
half plane. Indeed, following [68, Proposition 2.3], we can prove the following
Proposition 1.2.1. If \( p(k) = \sum_{m=1}^{d} c_m(k)^m \) is a polynomial with real coefficients and roots in the upper half plane then the Gaussian process with spectral density \( |\tilde{\eta}(k)|^2 \) as in (1.2.6) is the solution of the SDE

\[
p\left( -i \frac{d}{dt} \right) x(t) = \frac{dW}{dt}, \tag{1.2.7}
\]
where \( W \) is a standard one dimensional Brownian motion.

Proof. We first prove a representation formula for the Gaussian process with spectral density \( |p(k)|^{-2} \), formula (1.2.8) below, and then we show that the process \( x(t) \) defined by (1.2.8) solves equation (1.2.7).

Define

\[
f(t) := \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{ikt} \frac{1}{p(k)} dk \quad \text{for } t > 0
\]
and \( f(t) = 0 \) for \( t \leq 0 \) (as the zeros of \( p \) are in the upper half plane). Then \( x(t) \) can be represented as

\[
x(t) = \int_{-\infty}^{+\infty} f(t-t')dW_{t'} = \int_{-\infty}^{t} f(t-t')dW_{t'} \tag{1.2.8}
\]

To prove the above representation formula we need to show that

\[
E(x(t)x(s)) = \int_{\mathbb{R}} e^{i(t-s)k} \frac{dk}{|p(k)|^2}.
\]

Indeed from (1.2.8) we can (formally) derive:

\[
E(x(t)x(s)) = \int_{\mathbb{R}} \int f(t-t')f(s-s')\delta(t'-s')dt'ds'
\]
\[
= \int_{\mathbb{R}} \int f(t-s')f(s-s')ds'
\]
\[
= \int_{\mathbb{R}} \int \int e^{ik(t-s')}e^{ik'(s-s')} \frac{1}{p(k)p(k')} ds'dkd'
\]
\[
= \int \int e^{ikt+k's} \delta(k+k') \frac{1}{p(k)p(k')} dkdk'
\]
\[
= \int e^{ikt} \frac{1}{p(k)p(-k)} dk,
\]

where in the first equality we used the delta correlation of white noise and in the third equality we expanded the expression of \( f(\cdot) \). In the above the domain of integration is always \( \mathbb{R} \). Now we need to check that (1.2.7) holds:

\[
p\left( -i \frac{d}{dt} \right) x(t) = \int_{-\infty}^{t} \int_{\mathbb{R}} p\left( -i \frac{d}{dt} \right) e^{ik(t-t')} \frac{1}{p(k)} dk dW_{t'}.
\]

Because, from straightforward calculations,

\[
\sum_{m=1}^{d} c_m \left( -i \frac{d}{dt} \right)^m e^{ik(t-t')} = p(k) e^{ik(t-t')},
\]
we obtain

\[ p \left( -i \frac{d}{dt} \right) x(t) = \int_{-\infty}^{t} \int_{\mathbb{R}} e^{ik(t-t')} dk dW_{t'} = \frac{dW_t}{dt}. \]

Notice that in particular, when \( p(k) = (ik + \alpha) \), we have

\[ \gamma(t) = \frac{\alpha}{\pi} \int \frac{e^{ikt}}{k^2 + \alpha^2} dk = e^{-\alpha t} \quad \text{(for } t > 0), \]

i.e. we are in the case that we will study in Chapter 3.

Finally, in view of Section 4.2, we find interesting the following remark, made in [39]. When the kernel of the GLE (1.1.2) is chosen to decay as a power law, \( \gamma(t-s) = (t-s)^{-\delta}, \delta > 0 \), and the potential is either quadratic or vanishing, the GLE can be explicitly solved. In these cases one can prove that the limiting behaviour of the GLE is subdiffusive \(^3\); also, the free particle (i.e. when \( V(q) = 0 \)) diffuses away, whereas in case of quadratic potential the equilibrium Gibbs distribution is approached at a subexponential (subdiffusive) rate.

---

\(^3\) A stochastic process is said to exhibit anomalous diffusion when its mean square displacement does not grow linearly in time, but as \( t^\zeta, \zeta \neq 1 \) and \( \zeta > 0 \). In particular, for \( 0 < \zeta < 1 \) we talk about subdiffusion, for \( 1 < \zeta < 2 \) we talk about superdiffusion, see [11, 59].
Chapter 2

Exponential Decay to equilibrium: Ergodicity, Hypocoercivity and Singular Space Theory

The purpose of nonequilibrium statistical mechanics is to explain irreversibility on the basis of microscopic dynamics, and to give quantitative predictions for dissipative phenomena. (Ruelle, [71])

In this chapter we present three different settings to study exponential convergence to equilibrium, namely the theory of geometric ergodicity for Markov processes, the theory of hypocoercivity and the singular space theory. The latter will also provide a method to determine the spectrum of quadratic hypoelliptic operators, hence the exact rate of convergence to equilibrium. The chapter is organized as follows: since we will be concerned with hypoelliptic systems, in Section 2.1 we recall some facts about hypoellipticity that we shall use in the following. Next, we start presenting the three main methods that we will employ in Chapter 3, 4 and 5 in order to study exponential convergence to equilibrium of system (0.0.5) and (0.0.3): Section 2.2 is devoted to the ergodic theory for hypoelliptic Markov processes; Section 2.3 provides an overview of hypocoercivity theory; Section 2.4 is split in two subsections, the first of which, Subsection 2.4.1, gives some basic definitions in semiclassical analysis and the second, Subsection 2.4.2, presents the main results obtained in [62] by the third author. At the end of the chapter, Section 2.5 contains comments and comparisons between the two functional analytic methods.
2.1 Hypoellipticity

Let $O$ be an open set of $\mathbb{R}^n$. A distribution $u$ on $O$ is a linear form on $C^\infty_0(O)$ such that for every compact set $K \subset O$ there exist constants $\beta$ and $\kappa$ such that

$$|u(\phi)| \leq \beta \sup |\partial^\alpha \phi|, \quad \forall \phi \in C^\infty_0(K).$$

The set of distributions on $O$ is denoted $\mathcal{D}'(O)$. Notice that a continuous function $g(x)$ can always be regarded as a distribution by considering $\int g(x)\phi(x), \phi \in C^\infty_0$; hence we have the chain of inclusions $C^\infty(O) \subset C(O) \subset \mathcal{D}'(O)$. The singular support of $u \in \mathcal{D}'(O)$ is the set of points in $O$ denoted by $\text{singsupp } u$ such that $u$ is a $C^\infty$ function on $O \setminus \text{singsupp } u$ and this is not true for any open set larger than $O \setminus \text{singsupp } u$. A linear differential operator $T$ on $\mathbb{R}^n$ with $C^\infty$ coefficients is hypoelliptic if $\text{sing supp } u = \text{sing supp } Tu$.

Consider now an SDE of the form

$$dx_t = b(x_t)dt + \sigma(x_t)dB_t, \quad x_t \in \mathbb{R}^n,$$

where $b$ is a vector-valued function, $\sigma$ is matrix valued and $B_t$ is an $n$-dimensional standard Brownian motion. Also, suppose that $b$ and $\sigma$ are smooth. The generator $L$ of (2.1.1) is in the "sum of squares" form $L = \sum_{j=1}^d X_j^2 + X_0$. Each $X_j$ is a vector field (first order differential operator) of the form $X_j = \sum_{k=0}^n a_{jk} \partial_k$, with $a_{jk}(x)$ a smooth function. Hence to $X_j$ we can associate the vector field $(a_{j1}(x), \ldots, a_{jn}(x))$ (which, for every fixed $x$, can be thought of as a tangent vector to $\mathbb{R}^n$ at the point $x$.) Given a collection of vector fields, $Y_1, \ldots, Y_d$, the Lie algebra generated by such fields is the smallest vector space that contains $Y_1, \ldots, Y_d$ and is closed under the commutator operation. We say that the Lie algebra generated by $Y_1, \ldots, Y_d$ is full at $x \in \mathbb{R}^n$ if the elements of the Lie algebra span $\mathbb{R}^n$ at $x$. (see [77]).

**Theorem 2.1.1** (Hörmander’s theorem). If the Lie algebra generated by $\{X_0, X_1, \ldots, X_d\}$, $\text{Lie}\{X_i\}_{0 \leq i \leq d}$, is full $\forall x \in \mathbb{R}^n$ then $L$ is hypoelliptic.

In the remainder of the section we follow [77]. One of the reasons why the concept of hypoellipticity is of great importance is because it allows to prove the existence of a smooth density (for the process (2.1.1)). To show this fact, let $f(t, x) \in C^\infty_0(\mathbb{R}^n \times (0, \infty))$. By Itô’s
2.2. Ergodic Theory for hypoelliptic diffusions

formula
\[ f(t, x_t) - f(0, x_0) = \int_0^t \left( \frac{\partial}{\partial s} + L \right) f(s, x_s) + \int_0^t \sigma(x_s) \frac{\partial f}{\partial x} dB_s. \]

So if \( t \) is big enough \( f(t, x_t) = f(0, x_0) = 0 \) and taking expectation on both sides of the above equality we get
\[ 0 = E \int_0^t \left( \frac{\partial}{\partial s} + L \right) f(s, x_s) = \int_0^t ds \int_{\mathbb{R}^n} \left[ \left( \frac{\partial}{\partial s} + L \right) f(s, x_s) \right] p(x_s \in dy). \]

Integrating by parts gives
\[ \int_0^t ds \int_{\mathbb{R}^n} \left[ \left( -\frac{\partial}{\partial s} + L' \right) p(x_s \in dy) \right] f(s, x_s) = 0 \quad \forall f \in C^\infty_0((0, \infty) \times \mathbb{R}^n), \quad (2.1.2) \]

where \( L' \) is the formal adjoint of \( L \). In other words \( \left(-\frac{\partial}{\partial s} + L' \right) p(x_s \in dy) = 0 \) (in distributional sense). Therefore, if \( \partial_t + L' \) is hypoelliptic, i.e. \( \text{Lie}\{\partial_t + X_0, X_1, \ldots, X_d\} \) is full on \( \mathbb{R}^n \times (0, \infty) \forall x \in \mathbb{R}^n \), then \( p(x_t \in dy) \) is a \( C^\infty \) function.

Lemma 2.1.2. \( \text{Lie}\{\partial_t + X_0, X_1, \ldots, X_d\} \) is full on \( \mathbb{R}^n \times (0, \infty) \forall x \in \mathbb{R}^n \) if and only if
\[ \text{Lie}\{X_1, \ldots, X_d, [X_1, X_0], \ldots, [X_d, X_0]\} \text{ is full on } \mathbb{R}^n, \forall x \in \mathbb{R}^n. \quad (2.1.3) \]

Theorem 2.1.3 (Smoothness of the density). Assume that \( x_t \) is the only solution to (2.1.1), whose coefficients do not depend on time, are infinitely differentiable with bounded partial derivatives of all orders and that condition (2.1.3) holds. Then the random vector \( x_t \) has an infinitely differentiable density \( \forall t > 0 \).

The above Theorem 2.1.3 can be found in [58].

2.2 Ergodic Theory for hypoelliptic diffusions

As we have already mentioned in the Overview, a process is said to be ergodic when it admits a unique invariant measure. It is said to be geometrically ergodic when it converges exponentially fast to the equilibrium measure, in a sense that will be made more precise by (2.2.3) (or (2.2.10)).

In this section we shall present some general results on the ergodic theory of Markov processes and we shall show how these results apply to the toy hypoelliptic system
\[ \dot{q} = p, \quad (2.2.1a) \]
\[ \dot{p} = -\nabla_q V(q) + r, \quad (2.2.1b) \]
\[ \dot{r} = -p - r + \dot{W}, \quad (2.2.1c) \]
2.2. Ergodic Theory for hypoelliptic diffusions

where \( W \) is a \( d \)-dimensional standard Brownian motion. We consider both the case \( (q, p, r) \in \mathbb{T}^d \times \mathbb{R}^d \times \mathbb{R}^d := X \) and \( (q, p, r) \in \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^d := Y \). The extension to the case \( r \in \mathbb{R}^{md} \) is straightforward (and we shall not present it) hence this section constitutes a proof of Theorem 3.1.2 as well, i.e. the techniques that we present in the following easily apply to prove ergodicity of the Markov process \( x(t) := \{q(t), p(t), z(t)\} \) given by (3.0.2). The main references for this section are [52, 54, 69, 51]. Let \( L \) be the generator of the process (2.2.1), namely

\[
L = p \cdot \nabla_q - \nabla_q V \cdot \nabla_p + r \cdot \nabla_p - (p + r) \cdot \nabla_r + \Delta_r,
\]

with \( \Delta_r = \sum_{i=1}^d \partial_{r_i}^2 \). and \( \| \cdot \| \) denote the Euclidean inner product and norm, respectively, and we shall use the notation \( x(t) = (q(t), p(t), r(t)) \).

Following [52] and [51], let \( P_t(x, A) \) be the transition kernel of the Markov process \( x(t) \). Consider the discretized process \( \{x_n\}_{n \in \mathbb{N}} \), obtained by sampling at the rate \( T > 0 \) and with transition kernel \( P(x, A) := P_T(x, A) \).

Lyapunov Condition: There exists a function \( G(x) : \mathbb{R}^{3d} \to [1, \infty) \) such that \( G(x) \to \infty \) as \( \| x \| \to \infty \) and

\[
LG(x) \leq -aG(x) + \tilde{d},
\]

(2.2.2)

for some \( a, \tilde{d} > 0 \).

Minorization condition: There exist \( T > 0, \eta > 0 \) and a probability measure \( \nu \), with \( \nu(C^c) = 0 \) and \( \nu(C) = 1 \) for some fixed compact set \( C \) in the phase space, such that

\[
P_T(x, A) \geq \eta \nu(A) \quad \forall A \in \mathcal{B}(\mathbb{R}^{3d}), x \in C.
\]

Consider now the set \( \mathcal{G} = \{ x \in \mathbb{R}^{3d} : G(x) \leq \frac{2\tilde{d}/a}{\gamma - e^{-\alpha T/2}} \} \) for some \( \gamma \in (e^{-\alpha T/2}, 1) \), \( G \), \( \alpha \) and \( \tilde{d} \) as in the Lyapunov condition. Notice that if \( \alpha \) and \( \tilde{d} \) are such that (2.2.2) is satisfied, then also \( \alpha \) and \( \tilde{d} \) will satisfy (2.2.2); therefore \( \tilde{d} \) can be chosen such that the quantity \( \frac{2\tilde{d}/a}{\gamma - e^{-\alpha T/2}} \) is bigger than 1. We will use the following result, the proof of which can be found in [54], in particular see [51, Theorem 3.2] or [54, Theorem 15.0.1].

**Theorem 2.2.1.** If there exists a function \( G \) satisfying the Lyapunov condition and there exists a sampling rate \( T > 0 \) such that the resulting chain \( \{x_n\}_{n \in \mathbb{N}} \) is an aperiodic Markov chain satisfying the minorization condition on the set \( \mathcal{G} \), then the process is geometrically ergodic, i.e. there exists a unique invariant measure \( \rho \) and

\[
|E^{x_0} g(x(t)) - \rho(g)| \leq k G(x_0) e^{-\lambda t}, \quad t \geq 0, \quad x_0 = x(0),
\]

(2.2.3)
for some $k, \lambda > 0$ and for any $g \in \mathcal{P} := \{ g : \mathbb{R}^3 \to \mathbb{R}, \text{measurable} : |g(x)| \leq G(x) \}$. 

**Assumption (⋆):** Let $B_s(y) \in \mathbb{R}^3$ be the ball of radius $s$ centered in $y$. For some fixed compact set $C$ we have

- $P_t(x, A)$ has a density $p_t(x, y)$ which is continuous $\forall (x, y) \in C \times C$, more precisely
  $$P_t(x, A) = \int_A p_t(x, y) \, dy \quad \forall A \in \mathcal{B}(\mathbb{R}^3) \cap \mathcal{B}(C), \forall x \in C;$$
- for some $x^* \in \text{int}(C)$ and $\forall \delta > 0$ we can find a $\bar{t} = \bar{t}(\delta)$ such that
  $$P_{\bar{t}}(x, B_\delta(x^*)) > 0, \quad \forall x \in C.$$

We have the following result.

**Lemma 2.2.2.** Assumption (⋆) $\implies$ Minorization Condition.

Some comments on the Lyapunov condition and on Assumption (⋆). For our toy process (2.2.1) we shall prove something stronger than Assumption (⋆). Namely, we shall show that the density of the transition probability function is not only continuous but also smooth (when the potential is smooth) and that the process is irreducible and aperiodic, i.e. $\forall x \in \mathbb{R}^3$ and for any open set $U$ of $\mathbb{R}^3$,

$$P_t(x, U) > 0, \quad \text{for some } t > 0. \quad (2.2.4)$$

The hypoellipticity of the process will serve both purposes. Indeed, the hypoellipticity of $L$ (and hence the hypoellipticity of $\partial_t + L$) implies both the smoothness of $p_t(x, y)$ (see Section 2.1) and the controllability of its dynamics (for a definition of controllability in the classic control theory setting see Appendix A.1). More precisely, (2.2.1) is an Ornstein-Uhlenbeck process, and for O-U processes hypoellipticity and controllability are equivalent (see [53]). We shall go back to this point in Chapter 5, see (i)-(iii) Section 5.1. For the time being, let us observe that once the controllability of (2.2.1) is ensured, Stroock-Varadhan Support Theorem (which can be found in Appendix A.1) gives irreducibility and aperiodicity of the process. However, Assumption (⋆) does not guarantee the existence of the invariant measure, it only ensures uniqueness. Existence is catered for by the Lyapunov condition. In its simplest form, such a condition reads as follows: there exists a positive function $G(x)$ with compact level sets such that

$$LG(x) \leq -c + b \, 1_K(x), \quad (2.2.5)$$
for some \( b, c > 0 \) and for some compact set \( K \). If such a function exists then the expected value of the hitting time of \( K \), \( E^x(\tau_K) \), is finite and we can explicitly construct an invariant measure. Moreover (2.2.5) and Assumption (*) imply convergence to the invariant measure. When the Lyapunov condition is satisfied in the stronger form (2.2.2) then the convergence is exponentially fast in the sense (2.2.3). We want to stress again that for an hypoelliptic O-U processes, the existence of a Lyapunov function (2.2.2) is enough to obtain geometric ergodicity. We shall extensively use this fact in Chapter 4. We now prove that Theorem 2.2.1 applies to system (2.2.1).

**Theorem 2.2.3** (Ergodicity). *The solution of (2.2.1) with \( x(t) \in X \) and \( V(q) \in C^\infty(\mathbb{T}^d) \) is geometrically ergodic. The same holds true when \( x(t) \in Y \), provided that the potential \( V(q) \in C^\infty(\mathbb{R}^d) \) is confining, has bounded Hessian, denoted \( \nabla^2 V \), and (2.2.8) holds.*

**Proof of Theorem 2.2.3.** Consider first the case \( x(t) \in X \). Let \( V(q) \) be a \( C^\infty(\mathbb{T}^d) \) potential, \( V(q) > -k \) for some positive constant \( k \). Consider the function

\[
G(x) = \dot{C} + \frac{B}{2} \|p\|^2 + \frac{C}{2} \|r\|^2 + DV(q) + H(p, r), \tag{2.2.6}
\]

where \( B, C, D, H \) and \( \dot{C} \) are positive constants to be chosen. We have that

\[
G(x) \geq \dot{C} + \frac{B}{2} \|p\|^2 + \frac{C}{2} \|r\|^2 - \frac{H}{2} \|p\|^2 - \frac{H}{2} \|r\|^2 - Dk, \tag{2.2.7}
\]

so we need \( B > H, C > H \) and \( \dot{C} > Dk \). Moreover

\[
LG(x) = D(\nabla_q V, p) - B(\nabla_q V, p) - H(\nabla_q V, r) + B(r, p) + H\|r\|^2 - C(p, r) - H\|p\|^2 - C\|r\|^2 - H(p, r) + C
\]

\[
\leq H\|r\|^2 + \frac{H}{4}\|\nabla_q V\|^2 - H\|p\|^2 - C\|r\|^2 + C + H\|r\|^2,
\]

where we have chosen \( B = D = C + H \). On the other hand, since \( V(q) \leq \hat{k} \),

\[
-aG(x) \geq -\frac{a}{2} B\|p\|^2 - \frac{a}{2} C\|r\|^2 - a\hat{k}B - a\frac{H}{2} \|r\|^2 - a\frac{H}{2} \|p\|^2
\]

so imposing also \( 2H - C \leq -\frac{a}{2} (C + H), -H \leq -\frac{a}{2} (B + H) \) for some \( a > 0 \), the Lyapunov condition is satisfied. One possible choice is \( a = 1/4, B = 13/16, C = 5/8 \) and \( H = 3/16 \). Consider now the case \( x(t) \in Y \). We introduce the Lyapunov function

\[
G(x) = \dot{C} + \frac{A}{2} \|q\|^2 + \frac{B}{2} \|p\|^2 + \frac{C}{2} \|r\|^2 + DV(q) + E(p, q) + F(q, r) + H(p, r) + M(\nabla_q V, p),
\]
for some constants $A, B, C, \hat{C}, D, E, F, H, M$ to be chosen. Therefore,

$$\nabla_q G = Aq + D\nabla_q V + Ep + Fr + M\nabla^2 V(q) \cdot p,$$

$$\nabla_p G = Bp + Eq + Hr + M\nabla_q V,$$

$$\nabla_r G = Cr + Fq + Hp.$$

Thus,

$$LG(x) = A(p, q) + D(\nabla_q V, p) + E\|p\|^2 + F(p, r) - B(\nabla_q V, p)$$

$$- E(\nabla_q V, q) - H(\nabla_q V, r) + B(p, r) + E(q, r) + H\|r\|^2$$

$$- C(p, r) - F(p, q) - H\|p\|^2 + M(p, \nabla^2 V(q) \cdot p)$$

$$- C\|r\|^2 - F(r, q) - H(p, r) + C - M\|\nabla_q V\|^2 + M(r, \nabla_q V).$$

From the boundedness of the Hessian of $V(x)$ it follows that there exist constants $\tilde{\beta}$ and $\tilde{\sigma}$ such that

$$\tilde{\sigma}\|q\|^2 - \tilde{\beta}\|\nabla_q V\|^2 \to -\infty \text{ as } \|q\|^2 \to +\infty. \quad (2.2.8)$$

Hence, with calculations analogous to those done in the periodic case, it follows that there exist constants $A, B, C, \hat{C}, D, E, F, H, M$ such that $G$ satisfies the Lyapunov condition.

As for Assumption $(*)$, first of all let us notice that, since the operator $\partial_t + L$ is hypoelliptic, the transition probability has a density (see Section 2.1, Theorem 2.1.3); because the SDE we consider has time independent coefficients the density is $C^\infty$ provided that $V(q) \in C^\infty$.

Moreover, studying the control problem associated with $dx = b(x)dt + \sigma dw$, namely $dZ = b(Z)dt + \sigma dU$ where $U(t)$ is a smooth control, and using the Stroock-Varadhan support Theorem, we can prove that $P_t(x, A) > 0 \forall x \in \mathbb{R}^{3d}, t > 0$ and for any open set $A \in \mathbb{R}^{3d}$.

We can also prove that if $G(x)$ is a Lyapunov function, then $G(x)^l$ is also a Lyapunov function, $\forall l \geq 1$. In other words, $\forall l \geq 1$ have

$$LG(x)^l \leq -a_l G(x)^l + \tilde{a}_l,$$  \quad (2.2.9)

for some suitable positive constants $a_l$ and $\tilde{a}_l$. Indeed,

$$\partial_q G(x)^l = lG(x)^l-1\partial_q G(x),$$

$$\partial_p G(x)^l = lG(x)^l-1\partial_p G(x),$$
and
\[
\partial_{\tau_i}^2 G(x) = \partial_{\tau_i} \left[lG(x)^{l-1} \partial_{\tau_i} G(x)\right] = l(l-1)G(x)^{l-2}(\partial_{\tau_i} G)^2 + lG(x)^{l-1} \partial_{\tau_i}^2 G(x).
\]
Furthermore, from (2.2.6) and (2.2.7), we have \((\partial_{\tau_i} G)^2 \leq c G\) for some \(c > 0\) so we obtain
\[
l(l-1)G(x)^{l-2}(\partial_{\tau_i} G)^2 \leq c_l G(x)^{l-1},
\]
and therefore
\[
L G(x)^l \leq lG(x)^{l-1} L G(x) + c_l G(x)^{l-1}.
\]
Using what we have proven in the case \(l = 1\), we get (2.2.9).

For any \(l \geq 1\), consider now the set
\[
P_l = \{g : \mathbb{R}^3d \rightarrow \mathbb{R}, \text{ measurable : } |g(x)| \leq G(x)^l\}.
\]
Then there exist constants \(k = k(l)\) and \(\lambda = \lambda(l)\), such that \(\forall g \in P_l\)
\[
|E_{x_0} g(x(t)) - \rho(g)| \leq k[G(x_0)]^l e^{-\lambda t}, \ t \geq 0, x_0 = x(0). \tag{2.2.10}
\]

## 2.3 Hypocoercivity

Hypocoercivity theory applies to evolution equations of the form
\[
\partial_t f + (A^* A + B) f = 0. \tag{2.3.1}
\]
Following [76], we first introduce the necessary notation. Let \(\mathcal{H}\) be a Hilbert space, real and separable, \(\mathcal{H}^m := \mathcal{H} \otimes \mathbb{R}^m\), \(\| \cdot \|\) and \((\cdot, \cdot)\) the norm and scalar product of \(\mathcal{H}\), respectively. Let \(A : \mathcal{H} \rightarrow \mathcal{H} \otimes \mathbb{R}^m\) and \(B : \mathcal{H} \rightarrow \mathcal{H}\) be unbounded operators with domains \(\mathcal{D}(A)\) and \(\mathcal{D}(B)\) respectively, and assume that \(B\) is antisymmetric, i.e. \(B^* = -B\), where \(*\) denotes adjoint in \(\mathcal{H}\). Notice that \(A\) can be thought of as an array of operators \(A_j : \mathcal{H} \rightarrow \mathcal{H}\), \(j = 1, \ldots, m\). We shall also assume that there exists a vector space \(S \subset \mathcal{H}\), dense in \(\mathcal{H}\), where all the operations that we will perform involving \(A\) and \(B\) are well defined (this will usually be the space of Schwartz functions). When we refer to the commutator between \(A\) and \(B\) what we actually mean is \([A, B] = AB - (B \otimes I)A\), where \(I\) is the identity on \(\mathbb{R}^m\), so that \([A, B] : \mathcal{H} \rightarrow \mathcal{H} \otimes \mathbb{R}^m\). Analogously \([A, A^*] : \mathcal{H} \rightarrow \mathcal{H} \otimes \mathbb{R}^m \otimes \mathbb{R}^m\) and \([A^*, A] : \mathcal{H} \rightarrow \mathcal{H}\). In Section 3.1.1 we give more details about this notation, applied to the case when \(\mathcal{H} = L^2(\rho)\), where \(\rho\) is the equilibrium measure. I say "the" equilibrium measure because in this context we typically assume ergodicity. We say that an unbounded linear operator \(S\) on \(\mathcal{H}\) is relatively
bounded with respect to the (linear unbounded) operators $T_1, \ldots, T_n$ if $D(S) \supset (\cap D(T_j))$ and
$\exists$ a constant $\alpha > 0$ s.t.
\[
\forall h \in D(S), \quad \|Sh\| \leq \alpha(\|T_1h\| + \ldots + \|T_nh\|).
\]
Finally, given the (unbounded) operators $C_0, \ldots, C_N$, we define the Sobolev norm associated to this family of operators as
\[
\|h\|_{H^1}^2 := \|h\|^2 + \sum_{j=0}^{N} \|C_jh\|^2, \quad \forall h \in (\cap D(C_j)). \tag{2.3.2}
\]
With this notation, consider the linear operator
\[
L = A^*A + B. \tag{2.3.3}
\]
In the cases of interest to us $A$ and $B$ will always be first order differential operators. The form $A^*A + B$ is more general than it might seem at first sight. Indeed notice that any linear operator with nonnegative symmetric part can be recast in this form (at least modulo technical issues). What we are looking at is a dissipative evolution equation, whose dynamic is generated by an operator $-L$ involving a conservative (antisymmetric) part and a (symmetric) degenerate diffusive one. Also, the operator $L$ can be rewritten in Hörmander sum of squares form. Indeed, the context we are working in, is to a certain extent similar to the one of hypoellipticity. We will further comment on this point in Section 2.5.

Suppose that the semigroup generated by $L$ is well defined (we shall deal with this problem for the case at hand in Proposition 3.2.1). Then the semigroup generated by $L$ is easily seen to be a contraction semigroup:
\[
\frac{1}{2} \left. \frac{d}{dt} \right|_{t=0^+} \|e^{-tL}h\|^2 = -\|Ah\|^2 \leq 0.
\]

**Definition 2.3.1** (Coercivity). Let $T$ be an unbounded operator on a Hilbert space $\mathcal{H}$, denote its kernel by $\mathcal{K}$ and assume there exists another Hilbert space $\tilde{\mathcal{H}}$ continuously and densely embedded in $\mathcal{K}$. If $\|\cdot\|_{\tilde{\mathcal{H}}}$ and $(\cdot, \cdot)_{\tilde{\mathcal{H}}}$ are the norm and scalar product on $\tilde{\mathcal{H}}$, respectively, then the operator $T$ is said to be $\lambda$-coercive on $\tilde{\mathcal{H}}$ if
\[
(Th, h)_{\tilde{\mathcal{H}}} \geq \lambda\|h\|_{\tilde{\mathcal{H}}}^2 \quad \forall h \in \mathcal{K} \cap D(T),
\]
where $D(T)$ is the domain of $T$ in $\tilde{\mathcal{H}}$. 
Without worrying about regularity issues, i.e. assuming that the operator $T$ generates a contraction semigroup, the following Proposition gives an equivalent definition of coercivity.

**Proposition 2.3.2.** With the same notation as in Definition 2.3.1, $T$ is $\lambda$-coercive on $\tilde{H}$ if and only if

$$
\| e^{-Tt}h \|_{\tilde{H}} \leq e^{-\lambda t} \| h \|_{\tilde{H}} \quad \forall h \in \tilde{H} \text{ and } t \geq 0.
$$

Let $K$ be the kernel of $L$ and notice that $Ker(A^*A) = Ker(A)$ and $K = Ker(A) \cap Ker(B)$. Suppose $KerA \subset KerB$; then $KerL = KerA$ then the coercivity of $L$ is equivalent to the coercivity of $A^*A$. So the case we are interested in is the case in which $A^*A$ is coercive and $L$ is not. In order for this to happen $A^*A$ and $B$ cannot commute; if they did, then $e^{-tL} = e^{-tA^*A}e^{-tB}$. Therefore, since $e^{-tB}$ is norm preserving, we would have $\| e^{-tL} \| = \| e^{-tA^*A} \|$.

This is the intuitive reason to look at commutators of the form $[A,B]$.

**Definition 2.3.3** (Hypocoercivity). With the same notation of Definition 2.3.1, assume $T$ generates a continuous semigroup. Then $T$ is said to be $\lambda$-hypocoercive on $\tilde{H}$ if there exists a constant $\kappa > 0$ such that

$$
\| e^{-Tt}h \|_{\tilde{H}} \leq \kappa e^{-\lambda t} \| h \|_{\tilde{H}} \quad \forall h \in \tilde{H} \text{ and } t \geq 0. \tag{2.3.4}
$$

We remark that the only difference between Definition 2.3.1 and Definition 2.3.3 is in the constant $\kappa$ on the right hand side of (2.3.4), when $\kappa > 1$. Thanks to this constant, the concept of hypocoercivity is invariant under a change of equivalent norm, as opposed to the concept of coercivity which relies on the choice of the Hilbert norm. Hence the basic idea employed in the proof of exponentially fast convergence to equilibrium for degenerate diffusions generated by operators in the form (2.3.3), is to appropriately construct a norm on $\tilde{H}$, equivalent to the existing one, and such that in this norm the operator is coercive. We now recast one of the main results of the hypocoercivity theory. Notice that in the following Theorem 2.3.4, assumption (2.3.5) is in the spirit of the UFG condition, see [56] and references therein.

**Theorem 2.3.4.** Let $L$ be an operator of the form $L = A^*A + B$, with $B^* = -B$, $K = KerL$ and assume there exists $N \in \mathbb{N}$ such that

$$
[C_{j-1}, B] = C_j + R_j \quad 1 \leq j \leq N + 1, \quad C_0 = A, \quad C_{N+1} = 0. \tag{2.3.5}
$$

Consider the following assumptions: for $k = 0, ..., N + 1$
1. \([A, C_k]\) is relatively bounded with respect to \([C_j]_{0 \leq j \leq k}\) and \([C_j A]_{0 \leq j \leq k-1}\).

2. \([C_k, A^*]\) is relatively bounded with respect to \(I\) and \([C_j]_{0 \leq j \leq k}\) (here \(I\) indicates the identity operator on \(\mathcal{H}\)).

3. \(R_k\) is relatively bounded with respect to \([C_j]_{0 \leq j \leq k-1}\) and \([C_j A]\)_{0 \leq j \leq k-1}.

If Assumptions 1 – 3 are satisfied then there exists a scalar product \(((\cdot, \cdot))\) on \(\mathcal{H}^1\) defining a norm equivalent to the \(\mathcal{H}^1\) norm (see (2.3.2)) and such that

\[
\forall h \in \mathcal{H}^1 / \mathcal{K}, \quad ((h, Lh)) \geq K \sum_{j=0}^{N} \|C_j h\|^2, \tag{2.3.6}
\]

for some constant \(K > 0\). Furthermore, if

\[
\sum_{j=0}^{N} C_j^* C_j \text{ is } \kappa\text{-coercive for some } \kappa > 0 \tag{2.3.7}
\]

then there exists a constant \(\lambda > 0\) such that

\[
\forall h \in \mathcal{H}^1 / \mathcal{K}, \quad ((h, Lh)) \geq \lambda (h, h). \tag{2.3.8}
\]

In particular, \(L\) is hypocoercive in \(\mathcal{H}^1 / \mathcal{K}\), i.e.

\[
\|e^{-tL}\|_{\mathcal{H}^1 / \mathcal{K} \to \mathcal{H}^1 / \mathcal{K}} \leq Ce^{-\lambda t}
\]

for some \(C, \lambda > 0\).

The scalar product \(((\cdot, \cdot))\) is constructed as follows:

\[
((h, h)) := \|h\|^2 + \sum_{k=0}^{N} a_k \|C_k h\|^2 + b_k (C_k h, C_{k+1} h),
\]

with \(a_k\) and \(b_k\) positive constants. So in the above theorem the statements about the scalar product \(((h, h))\) have to be interpreted to mean that there exist constants \(a_k\) and \(b_k\) such that the above scalar product is equivalent to the \(\mathcal{H}^1\) norm and (2.3.6) holds.

We would like to remark that the space \(\mathcal{K}^\perp\) is the same independent of whether we consider the \(\mathcal{H}\) norm, the Sobolev norm (2.3.2) or the homogeneous norm

\[
\|h\|^2_1 := \sum_{j=0}^{N} \|C_j h\|^2.
\]

This can be shown by simply observing that \(h \in \mathcal{K} \Rightarrow C_j h = 0\) for any \(j\), which is true by induction, combining the fact that \(\mathcal{K} = Ker A \cap Ker B\) with the definition of the \(C_j\)’s.
All the hypocoercivity theorems for linear operators (2.3.3) have the same structure: we first have some technical assumptions (namely, Assumptions 1-3) on the successive commutators between $A$ and $B$, on the implications of which I will come to in a moment. These requirements ensure that we can construct the norm in which $L$ will be coercive. Also, Though the crucial assumption is (2.3.7) which, not surprisingly, is the requirement that a Poincaré Inequality should hold. I also would like to stress how the only commutators needed in this context are the commutators between $A$ and $B$. Commutators like $[A_i, A_j]$ (which do not necessarily vanish) play a crucial role in the theory of hypoellipticity, though they are not needed when we are solely interested in the problem of exponential convergence to equilibrium. In other words, what we are looking at is the interaction between the stochastic and the deterministic part of the dynamics. However, Theorem 2.3.4 gives exponential convergence in $H^1$ (actually in $H^1/K$); what can we do in order to get the same result in $H$? It is the case that Assumptions 1-3 are enough in order to obtain a regularization result, which is the content of next theorem:

**Theorem 2.3.5.** With the same notation as in Theorem 2.3.4, if Assumptions 1-3 are satisfied then for any $t > 0$ we have

$$\|C_k e^{-tL} h\| \leq \frac{C}{t^{k+\frac{1}{2}}} \|h\|, \quad \forall k = 0, ..., N,$$

for all functions $h \in H$.

Estimate (2.3.8) is optimal for $t < 1$ and it is optimal in general also for $t \to \infty$; in fact, in order to obtain exponential convergence we also need a Poincaré inequality to hold, as we have already stressed. This fact can be shown either by direct computation or using Malliavin calculus based techniques (see [56]).

The strategy used in order to prove the bound (2.3.8) comes from a paper by Herau [27] and makes use of an appropriate Lyapunov functional; such a Lyapunov functional is precisely the (time dependent version of the) auxiliary scalar product $((\cdot, \cdot))$ used in Theorem 2.3.4.

We shall employ the same technique in Section 3.3 for the case at hand. It turns out that combining this technique with a semigroup approach leads to an analogous pointwise result for the derivatives of any order of the semigroup. We shall present this method in Chapter 6.

We would like to point out how this theory is typically applied with $H$ being the $L^2$ space weighted with some invariant measure $\rho$. Even if this is not an explicit requirement, the existence of a unique equilibrium state for the process is a sort of underlying assumption in the application of this theory. This is even more true in the nonlinear theory of hypocoercivity, see
[76, Part III], where the main abstract theorem requires the existence of a Lyapunov functional "which is dissipated by the equation and admits a unique absolute minimizer." In order to present some further results about the so called hypocoercivity in entropic sense, let $\mu$ and $\nu$ be two probability measures with densities $\rho$ and $f$ in $\mathbb{R}^d$; then the Boltzmann $H-$functional (or Kullback Information), is defined as

$$H_\rho(f) = \int f \log \left( \frac{f}{\rho} \right) \, dx = \int h \log h \, d\rho, \quad f = \rho h,$$

and the Fisher information $I_\rho(f)$ is given by

$$I_\rho(f) = \int |\nabla \log(h)|^2 \, dx = \int h |\nabla \log h|^2 \, d\rho, \quad f = \rho h.$$  \hfill (2.3.9)

**Theorem 2.3.6.** Let $V(x) \in C^2(\mathbb{R}^d)$ such that $\mu(dx) = e^{-V(x)} \, dx$ is a probability measure on $\mathbb{R}^d$ and assume that $\mathcal{L} = \sum_{j=0}^M A_j^* A_j + B$ generates a semigroup on a suitable space of positive functions. Here we assume $\{A_j\}_{1 \leq j \leq M}$ and $B$ to be first order differential operators with smooth coefficients; $^*$ denotes adjoint in the weighted $L^2(\mu)$ and as before $B = -B^*$. Assume there exists $N \in \mathbb{N}$ such that

$$[C_{j-1}, B] = C_j + R_j \quad 1 \leq j \leq N + 1, \quad C_0 = A, C_{N+1} = 0.$$  

If, for $0 \leq k \leq N + 1$ the following assumptions are fulfilled

1. $[A, C_k]$ is pointwise bounded with respect to $A$.
2. $[C_k, A^*]$ is pointwise bounded with respect to $I$ and $\{C_j\}_{0 \leq j \leq k}$.
3. $R_k$ is pointwise bounded with respect to $\{C_j\}_{0 \leq j \leq k-1}$.
4. $[A, C_k]^*$ is pointwise bounded relatively to $I$ and $A$.
5. there exists a positive constant $\lambda > 0$ such that $\sum_k C_k^* C_k \geq \lambda I$ pointwise on $\mathbb{R}^d$ ($I$ is the identity matrix on $\mathbb{R}^d$).
6. The probability measure $\mu$ satisfies a logarithmic Sobolev inequality.  \hfill 1

Then the Kullback information (2.3.9) and the Fisher information (2.3.10) decay exponentially fast to zero.

---

1 A probability measure $\nu$ satisfies a Logarithmic Sobolev Inequality if $H_\nu(\nu) \leq \frac{1}{2c} I_\nu(\nu)$ for some positive constant $c$ and for any probability measure $\nu$ (with the understanding that $H_\nu = I_\nu = \infty$ if $\nu$ is not absolutely continuous with respect to $\mu$).
2.4. Semiclassical Approach

In the above, pointwise bounded is referred to the vector field associated with the involved differential operator. More explicitly, suppose \( A = a(x) \nabla \), with \( a(x) \), \( \mathbb{R}^d \)-valued. The commutator \([A, C_k]\) is another differential operator, of the form say \( \xi_k(x) \nabla \). Then assumption 1 simply says that \(|\xi_k(x)|\) should be bounded, for all \( x \), by a multiple of \(|a(x)|\). Analogously for the other assumptions.

The importance of controlling the Boltzmann \( H \) functional is immediately seen in view of the following Kullback inequality

\[
\frac{1}{2} \| f - \rho \|_{L^1}^2 \leq H_\rho(f),
\]

which we shall be more specific about in the following (see (3.1.2)).

Now a regularization result in the same spirit as \((2.3.8)\).

**Theorem 2.3.7.** With the same notation as in Theorem 2.3.6, let \( V(x) \in C^2(\mathbb{R}^d) \) be such that \( \mu(dx) = e^{-V(x)}dx \) is a probability measure on \( \mathbb{R}^n \) and assume that \( \mathcal{L} \) generates a semigroup on a suitable space of positive functions. If Assumptions 1 – 4 of Theorem 2.3.6 are fulfilled, then the following bounds hold

\[
\int h_t \left| C_k \log h_t \right|^2 \, d\mu \leq \frac{C}{t^{2k+1}} \int h_0 \log h_0 \, d\mu \quad \forall k = 0, ..., N,
\]

where \( h_t = f_t/\rho \) and \( f_t \) is the density of the law of the process with generator \(-\mathcal{L}\).

2.4 Semiclassical Approach

2.4.1 Basics of Semiclassical Analysis

Following [50], we shall give in this section some basic definitions in semiclassical pseudodifferential calculus. The naive idea that we wish to convey is the following: a classical observable, the role of which will be played by symbols, is any smooth function defined on the phase space \( \mathbb{R}^{2n} \). A quantum observable, which in semiclassical language we shall call quantization of the symbol, is any selfadjoint operator on \( L^2(\mathbb{R}^n) \). Pseudodifferential calculus provides a correspondence between the space of classical observables and the space of quantum observables; the aim of the game is studying properties of the operator through properties of the associated function.

An order function on \( \mathbb{R}^N \) is a function \( g \in C^\infty(\mathbb{R}^N; \mathbb{R}_+) \) such that

\[
\partial_\alpha^2 g = O(\rho) \quad \forall \alpha \in \mathbb{N}^N,
\]
uniformly in \( \mathbb{R}^{2n} \). Typical examples of order functions are

\[
\langle x \rangle^m := (1 + |x|^2)^{m/2}, \quad x \in \mathbb{R}^N, \text{ for any fixed } m \in \mathbb{R}.
\]

Given an order function \( g \), the semiclassical space of symbols \( S_N(g) \) is defined as the space of smooth functions \( a(x) \) on \( \mathbb{R}^N \) such that \( \partial_x^n a(x) = \mathcal{O}(g) \) uniformly in \( x \). With this definition \( S_N(1) \) is the set of uniformly bounded \( C^\infty \) functions on \( \mathbb{R}^N \) with uniformly bounded derivatives of any order. If a function \( a(x) \) belongs to \( S_N(g) \) for some order function \( g \) then \( a(x) \) is said to be a symbol.

**Definition 2.4.1.** For \( a(x, y, \xi) \in S_{3n}(\langle \xi \rangle^m) \), we define the pseudodifferential operator of symbol \( a \) to be

\[
\text{Op}_a u(x) := \frac{1}{(2\pi)^n} \int_{\mathbb{R}^{2n}} e^{i(x-y)\xi} a(x, y, \xi) u(y) \, dy \, d\xi, \quad u \in C^\infty_0(\mathbb{R}^n).
\]

We shall use equivalently the notation \( \text{Op}_a \) and \( \text{Op}(a) \).

A priori the above definition is not well posed. Indeed, \( e^{i(x-y)\xi} a(x, y, \xi) u(y) \) is not integrable for any \( m \) and \( n \). In order to make sense out of the integral in (2.4.1), we use the Schwartz kernel Theorem (Theorem A.3.3). To this end, consider the integral

\[
I(a) = \int e^{i(x-y)\xi} a(x, y, \xi) d\xi,
\]

If \( a(x, y, \xi) \in S_{3n}(\langle \xi \rangle^m) \) then \( I(a) \) is absolutely convergent for any \( m < -n \). So, how to interpret \( I(a) \) when \( m \geq -n \)? Simple key observation: let \( P \) be the operator

\[
P(\xi, D_y) = \frac{1 - \xi D_y}{1 + \xi^2}, \quad D_y = \frac{\partial_y}{i}.
\]

Then \( P(e^{i(x-y)\xi}) = e^{i(x-y)\xi} \) so that for any positive integer \( k \), \( P^k(e^{i(x-y)\xi}) = e^{i(x-y)\xi} \); integrating by parts \( k \) times gives

\[
\int e^{i(x-y)\xi} a(x, y, \xi) u(y) \, dy \, d\xi = \int e^{i(x-y)\xi} P^k(a(x, y, \xi) u(y)) \, dy \, d\xi =: I_k(a) u(x)
\]

where

\[
P^k(a u) = \left( \frac{1 + \xi D_y}{1 + \xi^2} \right)^k (a u) = \mathcal{O}(\langle \xi \rangle^{m-k})
\]

uniformly for large \( \xi \). (2.4.3) implies that \( I_k(a)(u) \) is convergent for \( m < k - n, \forall k > 0 \). It can be shown that, for any \( k > m + n, I_k(a) \) defines a continuous linear operator from \( C^\infty_0(\mathbb{R}^n) \) to

\[\text{We are using the usual notation: given a function } h(x), f(x) = \mathcal{O}(h) \text{ uniformly in } \mathbb{R}^n \text{ if there exists } M > 0 \text{ s.t. } |f(x)| \leq M |h(x)|, \forall x \in \mathbb{R}^n.\]
2.4. Semiclassical Approach

$C^\infty(\mathbb{R}^n)$; this fact allows us to interpret $I(a)$ as the distribution kernel associated with $\text{Op}(a)$.\(^3\)

With this in mind, we have that by definition $\text{Op}_a : C_0^\infty(\mathbb{R}^n) \to C^\infty(\mathbb{R}^n)$ and $\text{Op}_a$ can be extended in a unique way to a linear continuous operator $\text{Op}_a : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$. Denoting by $\bar{a}$ the complex conjugate of $a$, we have that $\text{Op}_{\bar{a}}$ is the formal $L^2$ adjoint of $\text{Op}_a$. Also, if the symbol $a$ is polynomial in $\xi$, then $\text{Op}_a$ is a differential operator (intuitively, differentiating corresponds to multiplying by $\xi$ in Fourier space). Finally, notice that if $a \in S_{2n}(\langle \xi \rangle^m)$ then $a((1-t)x + ty, \xi) \in S_{3n}(\langle \xi \rangle^m)$, $t \in [0, 1]$, and we can define

$$\text{Op}(a, t) := \text{Op}(a((1-t)x + ty, \xi)).$$

When $t = 0$, $\text{Op}(a, 0)$ is called the standard or left quantization of the symbol $a$; when $t = 1$, $\text{Op}(a, 1)$ is called the right quantization; for $t = 1/2$ we obtain the Weyl quantization, which we shall extensively refer to in the next section. Here we just notice that the Weyl quantization of a real valued symbol gives an $L^2$-symmetric operator. In the next section we will denote the Weyl quantization of the symbol $q$ by $q^w$.

2.4.2 Singular Space theory and Spectral theory for quadratic hypoelliptic operators.

In this section we consider evolution equations associated with general quadratic operators

$$\begin{cases} \frac{\partial u}{\partial t}(t, x) + q^w(x, D_x)u(t, x) = 0 \\ u(t, \cdot)|_{t=0} = u_0 \in L^2(\mathbb{R}^n), \end{cases}$$

(2.4.4)

and address the problem of the exponential return to equilibrium for these systems. Quadratic operators are pseudodifferential operators, defined in the Weyl quantization

$$q^w(x, D_x)u(x) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^{2n}} e^{i(x-y) \cdot \xi} q \left( \frac{x+y}{2}, \xi \right) u(y) dy d\xi,$$

(2.4.5)

by symbols $q(x, \xi)$, with $(x, \xi) \in \mathbb{R}^n \times \mathbb{R}^n$, which are complex-valued quadratic forms

$$q : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{C}.$$

These operators are differential operators with simple and fully explicit expression. Indeed, the Weyl quantization of the quadratic symbol $x^\alpha \xi^\beta$, with $(\alpha, \beta) \in \mathbb{N}^{2n}$, $(\alpha + \beta) = 2$, is the differential operator

$$x^\alpha D_\xi^\beta + D_x^\beta x^\alpha, \quad D_x = i^{-1} \partial_x.$$

\(^3\)It can be shown that as $a$ varies in $S_{3n}(\langle \xi \rangle^m)$, $I(a)$ is continuous from $S_{3n}(\langle \xi \rangle^m)$ to $\mathcal{D}'(\mathbb{R}^n \times \mathbb{R}^n)$. Also, this definition of $I(a)$ is the only one that coincides with (2.4.2) when $m < -n$. 
Notice that quadratic operators are non-selfadjoint operators and that those with symbols having non-negative real parts are accretive (see Definition A.3.1). Since the classical work by J. Sjöstrand [72], a complete description for the spectrum of elliptic quadratic operators is known. Elliptic quadratic operators are quadratic operators whose symbols satisfy the ellipticity condition

\[(x, \xi) \in \mathbb{R}^{2n}, \quad q(x, \xi) = 0 \Rightarrow (x, \xi) = 0. \tag{2.4.6}\]

In a recent work [30], the spectral properties of non-elliptic quadratic operators, that is operators whose symbols may fail to satisfy the ellipticity condition on the whole phase space \(\mathbb{R}^{2n}\), were investigated. For any quadratic operator whose symbol has a real part with a sign, say here a symbol with non-negative real part

\[\text{Re } q \geq 0, \tag{2.4.7}\]

it was pointed out the existence of a particular linear vector space \(S\) in the phase space \(\mathbb{R}^{2n}\), \(S \subset \mathbb{R}^{2n}\), intrinsically associated to the symbol \(q\) and called singular space, \(S\), which plays a basic role in the understanding of the properties of this non-elliptic quadratic operator. In particular, this work [30] (Theorem 1.2.2) gives a complete description for the spectrum of any non-elliptic quadratic operator \(q^w(x, D_x)\) whose symbol \(q\) has a non-negative real part, \(\text{Re } q \geq 0\), and satisfies an assumption of partial ellipticity along its singular space,

\[(x, \xi) \in S, \quad q(x, \xi) = 0 \Rightarrow (x, \xi) = 0. \tag{2.4.8}\]

Under these assumptions, the spectrum of the quadratic operator \(q^w(x, D_x)\) is shown to be composed of a countable number of eigenvalues with finite multiplicity and the structure of the spectrum is similar to the one known for elliptic quadratic operators [72]. This condition of partial ellipticity is weaker than the condition of ellipticity in general, \(S \subset \mathbb{R}^{2n}\), and allows to deal with more degenerate situations. An important class of quadratic operators satisfying condition (2.4.8) are those with zero singular space \(S = \{0\}\). In this case, the condition of partial ellipticity trivially holds.

In this section we shall consider this class of quadratic operators with zero singular space and focus on the structure of the bottom of their spectra. More specifically, we shall see that the first eigenvalue in the bottom of their spectra has always algebraic multiplicity one with an eigenspace spanned by a ground state of exponential type. We shall also give an explicit formula for the spectral gap which is computable via a simple algebraic calculation; and finally
answer the question of long time behavior of the associated evolution equations by proving the property of exponential return to equilibrium for these quadratic systems. We begin by recalling miscellaneous facts and notations about quadratic operators. In all the following, $q : \mathbb{R}_x^n \times \mathbb{R}_\xi^n \rightarrow \mathbb{C}$, $(x, \xi) \mapsto q(x, \xi)$, stands for a complex-valued quadratic form with a non-negative real part

$$\text{Re } q(x, \xi) \geq 0, \ (x, \xi) \in \mathbb{R}^{2n}. \quad (2.4.9)$$

Associated to the quadratic symbol $q$ is the numerical range $\Sigma(q)$ defined as the closure in the complex plane of all its values,

$$\Sigma(q) = q(\mathbb{R}_x^n \times \mathbb{R}_\xi^n). \quad (2.4.10)$$

The Hamilton map $F \in M_{2n}(\mathbb{C})$ associated to the quadratic form $q$ is the unique map defined by the identity

$$q((x, \xi); (y, \eta)) = \sigma((x, \xi), F(y, \eta)), \ (x, \xi) \in \mathbb{R}^{2n}, (y, \eta) \in \mathbb{R}^{2n}, \quad (2.4.11)$$

where $q(\cdot, \cdot)$ stands for the polarized form associated to the quadratic form $q$ (see Definition A.3.5) and $\sigma$ is the canonical symplectic form on $\mathbb{R}^{2n}$,

$$\sigma((x, \xi), (y, \eta)) = \xi y - x \eta, \ (x, \xi) \in \mathbb{R}^{2n}, (y, \eta) \in \mathbb{R}^{2n}. \quad (2.4.12)$$

We denote by $\sigma$ both the canonical symplectic form and the spectrum of the operator $q'$; however, it should be clear from the context which of the two we are referring to. On a practical level, what (2.4.11) says is that given a quadratic operator we can associate to it, in a unique (and purely algorithmic) way, a $2n \times 2n$ matrix, the Hamilton map. The real and imaginary parts of the Hamilton map $F$,

$$\text{Re } F = \frac{1}{2}(F + F^\dagger) \quad \text{and} \quad \text{Im } F = \frac{1}{2i}(F - F^\dagger),$$

are the Hamilton maps associated to the quadratic forms $\text{Re } q$ and $\text{Im } q$, respectively. Analogously to what we said for hypocoercive operators, the real part of $F$ is related to the stochastic part of the dynamics, whereas $\text{Im } F$ comes from the conservative component of the evolution. Associated to the quadratic symbol $q$ is the singular space $S$ defined in [30] as the subvector space in the phase space equal to the following intersection of kernels

$$S = \bigcap_{j=0}^{+\infty} \text{Ker}[\text{Re } F(\text{Im } F)^j] \cap \mathbb{R}^{2n}.$$
2.4. Semiclassical Approach

Notice that the Cayley-Hamilton theorem applied to $\text{Im } F$ shows

$$(\text{Im } F)^k X \in \text{Vect}(X, ..., (\text{Im } F)^{2n-1} X), \ X \in \mathbb{R}^{2n}, \ k \in \mathbb{N};$$

where $\text{Vect}(X, ..., (\text{Im } F)^{2n-1} X)$ is the vector space generated by the span of the vectors $X, \ldots, (\text{Im } F)^{2n-1} X$. The singular space is therefore equal to the finite intersection of the kernels

$$S = \left( \bigcap_{j=0}^{2n-1} \text{Ker} \left[ \text{Re } F(\text{Im } F)^j \right] \right) \cap \mathbb{R}^{2n}. \quad (2.4.14)$$

As mentioned above, when the quadratic symbol $q$ satisfies a condition of partial ellipticity along its singular space $S$,

$$(x, \xi) \in S, \ q(x, \xi) = 0 \Rightarrow (x, \xi) = 0, \quad (2.4.15)$$

Theorem 1.2.2 in [30] gives a complete description for the spectrum of the quadratic operator $q^w(x, D_x)$ which is only composed of eigenvalues with finite algebraic multiplicity

$$\sigma(q^w(x, D_x)) = \left\{ \sum_{\lambda \in \sigma(F)} (r_\lambda + 2k_\lambda)(-i\lambda) : k_\lambda \in \mathbb{N} \right\}, \quad (2.4.16)$$

where $F$ is the Hamilton map associated to the quadratic form $q$, $r_\lambda$ is the dimension of the space of generalized eigenvectors of $F$ in $\mathbb{C}^{2n}$ belonging to the eigenvalue $\lambda \in \mathbb{C}$,

$$\Sigma(q|S) = \overline{q(S)} \text{ and } \mathbb{C}_+ = \{ z \in \mathbb{C} : \text{Re } z > 0 \}. $$

We are particularly interested in evolution problems associated to accretive quadratic operators $q^w(x, D_x)$ with zero singular space $S = \{0\}$. These operators where shown to be hypoelliptic [65]. Notice that in this case $\overline{q(S)} = \{0\}$ and the condition of partial ellipticity along the singular space trivially holds. The spectrum then reduces to

$$\sigma(q^w(x, D_x)) = \left\{ \sum_{\lambda \in \sigma(F)} (r_\lambda + 2k_\lambda)(-i\lambda) : k_\lambda \in \mathbb{N} \right\}. \quad (2.4.17)$$

Define

$$\mu_0 = \sum_{\lambda \in \sigma(F)} -i\lambda r_\lambda \quad (2.4.18)$$

and

$$\tau_0 = \min_{\lambda \in \sigma(F)} \text{Re}(2(-i\lambda)) = 2 \min_{\text{Im } \lambda > 0} \text{Im } \lambda, \quad (2.4.19)$$
2.4. Semiclassical Approach

Figure 2.1: Spectrum of $q^w$.

The following Theorem 2.4.2 shows that the first eigenvalue in the bottom of the spectrum $\mu_0$ has always algebraic multiplicity one with an eigenspace spanned by a ground state of exponential type and that the gap between the eigenvalues is constant and is exactly given by the positive rate $\tau_0 > 0$ (see Figure 2.1). The proof of the following two theorems, Theorem 2.4.2 and Theorem 2.4.3, can be found in [62].

**Theorem 2.4.2.** With the notation introduced so far, let $q^w(x, D_x)$ be a quadratic operator whose symbol is a complex-valued quadratic form with a non-negative real part $\text{Re} \, q \geq 0$ and a zero singular space $S = \{0\}$. Then, the first eigenvalue in the bottom of the spectrum defined in (2.4.18) has algebraic multiplicity one and the eigenspace

$$\text{Ker}(q^w(x, D_x) - \mu_0) = \mathbb{C}u_0,$$

is spanned by a ground state of exponential type

$$u_0(x) = e^{-a(x)} \in \mathcal{S}(\mathbb{R}^n),$$

where $a$ is a complex-valued quadratic form on $\mathbb{R}^n$ whose real part is positive definite. Moreover, the gap between the eigenvalues of the operator $q^w(x, D_x)$ is given by $\tau_0 > 0$ defined in (2.4.19),

$$\sigma(q^w(x, D_x)) \setminus \{\mu_0\} \subset \{z \in \mathbb{C} : \text{Re} \, z \geq \text{Re} \, \mu_0 + \tau_0\}.$$

Let us underline that operators satisfying the assumptions of Theorem 2.4.2 are not elliptic in general and that their symbol may fail to satisfy the ellipticity condition (2.4.8). On the
other hand, notice that the adjoint operator (see [33], p.426) is actually given by the quadratic operator
\[ q^w(x, D_x)^* = \overline{q}^w(x, D_x), \]
whose symbol is the complex conjugate symbol of \( q \). It is therefore a complex-valued quadratic form with non-negative real part and a zero singular space for which Theorem 2.4.2 applies. Then \( \mu_0 \) is also the first eigenvalue in the bottom of the spectrum for the quadratic operator \( q^w(x, D_x)^* \), and we know from Theorem 2.4.2 that the eigenspaces associated with the eigenvalue \( \mu_0 \) for both operators \( q^w(x, D_x) \) and \( q^w(x, D_x)^* \) are one-dimensional subvector spaces with ground states of exponential type. We shall assume further that the two operators have same ground state

\[ \text{Ker}(q^w(x, D_x) - \mu_0) = \text{Ker}(q^w(x, D_x)^* - \mu_0) = \mathbb{C}u_0 \subset S(\mathbb{R}^n), \]

(2.4.20)

with \( u_0(x) = e^{-a(x)} \), \( x \in \mathbb{R}^n \); where \( a \) is a positive definite quadratic form on \( \mathbb{R}^n \).

**Theorem 2.4.3.** Let \( q^w(x, D_x) \) be a quadratic operator satisfying the assumptions of Theorem 2.4.2. Assume that this quadratic operator is real and satisfies (2.4.20). Using the notations introduced in (2.4.18) and (2.4.19), we consider the operator
\[ Q = q^w(x, D_x) - \mu_0. \]

Then, for all \( 0 \leq \tau < \tau_0 \), there exists a positive constant \( C > 0 \) such that
\[ \forall t \geq 0, \forall u \in L^2(\mathbb{R}^n), \|e^{-tQ}u - c_u u_0\|_{L^2(\mathbb{R}^n)} \leq Ce^{-\tau t}\|u\|_{L^2(\mathbb{R}^n)}, \]

where \( c_u \) is the \( L^2(\mathbb{R}^n) \) scalar product of \( u \) and \( u_0/\|u_0\|_{L^2(\mathbb{R}^n)}^2 \).

Finally notice that this theory is developed on the flat \( L^2 \), as opposed to the theory of hypocoercivity, the natural setting of which is the weighted \( L^2 \). We shall show in Chapter 5 how to go from one space to the other.

\[ ^4 \text{We will show at the beginning of Chapter 5 that this condition is always fulfilled in the cases of interest to us.} \]
2.5 Comparison between hypocoercivity theory and singular space theory

Here we will be referring only to quadratic operators (unless otherwise specified) and we use the notation of the previous two sections. As we have already noticed both the hypocoercivity approach and the singular space approach are related to the notion of hypoellipticity. The former because of the use of commutators: we may naively observe that in order for the Poincaré inequality (2.3.7) to hold, $A = C_0, C_1, \ldots, C_N$ (notation of Theorem 2.3.4) must span $\mathbb{R}^n$ at each point. Therefore, if the operator is hypocoercive it is also hypoelliptic (and this is true even if the operator is not quadratic). On the other hand we have already mentioned how quadratic operators with trivial singular space are hypoelliptic. But this is surely not the only way in which the two approaches communicate to each other. We have already stressed (see comments before Definition 2.3.3) that if $A^*A$ and $B$ commute then the coercivity of $A^*A$ leads to the coercivity of $L$. Which is where the commutators $[A, B], [[A, B], B], [[[A, B], B], B]$ etc, come in. In other words, dissipation takes place because of the interaction between $A^*A$ and $B$; hence the use of commutators, which look at the interplay between the stochastic and the deterministic part of the dynamics. The singular space theory morally reaches the same conclusion. Indeed the definition of the relevant space $S$ (see Definition 2.4.14) on which to test the non degeneracy of the symbol, focuses again on the interaction between $\text{Re} F$ - dissipative - and $\text{Im} F$ - conservative.

In Chapter 5 we will show how the hypoellipticity of the O-U process and of its transpose are equivalent to the condition $S = \{0\}$. However, the singular space theory point of view gives extra information, which is not obtainable employing other methods, on the exact rate of convergence to the ground state. On a related note, in Chapter 4 we shall prove that for the Markovian approximation (0.0.3) (which is a quite general O-U process), hypoellipticity implies geometric ergodicity. This will be done by using classic ergodic techniques. We are quite sure that the same kind of proof can be done for operators of the form presented in Chapter 6, but this shall be the object of (near) future work.
Chapter 3

Asymptotic Analysis for the Generalized Langevin Equation

\textit{Nihil recte}

\textit{sine exemplo docetur.}

(Columella, De re Rustica)

In this chapter we study various qualitative properties of solutions to the generalized Langevin equation (GLE) in $\mathbb{R}^d$

$$\ddot{q} = -\nabla V(q) - \int_0^t \gamma(t - s) \dot{q}(s) \, ds + F(t), \quad (3.0.1)$$

when $\gamma_m(t) = \sum_{i=1}^m \lambda_i^2 e^{-\alpha_i t}$. We recall that $V(q)$ is a smooth potential (confining or periodic), $F(t)$ a mean zero stationary Gaussian process with autocorrelation function $\gamma(t)$. For this particular choice of the kernel the GLE (3.0.1) becomes (we drop the subscripts $m$ for notational simplicity)

\begin{align*}
\dot{q}(t) &= p(t), \quad q(0) = q_0, \quad (3.0.2a) \\
\dot{p}(t) &= -\nabla q V(q(t)) + \sum_{j=1}^m \lambda_j z_j(t), \quad p(0) = p_0, \quad (3.0.2b) \\
\dot{z}_j(t) &= -\lambda_j p(t) - \alpha_j z_j(t) + \sqrt{2\alpha_j \beta^{-1}} W_j, \quad z_j(0) \sim \mathcal{N}(0, \beta^{-1}) \quad (3.0.2c)
\end{align*}

for $j = 1, \ldots, m$. The process $\{q(t), p(t), z(t)\} \in \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^{md}$ is Markovian with generator
\(-\mathcal{L}\) given by

\[-\mathcal{L} = p \cdot \nabla_q - \nabla_q V(q) \cdot \nabla_p + \sum_{j=1}^{m} \lambda_j z_j(t) \cdot \nabla_p + \sum_{j=1}^{m} \left( -\lambda_j p \cdot \nabla z_j - \alpha_j z_j \cdot \nabla z_j + \alpha_j \beta^{-1} \Delta z_j \right).\]

(3.0.3)

This operator can be written in “sum of squares” form

\[\mathcal{L} = B + m \sum_{i=1}^{d} \sum_{j=1}^{d} A_{ij}^* A_{ij}\]

where, \(A_{ij} = -\sqrt{\beta^{-1} \alpha_i} \partial_{z_{ij}}\), \(A_{ij}^* = -\sqrt{\beta \alpha_i} z_i + \sqrt{\beta^{-1} \alpha_i} \partial_{z_{ij}}\) \(^1\) and

\[B = -p \cdot \nabla_q + \nabla_q V \cdot \nabla_p - m \sum_{j=1}^{m} \lambda_j \left( z_j \cdot \nabla_p - p \cdot \nabla z_j \right)\]

(for more details on this notation we refer the reader to Section 3.1.1). This is a degenerate second order elliptic differential operator of hypoelliptic type [31]. Convergence to equilibrium for models of the form (3.0.2) has been studied using functional analytic techniques [17, 14]. Similar results have also been proved using Markov chain techniques [52, 69]. In this chapter we present an alternative proof of exponentially fast convergence to equilibrium in relative entropy using the recently developed theory of hypocoercivity [76]. Our main results can be summarized as follows.

1. We prove well-posedness of the contraction semigroup generated by \(-\mathcal{L}\), Proposition 3.2.1.

2. We prove ergodicity and exponentially fast convergence to equilibrium for (3.0.2), Theorems 3.1.2, 3.1.3 and 3.1.4.

3. We obtain sharp estimates on derivatives of the Markov semigroup associated to the SDE (3.0.2), Theorem 3.1.6.

4. We prove a homogenization theorem (invariance principle) when the potential \(V(q)\) in (3.0.2) is periodic and we obtain estimates on the diffusion coefficient, Theorem 3.1.8. In order to prove these results we prove compactness of the resolvent of the generator of the SDE (3.0.2), Proposition 3.1.9.

\(^1\)\(A_{ij}^*\) is the adjoint of \(A_{ij}\) in the \(L^2\) space weighted with the invariant measure of the system.
5. We study the white noise limit of the GLE (3.0.1), i.e. the limit as the noise $F(t)$ in (3.0.1) (in the Markovian approximation (3.0.2)) converges to a white noise process. We show that in this limit the solution of (3.0.2) converges strongly to the solution of the Langevin equation

$$\ddot{q} = -\nabla V(q) - \gamma \dot{q} + \sqrt{2 \gamma \beta^{-1}} W$$

(3.0.4)

and we obtain a formula for the friction coefficient $\gamma$ in terms of the coefficients $\{\lambda_j, \alpha_j\}_{j=1}^m$, Theorem 3.1.10.

The rest of the chapter is organized as follows. In Section 3.1 we state our main results and we introduce the notation that we will be using. In Section 3.2 we prove exponentially fast convergence to equilibrium. In Section 3.3 we prove estimates on the derivatives of the Markov semigroup generated by $-\mathcal{L}$ defined in (3.0.3). In Section 3.4 we prove the homogenization theorem and the compactness of the resolvent of $\mathcal{L}$. In Section 3.5 we study the white noise limit.

### 3.1 Statement of Main Results

We will use the notation $X := T^d \times \mathbb{R}^d \times \mathbb{R}^{dm}$ and $Y := \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^{dm}$. We will also denote the process $\{q(t), p(t), z(t)\}$ by $x(t)$. When we study the dynamics (3.0.2) in $X$ the potential $V(q)$ is periodic, whereas when $x(t) \in Y$ the potential will be taken to be confining. The precise assumptions on the potential are given in Assumption 3.1.1 below.

By using a modification of the argument used in the proof of [26, Prop 5.5] (see also [76, Thm. A.5]) we can prove that $-\mathcal{L}$ defined in (3.0.3) generates a contraction semigroup, see Proposition 3.2.1.

Our fist result concerns the ergodicity of the SDE (3.0.2) in $X$ or in $Y$. To prove the ergodicity of the SDE in $Y$ we need to make the following assumptions on the potential.

**Assumption 3.1.1.**

(i) $V(q) \in C^\infty(\mathbb{R}^d)$ is a confining potential.

(ii) There exist strictly positive constants $\beta, \sigma$ such that $\langle \nabla_q V, q \rangle \geq \sigma V(q) + \beta \|q\|^2$ where $\langle \cdot, \cdot \rangle$ and $\|\cdot\|$ denote the Euclidean inner product and norm, respectively.

(iii) There exists a constant $c$ such that $\|\nabla^2 V\| \leq c$, where $\|\cdot\|$ denotes the Frobenious-Perron matrix norm and $\nabla^2$ the Hessian.
Notice that when $q \in \mathbb{R}^d$ then $(ii) \Rightarrow (i)$ in the above Assumption 3.1.1.

The density $\rho_\beta(q, p, z)$ of the invariant measure $\mu_\beta(dq \, dp \, dz)$ of the process (3.0.2), which is the unique solution of the stationary Fokker-Planck equation, is known:

$$\rho_\beta(q, p, z) = \frac{1}{Z_\beta} e^{-\beta \left( \frac{1}{2} |p|^2 + V(q) + \frac{1}{2} \|z\|^2 \right)},$$  

(3.1.1)

where $Z_\beta$ is the normalization constant. This invariant measure is unique and the law of the process (3.0.2) converges exponentially fast to $\mu_\beta$ (geometric ergodicity).

**Theorem 3.1.2 (Ergodicity).** The solution of (3.0.2) with $x(t) \in X$ and $V(q) \in C^1(\mathbb{T}^d)$ is geometrically ergodic. The same holds true when $x(t) \in Y$, provided that the potential $V(q) \in C^2(\mathbb{R}^d)$ satisfies Assumption 3.1.1.

The proof of this theorem, which is based on Markov chain-type arguments and which is similar to the proof presented in [69], see also [52], was given in Section 2.2.

We can prove exponentially fast convergence to equilibrium using tools from the theory of hypocoercivity [76]. We will use the notation $\mathcal{K} := \text{Ker}(L)$ and $H^1_\rho$ for the weighted Sobolev space $H^1$ with respect to $\rho_\beta$ on either $X$ or $Y$.

**Theorem 3.1.3.** Let $-L$ be the generator of the process $x(t) \in X$, the solution of (3.0.2) and assume that $V(q) \in C^2(\mathbb{T}^d)$. Then there exist constants $C$, $\lambda > 0$ such that

$$\|e^{-tL}\|_{H^1_\beta / \mathcal{K} \to H^1_\beta / \mathcal{K}} \leq Ce^{-\lambda t}.$$  

The same holds true when $x(t) \in Y$, provided that the potential satisfies Assumptions 3.1.1(i) and 3.1.1(iii).

Using the tools from [76] which we presented in Section 2.3 we can prove exponentially fast convergence to equilibrium in relative entropy. We recall that the relative entropy (or Kullback information) between two probability measures $\mu$ and $\nu$ with smooth densities $f$ and $\rho$, respectively, is defined as

$$H_\rho(f) = \int f \log \left( \frac{f}{\rho} \right) \, dx.$$  

We will measure the distance in relative entropy between the law of the process $x(t)$ at time $t$ and the equilibrium distribution. Since the operator $\frac{\partial}{\partial t} + L$ is hypoelliptic, the law of the process $x(t)$ in (3.0.2) has a smooth density with respect to Lebesgue which we will denote by $f_t$. 

3.1. Statement of Main Results

**Theorem 3.1.4 (Convergence to Equilibrium).** Let \( f_t \) be the density of the law of the process \( x(t) \) at time \( t \) and assume that \( H_\rho(f_0) < +\infty \) and \( V(q) \in C^2(\mathbb{T}^d) \). Then there exist constants \( C, \alpha > 0 \) such that

\[
H_\rho(f_t) \leq Ce^{-\alpha t}H_\rho(f_0).
\]

The same holds true when \( x(t) \in Y \), assuming that \( H_\rho(f_0) < +\infty \) and provided that the potential \( V(q) \) satisfies Assumption 3.1.1(i) and 3.1.1(iii).

**Remark 3.1.5.** In view of the Csiszar-Kullback (Pinsker) inequality

\[
\frac{1}{2}\|f_t - \rho\|_{L^1}^2 \leq H_\rho(f_t),
\]

Theorem 3.1.4 implies that, for initial data with finite relative entropy, we have exponentially fast convergence to equilibrium in \( L^1 \). For more details on inequality (3.1.2), we refer the reader to [4, 75, 49].

The proofs of Theorems 3.1.3 and 3.1.4 are presented in Section 3.2.

Estimates on the Markov semigroup associated to the Langevin equation and its derivatives can be proved using an appropriate Lyapunov function with time dependent coefficients [27, 23]. We introduce \( C_k, k = 0, 1, 2 \) with \( C_0 = A, C_1 = [A, B] \) and \( C_2 = [C_1, B] \) (see Section 3.1.1). We will use the notation \( L^2_\rho := L^2(\cdot; \mu_\beta(dx)) \) where \( \cdot \) is either \( X \) or \( Y \).

**Theorem 3.1.6 (Estimates on Derivatives of the Markov Semigroup).** Let \(-\mathcal{L}\) be the generator of the process \( x(t) \in X \), the solution of (3.0.2) with \( V(q) \in C^2(\mathbb{T}^d) \). Then the Markov semigroup \( e^{-t\mathcal{L}} \) satisfies the bounds

\[
\|C_k e^{-t\mathcal{L}}\|_{L^2_\rho \to L^2_\rho} \leq \frac{c}{t^{\frac{1+2k}{2}}} , \quad k = 0, 1, 2 \text{ and } t \in (0,1],
\]

for some (explicitly computable) positive constant \( c \). The same holds true when \( x(t) \in Y \), provided that the potential \( V(q) \) satisfies Assumption 3.1.1(i) and (iii).

**Remark 3.1.7.** The short time asymptotics (3.1.3) is typical of hypoelliptic evolution problems. Indeed, if we were to consider an elliptic operator, we would obtain that the short time behaviour of the space derivatives does not depend on the direction in which we are differentiating but only on the order of the derivative. For example, it was shown ([48], [41]-[44]) that for the semigroup \( T(t) \) generated by the elliptic Ornstein-Uhlenbeck process,

\[
\|D^\beta T(t)f\|_{L^2_\rho} \leq \frac{C}{t^{\frac{1+2|\beta|}{2}}} \|f\|_{L^2_\rho}, \quad f \in L^2_\eta(\mathbb{R}^n),
\]
where in the above inequality $\eta$ is the invariant measure of the semigroup (which we shall give more details about in Chapter 5) and $D^\beta$ denotes derivative of order $|\beta|$, for a multi-index $\beta$.

As noticed in [27], using estimate (3.1.3) together with the semigroup property and the contractivity of the semigroup we obtain that \( \forall h \in L^2_\rho \) and \( \forall t > 1/2 \)

\[
\|C_k e^{-tL}h\|_{L^2_\rho} = \left\| C_k e^{-\frac{1}{2}L} \left( e^{-\left(t-\frac{1}{2}\right)L}h\right) \right\|_{L^2_\rho} \\
\leq c 2^{\frac{1+2k}{2}} \|e^{-(t-1/2)L}h\|_{L^2_\rho} \\
\leq c \|h\|_{L^2_\rho},
\]

hence \( \|C_k e^{-tL}h\|_{L^2_\rho} \leq c \left( 1 + \frac{1}{t^{1+2k/2}} \right) \|h\|_{L^2_\rho} \quad k = 0, 1, 2, \; t > 0, \; h \in L^2_\rho \).

(3.1.3) can also be obtained by applying Theorem 2.3.5. Malliavin calculus-based arguments show that estimate (3.1.3) is sharp. Nonetheless for this relatively simple case the sharpness can be shown by hand.

When the potential $V(q)$ is periodic, the particle position, appropriately rescaled, converges weakly to a Brownian motion with a diffusion coefficient which can be calculated in terms of the solution of an appropriate Poisson equation. Results of this form have been known for a long time for the Smoluchowski (overdamped) equation [64, Ch. 13] as well as for the Langevin dynamics [63]. In this paper we prove a similar result for the generalized Langevin equation. We will use the notation $\phi^e := \phi \cdot e$, $p^e := p \cdot e$, where $e$ denotes an arbitrary unit vector in $\mathbb{R}^d$.

When $V(q)$ is 1-periodic then $q(t)$ enters in the definition of the process $x(t) = \{q(t), p(t), z(t)\}$ only mod 1, so we may replace $q(t)$ by $q(t) \in \mathbb{T}^d := \mathbb{R}^d / \mathbb{Z}$. The Markov process $x(t) = \{q(t), p(t), z(t)\}$ has state space $\mathbb{T}^d \times \mathbb{R}^d \times \mathbb{R}^{md}$ and, according to Theorem 3.1.2, it is an ergodic Markov process with invariant measure given by (3.1.1). For this process we prove the homogenization Theorem 3.1.8. To simplify the notation we shall drop the underbar from $\bar{x}(t)$ and $\bar{q}(t)$.

**Theorem 3.1.8 (Homogenization).** Let $x(t)$ be the solution of (3.0.2) with $V(q) \in C^\infty(\mathbb{T}^d)$ with stationary initial conditions. Then the rescaled process $q^e_c(t) := e \cdot c q(t/c^2)$ converges
weakly in $C([0, T], \mathbb{R})$ to a Brownian motion with diffusion coefficient $D$ with
\[
D^e := D e \cdot e = \beta^{-1} \sum_{j=1}^{m} \alpha_j \| \nabla_{z_j} \phi^e \|^2,
\]
where $\phi^e \in L^2_\mu$ is the unique, smooth, mean zero, periodic in $q$ solution of the Poisson equation
\[
L \phi^e = p^e
\]
on $X$. Furthermore, the following estimate holds
\[
0 < D^e \leq \frac{4}{\beta} \sum_{i=1}^{m} \frac{\alpha_i}{\lambda_i^2}.
\]
The proof of this theorem is based on a careful study of the Poisson equation (3.1.6). The well-posedness of this equation follows from the compactness of the resolvent of $L$.

**Proposition 3.1.9.** Let $-L$ be the generator of the process $x(t) \in X$, the solution of (3.0.2) and assume that $V(q) \in C^\infty(T^d)$. Then $L$ has compact resolvent on $L^2_\mu(T^d \times \mathbb{R}^d \times \mathbb{R}^{md})$.

Let $q(t)$ be the solution of the Langevin equation (3.0.4) and let $q^\gamma(t) := q(\gamma t)$. It is well known that this rescaled process converges strongly in the overdamped limit $\gamma \to +\infty$ to the solution of the Smoluchowski equation [57, Ch. 10]
\[
\dot{q} = -\nabla V(q) + \sqrt{2\beta^{-1}} W.
\]
We prove a result of this type for the convergence of solutions to the Markovian approximation of the GLE to the Langevin equation in the strong topology and obtain a formula for the friction coefficient that appears in the limiting Langevin equation.

Consider (3.0.1) with the rescaled noise process
\[
F^e(t) := \frac{1}{\sqrt{\epsilon}} F(t/\epsilon),
\]

---

2Let $B$ be a Banach space. A sequence $x_n$ converges weakly to $x$ in $B$ if $l(x_n) \to l(x)$ for every $l$ in the dual of $B$. In the statement of Theorem 3.1.8 the Banach space at hand is the space of real valued, continuous functions on $[0, T]$. We recall that the dual of $C([0, T], \mathbb{R})$ is the space of measures $\mu$ on $[0, T]$ with bounded total variation on $[0, T]$ (see [78], page 119); the total variation on $[0, T]$ is defined as
\[
\| \mu \|_{TV} := \sup_{f \in C([0, T])} \int_{[0, T]} f(s) \mu(ds).
\]
which is a mean zero stationary Gaussian process with autocorrelation function

\[ \gamma^\epsilon(t) = \frac{1}{\epsilon} \gamma(t/\epsilon). \quad (3.1.10) \]

For the memory kernel we are working with, \( \gamma^\epsilon(t) \) becomes

\[ \gamma^\epsilon_m(t) = \sum_{j=1}^m \frac{\lambda_j^2}{\epsilon} e^{-\alpha_j t}. \quad (3.1.11) \]

Consequently, the rescaled noise process (3.1.9) is obtained by rescaling the coefficients in (3.0.2) according to \( \lambda_j \to \frac{\lambda_j}{\sqrt{\epsilon}} \), \( \alpha_j \to \frac{\alpha_j}{\epsilon} \). Under this rescaling the SDEs become

\[
\begin{align*}
\text{eq}(t) &= \rho(t) dt, \\
\text{eq}(t) &= -\nabla q V(q(t)) dt + \frac{1}{\sqrt{\epsilon}} \sum_{i=1}^m \lambda_i z_i(t) dt, \\
\text{eq}(t) &= -\frac{\lambda_i}{\sqrt{\epsilon}} \rho(t) dt - \frac{\alpha_i}{\epsilon} z_i(t) dt + \sqrt{\frac{2\alpha_i \beta - 1}{\epsilon}} dW_i, \quad i = 1, \ldots, m. \tag{3.1.12c}
\end{align*}
\]

**Theorem 3.1.10 (The White Noise Limit).** Let \( \{q(t), p(t), z(t)\} \in X \) be the solution of (3.1.12) with \( V(q) \in C^2(\mathbb{T}^d) \) and initial conditions having finite moments of any order. Then the process \( \{q(t), p(t)\} \) converges strongly, as \( \epsilon \to 0 \), to the solution of the Langevin equation

\[
\begin{align*}
dQ(t) &= P(t) dt, \\
dP(t) &= \left(-\nabla q V(Q(t)) - \sum_{i=1}^m \lambda_i^2 P(t)\right) dt + \sum_{i=1}^m \sqrt{\frac{2\beta - 1}{\alpha_i}} dW_i, \\
\end{align*}
\]

with the same initial conditions as \( q \) and \( p \). Furthermore, for any \( n \geq 1 \), the following estimate holds

\[ \|q(t) - Q(t)\|_{n,\infty} + \|p(t) - P(t)\|_{n,\infty} \leq C\epsilon^{\frac{1}{2}}, \tag{3.1.14} \]

where \( \|f(t)\|_{n,\infty} := (\mathbb{E} \sup_{t \in [0, T]} |f(t)|^n)^{1/n} \). The same result holds true when \( \{q(t), p(t), z(t)\} \) \( \in Y \) provided that the potential \( V(q) \) satisfies Assumption 3.1.1(iii).

Consequently, the process \( \{q(t), p(t)\} \) converges weakly to the solution of the Langevin equation

\[
\begin{align*}
dQ(t) &= P(t) dt, \\
dP(t) &= \left(-\nabla q V(Q(t)) - \gamma P\right) dt + \sqrt{2\gamma \beta - 1} dW, \\
\end{align*}
\]

where the friction coefficient \( \gamma \) is given by the formula

\[ \gamma = \sum_{j=1}^m \frac{\lambda_j^2}{\alpha_i}. \tag{3.1.16} \]
3.1. Statement of Main Results

3.1.1 Notation

For \( x(t) = (q, p, z) \) in \( Y := \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^{dm} \) or \( x(t) \in X := \mathbb{T} \times \mathbb{R}^d \times \mathbb{R}^{dm} \) consider the operator \( \mathcal{L} \) defined in (5.0.2):

\[
\mathcal{L} = p \cdot \nabla_q - \nabla_q V(q) \cdot \nabla_p + \left( \sum_{j=1}^{m} \lambda_j z_j \right) \cdot \nabla_p + \sum_{j=1}^{m} \left( -\alpha_j z_j \cdot \nabla z_j - \lambda_j p \cdot \nabla z_j + \beta_j^{-1} \alpha_j \Delta z_j \right), \tag{3.1.17}
\]

with kernel \( \mathcal{K} := \text{Ker} \mathcal{L} \). The density of the invariant measure of the process \( x(t) \) is

\[
\rho_\beta(p, q, z) = \frac{1}{Z_\beta} e^{-\beta(V(q) + \frac{1}{2} \rho^2 + \frac{1}{2} |z|^2)}, \quad Z_\beta = \int e^{-\beta(V(q) + \frac{1}{2} \rho^2 + \frac{1}{2} |z|^2)} dp dq dz, \tag{3.1.18}
\]

where \( | \cdot | \) denotes either the Euclidean or the matrix norm. In (3.1.17), \( \nabla \) is the gradient (or the derivative when \( m = 1 \)) and \( \Delta \) the Laplacian. \( \nabla^2 \) denotes the Hessian and if \( O \) is an operator then \( O^* \) is its adjoint in \( L^2_\rho := L^2(\cdot ; \mu_\beta(d\mathcal{X})) \). Define

\[
B = -p \cdot \nabla_q + \nabla_q V(q) \cdot \nabla_p - \sum_{j=1}^{m} \lambda_j (z_j \cdot \nabla_p - \rho \cdot \nabla z_j). \tag{3.1.19}
\]

We easily check that \( B^* = -B \). To simplify the notation, we set \( \beta = \alpha_j = 1 \). When \( m = 1 \) then \( A_i = -\partial_{z_i} \) (derivative with respect to the \( i \)-th component of \( z \)) so that \( A_i^* = -z_i + \partial_{z_i} \) and we can write

\[
\mathcal{L} = B + \sum_{i=1}^{d} A_i^* A_i =: B + A^* A, \tag{3.1.20}
\]

where \( A \) is intended to be the row vector of operators \( (A_1, \ldots, A_d) \) (the same for \( A^* \)). More precisely, if \( m = 1 \) then: \( A : L^2_\rho \rightarrow L^2_\rho \otimes \mathbb{R}^d, B : L^2_\rho \rightarrow L^2_\rho \), \( [A^*, A] : L^2_\rho \rightarrow L^2_\rho \), being \( [A^*, A] := \sum_{j=1}^{d} [A^*_j, A_j] \); on the other hand \( [A, A^*] : L^2_\rho \rightarrow L^2_\rho \otimes \mathbb{R}^d \otimes \mathbb{R}^d \) is a matrix of operators whose \( ij \)-th component is given by \( [A, A^*]_{ij} := [A_i, A^*_j] \); in an analogous way \( [A, A] : L^2_\rho \rightarrow L^2_\rho \otimes \mathbb{R}^d \otimes \mathbb{R}^d \) is a matrix of operators with \( [A, A]_{ij} := [A_i, A_j] \); finally \( C := [A, B], C : L^2_\rho \rightarrow L^2_\rho \otimes \mathbb{R}^d \) is a vector of operators, \( C_i = [A_i, B], i = 1 \ldots d \), and the same holds for \( C_2 := [C, B], C_2 : L^2_\rho \rightarrow L^2_\rho \otimes \mathbb{R}^d \).

When \( m > 1 \) then (3.1.20) becomes

\[
\mathcal{L} = B + \sum_{i=1}^{m} \sum_{j=1}^{d} A^*_{ij} A_{ij} \tag{3.1.21}
\]

with \( A_{ij} = -\partial_{z_{ij}} \) i.e. the partial derivative with respect to the \( j \)-th component of \( z_i \), and \( A^*_{ij} = -z_{ij} + \partial_{z_{ij}} \). We will use the notation

\[
\mathcal{L} = B + A^* A, \tag{3.1.22}
\]
meaning either (3.1.20) or (3.1.21). As for the norms, unless otherwise specified, \( \| \cdot \| \) indicates the norm of \( L^2(\rho) \), \( \| \cdot \|_1 = \| A \cdot \|_2 + \| C \cdot \|_2 + \| C_2 \cdot \|_2 \) is a sort of homogeneous \( H^1(Y; \mu_\beta(dx)) =: H^1_\rho \) norm and \( \| \cdot \|_{H^1_\rho} = \| \cdot \|_2 + \| A \cdot \|_2 + \| C \cdot \|_2 + \| C_2 \cdot \|_2 \) is the usual inhomogeneous one. The inner products in these Hilbert spaces are denoted by \( (\cdot, \cdot)_1 \) and \( (\cdot, \cdot)_{H^1_\rho} \), respectively.

### 3.2 Convergence to Equilibrium

In this section we present the proofs of Theorems 3.1.3 and 3.1.4. As a preliminary result we show that \(-L\) given by (5.0.2) generates a contraction semigroup.

**Proposition 3.2.1.** Let \(-L\) be the generator of the process \( x(t) \in X \), the solution of (3.0.2) and assume that \( V(q) \in C^2(\mathbb{R}^d) \). Then \(-L\) generates a contraction semigroup.

**Proof.** The proof is almost identical to the proof of Proposition 5.5 in [26] and we will be very brief.\(^3\) Let \( L = B + A^*A \). To simplify the notation, we will set all the constants equal to 1 and will also consider the case \( d = m = 1 \). Clearly, \( L \) is an accretive operator. Furthermore, its domain of definition is dense in \( L^2(\rho) \). Thus, we can consider its closure, which we will still denote by \( L \). We define \( T = L + 2I \). From the Lumer-Phillips Theorem, i.e. Theorem A.3.2, in order to prove that \( L \) generates a contraction semigroup it is enough to show that the range of \( T \) is dense in \( L^2(\rho) \). To this end it is sufficient to prove that if

\[
(f, Tu) = 0 \quad \forall u \in C_0^\infty, \quad (3.2.1)
\]

then \( f = 0 \). Notice that Equation (3.2.1) is equivalent to \((A^*A - B + 2I)f = 0\) in the distributional sense. Hence, by hypoellipticity (see Equation (3.2.2)), this implies that \( f \) is a \( C^\infty \) function. Following the proof of [26, Prop. 5.5], we introduce a family of cut-off functions

\[
\zeta_k(q,p,z) := \zeta\left(\frac{q}{k}\right) \zeta\left(\frac{p}{\alpha(k)}\right) \zeta\left(\frac{z}{\omega(k)}\right), \quad \forall k \in \mathbb{N}_+,
\]

where \( \zeta \) is a \( C^\infty \) function satisfying \( \zeta \in [0,1] \), \( \zeta = 1 \) on \( B(0,1) \) and \( \text{supp} \zeta \in B(0,2) \), \( \alpha(k) \) and \( \omega(k) \) are positive functions which we will choose later on. With calculations analogous to

---

\(^3\)Note, however, that rather than transforming \( L \) into a Schrödinger operator and working in a flat \( L^2 \) space, we work with the generator in its original form in the weighted \( L^2 \) space.
3.2. Convergence to Equilibrium

those presented in [26, Prop. 5.5] we have that for any $u \in C^\infty$,

$$(f, T(\zeta_k^2 u)) - (\partial_z(\zeta_k f), \partial_z(\zeta_k u))$$

$$= (\partial_z f, \partial_z(\zeta_k^2 u)) + (f, B(\zeta_k^2 u)) + 2(f, \zeta_k^2 u) - (\partial_z(\zeta_k f), \partial_z(\zeta_k u))$$

$$= (\zeta_k \partial_z f, u \partial_z \zeta_k) - (f \partial_z \zeta_k, u \partial_z \zeta_k) - (f \partial_z \zeta_k, \zeta_k \partial_z u) + 2(f, \zeta_k^2 u) + (f, B(\zeta_k^2 u)).$$

Let now $f$ be the solution of (3.2.1) and choose $u = f$ in the above identity to obtain

$$2\|\partial_z(\zeta_k f)\|^2 + \|\partial_z(\zeta_k f)\|^2 = \|\partial_z(\zeta_k f)\|^2 - (f, B(\zeta_k^2 f)).$$

We use now the identity $(f, B(\zeta_k^2 f)) = (\zeta_k f^2, B\zeta_k)$, which follows from the antisymmetry of $B$, to deduce

$$2\|\zeta_k f\|^2 \leq \|f \partial_z \zeta_k\|^2 - (\zeta_k f^2, B\zeta_k).$$

Setting $\tilde{C}(k) := \sup_{|q| \leq 2k} |\partial_q V(q)|$, we then have

$$2\|\zeta_k f\|^2 \leq \frac{1}{\omega^2(k)} \|f\|^2 + \|f\|^2 + \tilde{C}(k) \frac{k}{\alpha(k)} \|f\|^2 + \frac{k}{\omega(k)} \|f\|^2.$$ 

We now choose $\alpha(k)$ and $\omega(k)$ such that, as $k \to \infty$, $\omega(k) \to \infty$, $\tilde{C}(k)/\alpha(k) \to 0$ and $k/\alpha(k), k/\omega(k) \to 0$. So, letting $k \to \infty$, from the above inequality we obtain $\|f\|^2 = 0$, hence $f = 0$.

3.2.1 Hypocoercivity

Background material on hypocoercivity is presented in Section 2.3. Here we just remark that $S$, the class of Schwartz functions, is dense in $D(A) \cap D(B)$ as well as in $L^2 \rho$. This guarantees that all the operations performed with these (unbounded) operators are well defined. Set $m = 1 = d$, $\alpha = \lambda = \beta = 1$. The first two commutators are

$$C_1 = C = [A, B] = \partial_p \quad \text{and} \quad C_2 = [C, B] = \partial_z - \partial_q.$$ (3.2.2)

Hence the operator is hypoelliptic [31]. Furthermore,

$$[A, A] = 0 \quad [A, C] = 0 \quad [A, C_2] = 0,$$ (3.2.3a)

$$[A, A^*] = I \quad [C, A^*] = 0 \quad [C_2, A^*] = -I,$$ (3.2.3b)

$$[C_2, B] = -\partial^2 V \partial_p - \partial_p,$$ (3.2.3c)

$$[C, C^*] = I \quad [C_2^*, C_2] = -I - \partial^2 q V,$$ (3.2.3d)
where $I$ is the identity operator.

**Remark** For simplicity all the proofs of this chapter are done for the case $m = d = 1$. When $m, d > 1$ everything still holds. For the sake of clarity we write system (3.0.2) in a more extended form when $m, d > 1$:

\[
\begin{align*}
\dot{q}_j &= p_j, \\
\dot{p}_j &= \partial_{q_j} V(q) + \sum_{1 \leq i \leq m} \lambda_{ij} z_{ij}, \\
\dot{z}_{ij} &= -\lambda_{ij} p_j - \alpha_{ij} z_{ij} + \sqrt{2\alpha_{ij}} W_{ij},
\end{align*}
\]

where $1 \leq j \leq d$ and $1 \leq i \leq m$; in other words, $q_j \in \mathbb{R}$ and $p_j \in \mathbb{R}$ are the $j$-th components of $q \in \mathbb{R}^d$ and $p \in \mathbb{R}^d$, respectively, whereas $z_{ij} \in \mathbb{R}$ is the $j$-th component of the $i$-th oscillator. Also in the above $\alpha_{ij}, \lambda_{ij} > 0$ and $W_{ij}$ are independent standard Brownian motions for every $i, j$. The generator of this process is

\[
L_{md} = \sum_{j=1}^{d} \left[ p_j \partial_{q_j} - \partial_{q_j} V \partial_{p_j} + \sum_{i=1}^{m} \lambda_{ij} z_{ij} \partial_{p_j} \right] + \sum_{i=1}^{m} \sum_{j=1}^{d} (-\lambda_{ij} p_j - \alpha_{ij} z_{ij}) \partial_{z_{ij}} + \alpha_{ij} \partial^2_{z_{ij}},
\]

the antisymmetric (in $L^2_\mu$) part of which is

\[
B_{md} = \sum_{j=1}^{d} \left[ p_j \partial_{q_j} - \partial_{q_j} V \partial_{p_j} + \sum_{i=1}^{m} \lambda_{ij} z_{ij} \partial_{p_j} \right] - \sum_{i=1}^{m} \sum_{j=1}^{d} \lambda_{ij} p_j \partial_{z_{ij}},
\]

Therefore

\[
[\partial_{s_k}, B_{md}] = \lambda_{hk} \partial_{p_k}, \quad [\partial_{p_k}, B_{md}] = \partial_{q_k} - \lambda_{ik} \partial_{z_{ik}},
\]

hence the operator is hypoelliptic. From now on we go back to the case $m = d = 1$.

### 3.2.2 Proof of Theorem 3.1.3

**Proof.** We will use Theorem 2.3.4. To this end, set

\[
P = A^* A + C^* C + C_2^* C_2
\]

and notice that $\text{Ker}(P) = \mathcal{K} =: \text{Ker}\mathcal{L}$ contains only constants; in fact

\[
\text{Ker}(P) = \text{Ker}(A^* A) \cap \text{Ker}(C^* C) \cap \text{Ker}(C_2^* C_2) = \text{Ker}(A) \cap \text{Ker}(C) \cap \text{Ker}(C_2).
\]

To show that $\mathcal{K} = \text{Ker}(A^* A) \cap \text{Ker}(C^* C) \cap \text{Ker}(C_2^* C_2)$: the inclusion $\supseteq$ is obvious (if $h$ is constant then it is in the kernel of $\mathcal{L}$). For the other inclusion: if $h \in \mathcal{K}$ then $\|Ah\|^2 + \|Ch\|^2 + \|C_2 h\|^2 = 0 \Rightarrow Ah = Ch = C_2 h = 0$. 


Theorem 2.3.4 requires two sets of hypotheses to be fulfilled. Hypothesis 1, 2 and 3 in Theorem 2.3.4 are quantitative assumptions, which are satisfied in our case with $N = 2$, $C_0 = A$, $C_1 = C$, $R_1 = R_2 = 0$, $R_3 = [C_2, B]$ (this is to have $C_3 = 0$) and thanks to Assumption 3.1.1(iii). Hypothesis 4 requires, in our case, for the operator $P$ to be $\kappa$-coercive on $K^\perp \cong L^2_\rho/K$. The coercivity of $P$ is equivalent to

$$\|Ah\|^2 + \|Ch\|^2 + \|C_2h\|^2 \geq \kappa\|h\|^2,$$

that is, more explicitly,

$$\|\nabla z_h\|^2 + \|\nabla p_h\|^2 + \|\nabla_q h\|^2 \geq \frac{1}{3}(\|\nabla z_h\|^2 + \|\nabla p_h\|^2 + \|\nabla_q h\|^2)$$

so we just need

$$\|\nabla z_h\|^2 + \|\nabla p_h\|^2 + \|\nabla_q h\|^2 \geq \kappa\|h\|^2$$

to hold true. Since $\mu_\beta$ is a product measure, we only need to verify that

$$\int |\nabla_q h|^2 e^{-V(q)} dq \geq \mu \int (h - \langle h \rangle)^2 e^{-V(q)} dq$$

holds true for some constant $\mu$, where the notation $\langle h \rangle := \int h e^{-V(q)} dq$ has been used. It is a standard result that if $V(q) \in C^2(\mathbb{R}^d)$ is such that $e^{-V(q)}/Z$ is a probability density and

$$\frac{|\nabla V(q)|^2}{2} - \Delta V(q) \xrightarrow{|q| \to \infty} +\infty$$

then $e^{-V(q)}/Z$ satisfies a Poincaré inequality (see, e.g., [76, Thm. A.1]). From Assumption 3.1.1(iii), Condition (3.2.4) is satisfied. We can conclude that there exist a scalar product $((\cdot, \cdot))$ inducing a norm equivalent to the inhomogeneous norm of $H^1_\rho$ and a constant $\hat{\lambda} > 0$ such that $L$ is coercive in this norm:

$$\forall h \in L^2_\rho/K, \quad ((h, Lh)) \geq \hat{\lambda}((h, h)). \quad (3.2.5)$$

The auxiliary scalar product is in this case

$$((h, g)) = (h, g) + a(Ah, Ag) + b(C_1 h, C_1 g) + c(C_2 h, C_2 g) + d(Ah, C_1 g) + e(C_1 h, C_2 g),$$

\footnote{To simplify the notation we have set $\beta = 1.$}
for some appropriate positive constants $a, b, c, d, e$. (3.2.5) gives the coercivity of $L$ in norm $(\cdot, \cdot)$, hence $L$ is hypocoercive on $L^2_{\rho}/K$ endowed with the $\| \cdot \|_{H^1_{\rho}}$ norm:

$$\|e^{-tL}h_0\|_{H^1_{\rho}} \leq Ce^{-\lambda t} \|h_0\|_{H^1_{\rho}}.$$  \hfill (3.2.6)

**Remark 3.2.2.** The orthogonal space to $K$ is the same with respect to both the $L^2_{\rho}$, the $(\cdot, \cdot)_1$ and the $(\cdot, \cdot)_{H^1_{\rho}}$ norms; moreover, since $P$ is coercive, these two norms are equivalent.

**Remark 3.2.3.** Theorem 2.3.5 allows us to state a similar result when the initial datum is in $L^2_{\rho}$. In fact, using Remark 3.2.2,

$$\|e^{-tL}h\|_{H^1_{\rho}} \leq \frac{c}{t^2} \|h\|, \quad t \in (0, 1].$$  \hfill (3.2.7)

So, putting together (3.2.6) and (3.2.7) we get, for $0 < t_0 < t, \ t_0 < 1$:

$$\|e^{-tL}h_0\|_{H^1_{\rho}} = \|e^{-(t-t_0)L}e^{-t_0L}h_0\|_{H^1_{\rho}} = \|e^{-(t-t_0)L}h_{t_0}\|_{H^1_{\rho}} \leq c e^{-\lambda(t-t_0)} \|h_{t_0}\|_{H^1_{\rho}} = c e^{-\lambda(t-t_0)} \|e^{-t_0L}h_0\|_{H^1_{\rho}} \leq c e^{-\lambda(t-t_0)} \|h_{t_0}\|,$$  \hfill (3.2.8)

where the notation $e^{-t_0L}h_0 =: h_{t_0}$ has been used.

**Remark 3.2.4.** The proof is identical when $m, d > 1$. In this case we can think of $A$ as a matrix of operators, see (3.1.21).

### 3.2.3 Proof of Theorem 3.1.4

**Proof.** For simplicity we present the proof of this result in one dimension, i.e. $d = 1$, and for $m = 1$; we also set $\alpha = \beta = 1$. The extension to arbitrary dimensions is straightforward. Let $f_t$ denote the density of the law of the process $x(t)$, i.e. the solution of the Fokker-Plank equation

$$\partial_t f_t + L' f_t = 0,$$

where $L'$ denotes the (flat) $L^2$ adjoint of $L$, namely

$$L' = p\partial_q - \partial_q V \partial_p + z \partial_p - p \partial_z - \partial_z (z \cdot) - z^2.$$


If we set \( f_t = \rho h_t \), then \( h_t \) satisfies the equation

\[
\partial_t h_t = B h_t - A^* A h_t.
\]  
(3.2.9)

We apply Theorem 2.3.6 to the operator \( \mathcal{F} = -B + A^* A \) with

\[
A = -\partial_z, \quad C_1 = -\partial_p, \quad C_2 = -\partial_q, \quad Z_2 = I, \quad R_2 = -\partial_z.
\]

Furthermore Assumption 3.1.1 (i) and (iii) together with the Bakry-Emery criterion \(^5\) imply that \( Z^{-1} e^{-V(q)} \) satisfies a Logarithmic Sobolev Inequality (LSI). Hypotheses 1, 2 and 4 are automatically satisfied. We put \( C_2 = \partial_q \) and we added the remainder \( R_2 \) in order to fulfill hypothesis 4. Hypothesis 3 is satisfied on account of Assumption 3.1.1 (iii). Now consider the relative entropy \( H_\rho(f) \),

\[
H_\rho(f) = \int f \log \left( \frac{f}{\rho} \right) dq dp dr = \int h \log h \, d\rho, \quad f = \rho h
\]  
(3.2.10)

and the Fisher information \( I_\rho(f) \)

\[
I_\rho(f) = \int |\nabla \log(h)|^2 dq dp dr = \int h |\nabla \log h|^2 \, d\rho, \quad f = \rho h.
\]  
(3.2.11)

Then if the initial datum has finite relative entropy, we obtain that

\[
H_\rho(f_t) = \mathcal{O}(e^{-t\alpha})
\]  
(3.2.12)

for some \( \alpha > 0 \) and for \( t > 0 \). If the initial datum has also finite Fisher information then

\[
I_\rho(f_t) = \mathcal{O}(e^{-t\alpha}),
\]  
(3.2.13)

as well. \( \square \)

**Remark 3.2.5.** We remark that (3.2.13), together with the LSI, implies (3.2.12).

**Remark 3.2.6.** In view of the LSI, it is interesting to notice that, by applying Theorem 2.3.7, we get the following bounds

\[
\int h_t |C_k \log h_t|^2 d\rho \leq \frac{c}{t^{2k+1}} \int h_0 \log h_0 \, d\rho,
\]  
(3.2.14)

for \( k = 0, 1, 2 \) and \( c \) an explicitly computable positive constant.

\(^5\)Let \( e^{-V(x)} \) be a probability measure on \( \mathbb{R}^n \) such that \( \nabla^2 V \geq k I_n \), with \( k > 0 \) and \( I_n \) the identity matrix of \( \mathbb{R}^n \). Then \( e^{-V} \) satisfies a Logarithmic Sobolev inequality.
3.3 Bounds on the derivatives of the Markov semigroup

Throughout this section we will use the notation \( u = e^{-t\mathcal{L}}u_0 \). We introduce the Lyapunov function, defined for \( t \in (0, 1] \):

\[
F(t) = a_0 t \|Au\|^2 + a_1 t^3 \|Cu\|^2 + a_2 t^5 \|C_2u\|^2 + b_0 t^2 (Au, Cu) + t^4 b_1 (Cu, C_2u) + b_2 \|u\|^2,
\]

where \( a_j, b_j, j = 0, 1, 2 \) are positive constants to be chosen. Also, we will make systematic use of Young’s inequality

\[
|ab| \leq \frac{a^2}{2\delta} + \frac{\delta b^2}{2}, \quad \forall a, b \in \mathbb{R}, \delta > 0.
\]

**Lemma 3.3.1.** There exist constants \( a_j, b_j, j = 0, 1, 2 \) such that the time derivative \( \partial_t F \) of the Lyapunov function along the semigroup is negative.

**Proof.** We will calculate the time derivative of each term in (3.3.1) separately and make use of the explicit relations (3.2.3):

\[
\begin{align*}
\partial_t \|u\|^2 &= -2(\mathcal{L}u, u) = -2\|Au\|^2, \\
\partial_t (Au, Au) &= -2(Cu, Au) - 2\|A^*Au\|^2 = -2(Cu, Au) - 2\|Au\|^2 - 2\|A^2u\|^2, \\
\partial_t (Cu, Cu) &= -2\|AC^2u\|^2 - 2(C_2u, Cu), \\
\partial_t (C_2u, C_2u) &= (2 + \delta^2 V)C_2u, Cu) - 2\|AC_2u\|^2 + 2(Au, C_2u), \\
\partial_t (Au, Cu) &= -2(A^2u, ACu) - (Au, Cu) - \|Cu\|^2 - (Au, C_2u), \\
\partial_t (Cu, C_2u) &= -\|C_2u\|^2 - 2(ACu, C_2u) + 2\|Cu\|^2 + (Cu, Au).
\end{align*}
\]

Putting everything together we obtain

\[
\partial_t F(t) = -2a_0 t \|A^2u\|^2 - 2a_1 t^3 \|ACu\|^2 - 2a_2 t^5 \|AC_2u\|^2 \\
-2b_0 t^2 (A^2u, ACu) - 2b_1 t^4 (ACu, AC_2u) \\
+(-2a_0 t + a_0 - 2b_2)\|Au\|^2 + (3a_1 t^2 + 2b_1 t^4 - b_0 t^2)\|Cu\|^2 \\
+(5a_2 t^4 - b_1 t^4)\|C_2u\|^2 + (2b_0 t - 2a_0 t - b_0 t^2 + b_1 t^4)(Au, Cu) \\
+(4b_1 t^3 - 2a_1 t^3 + 2a_2 t^5)(Cu, C_2u) + (2a_2 t^5 - b_0 t^2)(Au, C_2u).
\]

Now we estimate the sum of the first and of the second line (i.e. the sum of all the terms where \( A^2, AC \) and \( AC_2 \) appear). For \( t \in (0, 1] \) we have

\[
(3.3.3a) + (3.3.3b) \leq -2a_0 t \|A^2u\|^2 + 2b_0 t^2 \|A^2u\| \|ACu\|
\]
Choosing the constants in such a way that \(a \gg b_0 \gg b \gg a_1 \gg b_1 \gg a_2 > 1/c\), where \(c\) is a constant depending on the bound on the second derivative of the potential, we obtain that \(\partial_t F < 0\) \(\forall t \in (0, 1]\).

**Proof of Theorem 3.1.6.** We use the previous Lemma to deduce

\[
a_0 t \|Au\|^2 + a_1 t^3 \|Cu\|^2 + a_2 t^5 \|C_2 u\|^2 + b_0 t^2 (Au, Cu) + t^4 b_1 (Cu, C_2 u) + b_2 \|u\|^2 < b_2 \|u_0\|^2. \tag{3.3.4}
\]

This, in turn, implies that

\[
\|\nabla_z u\|^2 = \|Au\|^2 < \frac{\kappa}{t} \|u_0\|^2, \\
\|\nabla_p u\|^2 = \|Cu\|^2 < \frac{\kappa}{t^3} \|u_0\|^2, \\
\|\nabla_q u\|^2 - \frac{\|\nabla_z u\|^2}{3} \leq \|\nabla_q u - \nabla_z u\|^2 = \|C_2 u\|^2 < \frac{\kappa}{t^5} \|u_0\|^2 \\
\Rightarrow \|\nabla_q u\|^2 \leq \frac{\kappa}{t^5} \|u_0\|^2,
\]

\[\text{Here and in the following by } a \gg b \text{ we mean that } a \text{ is bigger then (multiples of) } b \text{ and } b^2.\]
where \( \kappa \) is an explicitly computable positive constant. The previous inequalities are justified by the fact that

\[
\begin{align*}
& a_0 t \| Au \|^2 + a_1 t^3 \| Cu \|^2 + a_2 t^5 \| C_2 u \|^2 + b_0 t^2 (Au, Cu) + t^4 b_1 (Cu, C_2 u) \\
& \geq (a_0 t - b_0 t^2) \| Au \|^2 + (a_1 t^3 - t^3 b_0^2 \frac{t^2}{2} - t^3 b_1^2) \| Cu \|^2 + (a_2 t^5 - t^5 \frac{t^2}{2}) \| C_2 u \|^2
\end{align*}
\]

and the second line is positive thanks to the choice of the constants we made.

Remark 3.2.4 holds also in this case.

Remark 3.3.2. From the estimates (3.1.3), similar estimates on \( A^*e^{-tL^*}, e^{-tL^*} A^*, C^*e^{-tL^*}, e^{-tL^*} C^*, C^*e^{-tL^*} \) and \( e^{-tL^*} C_2^* \) follow, where \( \ast \) and \( \bullet \) stand for either the \( L^2_{\rho} \)-adjoint or nothing. In fact:

(i) \( (Ae^{-tL} f, g) = (f, e^{-tL^*} A^* g) \leq \| Ae^{-tL} f \| \| g \| \leq \frac{\kappa}{\sqrt{t}} \| f \| \| g \| \)

\[ \Rightarrow (f, e^{-tL^*} A^* g) \leq \frac{\kappa}{\sqrt{t}} \| f \| \| g \|, \]

choose \( f = e^{-tL^*} A^* g \) and the result on \( e^{-tL^*} A^* \) follows.

(ii) Using \( [A, A^*] = I \) we have \( \| A^*e^{-tL} u_0 \|^2 = \| A^* u \|^2 = \| Au \|^2 + \| u \|^2 \), hence the estimate for \( A^*e^{-tL} \). Taking the adjoint as in (i) we get the result for \( e^{-tL^*} A \).

(iii) For \( Ae^{-tL^*} \) we can just repeat the proof we wrote for \( Ae^{-tL} \), since the only thing that changes when considering \( L^* \) is the sign of \( B \), which doesn’t play any role in the proof.

Now, by acting as in (i) and (ii), we obtain the results for \( e^{-tL} A^*, A^*e^{-tL^*} \) and \( e^{-tL} A \).

### 3.4 The Homogenization Theorem

In this section we prove Theorem 3.1.8. The proof of this theorem is based on standard techniques, namely the central limit theorem for additive functionals of Markov processes \([37, 45, 63]\), which in turn is based on the martingale central limit theorem. In order to apply these techniques we need to study the Poisson equation

\[
Lu = f. \tag{3.4.1}
\]

The boundary conditions for (3.4.1) are that \( u \in L^2_{\rho} \) and \( u \) is periodic in \( q \).

**Proposition 3.4.1.** Let \( f \in L^2_{\rho} \cap C^\infty(X) \) with \( \int_X f \mu_\beta(dx) = 0 \). Then the Poisson equation (3.4.1) has a unique smooth mean zero solution \( u \in L^2_{\rho} \cap C^\infty(X) \).
The proof of Theorem 3.1.8 follows now from the above proposition and from the Martingale Central Limit theorem, which we have recast, for the reader convenience, in Appendix A.2. More specifically we will make use of Corollary A.2.3.

**Proof of Theorem 3.1.8.** To simplify the notation we present the proof for $d = 1$. When $d > 1$ the same proof applies to the one-dimensional projections $q^e := q \cdot e$. In this case the diffusion coefficient $D$ is replaced by the projections of the diffusion tensor $D^e := De \cdot e$.

We consider the process $x(t)$ on $X$ with stationary initial conditions. Since $p \in L^2_\rho \cap C^\infty(X)$ and is centered with respect to the invariant measure $\mu$, Proposition 3.4.1 applies and there exists a unique mean zero solution $\phi \in L^2_\rho \cap C^\infty(X)$ to the problem

$$L\phi = p.$$ (3.4.2)

Recalling that $-L$ is the generator of the process, we use Itô’s formula to obtain

$$d\phi = -L\phi dt + \sum_{j=1}^{m} \sqrt{2\alpha_j \beta^{-1} \partial_{z_j} \phi} dW_j.$$ (3.4.3)

We combine this, together with (3.4.2) and the equations of motion to deduce

$$q^e(t) := eq(t/\epsilon^2)$$

$$= eq(0) + \epsilon \int_0^{t/\epsilon^2} p(s) \, ds$$

$$= eq(0) - \epsilon [\phi(q(t/\epsilon^2), p(t/\epsilon^2), z(t/\epsilon^2)) - \phi(q(0), p(0), z(0))]$$

$$+ \epsilon \int_0^{t/\epsilon^2} \sqrt{2\alpha_j \beta^{-1} \partial_{z_j} \phi} dW_j(s)$$

$$= \epsilon R^e + M^e,$$

where

$$R^e := q(0) - \left[ \phi(q(t/\epsilon^2), p(t/\epsilon^2), z(t/\epsilon^2)) - \phi(q(0), p(0), z(0)) \right],$$

$$M(t) := \sum_{j=1}^{m} \int_0^{t} \sqrt{2\alpha_j \beta^{-1} \partial_{z_j} \phi} dW_j(s),$$

so that $M^e(t) := \epsilon M(t/\epsilon^2)$. Our stationarity assumption, together with the fact that $\phi \in L^2_\rho$, imply that

$$\mathbb{E}|R^e|^2 \leq C.$$ (3.4.4)

To study the martingale term $M^e$ we use the above mentioned Martingale Central Limit theorem. We have that $M(0) = 0$, $M(t)$ has continuous sample paths and stationary increments.
3.4. The Homogenization Theorem

We will get to the boundedness of \( \phi \) in a moment; it will follow from the estimate (3.4.6) below, which gives \( \| \partial z_i \phi \| \leq C \). Furthermore, by ergodicity we deduce that

\[
\lim_{t \to \infty} \langle M \rangle_t / t = 2 \sum_{i=1}^{m} \alpha_i \beta^{-1} \| \partial z_i \phi \|^2 \quad \text{a.s.,}
\]

(3.4.4)

where \( \langle M \rangle_t \) denotes the quadratic variation of the martingale \( M \). (3.4.3) and (3.4.4) imply that the rescaled process \( q_\epsilon(t) := \epsilon q(t/\epsilon^2) \) converges weakly in \( C([0,t];\mathbb{R}) \) to a Brownian motion \( \sqrt{2D} W(t) \) where

\[
D = 2 \beta^{-1} \sum_{i=1}^{m} \alpha_i \| \partial z_i \phi \|^2.
\]

(3.4.5)

**Remark 3.4.2.** Notice that when \( d > 1 \) the convergence of the one dimensional projections \( q_\epsilon^i(t) := \epsilon q(t/\epsilon^2) \) does not imply the convergence of the process \( q_\epsilon(t) = \epsilon q(t/\epsilon^2) \) [24, Rem. 2.3]. The proof of the homogenization theorem in the multidimensional case, which is also based on the analysis of the Poisson equation, is very similar and it is omitted. Similar results for diffusion processes with periodic coefficients in arbitrary dimensions can be found in e.g. [7].

To prove estimate (3.1.7), we first show the upper bound and then the fact that the diffusion coefficient is bounded away from zero. We set \( \phi = g_i + \frac{1}{\lambda_i} z_i \) and use the Poisson equation (3.4.2) to obtain

\[
\mathcal{L} g_i = -\frac{\alpha_i}{\lambda_i} z_i,
\]

from which we obtain the estimate

\[
\alpha_i \beta^{-1} \| \partial z_i g_i \|^2 \leq \sum_{j=1}^{m} \alpha_j \beta^{-1} \| \partial z_j g_i \|^2 = (\mathcal{L} g_i, g_i)
\]

\[
= \frac{\alpha_i}{\beta \lambda_i} \int g_i \partial z_i \rho \, dx = -\frac{\alpha_i}{\beta \lambda_i} \int \rho \partial z_i g_i \, dx
\]

\[
\leq \frac{\alpha_i}{\beta \lambda_i} \| \partial z_i g_i \|.
\]

Consequently,

\[
\| \partial z_i g_i \| \leq \frac{1}{\lambda_i}.
\]

(3.4.6)

From this we obtain the following estimate on the diffusion coefficient \( D \)

\[
D = \sum_{i=1}^{m} \alpha_i \beta^{-1} \| \partial z_i \phi \|^2 = \frac{1}{\beta} \sum_{i=1}^{m} \alpha_i \left( \| \partial z_i g_i \| + \frac{1}{\lambda_i^2} \right)^2
\]

\[
\leq \frac{2}{\beta} \sum_{i=1}^{m} \alpha_i \left( \| \partial z_i g_i \|^2 + \frac{1}{\lambda_i^2} \right)
\]

\[
\leq \frac{4}{\beta} \sum_{i=1}^{m} \frac{\alpha_i}{\lambda_i^2}.
\]
The fact that $D > 0$ is easily seen by contradiction. Assume that $D = 0$. Then by (3.4.5),

$\|\partial z_i\phi\|^2 = 0 \forall i = 1 \ldots m$. Hence $\phi = \phi(q, p)$ and

$$L\phi = -p\partial_q \phi + \partial_q V \partial_p \phi + \sum_{i=1}^{m} \lambda_i z_i \partial_p \phi = p.$$ 

Multiplying both sides by $z_i e^{-z_i^2/2}$ and then integrating with respect to $z_i$ we get

$$-\int p \partial_q \phi z_i e^{-z_i^2/2} dz_i + \int \partial_q V \partial_p \phi z_i e^{-z_i^2/2} dz_i$$

$$+ \int \lambda_i \partial_p \phi z_i^2 e^{-z_i^2/2} dz_i + \sum_{j \neq i} \int \lambda_i z_i z_j \partial_p \phi e^{-z_i^2/2} dz_i$$

$$= \int p z_i e^{-z_i^2/2} dz_i,$$

from which we conclude that $\lambda_i \partial_p \phi = 0$ for all $i = 1 \ldots m$. Hence $\phi = \phi(q)$. By the same reasoning we get that $-p\partial_q \phi = p$, which does not have a periodic solution. 

We now prove Proposition 3.1.9

**Proof of Proposition 3.1.9.** We will use [15, Corollary 4.2], which we briefly restate here for the reader’s convenience (adapting the statement to our context and notation). Let $K$ be an operator on $L^2$ of the form

$$K = \sum_{i=1}^{r} X_i' X_i + X_0,$$

where we recall that $X_i'$ is the $L^2$ adjoint of $X_i$ and the $X_i$ are differential operators with $C^\infty$ coefficients for any $i \geq 0$. Suppose $K \in K_1$, where the class $K_1$ is defined in [15, Definition 2.2] and let

$$\Lambda^2 = 1 + p^2 + z^2.$$ 

(3.4.8)

If there exist constants $c, \epsilon > 0$ such that

$$\|\Lambda f\| \leq c(\|f\| + \|Kf\|)$$

(3.4.9)

then $K$ has compact resolvent on $L^2$. In the above inequality and for the rest of this proof $\|\cdot\|$ denotes the $L^2$ norm and $c$ will be a generic constant.

We intend to apply this statement to the generator (3.0.3); however the operator (3.0.3) acts on $L^2_p(T^d \times \mathbb{R}^d \times \mathbb{R}^{md})$. For the rest of the proof we set $d = m = 1$. In order to go from the weighted to the flat $L^2$ we apply the following unitary transformation

$$H = -\sqrt{p}L \left(\sqrt{p^{-1}} \cdot \right),$$

(3.4.10)

We will give more details about this transformation at the beginning of Chapter 5.
and we then reduce ourselves to the study of the operator \( K := H + 1/2 \), i.e.

\[
K = -p \partial_q + \partial_q V \partial_p - z \partial_p + p \partial_z + \frac{z^2}{4} - \partial_z^2.
\]

Therefore in our case \( r = 2 \) and

\[
X_1 = \partial_z, \quad X_2 = \frac{z}{2},
\]

\[
X_0 = -p \partial_q + \partial_q V \partial_p - z \partial_p + p \partial_z, \quad X'_0 = -X_0.
\]

\( K \) belongs to the class \( \mathcal{K}_1 \), so we need to prove that \( K \) satisfies the inequality (3.4.9), with \( \Lambda \) defined as in (3.4.8). To this end

\[
\| \Lambda f \| = (f, \Lambda^2 f) = (f, \Lambda^{2-2} \Lambda^2 f)
\]

so all we need is an estimate on \( \| \Lambda^{-1} p \| \) and \( \| \Lambda^{-1} z \| \) in terms of a constant multiple of \( (\| f \| + \| K f \|) := \mathfrak{B} \). \(^8\) \( z^2 \), as well as \( \partial_z^2 \) are in the symmetric part of the operator \( K \), which we will denote \( Sym(K) \). Therefore, being \( K = Sym(K) + X_0 \) with \( X_0 \) antisymmetric, we have

\[
\| z f \|^2 = (z^2 f, f) \leq |(Sym(K)f, f)| \leq |(K f, f)| \leq c \mathfrak{B}^2.
\]

Therefore

\[
\| z f \| \leq c \mathfrak{B} \tag{3.4.11}
\]

and, for the same reason

\[
\| \partial_z f \| \leq c \mathfrak{B}. \tag{3.4.12}
\]

Also, for any \( \beta_1 \geq 0 \),

\[
\| \Lambda^{-\beta_1} z f \| \leq c \mathfrak{B} \tag{3.4.13}
\]

and for any \( \alpha_1, \beta_2 \geq 1 \),

\[
\| \Lambda^{-\alpha_1} p f \| \leq c \| f \| \quad \text{and} \quad \| \Lambda^{-\beta_2} z f \| \leq c \| f \|. \tag{3.4.14}
\]

We also would like to stress, as we will be using it in the following, that for any \( \alpha > 0 \)

\[
D \Lambda^{-\alpha}, \; D^2 \Lambda^{-\alpha}, \; z \partial_z \Lambda^{-\alpha}, \; p \partial_z \Lambda^{-\alpha} \tag{3.4.15}
\]

\(^8\)We don’t need to take care of the q-variable because \( q \in \mathbb{T} \). This is also why q doesn’t appear in the definition of \( \Lambda^2 \).
are bounded functions, where $D$ denotes either $\partial_p$ or $\partial_z$. In other words the first and second derivatives of $\Lambda^{-\alpha}$ are bounded functions, along with $z \partial_z \Lambda^{-\alpha}$ and $p \partial_z \Lambda^{-\alpha}$. Now we are left with the more complicated part: knowing (3.4.11)-(3.4.15), find an estimate on $\|\Lambda^{\epsilon-1}f\|$ for some $0 \leq \epsilon < 1$. If we observe that $[X_0, z] = p$, we have

$$\|\Lambda^{\epsilon-1}f\|^2 = (\Lambda^{\epsilon-1}f, \Lambda^{\epsilon-1}pf) = ([X_0, z]f, \Lambda^{2\epsilon-2}pf)$$

and it all boils down to proving an inequality of the type

$$([X_0, z]f, \Lambda^{-\gamma}pf) \leq c\mathfrak{B}^2$$

(3.4.16) for some $0 < \gamma < 2$. Therefore the compactness of the resolvent of $K$ is shown once we prove (3.4.16). Let us write

$$([X_0, z]f, \Lambda^{-\gamma}pf) \leq |([X_0, z]f, \Lambda^{-\gamma}pf)| + |(z X_0 f, \Lambda^{-\gamma}pf)| := [1] + [2]$$

and start with estimating the term [2]. If we express $X_0 = K - \text{Sym}(K)$, we get

$$[2] \leq |(z K f, \Lambda^{-\gamma}pf)| + |(z \text{Sym}(K) f, \Lambda^{-\gamma}pf)| := [21] + [22].$$

The term [21] is easy to estimate, indeed

$$[21] = |(\Lambda^{-\gamma+1}z f, \Lambda^{-1}pf)| \leq c\mathfrak{B}^2$$

if $\gamma \geq 1$ (having used (3.4.13) and (3.4.14)). As for [22], for $\gamma_1$ and $\gamma_2$ such that $\gamma_1 + \gamma_2 = \gamma$, we further split it into

$$[22] = (\Lambda^{-\gamma_1} f, \Lambda^{-\gamma_2} pf) \leq |(\text{Sym}(K) \Lambda^{-\gamma_1}z f, \Lambda^{-\gamma_2}pf)| + |([\Lambda^{-\gamma_1}z, \text{Sym}(K)]f, \Lambda^{-\gamma_2}pf)| := [221] + [222].$$

We expand [222] into

$$[222] \leq |(\Lambda^{-\gamma_1}[z, \text{Sym}(K)] f, \Lambda^{-\gamma_2}pf)| + |([\Lambda^{-\gamma_1}, \text{Sym}(K)] z f, \Lambda^{-\gamma_2}pf)| := [2221] + [2222].$$

Once we calculate the involved commutators, [2221] and [2222] are easily estimated, in fact

$$[z, \text{Sym}(K)] = [\partial_z^2, z] = 2\partial_z$$

so if $\gamma \geq 1$, using (3.4.12) and (3.4.14),

$$[2221] = 2 |(\partial_z f, \Lambda^{-\gamma}pf)| \leq c\mathfrak{B}^2.$$
Also,

\[ [\Lambda^{-\gamma_1}, Sym(K)] = [\Lambda^{-\gamma_1}, \partial_z^2] = - (\partial_z^2 \Lambda^{-\gamma_1}) - 2(\partial_z \Lambda^{-\gamma_1}) \partial_z \]

and therefore, using (3.4.11), (3.4.12), (3.4.14) and (3.4.15), we can bound \( 222 \) as follows

\[
222 \leq \left| (z f, (\partial_z^2 \Lambda^{-\gamma_1}) \Lambda^{-\gamma_2} pf) \right| + 2 \left| ((\partial_z \Lambda^{-\gamma_1}) f, \Lambda^{-\gamma_2} pf) \right| + 2 \left| (\partial_z f, z(\partial_z \Lambda^{-\gamma_1}) \Lambda^{-\gamma_2} pf) \right| \leq cB^2,
\]

if \( \gamma_2 \geq 1 \). This concludes the bound for the term \( 222 \). As for \( 221 \), by the Cauchy-Schwartz inequality we get

\[
221 = \left| (\text{Sym}(K)^{1/2} \Lambda^{-\gamma_1} zf, \text{Sym}(K)^{1/2} \Lambda^{-\gamma_2} pf) \right|
\leq \left| (\text{Sym}(K) \Lambda^{-\gamma_1} zf, \Lambda^{-\gamma_1} zf) \right|^{1/2} \left( \| \Lambda^{-\gamma_2} pf \| + \| \partial_z (\Lambda^{-\gamma_2} pf) \| \right)^{1/2}
\leq cB^2 \left( 2211 + 2213 \right)^{1/2}.
\]

Let us study these terms separately:

\[
2211 \leq \left| (K \Lambda^{-\gamma_1} zf, \Lambda^{-\gamma_1} zf) \right|
\leq \left| (\Lambda^{-\gamma_1} zf K f, \Lambda^{-\gamma_1} zf) \right| + \left| ([\Lambda^{-\gamma_1} z, K] f, \Lambda^{-\gamma_1} zf) \right|
\leq \left| (\Lambda^{-2\gamma_1} zf K f, zf) \right| + \left| ([\Lambda^{-\gamma_1} z, X_0 f, \Lambda^{-\gamma_1} zf) \right| + \left| ([\Lambda^{-\gamma_1} z, \text{Sym}(K)] f, \Lambda^{-\gamma_1} zf) \right|
\leq cB^2 + \left( 2211_1 + 2211_2 \right)
\]

where the last inequality holds if \( \gamma_1 \geq 1/2 \) (as we need to use (3.4.14)). Now,

\[
[\Lambda^{-\gamma_1} z, \text{Sym}(K)] = [\Lambda^{-\gamma_1} z, \partial_z^2]
= \Lambda^{-\gamma_1} [z, \partial_z^2] + [\Lambda^{-\gamma_1}, \partial_z^2] z
= -2\Lambda^{-\gamma_1} \partial_z - (\partial_z^2 \Lambda^{-\gamma_1}) z - (\partial_z \Lambda^{-\gamma_1}) z \partial_z - (\partial_z \Lambda^{-\gamma_1}),
\]

so again \( 2211_2 \leq cB^2 \) because of (3.4.15), for any \( \gamma_1 > 0 \). As for \( 2211_1 \),

\[
[\Lambda^{-\gamma_1} z, X_0] = \Lambda^{-\gamma_1} p + [\Lambda^{-\gamma_1}, X_0] z.
\]

As a consequence of (3.4.15) and recalling that \( q \in T, [X_0, \Lambda^{-\gamma_1}] = -\gamma_1 \Lambda^{-\gamma_1 - 1} X_0(\Lambda) \) is bounded for any \( \gamma_1 > 0 \), so

\[
2211_1 \leq \left| ([\Lambda^{-\gamma_1} pf, zf) \right| + \left| ([\Lambda^{-\gamma_1} X_0] zf, \Lambda^{-\gamma_1} zf) \right| \leq cB^2
\]
if $\gamma_1 \geq 1/2$. Now $2212$ and $2213$ are easy to bound:

$2212 = \|\Lambda^{-\gamma_2} pzf\| \leq c\|zf\| \leq c\mathfrak{B}^2$

if $\gamma_2 \geq 1$. Also,

$2213 \leq \|\partial_z \Lambda^{-\gamma_2} pf\| + \|\Lambda^{-\gamma_2} p\partial_z f\| \leq c\mathfrak{B}^2$

under the same condition on $\gamma_2$ as above. Putting everything together we have shown that $2 \leq c\mathfrak{B}^2$. We are left with showing that the same bound holds for $\gamma = 3/2$.

$$1 = (zf, X_0 \Lambda^{-\gamma} pf)$$

$$= (zf, \Lambda^{-\gamma} pX_0 f) + (zf, [X_0, \Lambda^{-\gamma} pf])$$

$$\leq |(zf, \Lambda^{-\gamma} pX_0 f)| + |(zf, \Lambda^{-\gamma} qf)| + |(zf, [X_0, \Lambda^{-\gamma} pf])|$$


If $\gamma \geq 1$ then $[X_0, \Lambda^{-\gamma}]p$ is bounded. Therefore $12$ and $13$ are easily estimated by a constant multiple of $\mathfrak{B}^2$.

$$11 \leq |(zf, \Lambda^{-\gamma} pKf)| + |(zf, \Lambda^{-\gamma} p Sym(K)f)| \leq c\mathfrak{B}^2 + 112$$

if $\gamma \geq 1$. Now, again for $\gamma_1, \gamma_2$ such that $\gamma_1 + \gamma_2 = \gamma$,

$$112 = |(\Lambda^{-\gamma_1} zf, \Lambda^{-\gamma_2} pSym(K)f)|$$

$$\leq |(\Lambda^{-\gamma_1} zf, Sym(K)\Lambda^{-\gamma_2} pf)| + |(\Lambda^{-\gamma_1} zf, [Sym(K), \Lambda^{-\gamma_2} pf])|.$$  

The first addend of $112$ is equal to $221$, whereas the second can be treated analogously to $222$, hence we won’t repeat it. To conclude, if we choose $\gamma_1 = 1/2$ and $\gamma_2 = 1$, (3.4.16) holds for $\gamma = 3/2$ and therefore (3.4.9) holds with $\epsilon = 1/4$. This concludes the proof.  

**Proof of Proposition 3.4.1.** This is a consequence of the compactness of the resolvent of $\mathcal{L}$, which allows us to use Fredholm’s Theorem (see Appendix A.3). Recall that we are considering the Poisson equation $\mathcal{L}\phi = f$ where $f \in L_\mu \cap C^\infty(X)$ and centered with respect to the invariant measure $\mu_\beta(dx)$.

Set $\mathcal{L}_\gamma u = \gamma u + \mathcal{L}u$. Fredholm’s Theorem applies, so either the solution of

$$\left(\frac{1}{\gamma} I - \mathcal{L}_\gamma^{-1}\right) u = \tilde{h}, \quad \tilde{h} = \mathcal{L}_\gamma^{-1} f / \gamma$$

exists and is unique (and hence, by construction the solution to (3.4.1) is unique) or $\left(\frac{1}{\gamma} I - \mathcal{L}_\gamma^{-1}\right) u = 0$ admits a nonzero solution. We can rule out the latter option because $\left(\frac{1}{\gamma} I - \mathcal{L}_\gamma^{-1}\right) u = 0$ is
equivalent to $Lu = 0$; since we know that $\text{Ker} L$ contains only constants and we require the solution to have mean zero, we can conclude that the only solution of the equation $Lu = 0$ is $u = 0$.

### 3.5 The White Noise Limit

Throughout this section $C$ denotes a generic constant and $c(t)$ denotes a generic positive increasing continuous function bounded on compacts $[0, T]$; both $C$ and $c(t)$ are independent of $\epsilon$, even though they can depend on the coefficients $\{\lambda_i, \alpha_i\}_{i=1,\ldots,m}$ and they do depend on the exponent $n$ in estimate (3.1.14). To simplify the notation we present the proof in one dimension, i.e. $d = 1$ and we set $\beta = 1$. The proof is exactly the same in arbitrary dimensions.

Let $(Q(t), P(t)) \in \mathbb{R} \times \mathbb{R}$ be the solution to the system (3.1.13), and $(q(t), p(t), z(t))$ be the solution to the system (3.1.12), then

$$|q(t) - Q(t)| \leq \int_0^t |p(s) - P(s)| \, ds.$$  

From (3.1.12c)

$$\frac{1}{\sqrt{\epsilon}} \int_0^t ds z_i(s) = -\sqrt{\epsilon} \left( z_i(t) - z_i(0) \right) - \frac{\lambda_i}{\alpha_i} \int_0^t ds p(s) + \sqrt{\frac{2}{\alpha_i}} W_i(t),$$

so that, setting $\theta_i = \frac{\lambda_i^2}{\alpha_i}$, we have

$$p(t) - P(t) = \int_0^t \left( -\partial_q V(q(s)) + \partial_q V(Q(s)) \right) ds + \sum_{i=1}^m \theta_i \int_0^t \left( P(s) - p(s) \right) ds$$

$$-\sqrt{\epsilon} \sum_{i=1}^m \frac{\lambda_i}{\alpha_i} \left( z_i(t) - z_i(0) \right).$$

We use the Lipshitz continuity of $\partial_q V(q)$ together with Hölder’s inequality to obtain

$$\eta_n(T) := E \sup_{t \in [0, T]} \{ |q(t) - Q(t)|^n + |p(t) - P(t)|^n \}$$

$$\leq C T^{n-1} \int_0^T E \sup_{s \in [0, t]} |q(s) - Q(s)|^n \, dt$$

$$+ C \left( \sum_{i=1}^m \theta_i^n \right) T^{n-1} \int_0^T E \sup_{s \in [0, t]} |p(s) - P(s)|^n \, dt$$

$$+ C \epsilon^{\frac{n}{2}} \sum_{i=1}^m \left( \frac{\lambda_i}{\alpha_i} \right)^n E \sup_{t \in [0, T]} |z_i(t) - z_i(0)|^n.$$
From this we deduce
\[ \eta_n(T) \leq Cc(T) \int_0^T dt \eta_n(t) + C \varepsilon^2 \sum_{i=1}^m E \sup_{t \in [0,T]} |z_i(t) - z_i(0)|^n. \]

From Gronwall’s Lemma \(^9\) we then have
\[ \eta(T) \leq C \varepsilon^2 \sum_{i=1}^m E \sup_{t \in [0,T]} |z_i(t) - z_i(0)|^n + Cc(T)\varepsilon^2 \int_0^T dt \sum_{i=1}^m E \sup_{s \in [0,t]} |z_i(s) - z_i(0)|^n \]
and the result now follows from Proposition 3.5.1 below.

**Proposition 3.5.1.** With the same notation and assumptions as in Theorem 3.1.10 the following estimate holds true:
\[ \sum_{i=1}^m E \sup_{t \in [0,T]} |z_i(t) - z_i(0)|^n \leq Cc(T) \left[ \sum_{i=1}^m E[z_i(0)]^n + E[p(0)]^n + E[q(0)]^n + 1 \right], \]
where \(c(t)\) is a positive increasing continuous function bounded on compacts [0, T].

**Proof.** From (3.1.12c),
\[ z_i(t) = e^{-\alpha t}z_i(0) + \int_0^t e^{-(t-s)\frac{\alpha}{\varepsilon}} \left( -\frac{\lambda_i}{\sqrt{\varepsilon}}p(s)ds + \frac{2 \alpha_i}{2 \sqrt{\varepsilon}} dW_i(s) \right). \]  

(3.5.1)

So from (3.1.12a), (3.1.12b) and (3.5.1) we have
\[ q(t) + p(t) = -\int_0^t ds \partial_q V(q(s)) + \int_0^t ds p(s) + q(0) + p(0) \]
\[ + \frac{1}{\sqrt{\varepsilon}} \sum_{i=1}^m \lambda_i \left[ \int_0^t ds e^{-\alpha t} z_i(0) + \frac{1}{\sqrt{\varepsilon}} \int_0^t ds e^{-\alpha t} \int_0^s ds e^{-\alpha t} (-\lambda_i p(u)du + \sqrt{2 \alpha_i} dW_i(u)) \right]. \]

By integration by parts,
\[ \int_0^t ds e^{-\alpha t} \int_0^s ds e^{-\alpha t} (-\lambda_i p(u)du + \sqrt{2 \alpha_i} dW_i) \]
\[ = \frac{1}{\alpha_i} \int_0^t (e^{-(t-u)\frac{\alpha}{\varepsilon}} + 1) (-\lambda_i p(u)du + \sqrt{2 \alpha_i} dW_i(u)), \]
hence, using again the Hölder continuity of \(V(q)\), we obtain
\[ \xi_n(T) := E \sup_{t \in [0,T]} \{ |q(t)|^n + |p(t)|^n \} \]
\[ = Cc(T) \left[ \int_0^T dt \xi_n(t) + E (|q(0)|^n + |p(0)|^n) + C \varepsilon^2 \sum_{i=1}^m E[z_i(0)]^n + 1 \right]. \]

---

\(^9\)We apply Gronwall’s Lemma in the following form: suppose \(u(t) \leq a(t) + b(t) \int_0^t u(s)ds\). Then \(u(t) \leq a(t) + b(t) \int_0^t u(s)exp(\int_0^t b(r)dr).\)
and by Gronwall’s Lemma

\[
\xi_n(T) \leq C \left[ E \left( |q(0)|^n + |p(0)|^n \right) + C \epsilon^{\frac{2}{n}} \sum_{i=1}^{m} E |z_i(0)|^n \right] (1 + c(T)),
\]

which implies

\[
E \sup_{t \in [0,T]} |p(t)|^n \leq \left[ E \left( |q(0)|^n + |p(0)|^n \right) + C \epsilon^{\frac{2}{n}} \sum_{i=1}^{m} E |z_i(0)|^n \right] (1 + c(T)). \tag{3.5.2}
\]

Since by (3.5.1) we have

\[
E \sup_{t \in [0,T]} |z_i(t)|^n \leq c(T) \left( E |z_i(0)|^n + E \sup_{t \in [0,T]} |p(t)|^n + 1 \right), \tag{3.5.3}
\]

Proposition 3.5.1 follows from (3.5.2) and (3.5.3). \qed
Chapter 4

Markovian Approximation of Open Classical Systems

Consider the system

\[ \begin{align*}
 dq & = p \, dt \\
 dp & = -\partial_q V(q) \, dt + g \cdot u \, dt \\
 du & = (-p g - A u) \, dt + C \, dW(t),
\end{align*} \]

where \((q, p) \in \mathbb{R}^2, u \) and \( g \) are column vectors of \( \mathbb{R}^d \), \( \cdot \) denotes Euclidean scalar product, \( W(t) = (W_1(t), \ldots, W_d(t)) \) is a \( d \)-dimensional Brownian motion, \( V(q) \) is a potential and \( A \) and \( C \) are constant coefficients \( d \times d \) matrices, related through the fluctuation dissipation principle, which in the present case reads

\[ A + A^T = C C^T. \]

 Remark 4.0.2. Since \( \Sigma := C C^T \) is a semipositive definite symmetric matrix and we denote \( m = \text{Rank} \, \Sigma \). For the vector \( g \), we shall always assume that \( g \neq 0 \) (to avoid the uninteresting case in which there is no coupling between the heat bath and the particle). In the following, we shall denote \( x = (q, p, u) \in \mathbb{R}^N, N = d + 2 \). The generator of (4.0.1) is

\[ L = p \partial_q - \partial_q V \partial_p + \sum_{i=1}^d g_i u_i \partial_p - p \sum_{i=1}^d g_i \partial u_i - \sum_{i,j=1}^d a_{ij} u_j \partial u_i + \frac{1}{2} \sum_{i,j=1}^d \Sigma_{ij} \partial^2 u_i u_j. \]

 Remark 4.0.2. Since \( \Sigma \) is symmetric, there exists an orthogonal matrix \( T \) such that \( T^{-1} \Sigma T := \Delta \) is diagonal with \( \Delta = \text{diag}\{\lambda_1, \ldots, \lambda_m, 0, \ldots, 0\} \), \( \lambda_j > 0, j = 1 \ldots, m \). As noticed in [39],
4.1. Ergodicity of the general approximation

setting \( u = Tv, \ T^{-1}g = \tilde{g} \) and \( T^{-1}AT := \tilde{A} \), we have that system \((4.0.1)\) is equivalent to the following system \((4.0.4)\):

\[
\begin{align*}
 dq &= p \, dt \tag{4.0.4a} \\
 dp &= -\partial_q V(q) \, dt + \tilde{g} \cdot v \, dt \tag{4.0.4b} \\
 dv &= (-p \tilde{g} - \tilde{A}v) \, dt + T^{-1}C \, dW(t). \tag{4.0.4c}
\end{align*}
\]

The fluctuation-dissipation theorem for the above system is

\[ \tilde{A} + \tilde{A}^T = \Delta. \]

Under the assumption of bounded Hessian potential, we can repeat the analysis of system \((3.0.2)\) done in the previous chapter, for this more general approximation of the Generalized Langevin Equation. In particular, all the results about exponential convergence to equilibrium obtained through the theory of hypocoercivity still hold under some requirements on the coefficients of the matrix \(A\) and of the vector \(g\), which we shall make explicit in the next section.

What we want to focus on is the nice property of system \((4.0.4)\), of being ergodic under the sole assumption of hypoellipticity. We shall show this property in the next Section 4.1. We believe that the same result can be produced for a general O-U process. Section 4.2 deals with the largely open problem of estimating how close the solution of the approximating system is to the solution of the non-Markovian GLE; we shall mainly discuss the case in which the matrix \(A\) is diagonal. Section 4.3 gives the well-posedness, in an appropriate framework, of system \((4.0.4)\) when \(A\) is diagonal and \(d \to \infty\).

### 4.1 Ergodicity of the general approximation

Once we fix \(C\), there are many matrices \(A\) such that the fluctuation-dissipation relation \((4.0.2)\) is satisfied. Also, as we will see below, among such matrices there are still many such that the generator of \((4.0.1)\) is hypoelliptic. For a given \(C\), let us call \(A_C\) the set of matrices such that the fluctuation relation is satisfied and the process is hypoelliptic. Theorem 4.1.1 says that if \(\det C \neq 0\), then for any \(A \in A_C\) the process is ergodic; if \(\det C = 0\) then there exists (possibly more than one) \(A \in A_C\) such that the process is ergodic.

**Theorem 4.1.1.** Assume \(V(q) = q^2 / 2\) and \(\tilde{g} \neq 0\) (equivalently, \(g \neq 0\)). Also, assume that the generator of \((4.0.4)\) (equivalently, \((4.0.1)\)) is hypoelliptic. With the notation introduced so far,
4.1. Ergodicity of the general approximation

If the matrix $C$ has full rank, i.e. $\text{rank } C = d$, then the solution $x(t)$ of (4.0.4) (equivalently, (4.0.1)) is ergodic. If the matrix $C$ is degenerate, i.e. $\text{rank } C < d$, there exists $\tilde{A} \in \mathcal{A}_C$ (equivalently, $A$) such that the process is ergodic.

**Proof.** We shall prove ergodicity by using the Markov chain technique presented in Section 2.2. Here we just recall that the ergodicity of (4.0.1) is implied by the hypoellipticity of the generator $L$ together with the existence of a Lyapunov function, namely a function $G(x) : \mathbb{R}^n \to [1, \infty)$ such that $G(x) \to \infty$ as $|x| \to \infty$ and $LG(x) \leq -aG(x) + b$ for some $a, b > 0$. Instead of considering system (4.0.1), we make a change of coordinate and consider the equivalent system (4.0.4), see Remark 4.0.2. The fluctuation-dissipation relation (4.0.2), reads then

$$\tilde{A} + \tilde{A}^T = \text{diag}\{\lambda_1, \ldots, \lambda_m, 0, \ldots, 0\}, \quad \lambda_j > 0, \quad j = 1, \ldots, m.$$

Hence

$$\tilde{a}_{ij} = -\tilde{a}_{ji} \quad i \neq j,$$

$$2\tilde{a}_{ii} = \lambda_i > 0 \quad 1 \leq i \leq m$$

$$\tilde{a}_{ii} = 0 \quad m < i \leq d,$$

where we recall that $m = \text{Rank } \Sigma$. In these coordinates $L$ reads

$$L = p\partial_q - \partial_q V \partial_p + \sum_{i=1}^d \tilde{g}_i v_i \partial_p - p \sum_{i=1}^d \tilde{g}_i \partial_v - \sum_{i,j=1}^d \tilde{a}_{ij} v_j \partial_v + \frac{1}{2} \sum_{i=1}^m \lambda_i \partial_{v_i}^2$$

and can be put in the Hörmander’s sum of squares form

$$L = B + \frac{1}{2} \sum_{i=1}^m X_i^2,$$

where $X_i := \sqrt{\lambda_i} \partial_{v_i}$ and

$$B := p\partial_q - \partial_q V \partial_p + \sum_{i=1}^d \tilde{g}_i v_i \partial_p - p \sum_{i=1}^d \tilde{g}_i \partial_v - \sum_{i,j=1}^d \tilde{a}_{ij} v_j \partial_v.$$

Let us see what the hypoellipticity assumption implies on the structure of $\tilde{A}$ and $\tilde{g}$. First of all we recall that the operator $L$ is hypoelliptic if the Lie Algebra

$$\{B, X_1, \ldots, X_d, [X_i, X_j], [X_i, B], [X_j, [X_i, B]], \ldots\}_{i,j=1,\ldots,d}$$

spans $\mathbb{R}^n$ at each point. Because $[X_i, X_j] = 0 \quad \forall i, j = 1, \ldots, m$, the only way in which we can obtain the fields that are not already contained in the generator is by taking commutators with $B$. For the sake of clarity, let us distinguish the two cases in which $m = d$, i.e. $\Sigma$ has full rank, and $0 < m < d$. 
• Case $m = d$. The fields $\partial_{v_1}, \ldots, \partial_{v_d}$ are in the generator. For $k \in \{1, \ldots, d\}$ we have

$$[\partial_{v_k}, B] = \tilde{g}_k \partial_p - \sum_{i=1}^{d} \tilde{a}_{ik} \partial_{v_i}.$$  

Because by assumption $\tilde{g} \neq 0$, there exists at least one $k$ such that $\tilde{g}_k \neq 0$, hence we have obtained the field $\partial_p$.

$$[\partial_p, B] = \partial_q - \sum_{i=1}^{d} \tilde{g}_i \partial_{v_i}.$$  

In other words, when $\Sigma$ is non degenerate, $g \neq 0$ ensures hypoellipticity.

• Case $0 < m < d$. The generator contains $\partial_{v_1}, \ldots, \partial_{v_m}$. For $k \in \{1, \ldots, m\}$,

$$[\partial_{v_k}, B] = \tilde{g}_k \partial_p - \sum_{i=1}^{d} \tilde{a}_{ik} \partial_{v_i}.$$

Notice that $\partial_q$ cannot be recovered without having obtained $\partial_p$ first. From (4.1.4) it is clear that if $\tilde{g}_k \neq 0$ for at least one $k \in I_1 := \{1, \ldots, m\}$ then $\partial_p$ is contained in the Lie Algebra (4.1.3). Let

$$I_{gI} := \{ i \in I_1 \text{ s.t. } \tilde{g}_i \neq 0 \}.$$  

From the commutator (4.1.4) we might also have obtained some of the fields $\partial_{v_j}$, $j \in \{m+1, \ldots, d\}$, we shall get to that in a moment. Suppose $I_{gI}$ is not empty, then

$$[\partial_p, B] = \partial_q - \sum_{i=1}^{d} \tilde{g}_i \partial_{v_i}.$$  

(4.1.5)

In this way we obtain $\partial_q$. Looking at the second addend on the RHS of (4.1.4) and at the second addend on the RHS of (4.1.5), when $I_{gI}$ is not empty, the hypoellipticity assumption implies that we can find two sets of indices, $I_{gII}$ and $I_{a1}$, defined as follows:

$$I_{gII} := \{ m < i \leq d : \tilde{g}_i \neq 0 \},$$

$$I_{a1} := \{ m < i \leq d : \tilde{a}_{il_0} \neq 0 \text{ for some } l_0 \in I_1 \}.$$  

(4.1.6)

and such that $I_{gII} \cup I_{a1} = I_2 = \{m+1, \ldots, d\}$. However notice that, because all the elements in $I_{a1}$ are off-diagonal elements, we can always modify $I_{a1}$ (i.e. modify $\tilde{A}$) such that the fluctuation-dissipation relation (4.1.1) still holds and $I_{a1} = I_2$.

Going back to (4.1.4), if $I_{gI}$ is empty, then in order to have hypoellipticity the set
4.1. Ergodicity of the general approximation

$I_{a_1} = \{ l_1 \in \mathcal{I}_2 \text{ such that } \tilde{a}_{l_1 l_0} \neq 0 \text{ for some } l_0 \in \mathcal{I}_1 \}$ must be not empty. In this way we generate $\partial_{v_1}$ for any $l_1 \in I_{a_1}$ and we have

$$[\partial_{v_1}, B] = \tilde{g}_{l_1} \partial_p - \sum_{i=1}^{d} \tilde{a}_{l_1} \partial_{v_i}.$$ 

If $\tilde{g}_{l_1} = 0$ for any $l_1 \in I_{a_1}$ then we repeat the same argument as before. In the end, because $\tilde{g} \neq 0$, there will be at least an index $l_N$ such that $\tilde{g}_{l_N} \neq 0$ and hence we will have generated $\partial_p$. So, if $I_{g_1}$ is empty, the hypoellipticity condition on $\tilde{A}$ implies that the union of $I_{g_1}, I_{a_1}, \ldots, I_{a_N}$ gives the set $\mathcal{I}_2$, where

$$I_{a_2} = \{ l_2 \in \mathcal{I}_2 : a_{l_2 l_1} \neq 0 \text{ for some } l_1 \in I_{a_1} \text{ and } l_1 \neq l_2 \}$$

and for $2 < j \leq N$

$$I_{a_j} = \{ l_j \in \mathcal{I}_2 : a_{l_j l_{j-1}} \neq 0 \text{ for some } l_{j-1} \in I_{a_{j-1}} \text{ and } l_{j-1} \neq l_j \}.$$ 

To simplify the notation in what follows, set

$$I_a := I_{a_1} \cup \cdots \cup I_{a_N}$$

and

$$I_g := I_{g_1} \cup I_{g_1} = \{ i \in \{1, \ldots, d\} : \tilde{g}_i \neq 0 \}.$$ 

In general the hypoellipticity assumption can be rewritten as

$$I_g \neq \emptyset \quad \text{and} \quad I_a \cup I_{g_1} = \mathcal{I}_2.$$ 

However, as we noticed before, the elements $a_{l_j l_{j-1}}$ that we used to construct the set $I_a$, are all off-diagonal elements and hence they can be chosen (i.e. $\tilde{A}$ can be modified) in a way that the sole $I_a$ coincides with $\mathcal{I}_2$. With this in mind, we modify the matrix $\tilde{A}$ as follows: the cardinality of every set $I_{a_j}$ is exactly one and $I_a = \mathcal{I}_2$. From now on it is understood that $\tilde{A}$ is the modified $\tilde{A}$. \footnote{Notice that for less than a rotation of the coordinates appearing in the $I_{a_j}$’s for which $|I_{a_j}| > 1$, it is always possible to reduce ourselves to such an $\tilde{A}$.} Also, we denote by $s$ the first (according to the above construction) index for which $\tilde{g}_s \neq 0$. For simplicity (and without loss of generality), we assume $s \in I_1$ and $s \neq l_0$, where $l_0$ is as in (4.1.6).
Remark 4.1.2. The modified $\tilde{A}$ is, for example, of the type

$$\tilde{A} = \begin{vmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{vmatrix}$$

so that the overall drift matrix is, for example (taking $s = 1, \tilde{g}_s = 1$ and $\tilde{g}_j = 0 \ \forall j \neq s$)

$$\begin{vmatrix} 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{vmatrix}.$$

Notice that all the eigenvalues of the above drift matrix have strictly positive real part (cfr Section 5.1, page 85).

Set $d - m = M$ and notice that with this modification the generator of the process rewrites

$$L = p\partial_q - q\partial_p + \sum_{i=1}^d \tilde{g}_i v_i \partial_p - \sum_{i=1}^d \tilde{g}_i p \partial_v_i - \sum_{i,j=1}^{d} \tilde{a}_{ij} v_i \partial_v_j - \sum_{k=0}^{M-1} a_{ik+1} l_k (v_i \partial_{v_{ik+1}} - v_{ik+1} \partial_{v_i}) + \sum_{i=1}^{m} \lambda_i \partial_{v_i}^2.$$

We claim that the following $G(x)$ is a suitable Lyapunov function, for an appropriate choice of the positive constants $D, E, \{C_i\}_{i=1}^d, R, H, \{P_{l_{j+1}}\}_{j=0}^{M-1}$:

$$G(x) = D\frac{q^2}{2} + E\frac{p^2}{2} + \frac{1}{2} \sum_{i=1}^d C_i v_i^2 + R(p, q) + \tilde{g}_s H(p, v_s) + \mathcal{P} + 1$$

where

$$\mathcal{P} = \sum_{k=0}^{M-1} (P_{l_{k+1}} - P_{l_{k+1}}) a_{ik+1} l_k v_{ik+1} v_k.$$

Then

$$\partial_q G(x) = Dq + R_p, \quad \partial_p G = Ep + Rq + \tilde{g}_s Hv_s$$
In order to calculate \( LG(x) \) let us start with calculating \( \sum_{k=0}^{M-1} a_{k+1} l_k (v_k \partial v_{k+1} - v_{k+1} \partial v_k) P \):

\[
\begin{align*}
- \sum_{k=0}^{M-1} a_{k+1} l_k (v_k \partial v_{k+1} - v_{k+1} \partial v_k) P \\
= -a_{l_0} l_0 v_0 [(P_{l_0} - P_{l_0}) a_{l_0} l_0 v_0 + (P_{l_1} - P_{l_1}) a_{l_1} l_1 v_1] \\
+ a_{l_0}^2 l_0 v_0^2 (P_{l_0} - P_{l_0}) - a_{l_M l_{M-1}}^2 (P_{l_M l_{M-1}} - P_{l_{M-1} l_{M-1}}) v_{l_{M-1}}^2 \\
- \sum_{k=1}^{M-2} a_{l_k+1} l_k v_k [(P_{l_{k+1}} - P_{l_{k+1}}) a_{l_k} l_k v_k + (P_{l_{k+2}} - P_{l_{k+1}}) a_{l_{k+1}} l_{k+1} v_{k+1}] \\
+ \sum_{k=1}^{M-2} a_{l_k+1} l_k v_{l_k} \left[(P_{l_{k-1}} - P_{l_{k-1}}) a_{l_k} l_{k-1} v_{k-1} + (P_{l_{k+1}} - P_{l_{k+1}}) a_{l_k} l_k v_k \right] \\
+ a_{l_M l_{M-1}} l_M [(P_{l_{M-1} l_{M-2}} - P_{l_{M-2} l_{M-2}}) a_{l_{M-1}} l_{M-2} v_{M-2} + (P_{l_{M-1} l_{M-1}} - P_{l_{M-1} l_{M-1}}) a_{l_{M-1}} l_{M-1} v_{M-1}] \\
\end{align*}
\]

By using Young’s inequality as follows

\[
a_{k+1} l_k v_k (P_{l_{k+2} l_{k+1}} - P_{l_{k+1} l_{k+2}}) a_{k+2} l_{k+1} v_{k+2} \leq a_{k+1} l_k a_{k+2} l_{k+1} \left[(P_{l_{k+2} l_{k+1}} - P_{l_{k+1} l_{k+2}})^2 \frac{|v_k|^2}{2} + \frac{|v_{k+2}|^2}{2} \right]
\]

and

\[
a_{k+1} l_k v_{k+1} \left( P_{l_{k-1} l_{k-2}} - P_{l_{k-2} l_{k-1}} \right) a_{l_{k-1}} l_{k-2} v_{k-1} \leq a_{k+1} l_k a_{k+2} l_{k+1} \left[(P_{l_{k-1} l_{k-2}} - P_{l_{k-2} l_{k-1}})^2 \frac{|v_{k-1}|^2}{2} + \frac{|v_{k+1}|^2}{2} \right]
\]

we obtain

\[
\begin{align*}
- \sum_{k=0}^{M-1} a_{k+1} l_k (v_k \partial v_{k+1} - v_{k+1} \partial v_k) P \\
\leq c |v_M|^2 \left[a_{l_M l_{M-1}}^2 (P_{l_M l_{M-1}} - P_{l_{M-1} l_{M-1}}) \right] \\
+ c \sum_{k=1}^{M-1} |v_k|^2 \left[a_{l_k} l_k^2 (P_{l_{k-1} l_{k-2}} - P_{l_{k-2} l_{k-1}}) + \sum_{j=k}^{M-1} (P_{l_j l_{j+1}}^2 + P_{l_{j+1} l_j}^2) \right] \\
+ |v_0|^2 \left[-a_{l_0} l_0 (P_{l_1 l_0} - P_{l_0 l_1}) + \sum_{j=0}^{M-1} (P_{l_j l_{j+1}}^2 + P_{l_{j+1} l_j}^2) \right], \quad (4.1.8)
\end{align*}
\]
where in the above and from now on \( c \) is a generic strictly positive constant constant that does not depend on any of the constants appearing in the definition of the Lyapunov function \( G \), though it might depend on the entries of \( \tilde{A} \), on the components of \( \tilde{g} \) and on \( d \) and \( m \).

\[
LG(x) = Dpq + Rp^2 - Epq - Rq^2 - \tilde{g}_s H v_s q \\
+ \sum_{i=1}^d \tilde{g}_i v_i E p + \sum_{i=1}^d \tilde{g}_i v_i R q + \sum_{i=1}^d \tilde{g}_i v_i \tilde{g}_s H v_s \\
- d \sum_{i=1}^d C_i \tilde{g}_i p v_i - \tilde{g}_s^2 H p^2 - \tilde{g}_0 p (P_{i_1 l_0} - P_{i_0 l_1}) a_{i_1 l_0} v_{l_1} \\
- \sum_{k=1}^{M-1} g_{l_k} p \left[ (P_{i_{k+1} l_k} - P_{i_k l_{k-1}}) a_{i_{k+1} l_k} v_{l_{k-1}} + (P_{i_k l_{k-1}} - P_{i_{k-1} l_k}) a_{i_{k+1} l_k} v_{l_{k+1}} \right] \\
- g_{l_M} p (P_{M l_M l_{M-1}} - P_{M-1 l_{M-1}}) a_{M l_M l_{M-1}} v_{l_{M-1}} \\
- \sum_{i,j=1}^m a_{i j} v_j C_i v_i - \sum_{j=1}^m a_{i j} \tilde{g}_k H p - \sum_{j=1}^m a_{i j} v_j (P_{l_1 l_0} - P_{l_0 l_1}) a_{i_1 l_0} v_{l_1} \\
- \sum_{k=0}^{M-1} a_{ik+1 l_k} (v_{l_k} C_{ik} v_{l_{k+1}} - v_{l_k} C_{ik} v_{l_{k+1}}) \\
- \sum_{k=0}^{M-1} a_{ik+1 l_k} (v_{l_k} \partial v_{l_{k+1}} - v_{l_{k+1}} \partial v_{l_k}) P + \sum_{i=1}^m \lambda_i C_i.
\]

If we choose \( D = E = C_1 = \ldots = C_d = C \), recalling (4.1.1) we have

\[
LG(x) \leq -Rq^2 + c p^2 \left[ R - \tilde{g}_s^2 H + \sum_{j=0}^{M-1} \left( P_{l_j+l_j}^2 + P_{l_j-l_j}^2 \right) \right] \\
+ c \sum_{i=1}^m v_i^2 \left[ -C + H^2 + P_{l_1 l_0}^2 + P_{l_0 l_1}^2 + R^2 \right] \\
+ c v_{l_0}^2 \left[ -C + H^2 + R^2 + \sum_{j=0}^{M-1} \left( P_{l_j+l_j}^2 + P_{l_j-l_j}^2 \right) \right] \\
+ c \sum_{k=0}^{M-1} v_k^2 \left[ R^2 + a_{l_k l_{k-1}}^2 (P_{l_k l_{k-1}} - P_{l_{k-1} l_k}) + \sum_{j=k}^{M-1} \left( P_{l_j+l_j}^2 + P_{l_j-l_j}^2 \right) \right] \\
+ c v_{l_M}^2 \left[ R^2 + a_{l_M l_{M-1}} (P_{l_M l_{M-1}} - P_{l_{M-1} l_M}) + \sum_{i=1}^m \lambda_i C. \right]
\]

Once we choose

\[
R \ll P_{l_M l_{M-1}} \ll P_{l_{M-1} l_M} \ll \ldots \ll P_{l_{k-1} l_k} \ll P_{l_k l_{k-1}} \ll \ldots \\
\ldots \ll P_{l_2 l_1} \ll P_{l_1 l_2} \ll P_{l_1 l_0} \ll P_{l_0 l_1} \ll H \ll C,
\]
4.2 Open problem: estimating the "distance" between the GLE dynamics and the approximating Markovian process

We have not yet addressed two fundamental questions; let \( q_n \) be the solution of (4.0.1) when \( u \in \mathbb{R}^n \) and \( q \) be the solution of (3.0.1). Does \( q_n \) converge to \( q \) (in some appropriate norm) as \( n \to \infty \)? If yes, what is the rate of convergence? Both problems are still largely lacking of a satisfactory solution. In this section we shall rephrase these questions as questions in approximation theory of real functions of real variable and restrain ourselves to just make some comments on the problem.

Motivated by the discussion in Section 1.2, let \( \gamma(t) \) be an arbitrary function and let \( \gamma_n(t) \) be an approximating sequence for \( \gamma(t) \), i.e. \( |\gamma - \gamma_n| \to 0 \) as \( n \to \infty \), where \(|\cdot|\) denotes some norm, yet to be chosen. We might also know the rate of convergence as a function of \( n \). Let \( q \) denote the solution of the GLE (3.0.1) with kernel \( \gamma \) and \( q_n \) denote the solution of the GLE with kernel \( \gamma_n \). If we assume the initial conditions for the full dynamics to be the same as the ones for the approximating system, a straightforward calculation gives

\[
q(t) - q_n(t) = \int_0^t ds \int_0^s du [\partial_q V(q_n(u)) - \partial_q V(q(u))] \\
+ \int_0^t ds[q_n(s)\Gamma_n(t-s) - q(s)\Gamma(t-s)] + \tilde{F}(t) - \tilde{F}_n(t),
\]

where \( \Gamma(t) = \int_0^t dv \gamma(v) \) and \( \tilde{F}(t) = \int_0^t ds \int_0^s du F(u) \) (with analogous definitions for \( \Gamma_n \) and \( F_n \), respectively). Therefore if \( \|\gamma - \gamma_n\|_{L^2[0,T]} \) and \( E|F(t) - F_n(t)|^2 \) tend to zero as \( n \to \infty \), under a smoothness assumption on the potential it follows that also \( E\sup_{t \in [0,T]} |q(t) - q_n(t)|^2 \) tends to zero. The point is that we want the approximating dynamics to be Markovian.

As we discussed in Section 1.2, for some specific choices of the kernel \( \gamma(t) \) the GLE is equivalent to a finite dimensional Markovian system (in an extended state space). This happens at least in three cases: when \( \gamma(t) \) is a finite sum of exponentials; when the Fourier transform of \( \gamma(t) \) belongs to the set of functions \( \mathcal{F}_R := \{ \text{rational functions of the form } 1/|p(k)|^2, \text{where } p(k) \text{ is a polynomial with real coefficients and roots in the upper half plane } \} \); when the Laplace transform of \( \gamma(t) \) admits a continued fraction expansion. So a natural way of approaching the problem is the following: first, we approximate an arbitrary function \( \gamma(t) \), kernel of the GLE
4.2. Open problem: estimating the "distance" between the GLE dynamics and the approximating
Markovian process

(3.0.1), through a sequence of functions $\gamma_n(t)$ such that, $\forall n \in \mathbb{N}$, $\gamma_n(t)$ belongs to either one
of the above listed class of functions. Second, using the fluctuation dissipation principle, try
and obtain convergence of the noise $F_n$ to $F$. The success of the second step clearly relies on
the norm chosen to approximate $\gamma$.

In the following we want to point out some of the difficulties that one encounters when
dealing with this problem. Approximating the Fourier transform of $\gamma$ through functions in the
set $\mathcal{F}_R$ is a very difficult task. The case in which $\gamma_n$ is the truncation of a continued fraction
expansion seems more viable and we intend to consider this situation in [61]. In the remainder of
this section we choose to focus the discussion on the case in which $\gamma_n(t) = \sum_{i=1}^{n} \lambda_i^2 e^{-\alpha_i t}, t \geq 0, \alpha_i > 0 \forall i$. This choice is motivated by the possibility of obtaining detailed information for
system (3.0.2) and by the work [8], which tackles the problem of the $L^2[0, \infty)$ approximation of
functions through sums of exponentials. Indeed, let us start by summarizing some of the results
obtained in [8]. If $(\lambda_k)_{k \geq 0}$ is a sequence of positive real numbers which admits a subsequence
$(\lambda_k_j)_{j \geq 0}$ s.t. $\lambda_k_j \to \lambda > 0$ as $j \to \infty$, then linear combinations of the functions $(e^{-\lambda_k t})_{k \geq 0}, t \geq 0,$
are dense in $\mathcal{U}$, the space of functions in $L^2(\mathbb{R})$ with support contained in $[0, \infty)$. This can
be shown by using the principle of isolated zeros for analytic functions. In particular, linear
combinations of functions of the type $(e^{-(\alpha+q)k t})_{k \geq 0}, q \in (0, 1), \alpha > 0, t \geq 0$, are dense in $\mathcal{U}$. If $\alpha = 0$ explicit computations are easier to carry out but this family of functions is not
dense in $\mathcal{U}$ and we pay the price of restricting the set of target functions for which we seek an
approximation. More precisely, let $\mathcal{U}_n = \text{span}\{e^{-q^k t}, 0 \leq k \leq n, q \in (0, 1), t \geq 0\}$ and $U$ be
the closure in $L^2$ of the subspace $\mathcal{U}_1 + \mathcal{U}_2 + \ldots$. Then it is proved in [8] that if $f$ belongs to
$U$ and its Fourier transform is analytic then

$$\|f - f_n\|_{L^2} \leq C q^{n/2},$$

where $C$ is a computable constant and $f_n$ is the projection of $f$ on $\mathcal{U}_n$. Translating everything
to our context, we can write

$$\gamma_n(t) = \sum_{j=0}^{n} \frac{\langle \gamma, r_j \rangle}{\|r_j\|^2} r_j(t),$$

where $\| \cdot \|$ and $\langle \cdot, \cdot \rangle$ denote the $L^2$ norm and scalar product, respectively;

$$r_j(t) = \sum_{k=0}^{j} r_{jk} e^{-q^k t}, \quad r_{jk} \in \mathbb{R},$$

and the coefficients $r_{jk}$ are chosen in such a way that $r_n$ is orthogonal to $e^{-q^l t}, \forall i = 0, \ldots, n-1$. 

4.2. Open problem: estimating the "distance" between the GLE dynamics and the approximating Markovian process

and hence the $r_n$'s are mutually orthogonal. The expression for $\gamma_n$ can be rewritten as

$$\gamma_n(t) = \sum_{j=0}^{n} \sum_{k=0}^{j} \frac{(\gamma \cdot r_j)}{\|r_j\|^2} r_{jk} e^{-q_k t} = \sum_{k=0}^{n} \tilde{r}_{kn} e^{-q_k t},$$

having set $\tilde{r}_{kn} := \sum_{j=k}^{n} \frac{(\gamma \cdot r_j)}{\|r_j\|^2} r_{jk}$. Clearly, because of orthogonality, the coefficients $\tilde{r}_{kn}$ will not be all positive. The non positivity of these coefficients is somewhat an issue; indeed, if we use $\gamma_n$ given in (4.2.3) as correlation function of the approximating system and not all the $\tilde{r}_{kn}$ are positive, what we obtain is not system (0.0.5), i.e. we do not obtain independent Ornstein-Uhlenbeck processes as heat bath variables. We want to stress how $\gamma_n(t)$ could be positive even if some of the coefficients are negative and hence it could still be an acceptable correlation function; but in this case the approximating system would not be (0.0.5). Approximating a function (even a sufficiently smooth function) $\gamma(t)$ through sums of exponentials with strictly positive coefficients is an extremely hard task in approximation theory, mainly because we cannot use an orthogonal basis to decompose the function. Intuitively, we believe that restricting the class of target functions to negative powers, i.e. to functions of the form $t^{-a}$, $a > 0$, should simplify the problem but so far we haven’t obtained any results in this sense. However, once we have an $L^2$ approximation, we know that there exists a subsequence $\gamma_{n_k}(t)$ that converges pointwise to $\gamma(t)$ $\forall t \geq 0$ (because $\gamma(t)$ is assumed to be at least continuous). Recalling that $F_n$ is stationary, Gaussian and mean zero with autocorrelation function $E(F_n(t)F_n(s)) = \gamma_n(t-s)$ (here we have set the inverse temperature $\beta = 1$ for simplicity), an elementary calculation$^2$ shows that

$$E \left| F_{n_k}(t) - F(t) \right|^2 \leq \frac{1}{3} \left| \gamma(0) - \gamma_{n_k}(0) \right|,$$

and hence the subsequence $F_{n_k}$ converges to $F$ in quadratic mean. Though, when passing from the sequence $\gamma(t)$ to the subsequence $\gamma_{n_k}(t)$ we loose information on the rate of convergence. The next section deals with the first step towards the understanding of the approximation problem, i.e. the well posedness of system (3.0.2) when the number of heat bath variables tends to infinity.

$^2$ $\gamma_n(0) - \gamma(0) = E \left| F_n(t) \right|^2 - E \left| F(t) \right|^2 = E \left| F_n(t) - F(t) \right|^2 - 2E \left| F(t) \right|^2 + 2E \left| F_n(t) \right|^2 + 2E[F_n(t)(F_n(t) - F_n(t))] \Rightarrow 3E \left| F_n(t) - F(t) \right|^2 \leq (\gamma(0) - \gamma_n(0)) - \gamma(0)/2 \leq |\gamma(0) - \gamma_n(0)|$, assuming $\gamma(0) > 0$. 

4.3 Approximation by SDE in infinite dimensions

Following [10, 66] we present some background material on stochastic processes in infinite dimensions. We refer the reader to [66] for further details and proofs.

Let $(\Omega, \mathcal{F}, P)$ be a probability triple and $U$ and $H$ be separable Hilbert spaces. We denote by $L(U, H)$ the space of linear and bounded operators $T : U \to H$. $T \in L(U, H)$ is said to be trace class if $\text{Tr} T := \sum_{n \in \mathbb{N}} (T \varphi_n, \varphi_n) < \infty$, where $\{\varphi_n\}_{n \in \mathbb{N}}$ is an orthonormal basis of $U$ (and the definition can be shown to be independent of the chosen orthonormal basis). $T \in L(U, H)$ is said to be Hilbert-Schmidt if $\|T\|_2^2 := \sum_{n \in \mathbb{N}} (T \varphi_n, T \varphi_n) < \infty$ and $L^2(U, H)$ denotes the space of Hilbert-schmidt operators from $U$ to $H$. A centred Gaussian probability measure $\rho$ on $U$ is a Borel measure such that $f^* \rho = \rho(f(f^{-1}))$ is centred and Gaussian for any linear functional $f : U \to \mathbb{R}$. The covariance operator of $\rho$ is the symmetric and non-negative operator $A : U \to U$ such that $(Ah, g)_U = \int_U (x, h)_U (x, g)_U \rho(dx)$ and we say that $\rho$ has $\mathcal{N}(0, A)$-law. It can be shown that the covariance operator is trace class. Conversely, if $A : U \to U$, $A \in L(U)$, is a non-negative and symmetric operator with finite trace, then there exists a mean zero Gaussian measure with $\mathcal{N}(0, A)$ law; moreover, there exists an orthonormal basis of $U$, $\{e_k\}_{k \in \mathbb{N}}$, in which $A$ diagonalizes:

$$A e_k = \lambda_k e_k \quad \forall k \in \mathbb{N}, \quad (4.3.1)$$

with $\{\lambda_k\}_{k \in \mathbb{N}}$ a bounded sequence of non-negative real numbers. So, in the following we shall always assume that $A \in L(U)$ is a symmetric, non-negative operator with $\text{Tr} A < +\infty$, i.e.

$$\sum_{k=1}^{\infty} \lambda_k < +\infty, \quad (4.3.2)$$

A stochastic process $W(t) : \mathbb{R}_+ \to U$ with $W(0) = 0$ is an $A$-Wiener process if it has ($P$ a.s.) continuous trajectories, independent increments and the law of $W(t) - W(s)$ is $\mathcal{N}(0, (t-s)A)$. A $U$-valued $A$-Wiener process can be represented as an $L^2(\Omega, \mathcal{F}, P; U)$-converging series,

$$W(t) = \sum_{j=1}^{\infty} \sqrt{\lambda_j} \beta_j(t) e_j, \quad t \in [0, T], \quad (4.3.3)$$

where $\{\beta_j(t)\}_{j \in \mathbb{N}}$ are independent $\mathbb{R}$-valued standard Brownian motions,

$$\beta_j(t) = \frac{1}{\sqrt{\lambda_j}} \langle W(t), e_j \rangle_U. \quad (4.3.4)$$
(More precisely, the series (4.3.3) can be shown to converge also in $L^2(\Omega, \mathcal{F}, P; C([0, T], U))$.) The construction of the stochastic integral with respect to an infinite dimensional Wiener process can be done analogously to the finite dimensional case, so we will just sketch it. We start with the class $\mathcal{E}$ of $L(U, H)$-valued elementary processes with normal filtration $\mathcal{F}_t$, i.e. processes of the form

$$\Phi(t) = \sum_{m=0}^{k-1} \Phi_m \mathbf{1}_{(t_m, t_{m+1}]}(t), \quad t \in [0, T],$$

where $\Phi_m : \Omega \to L(U, H)$ is $\mathcal{F}_{t_m}$-measurable and takes only a finite number of values, $\forall m = 0, \ldots, k - 1$. Now everything goes as usual: we first define the stochastic integral for elementary processes:

$$\int_0^t \Phi(s) dW_s := \sum_{m=0}^{k-1} \Phi_m (W(t_{m+1} \wedge t) - W(t_m \wedge t)).$$

Then we prove that $\text{Int} : (\mathcal{E}, \| \cdot \|_T) \to (\mathcal{M}^2_T, \| \cdot \|_{\mathcal{M}^2_T})$ is an isometry, where $\mathcal{M}^2_T := \mathcal{M}^2_T(H)$ is the Banach space of $H$-valued square integrable martingales with norm $\|M\|^2_{\mathcal{M}^2_T} := E(\|M(T)\|^2_H)$ and

$$\|\Phi\|^2_T := E \int_0^T \|\Phi(s)A^{1/2}\|^2_{L_2} ds.$$

Hence $\text{Int}$ can be uniquely extended to $\bar{\mathcal{E}}$. In order to give an explicit representation of $\bar{\mathcal{E}}$, let us introduce the Hilbert space $U_0 := A^{1/2}U \subset U$ with scalar product

$$\langle u, v \rangle_0 := \langle A^{-1/2}u, A^{-1/2}v \rangle_U,$$

where $A^{-1/2}$ is in general the pseudo-inverse of $A^{1/2}$. Now let $L^0_{2} := L_2(U_0, H)$ be the space of Hilbert-Schmidt operators from $U_0$ to $H$; it is possible to show that

$$\|T\|^2_{L^0_2} = \|TA^{1/2}\|^2_{L_2}, \quad \forall T \in L^0_{2}.$$

It turns out that $\bar{\mathcal{E}}$ is the space $N_{W}^2([0, T], L^0_2(U_0, H))$, the space of predictable $L^0_2$-valued processes with finite $\| \cdot \|_T$ norm. To recap, if $W(t)$ is a $A$-Wiener process, the process

$$X(t) = X(0) + \int_0^t ds\phi(s) + \int_0^T \Phi(s) dW(s) \quad (4.3.5)$$

is well defined for $\Phi(t) \in N_{W}^2$, $\phi(t)$ an $H$-valued predictable process $P$-a.s. Bochner integrable in $[0, T]$ and $X(0)$ an $H$-valued $\mathcal{F}_0$-measurable random variable. The framework described

\footnote{A random variable $\phi$ on the probability triple $(\Omega, \mathcal{F}, P)$ and taking values in a Hilbert space $H$ is said to be Bochner integrable if $\int_{\Omega} \|\phi(\omega)\|_H P(d\omega) < \infty$.}
so far can be used to make sense out of the following system:

\[ dq(t) = p(t) \, dt \]  
\[ dp(t) = -\partial_q V(q(t)) \, dt + \sum_{i=1}^{\infty} \lambda_i z_i(t) \, dt \]  
\[ dz_i(t) = -\alpha_i z_i(t) \, dt - \lambda_i p(t) \, dt + \sqrt{\alpha_i} \, d\beta_i(t), \quad i \in \mathbb{N}, \]

where \( \{\alpha_i\}_{i \in \mathbb{N}}, \{\lambda_i\}_{i \in \mathbb{N}} \) are square summable sequences of positive real numbers, that is

\[ \sum_{j=1}^{\infty} \alpha_j^2 < +\infty, \quad \sum_{j=1}^{\infty} \lambda_j^2 < +\infty, \]  

and \( \beta_i(t) \) are as in (4.3.4). Also, \( q, p, z_i \in \mathbb{R}^d \). For simplicity we will fix \( d = 1 \) but everything we are going to say in the following holds for arbitrary finite \( d \geq 1 \). For the purpose of our analysis we shall take \( U \) to be the space of square summable sequences, \( U := \ell^2 \) with the usual \( \ell^2 \) scalar product, and \( H := \mathbb{R} \times \mathbb{R} \times \ell^2 \) with scalar product

\[ (h_1, h_2) := (q_1, q_2)_{\mathbb{R}} + (p_1, p_2)_{\mathbb{R}} + (l_1, l_2)_{\ell^2}, \quad \forall h_i = (q_i, p_i, l_i) \in H, \quad i = 1, 2. \]

We fix \( \{e_k\}_{k \in \mathbb{N}} \) to be an orthonormal basis for \( \ell^2 \), which we will assume to be the canonical basis (i.e. \( e_j \) is the sequence s.t. \( e_{j_k} = \delta_{jk} \)). Consider the operators \( A : \ell^2 \to \ell^2 \) and \( \Lambda : \ell^2 \to \ell^2 \) defined as follows

\[ Al := \sum_{j=1}^{\infty} \alpha_j l_j e_j, \quad l \in \ell^2 \]  
\[ \Lambda l := \sum_{j=1}^{\infty} \lambda_j l_j e_j, \quad l \in \ell^2. \]

We also define \( \sigma : H \to H \) and \( \Sigma : U \to H \) as

\[ \sigma = \begin{pmatrix} p \\ -\partial_q V(q) + \sum_{i=1}^{\infty} (\Lambda z,e_i)_U \\ -\Lambda z - p \sum_{i=1}^{\infty} \Lambda e_i \end{pmatrix} \]

and

\[ \Sigma = \begin{pmatrix} 0 \\ 0 \\ I_{d_{\infty}} \end{pmatrix}. \]
where \( \text{Id}_\infty \) is the identity on the space of linear operators from \( U \) into \( U \). Setting \( X(t) := (q(t), p(t), z(t)) \in H \), system (4.3.6) can now be recast in a more compact form, as follows

\[
dX(t) = \sigma(X(t))dt + \Sigma dW,
\]

where \( W(t) \) is an \( U \)-valued, \( A \)-Wiener process given by (4.3.3). In order for (4.3.12) to make sense, we need to check that \( \Sigma \) is well defined as an operator in \( L_2^0 \), which is the case as

\[
\|\Sigma\|_{L_2^0}^2 = \|\Sigma A^{1/2}\|_{L_2^0}^2 = \sum_{h=1}^{\infty} \langle A^{1/2}e_h, A^{1/2}e_h \rangle_U = \sum_{h=1}^{\infty} \alpha_h < +\infty,
\]

by assumption. Hence the process \( X(t) \) in (4.3.12) is well defined. We only need to prove that \( z(t) \) actually belongs to \( \ell^2 \), \( \forall t \geq 0 \). To this purpose, assume

\[
E \sum_{i=0}^{\infty} |z_i(0)|^2 < \infty. \tag{4.3.13}
\]

From (4.3.6b) and applying Itô’s formula to (4.3.6c), we get

\[
\partial_t(p^2(t)) = -2p(t)\partial_q V dt + 2 \sum_{i=0}^{\infty} \lambda_i z_i(t)p(t) dt \tag{4.3.14}
\]

and

\[
d(z_i^2(t)) = 2(-\alpha_i z_i(t) - \lambda_i p(t))z_i(t) dt + \alpha_i dt + 2\sqrt{\alpha_i}z_i(t)d\beta_i(t). \tag{4.3.15}
\]

Also, from (4.3.6a), it is straightforward to see that

\[
E |q(t)|^2 \leq C \left( E |q(0)|^2 + t \int_0^t E |p(t)|^2 \right), \tag{4.3.16}
\]

where in the above, as well as in the following, \( C \) is a generic constant. Let \( \alpha = \max_i \{\alpha_i\} \), \( \bar{\alpha} = \sum_{i=1}^{\infty} \alpha_i \), \( \lambda = \max_i \{\lambda_i\} \), \( \bar{\lambda} = \sum_{i=1}^{\infty} \lambda_i^2 \). Using the simple inequality

\[
\sum_{i=1}^{\infty} \lambda_i z_i(t)p(t) \leq \sum_{i=1}^{\infty} z_i^2(t) + \sum_{i=1}^{\infty} \lambda_i^2 p_i^2(t),
\]

it follows from (4.3.14) the bound

\[
E |p(t)|^2 \leq C \left( E |p(0)|^2 + (1 + \lambda) \int_0^t E |p(t)|^2 + \int_0^t E |\partial_q V|^2 + \int_0^t E \sum_{i=1}^{\infty} |z_i(t)|^2 \right). \tag{4.3.17}
\]

Analogously, from (4.3.15) we get

\[
E \sum_{i=1}^{\infty} |z_i(t)|^2 \leq C(\bar{\alpha}, \alpha, \lambda) \left( E \sum_{i=1}^{\infty} |z_i(0)|^2 + t + \int_0^t E |p(t)|^2 + \int_0^t E \sum_{i=1}^{\infty} |z_i(t)|^2 \right). \tag{4.3.18}
\]
Putting (4.3.16), (4.3.17) and (4.3.18) together,

\[ E|q(t)|^2 + E|p(t)|^2 + \sum_{i=0}^{\infty} |z_i(t)|^2 \leq C f(t) + C(1+t) \int_0^t ds \left( E \sum_{i=0}^{\infty} |z_i(s)|^2 + E|p(s)|^2 + E|q(s)|^2 \right), \]

where \( f(t) = E \sum_{i=0}^{\infty} |z_i(0)|^2 + E|p(0)|^2 + t + E|q(0)|^2 \) and \( C \) is a generic constant that depends on the \( \alpha_j \)'s and \( \lambda_j \)'s. By Gronwall's Lemma we then conclude

\[ E|q(t)|^2 + E|p(t)|^2 + \sum_{i=0}^{\infty} |z_i(t)|^2 \leq C f(t) + C(1+t) \int_0^t f(s)e^{C \int_s^t (1+r)dr} < \infty, \quad \forall t > 0. \]

Therefore, assuming (4.3.7) and (4.3.13), \( z(t) \in \ell^2 \forall t \geq 0. \)
Chapter 5

Spectral theory for Ornstein-Uhlenbeck Operators

In this section, some applications of the singular space theory to stochastic differential equations are discussed. Consider a stochastic differential equation in $\mathbb{R}^d$ of the form

$$dx(t) = b(x(t)) \, dt + \sigma \, dw(t),$$

(5.0.1)

where $w(t)$ is a $d$-dimensional standard Brownian motion and $Q = \sigma\sigma^T$ is a positive semidefinite matrix. The generator of this process is the second order operator

$$L = b(x) \cdot \nabla + \frac{1}{2} \sum_{i,j=1}^{d} Q_{ij} \frac{\partial^2}{\partial x_i \partial x_j}.$$  

(5.0.2)

The $L^2$-adjoint of this operator, i.e. the Fokker-Planck operator is

$$L' = -\nabla \cdot (b(x) \cdot \nabla) - \frac{1}{2} \sum_{i,j=1}^{d} Q_{ij} \frac{\partial^2}{\partial x_i \partial x_j}.$$  

(5.0.3)

In all the cases that we consider in this section the generator $L$ and its adjoint $L'$ are hypoelliptic operators and have smooth density with respect to the Lebesgue measure. Suppose that the diffusion process $x(t)$ is ergodic with invariant measure $\mu(dx)$. The density of this measure, which we shall denote by $\rho(x) := \exp(-\Phi(x))$, satisfies the stationary Fokker-Planck equation

$$L' \rho = 0.$$
5.1. Ornstein-Uhlenbeck operators

We will work in the following function spaces: we will study $L$ in $L^2(\mathbb{R}^d, \rho(x))$ and $L'$ in $L^2(\mathbb{R}^d, \rho^{-1}(x))$. Furthermore, we will consider the following two operators

$$H \cdot := \sqrt{\rho} L \left( \sqrt{\rho^{-1}} \cdot \right),$$

and

$$H' \cdot := \sqrt{\rho^{-1}} \left( L' \sqrt{\rho} \cdot \right).$$

We will study these two operators in the flat $L^2$ space, $L^2(\mathbb{R}^d)$. $H$ and $H'$, as operators acting on $L^2(\mathbb{R}^d)$ are unitarily equivalent to $L$ (acting on $L^2(\mathbb{R}^d, \rho(x))$) and $L'$ (acting on $L^2(\mathbb{R}^d, \rho^{-1}(x))$), respectively. In particular, the spectrum of $H$ ($H'$, respectively) in $L^2$ is the same as the spectrum of $L$ in $L^2_\rho$ ($L'$ in $L^2_{\rho^{-1}}$, respectively). This is easily seen by using (5.0.4).

Suppose $H f_n = \lambda_n f_n$; then setting $g_n = \sqrt{\rho^{-1}} f_n$ we have $L(g_n) = \lambda_n g_n$. Analogously one can show the other implication, that if $\lambda_n$ is an eigenvalue of $L$ then it is also an eigenvalue of $H$. Further details can be found in [28, 17, 14]. One last remark: in this setting the assumptions of Theorem 2.4.3 are fulfilled. Indeed, for an ergodic (dissipative) process we already know that 0 is the first eigenvalue at the bottom of the spectrum (this is the content of the Koopman-von Neumann Theorem, see Theorem 1.2.1 in [10]). Moreover, because $\rho$ is the ground state of $L'$, $\sqrt{\rho}$ is the ground state of $H'$ and it is the ground state of $H$ as well. Hence, by Theorem 2.4.3 we know that $\Phi(x)$ is a positive definite quadratic form.

5.1. Ornstein-Uhlenbeck operators

In this subsection, we study the Ornstein-Uhlenbeck (O-U) process

$$dX = -BX dt + \Sigma dW,$$ (5.1.1)

where $X(t) \in \mathbb{R}^n$, $B, \Sigma \in \mathbb{R}^{n \times n}$ and $W(t)$ is a $n-$dimensional standard Brownian motion. The generator of this process is

$$L = \frac{1}{2} Tr(Q \nabla_x^2) - \langle Bx, \nabla_x \rangle, \quad x = (x_1, x_2, ..., x_n) \in \mathbb{R}^n,$$ (5.1.2)

with $Q \in \mathbb{R}^{n \times n}$ a symmetric positive definite matrix, $Q = \Sigma \Sigma^T$, $\Sigma^T$ being the transpose matrix of $\Sigma$. If $X$ is a matrix, $Tr.X$ denotes the trace of $X$. The process (5.1.1) has a unique
invariant measure if and only if all the eigenvalues of the matrix $B$ have strictly positive real part \cite[Chapter 9]{47}. Hence we will assume that the spectrum of $B$, $\sigma(B)$, is contained in $\mathbb{C}^+ = \{ \lambda \in \mathbb{C} : \text{Re}\lambda > 0 \}$, so in particular $B$ is invertible. When $Q$ is non-degenerate, $\mathcal{L}$ is elliptic hence its spectral properties are well known. Here we want to focus on the case where $Q$ is degenerate, in particular we want to consider the case where $\mathcal{L}$ is hypoelliptic. This problem has been studied in \cite{53}. It is recalled in this work that the generator of the O-U process is hypoelliptic if and only if one of the following equivalent assertions holds (see also \cite{47}):

1. The kernel of $Q$ does not contain any non trivial invariant subspace of $B^T$;
2. $\forall t > 0$, the matrix $Q_t$
   \begin{equation}
   Q_t = \int_0^t e^{-sB}Qe^{-sB^T}ds
   \end{equation}
   is nonsingular, i.e. $\det Q_t > 0$;
3. $\text{Rank}\: [B|\Sigma] = n$, with $[B|\Sigma] = [\Sigma, B\Sigma, \ldots, B^{n-1}\Sigma]$ the $n \times n^2$ matrix obtained by writing consecutively the columns of the matrices $B^j\Sigma$. In the language of control theory, the condition $\text{Rank}\: [B|\Sigma] = n$ is also expressed by saying that the pair $(B, \Sigma)$ is a controllable pair.

We recall that the Weyl quantization of the quadratic symbol $x^\alpha\xi^\beta$, $(\alpha + \beta) = 2$, is the differential operator
\begin{equation}
\frac{x^\alpha D^\beta_x + D^\beta_x x^\alpha}{2}, \quad D_x = i^{-1}\partial_x.
\end{equation}
The Hamilton map $F \in M_{2n}(\mathbb{C})$ associated to the quadratic form $q$ is the unique $2n \times 2n$ matrix defined by the identity (2.4.11) and the singular space $S$ is defined as
\begin{equation}
S = \left( \bigcap_{j=0}^{2n-1} \text{Ker} [\text{Re}\: F(\text{Im}\: F)^j] \right) \cap \mathbb{R}^{2n}. \quad (5.1.3)
\end{equation}
Consider the operator
\begin{equation}
\mathcal{H} = -\sqrt{\rho}\mathcal{L}(\sqrt{\rho^{-1}}\cdot) + \frac{1}{2}\text{Tr}B. \quad (5.1.4)
\end{equation}
We want to apply Theorem 2.4.2 in order to determine the spectrum of the operator $\mathcal{H}$ in $L^2(dx)$ and hence the spectrum of $\mathcal{L}$ in $L^2(\rho(x)dx)$. The main result of this section is the following Proposition 5.1.1.
Proposition 5.1.1. Suppose $\sigma(B) \subset \mathbb{C}^+$. Then the singular space $S$ associated with $\mathcal{H}$ is trivial if and only if both the process

$$dX(t) = -BX(t) dt + \Sigma dW$$ (5.1.5)

and the process

$$dY(t) = -B^T Y(t) dt + \Sigma dW$$ (5.1.6)

are hypoelliptic.

Before getting to the proof of the above proposition, let us make some remarks. First of all notice that in general the hypoellipticity of (5.1.5) does not imply the hypoellipticity of (5.1.6) and viceversa. This is clear looking at condition (i). Indeed, even if $B$ and $B^T$ have the same eigenvalues, they have in general different eigenspaces, so $\text{Ker} Q$ might not contain invariant subspaces under $B^T$ and still contain invariant subspaces under $B$. Secondly, as we have mentioned in Section 2.4.2, we already know that if $S = \{0\}$ then (5.1.5) is hypoelliptic.

Let us now recall some basic facts about the process (5.1.5), which can be found in [47] and in [53]. Assuming that $X(t)$ is hypoelliptic, the unique invariant measure for the process (5.1.5) is

$$\rho(dx) = \frac{1}{\sqrt{(2\pi)^n \det Q_\infty}} e^{-\frac{1}{2}(Q_\infty^{-1}x,x)} dx,$$ (5.1.7)

where

$$Q_\infty = \int_0^\infty ds e^{-sB} Q e^{-sB^T}.$$ 

The matrix $Q_\infty$ is symmetric and strictly positive definite; in particular it is invertible thanks to the assumption of hypoellipticity (ii) and its inverse is symmetric. Therefore expression (5.1.7) is well posed. We will often make use of the steady state variance equation

$$Q = BQ_\infty + Q_\infty B^T,$$ (5.1.8)

which can be easily derived by differentiating in $t = 0$ the elementary relation

$$Q_t + e^{-tB} Q_\infty e^{-tB^T} = Q_\infty.$$

In order to obtain an expression for $\mathcal{H}$ we use the fact that the density $\rho(dx)$ defined in (5.1.7) is the unique (normalized) solution of the equation $\mathcal{L} \rho = 0$, where $\mathcal{L}'$ is the $L^2(\mathbb{R}, dx)$-adjoint of $\mathcal{L}$. Indeed

$$\mathcal{L}' = Tr B + \langle Bx, \nabla \cdot \rangle + \frac{1}{2} Tr(Q \nabla^2),$$
so it is easy to check that

\[
\mathcal{L}' \rho = 0 \iff \begin{cases} 
\text{Tr} B = \frac{1}{2} \text{Tr} \left( Q Q^{-1}_\infty \right) \\
\langle B x, Q^{-1}_\infty x \rangle = \frac{1}{2} \sum_{k,j} q_{kj} \sum_l a_{kl} x_l \sum_l a_{jl} x_l, 
\end{cases}
\]

(5.1.9)

where \( a_{hk} = \{ Q^{-1}_\infty \}_{hk} \). Using (5.1.9) we obtain

\[
\sqrt{\rho} \mathcal{L}(\sqrt{\rho^{-1}}) = -\frac{1}{4} \langle Q^{-1}_\infty x, B x \rangle + \frac{1}{2} \text{Tr} (Q \nabla^2) + \langle \left( \frac{1}{2} Q Q^{-1}_\infty - B \right) x, \nabla \rangle + \frac{1}{2} \text{Tr} B
\]

(5.1.10)

\[
= -\frac{1}{4} \sum_{h,k=1}^n a_{hk} x_k \sum_{j=1}^n b_{hj} x_j + \frac{1}{2} \sum_{h,k=1}^n q_{hk} \frac{\partial^2}{\partial h \partial k} + \frac{1}{2} \text{Tr} B
\]

(5.1.11)

Let us set

\[ P = Q^{-1}_\infty B, \quad M = \frac{1}{2} Q Q^{-1}_\infty - B. \]

(5.1.12)

Then

\[
\mathcal{H} = +\frac{1}{4} \langle Px, x \rangle - \frac{1}{2} \text{Tr} (Q \nabla^2) - \langle M x, \nabla \rangle
\]

(5.1.13)

\[
= +\frac{1}{4} \sum_{h,k=1}^n P_{hk} x_h x_k - \frac{1}{2} \sum_{h,k=1}^n q_{hk} \frac{\partial^2}{\partial h \partial k} - \sum_{h,k=1}^n m_{hk} x_k \frac{\partial}{\partial h}.
\]

By (5.1.12) and the steady state variance equation, \( M = -B/2 + (Q_\infty B^T Q^{-1}_\infty)/2 \) hence \( \text{Tr} M = 0 \); therefore the symbol (in the Weyl quantization) of \( \mathcal{H} \) can be expressed as

\[
q(x, \xi) = \frac{1}{4} \sum_{h,k=1}^n P_{hk} x_h x_k + \frac{1}{2} \sum_{h,k=1}^n q_{hk} \xi_h \xi_k - i \sum_{h,k=1}^n m_{hk} x_k \xi_h,
\]

\((x, \xi) \in \mathbb{R}^{2n}\). For general \( Q \) and \( B \) the symbol is not real (\( \mathcal{H} \) is not self adjoint). Moreover, recalling that a matrix is positive definite if and only if its symmetric part is positive definite, we have that \( \text{Re}(q) \geq 0 \), in fact

\[
\text{Re} q = \frac{1}{4} \langle Q^{-1}_\infty B \rangle + \frac{1}{2} Q,
\]

(5.1.14)

where \( Q \) is positive definite by assumption and, from (5.1.8), the symmetric part of \( (Q^{-1}_\infty B)^2 \)

\( (Q^{-1}_\infty Q Q^{-1}_\infty) / 2 \), which is positive definite as well. The Hamilton map \( F \in \mathbb{R}^{2n \times 2n} \) (defined

\(^2\)The symmetric part of a matrix \( M \) is \((M + M^T)/2\).
through relation (2.4.11)) associated with the symbol \( q(x, \xi) \) is

\[
F = \begin{vmatrix}
-\frac{i}{2}M & Q/2 \\
-(P + P^T)/8 & +\frac{i}{2}M^T
\end{vmatrix};
\]

(5.1.15)

hence, for \( j = 0, \ldots, 2n - 1 \),

\[
Ker[ReF(ImF)^j] = \{(x, \xi) \in \mathbb{R}^{2n} : (P + P^T)M^jx = 0 \text{ and } Q(M^T)^j\xi = 0\}
\]

and the singular space \( S \) associated with (5.1.15), is

\[
S = \bigcap_{j=0}^{2n-1} \{(x, \xi) \in \mathbb{R}^{2n} : (P + P^T)M^jx = 0 \text{ and } Q(M^T)^j\xi = 0\}.
\]

We will need the following Remark 5.1.2 and Lemma 5.1.3.

**Remark 5.1.2.** \( S = \{0\} \) if and only if the \( n^2 \times n \) matrices

\[
\begin{vmatrix}
(P + P^T)M \\
(P + P^T)M^2 \\
\vdots \\
(P + P^T)M^{2n-1}
\end{vmatrix}
\quad \begin{vmatrix}
Q \\
QM^T \\
\vdots \\
Q(M^T)^{2n-1}
\end{vmatrix}
\]

and

\[
\begin{vmatrix}
(P + P^T)x = 0, \quad (P + P^T)Mx = 0, \ldots, \quad (P + P^T)M^{2n-1}x = 0 \\
(Q\xi = 0, \quad QM^T\xi = 0, \ldots, \quad Q(M^T)^{2n-1}\xi = 0)
\end{vmatrix}
\]

Remark 5.1.2 is true by definition. Indeed, if \( x \) and \( \xi \) are column vectors of \( \mathbb{R}^n \) then

\[
S = \{0\} \iff \begin{pmatrix}
(P + P^T)x = 0, \quad (P + P^T)Mx = 0, \ldots, \quad (P + P^T)M^{2n-1}x = 0 \\
Q\xi = 0, \quad QM^T\xi = 0, \ldots, \quad Q(M^T)^{2n-1}\xi = 0
\end{pmatrix} \Rightarrow x = 0; \quad \xi = 0
\]

\[
\begin{pmatrix}
(P + P^T)M \\
(P + P^T)M^2 \\
\vdots \\
(P + P^T)M^{2n-1}
\end{pmatrix}
\quad \begin{pmatrix}
Q \\
QM^T \\
\vdots \\
Q(M^T)^{2n-1}
\end{pmatrix}
\]

\[
\begin{pmatrix}
(P + P^T)x = 0 \\
(Q\xi = 0)
\end{pmatrix} \Rightarrow x = 0; \quad \xi = 0
\]

have rank \( n \).
Lemma 5.1.3. Let $A$ and $C$ be $n \times n$ matrices and assume $\text{Rank} C = p$. Then

$$\text{Rank}[C, AC, \ldots, A^{2n-1}C] = n \iff \text{Rank}[C, AC, \ldots, A^{n-p}C] = n.$$ 

The proof of Lemma 5.1.3 can be done repeating the argument of the proof of Corollary 4.1 in [6], so we shall not repeat it here. Let us now observe that, if we define

$$\hat{S} = \bigcap_{j=0}^{n-1} \{(x, \xi) \in \mathbb{R}^{2n} : (P + P^T)M^j x = 0 \quad \text{and} \quad Q(M^j)^T \xi = 0\},$$

then $S = \{0\} \iff \hat{S} = \{0\}$. The implication $\hat{S} = \{0\} \Rightarrow S = \{0\}$ is obvious. On the other hand, assume $S = \{0\}$ then by Remark 5.1.2

$$\begin{align*}
\text{Rank} & \begin{bmatrix} P + P^T \\ (P + P^T)M \\ (P + P^T)M^2 \\ \vdots \\ (P + P^T)M^{2n-1} \end{bmatrix} = n \quad \text{and} \quad \text{Rank} & \begin{bmatrix} Q \\ QM^T \\ Q(M^2)^T \\ \vdots \\ Q(M^{2n-1})^T \end{bmatrix} = n
\end{align*}$$

so, taking the transpose,

$$\text{Rank}[P + P^T, (P + P^T)^T, \ldots, (M^{2n-1})^T(P + P^T)] = n \quad \text{and} \quad (5.1.16)$$

$$\text{Rank}[Q, MQ, \ldots, M^{2n-1}Q] = n. \quad (5.1.17)$$

Applying Lemma 5.1.3 to the matrices in (5.1.16) and (5.1.17), since $Q$ and $P + P^T$ have at least Rank 1, we have that

$$\text{Rank}[P + P^T, (P + P^T)^T, \ldots, (M^{n-1})^T(P + P^T)] = n \quad \text{and} \quad (5.1.18)$$

$$\text{Rank}[Q, MQ, \ldots, M^{n-1}Q] = n. \quad (5.1.19)$$

This implies $\hat{S} = \{0\}$.

Proof of Proposition 5.1.1. Recalling (iii), the hypoellipticity of (5.1.5) and (5.1.6) is equivalent to

$$\text{Rank}[\Sigma, B\Sigma, \ldots, B^{n-1}\Sigma] = n \quad (5.1.20)$$
and

\[ \text{Rank}[\Sigma, B^T \Sigma, \ldots, (B^T)^{n-1} \Sigma] = n, \]  
\[ (5.1.21) \]

respectively. Since \( S = \{0\} \iff \hat{S} = \{0\} \), we need to prove that (5.1.20) is equivalent to (5.1.19) and (5.1.21) is equivalent to (5.1.18). More precisely, we will show the following:

\[ (5.1.20) \iff \text{Rank} \begin{vmatrix} Q & \cdots & Q \end{vmatrix} = n \iff \text{Rank} \begin{vmatrix} Q & \cdots & Q \end{vmatrix} = n \]  
\[ (5.1.22) \]

and

\[ (5.1.21) \iff \text{Rank} \begin{vmatrix} Q & \cdots & Q \end{vmatrix} = n \iff \text{Rank} \begin{vmatrix} (P + P^T) & \cdots & (P + P^T) \end{vmatrix} = n. \]  
\[ (5.1.23) \]

We will show in detail the chain of double implications (5.1.23), the one in (5.1.22) can be proven in an analogous way. By taking the transpose we have that

\[ (5.1.21) \iff \text{Rank} \begin{vmatrix} \Sigma^T \\Sigma^T B \\cdots \\Sigma^T B^{n-1} \end{vmatrix} = n. \]  
\[ (5.1.24) \]

From the fact that \( \Sigma \Sigma^T = Q \), it is easy to see that \( \text{Ker} \Sigma^T = \text{Ker} Q \). \(^3\) This implies that if \( QB^j x = 0 \ \forall j = 0, \ldots, n - 1 \) then also \( \Sigma^T B^j x = 0 \ \forall j = 0, \ldots, n - 1 \) (and viceversa) and hence \( x = 0 \); this means that

\[ \text{Rank} \begin{vmatrix} \Sigma^T \\Sigma^T B \\cdots \\Sigma^T B^{n-1} \end{vmatrix} = n \iff \text{Rank} \begin{vmatrix} Q & \cdots & Q \end{vmatrix} = n. \]  
\[ (5.1.25) \]

Notice that from the steady state variance equation (5.1.8),

\[ (P + P^T) = Q^{-1} \Sigma Q^{-1}. \]  
\[ (5.1.26) \]

\(^3\)The inclusion \( \subseteq \) is obvious. For the other inclusion: if \( x \in \text{Ker} Q \) then \( 0 = \langle Q x, x \rangle = \langle \Sigma \Sigma^T x, x \rangle = \langle \Sigma^T x, \Sigma^T x \rangle \Rightarrow \Sigma^T x = 0.\)
So, if $z$ is a column vector of $\mathbb{R}^n$, then

\[
\begin{align*}
\text{Rank} & \begin{pmatrix} (P + P^T) \\ (P + P^T) B \\ \vdots \\ (P + P^T) B^{n-1} \end{pmatrix} = n \iff \begin{pmatrix} Q^{-1}Q^{-1} \\ Q^{-1}Q^{-1}B \\ \vdots \\ Q^{-1}Q^{-1}B^{n-1} \end{pmatrix} z = 0 \Rightarrow z = 0
\end{align*}
\]

\[
\iff \begin{pmatrix} Q^{-1} \\ Q^{-1}B \\ \vdots \\ Q^{-1}B^{n-1}Q \end{pmatrix} y = 0 \Rightarrow y = 0.
\]

It is a standard fact in control theory that given two matrices $X, Y \in \mathbb{R}^{n \times n}$, $(X, Y)$ is a controllable pair if and only if $(T^{-1}XT, C^{-1}YC)$ is a controllable pair, with $T, C$ any invertible $n \times n$ matrices (see [79], page 22). Using this fact, from the above we have

\[
\begin{align*}
\text{Rank} & \begin{pmatrix} (P + P^T) \\ (P + P^T) B \\ \vdots \\ (P + P^T) B^{n-1} \end{pmatrix} = n \iff \text{Rank} \begin{pmatrix} (P + P^T) \\ (P + P^T) M \\ \vdots \\ (P + P^T) M^{n-1} \end{pmatrix} = n.
\end{align*}
\]

Let us show the implication $\Rightarrow$. So, assume that the left hand side of (5.1.25) holds. Then, for $x \in \mathbb{R}^n$ and setting $\tilde{P} := (P + P^T)$, we have

\[
\begin{align*}
\tilde{P}x = 0 \Rightarrow Q^{-1}Q^{-1}Hx = 0 \Rightarrow Hx = 0.
\end{align*}
\]

Set also $H := Q^{-1}Q^{-1}$. Then from (5.1.24), $\tilde{P} = Q^{-1}Q^{-1}H$ and from (5.1.12) $M = \frac{H}{2} - B$. Suppose there exists $x \in \mathbb{R}^n$ s.t. $\tilde{P}x = 0, \tilde{P}Mx = 0, \ldots, \tilde{P}M^{n-1}x = 0$ then

\[
\begin{align*}
\tilde{P}x = 0 \Rightarrow Q^{-1}Q^{-1}Hx = 0 \Rightarrow Hx = 0,
\end{align*}
\]

\[
\begin{align*}
\tilde{P}Mx = 0 \Rightarrow \tilde{P} \left( \frac{H}{2} - B \right) x = 0 \Rightarrow \tilde{P}Bx = 0.
\end{align*}
\]
Notice that 

\[ \tilde{P} B x = 0 \Rightarrow H B x = 0. \quad (5.1.27) \]

We can repeat the same thing for \( \tilde{P} M^2 \):

\[ \tilde{P} M^2 x = 0 \Rightarrow \tilde{P} M \left( \frac{H}{2} - B \right) x = 0 \quad (5.1.26) \Rightarrow \tilde{P} M B x = 0 \]

\[ \Rightarrow \tilde{P} \left( \frac{H}{2} - B \right) B x = 0 \quad (5.1.27) \Rightarrow \tilde{P} B^2 = 0 \text{ and hence } H B^2 = 0. \quad (5.1.28) \]

Therefore, at each step of the iteration, we have that \( \tilde{P} M^{j-1} = 0 \Rightarrow \tilde{P} B^{j-1} = 0 \) and hence in particular \( H B^{j-1} = 0 \). So at step \( j \) we have

\[ \tilde{P} M^j x = 0 \Rightarrow \tilde{P} \sum_{k=0}^{j} c_{jk} \frac{H^{j-k}}{2^{j-k}} (-1)^k B^k x = 0, \]

where \( c_{jk} \) is the binomial coefficient. Using the fact that \( H B^k x = 0, \forall k \geq 0 \), the above sum becomes \( \tilde{P} B^j = 0 \). Summarizing, \( \forall j = 0, \ldots, n-1 \)

\[ \tilde{P} M^j x = 0 \Rightarrow \tilde{P} B^j x = 0 \]

and hence \( x = 0 \). So the right hand side of (5.1.25) has been proven. The other implication follows with a similar argument so we will not repeat it. This concludes the proof.

5.2 Calculation of the spectrum of \( \mathcal{L} \)

When the assumptions of Proposition 5.1.1 are satisfied, we can apply (2.4.17) to the operator \( \mathcal{H} \) defined in (5.1.4) and obtain that the spectrum of \( \mathcal{L} \) is just

\[ \sigma(\mathcal{L}) = \left\{ \text{Tr} B - \sum_{\substack{\lambda \in \sigma(F) \\text{Im}\lambda > 0 \\text{Tr}\lambda}} (r_{\lambda} + 2k_{\lambda})(-i\lambda), \quad k_{\lambda} \in \mathbb{N} \right\}, \quad (5.2.1) \]

with \( r_{\lambda} \) denoting the algebraic multiplicity of the eigenvalue \( \lambda \). We will show that the eigenvalues of \( F \) with positive imaginary part are of the form \( \frac{i}{2} \mu \), where \( \mu \) is an eigenvalue of \( B \) (recall that by assumption the eigenvalues of \( B \) have strictly positive real part). Hence the spectrum of \( \mathcal{L} \) is completely determined once the eigenvalues of \( B \) are known. This is the content of the following
5.2. Calculation of the spectrum of $L$

Theorem 5.2.1. With the notation introduced above, assume the singular space of $\mathcal{H}$ is zero. Then we know that $L$ is hypoelliptic and we also have

$$\sigma(L) = \left\{ - \sum_{\mu \in \sigma(B)} \mu k_{\mu}, \quad k_{\mu} \in \mathbb{N} \right\}. \quad (5.2.2)$$

This result had already been shown in [53] using different techniques. The remainder of the section is devoted to proving (5.2.2). In the following we shall denote by $\text{diag}\{\alpha_1, \ldots, \alpha_n\}$ the diagonal matrix having $\alpha_1, \ldots, \alpha_n$ as diagonal entries.

Lemma 5.2.2. For any $n \in \mathbb{N}$, let $M_n$ be an $n \times n$ matrix. Consider the $2n \times 2n$ block matrix

$$M_{2n} = \begin{bmatrix} iM_{A_n}/2 & M_{S_n}/2 \\ -M_{S_n}/2 & iM_{A_n}/2 \end{bmatrix},$$

where $M_{S_n}$ and $M_{A_n}$ are the symmetric and antisymmetric part of $M_n$, namely

$$M_{S_n} := \frac{M_n + M_n^T}{2}, \quad M_{A_n} := \frac{M_n - M_n^T}{2}. \quad (5.2.3)$$

Then $\sigma(M_{2n}) = \pm \frac{i}{2} \sigma(M_n)$, i.e. the eigenvalues of $M_{2n}$, counted with their multiplicity, are $\sigma(M_{2n}) = \{ \pm \frac{i}{2} \mu_j, j = 1, \ldots, n \}$ where $\mu_1, \ldots, \mu_n$ are the eigenvalues of $M_n$ (again, counted with their multiplicity).

Proof of Lemma 5.2.2. Suppose first that $M$ is diagonalizable, i.e. $\exists$ an invertible matrix $E$ such that $E^{-1} M E = \Delta$ is diagonal, with $\Delta = \text{diag}\{\mu_1, \ldots, \mu_n\}$. Recall that for a $2n \times 2n$ block matrix (with square blocks) $X = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$, if $A_{22}$ commutes with $A_{12}$ then $\det X = \det(A_{11}A_{22} - A_{21}A_{12})$. Let $E$ be the $2n \times 2n$ block diagonal matrix having $E$ in both the diagonal blocks, i.e.

$$E = \begin{bmatrix} E & 0 \\ 0 & E \end{bmatrix},$$

and let $\tilde{M}_{2n} = E^{-1} M_{2n} E$. Then

$$\det(\tilde{M}_{2n} - \lambda I_{2n}) = \det \begin{bmatrix} -\lambda I_n & \frac{\Delta}{2} \\ -\frac{\Delta}{2} & -\lambda I_n \end{bmatrix} = \Pi_{j=1}^n \left( \lambda^2 + \frac{1}{4} \mu_j^2 \right).$$

So the roots of $\det(\tilde{M}_{2n} - \lambda I_{2n}) = 0$ are $\lambda = \pm \frac{i\mu_j}{2}$. If $M_n$ is not diagonalizable, let $T$ the matrix that brings $M_n$ into its Jordan canonical form, i.e. $T^{-1} M_n T = \tilde{M}_n$ is an upper
5.2. Calculation of the spectrum of $L$

triangular matrix of the form

$$\tilde{M}_n = \begin{vmatrix} \mu_1 & 1 & 0 & \ldots & 0 \\ 0 & \mu_2 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \ldots & 0 & \mu_n \end{vmatrix}, \quad (5.2.4)$$

where the upper diagonal entries $(\tilde{M}_n)_{j,j+1}$, $j = 1 \ldots n-1$, are either 0 or 1, depending on the algebraic and geometric multiplicity of the eigenvalues of $M$ (but not all of them can be 0 otherwise we would be in the case in which $M$ diagonalizes). From (5.2.3) and (5.2.4), $\tilde{M}_{A_n}$ and $\tilde{M}_{S_n}$ are tridiagonal matrices, with $(\tilde{M}_{A_n})_{j,j+1} = 0$ or $1/4$ when $(\tilde{M}_n)_{j,j+1} = 0$ or 1, respectively. Same thing holds for the entries $(\tilde{M}_{S_n})_{j,j+1}$:

$$\tilde{M}_{A/2} = \begin{vmatrix} 0 & 1/4 & 0 & \ldots & 0 \\ -1/4 & 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ldots & 0 \end{vmatrix}, \quad \tilde{M}_{S/2} = \begin{vmatrix} \frac{\mu_1}{2} & 1/4 & 0 & \ldots & 0 \\ 1/4 & \frac{\mu_2}{2} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ldots & 0 \end{vmatrix}$$

Denote by $T$ the $2n \times 2n$ block diagonal matrix having $T$ in both the diagonal blocks. The similarity transformation $T^{-1}\tilde{M}_{2n}T = \tilde{M}_{2n}$ gives

$$\tilde{M}_{2n} = \begin{vmatrix} i\tilde{M}_{A_n}/2 & \tilde{M}_{S_n}/2 \\ -\tilde{M}_{S_n}/2 & i\tilde{M}_{A_n}/2 \end{vmatrix} \quad (5.2.5)$$

and we want to show that the roots of

$$p_\lambda(\tilde{M}_{2n}) := \det(\tilde{M}_{2n} - \lambda I_{2n}),$$

counted with their multiplicity, are $\pm \frac{i\mu_j}{2}$, $j = 1, \ldots, n$. We shall prove this fact by induction on $n$. For $n = 1$ this is trivially true but of no relevance for the problem at hand (there is no hypoellipticity in one dimension). So let us look at the case $n = 2$. If $M_2$ is a $2 \times 2$ matrix that does not diagonalize, its canonical Jordan form can only be

$$\tilde{M}_2 = \begin{vmatrix} \mu & 1 \\ 0 & \mu \end{vmatrix},$$
i.e. $M_2$ has one eigenvalue of algebraic multiplicity 2 and geometric multiplicity 1. Hence

$$p_\lambda(\tilde{M}_4) = \det \begin{vmatrix} -\lambda & \frac{i}{4} & \frac{\mu}{2} & \frac{1}{4} \\ \frac{i}{4} & -\lambda & \frac{\mu}{2} & \frac{1}{4} \\ -\frac{\mu}{2} & -\frac{1}{4} & -\lambda & i \frac{1}{4} \\ -\frac{i}{4} & -\frac{\mu}{2} & -i \frac{1}{4} & -\lambda \end{vmatrix} = \left(\lambda^2 + \frac{\mu^2}{4}\right)^2 = 0 \Leftrightarrow \lambda = \pm \frac{\mu}{2}.$$ 

Now assume this is true for $n$ and let us prove it for $n + 1$. If $\tilde{M}_{2n}$ is of the form (5.2.5), then $\tilde{M}_{2(n+1)} - \lambda I_{2(n+1)}$ can be either of the form

$$\tilde{M}_{2(n+1)} - \lambda I_{2(n+1)} = \begin{vmatrix} -\lambda & 0 & 0 \cdots 0 & \frac{\mu_1}{2} & 0 & 0 \cdots 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & i\tilde{M}_A - \lambda I_n & 0 & -\lambda & 0 & 0 \cdots 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ -\frac{\mu_1}{2} & 0 & 0 \cdots 0 & -\lambda & 0 & 0 \cdots 0 \\ 0 & i\tilde{M}_A - \lambda I_n & 0 & -\lambda & 0 & 0 \cdots 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots \end{vmatrix} \quad (5.2.6)$$

or of the form

$$\tilde{M}_{2(n+1)} - \lambda I_{2(n+1)} = \begin{vmatrix} -\lambda & \frac{i}{4} & 0 \cdots 0 & \frac{\mu_1}{2} & \frac{1}{4} & 0 \cdots 0 \\ \frac{i}{4} & -\lambda & 0 \cdots 0 & \frac{\mu_1}{2} & \frac{1}{4} & 0 \cdots 0 \\ 0 & i\tilde{M}_A - \lambda I_n & 0 & -\lambda & i \frac{1}{4} & 0 \cdots 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & -\lambda & -\frac{i}{4} & 0 \cdots 0 \\ 0 & i\tilde{M}_A - \lambda I_n & 0 & -\lambda & -\frac{i}{4} & 0 \cdots 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots \end{vmatrix} \quad (5.2.7)$$
for less than reordering. In the first case (5.2.6), by swapping columns \( n \) times and rows \( n \) times, we obtain

\[
p_\lambda(\tilde{M}_{2(n+1)}) = \det \begin{bmatrix} -\lambda & \mu_1 & 0 \ldots & 0 & 0 \ldots & 0 \\ -\mu_1 & -\lambda \end{bmatrix} = (\lambda^2 + \mu_1^2/4) p_\lambda(\tilde{M}_n),
\]

hence the statement follows by the inductive assumption. In the second case (5.2.7), by swapping columns and rows as before we end up with

\[
p_\lambda(\tilde{M}_{2(n+1)}) = \det \begin{bmatrix} -\lambda & \mu_1 & 0 \ldots & 0 & 0 \ldots & 0 \\ -\mu_1 & -\lambda \end{bmatrix} = (\lambda^2 + \mu_1^2/4) p_\lambda(\tilde{M}_n),
\]

If \( X \) is a matrix, we shall denote by \( X^\dagger \) the matrix obtained from \( X \) by eliminating the first column. So, if we expand the expression for the determinant by always using the first row, we obtain
5.2. Calculation of the spectrum of $\mathcal{L}$

\[
p_{\lambda}(\tilde{\mathcal{M}}_{2(n+1)}) = \left(\lambda^2 + \frac{\mu_1^2}{4}\right) p_{\lambda}(\tilde{\mathcal{M}}_{2n}) - \left(\frac{\lambda}{4} + \frac{i\mu_1}{8}\right) \det \begin{vmatrix}
\frac{1}{4} & 0 & \cdots & 0 & \left(\tilde{M}_A - \lambda I_n\right)^{1} & \tilde{M}_S_n \\
0 & \ddots & \ddots & \ddots & \vdots & \vdots \\
0 & \ddots & \ddots & \ddots & \vdots & \vdots \\
0 & \ddots & \ddots & \ddots & \vdots & \vdots \\
-\frac{i}{4} & 0 & \cdots & 0 & \left(-\tilde{M}_S\right)^{1} & \tilde{M}_A - \lambda I_n
\end{vmatrix}
\]

\[
+ \left(\frac{i\lambda}{4} - \frac{\mu_1}{8}\right) \det \begin{vmatrix}
-\frac{1}{4} & 0 & \cdots & 0 & \left(\tilde{M}_A - \lambda I_n\right)^{1} & \tilde{M}_S_n \\
0 & \ddots & \ddots & \ddots & \vdots & \vdots \\
0 & \ddots & \ddots & \ddots & \vdots & \vdots \\
0 & \ddots & \ddots & \ddots & \vdots & \vdots \\
-\frac{1}{4} & 0 & \cdots & 0 & \left(-\tilde{M}_S\right)^{1} & \tilde{M}_A - \lambda I_n \\
0 & \ddots & \ddots & \ddots & \vdots & \vdots \\
0 & \ddots & \ddots & \ddots & \vdots & \vdots \\
0 & \ddots & \ddots & \ddots & \vdots & \vdots \\
0 & \ddots & \ddots & \ddots & \vdots & \vdots \\
0 & \ddots & \ddots & \ddots & \vdots & \vdots \\
\end{vmatrix}
\]

\[
\mp \left(\frac{i\lambda}{4} - \frac{\mu_1}{8}\right) \det \begin{vmatrix}
\frac{1}{4} & 0 & \cdots & 0 & \tilde{M}_A - \lambda I_n & \left(\tilde{M}_S\right)^{1} \\
0 & \ddots & \ddots & \ddots & \vdots & \vdots \\
0 & \ddots & \ddots & \ddots & \vdots & \vdots \\
0 & \ddots & \ddots & \ddots & \vdots & \vdots \\
-\frac{1}{4} & 0 & \cdots & 0 & -\tilde{M}_S & \tilde{M}_A - \lambda I_n \\
0 & \ddots & \ddots & \ddots & \vdots & \vdots \\
0 & \ddots & \ddots & \ddots & \vdots & \vdots \\
0 & \ddots & \ddots & \ddots & \vdots & \vdots \\
0 & \ddots & \ddots & \ddots & \vdots & \vdots \\
\end{vmatrix}
\]
5.2. Calculation of the spectrum of $L$

\[ \pm \left( \frac{\lambda}{4} + \frac{i \mu_1}{8} \right) \det \left( \begin{array}{c|c} -\frac{i}{4} & \tilde{M}_{A_n} - \lambda I_n \\ \vdots & \left( \tilde{M}_{S_n} \right)^{-1} \\ 0 & \left( \tilde{M}_{A_n} - \lambda I_n \right)^{-1} \end{array} \right), \tag{5.2.11} \]

where the signs $\mp$ are such that the top sign is referred to the case in which $n$ is odd and the bottom one is refers to the case in which $n$ is even. Now notice that the matrix in (5.2.9) is obtained from the matrix in (5.2.8) by multiplying the first column of the latter by $-i$ and the matrix in (5.2.11) is obtained in the same way from the matrix in (5.2.10). Hence the determinant of the matrix in (5.2.9) (in (5.2.11), respectively) is just $-i$ times the determinant of the matrix in (5.2.8) (in (5.2.10), respectively). Therefore we can eliminate the last four addends in the expression for the characteristic polynomial of $p_{\lambda}(\tilde{M}_{2(n+1)})$ and we are left with

\[ p_{\lambda}(\tilde{M}_{2(n+1)}) = \left( \lambda^2 + \frac{\mu_1^2}{4} \right) p_{\lambda}(\tilde{M}_{2n}), \]

which concludes the proof of the Lemma.

Proof of Theorem 5.2.1. Now a similarity transformation shows that we can apply Lemma 5.2.2 to the Hamilton map $F$, defined in (5.1.15). Indeed, using (5.1.8) and (5.1.12), we can rewrite $F$ as

\[ F = \frac{1}{4} \begin{vmatrix} i(B - Q_\infty B^T Q_\infty^{-1}) & 2(BQ_\infty + Q_\infty B^T) \\ -\frac{1}{2}(Q_\infty^{-1}B + B^T Q_\infty^{-1}) & i(Q_\infty^{-1}BQ_\infty - B^T) \end{vmatrix}. \]

Let $G$ be the $2n \times 2n$ block matrix

\[ G = \begin{vmatrix} \frac{1}{\sqrt{2}} Q_\infty^{-1/2} & 0 \\ 0 & \sqrt{2} Q_\infty^{1/2} \end{vmatrix}. \]

If we denote $\hat{B} = Q_\infty^{-1/2} BQ_\infty^{1/2}$ and let $\hat{B}_A$ and $\hat{B}_S$ be the antisymmetric and symmetric part of $\hat{B}$, respectively, we have

\[ \hat{F} := G^{-1} F G = \begin{vmatrix} i \hat{B}_A/2 & \hat{B}_S/2 \\ \hat{B}_S/2 & i \hat{B}_A/2 \end{vmatrix}. \]

\[ ^4 \text{Thanks to Dr. G. Moore for suggesting this transformation.} \]
5.3 Example

A simple example on which to carry out explicit computations is the system that we studied in Chapter 3, which we recast here for the reader’s convenience.

\[
\begin{align*}
\frac{dx}{dt} &= y dt \\
\frac{dy}{dt} &= -\nabla_x V(x) dt + \sum_{j=1}^{m} \lambda_j z_j dt \\
\frac{dz_j}{dt} &= - (\lambda_j y + \alpha_j z_j) dt + \sqrt{2\alpha_j} dW_j, \quad j = 1, \ldots, d,
\end{align*}
\]

(5.3.1)

with \((x, y, z) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^d, z = (z_1, \ldots, z_d), z_j \in \mathbb{R}^d\) and \((W_j)_{j=1}^d\) independent standard Brownian motions. The density of the invariant measure associated to the dynamics (5.3.1) is

\[
\rho(x, y, z) = \frac{1}{Z} e^{-\frac{1}{2} (V(x) + y_2 + |z|^2)},
\]

with \(Z\) being a normalizing constant and \(|\cdot|\) the Euclidean norm on \(\mathbb{R}^d\). We also recall that the generator of system (5.3.1) is

\[
L = y \partial_x - \partial_x V(x) \partial_y + \left( \sum_{j=1}^{m} \lambda_j z_j \right) \partial_y - y \sum_{j=1}^{m} \lambda_j \partial z_j - \sum_{j=1}^{m} \alpha_j z_j \partial z_j + \sum_{j=1}^{m} \alpha_j \partial z_j^2,
\]

(5.3.2)

and its \(L^2\)-adjoint is given by

\[
L' = -y \partial_x + \partial_x V(x) \partial_y - \left( \sum_{j=1}^{m} \lambda_j z_j \right) \partial_y + y \sum_{j=1}^{m} \lambda_j \partial z_j + \sum_{j=1}^{m} \alpha_j \partial z_j (z_j') + \sum_{j=1}^{m} \alpha_j \partial z_j^2.
\]

Let \(V(x)\) be a quadratic potential, \(V(x) = \omega^2 \frac{x^2}{2}\). Under the transformation \(\sqrt{\rho^{-1}} L' (\sqrt{\rho})\), we obtain the operator

\[
\mathcal{H}' = -y \partial_x + \omega^2 x \partial_y - \left( \sum_{j=1}^{m} \lambda_j z_j \right) \partial_y + y \sum_{j=1}^{m} \lambda_j \partial z_j - \sum_{j=1}^{m} \alpha_j \frac{z_j^2}{4} + \sum_{j=1}^{m} \alpha_j \partial z_j^2 + \frac{1}{2} \sum_{j=1}^{m} \alpha_j.
\]

Let us now consider the operator

\[
\mathcal{H} = -\mathcal{H}' + \sum_{j=1}^{m} \alpha_j.
\]

(5.3.3)

The symbol of the operator \(\mathcal{H}\) is given by

\[
q(x, y, z, \xi, \eta, \zeta) = i(y \xi - \omega^2 x \eta) + i \sum_{j=1}^{m} (\lambda_j z_j \eta - y \lambda_j \xi_j) + \sum_{j=1}^{m} (\alpha_j \frac{z_j^2}{4} + \alpha_j \zeta_j^2),
\]
so that \( \text{Re}(q) \geq 0 \). The polarized form of \( q \) is
\[
\tilde{q} = \sum_{j=1}^{m} \alpha_j \left( \frac{z_j \tilde{z}_j}{4} + \zeta_j \tilde{\zeta}_j \right) \\
+ \frac{i}{2} \left( y \tilde{\xi} + \tilde{y} \xi - \omega^2 x \tilde{\eta} - \omega^2 \tilde{x} \eta \right) \\
+ \frac{i}{2} \sum_{j=1}^{m} \lambda_j \left( z_j \tilde{\eta} + \tilde{z}_j \eta - y \tilde{\zeta}_j - \tilde{y} \zeta_j \right).
\] (5.3.4)

The Hamilton map associated with (5.3.4) is then a \( 2(2 + d) \times 2(2 + d) \) matrix, namely
\[
F = \begin{bmatrix}
E & \tilde{D} \\
-\tilde{D}/4 & -E^T
\end{bmatrix},
\]
where \( B \) and \( \tilde{D} \) are \( (2 + d) \times (2 + d) \) matrices,
\[
E = \begin{bmatrix}
0 & \frac{i}{2} & 0 & \cdots & 0 \\
-\frac{\omega^2}{2} & 0 & \frac{i \lambda_1}{2} & \cdots & \frac{i \lambda_m}{2} \\
0 & -\frac{i \lambda_1}{2} & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & -\frac{i \lambda_m}{2} & 0 & \cdots & 0
\end{bmatrix}, \\
\tilde{D} = \begin{bmatrix}
0 & 0 & \cdots & 0 \\
0 & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & D \\
0 & 0 & \cdots & 0
\end{bmatrix},
\] (5.3.5)

with \( D \) an \( d \times d \) diagonal matrix with \( D_{jj} = \alpha_j \). It is easy to see that
\[
\text{Ker} \left( \text{Re}(F) \right) \cap \mathbb{R}^{2(2+d)} = \left\{ (x, y, z, \xi, \eta, \zeta) \in \mathbb{R}^{2(2+d)} : \bar{z} = \bar{\eta} = 0 \right\}, \quad \text{(5.3.6)}
\]
\[
\text{Ker} \left( \text{Re}(F) \text{Im}(F) \right) \cap \mathbb{R}^{2(2+d)} = \left\{ (x, y, z, \xi, \eta, \zeta) \in \mathbb{R}^{2(2+d)} : y = \eta = 0 \right\}, \quad \text{(5.3.7)}
\]
\[
\text{Ker} \left( \text{Re}(F) \text{Im}(F)^2 \right) \cap \mathbb{R}^{2(2+d)} = \left\{ (x, y, z, \xi, \eta, \zeta) \in \mathbb{R}^{2(2+d)} : x = \xi = 0 \right\}, \quad \text{(5.3.8)}
\]
so the singular space \( S \) is just zero, \( S = \{0\} \). When \( d = 1 \) the roots of the characteristic polynomial of \( F \) are explicitly computable so formula (2.4.17) gives an explicit description of the spectrum.
Chapter 6

Short Time Behaviour of Hypoelliptic Markov Semigroups

In Theorem 3.1.6 we proved a short time estimate for the Markov semigroup generated by an operator of the form $A^*A + B$, with $B$ an antisymmetric operator. This theorem is a particular case of Theorem 2.3.5. In the proof, we heavily used the antisymmetry of the operator $B$ in $L^2_\rho$, which helps getting rid of many terms that should otherwise be controlled in some way. However, differential operators on $\mathbb{R}^m$ of the form $A^*A + B$ can be recast in Hörmander sums of squares forms; more precisely, suppose $A = a(x)\nabla$ and $B = b(x)\nabla$, then

$$L = A^*A + B = -A^2 + (B + cA),$$

where $c$ is a coefficient that depends on the derivatives of $a(x)$ and on the invariant measure. So, if we are able to prove some bounds for operators in the form $A^*A + B$, with $B$ antisymmetric, we should be able to prove the same bounds for operators of the form $Z_0^2 + B$, with $B$ not (necessarily) antisymmetric. Clearly, in the latter case, the $L^2_\rho$ setting is completely unessential. In this Chapter we will focus on operators in Hörmander’s sum of squares, and we will obtain pointwise results in the spirit of Theorem 2.3.5. In particular, the trick of using a Lyapunov function (an auxiliary time-dependent scalar product) which contains lower order mixed terms, still works when supplemented with few basic tools from semigroup theory.
6.1 Assumptions and statement of main result

Consider the second order differential operator on \( \mathbb{R}^m \):

\[
\mathcal{L} = Z_0^2 + B. \tag{6.1.1}
\]

Assume that for some \( N \in \mathbb{N}, N \geq 1 \), there exist \( Z_1, \ldots, Z_N \) such that:

\[
[B, Z_j] = Z_{j+1} \quad j = 0, \ldots, N - 1. \tag{6.1.2}
\]

This filiform algebra model is analogous to the one considered in Chapter 3. Also, assume that

\[
[B, Z_N] = \sum_{j=0}^{N} \alpha_j Z_j, \quad \text{for some } \{ \alpha_j \}_{0 \leq j \leq N}, \alpha_j \in \mathbb{R} \tag{6.1.3}
\]

and

\[
[Z_i, Z_j] = \sum_{h=0}^{N} \beta_{(ij),h} Z_h, \quad 0 \leq i, j \leq N, \{ \beta_{(ij),h} \}_{0 \leq h \leq N}, \beta_{(ij),h} \in \mathbb{R}. \tag{6.1.4}
\]

In the following \( \| \cdot \|_\infty \) indicates the supremum norm and we will use the notation \( \mathcal{P}_t := e^{t\mathcal{L}} \) for the Markov semigroup generated by \( \mathcal{L} \).

**Proposition 6.1.1.** For \( t > 0 \), let \( f_t = e^{t\mathcal{L}} f_0 \) and assume (6.1.2), (6.1.3) and (6.1.4). Then, along the semigroup we have that \( \forall n \geq 1 \) and \( k_1, \ldots, k_n \in \{0, \ldots, N\} \), there exists a constant \( c \) (that possibly depends on \( n \) and on \( k_1, \ldots, k_n \) but it is independent of \( t \)) such that

\[
|Z_{k_1} \cdot \ldots \cdot Z_{k_n} f_t|^2 \leq \frac{c}{t^2(\sum_{j=1}^{n} k_j) + n} \|f_0\|_\infty^2, \quad \forall t > 0. \tag{6.1.5}
\]

**Proof.** In order to streamline the proof we assume, for the time being, that

\[
[B, Z_N] = 0. \tag{6.1.6}
\]

For the same reason we also assume

\[
[Z_j, Z_0] = 0 \quad \forall j = 0, \ldots, N, \tag{6.1.7}
\]

and

\[
[Z_i, Z_j] = \sum_{h=0}^{N} \beta_{(ij),h} Z_h, \quad 1 \leq i, j \leq N, \{ \beta_{(ij),h} \}_{0 \leq h \leq N}, \beta_{(ij),h} \in \mathbb{R}. \tag{6.1.8}
\]

At the end of the proof we will explain how to remove assumptions (6.1.6) and (6.1.7) and obtain the result of Proposition 6.1.1 when we assume (6.1.3) and (6.1.4), instead.
For some strictly positive constants $a_j, b_j$ and $d$ to be determined, define

\[ \Gamma_t^{(0)} f_t = d|f_t|^2, \]

\[ \Gamma_t^{(1)} f_t = \sum_{j=0}^{N} a_j t^{2j+1} |Z_j f_t|^2 + \sum_{j=0}^{N-1} b_j t^{2j+2} (Z_j f_t)(Z_{j+1} f_t) + d|f_t|^2, \]

and, for $n \geq 1,$

\[ \Gamma_t^{(n+1)} f_t = \Gamma_t^{(n)} f_t + \sum_{k_1, \ldots, k_{n+1}=0}^{N} a_{k_1, \ldots, k_{n+1}} t^{2(\sum_{j=1}^{n+1} k_j) + (n+1)} |Z_{k_1} \cdots Z_{k_{n+1}} f_t|^2 + \sum_{k_2, \ldots, k_{n+1}=0}^{N} \sum_{k_1=1}^{N} b_{k_1, \ldots, k_{n+1}} t^{2(\sum_{j=1}^{n+1} k_j) + n} (Z_{k_1-1} Z_{k_2} \cdots Z_{k_{n+1}} f_t)(Z_{k_1} \cdots Z_{k_{n+1}} f_t). \]

Strategy of proof: if we can prove that

\[ \int_0^t \frac{d}{ds} \mathcal{P}_{t-s} \left( \Gamma_s^{(n)} f_s \right) \, ds < 0 \quad \forall n \geq 0, \]

then, from the fundamental theorem of calculus,

\[ \mathcal{P}_0 \left( \Gamma_t^{(n)} f_t \right) = \Gamma_t^{(n)} f_t < \mathcal{P}_t \left( \Gamma_0^{(n)} f_0 \right) \leq \| \Gamma_0^{(n)} f_0 \|_\infty = d \| f_0 \|_\infty \]

hence

\[ \Gamma_t^{(n)} f_t < d \| f_0 \|_\infty. \quad (6.1.9) \]

For an appropriate choice of the constants that appear in the definition of $\Gamma_t^{(n)},$ (6.1.9) implies (6.1.5). Because

\[ \frac{d}{ds} \mathcal{P}_{t-s} \left( \Gamma_s^{(n)} f_s \right) = \mathcal{P}_{t-s} \left( -\mathcal{L} \Gamma_s^{(n)} f_s + \frac{d}{ds} \Gamma_s^{(n)} f_s \right), \]

and the semigroup $\mathcal{P}_t$ preserves positivity, the whole thing boils down to proving that for $n = 0$ there exists a constant $d > 0$ s.t. \(-\mathcal{L} + \frac{\partial}{\partial t} \Gamma_t^{(0)} f_t < 0, \forall t > 0 \) and that $\forall n \geq 1$ there exist strictly positive constants $\{a_{k_1}\}_{0 \leq k_1 \leq \infty}, \{a_{k_1, k_2}\}_{0 \leq k_1, k_2 \leq \infty}, \ldots, \{a_{k_1, \ldots, k_n}\}_{0 \leq k_1, \ldots, k_n \leq \infty, d}$ and $\{b_{k_1}\}_{0 \leq k_1 \leq \infty}, \{b_{k_1, k_2}\}_{0 \leq k_1, k_2 \leq \infty}, \ldots, \{b_{k_1, \ldots, k_n}\}_{0 \leq k_1, \ldots, k_n \leq \infty}$ such that

\[ \left( -\mathcal{L} + \frac{\partial}{\partial t} \right) \left( \Gamma_t^{(n)} f_t \right) < 0, \quad \forall t > 0. \quad (6.1.10) \]

We will prove (6.1.10) by induction on $n.$ The inductive basis, i.e. the proof that for $n = 0$ there exists $d > 0$ such that \(-\mathcal{L} + \frac{\partial}{\partial t} \Gamma_t^{(0)} f_t < 0, \forall t > 0, \) is straightforward. Indeed

\[ \left( -\mathcal{L} + \frac{\partial}{\partial t} \right) |f_t|^2 = -2|Z_0 f_t|^2 - B|f_t|^2 + 2f_t \mathcal{L} f_t = -2|Z_0 f_t|^2 < 0, \]
where we simply used the fact that $Z_0^2$ is a second order differential operator and $B$ is a first order differential operator. Assuming that for $n \geq 1$ there exist strictly positive constants 
\[ \{a_{k_1} \}_{0 \leq k_1 \leq N}, \{a_{k_1,k_2} \}_{0 \leq k_1,k_2 \leq N}, \ldots, \{a_{k_1,\ldots,k_n} \}_{0 \leq k_1,\ldots,k_n \leq n} \] and 
\[ \{b_{k_1} \}_{0 \leq k_1 \leq N}, \{b_{k_1,k_2} \}_{0 \leq k_1,k_2 \leq N}, \ldots, \{b_{k_1,\ldots,k_n} \}_{0 \leq k_1,\ldots,k_n \leq n} \] such that 
\[ (-L + \frac{\partial}{\partial t}) \left( \Gamma_{t}^{(n)} f_t \right) < 0, \quad \forall t > 0; \] we want to prove by induction that there exist strictly positive constants 
\[ \{a_{k_1,\ldots,k_n+1} \}_{0 \leq k_j \leq N} \] and 
\[ \{b_{k_1,\ldots,k_n+1} \}_{0 \leq k_j \leq N} \] such that 
\[ (-L + \frac{\partial}{\partial t}) \left( \Gamma_{t}^{(n+1)} f_t \right) < 0, \quad \forall t > 0. \]

To this end, let us observe the two following elementary facts:

\[
Z_0^2 |Z_{k_1} \cdot \ldots \cdot Z_{k_{n+1}} f_t|^2 = 2 \left( Z_0^2 Z_{k_1} \cdot \ldots \cdot Z_{k_{n+1}} f_t \right) \left( Z_{k_1} \cdot \ldots \cdot Z_{k_{n+1}} f_t \right) \\
+ 2 |Z_0 Z_{k_1} \cdot \ldots \cdot Z_{k_{n+1}} f_t|^2 \tag{6.1.11}
\]

and

\[
(Z_{k_1} \cdot \ldots \cdot Z_{k_{n+1}} L) = \left[ Z_{k_1} \cdot \ldots \cdot Z_{k_{n+1}}, L \right] + L Z_{k_1} \cdot \ldots \cdot Z_{k_{n+1}}. \tag{6.1.12}
\]

Recalling that for any three operators $X,Y$ and $W$,

\[
[X,Y,W] = X[Y,W] + [X,W]Y, \tag{6.1.13}
\]

from (6.1.2) we have that if $k_1 \neq N$ then

\[
[Z_{k_1} \cdot \ldots \cdot Z_{k_{n+1}}, B] = Z_{k_1} [Z_{k_2} \cdot \ldots \cdot Z_{k_{n+1}}, B] - Z_{k_1+1} Z_{k_2} \cdot \ldots \cdot Z_{k_{n+1}}. \tag{6.1.14}
\]

We denote $k = (k_1, \ldots, k_{n+1})$, with $k_j \in \{0,1,\ldots,N\}$, $Z_k = (Z_{k_1} \cdot \ldots \cdot Z_{k_{n+1}})$ and $e_j$ the $j$th vector of the canonical basis of $\mathbb{R}^{n+1}$. So iterating (6.1.14) and recalling (6.1.6), we obtain

\[
[Z_{k_1} \cdot \ldots \cdot Z_{k_{n+1}}, B] = - \sum_{j=1}^{n+1} Z_{k+e_j}. \tag{6.1.15}
\]

Because we are assuming that $(-L + \frac{\partial}{\partial t}) \left( \Gamma_{t}^{(n)} f_t \right) < 0$ for some appropriate choice of the constants, we want to look at

\[
\left(-L + \frac{\partial}{\partial t}\right) \left( \Gamma_{t}^{(n+1)} f_t - \Gamma_{t}^{(n)} f_t \right)
\]

\[
= \left(-L + \frac{\partial}{\partial t}\right) \sum_{k_1,\ldots,k_{n+1}=0}^{N} a_{k_1,\ldots,k_{n+1}} t^{2(\sum_{j=1}^{n+1} k_j) + (n+1)} |Z_{k_1} \cdot \ldots \cdot Z_{k_{n+1}} f_t|^2
\]

\[
+ \left(-L + \frac{\partial}{\partial t}\right) \sum_{k_2,\ldots,k_{n+1}=0}^{N} \sum_{k_1=1}^{N} b_{k_1,\ldots,k_{n+1}} t^{2(\sum_{j=1}^{n+1} k_j) + n} (Z_{k_1-1} Z_{k_2} \cdot \ldots \cdot Z_{k_{n+1}} f_t)(Z_{k_1} \cdot \ldots \cdot Z_{k_{n+1}} f_t).
\]
6.1. Assumptions and statement of main result

Using (6.1.1), (6.1.11) and (6.1.12), we obtain

\[
-L + \frac{\partial}{\partial t} \left( \Gamma^{(n+1)} f_t - \Gamma^{(n)} f_t \right) = \\
- \sum_{k_1, \ldots, k_{n+1} = 0}^{N} 2a_{k_1, \ldots, k_{n+1}} t^{2(\sum_{j=1}^{n} k_j) + (n+1)} |Z_0 Z_{k_1} \cdot \cdot \cdot Z_{k_{n+1}} f_t|^2 \\
- \sum_{k_2, \ldots, k_{n+1} = 0}^{N} \sum_{k_1 = 1}^{N} 2b_{k_1, \ldots, k_{n+1}} t^{2(\sum_{j=1}^{n} k_j) + n} (Z_0 Z_{k_1-1} \cdot \cdot \cdot Z_{k_{n+1}} f_t) (Z_0 Z_{k_1} \cdot \cdot \cdot Z_{k_{n+1}} f_t) \\
+ \sum_{k_1, \ldots, k_{n+1} = 0}^{N} a_{k_1, \ldots, k_{n+1}} \left[ 2 \left( \sum_{j=1}^{n+1} k_j \right) + (n+1) \right] t^{2(\sum_{j=1}^{n+1} k_j) + n} |Z_{k_1} \cdot \cdot \cdot Z_{k_{n+1}} f_t|^2 \\
+ \sum_{k_2, \ldots, k_{n+1} = 0}^{N} \sum_{k_1 = 1}^{N} b_{k_1, \ldots, k_{n+1}} \left[ 2 \left( \sum_{j=1}^{n+1} k_j \right) + n \right] t^{2(\sum_{j=1}^{n+1} k_j) + (n-1)} (Z_{k_1-1} Z_{k_2} \cdot \cdot \cdot Z_{k_{n+1}} f_t) (Z_{k_1} \cdot \cdot \cdot Z_{k_{n+1}} f_t) \\
+ \sum_{k_1, \ldots, k_{n+1} = 0}^{N} 2a_{k_1, \ldots, k_{n+1}} t^{2(\sum_{j=1}^{n+1} k_j) + (n+1)} (Z_{k_1} \cdot \cdot \cdot Z_{k_{n+1}} f_t) [Z_{k_1} \cdot \cdot \cdot Z_{k_{n+1}}, L] f_t \\
+ \sum_{k_2, \ldots, k_{n+1} = 0}^{N} \sum_{k_1 = 1}^{N} b_{k_1, \ldots, k_{n+1}} t^{2(\sum_{j=1}^{n+1} k_j) + n} \{ (Z_{k_1} \cdot \cdot \cdot Z_{k_{n+1}} f_t) [Z_{k_1-1} Z_{k_2} \cdot \cdot \cdot Z_{k_{n+1}}, L] f_t \\
+ (Z_{k_1-1} Z_{k_2} \cdot \cdot \cdot Z_{k_{n+1}} f_t) [Z_{k_1} \cdot \cdot \cdot Z_{k_{n+1}}, L] f_t \}.
\]


In the following we will be referring to terms of the form \((Z_{k_1} \cdot \cdot \cdot Z_{k_{n+1}} f_t)\) and \((Z_0 Z_{k_1} \cdot \cdot \cdot Z_{k_{n+1}} f_t)\) as terms of length \(n+1\) and terms of length \(n+2\) starting with \(Z_0\), respectively. These terms are the only ones that we need to control.


\[ |xy| \leq \frac{|x|^2}{2\delta} + \frac{\delta |y|^2}{2}, \quad \forall x, y \in \mathbb{R}, \quad \delta > 0, \]

which we shall systematically use in the following. In particular we will choose \(\delta\) to be a constant times an appropriate positive power of \(t\). Also, from now on \(C\) will be a generic
constant, depending on $n$ and on $\sum_j k_j$ but not on $t$ or on the constants $a_{k_1,\ldots,k_{n+1}}$ and $b_{k_1,\ldots,k_{n+1}}$.

\[ I \leq \sum_{k_1,\ldots,k_{n+1}=0}^{N} a_{k_1,\ldots,k_{n+1}} t^{2(\sum_{j=1}^{n+1} k_j)+(n+1)} \left( -2 |Z_0 Z_{k_1} \cdots Z_{k_{n+1}} f_t|^2 \right) \]  
(6.1.22)

\[ + \sum_{k_2,\ldots,k_{n+1}=0}^{N} \sum_{k_1=1}^{N} 2b_{k_1,\ldots,k_{n+1}} \left( b_{k_1,\ldots,k_{n+1}} t^{2(\sum_{j=1}^{n+1} k_j)+(n+1)} |Z_0 Z_{k_1-1} \cdots Z_{k_{n+1}} f_t|^2 \right) \]  
(6.1.23)

\[ + t^{2(\sum_{j=1}^{n+1} k_j)+(n+1)} \frac{|Z_0 Z_{k_1} \cdots Z_{k_{n+1}} f_t|^2}{b_{k_1,\ldots,k_{n+1}}} \]

We look separately at the terms with $k_1 = 0$ and at the terms with $k_1 > 0$. In doing so, we need to notice that terms of the form $|Z_0 Z_{k_2} \cdots Z_{k_{n+1}} f_t|$ (i.e. those with $k_1 = 0$) come from (6.1.22) when $k_1 = 0$ but also from (6.1.23) when $k_1 = 1$. Hence

\[ I \leq C \sum_{k_2,\ldots,k_{n+1}=0}^{N} \left( -a_{0,k_2,\ldots,k_{n+1}} + b_{1,k_2,\ldots,k_{n+1}} \right) t^{2(\sum_{j=1}^{n+1} k_j)+(n+1)} |Z_0 Z_{k_2} \cdots Z_{k_{n+1}} f_t|^2 \]  
(6.1.24)

\[ + C \sum_{k_2,\ldots,k_{n+1}=0}^{N} \sum_{k_1=1}^{N} \left( -a_{1,k_2,\ldots,k_{n+1}} + b_{2,k_1+1,k_2,\ldots,k_{n+1}} + 1 \right) t^{2(\sum_{j=1}^{n+1} k_j)+(n+1)} |Z_0 Z_{k_1} \cdots Z_{k_{n+1}} f_t|^2, \]  
(6.1.25)

with the understanding that $b_{k_j+1} = 0$ when $k_j = N$. We repeat the same kind of procedure for $[II]$, applying first Young’s inequality and then looking separately at the two cases $k_1 = 0$ and $k_1 > 0$.

\[ [II] \leq \sum_{k_1,\ldots,k_{n+1}=0}^{N} a_{k_1,\ldots,k_{n+1}} \left[ 2 \left( \sum_{j=1}^{n+1} k_j \right) + n + 1 \right] t^{2(\sum_{j=1}^{n+1} k_j)+(n+1)} |Z_{k_1} \cdots Z_{k_{n+1}} f_t|^2 \]  
(6.1.26)

\[ + \sum_{k_2,\ldots,k_{n+1}=0}^{N} \sum_{k_1=1}^{N} b_{k_1,\ldots,k_{n+1}} \left[ 2 \left( \sum_{j=1}^{n+1} k_j \right) + n + 1 \right] t^{2(\sum_{j=1}^{n+1} k_j)+(n+1)} \frac{|Z_{k_1} \cdots Z_{k_{n+1}} f_t|^2}{4} \]  
(6.1.27)

\[ \leq C \sum_{k_2,\ldots,k_{n+1}=0}^{N} \left( a_{0,k_2,\ldots,k_{n+1}} + b_{1,k_2,\ldots,k_{n+1}} \right) t^{2(\sum_{j=1}^{n+1} k_j)+(n+1)} |Z_0 Z_{k_2} \cdots Z_{k_{n+1}} f_t|^2 \]  
(6.1.28)

\[ + C \sum_{k_2,\ldots,k_{n+1}=0}^{N} \sum_{k_1=1}^{N} \left( a_{k_1,\ldots,k_{n+1}} + b_{k_1+1,k_2,\ldots,k_{n+1}} + \frac{b_{k_1,\ldots,k_{n+1}}}{4} \right) t^{2(\sum_{j=1}^{n+1} k_j)+(n+1)} |Z_{k_1} \cdots Z_{k_{n+1}} f_t|^2. \]  
(6.1.29)
6.1. Assumptions and statement of main result

Before turning to [III] notice that, because of (6.1.7),

\[ [Z_{k_1} \cdot \ldots \cdot Z_{k_{n+1}}, \mathcal{C}] = [Z_{k_1} \cdot \ldots \cdot Z_{k_{n+1}}, B]. \]

\[ [\text{III}] \leq C \sum_{k_1, \ldots, k_{n+1}=0}^{N} a_{k_1, \ldots, k_{n+1}} \left( \sum_{j=1}^{\infty} t^{2(\sum_{j=1}^{\infty} k_j) + n} \right) a_{k_1, \ldots, k_{n+1}} |Z_{k_1} \cdot \ldots \cdot Z_{k_{n+1}} f_t|^2 \]

\[ + \sum_{h=1, h \neq j \forall j}^{n+1} t^{2(\sum_{j=1}^{\infty} k_j) + (n+2)} \frac{|Z_{(k_1, \ldots, k_{n+1})} + e_h f_t|^2}{a_{k_1, \ldots, k_{n+1}}} \]

(6.1.30)

\[ + C \sum_{k_2, \ldots, k_{n+1}=0}^{N} \sum_{k_1=1}^{N} \left( -b_{k_1, \ldots, k_{n+1}} t^{2(\sum_{j=1}^{\infty} k_j) + n} \right) |Z_{k_1} \cdot \ldots \cdot Z_{k_{n+1}} f_t|^2 \]

(6.1.31)

\[ + C \sum_{k_2, \ldots, k_{n+1}=0}^{N} \sum_{b_{k_1, \ldots, k_{n+1}}=1}^{N} b_{k_1, \ldots, k_{n+1}} t^{2(\sum_{j=1}^{\infty} k_j) + n} \left( \frac{|Z_{k_1} \cdot \ldots \cdot Z_{k_{n+1}} f_t|^2}{4} + \right. \]

\[ + \sum_{h=2, h \neq j \forall 2 \leq j, k_j=N}^{n+1} |Z_{(k_1-1, k_2, \ldots, k_{n+1})} + e_h f_t|^2 \]

(6.1.32)

\[ + \sum_{k_2, \ldots, k_{n+1}=0}^{N} \sum_{k_1=1}^{N} b_{k_1, \ldots, k_{n+1}} \left( b_{k_1, k_2, \ldots, k_{n+1}} t^{2(\sum_{j=1}^{\infty} k_j) + (n-2)} \right) |Z_{k_1-1} Z_{k_2} \cdot \ldots \cdot Z_{k_{n+1}} f_t|^2 \]

\[ + \sum_{h=1, h \neq j \forall 2 \leq j, k_j=N}^{n+1} t^{2(\sum_{j=1}^{\infty} k_j) + (n+2)} \frac{|Z_{(k_1, \ldots, k_{n+1})} + e_h f_t|^2}{b_{k_1, \ldots, k_{n+1}}} \]

(6.1.33)

\[ \leq C \sum_{k_2, \ldots, k_{n+1}=0}^{N} \left( a_{0,k_2,\ldots,k_{n+1}}^2 + 1 + \sum_{h=2, h \neq j \forall 2 \leq j, k_j=N}^{n+1} b_{1,k_2,\ldots,k_{n+1}} - e_h + b_{2,k_2,\ldots,k_{n+1}}^2 \right) \]

\[ \cdot t^{2(\sum_{j=1}^{\infty} k_j) + n} |Z_0 Z_{k_2} \cdot \ldots \cdot Z_{k_{n+1}} f_t|^2 \]

(6.1.34)
In [I] appear only and all the terms of length $n + 2$ starting with $Z_0$. [II] and [III] contain terms of length $n + 1$, which can either be of the form $\{Z_0, Z_{k_2}, \ldots, Z_{k_{n+1}} f_l\}$, i.e. starting with $Z_0$, or of the form $\{Z_k, \ldots, Z_{k_{n+1}} f_l\}$ with $k_1 \geq 1$. The latter terms are those that we can control directly using (6.1.31) and through an appropriate choice of the constants $a_{k_1,\ldots,k_{n+1}}$ and $b_{k_1,\ldots,k_{n+1}}$, which we will make explicit in few lines. On the other hand, the terms $\{Z_0, Z_{k_2}, \ldots, Z_{k_{n+1}} f_l\}^2$ never appear multiplied by a negative constant. This problem is easily overcome. Indeed

$$\left( -\mathcal{L} + \frac{\partial}{\partial t} \right) (\Gamma_t^{(n)} f_l) \leq \left( -\mathcal{L} + \frac{\partial}{\partial t} \right) (\Gamma_t^{(n)} f_l) + [I] + [II] + [III],$$

so in order to control the terms $\{Z_0, Z_{k_2}, \ldots, Z_{k_{n+1}} f_l\}^2$ we can get some help from

$$\left( -\mathcal{L} + \frac{\partial}{\partial t} \right) (\Gamma_t^{(n)} f_l),$$

which will contain the addend

$$\sum_{k_1,\ldots,k_n=0}^{N} -2a_{k_1,\ldots,k_n} t^2 (\sum_{j=1}^{n} k_j) + n |Z_0, Z_{k_1}, \ldots, Z_{k_n} f_l|^2,$$

i.e., relabelling things

$$\sum_{k_1,\ldots,k_n=0}^{N} -2a_{k_2,\ldots,k_n} t^2 (\sum_{j=2}^{n} k_j) + n |Z_0, Z_{k_2}, \ldots, Z_{k_n} f_l|^2. \quad (6.1.35)$$

From (6.1.24) and (6.1.25), in order for [I] to be negative $\forall t > 0$ we need to impose:

when $k_1 \geq 1$ : $a_{k_1,\ldots,k_{n+1}} > b_{k_1+1,k_2,\ldots,k_{n+1}}^2 + 1$

when $k_1 = 0$ : $a_{0,k_2,\ldots,k_{n+1}} > b_{1,k_2,\ldots,k_{n+1}}^2$.

So imposing

$$a_{k_1,\ldots,k_{n+1}} \gg b_{k_1+1,k_2,\ldots,k_{n+1}}^2 + 1 \quad \forall k_1 \in \{0, \ldots, N\} \quad (6.1.36)$$

will do. Looking at (6.1.29) and (6.1.34), i.e. at the terms of length $n + 1$ that do not start with $Z_0$, we need to require

$$a_{k_1,\ldots,k_{n+1}}^2 + 2 - \frac{1}{2} b_{k_1,\ldots,k_{n+1}} + b_{k_1+1,k_2,\ldots,k_{n+1}}^2 + \sum_{h=2,h\neq j}^{n+1} b_{(k_1+1,k_2,\ldots,k_{n+1})-e_h} + a_{k_1,\ldots,k_{n+1}} + b_{k_1+1,k_2,\ldots,k_{n+1}} < 0.$$

Finally, putting (6.1.28), (6.1.33) and (6.1.35) together, we need

$$-a_{k_1,\ldots,k_{n+1}} + a_{0,k_2,\ldots,k_{n+1}} + b_{1,k_2,\ldots,k_{n+1}} + a_{0,k_2,\ldots,k_{n+1}}^2 + 1 + \sum_{h=2,h\neq j}^{n+1} b_{(1,k_2,\ldots,k_{n+1})-e_h} + b_{1,k_2,\ldots,k_{n+1}}^2 < 0.$$
Hence, we can choose the constants in the following way
\[ a_{k_2, \ldots, k_{n+1}} \gg b_{k_1, \ldots, k_{n+1}} \gg a_{k_1, \ldots, k_{n+1}} \gg b_{k_1+1, k_2, \ldots, k_{n+1}} \gg \{b_{(k_1+1, k_2, \ldots, k_{n+1})-e_0}\} h_{\geq 2} \gg 1. \]
\[
(6.1.37)
\]
Notice that this choice of the constants not only makes \((-\mathcal{L} + \frac{d}{dt}) \Gamma_{t}^{(n+1)} f_t < 0\) but it is also the right choice in order for \((6.1.9)\) to imply \((6.1.5)\).

We now show how to remove assumptions \((6.1.6)\) and \((6.1.7)\). If we remove \((6.1.7)\) and we assume that \((6.1.8)\) holds also for \(i, j = 0\) or, in other words, if we assume \((6.1.4)\), then \([Z_{k_1} \cdots Z_{k_{n+1}}, \mathcal{L}]\) is not anymore equal to \([Z_{k_1} \cdots Z_{k_{n+1}}, B]\) and we also need to consider the addend \([Z_{k_1} \cdots Z_{k_{n+1}}, Z_0^2]\). To this end observe that
\[
[Z_{k_j}, Z_0^2] = -Z_0[Z_0, Z_{k_j}] - [Z_0, Z_{k_j}] Z_0
\]
\[
= -Z_0 \left( \sum_{l=0}^{N} \beta_{(l k_j), l} \beta_{(l h), h} \right) \left( \sum_{l=0}^{N} \beta_{(l k_j), l} \right) Z_0.
\]
\[
(6.1.38)
\]
From the commutator relation \((6.1.13)\), iterating and using \((6.1.38)\) we get that for \(0 \leq k_1, \ldots, k_{n+1} \leq N,\)
\[
[Z_{k_1} \cdots Z_{k_{n+1}}, Z_0^2] = [Z_{k_1}, Z_0^2]Z_{k_2} \cdots Z_{k_{n+1}}
\]
\[
+ \sum_{j=2}^{n} Z_{k_1} \cdots Z_{j-1} [Z_{k_j}, Z_0^2]Z_{k_{j+1}} \cdots Z_{k_{n+1}} + Z_{k_1} \cdots Z_{k_n} [Z_{k_{n+1}}, Z_0^2]
\]
\[
= -Z_0 \left( \sum_{l=0}^{N} \beta_{(l k_1), l} \beta_{(l k_2), l} \right) Z_{k_2} \cdots Z_{k_{n+1}} - \left( \sum_{l=0}^{N} \beta_{(l k_1), l} \beta_{(l k_2), l} \right) Z_0 Z_{k_2} \cdots Z_{k_{n+1}}
\]
\[
(6.1.39)
\]
\[
- \sum_{j=2}^{n} Z_{k_1} \cdots Z_{j-1} Z_0 \left( \sum_{l=0}^{N} \beta_{(l k_j), l} \beta_{(l k_{j+1}), l} \right) Z_{k_{j+1}} \cdots Z_{k_{n+1}}
\]
\[
(6.1.40)
\]
\[
- \sum_{j=2}^{n} Z_{k_1} \cdots Z_{j-1} \left( \sum_{l=0}^{N} \beta_{(l k_j), l} \beta_{(l k_{j+1}), l} \right) Z_0 Z_{k_{j+1}} \cdots Z_{k_{n+1}}
\]
\[
(6.1.41)
\]
\[
- Z_{k_1} \cdots Z_{k_n} Z_0 \left( \sum_{l=0}^{N} \beta_{(l k_{n+1}), l} \beta_{(l k_1), l} \right) - Z_{k_1} \cdots Z_{k_n} \left( \sum_{l=0}^{N} \beta_{(l k_{n+1}), l} \beta_{(l k_1), l} \right) Z_0.
\]
\[
(6.1.42)
\]
As they are, the terms in \((6.1.39)-(6.1.42)\) are terms of length \(n+2\) and they do not necessarily start with \(Z_0\). Nonetheless, because
\[
Z_{k_j} Z_0 = [Z_{k_j}, Z_0] + Z_0 Z_{k_j} \overset{(6.1.4)}{=} \sum_{h=0}^{N} \beta_{(k_j, h), h} Z_h + Z_0 Z_{k_j},
\]
\[
(6.1.43)
\]
\[1\)
We remind the reader that, for positive constants \(c, d\), we are using the notation \(c \gg d\) to indicate \(c > d\) and \(c > d^2\).
each of them can be turned into the sum of a term of length \( n + 1 \) and a term of length \( n + 2 \) starting with \( Z_0 \). In the same way if, instead of (6.1.6), the more general (6.1.3) holds, iterating (6.1.14) we obtain

\[
[Z_{k_1} \cdots Z_{k_{n+1}}, B] = - \sum_{j=1}^{n+1} Z_{k+j} e_j \quad \text{when } k_j \neq N \forall j. \tag{6.1.44}
\]

Suppose now that \( k_j = N \) and \( k_h < N \forall h \neq j \), then from (6.1.3)

\[
[Z_{k_1} \cdots Z_{k_{n+1}}, B] = - \sum_{j \neq h=1}^{n+1} Z_{k+h} e_h + Z_{k_1} \cdots Z_{k_{j-1}} \left( \sum_{l=0}^{N} \alpha_l Z_l \right) Z_{k_{j+1}} \cdots Z_{k_{n+1}}. \tag{6.1.45}
\]

Putting (6.1.44) and (6.1.45) together we have

\[
[Z_{k_1} \cdots Z_{k_{n+1}}, B] = - \sum_{h=1}^{n+1} Z_{k+h} e_h + \sum_{j: k_j = N} Z_{k_1} \cdots Z_{k_{j-1}} \left( \sum_{l=0}^{N} \alpha_l Z_l \right) Z_{k_{j+1}} \cdots Z_{k_{n+1}}. \tag{6.1.46}
\]

The sum on the right hand side of (6.1.46) contains terms of length \( n + 1 \) and we know how to control them. This concludes the proof. \( \square \)

The result of Proposition 6.1.1 holds also for generators of the form

\[
\mathcal{L} = \sum_{j=1}^{M} X_j^2 + B - \alpha D
\]

where \( D \) is a dilation operator, meaning \([D, X_j] = -\lambda_j X_j \) for some \( \lambda_j > 0 \) and for all the operators \( X_j \) in the filiform algebra. \( \alpha = \alpha(x) \) is a bounded function with bounded derivatives.
Conclusions and future work

In this thesis we looked at the problem of return to equilibrium for hypoelliptic Markovian dynamics produced for example by the coupling of a small Hamiltonian system with an infinite dimensional heat reservoir. The techniques we used, however, are applicable to more general situations.

In everything we said, the potential $V(q)$ acting on the finite dimensional Hamiltonian system was assumed to be essentially quadratic. What happens for non quadratic potentials? The main difficulties in applying Theorem 2.3.4 in this case come from condition 3 and condition (2.3.7). The former imposes a bound on the Hessian of the potential; the latter, as we have already commented, is a Poincaré inequality for a measure of the form $e^{-V(q)}$. In [76], it is shown how to circumvent the problem arising from condition 3 in the case of the Fokker-Plank operator

$$F = -p \partial_q + \partial_q V \partial_p - p \partial_p + \partial_p^2.$$

In particular, the class of potentials is enlarged to $\mathcal{V} = \{ V \in C^2 : |\nabla^2 V| \leq c(1 + |\nabla V|) \}$, for some $c > 0$, which includes any polynomial potential. This is done by using a sort of "$U$-bounds" coming from classic semigroup theory and we believe that an analogous technique should give the same result in the case of system (0.0.3) and this is going to be the object of future work. As for condition (2.3.7), varying the potential means entering the research realm of functional inequalities. However, as long as $V \in \mathcal{V}$, a Poincaré inequality still holds. As for the singular space theory, if the potential is non-quadratic the symbol is not a quadratic form anymore, hence the theory does not apply at all and an explicit calculation of the spectrum is not possible. The moral of the story is that, at least for the problem at hand, obtaining exponential convergence for a wider class of potentials (say confining potentials) is mainly a technical problem. Nonetheless, in situations when the coupling is done with more that one heat reservoir, the result is in general not true [21, 22].

In Chapter 4 we have observed that the hypoellipticity condition suggests a recipe to construct the Lyapunov function used in the context of geometric ergodicity. We have tested this idea on system (4.0.1); hence a natural question to ask is how broad is the class of operators for which this trick does work. We believe such a class of operators should at least contain the set considered in Chapter 6, i.e. operators in Hörmander’s sum of squares form satisfying a strong hypoellipticity condition (6.1.4) over a filiform algebra.
The main problem which is left open in this thesis is the approximation problem. Let me recall here that it consists of two steps: the first step is a problem in approximation theory for scalar functions of real variable; the second is about the ”quality” of the approximation. We believe that, even if the approximation were possible, this would be in general a bad approximation of the GLE dynamics and surely not uniform in time. It has been shown in [39] that, assuming quadratic potential, the GLE exhibits anomalous diffusion behaviour when $\gamma(t)$ decays as an inverse power, i.e. $\gamma(t) \sim t^{-\alpha}$, $0 < \alpha < 1$; in particular, decay to equilibrium is sub-exponential. When $\gamma(t) \sim t^{-\alpha}$ it should be possible to approximate it by a sum of exponentials and the corresponding approximating Markovian system would be (0.0.5), for which decay is exponential.

Finally, future work is bound to go in the direction of infinite dimensional generators and ergodicity in infinite dimensions: suppose we have infinitely many finite dimensional systems of the kind ”particle + N heat bath molecules” (clearly you can also think of more general finite dimensional Markovian systems), interacting via short range interactions. The generator of this system is the sum of infinitely many copies of $\mathcal{L}$, the operator considered in Chapter 6, each of them placed on one of the nodes of the infinite dimensional lattice $\mathbb{Z}^d$, plus a local interaction term. Thanks to ”finite speed of propagation of information” (in other words, thanks to a localization procedure), the estimate of Proposition 6.1.1 are expected to be enough to prove the well posedness of the infinite dimensional semigroup. We will be interested in studying the ergodic properties of such a semigroup.
Appendix A

Some background material

A.1 Stroock-Varadhan Support Theorem

Consider the SDE

\[ dx_t = b(x_t) dt + \sigma dB_t, \quad x_t \in \mathbb{R}^n, \]

where \( b \) is a smooth vector valued function and \( \sigma \) is a matrix with constant coefficients. The deterministic control problem associated with such an SDE is the ODE

\[ \dot{\phi}_t = b(\phi_t) + \sigma \dot{u}_t, \quad \phi_0 = \overline{x}, \quad (A.1.1) \]

where \( u_t \) is a function with given regularity, called control. A point \( y \) is accessible from \( \overline{x} \) if there exists a control \( u_t \) such that the solution \( \phi_t = \phi_t(u) \) to (A.1.1) goes from \( \overline{x} \) to \( y \) in time \( t \), for some \( t > 0 \). In control theory, equation (A.1.1) is said to be controllable if for any \( t_0 \) and any pair of points \( \overline{x} = \phi_{t_0} \) and \( y \) there exists a control \( u_t \) and a time \( t > 0 \) such that \( \phi_t(u) = y \). Let us denote \( A_t(\overline{x}) \) the set of points accessible from \( \overline{x} \) in time \( t \) and \( C^{[0,t]}_\overline{x}(U) \) the set of all solutions of the control problem (A.1.1) as \( u_t \) varies in \( U \), where \( U \) is a class of functions with given regularity. Also, denote by \( S_{\overline{x}}^{[0,t]} \) the support of the diffusion \( x_s, s \in [0,t] \).

Stroock-Varadhan support theorem ([74]) states that

\[ S_{\overline{x}}^{[0,t]} = \overline{C^{[0,t]}_\overline{x}(U)} \]

which can be re-read as

\[ \text{supp} P_t(\overline{x},\cdot) = \overline{A_t(\overline{x})}, \]

with the overline denoting closure in \( \mathbb{R}^n \). Roughly speaking, the theorem says that the stochastic process goes wherever the deterministic solution can be lead to go, hence providing a link
between deterministic control theory and stochastic control theory.

### A.2 Convergence of Markov processes

**Definition A.2.1.** An \( \mathbb{R} \) valued, adapted, increasing continuous process \( Q(t), t \in [0, T] \), with \( Q(0) = 0 \), is the quadratic variation of an \( \mathbb{R} \) valued \( \mathcal{F}_t \) martingale \( M(t) \) if \( M(t)^2 - Q(t) \) is an \( \mathcal{F}_t \) martingale. We denote \( Q(t) \) by \( \langle M \rangle_t \).

The quadratic variation of an \( \text{Itô} \) integral of the form \( \int_0^t u(s)dW_s \) is given by \( \int_0^t u^2(s)ds \).

**Theorem A.2.2 (Martingale central limit theorem).** Let \( M(t) : \mathbb{R}_+ \to \mathbb{R} \) be a martingale on the probability space \((\Omega, \mathcal{F}, \mu)\) with respect to the filtration \( (\mathcal{F}_t)_{t \geq 0} \). Denote by \( \langle M \rangle_t \) its quadratic variation process. Assume the following:

- \( M(0) = 0 \)
- \( M(t) \) has continuous sample paths, is square integrable and has stationary increments
- there exists \( C > 0 \) such that
  \[
  \lim_{t \to \infty} E \left| \frac{\langle M \rangle_t}{t} - C \right| = 0.
  \]

Then the process \( \frac{M(t)}{\sqrt{t}} \) converges in distribution to a \( \mathcal{N}(0, C) \) random variable and the rescaled martingale \( M^\epsilon(t) := \epsilon M(t/\epsilon^2) \) converges weakly in \( \mathcal{C}([0, T], \mathbb{R}) \) to \( \sqrt{C}W(t) \), where \( W(t) \) is a standard Brownian motion.

All of the above holds also when \( M(t) \) is an \( \mathbb{R}^n \) valued martingale with \( C \) an \( n \times n \) symmetric matrix and \( W(t) \) an \( n \) dimensional Brownian motion. In Chapter 3 we used a corollary of the above Theorem A.2.2, namely

**Corollary A.2.3.** With the same notation as in Theorem A.2.2, let \( x(t) : \mathbb{R}_+ \to \mathbb{R} \) be a continuous ergodic Markov process, adapted to the filtration generated by \( W(t) \). Let \( f : \mathbb{R} \to \mathbb{R} \) be a smooth and bounded function and suppose the invariant measure of \( x(t) \) is \( \mu \). Define
\[
I(t) = \int_0^t f(x(s))dW(s).
\]

Then the rescaled stochastic integral \( I^\epsilon(t) := \epsilon I(t/\epsilon^2) \) converges weakly to \( \sqrt{C}W(t) \), with
\[
C = \int_{\mathbb{R}} f^2(y)\mu(dy).
\]
As the Martingale Central limit theorem, also the above corollary is valid in higher dimension, with \( f : \mathbb{R}^n \rightarrow \mathbb{R}^{n\times m} \). The proof of Theorem A.2.2 and of Corollary A.2.3 can be found in [45].

### A.3 Toolbox

This appendix contains miscellaneous facts and definitions.

**Definition A.3.1.** An unbounded operator \( T \) on a real Hilbert space \( H \), with domain \( D(T) \) is said to be **accretive** if

\[
(Tx, x) \geq 0 \quad \forall x \in D(T),
\]

where \((\cdot,\cdot)\) denotes the scalar product in \( H \).

The following theorem is a consequence of the more famous Hille -Yosida Theorem. The proof can be found in [67, Thm. X.48].

**Theorem A.3.2** (Lumer-Phillips). A closed operator \( T \) on a Hilbert space \( H \) is the generator of a contraction semigroup if and only if \( T \) is accretive and \( \text{Ran}(\lambda_0 + T) = H \) for some \( \lambda_0 > 0 \).

The Lumer-Phillips theorem holds also on Banach spaces (and in that case we need to define the notion of accretivity in a Banach context). In the above we stated it in its Hilbertian form, which is the form used in Chapter 3.

**Theorem A.3.3** (The Schwartz kernel Theorem). Given a linear and continuous operator \( A : C_0^\infty(\mathbb{R}^n) \rightarrow \mathcal{D}'(\mathbb{R}^n) \), there exists a unique kernel \( K \in \mathcal{D}'(\mathbb{R}^n \times \mathbb{R}^n) \) such that

\[
(Au, v)_{\mathcal{D}'(\mathbb{R}^n)} = (K, v \otimes u)_{\mathcal{D}'(\mathbb{R}^n)}
\]

where \( \otimes \) denotes tensorization and \((\cdot,\cdot)_{\mathcal{D}'(\mathbb{R}^n)}\) denotes the duality relation.

In other words, the Schwartz kernel Theorem says that an operator \( A : C_0^\infty(\mathbb{R}^n) \rightarrow \mathcal{D}'(\mathbb{R}^n) \) linear and continuous can always be regarded as

\[
A : C_0^\infty(\mathbb{R}^n) \rightarrow \mathcal{D}'(\mathbb{R}^n)
\]

\[
f(x) \rightarrow \int_{\mathbb{R}^n} K(x,y)f(y)dy.
\]
Theorem A.3.4 (Fredholm’s Alternative). Let $H$ be a Hilbert space with scalar product $(\cdot, \cdot)$ and $K$ a compact linear operator. Consider the following problems

\[(\lambda I - K)u = 0 \quad (\lambda I - K)u = f, \ f \in H \quad (A.3.1)\]

\[(\lambda I - K^*)u = 0 \quad (\lambda I - K^*)u = g, \ g \in H. \quad (A.3.2)\]

Then precisely one of the following statements holds:

(i) $u = 0$ is the only weak solution to $(A.3.1)_1$ and the solution to $(A.3.1)_2$ is unique;

(ii) $(A.3.1)_1$ admits a weak solution $u \neq 0$ and there exist $n$ linearly independent weak solutions to $(A.3.1)_1$, say $\{u_j\}_j$; in this case $(A.3.1)_2$ admits solutions iff $f$ is orthogonal to $u_j, \forall j$.

If (i) holds than the analogous statement is valid for the adjoint problem $(A.3.2)$; the same if (ii) holds.

Definition A.3.5. If $q(x, \xi), (x, \xi) \in \mathbb{R}^{2n}$ is a homogenous polynomial of degree two,

\[q(x, \xi) = \sum_{i=1}^{n} a_i x_i^2 + b_i \xi_i^2 + \sum_{1 \leq i \neq j \leq n} a_{ij} x_i x_j + \sum_{1 \leq i \neq j \leq n} b_{ij} \xi_i \xi_j + \sum_{1 \leq i, j \leq n} c_{ij} x_i \xi_j,\]

$a_i, b_i, a_{ij}, b_{ij}, c_{ij} \in \mathbb{C}$, its polarized form $q((x, \xi); (\tilde{x}, \tilde{\xi}))$ is a homogenous polynomial of degree 2 in $4n$ variables, defined as follows

\[q((x, \xi); (\tilde{x}, \tilde{\xi})) := \sum_{i=1}^{n} a_i \tilde{x}_i x_i + b_i \tilde{\xi}_i \xi_i + \sum_{1 \leq i \neq j \leq n} a_{ij} \frac{\tilde{x}_i x_j + \tilde{x}_j x_i}{2} + \sum_{1 \leq i \neq j \leq n} b_{ij} \frac{\tilde{\xi}_i \xi_j + \tilde{\xi}_j \xi_i}{2} + \sum_{1 \leq i, j \leq n} c_{ij} \frac{\tilde{x}_i \xi_j + \tilde{x}_j \xi_i}{2}.\]

Definition A.3.6. Let $\mathcal{P}_t$ be a Markov semigroup over a Polish space $E$. A probability measure $\mu$ is said to be invariant for $\mathcal{P}_t$ if for any bounded and measurable function $\varphi$,

\[\int_E (\mathcal{P}_t \varphi)(x) \mu(dx) = \int_E \varphi(x) \mu(dx).\]

- Some basic facts about ergodicity: the set of all invariant probability measures for $\mathcal{P}_t$ is convex and an invariant probability measure is ergodic if and only if it is an extremal point of such set. Hence, if a semigroup has a unique invariant measure that measure is ergodic. This justifies the definition of ergodic process that we gave in the Overview. It can be shown (see [12]) that, given a $C_0$ Markov semigroup, the following statements are equivalent:
(i) \( \mu \) is ergodic.

(ii) \( \varphi \in L^2(E, \mu) \) and \( \mathcal{P}_t \varphi = \varphi \ \forall t > 0, \mu \text{ a.s.} \Rightarrow \varphi \) is a constant (\( \mu \) a.s.).

(iii) for any \( \varphi \in L^2(E, \mu) \),

\[
\lim_{t \to \infty} \frac{1}{t} \int_0^t \mathcal{P}_s \varphi \, ds = \int_E \varphi(x) \mu(dx), \quad \text{in } L^2(E, \mu).
\]

Because the Markov semigroup preserves constants, condition (ii), read at the level of the generator of the semigroup, says that 0 is a simple eigenvalue of \( \mathcal{L} \), if and only if \( \mathcal{P}_t \varphi = \varphi \ \forall t > 0, \mu \text{ a.s.} \Rightarrow \varphi \) is a constant. In other words, \( \mathcal{L} \varphi = 0 \iff \varphi \) is constant (for \( \varphi \in L^2(E, \mu) \cap D(\mathcal{L}) \)).
Bibliography


