Geometric deformations of integrable systems and beyond

a thesis presented for the degree of

Doctor of Philosophy of Imperial College London

and the

Diploma of Imperial College

by

Alexis Arnaudon

Department of Mathematics

Imperial College

180 Queen's Gate, London SW7 2BZ

October 2017

I certify that this thesis, and the research to which it refers, are the product of my own work, and that any ideas or quotations from the work of other people, published or otherwise, are fully acknowledged in accordance with the standard referencing practices of the discipline.

Alexis Arnaudon

Copyright

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work.

Thesis advisor: Professor Darryl D. Holm Alexis Arnaudon

Geometric deformations of integrable systems and beyond

Abstract

The results of this work were obtained within a single mathematical framework which consists of using the geometrical structure of a given problem not only to study its properties but to implement geometrically consistent deformations. The present interest in such geometric deformations is not only to learn more about the original problem but also to explore new directions, beginning from a common and well-understood starting point and continuing at the interface between several mathematical and physical topics. This approach lays down the foundation of this work upon which various ideas for deformations of several standard systems will be developed and analysed. The mathematical framework that we will be relying on throughout these investigations is the use of symmetries in dynamical systems described by Lie groups which appear in many areas of mathematical physics and in particular in integrable systems, the second theme of this work. Indeed, the integrability of a dynamical system is often related to the existence of symmetries, but usually requires additional structures. The effect of the geometric deformations on the integrability of these systems will be one of the central questions that we will attempt to answer. To undertake such a program, we cannot restrict ourselves to the deformation of a single integrable system, but we must implement various ideas on several and different classic integrable equations in the aim of better characterising their effects and limitations. We will introduce noise, geometric dissipations and Sobolev norms in the classical integrable systems including the rigid body motion, the Toda lattice, the peakon system, partial differential equations such as KdV and NLS as well as the reduced Maxwell-Bloch equation.

To Camille, for her love and support

Acknowledgments

First and foremost, I am enormously grateful to my PhD advisor and friend Prof. Darryl Holm. He believed in me since even before our first encounter and kept providing me excellent pieces advice, a lot of enthusiasm and most of all, a contagious joy for science. My other main inspiration came from Prof. Tudor Ratiu, always available to discuss mathematics with great interests and thoughtful comments. I am very thankful to Andrew Hone and John Gibbon for agreeing to be my PhD examiners and, from a thorough reading of the manuscript, gave me many corrections and advice which substantially improved the quality of the text. I want to thank my collaborators Darryl Holm, Marco Castrillon-Lopez, Alex Castro, Rauan Alkylzanov, Rossen Ivanov, Stefan Sommer, John Gibbon, and Nader Ganaba with whom I had a very nice time working on various projects. I am also very grateful to my colleges at Imperial College London who helped me throughout my PhD, including Jaap Eldering, Henry O. Jacobs, Francois Demoures, Tom Tyranowski, Maximilien Engel, Ryan Barnett, John Elgin, and Wei Pan. It will not be possible for me to acknowledge all the people I met during my PhD, so I only give a selection of the most influential person to me. They include Carlos Tomei, Yuji Kodama, Martin Bauer, Martins Bruveris, Cesare Tronci, Ana-Bela Cruzeiro, Andrew Hone, and Francois Gay-Balmaz. Finally, I want to warmly thank all my friends in London and at home, as well as my family, and in particular my parents.

List of figures

1.1 Structure of the thesis 15

2.1 Energy distribution of the stochastic rigid body 44
2.2 Probability distribution of the stochastic rigid body 46
2.3 Stationary distribution of the stochastic rigid body 56
2.4 Lyapunov exponent of the stochastic rigid body 73
2.5 Random attractor of the stocahstic rigid body 75
2.6 Attractors of the periodically forced rigid body 78
2.7 Lyapunov exponent of the periodically forced rigid body . . . 79
2.8 Illustration of the stability of the stochastic rigid body 92
2.9 Effect of noise correlation on moments of the stochastic landmarks dynamic 123
2.10 Effect of particle interactions on moments of the stochastic landmarks dynamic 124
2.11 Moment evolution for several stochastic landmarks 125
2.12 Comparison of minimisation algorithms for noise estimation with a single landmark 126
2.13 Comparison of minimisation algorithms for noise estimation with a several landmarks 127

3.1 Standing wave solution of the CH-NLS equation 164
3.2 Solitary waves of the CH-NLS equation 166
3.3 Collision of solitary waves of the CH-NLS equation 167
3.4 Modulational instabilty of the CH-NLS equation 170
3.5 Peregrine solutions of the CH-NLS equation 171

3.6 Collision of solitons on the stochastic integrable NLS equation 184
3.7 Peregrine solution of the stochastic integrable NLS equations . 185
3.8 Kink solution of the HRMB equation 202
3.9 Collision of solitons of the original RMB equation 204
3.10 Linear stability analysis of the RMB equations 207
3.11 Typical solutions of the finite dimenional reduction of the RMB equations 209

B.1 Numerical convergence for the CH-NLS equation 234

Nomenclature

β	Inverse temperature
κ	Wavelength
Λ	Sobolev H^1 operator
λ	Loop group/algebra parameter
λ_+	Top Lyapunov exponent
λ_i	Lyapunov exponents
$\mathbb{I}$	Moment of inertia
$\mathbb{P}$	Probability distribution
$\mathbb{P}_\infty$	Stationary probability distribution
$\mathbf{\Omega}$	Angular velocity
$\mathbf{\Pi}$	Angular momentum
$\mathcal{A}$	Attractor
$\mathcal{A}_\omega$	Pullback attractor
$\mathcal{L}$	Kolmogorov operator
$\mathfrak{g}$	Lie algebra
μ	Element in the dual Lie algebra $\mathfrak{g}^*$
ω	Frequency
Φ_i	Stochastic potential

σ_i	Volatility
θ	Dissipation parameter
ξ	Element in the Lie algebra $\mathfrak{g}$
C	Casimir
dW_t^i	Standard Weiner process
G	Lie Group
h	Reduced Hamiltonian
J/K	Hamiltonian structures
L	Velocity loop algebra element
l	Reduced Lagrangian
$L\mathfrak{g}$	Loop algebra
LG	Loop group
M	Momentum dual loop algebra element
m	Momentum field
p	Momentum variable
q	Position variable
T	Temperature
u	Velocity field

Contents

1 Introduction 1

1.1 Motivations 1

1.2 Background 4

1.2.1 Geometric mechanics 4

1.2.2 Integrable systems 7

1.2.3 Stochastic processes 10

1.3 Structure of this thesis 15

1.3.1 By type of deformation 16

1.3.2 By type of systems 17

1.4 List of publications 18

2 Finite dimensional systems 22

2.1 Introduction 22

2.2 Stochastic deformation 24

2.2.1 The stochastic Euler-Poincaré equations 24

2.2.2 The Fokker-Planck equation and stationary distribution 36

2.2.3 The stochastic free rigid body 41

2.2.4 Summary and outlook 47

2.3 Dissipation and random attractors 48

2.3.1 Dissipation with double bracket 48

2.3.2 The Fokker-Planck equation and the Gibbs distribution 53

2.3.3 Statistical physics 56

2.3.4 Random attractors 59

2.3.5 Random attractor in the rigid body 72
2.3.6 Effect of periodic kicking on the rigid body 76
2.3.7 Summary and outlook 80
2.4 Stochastic nonlinear stability 82
2.4.1 Review on nonlinear stability 83
2.4.2 The stochastic Energy-Casimir method 86
2.4.3 Application to the rigid body 90
2.4.4 Summary and outlook 92
2.5 Stochastic Toda lattices 94
2.5.1 The deterministic Toda lattice 94
2.5.2 The stochastic Toda lattice 99
2.5.3 Dissipative stochastic Toda lattice 103
2.5.4 The isospectral stochastic Toda lattice 108
2.5.5 Summary and outlook 110
2.6 Stochastic landmarks 112
2.6.1 Stochastic dynamics 113
2.6.2 The Fokker-Planck equation and moment dynamics . . 116
2.6.3 Numerical examples 122
2.6.4 Summary and outlook 127

3 Infinite dimensional systems 130
3.1 Introduction 130
3.2 Lagrangian reduction and Sobolev deformation of integrable hierarchies 133
3.2.1 Preliminaries 134
3.2.2 Reduction theorems 139
3.2.3 Reduction for hierarchies 141
3.2.4 The AKNS hierarchy 147

3.2.5 The Sobolev deformation . . . 151
3.2.6 Deformed AKNS hierarchy . . . 153
3.2.7 Summary and outlook . . . 158
3.3 The CH-NLS equation . . . 160
3.3.1 Symmetries and non-integrability . . . 161
3.3.2 Solitary wave solutions . . . 163
3.3.3 Modulational instability . . . 169
3.3.4 Summary and outlook . . . 172
3.4 Stochastic hierarchies . . . 174
3.4.1 Stochastic constrained variational principle . . . 174
3.4.2 The stochastic AKNS hierarchy . . . 176
3.4.3 Stochastic integrable equations . . . 181
3.4.4 Summary and outlook . . . 188
3.5 Stochastic PDE with stochastic transport . . . 190
3.5.1 Classical transport noise . . . 190
3.5.2 Stochastic NLS via the Madelung variables . . . 192
3.5.3 Summary and outlook . . . 196
3.6 The reduced Maxwell-Bloch equations . . . 197
3.6.1 The four integrable reduced Maxwell-Bloch equations . 197
3.6.2 Finite dimensional reduction . . . 204
3.6.3 Modulational instability . . . 205
3.6.4 Application to strongly correlated Bose-Einstein condensates . . . 210
3.6.5 Summary and outlook . . . 215

4 Conclusion 217

4.1 Geometric deformations of integrable systems . . . 217
4.2 . . . and beyond . . . 220

Appendix A Moment equation for stochastic landmarks **223**

A.1 First moments 223

A.2 $\langle qq \rangle$ correlation 225

A.3 $\langle pp \rangle$ correlation 226

A.4 $\langle pq \rangle$ correlation 230

Appendix B Numerical scheme for the CH-NLS equation **232**

Appendix C Derivation of the RMB equations **235**

C.1 Standard RMB equations 236

C.2 Hyperbolic RMB equations 237

Index **241**

References **258**

1

Introduction

1.1 Motivations

Integrability of a dynamical system means that it can be integrated explicitly forward in time for some class of initial conditions. This is a rather strong condition not present for most dynamical systems but makes it remarkable from a mathematical point of view as well as useful from a practical point of view. The latter comes from the fact that computing explicit solutions provide a better understanding of their properties which can, in turn, be used in applications if, for example, the integrable equation models a physical process. The former point of view is what will interest us the most and concerns the mathematical reasons why these systems can be integrated. To integrate a system, we need to have a phase space as small as possible. This is achieved by having enough constant of motion, and if they are known, they will remove several dimensions of the phase space. In finite dimensions, if the phase space can be reduced in this way to half its original dimension, only the positions have to be determined and the momenta would have a direct algebraic relationship with them via the constants of motion. These particular relations can be proven to give enough information so that it becomes pos-

sible to directly solve the system forward using various methods, depending on the system. This definition also holds in infinite dimensions, although the condition for the number of conservation laws is different. Namely, we need an infinite number of independent constants of motion in involution. Our interest here is to study integrable systems in the framework of geometric mechanics which uses notions of classical physics together with differential geometry and Lie group theory to analyse and better understand mechanical systems with symmetries. In this language, a conservation law almost always corresponds to a symmetry of the system which will then be used to understand the structure of the dynamical equation. Indeed, for a system have symmetries, it must contain very specific and often nonlinear terms which will shed light on the properties of the solutions and possibly to the integrability of the system.

As the title suggests, the main purpose of this work is not to obtain a better understanding of integrability using geometric mechanics, but to study geometric deformation. We have the following two reasons for not only focusing on integrable equations:

- The theory of integrable systems is a well-developed research topic, and most of the classic examples are well-understood. Focusing on integrability would only allow us to make small contributions to this field.
- Integrability is a very specific feature for an equation and restricts its use for applications and in particular for the range of phenomena that an integrable equation can describe. Only focusing on integrability will make us miss opportunities to understand and study other non-integrable phenomena.

We thus want to follow another trajectory in this world at the interface of integrable systems and geometric mechanics. We want to use the geometric structure of integrable equations to deform them in a consistent and systematic way. More precisely, we want to modify or add extra terms to some standard integrable equations such that the geometrical structure of the original equation related to symmetries and conservations laws is preserved. By

doing this, we will be able to get around the two problems mentioned above to obtain more interesting results as well as new open problems. In particular, we extend them as follows:

- Geometric deformation will test the integrability of an integrable equation and will shed light on the robustness of integrability for this equation. If the system remains integrable after being deformed, we may even discover new integrable equations.
- Geometric deformation will allow us to leave the realm of integrability while remaining near it, therefore providing more freedom to explore phenomena not present in integrable systems as well as more applications of these deformed equations.

Following this direction of research, we will explore several classical integrable systems such as the free rigid body, the Toda lattice and the AKNS hierarchy by implementing geometric deformation including stochastic deformation and Sobolev deformation. Apart from special cases, the stochastic deformation do not preserve integrability of the finite dimensional systems, but provide, together with geometric dissipation, a vast playground for research in stochastic dynamical systems where the geometric conserved quantities remain conserved. In infinite dimensions, only a particular type of noise preserves integrability, but as these systems are infinite dimensional, there exist more possibilities to obtain new stochastic integrable equations. For Sobolev deformation, we will obtain both integrable and non-integrable equations depending on the flow of the AKNS hierarchy.

We will give more detail on the various applications of these deformed equations in section 1.3, but first we need to say a few words about the three most important mathematical topics for this work in section 1.2.

1.2 Background

In this section, we will paint a rough portrait of the three main topics of this text, which are geometric mechanics, integrable systems and stochastic processes. The first gives us the framework for this work. The second describes the specific properties of most of the systems which we will study and deform. The last is a particular type of deformation which we will often use.

1.2.1 Geometric mechanics

The field of geometric mechanics began in the late 60s with people such as V. Arnol'd and J. Marsden, but uses older ideas formulated by Hamilton, Lie, Lagrange, Poincaré, etc. In particular, one of the main equations which we will be using, the Euler-Poincaré equation, originates from Poincaré (1901). Much later, Arnol'd published a series of papers on how to use Lie groups to understand and study fluid mechanics, in particular, the fundamental paper Arnol'd (1966a) supplemented by stability results in Arnol'd (1966b). This was the starting point for numerous studies of finite and infinite dimensional mechanical systems and yielded classical results used in many other areas of mathematics. It would be impossible to list all of them here, so we only refer to some major breakthroughs. Less than five years after the first paper of Arnol'd, Ebin and Marsden (1970) obtained important analytical results on the incompressible Euler equation based on this geometrical interpretation of fluid mechanics.

In fact, the main tool of geometric mechanics is the use of Lie group structure to apply the theory of reduction by symmetry. One of the early theoretical work in this theory is Marsden and Weinstein (1974) where they show that the reduced space is always a symplectic manifold with a non-canonical symplectic form. These reduced spaces are the coadjoint orbits which we will be using throughout this text. Reduction by symmetries became an active topic of research and was applied to several different types of systems. The most important are systems with semi-direct product structures, first by Marsden et al. (1984) for Hamiltonian reduction and later by Holm et al. (1998a) for

Lagrangian reduction. These structures include systems obtained by breaking symmetries and form the building blocks of many theories in physics. Unfortunately, they are not general enough to include all such systems with broken symmetries. For this, the so-called Lagrange-Poincaré reduction was developed in Cendra et al. (2001) for the Lagrangian reduction and later by Marsden et al. (2007) in the Hamiltonian setting. The basic example of this theory is a particle in a magnetic field, but in fact generalises non-local gauge theories, where the gauge group corresponds to the symmetry group acting on the configuration space.

Many further developments includes the study of systems with constraints, internal degrees of freedom, dissipation, etc. We refer to the monographs Arnol'd (1989), Holm (2008), Marsden and Ratiu (1999) which contain excellent overviews of the scope of this theory with many additional references to the literature. In this work, we will mostly be using the simplest reduction by symmetry which produces the Euler-Poincaré equation. We will not use the Lagrange-Poincaré reduction here, but some results could be extended to this more general framework.

We conclude this very short introduction to geometric mechanics with a few definitions, important for this text.

We begin by one of the definitions of a semisimple Lie group and its Lie algebra.

Definition 1.2.1 (Lie group and Lie algebra actions)**.** *For a matrix Lie group G, its Lie algebra $\mathfrak{g}$ and $\xi, \eta \in \mathfrak{g}$ and $\mu \in \mathfrak{g}^*$, we have*

1. *the Adjoint action*

$$\mathrm{Ad}_g \xi = g \xi g^{-1}\,; \tag{1.2.1}$$

2. *the Coadjoint action*

$$\langle \mathrm{Ad}^*_{g^{-1}} \mu, \xi \rangle_{\mathfrak{g}\times\mathfrak{g}^*} = \langle \mu, \mathrm{Ad}_g \xi \rangle_{\mathfrak{g}\times\mathfrak{g}^*}\,; \tag{1.2.2}$$

3. the Lie algebra adjoint action

$$\mathrm{ad}_\xi \eta = [\xi, \eta]\,; \tag{1.2.3}$$

4. the Lie algebra coadjoint action

$$\langle \mathrm{ad}^*_\xi \mu, \eta \rangle_{\mathfrak{g}\times\mathfrak{g}^*} = \langle \mu, \mathrm{ad}_\xi \eta \rangle_{\mathfrak{g}\times\mathfrak{g}^*}\,. \tag{1.2.4}$$

Notice that these operations exist for all Lie groups, not necessarily for finite dimensional Lie groups. We can specify a convenient class of Lie groups which we will be using often.

Definition 1.2.2 (Semi-simple Lie group and Lie algebra)**.** *A semisimple Lie group is a Lie group with a semisimple Lie algebra. A semisimple Lie algebra has the property that the Killing form is non-degenerate, given by*

$$\kappa(\xi, \eta) = \mathrm{Tr}(\mathrm{ad}_\xi \mathrm{ad}_\eta)\,. \tag{1.2.5}$$

The advantage of these finite dimensional matrix Lie groups is that we can identify the Lie algebra with its dual, via the non-degenerate Killing form, that is $\mathfrak{g} \cong \mathfrak{g}^*$ and $\mathrm{ad}^* = -\mathrm{ad}$. Furthermore, for semi-simple Lie algebra, the Killing form is a matrix pairing (from the Frobenius norm) $\mathrm{Tr}(\mathrm{ad}_\xi \mathrm{ad}_\eta) = \epsilon \mathrm{Tr}(\xi \eta^T)$, where $\epsilon < 0$ is a number depending on the Lie algebra.

Coming back to the general case, we need to define the structure of coadjoint orbits in the dual of a Lie algebra.

Definition 1.2.3 (Coadjoint orbit)**.** *A coadjoint orbit $\mathcal{O}_\mu$ at $\mu \in \mathfrak{g}^*$ is the orbit of the Lie group G in the dual of the Lie algebra $\mathfrak{g}^*$ under the coadjoint action, that is*

$$\mathcal{O}_\mu := \{\mathrm{Ad}^*_g \mu,\ \forall g \in G\}\,. \tag{1.2.6}$$

The coadjoint orbits have the remarkable property of admitting a non-canonical symplectic structure.

Proposition 1.2.1 (Hamiltonian structure). *The coadjoint orbits admit the symplectic structure*

$$\omega_\mu(\xi,\eta) := \langle \mu, \mathrm{ad}_\xi, \eta\rangle_{\mathfrak{g}\times\mathfrak{g}^*} . \tag{1.2.7}$$

This structure defines a Hamiltonian structure on $\mathcal{O}_\mu$, or Lie-Poisson bracket for two functions $F, G \in C(\mathfrak{g}^*)$

$$\{F(\mu), G(\mu)\}(\mu) = \left\langle \mu, \left[\frac{\partial F}{\partial \mu}, \frac{\partial G}{\partial \mu}\right]\right\rangle . \tag{1.2.8}$$

The last important definition which we will give here is the possible existence of Casimir functions, useful to describe the geometry of the coadjoint orbit among other things.

Definition 1.2.4 (Casimir function). *A Casimir function $C \in C(\mathfrak{g}^*)$ is a function in the kernel of the Lie-Poisson bracket $\{\cdot,\cdot\}$, that is*

$$\{C, F\} = 0, \forall F \in C(\mathfrak{g}^*) . \tag{1.2.9}$$

Finding these function is not an easy task in general, but for semisimple Lie group, the squared norm of the momentum variable is always a Casimir.

1.2.2 Integrable systems

The second topic of this work is integrable systems, another well-developed area of mathematical physics which also started to be derived in the 60s with Gardner et al. (1967) and the important paper Lax (1968) on the Lax pair. A similar but later significant discovery was the bi-Hamiltonian structures by Magri (1978). These ideas of solitons and integrability emerged from an old discovery of solitary waves in a channel by Russell (1845), followed by the derivation of the KdV equation by Korteweg and De Vries (1895). Such a nonlinear equation was not popular at that time as many experts in fluid dynamics were convinced that only linear equations could model water waves.

We need to wait a century to see the re-emergence of these ideas, helped by

the discovery of a strange nonlinear lattice by Fermi et al. (1955), called the FPU lattice. This lattice would not relax to an equilibrium state, as predicted by fundamental principles of statistical mechanics as the nonlinearities are making this system near-integrable and the near-conserved quantities would prevent such a relaxation. After two decades an extension of the FPU called the Toda lattice was written down by Toda (1967) and then proven to be completely integrable by Flaschka (1974b). The proof of integrability relies on the theory of Lax and a change of variables, commonly referred to as Flaschka's change of variable.

At the same time, significant developments occurred and explained the remarkable properties of equations such as KdV, NLS or the sine-Gordon equation. First, Zakharov and Shabat (1972) solved the NLS equation, extending the concept of inverse scattering transform first introduced by Gardner et al. (1967). Shortly after, Ablowitz et al. (1973a,b) generalised this method to include the KdV and the modified KdV equation, creating the AKNS hierarchy. From this point onwards, the modern the field of integrable systems had emerged, and a large amount of work was published until now with many discoveries in the 80s. We refer to classic books such as Ablowitz and Segur (1981), Dodd et al. (1982), Newell (1985) for more detail about the developments of integrable systems in the 80s. For what interests us here, we want to mention two papers. One is the introduction of the R-matrix by Semenov-Tyan-Shanskiĭ (1983) and a similar other derivation of integrable systems by Flaschka et al. (1983). They provide an interesting geometric derivation of integrable hierarchies in the spirit of geometric mechanics but were never fully exploited in this context. These two papers inspired some results of the second chapter of this text. Finally, we want to mention another important discovery in integrable systems which is the existence of an integrable equation with peaked solitons by Camassa and Holm (1993), Fuchssteiner and Fokas (1981), called the Camassa-Holm equation. This paper also triggered numerous works and discoveries of such an equation.

We end this section with three important theorems in the theory of integrable systems, which we will be using later in the text.

The first theorem can be seen as a practical definition of finite-dimensional

integrable systems and is known as the Liouville integrability.

Theorem 1.2.1 (Liouville-Arnold, Arnol'd (1989)). *Suppose that we are given n functions F_i in involution (that is $\{F_i, F_j\} = 0$ for all $i, j = 1, \dots n$ on a symplectic $2n$-dimensional manifold. Consider the level set of these functions for $\mathbf{f} = (f_1, \dots, f_n)$,*

$$M_{\mathbf{f}} = \{x : F_i(x) = f_i, i = 1, \dots n\}\,, \tag{1.2.10}$$

and suppose that they are independent on it. We then have the following statements:

1. *$M_{\mathbf{f}}$ is a smooth manifold, invariant under the Hamiltonian flow with $H = F_1$.*
2. *If $M_{\mathbf{f}}$ is compact and connected, it is diffeomorphic to the n-dimensional torus with coordinates $\boldsymbol{\varphi} = (\varphi_1, \dots, \varphi_n)$.*
3. *The Hamiltonian flow associated to H determines a periodic motion on M_f of the form*

$$\frac{d\boldsymbol{\varphi}}{dt} = \boldsymbol{\omega}(\mathbf{f})\,. \tag{1.2.11}$$

4. *The original Hamiltonian system with Hamiltonian H can be integrated by quadratures.*

This theorem says that it is possible to construct the action-angle variables (F_i, φ_i), with a simple dynamics so that the system can be exactly solved.

Regarding infinite dimensional equations, we can say that a practical notion of integrability was coined by Lax (1968), and later developed into the theory of inverse scattering transform. We summarise it in the theorem below.

Theorem 1.2.2 (Lax (1968)). *A partial differential equation $u_t = F(u, u_x, u_{xx}, \dots)$ is completely integrable if it is equivalent to a zero curvature relation. That*

is there exists two differential operators L and M of the fields $u, u_x, u_{xx}, \ldots$ such that

$$L_t = [M, L] \tag{1.2.12}$$

holds true only if the field u satisfies the original differential equation.

The method to solve the original equation is to first solve the linear spectral problem

$$L\Phi = \lambda\Phi \tag{1.2.13}$$

for a given initial data $u(0)$. The time evolution of the eigenfunction Φ is then given by the equation

$$\frac{d\Phi}{dt} = M\Phi \tag{1.2.14}$$

and the eigenvalue is conserved in time. For a given time t, the solution $u(t)$ can then be reconstructed from the original scattering problem (1.2.13).

1.2.3 Stochastic processes

Stochastic processes, or the study of randomness in general, is a vast subject in mathematics and physics with an interesting history. We will not retrace it here but only refer to landmark papers from which the various tools of stochastic processes emerged. The primary ingredient of stochastic processes is the Brownian motion, first observed by Robert Brown in the 19th century and then theoretically explained by Albert Einstein in Einstein (1905). This paper had a significant impact in science and started the field of statistical physics. In mathematics, another direction was taken, namely, the one that tries to set up a rigorous mathematical framework for describing randomness. An important early work is Itô (1944) which for the first time explained what a stochastic integral is. The result of this paper is fundamental in this field and is referred to as the Itô calculus of stochastic processes. Later, another type of stochastic integral was discovered by Stratonovich (1966) with the

remarkable property that the usual notion of calculus could be used instead of the Itô calculus, provided an extra correction term is added to the original Itô stochastic process. Much later, the theory of rough paths of Lyons (1998) was developed and says that the Itô or Stratonovich calculus are only two extreme cases of an infinite number of possible stochastic integrals.

In the meantime, many properties and use of stochastic processes were discovered. In particular, the link between stochastic processes and diffusion equation such as the Fokker-Planck (Fokker (1914), Planck (1917)), also known as the Kolmogorov equation (Kolmogoroff (1931)) was discovered. We will only refer to classic books in stochastic analysis for more detail about this topic. We selected Da Prato and Zabczyk (2014), Gardiner (2009), Ikeda and Watanabe (1981), Øksendal (2003), Pavliotis (2014) for theoretical works and Chirikjian (2009, 2012) for applications in engineering or modelling of physical processes such as in biology, physics, chemistry, etc. In parallel, randomness started to be used to perturb dynamical systems, and many striking phenomena were discovered. We refer to the two monographs Arnold (1995), Kloeden and Rasmussen (2011) for the use of randomness in finite dimensional dynamical systems, or random dynamical systems.

Finally, we need to mention the developments of stochastic geometric mechanics. The first known attempt to include noise consistently with mechanics goes back to Bismut (1982) where noise in canonical Hamiltonian systems was considered. These results were recently extended to systems with reduction by symmetry in Lázaro-Camí and Ortega (2008). Another approach of perturbing paths on a Lie group with noise, and using reduction with an expected value to obtain dissipative terms in the reduced equation was introduced and studied by Arnaudon et al. (2014), Chen et al. (2015). The latest development in the direction of stochastic geometric mechanics is to include noise in the process of reduction to obtain a stochastic reduced equation. This was done for infinite dimensional systems by Holm (2015) and in finite dimensions by Arnaudon et al. (2017a).

As for the other sections of this introduction, we will now give a few basic definitions which will be used throughout this text. We begin with the definition of the object W_t, central in stochastic processes.

Definition 1.2.5 (Wiener process). *A Wiener process W_t, where t indexes time has the following properties:*

1. $W_0 = 0$;
2. W_t *has increments* dW_t *independent on the past and the future values;*
3. W_t *has Gaussian increments:* $W_{t+s} - W_t \propto N(0, s)$;
4. W_t *is continuous in* t.

The second important definition is the notion of time integral of Wiener processes, or more generally a martingale. Martingales are generalisations of Wiener processes for an amplitude possibly given by a function of the process itself. More precisely, we have the definition of a martingale.

Definition 1.2.6 (Martingale). *A martingale is a stochastic process X_t such that $\mathbb{E}(X_t) = 0$.*

For example, given a process $X \in \mathbb{R}^n$, we may write it as

$$dX = f(X)dW_t\,, \tag{1.2.15}$$

where $f : \mathbb{R}^n \to \mathbb{R}^n$. This equation is the integral equation

$$\int_0^T dX = \int_0^T f(X)dW_t\,, \tag{1.2.16}$$

where the integration must be defined appropriately.

The time integration of a stochastic process of the form (1.2.16) can be defined in two different ways. Let $\{\Delta_i\}$ be a sequence of partition of $[0, T]$ with mesh tending to 0 as $n \to \infty$.

Definition 1.2.7 (Itô integral). *The Itó integral is*

$$\int f(X)dW_t := \lim_{n\to\infty} \sum_{[t_i, t_{i+1}]\in\pi_n} f(X_{t_i})(W_{t_{i+1}} - W_{t_i})\,. \tag{1.2.17}$$

Definition 1.2.8 (Stratonovich integral). *The Stratonovich integral is*

$$\int f(X) \circ dW_t := \lim_{n\to\infty} \sum_{[t_i,t_{i+1}]\in\pi_n} \frac{f(X_{t_{i+1}}) + f(X_{t_i})}{2}(W_{t_{i+1}} - W_{t_i})\,. \quad (1.2.18)$$

These two definitions are equivalent, provided a correction term is added to one or the other definition. For this, we need Itô's lemma.

Lemma 1.2.1 (Itô's lemma). *For a stochastic process as before, but with $n = 1$ for simplicity, the chain rule becomes Itô's lemma for an arbitrary function $g : \mathbb{R} \to \mathbb{R}$*

$$dg(X_t) = g'(X_t)f(X_t)dW_t + \frac{1}{2}g''(X_t)f(X_t)^2dt\,. \quad (1.2.19)$$

The correction term to be added is simply the differential of $f(X)$ given by Itô's lemma, multiplied by the Wiener process such that only the stochastic part of the differential remains. The equivalence is thus, again in the case $n = 1$,

$$\int_0^T f(X) \circ dW_t = \int_0^T f(X)dW_t + \frac{1}{2}\int_0^T f'(X)f(X)dt\,. \quad (1.2.20)$$

This correction term is important in the theory of stochastic processes, as it is directly related to the generator of the stochastic process, in the following sense.

Definition 1.2.9 (Kolmogorov operator). *The time evolution of the probability distribution $\mathbb{P}$ of the stochastic process X_t (for $n - 1$) is given by*

$$\frac{d}{dt}\mathbb{P}(X,t) = \mathcal{L}^*\mathbb{P}(X,t)\,, \quad (1.2.21)$$

where $\mathcal{L}^*$ is the L^2 adjoint of the Kolmogorov operator

$$\mathcal{L} = f(X)\partial_X + \frac{1}{2}f(X)^2\partial_X^2\,. \quad (1.2.22)$$

The equation (1.2.21) *is also known as the Fokker-Planck equation.*

Of course, all these definitions extends to arbitrary dimensions and an arbitrary number of stochastic terms. A deterministic, or drift term can also be added without any further complications, that is we can have a more general stochastic process of the form

$$dX_t = f_0(X_t)dt + \sum_{i=1}^{N} f_i(X_t)dW_t^i\,, \tag{1.2.23}$$

written here in Itô form for $X_t \in \mathbb{R}^n$ and N functions $f_i : \mathbb{R}^n \to \mathbb{R}^n$. A further generalization would be to write stochastic processes on a manifold, see for example Ikeda and Watanabe (1981) or Elworthy (1982). In this text, we will not work on manifolds but on Lie algebra, or dual of Lie algebra which can be viewed as embedded in the Euclidean space.

1.3 Structure of this thesis

This thesis is organised into two main chapters. We treat finite dimensional systems in chapter 2 and infinite dimensional systems in chapter 3. Within these chapters, several sections consider various systems and various types of deformation. Although it is originally written to be read linearly, the reader has the freedom to change the order of the sections to some extent. We will describe here two alternative orders of reading. See also the appendix at the end of the text.

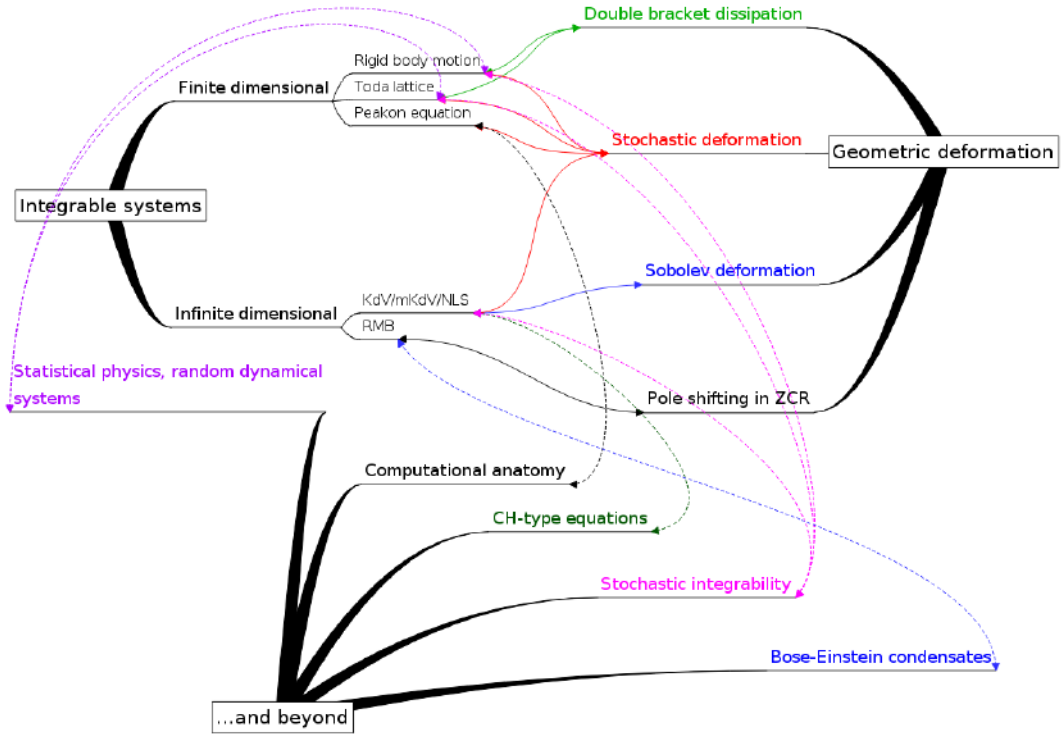


Figure 1.1: This diagram represents the structure of this thesis, where the connections between the tree labelled Integrable systems and the tree labelled geometric deformation indicates which deformation will be applied to which integrable system. The third tree represents the main outcomes of the geometric deformations of integrable systems.

1.3.1 By type of deformation

One can only read on a particular type of deformation of integrable systems. They can be classified into four categories.

1. The stochastic deformation:

   (a) The stochastic deformation is one of the main types of deformation used here. In section 2.2 we explain how to introduce it to dynamical systems with symmetries, and in particular for finite dimensional systems with the rigid body example.

   (b) As a direct extension, the stochastic deformation can be implemented for infinite dimensional systems, as shown in section 3.5.

   (c) This noise has the remarkable property that singular solutions remain for the Camassa-Holm equation. We use this fact to study the stochastic finite dimensional reduction of the original stochastic equation in section 2.6, with application to computational anatomy.

   (d) If one wants to ensure that integrability is preserved in the infinite dimensional systems, the noise must be very specific, as studied in section 3.4 based on the Lagrangian reduction of section 3.2.

   (e) Finally, we apply the stochastic deformation to another interesting system which is the Toda lattice in section 2.5. In particular, we try to understand if noise preserves the coadjoint orbit or the integrability of the Toda lattice.

2. Other extensions of stochastic systems:

   (a) The noise corresponds to diffusion in the phase space, and of the value of the energy of the system. To balance this energy input, it is interesting to add geometric dissipation. We did that in section 2.3 to study some properties of this system such as Gibbs invariant measures and random attractors. We also added a modified version of the geometric dissipation to the Toda lattice in section 2.5

which produces an interesting new dissipative stochastic equation in the continuum limit, or stochastic inviscid Burger's equation.

(b) Section 2.4 is not an extension of stochastic systems but studies their stability, and show that a stable equilibrium remains stable in probability for small times.

3. The Sobolev deformation:

   (a) To introduce the Sobolev deformation in integrable hierarchies and obtain equations such as the Camassa-Holm equation, we need a theory of Lagrangian reduction. Both are explained in section 3.2.

   (b) One of the results of this deformation is the new CH-NLS equation which we will study in section 3.3.

4. Other deformation preserving integrability:

   (a) In section 3.6, we deform the isospectral problem to generate four integrable equations, including the reduced Maxwell-Bloch equation. One of them, the hyperbolic RMB equation is shown to arise from the theory of strongly correlated Bose-Einstein condensates.

1.3.2 By type of systems

One can also read this text by selecting a particular type of system such as the free rigid body and the AKNS hierarchies which are the main examples in this work.

1. Tigid body motion:

   (a) Its stochastic deformation is discussed in section 2.2.

   (b) With dissipation, invariant measures and random attractors are derived in section 2.3.

   (c) The stochastic stability of stable equilibria is derived in 2.4.

(d) Finally, we treated a system with a similar geometrical structure but with a rather different symmetry group. It is the Toda lattice in section 2.5.

2. The AKNS hierarchy:

   (a) The Lagrangian reduction and the Sobolev deformation are explained in section 3.2.

   (b) We studied the CH-NLS equation in section 3.3 with the help of numerical simulations.

   (c) We implemented the stochastic perturbation of the hierarchies in section 3.4 to obtain stochastic integrable equations.

   (d) We deformed the negative flow of the AKNS hierarchy in section 3.6 to obtain four integrable RMB equations.

3. Integrable PDEs and singular solutions:

   (a) In section 3.5, we introduced noise as described for finite dimensional systems in section 2.2.

   (b) Remarkably, the stochastic CH equation admits the same singular solutions as the deterministic CH equation, but with stochastic Hamiltonian dynamic. These stochastic peakons, or landmarks are investigated in section 2.6 with the help of moments of the Fokker-Planck equation.

1.4 List of publications

We list here the publications written during this work. We will only summarise their content and explain where they appear in this thesis. They are order by appearance in the text. The last three papers are not part of this thesis.

Arnaudon, A., De Castro, A. L., and Holm, D. D. (2017a). Noise and dissipation on coadjoint orbits. To appear in: Journal of Nonlinear Science

This paper discusses the stochastic deformation of finite dimensional systems with symmetries. It also considers geometric dissipation and the existence of random attractors in these systems. Section 2.2 and 2.3 comprise the main results of this paper with a focus on the Euler-Poincaré equation and the rigid body. This paper also contain the extension to semi-direct products with the heavy top as an example.

Arnaudon, A., De Castro, A. L., and Holm, D. D. (2016). Noise and dissipation in rigid body motion. In Stochastic Geometric Mechanics, Springer

This paper is a conference proceedings summarising some results of the previous one Arnaudon et al. (2017a), focusing only on the rigid body.

Arnaudon, A., Ganaba, N., and Holm, D. D. (2017b). The stochastic energy-Casimir method. arXiv:1702.03899

This paper extend the energy-Casimir method to the previous stochastic systems. In addition to systems of finite dimensions, it discusses the application of the methods to simple stochastic fluid equations such as the compressible 2D Euler equation. The theory and the finite dimensional example are contained in section 2.4.

Arnaudon, A., Holm, D. D., and Sommer, S. (2017e). A geometric framework for stochastic shape analysis. arXiv preprint arXiv:1703.09971

This paper applies the stochastic deformation on image matching with application to stochastic landmarks, and describes two methods for the estimation of the noise fields from a given set of initial and final images. The stochastic landmark example together with on of these methods based on the Fokker-Planck equation are described in section 2.6.

Arnaudon, A., Holm, D. D., Pai, A., and Sommer, S. (2017d). A stochastic large deformation model for computational anatomy. In Information Processing in Medical Imaging*, pages 571–582*

This is a conference paper summarising Arnaudon et al. (2017e), with an additional real data application.

Arnaudon, A. (2016b). On a lagrangian reduction and a deformation of

completely integrable systems. Journal of Nonlinear Science, *26(5):1133–1160*

This paper explains the Lagrangian reduction for integrable hierarchies and implements the Sobolev deformation to obtain the CH equation as well as the CH-NLS equation. It is the basis of section (3.2).

Arnaudon, A. (2016a). On a deformation of the nonlinear Schrödinger equation. Journal of Physics A: Mathematical and Theoretical, *49(12):125202*

This paper studies the CH-NLS equation with mainly some numerical simulations with the aim of understanding its solitary waves solution and its non-integrability. It is contained in section 3.3.

Arnaudon, A. (2015). The integrable stochastic AKNS hierarchy. arXiv preprint arXiv:1511.07080

This paper derives the stochastic integrable hierarchies based on Arnaudon (2016a) and study some of the obtained stochastic equations. This preprint was used to write section 3.4.

Arnaudon, A. and Gibbon, J. D. (2017). Integrability of the hyperbolic reduced Maxwell-Bloch equations for strongly correlated Bose-Einstein condensates. Physical Review A, *96(1):013610*

This paper discusses the RMB equations, and in particular the hyperbolic RMB equation in the context of strongly correlated Bose-Einstein condensates. Section 3.6 is based on this paper.

Arnaudon, A., Lopez, M. C., and Holm, D. D. (2017f). Un-reduction in field theory. To appear in: Letters in Mathematical Physics

This paper extends the notion un-reduction of classical mechanics in the context of field theory. It is based on the Lagrange-Poincaré reduction. We will not use its results in this work.

Arnaudon, A., Castrillón López, M., and Holm, D. D. (2015). Covariant unreduction for curve matching. Mathematical Foundation of Computational Anatomy 2015, arXiv:1508.05325

This paper is a conference paper which summarises the previous work Ar-

naudon et al. (2015), with numerical examples.

Arnaudon, A., Holm, D. D., and Ivanov, R. I. (2017c). G-strands on symmetric spaces. Proc. R. Soc. A, *473(2199):20160795*

This paper studies the G-strand construction on symmetric spaces, with the discovery of some possibly new integrable equations.

2

Finite dimensional systems

2.1 Introduction

The study of finite dimensional integrable systems began several centuries ago and is a very rich and vast area of research. For us, integrability will always be present and will provide us with a robust framework and ideas for geometric deformation. Our primary focus will not be to investigate integrability of finite dimensional systems but to use it as a starting point for implementing deformation and understanding their effects on various equations. The main deformation of this chapter will be the introduction of noise compatible with the geometrical structure of the equations, and possibly with their integrability. This type of deformation has been used many times in the past as it brings the theory of stochastic processes into dynamical systems. This connection between dynamical systems and stochastic processes is the subject of the theory of random dynamical systems (see the classic monograph Arnold (1995)). The novelty here is the type of noise which will not be a simple additive or multiplicative noise but which we will obtain from a systematic and geometric derivation. In this section, we will present this theory developed in the recent papers Arnaudon et al. (2017a), Cruzeiro

et al. (2016), Holm (2015). We will mostly follow Arnaudon et al. (2017a). This new type of noise turns out to have interesting properties which can be derived in a rather general setting and applied to other systems such as the Toda lattice.

PLAN OF THIS CHAPTER We will introduce stochastic deformation of mechanical systems in section 2.2 which we will then use throughout this work and even for infinite dimensional systems in chapter 3. On top of the noise, we will, in section 2.3 add a certain geometric dissipation to study its interaction with the noise. The main result is the existence of so-called random attractors and stationary distribution of Gibbs types. These investigations are always for asymptotic behaviour of the system for large times, so in section 2.4 we turn to the effect of noise for small times when the system is close to a stable equilibrium. In particular, we extend the energy-Casimir method to assess the stochastic stability of these perturbed systems. Up to this section, we will only have considered the free rigid body as our running example, but in section 2.5, we will apply and extend some of the previous theory to the Toda lattice, which is another well-known integrable system. The Toda lattice is interesting for us as it lies between finite and infinite dimensional systems as it consists of a chain of interacting particles converging to the KdV, or the inviscid Burger's equation in the continuum limit.

2.2 Stochastic deformation

The main topic of this first section is the introduction of stochastic perturbations in mechanical systems with symmetries such that the geometrical structures are preserved. The most important structure is the coadjoint orbit, described by invariant functions called Casimir functions which should remain conserved after adding noise. In the case of integrable systems, other invariants must be preserved to retain the complete integrability of the original system. We will see that this additional condition will often be too restrictive to obtain interesting stochastic systems but will still give some information about the nature of integrability. In this section, we will thus introduce this particular type of noise in mechanical systems with symmetries following the recent works by Arnaudon et al. (2017a), Cruzeiro et al. (2016), Holm (2015). We will only use the free rigid body as our example, but semi-direct product structures can also be deformed in the same way. We refer to Arnaudon et al. (2017a) for some results for these systems, such as the heavy top.

2.2.1 The stochastic Euler-Poincaré equations

We begin with the derivation of the stochastic Euler-Poincaré equation which will be our central equation of this chapter. We will show three different derivations for the stochastic deformation of this equation. The first one is based on a constrained variational principle with Clebsch variables, which is the original derivation of Holm (2015), further generalised by Cruzeiro et al. (2016). The second derivation is based on stochastic constrained variations and will touch upon the open problem of existence of a stochastic Euler-Lagrange equation whose solutions should project onto solutions of the stochastic Euler-Poincaré equation. We finally show how to use another type of constrained variational principle, the Hamilton-Pontryagin principle, which will provide a physically more consistent derivation and works better with dissipation or semi-direct product structures.

The Clebsch derivation

The Clebsch derivation of dynamical equations starts from a constrained variational principle but with free variations. This method has its use in optimal control but also produce the Euler-Poincaré equation, see Gay-Balmaz and Ratiu (2011). For imposing the correct constraint, we need additional variables which will be called the Clebsch variables. They, in fact, correspond to generalised Lagrange multipliers. We will only treat here the simple case of the Euler-Poincaré equation by considering the Clebsch variables to be a pair $(q, p) \in \mathfrak{g} \times \mathfrak{g}^*$. We then write the following constrained variational principle δS, with action functional

$$S(\xi, q, p) = \int l(\xi)\, dt + \int \langle p, \circ dq + \mathrm{ad}_\xi q\, dt \rangle + \int \sum_{i=1}^{n} \Phi_i(q, p) \circ dW_t^i\,, \quad (2.2.1)$$

where $l : \mathfrak{g} \to \mathbb{R}$ is the reduced Lagrangian and the second and third terms encode the constraint for the stochastic evolution of the q variable. Here, $\Phi_i(q, p)$ are arbitrary functions of the Clebsch variables and will be called stochastic potentials, as they are associated with the Wiener processes dW_t^i.

This way of writing a constrained variational principle originates from optimal control systems where q is the control variable evolving according to a dynamics prescribed by the current state ξ and one wants to minimise a cost function, or a Lagrangian $l(\xi)$. The constraint on the dynamics of q is thus enforced directly into the variational principle via the Lagrange multiplier p. We refer to Bloch et al. (2003), Gay-Balmaz and Ratiu (2011) for a modern account of this approach of optimal control and to Holm and Kupershmidt (1983), Marsden and Weinstein (1983) for the original papers introducing the Clebsch method in geometric mechanics.

We now proceed further to derive the equations of motion which are critical points of the variational principle (2.2.1). In the process, we will make a particular choice for the stochastic potentials Φ_i to be able to write the stochastic Euler-Poincaré equation only in term of the variable ξ.

Proposition 2.2.1. *The variational principle* (2.2.1) *with free variations*

yields the stochastic Euler-Poincaré equation

$$d\frac{\delta l(\xi)}{\delta \xi} + \mathrm{ad}^*_\xi \frac{\delta l(\xi)}{\delta \xi} dt + \sum_i \mathrm{ad}^*_{\frac{\delta \Phi_i(\mu)}{\delta \mu}} \frac{\delta l(\xi)}{\delta \xi} \circ dW_t^i = 0\,, \tag{2.2.2}$$

where $\mu = \frac{\delta l}{\delta \xi} \in \mathfrak{g}^*$ *and* $\Phi(\mu) := \Phi(\mathrm{ad}_q p)$.

Proof. We will do the proof in the case of semi-simple Lie algebra, or when the Lie algebra has a bi-invariant pairing. Notice that this computation works for more general systems, as shown in Cruzeiro et al. (2016). We start by computing the free variations of all the variables to obtain

$$\begin{aligned}
&\delta\xi : \mu := \mathrm{ad}^*_q p = \frac{\delta l}{\delta \xi} \\
&\delta q : dp = \mathrm{ad}^*_\xi p\, dt + \sum_i \frac{\delta \Phi_i}{\delta q} \circ dW_t^i \\
&\delta p : dq = -\mathrm{ad}_\xi q - \sum_i \frac{\delta \Phi_i}{\delta p} \circ dW_t^i\,.
\end{aligned}$$

We then compute the time evolution of the momentum variable μ. To get rid of the Clebsch variables we need to restrict the form of the stochastic potential by letting it only depend on the momentum variable μ, that is we choose $\Phi(p,q) = \Phi(\mathrm{ad}^*_q p)$. We then compute the time evolution of the momenta using an arbitrary element $\eta \in \mathfrak{g}$ to obtain

$$\begin{aligned}
\langle \eta, d(\mathrm{ad}^*_q p)\rangle &= \langle \eta, \mathrm{ad}^*_{dq} p\rangle + \langle \eta, \mathrm{ad}^*_q dp\rangle \\
&= \left(-\langle \eta, \mathrm{ad}^*_{\mathrm{ad}_\xi q} p\rangle + \langle \eta, \mathrm{ad}^*_q \mathrm{ad}^*_\xi p\rangle\right) dt \\
&+ \left(-\langle \eta, \mathrm{ad}^*_{\delta_p \Phi_i} p\rangle + \langle \eta, \mathrm{ad}^*_q \delta_q \Phi_i\rangle\right) \circ dW_t^i \\
&= \left(\langle [\eta, [\xi, q]], p\rangle + \langle [\xi, [q, \eta]], p\rangle\right) dt \\
&+ \left(-\langle [\eta, [q, \delta_\mu \Phi_i]], p\rangle \circ dW_t^i + \langle [\delta_\mu \Phi_i, [q, \eta]], p\rangle\right) \circ dW_t^i \\
&= -\langle \eta, \mathrm{ad}^*_\xi \mu\rangle\, dt - \langle \eta, \mathrm{ad}^*_{\delta_\mu \Phi_i} \mu\rangle \circ dW_t^i\,,
\end{aligned}$$

where we used twice the Jacobi identity, the chain rule for the stochastic potential and the notation $\delta_q \Phi_i := \frac{\delta \Phi_i}{\delta q}$ for simplicity. From this computation,

we obtain the stochastic Euler-Poincaré equation when the Lie algebra admits a bi-invariant pairing. ■

It is sometimes convenient to rewrite the stochastic process (2.2.2) using the auxiliary process

$$dX := \xi dt + \sum_i \sigma_i \circ dW_t^i \,,, \tag{2.2.3}$$

where

$$\sigma_i := \frac{\partial \Phi_i(\mu)}{\partial \mu} \,. \tag{2.2.4}$$

In this case, (2.2.2) takes the simple form

$$d\mu + \mathrm{ad}^*_{dX}\mu = 0 \,, \tag{2.2.5}$$

which highlights the fact that this equation is a coadjoint motion with the stochastic velocity X.

Reduction with stochastic variations

The stochastic Euler-Poincaré equation (2.2.2) can also be derived from a stochastic modification of the Euler-Poincaré theorem. We first recall this theorem in the deterministic case to clarify where the noise will be introduced.

Theorem 2.2.1 (Euler-Poincaré reduction). *Provided the Lagrangian $L : TG \to \mathbb{R}$ is G-invariant, the following statements are equivalent in the sense that their solutions coincide.*

1. *Hamilton's variational principle for the Lagrangian $L : TG \to \mathbb{R}$:*

$$\delta \int L(g, \dot{g}) dt = 0 \,. \tag{2.2.6}$$

2. *The Euler-Lagrange equation holds*

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{g}} = \frac{\partial L}{\partial g}\,. \tag{2.2.7}$$

3. *Reduced Hamilton's principle for the reduced Lagrangian* $l : \mathfrak{g} \to \mathbb{R}$*:*

$$\delta \int l(\xi) dt = 0\,, \tag{2.2.8}$$

where $\xi = g^{-1}\dot{g} \in \mathfrak{g}$ *and the constrained variations are* $\delta\xi = \dot{\eta} + [\eta, \xi]$.

4. *The Euler-Poincaré equation holds*

$$\frac{d}{dt}\frac{\partial l}{\partial \xi} = \mathrm{ad}^*_\xi \frac{\partial l}{\partial \xi}\,. \tag{2.2.9}$$

When using this theorem to solve a variational equation on a Lie group, one should start by solving the Euler-Poincaré equation and then reconstruct the solution on the group using $\dot{g} = g\xi$. The relation between the solution on the Lie group and the solution on the Lie algebra is called the reconstruction relation, as it allows one to find the position of the system from the evolution of its speed ξ. The reconstruction relation is thus the obvious place where one would want to introduce noise to retain most of the results of this theorem. Hence, we replace the reconstruction relation by the following stochastic process

$$dg = g\xi dt + \sum_i g\sigma_i \circ dW_t^i\,, \tag{2.2.10}$$

where σ_i are given Lie algebra elements. Notice its similarity with the process X in (2.2.3), that is $dg = g\,dX$. We now proceed by computing the stochastic Euler-Poincaré equation and will give further comments on this stochastic reconstruction relation later. Using the relation (2.2.10), we can compute the associated stochastic variations

$$\delta\xi = d\eta + \mathrm{ad}_\xi \eta\, dt + \sum_i \mathrm{ad}_{\sigma_i}\eta \circ dW_t^i\,, \tag{2.2.11}$$

where $\eta = g^{-1}\delta g$, for the arbitrary variations δg. Indeed, we follow the deterministic derivation, and we compute

$$d(g^{-1}\delta g) = g^{-1}d\delta g - g^{-1}dgg^{-1}\delta g$$
$$\delta(g^{-1}dg) = g^{-1}\delta dg - g^{-1}\delta gg^{-1}dg\,.$$

These formulae yield (2.2.11) after subtracting them and using the fact that $d\delta g = \delta dg$. From this computation, we directly have the following proposition.

Proposition 2.2.2. *The solution of Hamilton's principle* (2.2.8) *for the reduced action functional under the stochastic reconstruction relation* (2.2.10) *is the stochastic Euler-Poincaré equation* (2.2.2).

This derivation of the stochastic Euler-Poincaré equation is the closest derivation to the standard deterministic one. By being inserted into the reconstruction relation, the noise can be interpreted as possible errors made when the position g of the system is computed from the velocity ξ. Interestingly, the noise also appears in the equation for the velocity, that is the Euler-Poincaré equation as one would think that this stochastic perturbation would only affect the position and not the speed. It is for this reason that variational principles are important and make a theory consistent by, in this case, making sure that the noise in the velocity equation is compatible with the stochastic trajectory on the group. In physical terms, we can think about this effect in the following manner. The noise in the reconstruction relation will eventually bring the position of the system to a place that can never be visited by the deterministic solution, and where the dynamics can be rather different from the expected one without noise. It is then necessary to use a different dynamics for the velocity which will depend on where the perturbation brings the system. This effect is modelled by the noise in the reduced equation. Notice that the noise realisation for the reconstruction relation and the stochastic Euler-Poincaré equation must be the same for this theory to make sense.

This derivation also justifies the simplification of the stochastic potentials $\Phi_i(q,p) = \Phi_i(\mu)$ used in the Clebsch derivation. There exists nevertheless a difficult issue to solve, that we now want to explain here. The previous

proposition is only half of the deterministic Euler-Poincaré reduction theorem 2.2.1. What we know is that if we can solve the stochastic equation (2.2.2), then the stochastic reconstruction relation (2.2.10) can be integrated forward to give a stochastic trajectory on the group that we will denote g_t. The question is now about g_t being a solution of a stochastic process corresponding to the Euler-Lagrange equation and if this stochastic process can be derived from a stochastic variational principle on the Lie group. First, the process g_t is non-Markovian, that is any state of the system depends on its past. We can see that by writing the reconstruction relation in integral form as

$$g_T = g_0 + \int_0^T g_t \left(\xi_0 + \int_0^t \mathrm{ad}^*_{\xi_s} \mu_s ds + \sum_i \int_0^t \mathrm{ad}^*_{\sigma_i} \mu_s \circ dW_s^i \right) dt + \sum_i \int_0^T g_t \sigma_i \circ dW_t^i \,, \tag{2.2.12}$$

where the s time variable is for the integrals associated to the reduced equation and the t variable for the group valued solution, evaluated at time T.

This process can also be interpreted as having a stochastic drift. Similar stochastic processes appear in mathematical finance with for example the Longstaff-Swartz model Longstaff and Schwartz (1992) for the description of short interest rates. The major difference between these models in mathematical finance and our system is that the Wiener processes for ξ and g are the same, whereas in finance they use independent Wiener processes.

We will not go further in studying this stochastic process on Lie groups, but rather give pieces of evidence and suggest an extension of this theory to the Lie group that would remain compatible with what we said so far. First, we have the following facts:

1. the noise was introduced preserves the coadjoint orbits obtained by the reduction theory,

2. for a solution of the stochastic Euler-Poincaré equation, there exists a corresponding stochastic path on the Lie group G,

3. if the noise vanishes, we recover the standard reduction theory of Euler-

Poincaré.

4. the noise is the same in both the reconstruction relation and the stochastic Euler-Poincaré equation.

Conjecture 2.2.1. *The stochastic path on the Lie group coming from the reconstruction of a stochastic path on the Lie algebra is obtained from a stochastic variational principle and satisfies a stochastic Euler-Lagrange equation.*

Derivation from the Hamilton-Pontryagin principle

We will now describe yet another derivation of the Euler-Poincaré equation which sits between the two previous approaches and will be useful later on with dissipative forces in section 2.3. The idea behind the Hamilton-Pontryagin principle is similar to the Clebsch principle. It is to enforce the constrained variations directly into the variational principle via a generalised Lagrange multiplier. In this case, we do not introduce Clebsch variables and directly use the Lie group elements and the momentum variable to enforce the constraint. In the present stochastic context, the constraint will become stochastic.

Proposition 2.2.3. *The stochastic Hamilton-Pontryagin variational principle*

$$\delta\left(\int l(\xi)\,dt + \sum_i \int \langle \mu, g^{-1}dg - \xi dt - \sigma_i \circ dW_t^i \rangle\right) = 0\,, \tag{2.2.13}$$

for free variations $\delta g, \delta\xi$ and $\delta\mu$ yields the stochastic Euler-Poincaré equation (2.2.2)*. Here, σ_i are arbitrary Lie algebra elements.*

Proof. We compute the variations $\delta\xi$ and $\delta\mu$

$$\begin{aligned}
\delta\xi &: \frac{\partial l(\xi)}{\partial \xi} = \mu \\
\delta\mu &: g^{-1}dg = \xi dt + \sum_i \sigma_i \circ dW_t^i\,,
\end{aligned}$$

and the δg variation gives

$$\begin{aligned}\delta_g\langle\mu, g^{-1}dg\rangle &= \langle\mu, -g^{-1}\delta g g^{-1}dg + g^{-1}d\delta g\rangle \\ &= -\langle d\mu, g^{-1}\delta g\rangle + \langle\mu, [g^{-1}dg, g^{-1}\delta g]\rangle \\ &= \langle -d\mu + \mathrm{ad}^*_{g^{-1}dg}\mu, g^{-1}\delta g\rangle\,,\end{aligned}$$

where we integrated by parts. Combining terms gives equation (2.2.2). ■

On semi-direct product extension

The first step towards leaving the realm of geodesic equations and kinetic Lagrangians is to introduce additional geometrical structures which can describe broken symmetries. The best and simplest tool is the theory of semi-direct products which can describe most of the interesting examples in finite and infinite dimensions. We refer to Holm et al. (1998a) for extensive works on semi-direct product structures in the context of dynamical systems and fluid dynamics. Most of what we will say in this chapter can be extended to include semi-direct products, where the example of the rigid body corresponds to the heavy top, with gravity breaking the rotation symmetry. For this section on stochastic deformation of the rigid body motion, we derived the equations and studied them to some extent in Arnaudon et al. (2017a). We will not present these results here as they do not provide more insights into the effect of the noise but only complications. We will focus on the Euler-Poincaré equation where the phase space is the tangent bundle of a Lie group.

Itô process and Lie-Poisson equation

We previously derived the stochastic Euler-Poincaré equation and it is now time to study it in more detail. We begin by computing the Itô correction term that allows us to write the stochastic process in Itô form, more appropriate to perform stochastic calculus than standard calculus. We can, in particular, take expectations and compute the Fokker-Planck equation. A direct computation of the Itô correction of the stochastic Euler-Poincaré

equation (2.2.2) gives the following Itô process

$$d\frac{\partial l(\xi)}{\partial \xi} + \mathrm{ad}^*_\xi \frac{\partial l(\xi)}{\partial \xi} dt + \sum_i \mathrm{ad}^*_{\sigma_i} \frac{\partial l(\xi)}{\partial \xi} dW_t^i \\ - \frac{1}{2} \sum_i \mathrm{ad}^*_{\sigma_i} \left(\mathrm{ad}^*_{\sigma_i} \frac{\partial l(\xi)}{\partial \xi} \right) dt = 0 \,, \tag{2.2.14}$$

where the last term is the Itô correction. The important point here is that the correction has a double ad* with a sum over all σ_i fields. The fact that the index i is the same for both coadjoint actions comes from the fact that the Wiener processes W_t^i are independent for each i. If they were not independent, some mixed terms would appear.

From these considerations, we can study the Lie-Poisson formulation of this equation in both Stratonovitch, and Itô form. In Stratonovitch form, the stochastic Lie-Poisson equation is

$$\begin{aligned} df(\mu) &= \left\langle \mu, \left[\frac{\partial f}{\partial \mu}, \frac{\partial h}{\partial \mu} \right] \right\rangle dt + \sum_i \left\langle \mu, \left[\frac{\partial f}{\partial \mu}, \frac{\partial \Phi_i}{\partial \mu} \right] \right\rangle \circ dW_t^i \\ &=: \{f, h\}\, dt + \sum_i \{f, \Phi_i\} \circ dW_t^i \,, \end{aligned} \tag{2.2.15}$$

where $f : \mathfrak{g}^* \to \mathbb{R}$ is an arbitrary function and $h : \mathfrak{g}^* \to \mathbb{R}$ is the Hamiltonian, obtained by a Legendre transform of the reduced Lagrangian l, provided that it is hyper–regular. This formula exhibits the Lie-Poisson bracket which appears in both the drift and the stochastic term. We can now compute the corresponding Itô process to get

$$df(\mu) =: \{f, h\}\, dt + \sum_i \{f, \Phi_i\} dW_t^i + \frac{1}{2} \sum_i \{\Phi_i, \{\Phi_i, f\}\}\, dt \,. \tag{2.2.16}$$

The double coadjoint action term is now a double Lie-Poisson bracket. We will use this equation later to study the Fokker-Planck equation, but for now, we will focus on the properties of the stochastic system and in particular the preserved geometric features.

Proposition 2.2.4. *All solutions of stochastic Euler-Poincaré equation* (2.2.2) *are restricted to a coadjoint orbit* $\mathcal{O}_{\mu_0}$ *where* $\mu_0 = \mu(0)$ *is the initial value of*

the process.

Proof. Without loss of generality, we can assume that $W_0^i = 0$ for all i, thus the coadjoint orbit is parametrised by the value of the initial momentum μ_0. One directly check that the Casimirs are preserved by the flow as both the drift and the volatility coefficients have this same Lie-Poisson structure. ■

Energy conservation and integrability

We just investigated the preservation of geometric invariants under the stochastic perturbation of the Euler-Poincaré equation. We will now study the dynamical invariants, which are not expected to be preserved in general.

In general, the energy $h(\mu)$ is not conserved by the solutions of the stochastic Euler-Poincaré equation (2.2.2). Indeed, using Itô's lemma 1.2.1, we can compute the time evolution of the energy to get

$$dh(\mu_t) = \langle \xi_t, d\mu_t \rangle = \sum_i \langle \xi_t, \mathrm{ad}^*_{\sigma_i} \mu_t \rangle \circ dW_t^i \,,$$

which is not conserved unless the dynamics of the drift is compatible with the fields σ_i in a specific sense. In the previous computation, we see that if $\mathrm{ad}_{\sigma_i}\xi \perp \mu$ for all time and index i where $\perp$ is orthogonality with respect to the pairing, then the energy is conserved. Notice that the drift dynamics depends on the Hamiltonian used to compute ξ_t in term of μ_t. This condition is thus a compatibility condition between the Hamiltonian and the σ_i fields. One could hope that the expected value of the energy would be conserved, but one can verify that it is not the case in general. Nevertheless, the energy is conserved when the functional h and Φ_i are compatible in the following sense. First, recall that the dynamics of our system is tied to the Poisson structure encoded in the Poisson bracket $\{\cdot,\cdot\}$. We then say that the functional h and Φ_i are compatible if

$$\{h, \Phi_i\} = 0, \quad \text{and} \quad \{\Phi_i, \Phi_j\} = 0 \quad \forall i, j \,. \tag{2.2.17}$$

Using only this property we can directly show that the stochastic process has

h and Φ_i as conserved quantities. Indeed, from the Lie-Poisson formulation (2.2.15), we have

$$dh(\mu_t) = \{h, h\}(\mu_t)\, dt + \sum_j \{h, \Phi_j\}(\mu_t) \circ dW_t^j = 0$$
$$d\Phi_i(\mu_t) = \{\Phi_i, h\}(\mu_t)\, dt + \sum_j \{\Phi_i, \Phi_j\}(\mu_t) \circ dW_t^j = 0\,.$$

Proposition 2.2.5. *For a stochastic deformation of an deterministic integrable system, if h and all Φ_i commute with respect to the Poisson bracket of the original system, then this system can be integrated by quadratures.*

Proof. We assume that the deterministic system is integrable in the Liouville sense, that is it can be integrated by quadrature via the action-angle variables, see Theorem 1.2.1. The functions Φ_i must be linearly dependent on the deterministic conserved quantities of the original system. The construction of the action-angle variables, which involves the quadratures is thus not affected but only the dynamical equation of these variables. Recall that the Φ_i are conserved by the stochastic flow. The action variables hence remain constants of motion for the stochastic equation. The angle variables will be affected by the noise, with a stochastic frequency, depending only on the constant action variables. This holds as the Φ_i are in involution with all the other conserved quantities. ■

This proposition leads to the following definition.

Definition 2.2.1 (Stochastic integrable system)**.** *The stochastic process* (2.2.2) *is called a stochastic integrable system if h and all Φ_i commute with respect to the Poisson bracket generating this process and if the deterministic system is already integrable.*

This ensures that the stochastic dynamics is restricted on the invariant tori of the original integrable system. The dynamics on them is stochastic, but as it is just an additive noise on a straight line motion, it can also be integrated explicitly.

Stochastic integrable systems were first introduced in Li (2008) with a more general definition which includes other systems than the present stochastic Euler-Poincaré equations. The present definition is rather abstract and not very explicit, so we will need further refinements by looking at specific examples later in the text. We can nevertheless make a few interesting comments. First, recall that for the previous derivation of the stochastic equations to make sense, we must have $\Phi_i(\mu) = \langle \mu, \sigma_i \rangle$ where σ_i does not depend on μ. This restriction has a direct impact on the possible integrable stochastic systems that we can derive. Indeed, the dynamics of an integrable system lies at the intersection of the level sets of the conserved quantities. The noise must then produce dynamics only along these level sets to preserve integrability. The restriction of being linear in the momentum is in general to restrictive for this to be achieved. Relaxing this assumption of linearity will correspond to the definition of Li (2008) where only stochastic processes with commuting vector field are considered, without any assumption on their form. This generalisation will also be used here in section 3.4 when we treat stochastic integrable hierarchies. In finite dimensions, there is hopefully enough possibilities in the choice of Hamiltonian systems to derive several different integrable stochastic systems. We will study a few of these examples throughout this work such as the free rigid body in 2.2.3 and the stochastic Toda lattice in 2.5.

2.2.2 The Fokker-Planck equation and stationary distribution

Once we have a stochastic process, we can study its associated Fokker-Planck (FP) equation which describes the dynamics of the probability density of the stochastic process. Before going any further, we have to mention that although most of what we said before can be applied to infinite dimensional systems (see for example Holm (2015)), the study of the Fokker-Planck equation is limited to finite dimensional systems for obvious dimensionality reasons.

The first step in the derivation of the Fokker-Planck equation is to compute the Kolmogorov operator or the infinitesimal generator of the stochastic process, see definition 1.2.9. This operator can directly be computed by taking the conditional expectation of the Lie-Poisson formulation in equation (2.2.15). For this, we need the Itô correction, then, the stochastic term can be removed by using the conditional expectation to obtain the Kolmogorov operator

$$\{(\mu) := \left\langle \mathrm{ad}^*_\xi \mu, \frac{\partial f}{\partial \mu} \right\rangle - \frac{1}{2} \sum_i \left\langle \mathrm{ad}^*_{\sigma_i} \mu, \frac{\partial}{\partial \mu} \left\langle \mathrm{ad}^*_{\sigma_i} \mu, \frac{\partial f}{\partial \mu} \right\rangle \right\rangle . \tag{2.2.18}$$

As before, we can explicitly use the Hamiltonian and stochastic potential to rewrite the Kolmogorov operator in a Lie-Poisson form as

$$\mathcal{L} f(\mu) = \{h, f\} - \frac{1}{2} \sum_i \{\Phi_i, \{\Phi_i, f\}\} . \tag{2.2.19}$$

Notice that it is the same object as the Itô correction of the Lie-Poisson form of the equation derived previously. The last step is to compute L^*, the adjoint of the Kolmogorov operator, or the so-called forward Kolmogorov operator, using the definition of adjoint

$$\int f(\mu) \mathcal{L}^* \mathbb{P}(\mu) d\mu = \int \mathbb{P}(\mu) \mathcal{L} f(\mu) d\mu .$$

Proposition 2.2.6. *The forward Kolmogorov operator is*

$$\mathcal{L}^* \mathbb{P}(\mu) = -\{h, \mathbb{P}\} - \frac{1}{2} \sum_i \{\Phi_i, \{\Phi_i, \mathbb{P}\}\} . \tag{2.2.20}$$

Proof. We first consider the drift term of (2.2.19) and use the Leibniz property of the Lie-Poisson bracket to get

$$\int \mathbb{P}\{h, f\} d\mu = -\int f\{h, \mathbb{P}\} d\mu + \int \{h, f\mathbb{P}\} d\mu . \tag{2.2.21}$$

Recall that a non-degenerate Poisson bracket can be expressed in term of a

symplectic form as $\{F, G\} = \omega(X_F, X_G)$ where X_F and X_G are the Hamiltonian vector fields associated to the function F and G. Now, the symplectic form on the coadjoint orbit is an exact 2-form, that is $\omega = d\theta$ for some 1-form θ. These facts mean that provided that we have vanishing or periodic boundary conditions on the coadjoint orbits, the integration of the Lie-Poisson bracket corresponds to the integration of the exterior differential of a 1-form. This integral vanishes by using Stokes theorem and suitable boundary conditions. For the double bracket term, we similarly obtain

$$\begin{aligned}\int \mathbb{P}\{\Phi_i, \{\Phi_i, f\}\}d\mu &= -\int \{\Phi_i, f\}\{\Phi_i, \mathbb{P}\}d\mu + \int \{\Phi_i, \{\Phi_i, f\}\mathbb{P}\}d\mu \\ &= \int f\{\Phi_i, \{\Phi_i, \mathbb{P}\}\}d\mu - \int \{\Phi_i, f\{\Phi_i, \mathbb{P}\}\}d\mu \\ &\quad + \int \{\Phi_i, \{\Phi_i, f\}\mathbb{P}\}d\mu\,,\end{aligned}$$

where the last two terms vanishes from the same argument as before, that is using integration by parts. Collecting terms give the forward Kolmogorov operator (2.2.20). ■

From the forward Kolmogorov operator, the Fokker-Planck equation for the probability distribution $\mathbb{P}(t, \mu)$ is

$$\frac{d}{dt}\mathbb{P} = \{h, \mathbb{P}\} + \frac{1}{2}\sum_i \{\Phi_i, \{\Phi_i, \mathbb{P}\}\}\,. \tag{2.2.22}$$

The second term is new and describes the diffusion generated by the noise. For illustration, in the case of one-dimensional Brownian motion, the Fokker-Planck equation corresponds to the heat equation where the double bracket term is the Laplacian.

The Hörmander condition and ergodicity

The Fokker-Planck equation is a linear but high dimensional PDE which cannot, in general, be solved analytically. One interesting feature is the existence of a unique stationary distribution which relies on the property

that the drift and volatility vector fields of the stochastic system satisfy the Hörmander condition. We refer, for example, to the introduction of Hairer (2011) for more detail about this topic. We first state its standard definition based on a generic stochastic process $x_t \in \mathbb{R}^n$

$$dx_t = V_0(x_t)dt + \sum_i V_i(x_t) \circ dW_t^i \,, \tag{2.2.23}$$

for vector fields $V_i(x) : \mathbb{R}_n \to \mathbb{R}_n$. We will need the Lie bracket between two vector fields, given by

$$[V, W](x) = DV(x) \cdot W(x) - DW(x) \cdot V(x) \,, \tag{2.2.24}$$

where DV is the matrix derivative $(DV)_{ij} = \partial_i V_j$.

Definition 2.2.2 (Parabolic Hörmander condition)**.** *Define the collections of vector fields*

$$\mathcal{V}_0 = \{V_i | i > 0\} \tag{2.2.25}$$

$$\mathcal{V}_{k+1} = \{[U, V_i], U \in V_k, i \geq 0\} \,. \tag{2.2.26}$$

If $\mathcal{V}(x) = \bigcup_k \mathcal{V}_k(x) = \mathbb{R}^n$*, then* (2.2.23) *satisfies the parabolic Hörmander condition.*

This condition means that the vector fields V_i and all their commutators must span the entire phase space of the stochastic process (2.2.23). With this property, one can prove the Hörmander theorem of Hörmander (1967).

Theorem 2.2.2 (Hörmander theorem)**.** *A stochastic process with the Hörmander conditions has solutions which admit a smooth stationary distribution with respect to the Lebesgue measure on the phase space.*

In our case, we need to adapt this definition, or none of our stochastic processes will fulfil the Hörmander condition. Indeed, we already saw that the multiplicative noise that we use allows us to preserve the coadjoint orbits of the deterministic system which is a subset of the full phase space. We thus need to modify Hörmander condition to $\mathcal{V}(\mu) = T_\mu \mathcal{O}$. Equivalently, we

require that $\mathcal{V}(\mu)$ span the tangent space to the coadjoint orbit at the point μ. The tangent space of a coadjoint orbit is computed from the definition of the coadjoint orbit and is given as

$$T_\mu \mathcal{O} = \{\mathrm{ad}^*_\xi \mu \,|\, \xi \in \mathfrak{g}\}\,. \tag{2.2.27}$$

From this we see that Hörmander condition is satisfied if $\mathrm{span}(\sigma_i) = \mathfrak{g}$, in this case only $\mathcal{V}_0$ is needed as $V_i = \mathrm{ad}^*_{\sigma_i}\mu$. If this is not the case, we need to go further and use the explicit form of the vector fields V_i to compute

$$[\mathrm{ad}^*_{\sigma_i}\mu, \mathrm{ad}^*_{\sigma_j}\mu] = \left(\mathrm{ad}^*_{\sigma_i}\mathrm{ad}^*_{\sigma_j} - \mathrm{ad}^*_{\sigma_j}\mathrm{ad}^*_{\sigma_i}\right)\mu\,,$$

and verify that $T_\mu\mathcal{O}$ is spanned by these vector fields, or higher commutators.

We now need to comment on this feature of stochastic processes in the case of integrable stochastic systems. As we said earlier, the integrability property corresponds to the system being restricted to a submanifold of the phase space which is the intersection of the coadjoint orbit with the level set of additional conserved quantities. The noise must then preserve this subset and the Hörmander condition will thus never be satisfied for the entire coadjoint orbit but only for this particular subset.

Stationary distribution

We now describe the stationary distribution, or the stationary solution of the Fokker-Planck equation, given as the unique solution of

$$\{h, \mathbb{P}_\infty\} = -\sum_i \{\Phi_i, \{\Phi_i, \mathbb{P}_\infty\}\}\,. \tag{2.2.28}$$

Proposition 2.2.7. *The stationary distribution is constant on the coadjoint orbit if the Hörmander condition for the underlying stochastic process holds on the same orbit and the system is not integrable.*

Proof. First, recall that the Hörmander theorem states that the stationary distribution stationary distribution exists and is smooth with respect to the

Lebesgue measure on the coadjoint orbit. Second, because the system is not integrable, that is h and σ_i are arbitrary, we need $\{h, \mathbb{P}\}$ and $\{\Phi_i, \mathbb{P}\}$ to vanish independently. They vanish only if $\mathbb{P}$ is in the kernel of the Poisson bracket, that is if $\mathbb{P}$ is an arbitrary function of the Casimir or invariant function on the coadjoint orbit. $\mathbb{P}$ must only depend on the initial condition μ_0 which defines the orbit $\mathcal{O}_{\mu_0}$. This dependence is then removed after the normalisation of $\mathbb{P}$ by imposing that it is a probability distribution, so that $\int_{\mathcal{O}_{\mu_0}} \mathbb{P}_\infty(\mu) d\mu = 1$. Furthermore, we know from the Hörmander theorem that the stationary distribution is unique, and we found one, so we proved the result of this proposition. ■

Notice that if the volume of the coadjoint orbit is infinite, the stationary distribution is not defined and we need to use more general measures. We will avoid by using only compact coadjoint orbits corresponding to compact Lie groups.

2.2.3 The stochastic free rigid body

We will now apply this theory to the free rigid body which is our main example to illustrate the stochastic deformation of geometric mechanics in finite dimensions.

The free rigid body

We will only give the necessary details for a basic understanding of the dynamics of the free rigid body, and we refer to Marsden and Ratiu (1999) for a complete description of this system. The Lie group that describes the rigid body is $SO(3)$, the group of rotations in three dimensions. This group naturally acts on vectors in $\mathbb{R}^3$ which represent the positions of each infinitesimal element of the solid. The dynamics of the body can thus be described by a path on $R(t) \in SO(3)$. The body being free, that is there is no external force, the Lagrangian is purely kinetic. The Lie group has a natural action on its tangent vector $\dot{R}(t) \in T_R SO(3)$ corresponding to the velocity vector. On top of that, the group action preserves the norm of these vectors and

the kinetic energy so the theory of reduction by symmetry can be applied. The reduced velocity variable is thus $\Omega := R^{-1}\dot{R} \in \mathfrak{so}(3)$ and the reduced Lagrangian is

$$l(\boldsymbol{\Omega}) = \frac{1}{2}\boldsymbol{\Omega} \cdot \mathbb{I}\boldsymbol{\Omega} := \frac{1}{2}\boldsymbol{\Omega} \cdot \boldsymbol{\Pi}\,, \tag{2.2.29}$$

where $\mathbb{I} = \mathrm{diag}(\mathbb{I}_1, \mathbb{I}_2, \mathbb{I}_3)$ is the tensor of inertia which captures the mass distribution of the solid in three principal moments of inertia. Geometrically speaking, it is a map $\mathbb{I} : \mathfrak{g} \to \mathfrak{g}^*$, and defines the reduced momentum variable $\Pi = \mathbb{I}\Omega \in \mathfrak{g}^*$. We are using bold fonts in the definition of the Lagrangian which means that in fact $\boldsymbol{\Omega} \in \mathbb{R}^3 \cong \mathfrak{so}(3)$ via the isomorphism between 3-dimensional anti-symmetric matrices and $\mathbb{R}^3$. With this isomorphism, the trace pairing becomes the dot product, the matrix commutator the cross product and $\mathrm{ad} = -\mathrm{ad}^*$. We thus obtain the following Euler-Poincaré equation on the group of rotations

$$\dot{\boldsymbol{\Pi}} = \boldsymbol{\Pi} \times \boldsymbol{\Omega}\,. \tag{2.2.30}$$

We will not study the dynamics of this system in detail here but only mention that the coadjoint orbit is the momentum sphere with Casimir $C(\boldsymbol{\Pi}) = \|\boldsymbol{\Pi}\|^2$ and the trajectories lie in the intersection of this sphere and the ellipsoids given by the conservation of energy. We will see more detail on the dynamics with the noise later on in this section.

The stochastic free rigid body

We will now follow the present theory of stochastic deformation and apply it to the free rigid body. We first need a stochastic potential linear in the momentum, so we choose

$$\Phi_i(\boldsymbol{\Pi}) = \boldsymbol{\sigma}_i \cdot \boldsymbol{\Pi}\,, \tag{2.2.31}$$

where $\sigma_i \in \mathfrak{so}(3)$ are orthogonal fixed elements. The stochastic free rigid body equation is then, from the stochastic Euler-Poincaré equation (2.2.2),

$$d\mathbf{\Pi} + \mathbf{\Pi} \times \mathbf{\Omega}\, dt + \sum_i \mathbf{\Pi} \times \boldsymbol{\sigma}_i \circ dW_t^i = 0\,, \tag{2.2.32}$$

and preserves the coadjoint orbit as it is easily checked that the Casimir $C(\mathbf{\Pi}) = \|\mathbf{\Pi}\|^2$ is preserved by the flow. The equivalent Itô formulation of (2.3.44) is given by

$$d\mathbf{\Pi} + \mathbf{\Pi} \times \mathbf{\Omega}\, dt + \frac{1}{2}\sum_i (\mathbf{\Pi} \times \boldsymbol{\sigma}_i) \times \boldsymbol{\sigma}_i\, dt + \sum_i \mathbf{\Pi} \times \boldsymbol{\sigma}_i\, dW_t^i = 0\,, \tag{2.2.33}$$

where again one can verify that the Casimir is conserved using Itô calculus.

It is now interesting to study the energy conservation of this system, as this is the only other deterministically conserved quantity apart from the Casimir. Using Itô's Lemma, we compute the time evolution of the energy to obtain

$$\begin{aligned} dh(\mathbf{\Pi}) = \sum_i (\mathbf{\Pi} \times \boldsymbol{\sigma}_i) \cdot [\mathbb{I}^{-1}(\mathbf{\Pi} \times \boldsymbol{\sigma}_i) - (\mathbf{\Omega} \times \boldsymbol{\sigma}_i)]\, dt \\ + 2\sum_i (\mathbf{\Pi} \times \boldsymbol{\sigma}_i) \cdot \mathbf{\Omega}\, dW_t^i\,. \end{aligned} \tag{2.2.34}$$

From this formula, we can see that the energy is not conserved in general. In order to study the energy evolution we numerically computed the probability density for the energy, that is the solution of the Fokker-Planck equation associated to the process (2.2.34) and display the result in figure 2.1, for two different initial conditions. In practice, we ran many independent simulations of the system with different noise realisations and then counted the number of rigid bodies for each energy bin between the two possible extreme values of the energy. The latter are found from the conservation of the momentum sphere to be at the two stable equilibria associated with the directions of the largest and lowest moment of inertia. For large times, the initial conditions do not affect the final distribution, demonstrating the ergodicity of this system. The second remark is that this final density is not constant. This is explained by the fact that the volume of the coadjoint orbit for each value of the energy

is different, thus the number of rigid bodies, as they are uniformly distributed on the momentum sphere. Recall the energy is not conserved unless we make

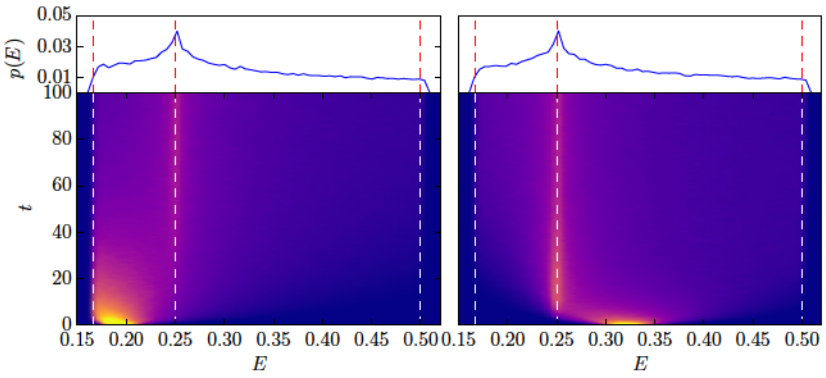


Figure 2.1: We display in this figure the density distribution of the stochastic process (2.2.34) with $\mathbb{I} = \mathrm{diag}(1,2,3)$ for two different initial conditions. The blue colour means that the probability vanishes, and the yellow colour that the probability is high. This figure demonstrates the ergodicity of the system and the property that the constant energy regions of the momentum sphere have different volumes. The numerical scheme used here is the Heun method for the stochastic term and the standard odeint method of the Scipy package in python.

a particular choice for $\mathbb{I}$ and $\boldsymbol{\sigma}_i$ such that $(\boldsymbol{\sigma}_i \times \boldsymbol{\Pi}) \perp (\mathbb{I}^{-1}\boldsymbol{\Pi})$ holds. One can easily see that this the case only if there is a single Wiener process associated to the fixed vector $\sigma\boldsymbol{n}$, where $\boldsymbol{n}$ is the unit vector and σ a real number. On top of that, the moment of inertia must be uniform in the subspace perpendicular to $\boldsymbol{n}$, that is, there is a number I such that

$$\mathbb{I}(\boldsymbol{e}-\boldsymbol{n}) = \mathbb{I}_1(\boldsymbol{e}-\boldsymbol{n})\,,$$

where $\boldsymbol{e} = (1,1,1)$. Without loss of generality, we can set $\boldsymbol{n} = (0,0,1)^T$ and $\mathbb{I} = \mathrm{diag}(\mathbb{I}_1, \mathbb{I}_1, \mathbb{I}_3)$ such that the dynamics takes place in circles perpendicular to the axis $(1,0,0)$. One can then verify that under this condition, the energy is conserved along the stochastic paths that solve (2.3.44).

In fact, this condition which gives integrability of this system is equivalent to the condition that the Hamiltonian and the stochastic potential are commuting integral of motions. In this simple case, we can go even go further

by writing the stochastic free rigid body equation in components.

$$
\begin{aligned}
d\Pi_1 &= \Pi_2(a\Pi_3 dt + \sigma \circ dW) \\
d\Pi_2 &= -\Pi_1(a\Pi_3 dt + \sigma \circ dW) \\
d\Pi_3 &= 0\,,
\end{aligned}
$$

where $a := \frac{I-I_3}{II_3}$. This equation is the classical Kubo oscillator introduced by Kubo et al. (1991) in the context of statistical physics. Its solution can be explicitly given in term of the Wiener process W_t

$$
\begin{aligned}
\Pi_1(t) &= \Pi_1(0)\cos(a\Pi_3 t + \sigma W_t) - \Pi_2(0)\sin(a\Pi_3 t + \sigma W_t) \\
\Pi_2(t) &= \Pi_2(0)\cos(a\Pi_3 t + \sigma W_t) + \Pi_1(0)\sin(a\Pi_3 t + \sigma W_t)\,.
\end{aligned}
$$

The outcome of this analysis is that integrability is rare for these stochastic systems as it only appears for a single set of combinations of the parameters.

The Fokker-Planck equation and stationary distributions

We can now study the Fokker-Planck equation in more detail. We start with the conditions on the $\boldsymbol{\sigma}_i$ such that the Hörmander condition is satisfied. In the generic case of isotropic noise described by three Wiener processes with $\boldsymbol{\sigma}_i = \sigma \boldsymbol{e}_i$, where $\boldsymbol{e}_{ii}$ is an orthonormal basis of $\mathbb{R}^3$ and σ a real number, the Hörmander is satisfied. More noise fields could be added or removed without changing this result. In fact, only one noise field $\boldsymbol{\sigma}$ is needed if the moments of inertia are all distinct. Indeed, the noise will always be in along the same direction, and the dynamics will bring enough mixing to obtain ergodicity. In the case of degeneracy for the moments of inertia, if the noise amplitude is not as in the stochastically integrable case, the Hörmander condition will be fulfilled. This feature demonstrates that the Hörmander condition on the full coadjoint orbit prevents the system from being integrable. On the contrary, if the system is integrable, the Hörmander condition will hold only on the 1-dimensional subspace of solutions, as one can see for the Kubo oscillator.

We now compute the Fokker-Planck equation for the stochastic free rigid

body to get

$$\frac{d}{dt}\mathbb{P} + (\mathbf{\Pi} \times \mathbf{\Omega}) \cdot \nabla \mathbb{P} + \frac{1}{2} \sum_i (\mathbf{\Pi} \times \boldsymbol{\sigma}_i) \cdot \nabla [(\mathbf{\Pi} \times \boldsymbol{\sigma}_i) \cdot \nabla \mathbb{P}] = 0\,, \qquad (2.2.35)$$

where ∇ is the gradient with respect to the independent variable $\mathbf{\Pi} \in \mathbb{R}^3$. As we proved for the general case in Proposition 2.2.7, the stationary distribution of the Fokker-Planck equation is a constant on the coadjoint orbits, which is here the momentum sphere.

The interesting thing not captured in the stationary distribution is the dynamics of the probability density $\mathbb{P}$ when it is not stationary and where the combination of the advection and the diffusion due to the noise are visible. To illustrate these effects, we solved the Fokker-Planck equation numerically with a Monte-Carlo sampling method with a delta initial distribution. We display several snapshots of the solution in figure 2.2. The solution was obtained by a Monte-Carlo sampling of the Fokker-Planck equation, that is many simulations of the stochastic rigid body were computed with a different realisation of the noise each time. We used a split step method with the exact solution for the stochastic term and the plot represents the density of rigid bodies on the momentum sphere at a given time.

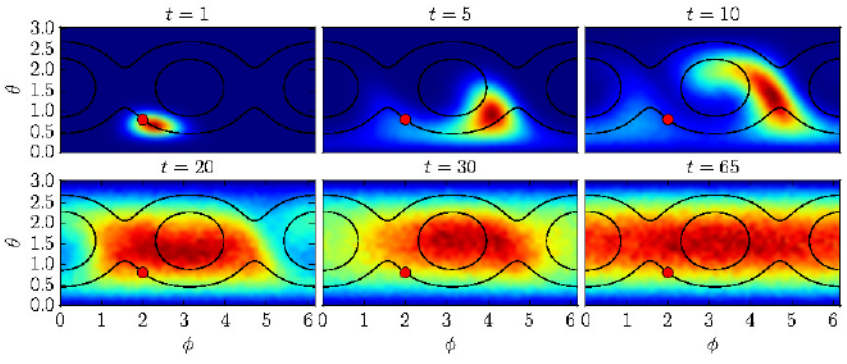


Figure 2.2: In this figure we display six snapshots of a numerical simulation of the Fokker-Planck evolution of the free rigid body, with noise amplitude $\boldsymbol{\sigma}_i = \sigma e_i$ and moment of inertia $\mathbb{I} = \mathrm{diag}(1, 2, 3)$. The red dot is the initial delta distribution that converges to the uniform distribution as time goes to infinity. During the process, the advection term coming from the underlying rigid body dynamics accelerates the mixing by transporting the solution in various regions of the momentum sphere.

2.2.4 Summary and outlook

Most of this section was devoted to the introduction of noise in mechanical systems with symmetries. We saw three different and almost equivalent derivations in section 2.2.1: the Clebsch variables, stochastic forces with the Lagrange-d'Alembert principle and the stochastic Hamilton-Pontryagin variational principle. The most general derivation turned out to be the last which can even allow a noise to depend on advected parameters. We did not investigate these systems here and refer to Arnaudon et al. (2017a) for some results about them. One of the interesting features is that the coadjoint orbits are not compact and the study of the stationary distribution needs to be done via marginal distributions obtained by integrating over parts of the coadjoint orbit. To study stationary distributions, we obtained a general and simple formulation of the Fokker-Planck equation in term of the Lie-Poisson bracket and directly obtained the result of a constant stationary distribution on the coadjoint orbit.

Regarding the integrability of the stochastic deformation, the deterministic free rigid body is integrable for all values of the moment of inertia, but with noise only a few configurations of the system preserve integrability. This is the case when the body has a cylindrical symmetry with the noise along the axis of symmetry. This system is, in fact, the stochastic Kubo oscillator as the dynamics is restricted to circles on the plane, with explicit exact solutions in term of the noise. For the heavy top which is the classic example of the semi-direct product extension, a similar result holds where the noise must be in the axis of symmetry of the Lagrange top (a heavy top with cylindrical symmetry), and explicit solutions can also be computed, see Arnaudon et al. (2017a).

At this point, there are no major open problems but rather questions of how to extend this approach to other types of systems. We already mentioned semi-direct products but other structures could be considered such as the Lagrange-Poincaré reduction, cocycles, reduction with non-holonomic constraints, etc. Some of these extensions are work in progress and some have not yet been investigated.

2.3 Dissipation and random attractors

So far we studied stochastic Hamiltonian systems which can be seen as non-dissipative systems despite the noise. Indeed, the noise introduces diffusion rather than dissipation as we saw in the energy behaviour of the stochastic rigid body. This difference is important especially if we want to understand the interaction of noise and energy dissipation in dynamical systems and in particular the possible equilibrium regime resulting from the balance between these two effects. The aim of this section is to introduce dissipation in the stochastic Euler-Poincaré equation to perform such a study without destroying the coadjoint orbit structure. For this, we will introduce in the next section 2.3.1 a nonlinear dissipation term of double bracket type. With this dissipation, we will have two directions to study its interaction with the noise. We can either average over the realisations of the noise or fix a realisation and average over the initial conditions. We will start in section 2.3.2 by deriving the Fokker-Planck equation and compute its stationary distributions to show that they are the Gibbs distribution of statistical physics. We will then briefly investigate this connection with statistical physics in section 2.3.3. We will then take the second point of view in section 2.3.4 to study the attractors of these systems. This will lead us to the notion of random attractors, chaos and sensitivity to initial conditions, a central theme in the theory of random dynamical systems.

2.3.1 Dissipation with double bracket

The introduction of dissipation in mechanical systems with symmetries should be carefully done to retain important geometrical properties such as preservation of coadjoint orbits. Such a theory of geometric dissipation has been developed in recent years and can be applied to a large class of mechanical systems, including finite or infinite dimensional as well as semi-direct products systems. The dissipative term that we will add here is a generalisation of the double bracket dissipation first introduced by Brockett (1991) in the context of diagonalizations of matrices in linear programming and shortly after by Bloch et al. (1992) for a different description of the completely integrable

Toda lattice. The double bracket theory was then further extended to mechanical systems in Bloch et al. (1996) and more recently in Gay-Balmaz and Holm (2013, 2014) where both energy and Casimir dissipations were studied and applied to fluid mechanics. We will not discuss further the use of double brackets for fluid dynamics here as we will only consider double brackets for finite dimensional systems. As studied in Gay-Balmaz and Holm (2013), the dissipation of the Casimir while preserving the energy is often used to model turbulence cascade in fluid dynamics. We want here to do the reverse, that is to preserve the Casimirs and dissipate the energy so that we have a dissipative mechanics on the coadjoint orbits.

The double bracket term

We begin by introducing this dissipation in the classical example of the rigid body which also illustrates the fact that the Casimir and energy dissipation have a similar form. Indeed, for the Casimir dissipation, the equation of motion is

$$\dot{\mathbf{\Pi}} = \mathbf{\Pi} \times \mathbf{\Omega} + \theta\, \mathbf{\Omega} \times (\mathbf{\Omega} \times \mathbf{\Pi})\,, \tag{2.3.1}$$

and for energy dissipation it reads

$$\dot{\mathbf{\Pi}} = \mathbf{\Pi} \times \mathbf{\Omega} + \theta\, \mathbf{\Pi} \times (\mathbf{\Pi} \times \mathbf{\Omega})\,, \tag{2.3.2}$$

where the inverse timescale $\theta \in \mathbb{R}$ parametrises the dissipation strength. One can directly verify that the Casimir or the energy decay until the system reaches a position $\mathbf{\Omega}_e$ where $\|\mathbf{\Omega}_e \times \mathbf{\Pi}_e\|^2 = 0$. This condition is satisfied when the rigid body is at an equilibrium position, i.e. aligned with a principal axis of inertia.

We can now introduce the general dissipative term of Gay-Balmaz and Holm (2013, 2014) to get the dissipative Euler-Poincaré equation

$$\frac{d}{dt}\mu + \mathrm{ad}^*_{\frac{\delta h}{\delta \mu}}\mu + \theta\, \mathrm{ad}^*_{\frac{\delta C}{\delta \mu}} \left[\frac{\delta C}{\delta \mu}, \frac{\delta h}{\delta \mu}\right]^\flat = 0\,, \tag{2.3.3}$$

where the isomorphism $\flat : \mathfrak{g} \to \mathfrak{g}^*$ is associated to a given pairing, that can in general be different from the natural pairing of the Lie algebra. Here, we will always use the natural pairing, but for fluid dynamics, other choices may be possible, see Gay-Balmaz and Holm (2013).

In this general case, the energy decay is given by

$$\frac{d}{dt}h(\mu) = -\theta \left\| \left[\frac{\delta C}{\delta \mu}, \frac{\delta h}{\delta \mu} \right] \right\|^2 , \tag{2.3.4}$$

where the norm is associated to the pairing used in the definition of the flat operation. This equation shows that the system is forced towards a position where the right-hand side vanishes. This condition is compatible with an equilibrium solution of the original deterministic system, see Gay-Balmaz and Holm (2014) for an extensive discussion on this topic. We will come back to this point in section 2.4 where we will study the stochastic stability of these equilibrium solutions via the energy-Casimir method.

Derivation from a variational principle

This additional dissipative term which we introduced by hand can also be derived from a variational principle. To do this derivation we need to use a forcing term which would not come from a potential energy, as it is a dissipative force. One of the best ways to do this is to use the reduced Lagrange-d'Alembert principle. We refer to the monograph Bloch et al. (2003) for more detail on this theory and will only briefly explain it in our particular context. This variational principle is used to include forces in the Euler-Poincaré equation and is given as

$$\delta \left(\int l(\xi)\, dt \right) + \int \langle F(\xi), \eta \rangle \, dt = 0 \,, \tag{2.3.5}$$

where the constrained variations $\delta\xi = \dot{\eta} - [\xi, \eta]$ are only used for the variation of the first term. The second term includes the force $F : \mathfrak{g} \to \mathfrak{g}^*$ and the free variations η. In the stochastic Euler-Poincaré equation with energy

dissipation, the Lagrange-d'Alembert variational principle is

$$\delta\left(\int l(\xi)\,dt\right)+\theta\int\left\langle\left[\frac{\partial C}{\partial\mu},\delta\xi\right],\left[\frac{\partial C}{\partial\mu},\eta\right]\right\rangle\,dt=0\,,\tag{2.3.6}$$

where the stochastic constrained variations are $\delta\xi=d\eta+[\xi,\eta]\,dt+\sum_i[\sigma_i,\eta]\circ dW_t^i$. We can then compute the dynamical equation that extremizes this variational principle to get

$$\left\langle\frac{\partial l(\xi)}{\partial\xi},\delta\xi\right\rangle+\theta\left\langle\mathrm{ad}^*_{\frac{\partial C}{\partial\mu}}\left[\frac{\partial C}{\delta\mu},\frac{\partial h}{\partial\mu}\right]^\flat,\delta\xi\right\rangle=0\,.\tag{2.3.7}$$

Substituting the formula for the variations $\delta\xi$, integrating by parts and isolating the variable η we get, in the Hamiltonian form

$$d\mu+\mathrm{ad}^*_{\frac{\partial h}{\partial\mu}}\mu\,dt+\theta\,\mathrm{ad}^*_{\frac{\partial C}{\partial\mu}}\left[\frac{\partial C}{\partial\mu},\frac{\partial h}{\partial\mu}\right]^\flat dt+\sum_i\mathrm{ad}^*_{\sigma_i}\mu\circ dW_t^i=0\,,\tag{2.3.8}$$

where $\xi=\frac{\partial h}{\partial\mu}$.

The complete stochastic model

We now come back to the equation (2.3.8) and notice the following fact. Introducing noise into this equation amounts in replacing the velocity field ξ by $\xi dt+\sum_i\sigma_i\circ dW_t^i$. This replacement is only done in the advection term and not in the dissipation term which also contains an explicit ξ variable. Physically speaking, this equation is not consistent unless an extra stochastic term is added in the dissipation. Notice that this introduction of noise can also be seen as replacing the Hamiltonian h with the stochastic quantity $h\,dt+\Phi_i\circ dW_t^i$ in both the advection and dissipation. Using a variational principle, we will derive the physically consistent equation

$$\begin{aligned}d\mu+\mathrm{ad}^*_{\frac{\partial h}{\partial\mu}}\mu\,dt+\theta\,\mathrm{ad}^*_{\frac{\partial C}{\partial\mu}}\left[\frac{\partial C}{\partial\mu},\frac{\partial h}{\partial\xi}\right]^\flat dt+\theta\sum_i\mathrm{ad}^*_{\frac{\partial C}{\partial\mu}}\left[\frac{\partial C}{\partial\mu},\sigma_i\right]^\flat\circ dW_t^i\\+\sum_i\mathrm{ad}^*_{\sigma_i}\mu\circ dW_t^i=0\,.\end{aligned}\tag{2.3.9}$$

We will use a combination of Lagrange-d'Alembert for the dissipative force and Hamilton-Pontryagin principle for the noise, as introduced in 2.2.1.

Proposition 2.3.1. *The variational principle*

$$\begin{aligned}\delta\left(\int l(\xi)\,dt\right) + \theta\int\left\langle\left[\frac{\partial C}{\partial\mu}, g^{-1}dg\right], \left[\frac{\partial C}{\partial\mu}, g^{-1}\delta g\right]\right\rangle \\ + \delta\left(\sum_i\int\langle\mu, g^{-1}dg - \xi dt - \sigma_i\circ dW_t^i\rangle\right) = 0\,,\end{aligned} \tag{2.3.10}$$

yields the stochastic process (2.3.9) *for free variations* $\delta g, \delta\xi$ *and* $\delta\mu$.

Proof. From this variational principle, the first and last term are treated as in proposition 2.2.3 using the Hamilton-Pontryagin theory. We do the computation for the second term only, to get, using the form of $g^{-1}dg$ from the last term,

$$\int\left\langle g^{-1}\delta g, \mathrm{ad}^*_{\frac{\partial C}{\partial\mu}}\left[\frac{\partial C}{\partial\mu}, g^{-1}dg\right]^\flat\right\rangle = \int\left\langle g^{-1}\delta g, \mathrm{ad}^*_{\frac{\partial C}{\partial\mu}}\left[\frac{\partial C}{\partial\mu}, \xi dt + \sigma_i\circ dW_t^i\right]^\flat\right\rangle . \tag{2.3.11}$$

The variations δg being arbitrary, we obtain the forced stochastic Euler-Poincaré equation (2.3.9). ■

Notice that such a derivation will not work in the case of Casimir dissipation with preservation of energy. Indeed, we would need a different Lagrange-d'Alembert term which should contain two $g^{-1}dg$ terms instead of two Casimir dependent terms. This different term cannot work in the stochastic setting unless we use ξ instead of $g^{-1}dg$. We can explain this issue by the fact that there is no energy conservation in the stochastic system, so trying to dissipate the Casimir while preserving the energy does not make sense.

Although this derivation is physically more consistent, we will not use it here, and prefer the simpler derivation which produces equation (2.3.8). One can argue for this choice by assuming that the amplitude of the dissipation and

the noise are taken to be small enough so that the extra term proportional to $\sigma\theta$ can be neglected, as well as its Itô correction terms, proportional to $\sigma^2\theta$ and $\theta^2\sigma^2$.

It is interesting that only the Hamilton-Pontryagin principle gives a physically consistent dissipative equation. This fact may be an indication that this variational principle is the most general to include noise into these dynamical systems with symmetries. This can also be seen for semi-direct product structures, where the Clebsch derivation does not allow the noise fields to depend on the advected quantities whereas the Hamilton-Pontryagin principle does not have this restriction.

2.3.2 The Fokker-Planck equation and the Gibbs distribution

After having added the double bracket dissipation to the stochastic mechanical systems, we will study the evolution of the probability distribution via the Fokker-Planck equation and in particular the stationary distribution of this process.

The Fokker-Planck equation for the dissipative stochastic process (2.3.8) is straightforward to compute because the extra term only modifies the advection part of the Fokker-Planck equation. We obtain

$$\begin{aligned}\frac{d}{dt}\mathbb{P}(\mu) + \{h, \mathbb{P}\} + \theta \left\langle \left[\frac{\partial \mathbb{P}}{\partial \mu}, \frac{\partial C}{\partial \mu}\right], \left[\frac{\partial h}{\partial \mu}, \frac{\partial C}{\partial \mu}\right] \right\rangle \\ - \frac{1}{2}\sum_i \{\Phi_i, \{\Phi_i, \mathbb{P}\}\} = 0\,.\end{aligned} \tag{2.3.12}$$

Already from the Fokker-Planck equation, we can guess the form of the invariant solution which will be a balance between the dissipative and the diffusion term. The dissipative term contains only one derivative of $\mathbb{P}$ whereas the diffusion has two derivatives. We thus expect to obtain a solution proportional to the exponential of a function of μ. In the following theorem, we compute the form of the stationary distribution which turns out to be the Gibbs distribution of statistical physics.

Theorem 2.3.1. *Let $\mathfrak{g} \cong \mathfrak{g}^*$ be a compact semi-simple Lie algebra and e_i an*

orthonormal basis of its underlying vector space. For an isotropic noise, that is $\sigma_i = \sigma e_i$ *for a given* $\sigma \in \mathbb{R}$ *and a basis* $\{e_i\}_i$*, the invariant distribution of the Fokker-Planck equation* (2.3.12) *associated to* (2.3.8) *with Casimir* $C = \frac{1}{2}\|\mu\|^2$ *is the Gibbs distribution*

$$\mathbb{P}_\infty(\mu) = Z^{-1} e^{-\frac{2\theta}{\sigma^2} h(\mu)}, \tag{2.3.13}$$

where Z *is the normalisation constant that enforces* $\int_{\mathcal{O}} \mathbb{P}_\infty(\mu) d\mu = 1$*, or partition function in the language of statistical mechanics.*

Proof. The first term vanishes if we choose the invariant solution in the kernel of the Hamiltonian term, or of the form $\mathbb{P}_\infty(\mu) = f(h(\mu))$ where the function f is to be computed. We first use the bi-invariance of the pairing, guaranteed by the semi-simple property of the Lie algebra to rewrite the dissipative term of the Fokker-Planck equation as

$$\begin{aligned} \theta \left\langle \left[\frac{\partial \mathbb{P}}{\partial \mu}, \frac{\partial C}{\partial \mu}\right], \left[\frac{\partial h}{\partial \mu}, \frac{\partial C}{\partial \mu}\right] \right\rangle &= \theta \left\langle \frac{\partial \mathbb{P}}{\partial \mu}, \mathrm{ad}_{\frac{\partial C}{\partial \mu}} \left[\frac{\partial C}{\partial \mu}, \frac{\partial h}{\partial \mu}\right] \right\rangle \\ &= \theta\, \mathbf{d} \left(f(h) \mathrm{ad}_{\frac{\partial C}{\partial \mu}} \left[\frac{\partial C}{\partial \mu}, \frac{\partial h}{\partial \mu}\right] \right), \end{aligned} \tag{2.3.14}$$

where $\mathbf{d}$ stands for the divergence operator and we used partial derivatives instead of variational derivatives. Notice that we also needed to use the fact that the coadjoint action for semi-simple Lie algebras is divergence-free. We already proved this point in the computation of the Fokker-Planck equation from the backwards Kolmogorov operator. We now consider the diffusion term, and we manipulate it using the quadratic Casimir, that is $\mu = \frac{\partial C}{\partial \mu}$, to

obtain, for $\Phi_i = \langle \mu, \sigma_i \rangle$

$$
\begin{aligned}
-\sum_i \{\Phi_i, \{\Phi_i, \mathbb{P}\}\} &= -\sigma^2 \sum_i \left\langle \mathrm{ad}_{e_i} \frac{\partial C}{\partial \mu}, \frac{\partial}{\partial \mu} \left\langle \mathrm{ad}_{e_i} \frac{\partial C}{\partial \mu}, \frac{\partial \mathbb{P}}{\partial \mu} \right\rangle \right\rangle \\
&= \sigma^2 \sum_i \mathbf{d} \left(f'(h) \mathrm{ad}_{e_i} \frac{\partial C}{\partial \mu} \left\langle \mathrm{ad}_{\frac{\partial h}{\partial \mu}} \frac{\partial C}{\partial \mu}, e_i \right\rangle \right) \\
&= \sigma^2 \mathbf{d} \left(f'(h) \mathrm{ad}_{\left[\frac{\partial h}{\partial \mu}, \frac{\partial C}{\partial \mu}\right]} \frac{\partial C}{\partial \mu} \right) \\
&= \sigma^2 \mathbf{d} \left(f'(h) \mathrm{ad}_{\frac{\partial C}{\partial \mu}} \mathrm{ad}_{\frac{\partial C}{\partial \mu}} \frac{\partial h}{\partial \mu} \right) .
\end{aligned}
\tag{2.3.15}
$$

The result then follows by choosing $f(x) = e^{-2\theta x/\sigma^2}$ and comparing the two terms (2.3.14) and (2.3.15). To conclude the proof the uniqueness of this stationary distribution is provided by the Hörmander theorem. ■

Before going into the physical interpretation of this stationary distribution, we want to comment on a few important restrictions on the noise and the Lie algebra we had to use for this proof. First, the noise should be isotropic, that is $\sigma_i = \sigma e_i$ where the e_i span $\mathfrak{g}$. If this condition is relaxed, the diffusion and the dissipation terms will not be directly proportional to each other and the usual Gibbs distribution will not be a valid solution. The other requirement of the Lie algebra being semi-simple is needed for a similar reason. Indeed, the selective decay term is in general

$$
\theta \, \mathbf{d} \left(f(h) \mathrm{ad}^*_{\frac{\partial C}{\partial \mu}} \left[\frac{\partial C}{\partial \mu}, \frac{\partial h}{\partial \mu} \right]^\flat \right) ,
$$

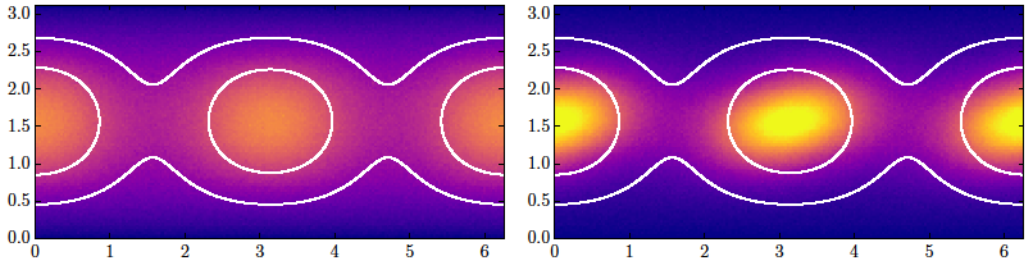
and the diffusion

$$
\sum_i \mathbf{d} \left(f'(h) \mathrm{ad}^*_{e_i} \frac{\partial C}{\partial \mu}^\flat \left\langle \mathrm{ad}^*_{\frac{\partial h}{\partial \mu}} \frac{\partial C}{\partial \mu}^\flat, e_i \right\rangle \right) = \mathbf{d} \left(f'(h) \mathrm{ad}^*_{\left(\mathrm{ad}^*_{\frac{\partial h}{\partial \mu}} \frac{\partial C}{\partial \mu}^\flat \right)^\sharp} \frac{\partial C}{\partial \mu}^\flat \right) .
$$

These two terms are not compatible term unless the pairing is ad-invariant, that is $\mathrm{ad}^* = -\mathrm{ad}$ and the $\flat$ and $\sharp$ musical operators can be dropped (where $\sharp$ is the inverse of $\flat$). In both cases, we know that there is an stationary

distribution, if the Hörmander condition is satisfied, but it is in general difficult to find an explicit form for it. This difficulty can be directly seen in the equation used above for computing the stationary distribution as if the noise is non-isotropic, the ansatz $\mathbb{P}_\infty = f(h(t))$ does not work. The stationary distribution is then a subtle balance between the advection, the dissipation and the diffusion terms.

We display the Gibbs distribution in figure 2.3 for two different noise amplitudes. The isotropic noise exhibits the Gibbs distribution computed previously, and the non-isotropic noise produces a non-symmetric distribution but still close to the Gibbs distribution.



(a) Isotropic noise amplitude with noise amplitudes $\sigma(1,0,0), \sigma(0,1,0)$ and $\sigma(0,0,1)$.

(b) Non-isotropic noise amplitude with noise amplitudes $\sigma(1,0,0)$, and $\sigma(0,0,1)$.

Figure 2.3: This figure displays two stationary distributions of the Fokker-Planck equation for the stochastic free rigid body. In yellow are the high density regions and in dark blue the low density regions. The top panel is a simulation with an isotropic noise which shows the Gibbs distribution and the bottom panel has a non-isotropic noise. The latter has a less symmetric profile and a higher density around the lowest energy equilibrium position. The white lines show a few deterministic trajectories of the rigid body.

2.3.3 Statistical physics

This result of obtaining the Gibbs measure as the stationary distribution is important and connects with the classical theory of statistical physics. This theory begins with the description of gases or lattices in contact with heat baths and allows a systematic formulation of statistical quantities such as temperature or entropy of the system. The important point here is that statistical physics only treat systems in equilibrium, which for us corresponds

to a system at a stationary state (the measure of the process is the stationary distribution). Here, we will not go very deep into this theory but only give a few definitions to be able to derive the Gibbs distribution from a certain constrained variational principle. We refer to Chandler (1987), Landau and Lifshitz (1980) for more details on the theory of statistical physics.

In statistical physics, the most important quantity is the partition function Z, defined as a normalization constant for the Gibbs distribution. For us, it was defined in (2.3.13) by integrating only on the coadjoint orbit, and not over the entire dual of the Lie algebra. The partition function can also be viewed as a function of the inverse temperature

$$\beta := \frac{2\theta}{\sigma^2}\,, \tag{2.3.16}$$

or any other parameters of the system. In the next paragraph we will define useful statistical quantities and show how they can be computed directly from the partition function but first, we need to comment on the inverse temperature. We mentioned the connection with heat baths before, and the inverse temperature $\beta = \frac{1}{k_B T}$ makes it clear as T is the temperature of the heat bath and k_B the Boltzmann constant. The definition of the inverse temperature (2.3.16) can be re-interpreted as the Einstein relation which gives a relation between the dissipation θ and the fluctuation coefficient σ such that the system is at equilibrium at temperature T. This is also present in the classical, and simplest version of the fluctuation-dissipation theorem, which can be extended to quantum systems and turns out to be very important in many applications. We only refer to the classic short review of Kubo (1966) and references therein for a more precise treatment.

We begin by the definition of the average energy of the system

$$\langle E\rangle := \int_{\mathcal{O}} h(\mu)\mathbb{P}_\infty(\mu)d\mu = Z^{-1}\int_{\mathcal{O}} h(\mu)e^{-\beta h(\mu)}d\mu\,, \tag{2.3.17}$$

and can directly be obtained from the partition function as

$$\langle E\rangle = -\frac{\partial}{\partial\beta}\log(Z)\,. \tag{2.3.18}$$

In the case of compact coadjoint orbits, such as the rigid body, the average energy is bounded from above. Indeed, as $T \to \infty$, the Gibbs measure $\mathbb{P}_\infty \to 1$ and we have the inequality

$$\langle E \rangle = \int_{\mathcal{O}} h(\mu) \mathbb{P} d\mu \leq \int_{\mathcal{O}} h(\mu) d\mu \leq \infty \,. \tag{2.3.19}$$

Similarly, the energy fluctuation around the mean energy $\langle E \rangle$ can be expressed as a function of the partition function as

$$\Delta_2 \langle E \rangle = \langle E^2 \rangle - \langle E \rangle^2 = \frac{\partial^2}{\partial \beta^2} \log Z = -\frac{\partial \langle E \rangle}{\partial \beta} \,. \tag{2.3.20}$$

As for the mean energy, it saturates as $T \to \infty$ if the coadjoint orbit is compact.

We finally arrive at the definition of the entropy on the coadjoint orbits, given by the famous relation

$$S = -k_B \int_{\mathcal{O}} \mathbb{P} \log(\mathbb{P}) \,, \tag{2.3.21}$$

or, in term of the partition function

$$S = k_B \frac{\partial}{\partial T} (T \log Z) \tag{2.3.22}$$

If the orbit is compact, the entropy saturate as well for large temperatures T.

The entropy is another important quantity as it is maximized by the Gibbs distribution, under the constraint that the average energy is fixed. This reflects the second law of thermodynamics, which states that a system increases its entropy until it reaches a statistical equilibrium. The condition of fixed mean energy is in a sense the definition of the canonical ensemble of statistical mechanics, that is the temperature is fixed by the exchange of energy with an external heat bath.

We enforce the constraint $\langle E \rangle = E_0$, for a given mean energy E_0, with the

Lagrange multiplier β in the entropy functional

$$\mathbb{S}(\mathbb{P}) = \int \Big[\mathbb{P}(\mu)\log \mathbb{P}(\mu) + \beta\,(h(\mu)\mathbb{P}(\mu) - E_0)\Big] d\mu\,. \tag{2.3.23}$$

The variation with respect to $\mathbb{P}$ gives

$$\log \mathbb{P}_\infty(\mu) + 1 + \beta h(\mu) = 0\,, \tag{2.3.24}$$

then, by normalizing, we obtain the Gibbs distribution with inverse temperature β

$$\mathbb{P}_\infty(\mu) = Z^{-1} e^{-\beta h(\mu)}\,. \tag{2.3.25}$$

This short section on the connection between the stochastic dissipative processes on coadjoint orbits and classical statistical mechanics is to be extended in future works. In particular, the treatment of interacting gas, or lattices of coadjoint orbits could be of interest, as some phase transitions would be expected to occur.

2.3.4 Random attractors

After the study of the Fokker-Planck equation and its stationary distributions for the dissipative stochastic processes that we considered so far, we are going to explore a rather different topic: random attractors. We will first explain what random attractors are and then show their existence in the present stochastic systems.

What are random attractors?

In deterministic dissipative dynamical systems, the study of attractors is a vast area of research, with numerous works related to the existence and various properties of these sets. They are important objects as they will contain most of the dynamics of the system for large times due to their attractive property. An attractor is the subset $\mathcal{A}$ of the phase space such that

for a dynamical system $\dot{x} = f(x)$, this set is invariant under the dynamics. More precisely, if φ is the flow of this system, we have $\varphi(\mathcal{A}) \subset \mathcal{A}$ and for any $x_0 \in B(\mathcal{A})$, the basin of attraction of $\mathcal{A}$, for any open neighborood $N(\mathcal{A})$ of the attractor, there exists T such that $x(t) \in N(\mathcal{A})$ for all $t > T$. The definition can also be written as a limit as $t \to \infty$ instead of a time T, that is for $t \to \infty$, $y(t)$ will be in the set $\mathcal{A}$. In most cases, attractors consist of points, limit cycles or the so-called strange attractors. The first type is of 0 dimension and the second type can have 1 or more dimensions. Strange attractors can have a fractal dimension and are the most complex and interesting attractors to study. Well-known examples such as the Smale horseshoe or the Lorenz attractors triggered much research in recent years. We refer for more detail on this topic to the review Eckmann and Ruelle (1985).

We are not going to study deterministic attractors as they are only points in our previous examples with a double bracket dissipation. Indeed, the dissipation will bring the system to the state of lowest energy, which consists of two points in the rigid body example. To get interesting attractors, forcing must be including as the conservative Hamiltonian dynamics cannot balance the dissipation by itself. The noise will play the role of the external forcing, but when introduced, the notion of attractor must be modified, as the limit for large times will not be time independent anymore. To find and study a fixed attractor we need the notion of pullback attractors. In words, a pullback attractor is the invariant set observed at a fixed time by pulling back to $-\infty$ a set of initial conditions. If these initial conditions are in the attractor, they will remain in it almost surely.

We will now give the mathematical definition of a pullback attractor, but we first need to introduce some elements of random dynamical systems. The solution of a random dynamical system is given by a skew-product flow on the space of solutions X and probability space Ω. In this view, the noise realisation at time t is denoted by $\theta_t\omega$, for $\omega \in \Omega$, and the solution of the dynamical system is given by $x(t) = \varphi(t, \omega)x(0)$. One can think of θ_t of being the flow of the noise realisation and φ the flow of the entire system, also denoted cocycle in the language of random dynamical systems. The

definition of pullback attractor is then the following.

Definition 2.3.1 (Pullback attractor). *A pullback attractor $\mathcal{A}_\omega$ is a set-valued map $\mathcal{A}_\omega : \Omega \to X$ invariant under the flow, i.e. $\varphi(t,\omega)\mathcal{A}_\omega = \mathcal{A}_{\theta_t\omega}$, and for every compact set $B \subset X$, we have almost surely*

$$\lim_{t\to\infty} \sup_{x\in B} d(\varphi(t,\theta_{-t}\omega)x, \mathcal{A}_\omega) = 0\,,$$

for an appropriate distance function $d(\cdot,\cdot)$.

From this definition, iit is clear that f the noise is absent, the attractor will correspond to the usual definition of a deterministic attractor as only the noise is pulled back to $-\infty$. In practice, to numerically compute this attractor one just pull back the initial conditions to $-\infty$ for a fixed realisation of the noise and save the snapshots at for example $t = 0$ which are computed by evolving the same initial conditions from earlier times.

Before going any further, we need to comment on the fact that the attractor depends on the noise realisation ω. This fact is crucial and makes this study rather different from the previous stationary distributions analysis that we did. In practice, this means that instead of averaging over the noise realisation when we obtained the stationary distribution, we now average over the initial conditions with the same noise realisation for each initial condition. From this point of view, it is clear that if we also average over the noise realisation, we would recover the stationary distribution, that is, we have the relation

$$\mathbb{E}(\mathcal{A}_\omega) = \mathbb{P}_\infty\,, \tag{2.3.26}$$

valid for any form of the stationary distribution $\mathbb{P}_\infty$. This was first noticed by in Crauel (1991) and later formalized in a theorem of Crauel and Flandoli (1998).

As for deterministic attractors, there exist several types of random attractors. Point attractors are the most common and can even be created from a random perturbation of a deterministic limit cycle. A recent work on this topic of synchronisation by noise is for example Flandoli et al. (2016). Here, we

begin with a deterministic point attractor and study its properties under stochastic perturbations, so we can only expect here the creation of a higher dimensional attractor and not synchronisation. We will in fact numerically show that we can obtain the so-called random attractors for particular values of the parameters of the system. A random attractor is a pullback attractor which admits chaotic motions and a non-singular measure, called the Sinai-Ruelle-Bowen measure (SRB). The proof of the existence of such a measure will involve showing the existence of chaos using Lyapunov exponents. We will treat this in more detail later. We will proceed here in the framework of geometric mechanics and using basic notions of random dynamical systems. We refer the interested reader to the classic monograph Arnold (1995) or Kloeden and Rasmussen (2011) for a general account on RDS , and to Crauel and Flandoli (1994) for random attractors.

Existence of attractors

The first step towards proving the existence of random attractors with a non-singular SRB measure is to show that the system contains random attractive sets. We will follow Crauel and Flandoli (1994) and Schenk-Hoppé (1998) for the exposition of attractors in the context of random dynamical systems. We begin with the definition of an attracting set for a generic phase space denoted by X.

Definition 2.3.2 (Limit set). *The limit set $\Omega_D(\omega)$ of a random set $D \in \Omega \times X$ is*

$$\Omega_D(\omega) := \bigcap_{T\geq 0} \overline{\bigcup_{t\geq T} \varphi(t, \theta_{-t}\omega) D(\theta_{-t}\omega)}\,.$$

This set corresponds to a subspace of the phase space X where all the points that started in the larger subspace D will eventually end up. It can be shown that the random attractor is $\mathcal{A}(\omega) = \bigcup_{D\in\mathcal{D}(\mathcal{A})} \Omega_D(\omega)$, where $\mathcal{D}(\mathcal{A})$ is the basin of attraction of the random attractor $\mathcal{A}$. We thus need to construct explicitly these limit sets which will, in turn, give us some information on the random attractor. We will achieve this by using Lyapunov functions that

have the property to map a compact set of the phase space to a compact set of $\mathbb{R}$. With such functions, we can transfer this problem from the phase space X to $\mathbb{R}$, where it is easier to show the existence of limit sets. In the present case, the Lyapunov function that we will use is the deterministic Hamiltonian of the system. Remarkably, the proof of the existence of these attractive sets only relies on a few assumptions. In particular, it seems to be valid for any Lie groups, even of infinite dimensions.

Theorem 2.3.2. *The stochastic process* (2.3.8) *admits a random attractor for every finite dimensional Lie group G.*

Proof. We begin the proof by recasting the dissipative stochastic process (2.3.8) as a random dynamical equation (RDE) by using vector Wiener processes z_i defined as

$$dz_i = \sigma_i \circ dW_t^i \,,$$

where z_i is understood as a vector process in the direction along arbitrary $\sigma_i \in \mathfrak{g}$. Notice that we will always use the Stratonovitch integrals here. The process $z(t,\omega) = \sum_i z_i(t,\omega) \in \mathfrak{g}$ thus defines a stochastic path in the Lie algebra $\mathfrak{g}$ and a stochastic path in the group G, given as $g(t,\omega) = \exp(z(t,\omega))$ with the help of the exponential map $\exp : \mathfrak{g} \to G$.

We now apply the change of variables $\widetilde{\mu}(t) = g(t)\mu(t) := \mathrm{Ad}^*_{g(t)}\mu(t)$ and we have, from the original stochastic process

$$d\mu + \mathrm{ad}^*_{\frac{\delta h}{\delta \mu}}\mu \, dt + \theta \, \mathrm{ad}^*_{\frac{\delta C}{\delta \mu}} \left[\frac{\delta C}{\delta \mu}, \frac{\delta h}{\delta \mu}\right]^\flat dt + \sum_i \mathrm{ad}^*_{\sigma_i}\mu \circ dW_t^i = 0 \,,$$

written as $d\mu = F(\mu)dt + G_i(\mu) \circ dW_i$ for convenience,

$$\begin{aligned} d\widetilde{\mu}(t) &= \mathrm{Ad}^*_{g(t)} \left(-\sum_i \mathrm{ad}^*_{\sigma_i}\mu \circ dW + d\mu \right) \\ &= \mathrm{Ad}^*_{g(t)} \left(F(\mathrm{Ad}^*_{g(t)^{-1}}\widetilde{\mu}(t)) \right) dt \,. \end{aligned}$$

From here, we can write the RDE corresponding to the process (2.3.8) as

$$\frac{d}{dt}\widetilde{\mu}(t) = \widetilde{F}(\widetilde{\mu}(t), g(t))\,, \tag{2.3.27}$$

where $\widetilde{F}$ is

$$\widetilde{F} := F(\mathrm{Ad}^*_{g(t)^{-1}}\widetilde{\mu}(t)\,.$$

Recall that from the theory of selective decay we have the following energy dissipation rate

$$\frac{d}{dt}h(\mu) = -\theta \left\| \left[\frac{\partial C}{\partial \mu}, \frac{\partial h}{\partial \mu} \right] \right\|^2 ,$$

and, due to the G-invariance of the Hamiltonian h, we have $h(\mathrm{Ad}^*_g\mu) = h(\mu)$. We can then rewrite the energy dissipative for the stochastic process as

$$\frac{d}{dt}h(\mu) = -\theta \left\| \left[\mathrm{Ad}_{g^{-1}}\frac{\partial C(\widetilde{\mu})}{\partial \widetilde{\mu}}, \mathrm{Ad}_{g^{-1}}\frac{\partial h(\widetilde{\mu})}{\partial \widetilde{\mu}} \right] \right\|^2 \leq 0\,. \tag{2.3.28}$$

This inequality assures that the energy decays at a random non-positive rate. The existence of the random attractor then follows from standard arguments that can be found in Schenk-Hoppé (1998). ■

Existence of SRB measures

We already mentioned Sinai-Ruelle-Bowen (SRB) measures which are particular measures supported on the random attractors. In this work, we will use the convention that SRB measures are non-singular and thus describe a chaotic dynamics on the random attractor. In particular, in dissipative systems, there is no notion of dynamical volume for the flow in directions where the dynamics is contracting. The SRB measure is thus a measure of the volume along the unstable directions which are not contracting. We will not go further here and only refer to Young (2002) for a short review on this topic. To show that there exists chaotic dynamics, one needs the existence of positive Lyapunov exponent, which we will show in the next section. Assum-

ing positiveness of the top Lyapunov exponent, we have the following result, based on Ledrappier and Young (1988).

Theorem 2.3.3. *If the largest Lyapunov exponent of* (2.3.8) *is positive,* (2.3.8) *admits a random attractor supported by an SRB measure.*

Proof. The proof uses the Theorem B in Ledrappier and Young (1988) which assumes the existence of a random attractor, proved in theorem 2.3.2, a positive top Lyapunov exponent that we will study next, and ergodicity which we studied in section 2.2.2. Then this theorem ensures that the random measure associated to the pullback attractor is continuous with respect to the Lebesgue measure, the property of it being a SRB measure. The key point is to have a positive top Lyapunov exponent in order to have enough chaos on the attractor such that it is non-singular. ■

Lyapunov exponents

We just saw that the proof of the existence of an SRB measure and thus the existence of chaos involves showing that there exists a positive Lyapunov exponent. A Lyapunov exponent can be understood as a measure of the sensitivity of the system to initial conditions and is defined in the following manner. Given two nearby initial conditions $\mu_1(0)$ and $\mu_2(0)$, i.e. $\mu_2(0) = \mu_1(0) + \delta\mu(0)$ where $\delta\mu(0)$ is small, we can assume that for large time the norm of this difference behaves as

$$\|\delta\mu(t)\| \approx \|\delta\mu(0)\|e^{\lambda t},$$

for a constant λ called the Lyapunov exponent. The complete definition will be given in the next theorem 2.3.4. At this point, we can already mention that if $\lambda < 0$, the two solutions will eventually be indistinguishable whereas if $\lambda > 0$ the two solutions will diverge. The second case is an indication of a chaotic behaviour and will be the one of interest for us as it is a necessary condition for the existence of the SRB measure. For multi-dimensional systems, there will be a positive or a negative Lyapunov exponent associated with each spatial direction.

This notion of Lyapunov exponents is a well-studied and well-used method in the theory of deterministic dynamical systems, but it is also used in the context of stochastic systems, see for example the book Viana (2014) or the review Young (2013). In particular, there is a classical theorem which, provided that the system is ergodic, ensures that the Lyapunov exponents are global quantities. This means that they do not depend on the initial conditions nor the realisation of the noise. This result is the multiplicative ergodic theorem (MET) due to Oseledec (1968). We give here an adapted version of it and refer to the book Arnold (1995) or an earlier paper Ruelle (1979) for more detail and the proof of this theorem.

Theorem 2.3.4 (MET theorem on dual of Lie algebra). *For an ergodic stochastic process μ on the dual of a Lie algebra $\mathfrak{g}^*$ with stationary distribution $\mathbb{P}_\infty$ there exists a flow invariant subset $\Delta \subset \mathfrak{g}^* \times \Omega$ of the n-dimensional phase space and a sequence of ordered Lyapunov exponents λ_i for $i = 1, \ldots, n$ such that the following properties hold for all $(\mu, \omega) \in \Delta$:*

1. *The dual of the Lie algebra can be decomposed into a direct sum*

$$T_\mu \mathfrak{g}^* = E_1(\omega, \mu) \oplus \ldots \oplus E_n(\omega, x),,$$

   where

$$\delta\mu \in E_i \Leftrightarrow \lim_{t\to\infty} \frac{1}{t} \log \|DF(t, \omega, \mu)\delta\mu\| = \lambda_i \,. \tag{2.3.29}$$

   $d(\delta\mu) = DF(t, \omega, \mu)\delta\mu$ is the linearization of the flow equation $d\mu = F(t, \omega, \mu)$ along a particular solution $\mu(t)$.

2. *For a generic $\delta\mu$, the associated Lyapunov exponent is $\lambda(\omega, x, v) = \lambda_+$, the largest one almost surely.*

3. *The sum of the Lyapunov exponents is*

$$\sum_i \lambda_i = \lim_{t\to\infty} \frac{1}{t} \log \det DF(t, \omega, \mu) \,. \tag{2.3.30}$$

The first property says that the tangent space of the phase space can be

decomposed into several independent one-dimensional subspaces with a Lyapunov exponent associated to each subspace. This decomposition will give the contracting and expanding directions of the dynamics at each position in the phase space. Notice that these subspaces are difficult to find in practice as they depend on the noise realisation and the position in the phase space. The second property says that, because of ergodicity and the fact that the Lyapunov exponents are defined in the limit of large times, only the largest exponent can be observed. The last property then gives information on the total infinitesimal volume of the tangent space around each point via the sum of the Lyapunov exponents. It can be shown that for a compact phase space, the sum is always negative, see for example Baxendale (1991), Le Jan (1987).

We will then start our investigation of the Lyapunov exponents for our systems by computing the sum of them in the next proposition.

Proposition 2.3.2. *For a compact semi-simple Lie algebra $\mathfrak{g}$, the sum of the Lyapunov exponents for the stochastic process* (2.2.2) *with isotropic noise is estimated from below by*

$$\sum_i \lambda_i \geq -\frac{1}{2}|\epsilon| n\sigma^2 - \theta|\epsilon| \left(c^2 \mathbb{I}_{\min}^{-1} - \mathbb{E}_\infty h(\mu)\right) , \tag{2.3.31}$$

where $c = \|\mu(0)\|$, ϵ is the Killing form constant (see (1.2.5) *or in the proof), n is the number of $\sigma_i = \sigma e_i$ forming the isotropic noise, that is $n = \dim(\mathfrak{g})$. The quantity $\mathbb{I}_{\min}^{-1}$ is the largest eigenvalue of the Hessian of the Hamiltonian. The expectation $\mathbb{E}_\infty$ is taken with respect to the stationary distribution $\mathbb{P}_\infty$. An estimation from above is also available upon using $\mathbb{I}_{\max}^{-1}$ instead of $\mathbb{I}_{\min}^{-1}$.*

Proof. To simplify the proof, we will denote our stochastic process (2.3.8) which includes dissipation as

$$d\mu = F(\mu)dt + \sum_i G_i(\mu) dW_t^i \, .$$

where $F(\mu)$ contain the advection, dissipation, and Itô correction terms. We now apply the multiplicative ergodic theorem 2.3.4 that gives the following

formula to compute the sum of the Lyapunov exponents

$$\sum_i \lambda_i = \lim_{t\to\infty} \frac{1}{t} \log\det DF(t,\omega,x)\,. \tag{2.3.32}$$

From here, we need to rewrite this formula explicitly in term of F and G. By using the Jacobi formula, (2.3.32) can be written as

$$\lim_{t\to\infty} \frac{1}{t} \log\det DF(t,\omega,x) = \lim_{t\to\infty} \frac{1}{t} \mathrm{Tr} \int_0^t DF(\varphi(t,x,\omega))ds\,. \tag{2.3.33}$$

Notice that we did not include a term with DG, the linearization of the noise amplitude, since this term is in fact the trace of the adjoint action and always vanishes. Then, using the ergodicity property of this process on the coadjoint orbits, we can replace time integral with space averaging to get

$$\sum_i \lambda_i = \int \mathrm{Tr}(DF(\mu))\mathbb{P}_\infty(\mu)d\mu\,, \tag{2.3.34}$$

where $\mathbb{P}_\infty$ is the stationary distribution of the stochastic process for μ.

We will now obtain bounds for this integral by using the explicit form of F. First, notice that the calculation of the trace simplifies when $\mathfrak{g}$ is a compact semi-simple Lie algebra. Indeed, recall that the Killing form is in this case $\mathrm{Tr}(\mathrm{ad}_A\mathrm{ad}_B) = \epsilon A\cdot B$, where $\epsilon<0$ is a number depending on the Lie algebra. We first recall the explicit form of F for semi-simple Lie algebra $\mathfrak{g}$

$$F(\mu) = \mathrm{ad}_{\frac{\partial h}{\partial \mu}}\mu + \theta\,\mathrm{ad}_\mu\mathrm{ad}_\mu\frac{\partial h}{\partial \mu} + \frac{1}{2}\sum_i \mathrm{ad}_{\sigma_i}\mathrm{ad}_{\sigma_i}\mu\,,$$

and compute

$$\begin{aligned}\mathrm{Tr}(DF(\mu)) &= \mathrm{Tr}\left(-\theta\,\mathrm{ad}_\mu\mathrm{ad}_{\frac{\partial h}{\partial\mu}} + \theta\,\mathrm{ad}_\mu\mathrm{ad}_\mu\frac{\partial^2 h}{\partial\mu^2} + \frac{1}{2}\sum_i \sigma^2\mathrm{ad}_{e_i}\mathrm{ad}_{e_i}\right)\\ &= |\epsilon|\theta h(\mu) + \theta A(\mu,\mu) - \frac{1}{2}|\epsilon| n\sigma^2\,,\end{aligned} \tag{2.3.35}$$

where n is the number of σ_i fields and

$$A(\mu,\mu) = \mathrm{Tr}\left(\theta \mathrm{ad}_\mu \mathrm{ad}_\mu \frac{\partial^2 h}{\partial \mu^2}\right) \tag{2.3.36}$$

is a term that depends on the Lie algebra structure constants. There is no explicit form for each Lie algebra, but in general, we can estimate it using

$$|\epsilon|\theta\kappa(\mu,\mu)\mathbb{I}_{\max}^{-1} \geq \theta \mathrm{Tr}\left(\mathrm{ad}_\mu \mathrm{ad}_\mu \frac{\partial^2 h}{\partial \mu^2}\right) \geq |\epsilon|\theta\kappa(\mu,\mu)\mathbb{I}_{\min}^{-1}\,. \tag{2.3.37}$$

Finally, after collecting terms we arrive at the lower and upper bound for the sum of the Lyapunov exponents of equation (2.3.31). ■

We now need to comment on this proposition. First, the assumption of a compact Lie algebra is necessary to make sense of this formula. Indeed, in the non-compact case, the Killing form is not positive definite, so it does therefore not make sense to use c^2 and no bound can be derived for the sum of the exponents. The non-compact case is thus more complicated and will not be treated it in this work. Second, the lower bound on the sum of the Lyapunov exponent is always negative, as seen from

$$\mathbb{E}_\infty h(\mu) = \int h(\mu) e^{-\frac{2\theta}{\sigma^2}h(\mu)} d\mu \leq c^2 \mathbb{I}_{\min}^{-1}\,.$$

This means that we cannot obtain any relevant information on the top Lyapunov exponent and we thus need to proceed further by using the next formula in the MET theorem. The third remark is about how general our formula is. Recall that we needed the semi-simple assumption to simplify the computation and obtain an explicit formula for this bound. The computation can in principle be extended to more general groups, although one may have to explicitly use the Lie algebra structure constants to obtain an explicit bound. Following this discussion, we also used an isotropic noise, but a more general noise could be used in principle. In any case, these generalisations will not change the overall sign of the right-hand side and will only complicate it. We will thus not derive any general formula here and proceed with the estimation of the top Lyapunov exponent.

As before, we will use the following notation for our stochastic process, which gives the linearized flow

$$d\delta\mu = DF(\mu)\delta\mu\, dt + \sum_i DG_i(\mu)\delta\mu\, dW_t^i\,, \tag{2.3.38}$$

where for example $DG_i(\mu) = \mathrm{ad}_{\sigma_i}$. We will then use the Furstenberg-Khasminskii formula which is the best way to estimate the top Lyapunov exponent. We refer to Arnold (1995) for a complete exposition of this formula in various settings. We will use here a simplified version of this formula. We first need to introduce polar coordinates

$$R = \|\delta\mu\| \in \mathbb{R} \quad \text{and} \quad \Psi = \frac{\delta\mu}{\|\delta\mu\|} \in \mathbb{S}^{n-1}\,, \tag{2.3.39}$$

where $\mathbb{S}^{n-1}$ denotes the $(n-1)$-sphere. The interest of having these variables is to be able to average over the directions Ψ and only capture the expansions around each point with the variable R. For doing this, we need the stochastic process for these variables which are readily computed by applying the projections to the stochastic process $\delta\mu$. We obtain

$$\begin{aligned} dR &= \langle \Psi, (DF(\mu)dt + DG_i dW_t^i)\Psi\rangle R := Q(\mu,\Psi)R\,,\\ d\Psi &= (DF(\mu)dt + DG_i dW^i)\Psi - \langle \Psi, (DF(\mu)dt + DG_i dW^i)\Psi\rangle \Psi\\ &:= P(\mu,\Psi)dt + \sum_i P_i(\mu,\Psi)dW_t^i\,. \end{aligned} \tag{2.3.40}$$

We then average over the variable Ψ and the position μ to estimate the top Lyapunov exponent. We obtain the Furstenberg-Khasminskii formula that can be formally proved using ergodicity and the MET theorem. It reads

$$\lambda_+ = \int Q(\mu,\Psi)\mathbb{Q}(\mu,\Psi)d\mu d\Psi\,, \tag{2.3.41}$$

where $\mathbb{Q}(\mu,\Psi)$ is the joint stationary distribution of the processes μ and Ψ. Notice that the process Ψ depends on μ, but not the opposite and also that computing $\mathbb{Q}$ involves solving the corresponding Fokker-Planck equation. It is therefore not possible to compute this stationary distribution explicitly in

the general case and not even for the rigid body. The first thing to do is to try to estimate this integral from below, but one will see that only a negative bound can be found which is not useful for us as we want to show that this integral can be positive for certain values of the parameters.

In fact, we are not able to obtain an interesting formula for a lower bound of this integral, but we can go a little further than before and show some interesting features about this formula. We start by evaluating the quantity Q at the equilibrium positions (μ_i, Ψ_j), where we use the basis corresponding to the eigenvalues of the Hessian of h. For simplicity, we consider a rigid body motion in an arbitrary semi-simple Lie algebra where we ordered the eigenvalues as $\mathbb{I}_1 > \ldots > \mathbb{I}_n$. We will also use the fact that the dynamics of μ takes place on the coadjoint orbit. Hence the linearization is tangent to the orbits and implies that we can only consider the case when $i \neq j$. The following calculation gives an interesting formula for the quantity $Q(\mu_i, \Psi_j)$

$$\begin{aligned}
Q(\mu_i, \Psi_j) &= \langle \Psi_j, DF(\mu_i)\Psi_j \rangle \\
&= \langle \mathrm{ad}_{\mu_i}\Psi_j, \mathbb{I}^{-1}\Psi_j \rangle - \frac{1}{2}\sigma^2 \sum_k \langle \mathrm{ad}_{e_k}\Psi_j, \mathrm{ad}_{e_k}\Psi_j \rangle \\
&+ \theta \langle \mathrm{ad}_{\mu_i}\Psi_j, \mathrm{ad}_{\mathbb{I}^{-1}\mu_i}\Psi_j \rangle - \theta \langle \mathrm{ad}_{\mu_i}\Psi_j, \mathrm{ad}_{\mu_i}\mathbb{I}^{-1}\Psi_j \rangle \\
&= \langle \mu_i, \mathrm{ad}_{\Psi_j}\mathbb{I}^{-1}\Psi_j \rangle - \frac{1}{2}\sigma^2 \sum_k \|\mathrm{ad}_{e_k}\Psi_j\|^2 \\
&+ \theta \langle \mathrm{ad}_{\mu_i}\Psi_j, \mathrm{ad}_{\mu_i}(\mathbb{I}_i^{-1}\mathrm{Id} - \mathbb{I}^{-1})\Psi_j \rangle \\
&= -\frac{1}{2}\sigma^2 \sum_k \|\mathrm{ad}_{e_k}\Psi_j\|^2 + \theta\,(\mathbb{I}_i^{-1} - \mathbb{I}_j^{-1})\|\mathrm{ad}_{\mu_i}\Psi_j\|^2 \\
&= -\frac{n-1}{2}\sigma^2 + \theta\, c^2(\mathbb{I}_i^{-1} - \mathbb{I}_j^{-1})\,.
\end{aligned}$$

We have computed this formula explicitly in the case of rigid body motion to illustrate an important point about the nature of the random attractor. To obtain a positive top Lyapunov exponent, it is necessary to have $Q(\mu, \Psi)$ positive in a large enough region of the space $\mathfrak{g}^* \times \mathbb{S}^{n-1}$ This will be the case only if some extremal positions have this property. Some will be positive if the difference between two moments of inertia is large enough compared to

the ratio between the noise and the damping, that is

$$\mathbb{I}_1^{-1} - \mathbb{I}_2^{-1} > \frac{(n-1)\sigma^2}{2\theta c^2}\,.$$

The difference between the moments of inertia is an indication of the strength of the original deterministic dynamics which mainly consists in a shear effect. This formula thus indicates that the shear induced by the deterministic dynamics must be large enough compared to the temperature of the system (recall that it is given by $T^{-1} = 2\theta/\sigma^2$). This condition is of course not sufficient as a small noise will satisfy such relation but not produce an SRB measure because the region of the space where Q is positive is too small to balance contribution of the region where Q is negative. It is then clear that the existence of the SRB measure or positive top Lyapunov exponent is due to a delicate balance between all the parameters of the system and that a rigorous proof with a precise lower bound is completely out of reach. The importance of the shear effect and the difficulty of describing mathematically has been noticed before and is an active area of research in dynamical systems. We refer for example to Lin and Young (2008), Wang and Young (2003), or more recently to Engel et al. (2016) for another example.

In the next section 2.3.5, we will give more detail on this issue with the help of numerical simulations and then study a simpler case of periodic kicking in the rigid body in section 2.3.6.

2.3.5 Random attractor in the rigid body

Our example is, as before, the stochastic free rigid body with double bracket dissipation. It is simple enough so that we can have precise estimates on the sum of the Lyapunov exponents and we can perform reliable numerical experiments to visualise the attractor and estimate the top Lyapunov exponent.

We can first obtain an exact bound for the Theorem 2.3.2, using the explicit form of the Lie algebra $\mathfrak{so}(3)$.

Proposition 2.3.3. *For the stochastic dissipative rigid body, the sum of the*

Lyapunov exponents is given by

$$\sum_i \lambda_i = -3\sigma^2 - \theta\left(c^2 \mathrm{Tr}\,\mathbb{I}^{-1} - 6\mathbb{E}_\infty h\right)\,, \tag{2.3.42}$$

where, as before, c is the radius of the momentum sphere.

Proof. We explicitly compute the term A (2.3.36) of Proposition 2.3.2 using the structure constants $c_k^{ij} = \epsilon_{ijk}$ of $\mathfrak{so}(3)$ to get

$$\begin{aligned} -A(\Pi,\Pi) &= -\mathrm{Tr}(\mathrm{ad}_\Pi \mathrm{ad}_\Pi \mathbb{I}^{-1}) = -c_n^{im} c_m^{jn} \mathbb{I}^{-1} \Pi_i \Pi_j \\ &= \Pi_1^2(\mathbb{I}_2^{-1} + \mathbb{I}_3^{-1}) + \Pi_2^2(\mathbb{I}_1^{-1} + \mathbb{I}_3^{-1}) + \Pi_3^2(\mathbb{I}_1^{-1} + \mathbb{I}_2^{-1})\,, \end{aligned}$$

which directly gives equation (2.3.42). ∎

We can now proceed by estimating the value of the top Lyapunov exponent and find the region of the parameter space (σ, θ) where it is positive. This corresponds to a chaotic motion on the non-singular random attractor. We show in figure 2.4 the smoothed result of this computation, where we estimated the value of the top Lyapunov exponent on a grid in (σ, θ) with steps of size 0.1. We did not show any error bars in this plot, but we need to

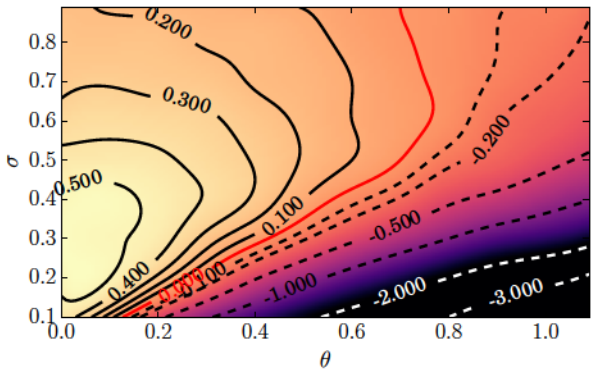


Figure 2.4: This plot shows the value of the top Lyapunov exponent in the parameter space (σ, θ). The black lines are iso-lines, and the red line is the line where $\lambda_+ = 0$. A clear and large region supports a positive Lyapunov exponent, indicating chaotic motion on the non-singular random attractor.

mention that the errors are not constant throughout the plot. In particular,

the regions with a noise amplitude near 0 or 1 contain more variations in the simulation, thus larger errors. This can be explained as a large noise requires smaller time steps for the same precision, and a smaller noise loses ergodicity which requires longer simulation times. The only thing that we cannot guarantee is that the observed chaos is not a transient chaos. This is a chaotic motion which does not persist in time. To be more certain about this, we performed a couple of longer runs which suggest that the positive top Lyapunov exponent will remain positive at all times. Despite these errors, the global trend of the plot is reliable and provides interesting information about the random attractor. The most important result is that a large region of this space has a positive top Lyapunov exponent, but we can read more from this plot. We describe below some important such features.

- Let us first look at the position where $\theta = 0$. By gradually increasing the noise, we first observe an increase and then a decrease of chaotic motion. This increase of chaos with the noise is to be expected, but the decrease is a particular feature of the rigid body. Recall that the driving mechanism of this type of random attractors is the shear of the dynamics and is bounded by a value proportional to the difference between the two extreme moments of inertia. When the noise amplitude is comparable to the size of the momentum sphere, the maximum shear effect is attained, and the chaos cannot be larger. This effect is clearly only present for compact systems. For example, shear induced attractors on planar systems were studied numerically by Lin and Young (2008) and analytically by Engel et al. (2016). In these cases, the top Lyapunov exponent is not bounded for large noise amplitudes.

- Let us look at the other critical position, where $\sigma = 0$ and $\theta = 0$. The Lyapunov exponent cannot be computed because the system is the original Hamiltonian rigid body dynamics, and the MET theorem does not apply. If now $\theta > 0$, the top Lyapunov exponent converges to a single negative value related to the strength of the damping.

- Another interesting feature is that the region of positive Lyapunov exponent has slow variations, whereas the other region shows a steep

decrease (in the bottom right). On top of that, the slope of the line where, for example, $\lambda^+ = -1$, is approximately 0.4. This number is smaller than unity which means that the dissipation must increase faster than the noise in order to observe a random attractor with a similar amount of chaos.

This analysis of figure 2.4 gave us interesting insights into the mechanisms of creation of the random attractor as well as a numerical proof of the existence of SRB measures on the momentum sphere. We will end this study by a numerical computation of the SRB measure displayed in Figure 2.5.

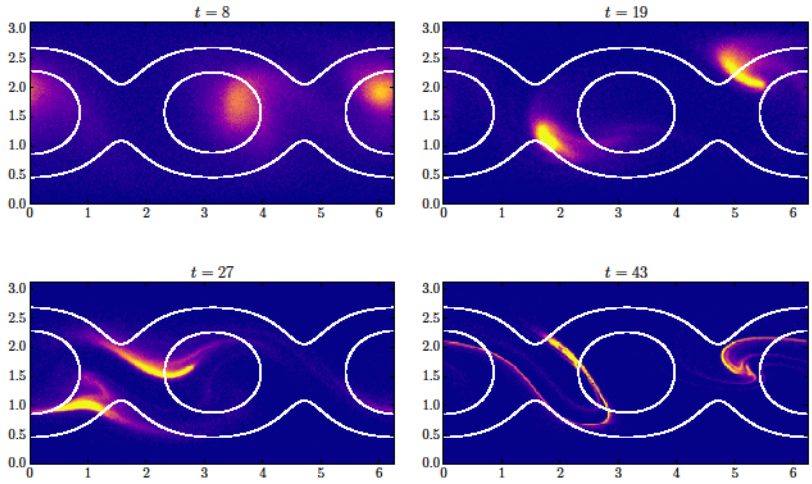


Figure 2.5: These four panels show a realisation of the SRB measure for the free rigid body in log scale. The bright colour indicates a larger density of the SRB measure and the blue regions are empty.

We extracted four snapshots of the same attractor at different times. The bright colour indicates a higher density of the SRB measure, which is plotted in log scale so that we can see the fine structure of the SRB measure. From these snapshots, the mechanism of creation of the random attractor can be viewed as a stretching and folding process triggered by the natural shear of the rigid body. In most of the standard examples such as the Smale horseshoe of the Duffing attractor, this mechanism produces a strange attractor with a fractal structure. We can observe the creation of such structure in the light regions of these figures but it remains a difficult open problem to better

characterise these structures in term of the fractal dimension, etc. In the next section, we will study a simplification of this system where the noise is replaced by a periodic kicking and which can still produce chaotic motion.

2.3.6 Effect of periodic kicking on the rigid body

Up to this point, we only considered stochastic perturbations of the Euler-Poincaré equation, but nothing prevents us from using a deterministic forcing instead of the noise. In this section, we will replace the noise by a periodic kicking and show that chaotic behaviours can exist. This study is inspired by various works in dynamical systems which use this deterministic periodic forcing to detect chaos and other strange attractors. We refer to Lin and Young (2008), Wang and Young (2003) and references therein for such studies.

We will here only consider the free rigid body and will not go into technical details but rather try to qualitatively understand how the periodic kicking modifies the free rigid body dynamics. We use the same perturbation of the rigid body, but we replace the Wiener process W_i by

$$dW_t^i \Rightarrow \sigma_i \sum_{n=1}^{\infty} \delta(t - nT)\,, \tag{2.3.43}$$

where $T \in \mathbb{R}$ is the period for the kicking and $\sigma \in \mathbb{R}^3$ still gives the amplitude and direction of the kick. We will use here the generic case $\sigma_i = \sigma e_i$ so the dynamical equation is

$$\frac{d}{dt}\mathbf{\Pi} = \mathbf{\Pi} \times \mathbf{\Omega} + \theta \mathbf{\Pi} \times (\mathbf{\Pi} \times \mathbf{\Omega}) + \sigma \left(\sum_i \mathbf{\Pi} \times e_i \right) \left(\sum_{n=1}^{\infty} \delta(t - nT) \right) . \tag{2.3.44}$$

There is an important remark to make at this point. In the case of stochastic forcing, this form of the σ_i means that the noise is isotropic but in this case it only means that the kicking axis is fixed in the $(1, 1, 1)$ direction and with amplitude σ.

We now want to study the existence of non-singular attractors supported by possible chaotic motion. For this, we need to observe the system at intervals T to obtain a time independent attractor. Indeed, between each kick the system will evolve and result in a T periodic motion when the attractor is reached. We will only do numerical simulations of this system and analyse its behaviours by varying the kicking amplitude. We use $T = 1$, $\theta = 0.2$, $\sigma = \|\sigma\|(1, 1, 1)$ and $\mathbb{I} = \text{diag}(1, 2, 3)$. Notice that σ is not aligned with an eigenvalue of the moment of inertia $\mathbb{I}$ such that the system will exhibit many possible behaviours by just varying the kicking amplitude σ.

In figure 2.6 we display several attractors of the kicked rigid body for 6 values of σ. All the attractors are different and we will try to understand their properties in detail. First recall that in the deterministic case, there are two singular attractors corresponding to the two lowest energy equilibrium position of the rigid body. With the periodic kicking, these two attractors will most likely be different for the following reason. The dynamics of the rigid body is anti-symmetric with respect to a plane perpendicular to the axis of rotation of the lowest energy equilibrium and the kicking is symmetric. The interaction of the kicking and the rigid body dynamics will thus be different in the opposite side of the momentum sphere. In particular, the kicking and the dynamics are in the same direction in the lower right region of the plots in figures 2.6, and in the opposite direction for the upper left parts. In the sequel, we will denote by $\mathcal{A}^+$ the attractor where the dynamical flow and the kicking are in the same direction and by $\mathcal{A}^-$ the other attractor. They will be understood as being the same if only one attractor is observed. We will analyse the attractors and classify the different regimes according to the values of $\|\sigma\|$. To help us understanding these different behaviours, we estimated the top Lyapunov exponents of the system for each value of $\|\sigma\|$, which we displayed in figure 2.7. From the estimations of the Lyapunov exponents and the simulation 2.6, we can describe each attractor in more detail.

1. $\|\sigma\| < 0.02$: This is the simplest case and the closest to the original dynamics. The two attractors $\mathcal{A}^+$ and $\mathcal{A}^-$ are singular and are near their original positions.

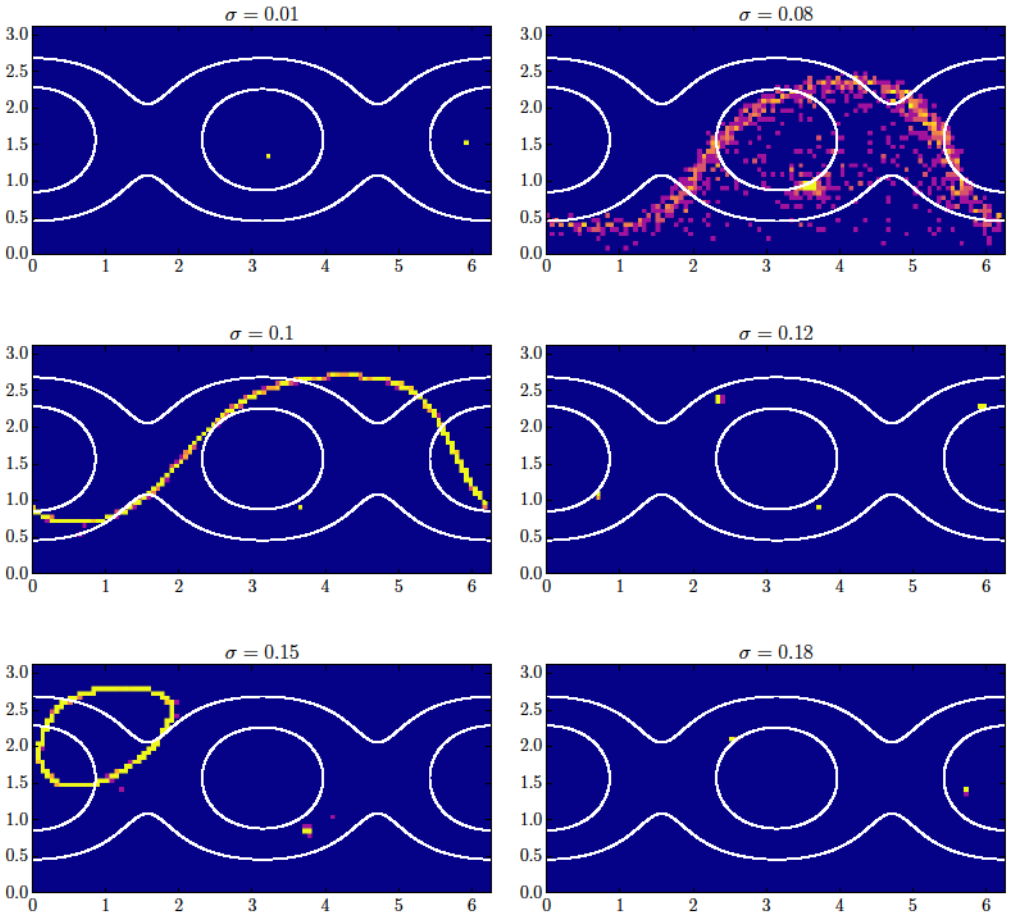


Figure 2.6: 6 attractors of the periodically kicked free rigid body are displayed here. They are all different but with three main types. Either fixed point, limit cycle or alternating fixed points. Chaos is observed in some limit cycles as well, see figure 2.7. A transient chaos is observed when $\sigma = 0.08$, where the simulation was stopped before the chaotic motion relaxed to a unique fixed point.

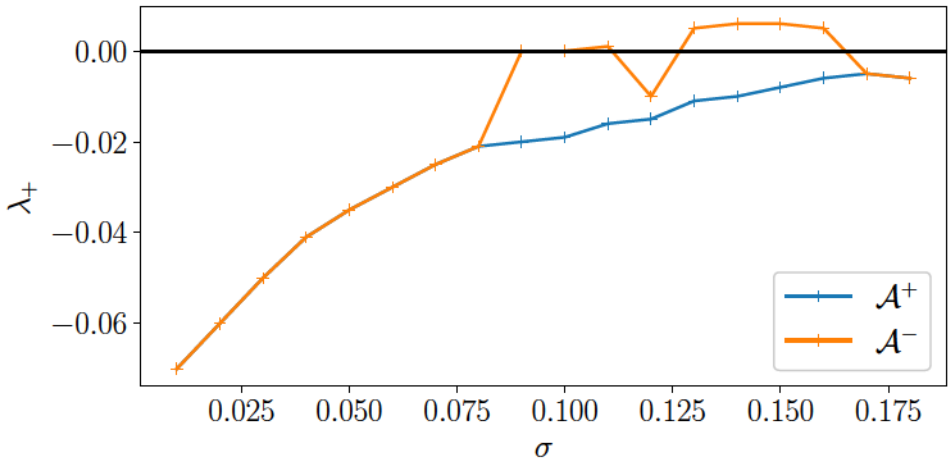


Figure 2.7: This plot displays the value of the top Lyapunov exponent of the various attractors of figure 2.6. If two attractors are observed, we computed the top Lyapunov exponent for both.

2. $0.02 < \|\sigma\| < 0.08$: This snapshot was taken before the complete convergence to the attractor, which in this case consists in the yellow dot in the centre. This is to illustrate the phenomenon of transient chaos. The simulations which started near the attractor $\mathcal{A}^-$ of the previous panel will eventually reach $\mathcal{A}^+$, but before being captured in a ring, they show a chaotic motion for some time.

3. $0.08 < \|\sigma\| < 0.11$: Slightly increasing the kicking amplitude stabilises the ring, or equivalently the attractor $\mathcal{A}^-$. This is a limit cycle and it passes near the both equilibria of minimum energy, experiencing kicks from both sides of the sphere with opposite interaction with the rigid body dynamics. The chaos on this limit cycle is not clear from the Lyapunov exponent computation as the top Lyapunov exponent is close to 0, see figure 2.7. The other attractor $\mathcal{A}^+$ is singular and remains near its previous position.

4. $0.11 < \|\sigma\| < 0.13$: By looking at intermediate values of σ, we observed that increasing σ moves the limit cycle in the opposite direction from the other attractors. At some point, it crosses the axis of the body with

maximal energy (top and bottom line of the plot) and breaks into four parts. These 4 points are only two different attractors. The attractor $\mathcal{A}^+$ remains close to its previous position, and the 3 other singular points are at equal distances from the kicking axis $(-1,-1,-1)$ and form $\mathcal{A}^-$. After each kick, the singular attractors cyclically jump from one to the other in $\mathcal{A}^{-1}$.

5. $0.13 < \|\sigma\| < 0.17$: After crossing the smallest axis of rotation, the limit cycle is again present but remains close to the region of the momentum sphere where the kicking is of opposite direction to the motion of the body. As compared to the previous limit cycle, this motion is more chaotic and we can even be certain that the top Lyapunov is positive from the errors of the simulations.

6. $0.17 < \|\sigma\| < 0.18$: Finally, after more increase of the kicking amplitude, one kick will almost correspond to half a rotation of the rigid body. It thus creates a single attractor consisting of two singular points periodically switching between each other.

To summarise this brief investigation of the periodically kicked rigid body, we want to stress that the interesting feature here is not only the existence of chaos but also the existence of limit cycle and cyclic singular attractors. These attractors only appear for a subtle combination of the geometry and the kicking amplitude and direction. We will not investigate this system further here and leave it open for future numerical or analytical investigations.

2.3.7 Summary and outlook

This section builds upon the stochastic deformation of the Euler-Poincaré equation developed in the previous section 2.2. The new element is the geometric dissipation which we included following the general approach of Gay-Balmaz and Holm (2013). The interaction of noise and dissipation provided us with a richer system. One of the main results is related to the fact that the stationary distribution of the Fokker-Planck equation is not anymore constant on the coadjoint orbit but more concentrated in the regions

of lowest energy. In fact, provided that the noise is isotropic, we were able to compute the exact form of the stationary distribution as a Gibbs distribution, often found in statistical physics for open systems interacting with an external heat bath. This result suggests that this system has a strong connection with statistical physics which could give an interesting physical motivation for the study of such systems. We briefly showed this connection by extending a few classical definitions of statistical quantities to coadjoint orbits, and showed that the Gibbs measure can be obtained as a distribution of maximal entropy, for a fixed average energy. As in the first section, we did not study the semidirect product case which has similar stationary distribution but only for the marginal distribution as the coadjoint orbit is not compact. We refer to Arnaudon et al. (2017a) for more detail about this and the heavy top example.

The core of this section was devoted to the study of the so-called random attractors. These attractors are observed only when the realisation of the noise is fixed and can support chaotic motions. This chaos comes from a subtle interaction of noise, dissipation and shear provided by the underlying deterministic dynamics. If the motion is chaotic on the attractor, it supports an important measure in the theory of random dynamical processes called the SRB measure. The difficulty here is only to prove that there is a chaotic motion. We could not achieve this analytically, and this problem is in fact hard even in simpler cases. We thus performed numerical estimations of the top Lyapunov exponent, and by analysing this data we obtained a better insight into the interaction of noise and dissipation in this system. We even considered a simpler deformation with a periodic kicking which exhibited similar behaviours of chaotic motion, but on limit cycles. One could push further these investigations by either considering more complex systems, such as the heavy top, or to obtain a better analytical understanding of them. At this point of the work, we leave these as open problems within the realm of random dynamical systems.

2.4 Stochastic nonlinear stability

In this section, we will study another important aspect of mechanical systems with symmetries, that is the stability of relative equilibria. First, recall that the Euler-Poincaré equation only describes parts of the complete dynamics of the system which is a path in the Lie group itself. This means that a fixed point of the Euler-Poincaré equation will not be a static equilibrium of the system but only an equilibrium relative to the moving frame of the reduction by symmetry. To assess the stability of such solutions, the most practical method is the so-called energy-Casimir method. This method is rather powerful as it proves the nonlinear stability of some relative equilibria. It is based on the classical Lyapunov method and takes into account the geometrical structure of the equation to construct a proper Lyapunov functional. This theory started in the early days of geometric mechanics by Arnold (1965), Arnol'd (1966b) and was later extended by Holm et al. (1985). This latter paper will be our main reference for the energy-Casimir method. This method was further extended to encompass more general systems which do not have Casimir functions by Simo et al. (1991). It is called the energy-momentum method and in this theory the Lyapunov functional is constructed using the momentum map instead of directly using the Casimir functions.

In a nutshell, the energy-Casimir method relies on the construction of an extended Hamiltonian $H_C = H + \lambda C$, where C is a Casimir of the system and λ a parameter. This functional serves as a Lyapunov functional for assessing the stability of a given relative equilibrium via the sign-definiteness of its second variation at the equilibrium position. The Casimir is a crucial part of this construction as it takes into account the moving frame due to the reduction by symmetry, thus the geometry of this system. If stochastic perturbations are included such that the geometrical structure of these equations is preserved, the energy-Casimir method can be extended to this context and provide stability result in a probabilistic sense. This is the aim of the section which is based on Arnaudon et al. (2017b).

2.4.1 Review on nonlinear stability

We begin with a short review of stability for mechanical systems. There exist several non-equivalent notions of stability suited for different purposes, but we will only use the notion of nonlinear stability. We refer to Abraham et al. (1978) for a classical exposition of stability in mechanics.

Deterministic stability

We will use the following definition of deterministic nonlinear stability of an equilibrium solution of a generic dynamical system of the form

$$\dot{x} = f(x)\,, \tag{2.4.1}$$

where $x \in \mathbb{R}^n$, and $f : \mathbb{R}^n \to \mathbb{R}^n$. We assume that there exists a solution x_e which is an equilibrium position, i.e. $f(x_e) = 0$.

Definition 2.4.1 (Nonlinear stability). *An equilibrium position x_e of a dynamical system is nonlinearly stable if for every neighbourhood U of x_e there exists a neighbourhood V such that trajectories starting in V will never leave U. In mathematical terms, given an appropriate norm $\|\cdot\|$, $\forall \epsilon > 0$, $\exists \delta > 0$ such that if $\|x(0) - x_e\| < \delta$, then $\|x(t) - x_e\| < \epsilon$ for $t > 0$.*

This notion of stability states that the system will remain in a neighbourhood of the equilibrium position for all time if the initial condition is taken to be near this position. In finite dimensional systems, there is an equivalent notion, the formal stability which requires the existence of a conserved integral of the motion, provided by a Casimir function.

Definition 2.4.2 (Formal stability). *An equilibrium point x_0 is formally stable if a conserved quantity exists with a vanishing first variation and a definite second variation evaluated at the equilibrium position.*

This notion is important for the energy-Casimir method as the construction of the correct conserved quantity will give formal stability for the finite di-

mensional systems, which will, in turn, provide nonlinear stability. The next proposition can be proved in finite dimensions only.

Proposition 2.4.1. *In finite dimensions, formal stability is equivalent to nonlinear stability.*

The energy-Casimir method

Before considering stochastic systems, we will review the energy-Casimir method in finite dimensions. This method consists of four steps and assumes that we have a Euler-Poincaré system with Hamiltonian h, Casimir C, momentum variable μ and an equilibrium solution μ_e. The most important object is the extended Hamiltonian

$$H_C(\mu) = h(\mu) + C(\mu)\,. \tag{2.4.2}$$

Notice that there is some freedom in the choice of the Casimir function, as any function of a Casimir is still a Casimir. We will need the first and second variations of this Hamiltonian. The first variation is $DH_c : \mathfrak{g} \to \mathfrak{g}$, that is $DH_c(\mu) \in \mathfrak{g}$ is the variational derivative of H_C. The second variation is the variational derivative of DH_C with respect to its argument. It is another map $D^2H_C : \mathfrak{g} \to \mathfrak{g} \times \mathfrak{g}$ such that $D^2H_C(\mu) : \mathfrak{g}^* \times \mathfrak{g}^* \to \mathbb{R}$, or $D^2H_C(\mu) : \mathfrak{g}^* \to \mathfrak{g}$. The four steps of the energy-Casimir method are then:

1. Restrict the form of the Casimir function C such that

$$\langle DH_C(\mu_e), \delta\mu\rangle = 0, \quad \forall \delta\mu \in \mathfrak{g}^*\,, \tag{2.4.3}$$

   where $DH_C(\mu) : \mathfrak{g}^* \to \mathbb{R}$ stands for the first variation of the extended Hamiltonian. This first step ensures that μ_e is a critical position of the functional $H_C(\mu)$.

2. Compute the second variation of the extended Hamiltonian and try to restrict further the form of the Casimir such that it is sign-definite, i.e.,

$$\pm\langle \delta\mu, D^2HC(\mu_e)\delta\mu\rangle \geq 0, \quad \forall \delta\mu\,, \tag{2.4.4}$$

where the pairing is the standard pairing on the Lie algebra $\mathfrak{g}$.

3. If the second variation is sign-definite, then define the norm $\|\delta\mu\|^2_{H_C}$ and use it to obtain the formal stability of μ_e. In general, the previous step will give certain conditions on the parameter of the system to ensure that the second variation is sign-definite.

Although we discussed the Euler-Poincaré setting, semi-direct products systems will follow the same procedure.

Stochastic stability

We now turn to the definition of stochastic stability corresponding to the previous notion of nonlinear stability. We consider an arbitrary stochastic extension of the dynamical system (2.4.1) of the form

$$dx_t = f(x_t)\,dt + \sum_i \sigma_i(x_t) \circ dW_t^i\,, \tag{2.4.5}$$

for a given set of vector fields $\sigma_i : \mathbb{R}^n \to \mathbb{R}^n$. As before, we assume that x_e is a equilibrium solution, that is $f(x_e) = 0$ and $\sigma_i(x_e) = 0$ for all i. The latter condition will be a restriction on the noise vector fields that will be compatible with the equilibrium solution of the deterministic system. We will discuss this later in the examples and only assume here that such a stationary solution x_e exists. We can then define our notion of stochastic stability as an extension of the usual notion of stochastic stability, found for example in the monograph Khasminskii (2011).

Definition 2.4.3 (Transient stochastic stability)**.** *For a suitable norm $\|\cdot\|$, a fixed point x_e is weakly stable in probability with stopping time T if for every $\epsilon, \sigma > 0$ there exists $\delta > 0$ and $T(\epsilon, \delta, C) > 0$ such that if $\|x(0) - x_e\| < \delta$ and $T > t > 0$*

$$P\left(\|x(t) - x_e\| > \epsilon\right) < \sigma\,. \tag{2.4.6}$$

The difference between this concept of transient stochastic stability and the

standard one is the use of the stopping time T. If $T \to \infty$, there is no stopping time and we recover the usual definition of stochastic stability. For us, this stopping time is important as it will give extra information about how the noise affects the stability of the deterministic equilibrium solution, and in particular for how long the stochastic stability persists. We will see that this stopping time will directly be related to the amplitude of the noise.

2.4.2 The stochastic Energy-Casimir method

In the stochastic context, we will see that the deterministic energy-Casimir method applies up until the last step which we did not describe in detail in the previous section. It consists in using the norm $\|\delta\mu\|_{H_C}$ in the definition of the nonlinear stability. We will see that this norm is no more preserved by the stochastic dynamics but that we can use the previous notion of stochastic stability. We will proceed with the usual energy-Casimir method following the appendix A of Holm et al. (1985).

We first need to find conditions on the Casimir such that the first variation of the extended Hamiltonian (2.4.2) vanishes at μ_e, that is $\langle DH_C(\mu_e), \delta\mu\rangle = \langle Dh(\mu_e), \delta\mu\rangle + \langle DC(\mu_e), \delta\mu\rangle = 0$, for all $\delta\mu \in \mathfrak{g}^*$. The next step is to check the sign-definiteness of the second variation of H_C when evaluated at μ_e. Both conditions are purely geometric and do not depend on the dynamics of the system, thus on the noise. We can use the deterministic results and assume that μ_e is deterministically stable. Recall that the norm is

$$\|\xi\|^2_{H_C} := \langle \xi, D^2 H_C(\mu_e)\xi\rangle\,, \tag{2.4.7}$$

which is a well-defined norm on the space of variations on paths on $\mathfrak{g}^*$. In the deterministic case, this norm is preserved by the linearized flow of the system, and this leads to the nonlinear stability of μ_e. In the presence of stochastic perturbation, we must investigate how this norm is not preserved by the linearized stochastic flow. For this we first compute the linearized

flow using $\mu_t = \mu_e + \delta\mu_t$ and obtain

$$d\delta\mu_t = \mathrm{ad}^*_{Dh(\mu_e)}\delta\mu_t\,dt + \mathrm{ad}^*_{D^2H_C\delta\mu_t}\mu_e\,dt + \sum_i \mathrm{ad}^*_{\sigma_i}\delta\mu_t \circ dW_t^i\,. \tag{2.4.8}$$

We then use the extended Hamiltonian to compute the evolution of the associated norm, and the previous equation simplifies to

$$d\delta\mu_t = \mathrm{ad}^*_{D^2H_C\delta\mu_t}\mu_e\,dt + \sum_i \mathrm{ad}^*_{\sigma_i}\delta\mu_t \circ dW_t^i\,. \tag{2.4.9}$$

From this computation it is clear that the norm (2.4.7) will in general not be preserved by this flow due to the stochastic term. In order to go further, we compute the time evolution of the H_C norm of the stochastic path $\delta\mu_t$

$$\begin{aligned} d\|\delta\mu_t\|^2_{H_C} &= 2\langle D^2H_C\delta\mu_t, d\delta\mu_t\rangle \\ &= \sum_i 2\langle D^2H_C\delta\mu_t, \mathrm{ad}^*_{\sigma_i}\delta\mu_t\rangle \circ dW_t^i\,, \end{aligned}$$

where we used the symmetry of the operator D^2H_C. This term gives us the exact dynamics of the H_C norm (2.4.7), but we want the average dynamics to obtain the transient stability in probability. For this, we need to use the infinitesimal generator of a stochastic process μ. It is defined as

$$\mathcal{D}_\mu f(\mu(t)) := \lim_{\Delta t\to 0} \mathbb{E}\left(\frac{f(\mu(t+\Delta t)) - f(\mu(t))}{\Delta t}\middle|\,\mathcal{F}_t\right)\,, \tag{2.4.10}$$

for any function $f : \mathfrak{g}^* : \mathbb{R}$. This definition can for example be found in Arnaudon et al. (2014) when defined directly on Lie groups. When this derivative is viewed as the infinitesimal generator of the process μ, it generates the Fokker-Planck equation for the process μ

$$\partial_t\mathbb{P}(\mu,t) = \mathcal{D}^*_\mu\mathbb{P}(\mu,t)\,, \tag{2.4.11}$$

where $\mathcal{D}^*_\mu$ is the adjoint of $\mathcal{D}_\mu$ in the L^2 pairing. In this case, this is the Kolmogorov operator which we used so far to compute the Fokker-Planck equation, but which can be applied to any other function of the stochastic process. We also refer to Arnaudon et al. (2014), Zambrini (2009, 2015) for

more for the use of this operator in the context of geometric mechanics and stochastic Euler-Lagrange equation.

We can now apply this derivative for the process $\delta\mu$ to the H_C norm of $\delta\mu_t$ to get

$$\mathcal{D}_{\delta\mu}\|\delta\mu_t\|^2_{H_C} = \sum_i \langle D^2 H_C \mathrm{ad}^*_{\sigma_i}\delta\mu_t + \mathrm{ad}_{\sigma_i} D^2 H_C \delta\mu_t, \mathrm{ad}^*_{\sigma_i}\delta\mu_t\rangle \,. \tag{2.4.12}$$

This formula is interesting in itself as it is close to a commutator of two operators. Indeed, recall that $D^2H_C : \mathfrak{g}^* \to \mathfrak{g}$ and that in most cases considered here $\mathrm{ad}^* \cong -\mathrm{ad}$. For bi-invariant pairings, the commutator structure is more apparent as (2.4.12) simplifies to

$$\mathcal{D}_{\delta\mu}\|\delta\mu_t\|^2_{H_C} = \sum_i \langle \mathrm{ad}_{\sigma_i} D^2 H_C \delta\mu_t - D^2 H_C \mathrm{ad}_{\sigma_i}\delta\mu_t, \mathrm{ad}_{\sigma_i}\delta\mu_t\rangle \,. \tag{2.4.13}$$

The average lack of conservation of the H_C norm is thus intimately linked to the interaction between the noise fields σ_i and the dynamics of the system through D^2H_C. As we will see in the examples, this formula can be further simplified.

What we can do to proceed further and obtain a general theory for stochastic stability is consider a system such that we can bound (2.4.12) as

$$D_{\delta\mu}\|\delta\mu_t\|^2_{H_C} \leq \Sigma^2 \|\delta\mu_t\|^2_{H_C} \,, \tag{2.4.14}$$

for a constant Σ. As we will see in the examples, this assumption seems to be valid in most standard cases, and the constant Σ can directly be estimated. At this point, it is expected that this constant will be proportional to the noise amplitude. Indeed, if the noise vanishes we should recover the conservation of this norm. From this estimate, we can obtain an exponential upper bound for the evolution of the energy.

Proposition 2.4.2. *Assuming that* (2.4.14) *holds for some* Σ*, we have the following time-dependent bound for the expected value of the* H_C *norm of* $\delta\mu_t$

$$\mathbb{E}(\|\delta\mu_t\|^2_{H_C}) \leq \|\delta\mu_0\|^2_{H_C} e^{\Sigma^2 t} \,. \tag{2.4.15}$$

Proof. We compute

$$\begin{aligned}
\frac{d}{dt}\mathbb{E}(\|\delta\mu_t\|^2_{H_C}) &= \frac{d}{dt}\int \|\delta\xi\|^2_{H_C}\mathbb{P}(\xi,t)\,d\xi \\
&= \int \|\xi\|^2_{H_C}\frac{d}{dt}\mathbb{P}(\xi,t)\,d\xi \\
&= \int \|\xi\|^2_{H_C}D^*_\xi\mathbb{P}(\xi,t)\,d\xi \\
&= \int D_\xi\|\xi\|^2_{H_C}\mathbb{P}(\xi,t)\,d\xi \\
&\leq \Sigma^2\int \|\xi\|^2_{H_C}\mathbb{P}(\xi,t)\,d\xi \\
&= \Sigma^2\,\mathbb{E}(\|\delta\mu_t\|^2_{H_C})\,,
\end{aligned}$$

where $\mathbb{P}$ is the probability density of the process $\delta\mu$ and D_ξ its generator (2.4.12). Finally, a direct time integration gives (2.4.15). ■

From this inequality, we arrive at the notion of stochastic stability with stopping time, summarised in the following theorem.

Theorem 2.4.1 (Finite dimension)**.** *For the stochastic Euler-Poincaré equation considered here, a fixed point μ_e is weakly stable in probability with stopping time if it is stable for the deterministic dynamics in the sense of the energy-Casimir method.*

Proof. Given the constant Σ estimated before, we apply Markov inequality to the norm $\|\delta\mu\|^2_{H_C}$, that is

$$\begin{aligned}
P(\|\delta\mu_t\|^2_{H_C} > \epsilon) &\leq \frac{1}{\epsilon}\mathbb{E}(\|\delta\mu_t\|^2_{H_C}) \\
&\leq \frac{1}{\epsilon}\|\delta\mu(0)\|^2_{H_C}e^{\Sigma^2 t}\,,
\end{aligned}$$

to obtain the bound

$$P(\|\delta\mu_t\|^2_{H_C} > \epsilon) \leq \frac{1}{\epsilon}\|\delta\mu(0)\|^2_{H_C}e^{\Sigma^2 t}\,.$$

Thus, for all $\epsilon > 0$ we can find $\delta = \|\delta\mu(0)\|^2_{H_C}$ and $\sigma \geq \frac{1}{\epsilon}\|\delta\mu(0)\|^2_{H_C} e^{\Sigma^2 t}$ up to time $T_{\max} = \frac{1}{\Sigma^2}\ln\left(\frac{\epsilon}{\delta}\right)$. ■

This theorem then states that the multiplicative noise that we used to perturb the original mechanical systems not only preserves the geometrical structure of the equation but also preserves more subtle properties such as stability. In fact, this result is a direct consequence of the fact that the noise preserves the geometrical structure of the equation and allows the use of the energy-Casimir method. One could, for example, consider additive noise in the rigid body, and no such stability result would hold.

2.4.3 Application to the rigid body

We now apply this theory to our traditional example of the stochastic free rigid body to study the stochastic stability of the deterministically stable equilibrium. Using the deterministic energy-Casimir method, one can show that the equilibrium of lowest and highest energy are nonlinearly stable. They are the one we will study now. Without loss of generality, we assume that the position Π_1 is stable, that is we choose $\mathbb{I}_1 > \mathbb{I}_2 > \mathbb{I}_3$. A noise compatible with this stable position is a single Weiner process with amplitude $\boldsymbol{\sigma} = \sigma e_1$. In fact, it is the only compatible noise as one can directly check from the equations of motion. Recall that the Casimir is the invariant function of the momentum sphere and that $C(\boldsymbol{\Pi}) = \Psi(\frac{1}{2}\|\boldsymbol{\Pi}\|^2)$ for an arbitrary function $\Psi : \mathbb{R} \to \mathbb{R}$ is still a Casimir. This function is used to enforce that $DH_C = 0$ and D^2H_C is positive definite, where

$$H_C(\Pi) = \frac{1}{2}\boldsymbol{\Pi} \cdot \boldsymbol{\Omega} + \Psi\left(\frac{1}{2}\|\boldsymbol{\Pi}\|^2\right). \tag{2.4.16}$$

We refer to the complete treatment of the rigid body in Holm et al. (1985) for more detail on this function. The second variation of the Hamiltonian is directly computed and is

$$D^2H_C(\Pi_e)(\delta\boldsymbol{\Pi})^2 = \left(\frac{1}{\mathbb{I}_2} - \frac{1}{\mathbb{I}_1}\right)\delta\Pi_2^2 + \left(\frac{1}{\mathbb{I}_3} - \frac{1}{\mathbb{I}_1}\right)\delta\Pi_3^2 + \Psi''\left(\frac{1}{2}\right)\delta\Pi_1^2,$$

where we used the condition for $DH_C = 0$ which gives that Ψ should be evaluated at $\frac{1}{2}$. The extra freedom of Ψ yields the positive definiteness of this norm, thus the deterministic stability of these equilibria. We now derive the constant Σ for the stochastic system and find, by directly computing (2.4.13), that

$$\begin{aligned} D_{\delta\Pi}\|\delta\mathbf{\Pi}\|^2_{H_C} &= \sigma^2\left(\delta\Pi_3^2(\mathbb{I}_3^{-1}-\mathbb{I}_2^{-1})+\delta\Pi_2^2(\mathbb{I}_2^{-1}-\mathbb{I}_3^{-1})\right) \\ &\leq \sigma^2\left(\delta\Pi_3^2(\mathbb{I}_3^{-1}-\mathbb{I}_1^{-1})+\delta\Pi_2^2(\mathbb{I}_2^{-1}-\mathbb{I}_1^{-1})\right) \\ &\leq \sigma^2\|\delta\mathbf{\Pi}\|^2_{H_C}\,, \end{aligned}$$

where the last inequality is valid only for the positive definite case. For the negative definite case, one just needs an overall minus sign in the computation to obtain the same result. We therefore find that $\Sigma^2 = \sigma^2$ and does not depend on the geometry of the rigid body.

Notice that in the integrable case, that is if $\mathbb{I}_2 = \mathbb{I}_3$, the stability result is much stronger as we have $\Sigma = 0$. This result is to be expected as this system is now equivalent to the Kubo oscillator in the e_1 direction and the noise cannot drive the systems toward larger orbits in this direction. For this integrable system, the stability is even stronger and corresponds to the usual stability in probability.

We now illustrate this effect of progressively losing stability with numerical simulations. We compute the marginal probability density of the rigid body along Π_1, that is

$$\mathbb{P}_1(\Pi_1,t) := \int \mathbb{P}(\Pi_1,\Pi_2,\Pi_3,t)d\Pi_2 d\Pi_3\,, \tag{2.4.17}$$

and display it as a function of time in figure 2.8. In the deterministic case, the rigid body would oscillate near the top line of the plot. The solid line represents the mean value of Π_1, and one can clearly see that after some time, the mean of this distribution converges to 0 and the distribution becomes uniform. From the theory developed before, if one fixes a value for Π_1, it is possible to estimate with Σ the time when the solid line becomes smaller than this chosen value.

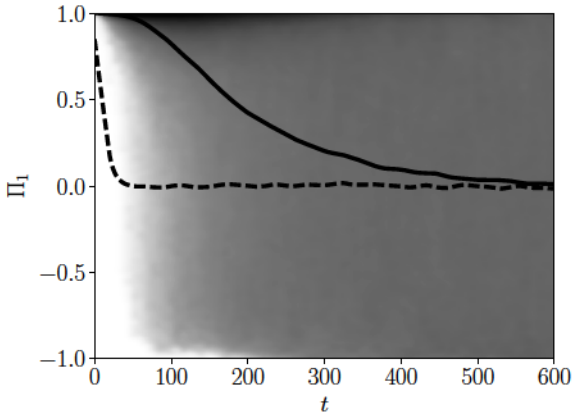


Figure 2.8: This figure displays the marginal probability density of the rigid body along Π_1. The initial condition is taken close to the equilibrium solution $\mathbf{\Pi} = (1, 0, 0)$, and the deterministic path would remain close to it. The solid line is the mean path and highlights the transition between the deterministic stable regime near $\Pi_1 = 1$ and the uniform distribution explained by ergodicity of the system. The dashed line is the mean path of the stochastic rigid body with an isotropic noise of the same amplitude, illustrating the immediate loss of stability.

2.4.4 Summary and outlook

This section can be seen as complementary to the first two sections of this chapter as the main focus here was on the dynamics of the stochastic Euler-Poincaré equation for short times. We previously considered the effect of noise and dissipation for large times through the stationary distribution of the Fokker-Planck equation or the random attractors. The main interest of stochastic dynamics for small times is in the stability of equilibrium. Indeed, we know from the ergodicity property that the system will eventually relax to form a uniform probability density on the coadjoint orbit. The main result obtained regarding the stability of a deterministically stable relative equilibria is that they remain stable in probability for a short period of time. The timescale for which it remains stable can even be estimated from above and is proportional to the amplitude of the noise in the case of rigid body motion. This result holds only if the noise amplitude is compatible with the equilibrium position such that the equilibrium persists in the stochastic system. This condition is rather strict but necessary. For more general

noises such as isotropic noises, the numerics suggest that there is no possible notion of stability as the system quickly leaves the neighbourhood of the initial position. The other important open problem is in the estimation of the timescale. In all the examples we can estimate it but we do not have a general explicit formula for it (or rather for the coefficient Σ). Here, we only considered the rigid body example, but in Arnaudon et al. (2017b) we also investigated the heavy top and the stochastic fluid dynamics and in particular the compressible two dimensional Euler equations. Surprisingly, the value of Σ for the fluid equation depends on the gradient of the noise amplitude with respect to the space variable instead of its amplitude as in the rigid body or the heavy top example. From this investigation, several research directions are possible. We can mention the extension of this theory to the more general energy-momentum method, the inclusion of geometric dissipation, or its use in control systems for stabilisation of unstable equilibria.

2.5 Stochastic Toda lattices

Until now, we studied stochastic perturbations of mechanical systems such as the rigid body and the heavy top. We will go a little further here by looking at the Toda lattice. This system can be high dimensional and has a more subtle Lie group structure than the free rigid body. It is interesting from the point of view of integrable systems and geometric mechanics as it is completely integrable and has a coadjoint orbit structure. We will apply the theory of stochastic deformation of the first section 2.2 as well as introduce a double bracket dissipation similar the one of section 2.3.4. We will then discuss how these deformation affect the Toda lattice and we will obtain results slightly different from the ones of the two previous sections.

We will begin by a review of the geometric structure of the Toda lattice in section 2.5.1, then introduce noise in section 2.5.2, dissipation in section 2.5.3 to end with a different noise which preserves isospectrality in section 2.5.4.

2.5.1 The deterministic Toda lattice

In this section, we begin with a short review of the Toda lattice with a focus on its geometrical structure. The Toda lattice has its roots in the origin of the theory of integrable systems with the discovery of recurrences in the FPU lattice by Fermi et al. (1955), with numerical simulations performed by M. Tsingou, see Dauxois (2008) for more on this story. The surprise was that for large times, no statistical equilibrium emerged for this lattice despite theoretical predictions of statistical physics. It turns out that the FPU chain is an approximation of the Toda lattice, which is an integrable system and has no thermal equilibrium. In the FPU chain, if one waits for a long enough time such that the non-integrable effects appear, this expected equilibrium can be observed in precise numerical simulations, see for example Benettin et al. (2013), Benettin and Ponno (2011).

The Toda lattice was first formulated in Toda (1967) (see also Toda (1989)) and later proved to be integrable by Flaschka (1974a,b) and Henón (1974). These works triggered numerous studies on such lattices, with various con-

nections to geometry.

The open Toda lattice

The Toda lattice consists of n particles with positions x_i and velocity y_i on a line. The only force in this lattice is the exponential interaction between neighbours described by the following Hamiltonian function

$$H_p = \sum_{i=1}^{n} \left(\frac{1}{2} y_i^2 + e^{x_{i+1} - x_i} \right) , \tag{2.5.1}$$

where we removed all the physical constants. One can add a constant in the exponential to vary the strength of the interaction or a mass in the kinetic energy term, but we will stick to this simple case here.

The open Toda lattice is a canonical Hamiltonian system for the variables $(x_i, y_i) \in T^*\mathbb{R}^n$ with Hamiltonian equations

$$\begin{aligned} \dot{x}_i &= y_i \\ \dot{y}_i &= e^{x_i - x_{i-1}} - e^{x_{i+1} - x_i} . \end{aligned} \tag{2.5.2}$$

If one approximates the exponential interaction by a Taylor expansion, we obtain the FPU lattice by keeping only the linear and quadratic terms.

We now introduce the change of variables due to Flaschka (1974b) used to prove the complete integrability of this system. The new variables are given by

$$\begin{aligned} a_k &= \frac{1}{2} y_k \\ b_k &= \frac{1}{2} e^{(x_{k+1} - x_k)/2} , \end{aligned} \tag{2.5.3}$$

and the equations of motion become

$$\begin{aligned} \dot{b}_i &= b_i (a_i - a_{i+1}) \\ \dot{a}_i &= 2(b_{i-1}^2 - b_i^2) . \end{aligned} \tag{2.5.4}$$

We did not mention the boundary conditions which can be open or periodic without affecting the integrability of the lattice. The simplest proof of complete integrability is achieved by constructing the Lax pair with

$$L = \begin{bmatrix} a_1 & b_1 & 0 & \dots \\ b_1 & a_2 & b_2 & \dots \\ 0 & b_2 & a_3 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \tag{2.5.5}$$

and $M = L^+ - L^-$, where L^+ is the upper part of the matrix and L^- the lower part. Integrability is then a consequence of the Lax equation $\dot{L} = [L, M]$.

The aim now is to introduce more geometry to understand this Lax pair construction as a coadjoint motion on the dual of a certain Lie algebra. Only then it will be possible to introduce noise and dissipation in a consistent manner.

Geometry of the Toda lattice

Many interesting works exist on the geometry of the Toda flow. We will follow here the derivation of Deift et al. (1989), but other works include Deift et al. (1986) or even Kostant (1979) with applications to representation theory.

This geometrical structure of the Toda lattice arises from the standard LQ decomposition of matrices $g \in GL(n)$ with only positive determinants. This decomposition produces $g_Q = SO(n)$ and $g_L \in L(n)$ such that $g = g_Q g_L$, where $L(n)$ is the group of lower triangular matrices and $SO(n)$ the special rotation group. From this decomposition, there is a group multiplication defined as $h * g := g_Q h g_L$ which makes the space of matrices with positive determinant, denoted G_{QL}, a Lie group. There also exists an isomorphism $m : G_{QL} \to SO(n) \times L(n)$ given by $g \mapsto (g_Q^{-1}, g_L)$ which can be used to construct a Lie bracket on the Lie algebra $\mathfrak{g}_{QL}$. We first define by $\mathfrak{l}$ the Lie algebra of $L(n)$, by $\mathfrak{q}$ the Lie algebra of $SO(n)$ and two projections onto these

Lie algebras by

$$\pi_{\mathfrak{q}} L = L_+ - L_+^T \quad \text{and} \quad \pi_{\mathfrak{l}} L = L_- + L_0 + L_+^T \,, \tag{2.5.6}$$

for a generic element $L \in \mathfrak{g}_{QL}$. The subscripts $+$ means the strict upper triangular part, the subscript $-$ the strict lower triangular part and the subscript 0 the diagonal of the matrix L. The Lie bracket on $\mathfrak{g}_{QL}$ is then defined as

$$[\xi, \eta]_{\mathfrak{g}_{QL}} = [\pi_{\mathfrak{l}} \xi, \pi_{\mathfrak{l}} \eta] - [\pi_{\mathfrak{q}} \xi, \pi_{\mathfrak{q}} \eta] \,, \tag{2.5.7}$$

for $\xi, \eta \in \mathfrak{g}_{QL}$. We now recall the adjoint and coadjoint action of the group G_{QL} on its Lie algebra

$$\mathrm{Ad}_g(\xi) = g_Q^{-1}(\pi_{\mathfrak{q}} \xi) g_Q + g_L^{-1}(\pi_{\mathfrak{l}} \xi) g_L \tag{2.5.8}$$

$$\mathrm{Ad}^*_{g^{-1}}(\xi) = \pi_{\mathfrak{l}^T} g_Q^{-1} \xi g_Q + \pi_{\mathfrak{q}^T} g_L^{-1} \xi g_L \,, \tag{2.5.9}$$

where the transpose and adjoint are defined with respect to the trace pairing

$$\langle \xi, \eta \rangle = \mathrm{Tr}(\xi^T \eta) \,, \tag{2.5.10}$$

that is we have

$$\pi_{\mathfrak{l}^T} L = L_+ - L_-^T, \qquad \pi_{\mathfrak{q}^T} L = L_- + L_0 + L_-^T \,. \tag{2.5.11}$$

We will now need the classical R-bracket, defined from the R-matrix

$$R = \frac{1}{2}(\pi_{\mathfrak{q}} - \pi_{\mathfrak{l}}) = \pi_{\mathfrak{q}} - \frac{1}{2} \,. \tag{2.5.12}$$

We can then rewrite the bracket (2.5.7) as the R-bracket

$$\mathrm{ad}^R_\xi \eta := [\xi, \eta]_R = [R\xi, \eta] + [\xi, R\eta] \,, \tag{2.5.13}$$

and define the adjoint of the R-matrix as

$$\langle \xi, R\eta \rangle =: \langle R^*\xi, \eta \rangle \,, \tag{2.5.14}$$

or in term the trace of matrices

$$\mathrm{Tr}(\xi^T R\eta) = \mathrm{Tr}(\eta^T R^T \xi) \,.$$

We thus have

$$R^* = R^T = \frac{1}{2}(\pi_{\mathfrak{l}^T} - \pi_{\mathfrak{q}^T}) = \pi_{\mathfrak{q}^T} - \frac{1}{2} \,, \tag{2.5.15}$$

because $\mathfrak{l}^\perp$ is adjoint of $\mathfrak{q}$ and $\mathfrak{q}^\perp$ is adjoint of $\mathfrak{l}$ under the trace pairing. We can now compute the coadjoint action on this Lie algebra.

Proposition 2.5.1. *The coadjoint action of $\mathfrak{g}_{GL}$ on its dual in term of the R-matrix* (2.5.12) *is*

$$\mathrm{ad}^{R,*}_\xi \mu = [(R\xi)^T, \mu] + R^T[\xi^T, \mu] \,, \tag{2.5.16}$$

for $\xi \in \mathfrak{g}_{QL}$ and $\mu \in \mathfrak{g}^*_{QL}$.

Proof. We compute

$$\begin{aligned}
\mathrm{Tr}(\mu^T[\xi,\eta]_R) &= \mathrm{Tr}(\mu^T[R\xi,\eta] + \mu^T[\xi, R\eta]) \\
&= \mathrm{Tr}(\mu^T R\xi\eta - \mu^T\eta R\xi + \mu^T\xi R\eta - \mu^T R\eta\xi) \\
&= \mathrm{Tr}(-\mu(R\xi)^T\eta^T + \mu\eta^T(R\xi)^T + \mu^T\xi R\eta - \xi\mu^T R\eta) \\
&= \mathrm{Tr}([(R\xi)^T,\mu]\eta^T + [\mu^T,\xi]R\eta) \\
&= \mathrm{Tr}([(R\xi)^T,\mu]\eta^T + \eta^T R^T[\xi^T,\mu]) \,,
\end{aligned}$$

and get the result for the coadjoint action. ∎

The corresponding R-Lie-Poisson bracket is the usual bracket

$$\{F,G\}(L) = \mathrm{Tr}\left(L^T\left[\frac{\delta F}{\delta L}, \frac{\delta G}{\delta L}\right]_R\right) = \mathrm{Tr}\left(\mathrm{ad}^{R,*}_{\frac{\delta F}{\delta L}} L \frac{\delta G}{\delta L}^T\right) \,. \tag{2.5.17}$$

For this Lie-Poisson bracket, one can see that all traces of powers of L, i.e. $H_i = \frac{1}{n}\mathrm{Tr}(L^i)$ are commuting integral of motion. Notice that for n particles, only n integrals are independent. Now, the Lie-Poisson equation is for any Hamiltonian H_i given by

$$\dot{L} = \left[\left(R\frac{\delta H}{\delta L}\right)^T, L\right] + R^T\left[\frac{\delta H}{\delta L}^T, L\right]. \tag{2.5.18}$$

The usual Toda flow is recovered by using $H_1 = \frac{1}{2}L^2$, so that $\delta H_1 = L$. As L is a symmetric matrix, the above second term drops. Finally, using the definition of the R-matrix in term of $\pi_{\mathfrak{q}}$, we arrive at the Toda flow

$$\dot{L} = [(RL)^T, L] = [(\pi_{\mathfrak{q}}L)^T, L] = -[\pi_{\mathfrak{q}}L, L]\,. \tag{2.5.19}$$

2.5.2 The stochastic Toda lattice

We are now able to directly apply the present theory of stochastic deformation. We will do it in the Lie-Poisson form and write

$$dF(L) = \{F, H_1\}\,dt + \sum_i \{F, \Phi_i\} \circ dW_t^i\,, \tag{2.5.20}$$

where we use a set of Wiener processes indexed by i. The stochastic potentials $\Phi_i : \mathfrak{g}^* \to \mathbb{R}$ are also taken to be linear in the momentum variable L, that is $\Phi_i = \mathrm{Tr}\left(L^T\sigma_i\right)$, where the set of σ_i span the symmetric tri-diagonal matrices with zeros on the diagonal. The choice of the σ_i fields is, in general, free but we will use here the simplest type of noise, that is the uniform noise where $\|\sigma_i\| = \sigma$ for all i. The stochastic Toda lattice is then given by

$$dL = -\left[\pi_{\mathfrak{q}}L, L\right]dt + \sum_i \mathrm{ad}^{R,*}_{\sigma_i}L \circ dW_t^i\,, \tag{2.5.21}$$

or, more explicitly in term of the R-matrix

$$dL = -\left[\pi_{\mathfrak{q}}L, L\right]dt + \sum_i ([(R\sigma_i)^T, L] + R^T[\sigma_i, L]) \circ dW_t^i\,. \tag{2.5.22}$$

By construction, this flow preserves the coadjoint orbit structure, namely the tri-diagonal symmetric matrices.

Proposition 2.5.2. *The stochastic Toda lattice* (2.5.22) *preserves the tridiagonal symmetric form of the matrix L. Furthermore, the stochastic processes for a_i and b_i are given by*

$$\begin{aligned} \dot{b}_i &= b_i(a_i - a_{i+1}) \\ da_i &= 2(b_{i-1}^2 - b_i^2)dt + 2\sigma_{i-1}b_{i-1}dW_t^{i-1} - 2\sigma_i b_i dW_t^i \,. \end{aligned} \tag{2.5.23}$$

Proof. A direct calculation from (2.5.22) gives equations (2.5.23). Not having any noise in the b_i equation implies that the Itô correction term vanishes and the Itô or Stratonovitch interpretation of the stochastic integral have the same form. ■

Then, with Flaschka's change of variable (2.5.3), the stochastic Toda lattice is a stochastic Hamiltonian system for (x_i, y_i) given by

$$\begin{aligned} \dot{x}_i &= y_i \\ dy_i &= \left(e^{x_i - x_{i-1}} - e^{x_{i+1}-x_i}\right) dt + \\ &\quad + 2\sigma_{i-1}e^{(x_i - x_{i-1})/2}dW_t^{i-1} - 2\sigma_i e^{(x_{i+1}-x_i)/2}dW_t^i \,. \end{aligned} \tag{2.5.24}$$

This stochastic process is in fact a type of Langevin equation with multiplicative noise as it can be written in term of the position variable only as

$$\begin{aligned} \ddot{x}_i &= e^{x_i - x_{i-1}} - e^{x_{i+1}-x_i} \\ &\quad + 2\sigma_{i-1}e^{(x_i - x_{i-1})/2}\dot{W}_t^{i-1} - 2\sigma_i e^{(x_{i+1}-x_i)/2}\dot{W}_t^i \,. \end{aligned} \tag{2.5.25}$$

This Langevin equation can also be written in a variational principle, given as

$$S = \int \sum_i \left(\dot{x}_i^2 - e^{x_{i+1}-x_i}\right) dt + \int \sum_i 4\sigma_i e^{(x_{i+1}-x_i)/2} \circ dW_t^i \,, \tag{2.5.26}$$

which will correspond to the associated Euler-Lagrange equations. From this,

the stochastic potentials are given in these variables by

$$\Phi_i = 4\sigma_i e^{(x_{i+1}-x_i)/2}\,. \tag{2.5.27}$$

This system is of course a stochastic Hamiltonian system and can be written as

$$dx_i = \frac{\delta H_1}{\delta y_i}dt\,, \qquad dy_i = -\frac{\delta H_1}{\delta x_i}dt - \sum_i \frac{\delta \Phi_i}{\delta x_i}\circ dW_t^i\,. \tag{2.5.28}$$

Notice that there is no noise in the x_i variable because the stochastic potential does not depend on y explicitly. Interestingly, it is not possible to add noise to the x equation while remaining consistent with this theory of stochastic deformation. First, one can check that the noise term can never have a non-diagonal part, and second that σ_i must not depend on the dynamical variables.

Conservation laws

It can easily be seen from the equations of motion (2.5.24) that $H_0 = \sum_i y_i$ remains an exact conserved quantity, even in the stochastic context. The other deterministically conserved quantities H_i are not conserved by the stochastic flow. This is to be expected as the noise only conserves the coadjoint orbit structure, and H_i are not Casimirs of this orbit, but just dynamically conserved quantities. To understand this better, we compute their time evolution

$$\begin{aligned}\frac{1}{k+1}dL^{k+1} &= \langle L^k, \mathrm{ad}^*_{\sigma_i}L\rangle \circ dW_t^i \\ &= \frac{1}{2}\langle \mathrm{ad}_{\sigma_i}L^k, \mathrm{ad}^*_{\sigma_i}L\rangle dt + \frac{k}{2}\langle \mathrm{ad}_{\sigma_i}(L^{k-1}\mathrm{ad}^*_{\sigma_i}L), L\rangle dt + \langle \mathrm{ad}_{\sigma_i}L^k, L\rangle dW_t^i \\ &= \frac{k}{2}\langle L^{k-1}\mathrm{ad}^*_{\sigma_i}L, \mathrm{ad}^*_{\sigma_i}L\rangle dt + \langle \mathrm{ad}_{\sigma_i}L^k, L\rangle dW_t^i\,.\end{aligned}$$

In the case of the energy for $k=1$, the mean path of the energy satisfies

$$D_L H_1 = \frac{1}{2}\|\mathrm{ad}^*_{\sigma_i}L\|^2\,, \tag{2.5.29}$$

where D_L is the infinitesimal generator of the process for L defined in section 2.4 in the context of stochastic stability. Recall that it is the conditional expectation of a time derivative of a stochastic process. Let us see in physical coordinates the form of this equation by a direct computation.

Proposition 2.5.3. *If the noise is uniform, that is $\|\sigma_i\| = \sigma$ for all i, the mean energy increases in time as*

$$D_L H_1(L(t)) \leq 4\sigma^2 \mathbb{E}(V(L(t))|\mathcal{F}_t)\,, \tag{2.5.30}$$

where V denotes the deterministic potential energy the expectation on the right hand side is the conditional expectation, this equation is thus still stochastic.

Proof. We compute the Itô correction to get

$$\begin{aligned}\sum_i y_i dy_i &= \sum_i 2\sigma y_i e^{(x_i - x_{i-1})/2} \circ dW_t^{i-1} - \sum_i 2\sigma y_i e^{(x_{i+1}-x_i)/2} \circ dW_t^i \\ &= \sum_i 2\sigma y_i e^{(x_i - x_{i-1})/2} dW_t^{i-1} - \sum_i 2\sigma y_i e^{(x_{i+1}-x_i)/2} dW_t^i \\ &+ 2\sigma^2 \sum_i \left(e^{x_i - x_{i-1}} + e^{x_{i+1}-x_i}\right) dt - 4\sigma^2 \sum_i e^{(x_{i+1}-x_{i-1})/2} dt\,,\end{aligned}$$

yielding the result after discarding the negative and stochastic terms. ■

Similarly, with one of our result on the stochastic stability of section 2.4, we have an upper bound in time for the expected value of the energy.

Proposition 2.5.4. *We have for uniform noise the following time-dependent upper bound for the mean energy*

$$\mathbb{E}(H_1(L(t))) \leq H_1(L(0)))\, e^{4\sigma^2 t}\,. \tag{2.5.31}$$

Proof. The proof is similar to the computation for the stability result of proposition 2.4.2 in section 2.4 where we proved a similar formula. ■

This proposition demonstrates that the mean effect of the noise in this system is to introduce energy into the system, which can, in turn, be unbounded.

This property suggests that we would need a dissipative mechanism that would balance this energy input. This will be done in the next section 2.5.3. From the non-conservation of these quantities, it is also clear that this equation cannot be stochastic integrable in any sense. We will study this fact in more details in section 2.5.4.

2.5.3 Dissipative stochastic Toda lattice

We want to introduce a nonlinear dissipation term in the Toda lattice which is compatible with the geometrical structure used to add noise. The best way to do this is to introduce, as before, a double bracket dissipation. Recall that the generalised double bracket dissipation was constructed using the Casimir of the Lie-Poisson bracket. Here, although the Toda lattice has many conserved quantities, there is no Casimir of the Lie-Poisson bracket derived with the R-matrix. We thus need a slightly different term for the double bracket dissipation.

The double bracket dissipation

We are going to show that the following double bracket dissipation

$$\dot{L} = \mathrm{ad}^{R,*}_L L + \theta \mathrm{ad}^{R,*}_{\mathrm{ad}^{R,*}_L L} L \tag{2.5.32}$$

preserves the coadjoint orbit and is a dissipative term. Recall that $\mathrm{ad}^{R,*}$ is the coadjoint action with respect to the R-matrix. For simplicity, we will drop the upper R for the adjoint and coadjoint actions in the rest of this section. Notice that we do not need to convert the second ad^* operator to the Lie algebra as it is a symmetric matrix and the transpose operation corresponds to the isomorphism $\flat : \mathfrak{g}^* \to \mathfrak{g}$. First, the Lie-Poisson formulation of (2.5.32) is

$$\dot{F}(L) = \{F, H_1\}_R - \theta \left\langle \mathrm{ad}^*_{\frac{\delta H_1}{\delta L}} \frac{\delta H_1}{\delta L}, \mathrm{ad}^*_{\frac{\delta F}{\delta L}} \frac{\delta H_1}{\delta L} \right\rangle . \tag{2.5.33}$$

Indeed, we just compute

$$\begin{aligned}\langle \nabla F, \mathrm{ad}^*_{\mathrm{ad}^*_L L} L\rangle &= \left\langle \mathrm{ad}_{\mathrm{ad}^*_L L}\frac{\delta F}{\delta L}, L\right\rangle \\ &= -\left\langle \mathrm{ad}_{\frac{\delta F}{\delta L}}\mathrm{ad}^*_L L, L\right\rangle \\ &= -\left\langle \mathrm{ad}^*_L L, \mathrm{ad}^*_{\frac{\delta F}{\delta L}} L\right\rangle ,\end{aligned}$$

and use $\frac{\delta H_1}{\delta L} = L$ to get (2.5.33). Notice that it is different from the usual double bracket dissipation where the adjoint action is used instead of the coadjoint action here.

In term of a_i, b_i we have the following dissipative Toda lattice

$$\begin{aligned}\dot{b}_i &= b_i(a_i - a_{i+1}) + 2\theta b_i(b_{i-1}^2 - 2b_i^2 + b_{i+1}^2) \\ \dot{a}_i &= 2(b_{i-1}^2 - b_i^2) + 2\theta\left(b_i^2(a_{i+1} - a_i) - b_{i-1}^2(a_i - a_{i-1})\right) .\end{aligned} \tag{2.5.34}$$

Written in this form, the new terms have a clear meaning of dissipation as they both behave as discrete nonlinear second order operators as we will see in more detail in the next section.

In term of the physical coordinates, the dissipative stochastic equation is

$$\begin{aligned}\dot{x}_i &= y_i + \frac{1}{2}\theta(e^{x_i - x_{i-1}} - 2e^{x_{i+1}-x_i} + e^{x_{i+2}-x_{i+1}}) \\ dy_i &= \left(e^{x_i - x_{i-1}} - e^{x_{i+1}-x_i}\right) dt + \\ &\quad + 2\sigma e^{(x_i - x_{i-1})/2} \circ dW_t^{i-1} - 2\sigma e^{(x_{i+1}-x_i)/2} \circ dW_t^i \\ &\quad + \frac{1}{2}\theta\left((y_{i+1} - y_i)e^{x_{i+1}-x_i} - (y_i - y_{i-1})e^{x_i - x_{i-1}}\right) dt .\end{aligned} \tag{2.5.35}$$

Let's now show that the energy decays when $\theta > 0$.

Proposition 2.5.5. *The conserved quantities of the Toda lattice satisfy the differential equation*

$$\frac{d}{dt}\frac{1}{i+1}\mathrm{Tr}(L^{i+1}) = -\theta\langle \mathrm{ad}^*_{\nabla H}\nabla H, \mathrm{ad}^*_{\nabla H_i}\nabla H\rangle , \tag{2.5.36}$$

when the double bracket dissipation (2.5.32) *is added.*

Proof. This formula is a direct consequence of the Lie-Poisson form (2.5.33). ∎

For the Hamiltonian H_1 we have

$$\frac{d}{dt}H_1 = -\theta\|\mathrm{ad}^*_{\nabla H}\nabla H_1\|^2\,. \tag{2.5.37}$$

The negative right-hand sign of this equation shows that the energy decays. We can obtain the same result by a more direct computation.

Proposition 2.5.6. *The energy dissipation, in physical coordinates, is given by*

$$\frac{d}{dt}H_1 = -2\theta\sum_i \left((y_i - y_{i+1})^2 e^{x_{i+1}-x_i} + (e^{x_{i+1}-x_i} - e^{x_i - x_{i-1}})^2\right)\,. \tag{2.5.38}$$

Proof. We first compute

$$\begin{aligned}
\sum_i y_i dy_i &= \theta\sum_{i=1}^{N} y_i\left((y_i - y_{i-1})e^{x_i-x_{i-1}} + \theta\sum_{i=0}^{N-1}(y_i - y_{i+1})e^{x_{i+1}-x_i}\right)dt \\
&= \theta\sum_{i=0}^{N-1} y_{i+1}(y_{i+1} - y_i)e^{x_{i+1}-x_i} + \theta\sum_{i=0}^{N-1} y_i(y_i - y_{i+1})e^{x_{i+1}-x_i}dt \\
&= \theta\sum_{i=0}^{N-1}(y_i^2 + y_{i+1}^2 - 2y_iy_{i+1})e^{x_{i+1}-x_i}dt \\
&= \theta\sum_{i=0}^{N-1}(y_i - y_{i+1})^2 e^{x_{i+1}-x_i}dt\,.
\end{aligned}$$

The second term is given by similarly computing $x_i\dot{x}_i$. ∎

Without noise, this system will tend to a steady position (x_i^s, y_i^s) satisfying

$$\sum_i (y_i^s - y_{i+1}^s)^2 e^{x_{i+1}^s - x_i^s} + \sum_i (e^{x_{i+1}^s - x_i^s} - e^{x_i^s - x_{i-1}^s})^2 = 0\,.$$

The solution of this positive definite quantity is $x_i = x$ and $y_i = y$ for some number y and x. The value of y can be found from the conservation of total

momentum, but the value x does not have any direct dependence on the initial conditions if the lattice is periodic. Finding the equilibrium solution as a function of the initial condition is still an open problem for the periodic lattice. In the open Toda lattice, the first and last particles are at $-\infty$ and ∞ and the equilibrium solution is $x = \pm\infty$.

Continuum limit

One of the interesting features of the Toda lattice or the FPU lattice is that in the limit of an infinite number of particles we recover the Burger's equation or even the KdV equation if we include higher order terms in the approximation. The treatment of this limit is a rather subtle topic as the Toda lattice can have shocks and the Burger's model will no longer be valid after a shock. This was extensively studied by Deift and McLaughlin (1998) where various after-shocks scenarios were considered. We want to derive here the continuum limit of the dissipative stochastic Toda lattice and we will not consider the formation of shocks as the dissipation will prevent them from being formed. We will follow Deift and McLaughlin (1998) for the derivation of the continuum limit and start by defining the continuous functions interpolating through the discrete system as

$$a(\epsilon k, t) = a_k(t) \qquad \text{and} \qquad b(\epsilon k, t) = b_k(t)\,,$$

where ϵ corresponds to the rest distance between two particles. We then need to rescale the time as $t \to t/\epsilon$ and the diffusion coefficient (which we recall is a time scale) $\theta \to \theta/\epsilon$ to obtain from (2.5.34) the continuous equations

$$\begin{aligned} \dot{b} &= b(a_x + 2\theta\, b_{xx}) \\ \dot{a} &= 2\left(b^2(1+\theta\, a_x)\right)_x\,. \end{aligned} \tag{2.5.39}$$

In the non-dissipative case, setting $a = 2b$ recovers the inviscid Burger's equation, but not in this case.

We now add the noise into this equation and formally derive the continuum limit. For this we first assume a isotropic noise and rescale it accordingly as

$\sigma_i = \sigma/\sqrt{\epsilon}$ for all i. We then propose the following continuum approximation of the complete stochastic dissipative Toda lattice

$$\begin{aligned} \dot{b} &= b(a_x + 2\theta b_{xx}) \\ da &= 2\left(b^2(1+\theta a_x)\right)_x \, dt + 2\sigma b_x \circ dW_{t,x}\,, \end{aligned} \tag{2.5.40}$$

where the noise is white in space and time, or delta correlated, that is $\mathbb{E}[W_{x,t}, W_{x',t'}] = \delta(x-x')\delta(t-t')$. We will not rigorously prove that this stochastic PDE is the continuum limit of the stochastic Toda lattice as this will be beyond the scope of this investigation, but we want to verify that it make sense to have such a noise term by computing the limit of the auto-covariance function.

Proposition 2.5.7. *The auto-covariance of the discrete system* (2.5.35) *converges to the auto-covariance of the continuous model* (2.5.40).

Proof. We compute the auto-covariance of (2.5.40) to get

$$\begin{aligned} \mathrm{Cov}(a_k, a_{k'}) &= \frac{4\sigma^2}{\epsilon}\mathbb{E}\left(\int (b_{k-1}dW_t^{k-1} - b_k dW_t^k)\int (b_{k'-1}dW_t^{k'-1} - b_{k'}dW_t^{k'})\right) \\ &= \frac{4\sigma^2}{\epsilon^2}\mathbb{E}\int (b_k(\delta_{k,k'}b_{k'} - \delta_{k,k'-1}b_{k'-1}) \\ &\qquad - b_{k-1}(\delta_{k-1,k'}b_{k'} - \delta_{k-1,k'-1}b_{k'-1}))\,dt\,, \end{aligned}$$

where used the fact that the Wiener processes are not correlated and the Itô isometry, for a process μ_t

$$\mathbb{E}\left(\left(\int \mu_t dW\right)^2\right) = \mathbb{E}\left(\int \mu_t^2 dt\right)\,. \tag{2.5.41}$$

In the limit $\epsilon \to 0$, we arrive at

$$\begin{aligned} \lim_{\epsilon\to 0}\mathrm{Cov}(a_k, a_{k'}) &= 4\sigma^2\mathbb{E}\left(\int \delta(x-x')b_x(x)b_{x'}(x')dt\right) \\ &= \mathrm{Cov}(a(x), a(x'))\,, \end{aligned}$$

where the last step is computed from the process (2.5.40). ∎

This proposition suggests the following conjecture which remains to be proved.

Conjecture 2.5.1. *The solutions of the stochastic dissipative Toda lattice converge to the continuous process* (2.5.40) *almost surely.*

2.5.4 The isospectral stochastic Toda lattice

We end this study on the question of integrability of the stochastic Toda lattice. We saw that the noise introduced previously does not conserve any of the deterministic integrals of motion except the total momentum. We thus do not expect any stochastic integrability of this stochastic Toda lattice. It is also interesting to remark that the noise is not proportional to any other flow of the deterministic Toda lattice, or is not generated by a vector field compatible with the Toda flow (in the sense of commuting flows). Such a stochastic Toda lattice could be constructed but would not have a different Wiener process W_i for each particle, but rather a single Wiener process proportional to another flow of the Toda hierarchy. We will see in section 3.4 that a such stochastic deformation preserves integrability and appears naturally in the context of stochastic deformation of hierarchies of integrable PDEs.

We can nevertheless obtain an stochastic equation which seems integrable if we remove the term of the noise proportional to R^* in the stochastic Toda lattice equation (2.5.22) we obtain

$$dL = -[\pi_{\mathfrak{q}} L, L]\, dt + \sum_i [(R\sigma_i)^T, L] \circ dW_t^i \,. \tag{2.5.42}$$

This equation is isospectral as the deterministic Toda lattice but with a stochastic spectral problem similar to the notion of stochastic integrable systems which we will study in section 3.4.2.

Proposition 2.5.8. *The stochastic process* (2.5.42) *is isospectral with the*

following stochastic spectral problem

$$\begin{aligned} L\psi &= \lambda\psi \\ d\psi &= -\pi_{\mathfrak{q}} L\psi dt + \sum_i (R\sigma_i)^T \psi \circ dW_t^i \,. \end{aligned} \tag{2.5.43}$$

Proof. We compute

$$\begin{aligned} d(L\psi) &= dL\psi + Ld\psi \\ &= dL\psi - L\pi_{\mathfrak{q}} L\psi dt + L\sum_i (R\sigma_i)^T \psi \circ dW_t^i \,, \end{aligned}$$

and

$$\begin{aligned} d(\lambda\psi) &= d\lambda\psi + \lambda d\psi \\ &= -\lambda\pi_{\mathfrak{q}} L\psi dt + \lambda\sum_i (R\sigma_i)^T \psi \circ dW_t^i \\ &= -\pi_{\mathfrak{q}} LL\psi dt + \sum_i (R\sigma_i)^T L\psi \circ dW_t^i \,. \end{aligned}$$

where we used $L\psi = \lambda\psi$ for the last equality. The two computations are equal by using $L\psi = \lambda\psi$, so we collect terms and rearrange to obtain the stochastic Toda equation (2.5.42). ■

One can directly check that the conserved quantities H_i remain conserved for all i by this stochastic flow. The drawback of removing a stochastic term from the original equation is that the coadjoint orbit is not preserved any more by the stochastic flow. In fact, the matrix L is full even for tri-diagonal initial conditions, and the connection with the physical Toda lattice via Flaschka's change of variables is lost. This raises the question of what are the other necessary conserved quantities of the stochastic Toda lattice. In Deift et al. (1989), a generalised Toda flow on the coadjoint orbit of full symmetric matrices was proven to be completely integrable, and formulae for the extra conserved quantities were derived. We will not investigate this further here, but it may be possible that this stochastic Toda lattice is a stochastic integrable system on the coadjoint orbit of symmetric matrices.

This feature of the Toda lattice is different from before, where the coadjoint orbit could not be modified when changing stochastic terms.

2.5.5 Summary and outlook

In this section, we implemented the stochastic deformation and the double bracket dissipation in the integrable Toda lattice. Although this system is of the Euler-Poincaré form, its Lie algebra structure based on the R-matrix makes it rather different from the free rigid body. In particular, its dimension can be as large as we want and lead to partial differential equations in the continuum limit. As we saw in the previous sections, the non-compactness of a coadjoint orbit makes difficult the study of stationary distributions of the Fokker-Planck equation corresponding to the stochastic Euler-Poincaré equation. It is also the case here, and we, therefore, did not investigate this aspect of the stochastic Toda lattice. The coadjoint orbit of the Toda lattice is the main difference with other systems as it consists of symmetric matrices and does not have Casimir function. The noise which we introduced preserves this structure and allows the use of Flaschka change of variables to go back to the physical variables. Being able to preserve this change of variable allows us to transform the abstract equation on symmetric matrices back to a standard Langevin equation. This analogy with the Langevin equation suggests that more can be done in the direction of statistical physics, as we already suggested in the case of the Gibbs distribution of the rigid body.

During the course of this work, we also tried simple numerical experiments, but nothing worth mentioning here came out of them apart from the fact that the system reaches a statistical equilibrium if there is dissipation. Other investigations with numerical simulation may include the 3-particle Toda lattice and possible random attractors visualised with the Poincaré section. Another intriguing finding is the continuum limit of the dissipation and the noise. They both have a rather unusual form and may deserve more investigations which are out of the scope of this work.

What is in fact more interesting for us here is the question of complete integrability of the stochastic Toda lattice, with no dissipation. The stochastic

term is different from being of the form of a higher order Toda flow multiplying a Wiener process. In this case, the system would be integrable, as we will see for infinite dimensional systems later. Instead, the stochastic term has a different Wiener process for each particle, which makes the system non-integrable. In fact, we cannot recast the geometric stochastic deformation of this system into a stochastic isospectral problem unless one of the stochastic terms is discarded. In this case, the coadjoint orbit is not preserved, and a tri-diagonal initial condition will quickly fill the entire symmetric matrix, but we can write a stochastic iso-spectral problem. A question still remains, which is if we can solve this stochastic spectral problem.

2.6 Stochastic landmarks

This section is the last section on stochastic deformation of finite dimensional systems. We will consider another type of Hamiltonian systems describing the interaction of particles, different from the Toda lattice. This system describes the landmark dynamics arising from a finite dimensional reduction of a particular type of fluid equation of Euler-Poincaré form. We will not study this infinite dimensional equation in detail here but only briefly later in section 3.5.1. Here, we will instead directly derive the finite dimensional stochastic landmark equations.

The solution of the landmark system is a solution of the original fluid equation but has the advantage of being only finite dimensional. In the case of the one-dimensional fluid equation with the Sobolev H^1 norm, the landmark equation is the peakon equation of the integrable Camassa-Holm equation, which is also an integrable finite dimensional system. We will consider here a more general setting of arbitrary dimensions and norm as the main application of this theory is in computational anatomy where the fluid velocity field u describes a diffeomorphic deformation of an image. We refer to the monograph Younes (2010) for a complete exposition of such geometrical methods for image analysis and computational anatomy and to Holm and Marsden (2005) for the theory of singular reductions of the Euler-Poincaré equation on the diffeomorphism group. The results of this section are contained in Arnaudon et al. (2017d,e). Notice that only the case of peakon is completely integrable, and most of other shapes of landmark will destroy integrability.

We will begin by introducing noise into the landmark equation in the same way as we previously did and we will then proceed by studying its effect on the landmark dynamics. The main feature which allows us to retain the reduction procedure from the stochastic infinite dimensional system to a set of finite dimensional stochastic equations is the preservation by the noise of the geometrical structures of the deterministic equation and in particular the singular momentum map. We will then derive a finite dimensional approximation of the Fokker-Planck equation to solve the problem of estimating the noise amplitude from real data. We will end this section with several

numerical experiments to illustrates this theory.

2.6.1 Stochastic dynamics

We start by deriving the stochastic landmark equation from a stochastic variational principle. We will denote the position of the N landmarks by $\mathbf{q} = (\mathbf{q}_1, \dots, \mathbf{q}_N)$, where $\mathbf{q}_i \in \mathbb{R}^d$ for a d dimensional space. In practice, we will only use $d = 2$ to describe images but the equations are valid for any dimensions. We will also often use the coordinate notation q_i^α where α runs on the image dimension. Each particle has a velocity, and as we will describe them in a Hamiltonian framework, so we introduce their momentum $\mathbf{p} = (\mathbf{p}_0, \dots, \mathbf{p}_N)$, where $\mathbf{p}_i \in \mathbb{R}^d$.

From a configuration of these landmarks, it is possible to reconstruct the entire deformation vector fields $\mathbf{u} : \mathbb{R}^d \to \mathbb{R}^d$ as follows

$$\mathbf{u}(\mathbf{x}) = \sum_i \mathbf{p}_i K(\mathbf{q}_i - \mathbf{x}) \,, \tag{2.6.1}$$

where $K(\mathbf{x})$ is a given symmetric kernel function and $\mathbf{x} \in \mathbb{R}^d$ is the position on the image. The corresponding momentum $\mathbf{m} : \mathbb{R}^d \to \mathbb{R}^d$ is given by

$$\mathbf{m}(\mathbf{q}, \mathbf{p})(\mathbf{x}) = \sum_i \mathbf{p}_i \delta(\mathbf{q}_i - \mathbf{x}) \,. \tag{2.6.2}$$

This is a singular momentum map for the action of the diffeomorphism group on the canonically conjugate variables $\mathbf{q}, \mathbf{p}$. By using this momentum variable in term of the spatial coordinate $\mathbf{x}$ and not $\mathbf{q}$ and $\mathbf{p}$, m satisfies the Euler-Poincaré equation

$$\dot{\mathbf{m}}(\mathbf{x}) = \mathrm{ad}^*_{\mathbf{u}(\mathbf{x})} \mathbf{m}(\mathbf{x}) \,, \tag{2.6.3}$$

where ad^* is the coadjoint action of the diffeomorphism group for the kinetic Lagrangian $l(\mathbf{u}) = \frac{1}{2} \int \mathbf{u} \cdot \mathbf{m} \, d\mathbf{x}$. We will not go further in this direction but instead use the fact that (2.6.2) is a momentum map for the action of the diffeomorphism group on the Clebsch variables $(\mathbf{q}_i, \mathbf{p}_i)$, see Holm and Marsden

(2005). We start with the following stochastic Clebsch action functional

$$S(\mathbf{u},\mathbf{q},\mathbf{p}) = \iint l(\mathbf{u})\,d\mathbf{x}\,dt + \sum_i \int \mathbf{p}_i \cdot \left(\circ d\mathbf{q}_i - \mathbf{u}(\mathbf{q}_i)\,dt + \sum_l \sigma_l(\mathbf{q}_i) \circ dW_t^l \right), \tag{2.6.4}$$

where $\sigma_i : \mathbb{R}^d \to \mathbb{R}^d$ are arbitrary vector fields. We will now write this variational principle only in term of $(\mathbf{q}_i, \mathbf{p}_i)$ by first noticing that from (2.6.2) we have

$$2l(\mathbf{u}) = \int \mathbf{m}(\mathbf{q},\mathbf{p})(\mathbf{x}) \cdot \mathbf{u}(\mathbf{x}) d\mathbf{x} = \sum_i \mathbf{p}_i \cdot \mathbf{u}(\mathbf{q}_i)\,.$$

Combining the first two terms in the Hamiltonian function $h(\mathbf{q},\mathbf{p})$ we have the finite dimensional functional

$$S(\mathbf{q},\mathbf{p}) = h(\mathbf{q},\mathbf{p}) + \sum_i \int \mathbf{p}_i \cdot \left(\circ d\mathbf{q}_i + \sum_l \sigma_l(\mathbf{q}_i) \circ dW_t^l \right), \tag{2.6.5}$$

where the Hamiltonian is

$$h(\mathbf{q},\mathbf{p}) = \frac{1}{2} \sum_{i,j=1}^{N} (\mathbf{p}_i \cdot \mathbf{p}_j) K(\mathbf{q}_i - \mathbf{q}_j)\,. \tag{2.6.6}$$

This variational principle is often called a phase space variational principle as it is written in term of the Hamiltonian and includes a Legendre transformation with the term $\mathbf{p}_i \cdot d\mathbf{q}_i$. The noise appears only once in the variational principle and corresponds to the stochastic potentials

$$\phi_l(\mathbf{q},\mathbf{p}) = \sum_i \mathbf{p}_i \cdot \sigma_l(\mathbf{q}_i)\,. \tag{2.6.7}$$

By extending the landmark Hamiltonian (2.6.6) by these stochastic potentials, we can rewrite the variational principle (2.6.5) as a canonical Hamiltonian system with Hamiltonian $h dt + \sum_l \phi_l \circ dW_t^l$. By either taking free

variations of (2.6.5) or the Hamiltonian equation with the stochastic Hamiltonian, we obtain the following stochastic Hamiltonian equation for landmark dynamics

$$\begin{aligned} d\mathbf{q}_i &= \frac{\partial h}{\partial \mathbf{p}_i} dt + \sum_l \sigma_l(\mathbf{q}_i) \circ dW_t^l \\ d\mathbf{p}_i &= -\frac{\partial h}{\partial \mathbf{q}_i} dt - \sum_l \frac{\partial}{\partial \mathbf{q}_i} \left(\mathbf{p}_i \cdot \sigma_l(\mathbf{q}_i)\right) \circ dW_t^l \,, \end{aligned} \tag{2.6.8}$$

or, in coordinates,

$$\begin{aligned} dq_i^\alpha &= \frac{\partial h}{\partial p_i^\alpha} dt + \sum_l \sigma_l^\alpha(\mathbf{q}_i) \circ dW_t^l \\ dp_i^\alpha &= -\frac{\partial h}{\partial q_i^\alpha} dt - \sum_{l,\beta} \frac{\partial \sigma_l^\beta(\mathbf{q}_i)}{\partial q_i^\alpha} p_i^\beta \circ dW_t^l \,. \end{aligned} \tag{2.6.9}$$

It is worthwhile to mention that a solution of this stochastic ODE gives a solution of the stochastic Euler-Poincaré equation

$$d\mathbf{m}(\mathbf{x}) = \mathrm{ad}^*_{\mathbf{u}(\mathbf{x})}\mathbf{m}(\mathbf{x})\, dt + \sum_l \mathrm{ad}^*_{\sigma_l(\mathbf{x})}\mathbf{m}(\mathbf{x}) \circ dW_t^l \,, \tag{2.6.10}$$

which itself gives a stochastic path on the diffeomorphism group via the reconstruction relation

$$dg = u_t(g)\, dt + \sigma_l(g) \circ dW_t^l \,.$$

We refer to Arnaudon et al. (2017e) for more detail about this construction.

To go further in the study of the stochastic landmark dynamics, we need to specify a particular form for the noise fields σ_l. In principle, any set of $C^1(\mathbb{R}^d)$ functions could be used, but to have interesting and tractable solutions of these equations, we will choose Gaussian kernels at fixed positions with a

generic length scale α_l, amplitude $\lambda_l \in \mathbb{R}^d$ and position $\delta_l \in \mathbb{R}^d$ that is

$$\sigma_l^\alpha(\mathbf{q}_i) = \lambda_l^\alpha e^{-\frac{\|\mathbf{q}_i - \delta_l\|^2}{2\alpha_l^2}}. \tag{2.6.11}$$

With this choice of σ_l we can easily cover the entire image with noise and finely control its the spatial correlation.

This form of the noise thus represents local uncertainties in the image in such a way that the equation remains consistent with the variational principle and produces a stochastic path on the diffeomorphism group. In particular, there is no possible crossing of landmarks, although this is only interesting for $d = 1$, (see Holm and Tyranowski (2016) for such a study). Recently, stochastic landmark dynamics were investigated with only a stochastic forcing, that is an additive noise in the $\mathbf{p}$ equation. See for example Marsland and Shardlow (2017), Trouvé and Vialard (2012), Vialard (2013). For the present system, the noise appears as almost additive in the $\mathbf{q}$ equation and almost linear in the $\mathbf{p}$ equation. By this we mean that the noise still depends on the position of the landmark but if $\alpha_l \to \infty$ the noise becomes additive and all landmarks will have the same noise. In this case, the noise on the $\mathbf{p}$ equation also vanishes. We thus observe interested effects only for localised noise fields, i.e. $\alpha_l < \infty$. For example, the momentum of the landmarks is then coupled to the gradient of σ_l in the momentum equation.

2.6.2 The Fokker-Planck equation and moment dynamics

Before solving the stochastic landmark system numerically in section 2.6.3, we will study the corresponding Fokker-Planck equation and in particular derive a finite dimensional approximation of its dynamics.

The first point to notice here is the high dimensionality of this PDE, that is $\mathbb{R}^{2dN}$ where usually $d = 2$ and N is the number of landmarks that can be arbitrarily large for practical applications. The second point is that the phase space is not compact and it will thus not be interesting to study the stationary distribution of this Fokker-Planck equation. Instead, what is interesting to study in this case is the time evolution of the covariance of the error of each

landmark and how it depends on the noise and the landmark interactions. To better understand the evolution of the covariance, we will use a finite dimensional approximation of the Fokker-Planck equation to describe it as well as the mean trajectory of the landmarks.

The Fokker-Planck Equation

As in the previous sections, we denote the probability density of the stochastic process of the landmarks by $\mathbb{P}(\mathbf{q}, \mathbf{p})$. Recall that the stochastic process is Hamiltonian and can be written with the canonical Poison bracket of the form

$$\{F, G\}_{\text{can}}(\mathbf{q}, \mathbf{p}) = \sum_i \frac{\partial F}{\partial \mathbf{q}_i} \cdot \frac{\partial G}{\partial \mathbf{p}_i} - \sum_i \frac{\partial F}{\partial \mathbf{p}_i} \cdot \frac{\partial G}{\partial \mathbf{q}_i}, \tag{2.6.12}$$

where the derivatives are gradients and the scalar products are in the R^d directions. Using this bracket, the Fokker-Planck equation is directly computed following the theory of the previous section.

Proposition 2.6.1. *The Fokker-Planck equation of the process* (2.6.8) *is*

$$\frac{d}{dt}\mathbb{P}(\mathbf{q}, \mathbf{p}, t) = \{h, \mathbb{P}\}_{\text{can}} + \frac{1}{2}\sum_l \{\phi_l, \{\phi_l, \mathbb{P}\}_{\text{can}}\}_{\text{can}} := \mathcal{L}^*\mathbb{P}\,. \tag{2.6.13}$$

We do not write the proof, which is the same as before, but rather write the Fokker-Planck equation explicitly. We obtain

$$\begin{aligned}\frac{d}{dt}\mathbb{P}(\mathbf{q}, \mathbf{p}, t) = {} & \partial_{q_i}\mathbb{P}\partial_{p_i}h - \partial_{p_i}\mathbb{P}\partial_{q_i}h \\ & + \frac{1}{2}\sum_l \Big(\sum_{i,j} \partial_{q_j}\phi_l \partial q_i \phi_l \partial_{p_j}\partial_{p_i}\mathbb{P} - 2\sum_{i,j} \partial_{q_j}\phi_l \partial_{p_i}\phi_l \partial_{p_j}\partial_{q_i}\mathbb{P} \\ & + \sum_{i,j} \partial_{p_j}\phi_l \partial_{p_i}\phi_l \partial_{q_j}\partial_{q_i}\mathbb{P} + \sum_i \partial_{q_i}\phi_l \partial_{p_i}\partial_{q_i}\phi_l \partial_{p_i}\mathbb{P} \\ & - \sum_i \partial_{p_i}\phi_l \partial_{q_i}\partial q_i \phi_l \partial_{p_i}\mathbb{P} + \sum_i \partial_{p_i}\phi_l \partial_{q_i}\partial p_i \phi_l \partial_{q_i}\mathbb{P} \Big)\,.\end{aligned} \tag{2.6.14}$$

In this form, it is difficult to understand the various interactions between

the landmarks and the noise, but the finite dimensional approximation will bring more insights. The only thing that we can see here is the fact that the double bracket terms contain second and first order derivatives of the probability density $\mathbb{P}$, as expected for a diffusion term.

Moments of the Fokker-Planck equation

From a practical point of view, the interesting quantities to observe in landmark dynamics are the mean position and the covariance of each landmark. We define these two quantities as averages with respect to the probability density $\mathbb{P}$, that is

$$\langle q_i^\alpha \rangle (t) := \int q_i^\alpha \mathbb{P}(\mathbf{q}, \mathbf{p}, t)\, d\mathbf{q}d\mathbf{p} \tag{2.6.15}$$

$$\langle q_i^\alpha q_j^\beta \rangle (t) := \int q_i^\alpha q_j^\beta \mathbb{P}(\mathbf{q}, \mathbf{p}, t)\, d\mathbf{q}d\mathbf{p}\,. \tag{2.6.16}$$

Other moments can be defined accordingly such as those involving the momentum variable, or higher order moments. The aim is now to find a dynamical equation for these quantities. We proceed by computing the time derivative of a moment. For example, we have

$$\begin{aligned}\frac{d}{dt}\langle q_i^\alpha \rangle (t) &= \int q_i^\alpha \mathcal{L}^* \mathbb{P}(\mathbf{q}, \mathbf{p}, t)\, d\mathbf{q}d\mathbf{p} \\ &= \int \mathcal{L} q_i^\alpha \mathbb{P}(\mathbf{q}, \mathbf{p}, t)\, d\mathbf{q}d\mathbf{p}\,,\end{aligned}$$

where in the second step we used the fact that the backward Kolmogorov operator $\mathcal{L}$ with the dual of $\mathcal{L}^*$ in the L^2 pairing. We thus need to compute $\mathcal{L}q_i^\alpha$ to obtain

$$\begin{aligned}\mathcal{L}q_i^\alpha &= -\{h, q_i^\alpha\}_{\text{can}} + \frac{1}{2}\sum_l \{\phi_l, \{\phi_l, q_i^\alpha\}_{\text{can}}\}_{\text{can}} \\ &= \frac{\partial h}{\partial p_i^\alpha} + \frac{1}{2}\frac{\partial \sigma_l^\alpha(\mathbf{q}_i)}{\partial q_i^\beta}\sigma_l^\beta(\mathbf{q}_i)\,,\end{aligned} \tag{2.6.17}$$

where we used the fact that the double bracket is symmetric and the advection term is anti-symmetric. Similarly, for the momentum evolution, we obtain

$$\begin{aligned}\mathcal{L}p_i^\alpha &= -\{h, p_i^\alpha\}_{\text{can}} + \frac{1}{2}\sum_l \{\phi_l, \{\phi_l, p_i^\alpha\}_{\text{can}}\}_{\text{can}} \\ &= -\frac{\partial h}{\partial q_i^\alpha} + \frac{1}{2} p_i^\gamma \frac{\partial \sigma_l^\gamma(\mathbf{q}_i)}{\partial q_i^\beta}\frac{\partial \sigma_l^\beta(\mathbf{q}_i)}{\partial q_i^\alpha} - \frac{1}{2} p_i^\beta \frac{\partial^2 \sigma_l^\beta(\mathbf{q}_i)}{\partial q_i^\alpha \partial q_i^\gamma}\sigma_l^\gamma(\mathbf{q}_i)\,.\end{aligned} \tag{2.6.18}$$

Finally, we need to compute the average of this quantity with respect to $\mathbb{P}$, which corresponds to taking expectation, that is $\langle \mathcal{L}q_i^\alpha \rangle$. A first order approximation of this quantity is to replace the expectation of products with the product of expectations as for example $\langle q_i^\alpha q_j^\beta \rangle \approx \langle q_i^\alpha \rangle \langle q_j^\beta \rangle$. In this case, the double bracket term will correspond to the Itô correction of the original Stratonovich stochastic process.

This problem of finding a good approximation of this expectation is difficult and must be solved case by case. Indeed, the natural method is to include higher order correlations up to a certain chosen order. Closing this hierarchy of higher order correction is the called the closure problem of the BBGKY hierarchy which mostly arises in many-body problems in quantum mechanics. We refer to Holm et al. (2007) and references therein for the application of these ideas in the context of the Vlasov equation with a similar double bracket dissipation term. We cannot use this geometrical method here for the following reason. This scheme relies on the approximation of the expectation of the Hamiltonian in terms of only a finite set of moments. In our case, the kernel function cannot be expended in a polynomial of the position with a small number of terms. Indeed, a truncation at any order will give unbounded approximation in the $\mathbf{q}$ variable which will result in a diverging interaction of far apart landmarks. We will thus use a less systematic method to derive the equation for the first moments where we can get rid of this problem.

Before going into the technical part of this approximation, we first look at

the dynamics of the second moment, that is

$$\mathcal{L}(q_i^\alpha q_j^\beta) = -q_i^\alpha \frac{\partial h}{\partial p_j^\beta} + \sum_l \sigma_l^\alpha(\mathbf{q}_i)\sigma_l^\beta(\mathbf{q}_j) + \frac{1}{2}\sum_l q_i^\alpha \sigma_l^\gamma(\mathbf{q}_j)\frac{\partial \sigma_l^\beta(\mathbf{q}_j)}{\partial q_j^\gamma} + t.\,, \tag{2.6.19}$$

where $t.$ stands for the transpose of the previous terms in the sense that $i \leftrightarrow j$ and $\alpha \leftrightarrow \beta$. This quantity is the complete covariance between each landmark, but it is more natural to consider the centred covariances

$$\Delta_2 \left\langle q_i^\alpha q_j^\beta \right\rangle := \left\langle q_i^\alpha q_j^\beta \right\rangle - \left\langle q_i^\alpha \right\rangle \left\langle q_j^\beta \right\rangle . \tag{2.6.20}$$

For $i = j$ they correspond to the covariance of the probability distribution for the landmark i, the quantity of interest for us. The time evolution of the centred covariance, or correlation, is directly computed as

$$\begin{aligned}
\frac{\partial}{\partial t}\Delta_2 \left\langle q_i^\alpha q_j^\beta \right\rangle &= \frac{\partial}{\partial t}\left\langle q_i^\alpha q_j^\beta \right\rangle - \left\langle q_i^\alpha \right\rangle \frac{\partial}{\partial t}\left\langle q_j^\beta \right\rangle + t. \\
&= \sum_l \left\langle \sigma_l^\alpha(\mathbf{q}_i)\sigma_l^\beta(\mathbf{q}_j) \right\rangle + \left\langle q_i^\alpha \frac{\partial h}{\partial p_j^\beta} \right\rangle - \left\langle q_i^\alpha \right\rangle \left\langle \frac{\partial h}{\partial p_j^\beta} \right\rangle \\
&+ \frac{1}{2}\sum_l \left(\left\langle q_i^\alpha \sigma_l^\gamma(\mathbf{q}_j)\frac{\partial \sigma_l^\beta(\mathbf{q}_j)}{\partial q_j^\gamma} \right\rangle - \left\langle q_i^\alpha \right\rangle \left\langle \sigma_l^\gamma(\mathbf{q}_j)\frac{\partial \sigma_l^\beta(\mathbf{q}_j)}{\partial q_j^\gamma} \right\rangle \right) + t..
\end{aligned}$$

This equation contains three different types of terms. The first term is just the square of the noise amplitude and will linearly increase the correlation in the direction of the noise whereas the other two terms account for respectively the interaction between landmarks and the interaction between landmarks and the gradient of the noise. To proceed further and approximate these equations we need to explain how we expand nonlinear terms using the cluster expansion method. This approximation is rather rough and does not converge to the true solution in any sense, but will only be used for numerical simulations. We will study later the bias induced by this approximation.

Before deriving the approximate equations of motion for the moments, we need to explain how to proceed in general. We will follow the so-called cluster expansion method used for example in quantum semiconductor physics as explained in Kira and Koch (2011). In an abstract form, the aim is to approximate the dynamics of the expected value of the product of N particles which we will write $\langle N \rangle$ in terms of lower order particle correlations. For this, we apply the cluster expansion, which consists in the decomposition

$$\langle 2 \rangle = \langle 2 \rangle_s + \Delta_2 \langle 2 \rangle \ , \tag{2.6.21}$$

where $\langle 2 \rangle_s = \langle 1 \rangle \langle 1 \rangle$. In term of quantum mechanics, $\langle 1 \rangle$ is a singlet and $\langle 2 \rangle$ a doublet, where $\Delta_2 \langle 2 \rangle$ is the doublet correlation. The decomposition of the triplet $\langle 3 \rangle$ is written

$$\langle 3 \rangle = \langle 3 \rangle_s + \langle 1 \rangle \Delta_2 \langle 2 \rangle + \Delta_3 \langle 3 \rangle \ , \tag{2.6.22}$$

where we introduced $\Delta_3 \langle 3 \rangle$ the triplet correlation. This decomposition can be done with any expectation of the product of particles to express these products only with singlets $\langle 1 \rangle$ and correlations up to a given order. Here, we will only consider the singlet and doublet correlations, so $\Delta_i i = 0$ for $i > 2$ and, for example, the term $\langle N \rangle$ expands as

$$\begin{aligned} \langle N \rangle = \langle N \rangle_s + \langle N-2 \rangle_s \Delta_2 \langle 2 \rangle + \langle N-4 \rangle_s \Delta_2^2 \langle 2 \rangle \\ + \sum_i \langle N - 2i \rangle_s \Delta_2^i \langle 2 \rangle + O(\Delta_3) \, . \end{aligned}$$

As an illustration to the landmark dynamics, we compute two approximations for doublets and triplets

$$\begin{aligned} \langle q_i^\alpha p_j^\beta \rangle &= \langle q_i^\alpha \rangle \langle p_j^\beta \rangle + \Delta_2 \langle q_i^\alpha p_j^\beta \rangle \\ \langle q_i^\alpha q_j^\beta p_k^\gamma \rangle &\approx \langle q_i^\alpha \rangle \langle q_j^\beta \rangle \langle p_k^\gamma \rangle + \langle q_i^\alpha \rangle \Delta_2 \langle q_j^\beta p_k^\gamma \rangle + \langle q_j^\beta \rangle \Delta_2 \langle q_i^\alpha p_k^\gamma \rangle + \langle p_k^\gamma \rangle \Delta_2 \langle q_i^\alpha q_j^\beta \rangle \ . \end{aligned}$$

In the sequel, we will use this method to expand the nonlinear terms and approximate their expected value. To tackle the problem of approximating

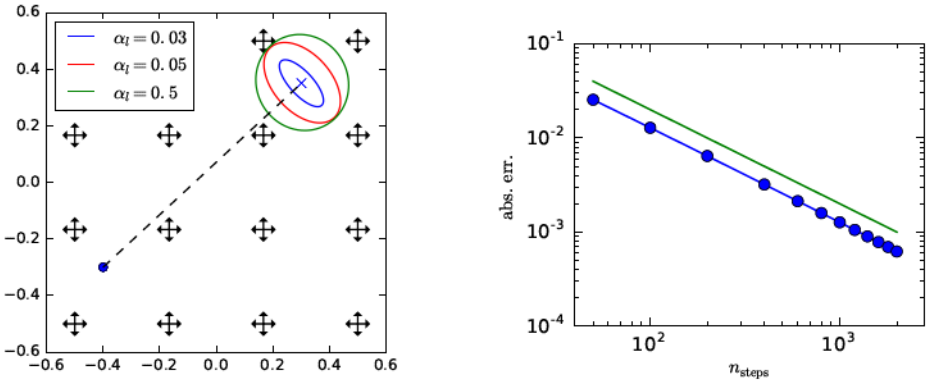
the kernel with a finite number of moments, we will approximate $\langle \sigma_l(\mathbf{q}_i) \rangle \approx \sigma_l(\langle \mathbf{q}_i \rangle)$, or $\langle K(\mathbf{q}_i - \mathbf{q}_j) \rangle \approx K(\langle \mathbf{q}_i \rangle - \langle \mathbf{q}_j \rangle)$. This approximation turns out to be good enough for this work. Higher order approximation will complicate the final equation a lot for only small improvements. We need to recall here that we are also making the choice of Gaussian kernels such that its derivative has a simple form. In appendix A we computed the approximate equations of motion for the singlets $\langle q_i^\alpha \rangle$ and $\langle p_i^\alpha \rangle$ in terms of doublet correlations.

2.6.3 Numerical examples

In this last section on stochastic landmark dynamics, we will use numerical simulations of the moment equations to study the effect of noise on the landmarks and determine the amplitude of the noise from a real set of data. We will only present parts of these results here and refer to Arnaudon et al. (2017d,e) for more detail and another method for determining the amplitude of the noise based on maximum likelihood expectation.

Solution of the Fokker-Planck equation

We begin with the simplest situation which consists of a single landmark subject to noise where the noise can by a square array of Gaussian kernels. In this setting, there is no interaction between landmarks and the deterministic dynamics is a straight line in the direction of the initial momentum of this landmark. We display such an experiment in figure 2.9 where we have two arrays of four by four noise fields with $\lambda_l = (0.08, 0)$ or $\lambda_l = (0, 0.08)$. We run this experiment three times and vary the parameter $\alpha_l = 0.5, 0.05, 0.03$. For large values of α_l, the noise is rather uniform, and the gradient of σ_l is small. The only term which contributes to the increase of the covariance of the landmark is the $\left\langle \sigma_l^\alpha(\mathbf{q}_i)\sigma_l^\beta(\mathbf{q}_j) \right\rangle$ term. This term comes from the noise in the $\mathbf{q}$ landmark equation and has the effect of only linearly increasing the covariance. The final variance is thus a circle as the noise has equal strength in both the x and y coordinates. When decreasing the parameter α_l, the gradient of σ_l becomes large enough for the other terms in the moment equation to have a non-negligible effect. This effect comes from the coupling

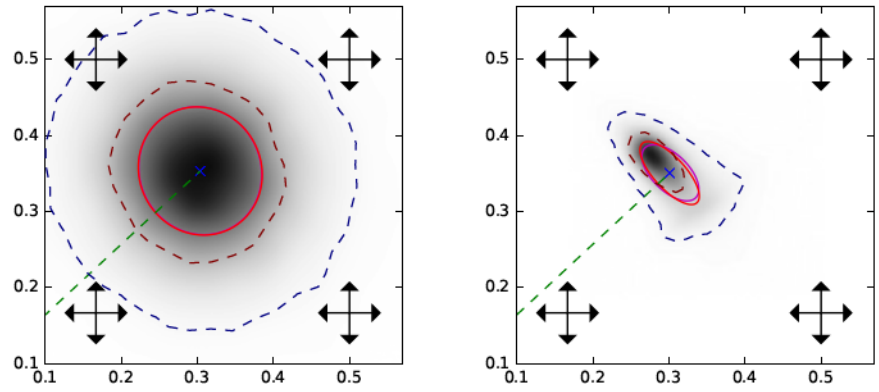


(a) Moment dynamics (b) Convergence of moments

Figure 2.9: On the left panel we display three simulations of a single landmark dynamics in an array of Gaussian noise fields with parameters $\lambda_l = (0.08, 0)$ or $\lambda_l = (0, 0.08)$. We vary the length-scale of the noise fields to see the effect of small or large Gaussian fields σ_l on the covariance and mean path of the landmark. We used 5000 timesteps to integrate the moment equation forwards to $t = 1$. The initial momentum is found by the deterministic shooting method with no Itô correction to match a particular final position. On the right panel, we display the numerical convergence (of order one) of the final moment of this experiment with $\alpha_l = 0.05$ with respect to an exact solution with $0.5e6$ timesteps.

in the $\mathbf{p}$ equation between the momentum of the landmark and the gradient of σ_l. This effect can be seen on figure 2.9 where the final covariance has a large value in the direction of the gradient of σ_l. Indeed, recall that this coupling is of the form $\frac{\partial}{\partial q_i}(\sigma_l(\mathbf{q}_i) \cdot \mathbf{p}_i)$, thus in the direction of the gradient, not the momenta.

We now want to assess the validity of the approximations we used in the derivation of the moment equation. We take the previous simulation with $\sigma_l = 0.03$ and compare the final covariance computed from the moment equation with the one estimated from a Monte-Carlo simulation with 50 000 stochastic paths. We display this result on the right panel of figure 2.10 which shows a good agreement between them but with some noticeable differences. We can compare this result to the same simulation with $\sigma_l = 0.5$ on the left panel which shows a small error (the red and magenta curves are on top of each other). This test shows the limitation of the current form of the moment equation which only considers the correlations up to second order.

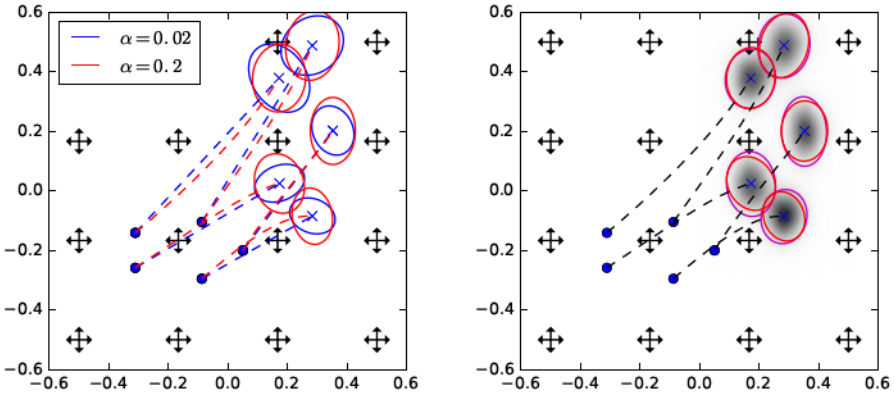


(a) $\alpha_l = 0.5$ (b) $\alpha_l = 0.03$

Figure 2.10: We display in this figure two comparisons between the estimation of the final covariance from a Monte-Carlo sampling and from the solution of the moment equation. The parameters are the same as in figure 2.9 but with different values of α. The black density represents the probability distribution of the landmark and the dashed lines a few level sets. The magenta ellipse is the estimation of the covariance from the sampling and the red ellipse from the moment equation.

This particular test on the right panel of (2.10) clearly shows a strong third moment or skewness in its final distribution. Although we cannot capture this effect with the moment equation, the estimation for the second moment remains remarkably reliable. It is in principle possible to extend the moment equation to include skewness effects such as this one, but it will require more terms in the already long equation. Indeed, this case is rather extreme, and we expect the noise to be in practice rather uniform throughout an image.

We perform another experiment with 5 landmarks arranged in an ellipse and matching another ellipse displayed in figure 2.11. The noise is the same as before, with $\alpha_l = 0.2$ but we now run it twice and change the landmark interaction scale α. For $\alpha = 0.02$, the landmarks almost do not interact, but for the second simulation with $\alpha = 0.2$, the landmark interactions become non-negligible. To only study this effect on the covariance, we performed the shooting again to land at the same final position. Of course, the trajectories are slightly different, and the landmarks will be subject to a different noise, but the overall effect is to reduce the differences between the covariances of



(a) Landmark interaction (b) Sampling comparison

Figure 2.11: This figures present a similar study as we did in figure 2.9 and 2.10 but with 4 landmarks. On the left panel, we illustrate the effect of varying the landmark length-scale α, and on the right panel, we compare a simulation with $\alpha = 0.1$ and a Monte-Carlo simulation. The black density plot show the final probability density of the landmarks.

interacting landmarks. This effect is explained by the fact that the noise affecting a landmark will also affect another landmark indirectly by their mutual interaction, thus averaging their differences in covariance.

Estimating the noise amplitude

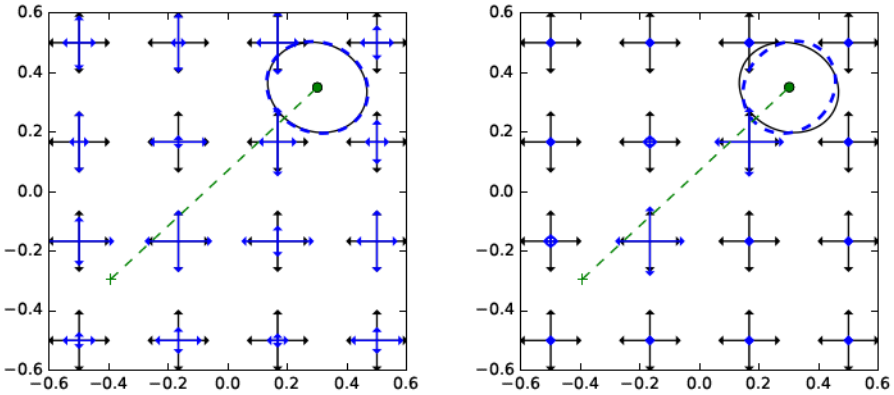
The main goal of having a finite dimensional approximation of the dynamics of the probability distribution of the landmarks is to be able to estimate the noise fields σ_l from a set of observed landmarks. This application would be useful in medical imaging to characterise the variability among images of similar patients. We will here only briefly show how to achieve this and refer to Arnaudon et al. (2017d,e) for more detail.

The algorithm is rather simple and corresponds to an extension of the shooting method for deterministic landmarks, but instead trying to find the initial momentum and the parameters λ_l of the noise fields such that the following

cost functional is minimised

$$C(\langle \mathbf{p} \rangle (0), \lambda_l) = \frac{1}{\gamma_1} \left\| \langle \mathbf{q} \rangle_T - \langle \mathbf{q} \rangle (1) \right\|^2 + \frac{1}{\gamma_2} \left\| \Delta_2 \langle \mathbf{qq} \rangle_T - \Delta_2 \langle \mathbf{qq} \rangle (1) \right\|^2 . \tag{2.6.23}$$

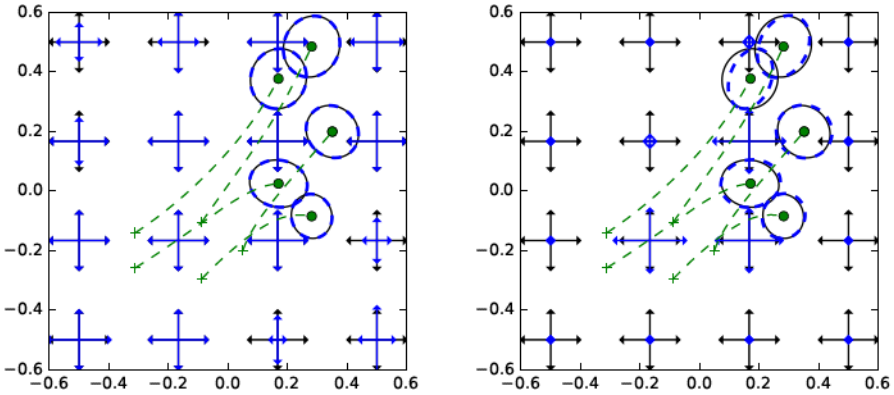
The first term is the usual matching for the mean final position and the second enforces the matching of the covariance. The parameters γ_1 and γ_2 serve as weights but, in practice,we do the matching in two steps by assuming that the mean matching does not depend on the noise.



(a) Genetic algorithm (b) Gradient descent

Figure 2.12: This figure shows two methods for estimating the σ_l fields from only the final position and final variance of the landmarks. The original σ_l are in black and the estimated ones in blue. The error in the final covariance for the genetic algorithm is of the order of 10^{-10} and for the BFGS algorithm is of the order of $5 \cdot 10^{-2}$.

The minimization of this function is a difficult task and standard usual gradient descent algorithms such as the BFGS algorithm do not perform well compared to genetic algorithms. We used here a genetic algorithm called the differential evolution method introduced by Storn and Price (1997) which is supposed to be able to find the global minimum of a function. We will not describe this method in detail here, as it is not interesting for this presentation. We will only compare these two algorithms in the test cases that we studied previously. In figure 2.12 we show the result for the single landmark experiment and in figure 2.13 for the ellipse matching experiment. The



(a) Genetic algorithm (b) Gradient descent

Figure 2.13: Same as figure 2.12 but for the ellipse experiment. The error in the final covariance for the differential equation genetic algorithm is of the order of 10^{-9} and for the BFGS algorithm it is of the order of $5 \cdot 10^{-3}$.

BFGS algorithm uses the first gradients of the cost C in (2.6.23) computed symbolically using the Theano package of Theano Development Team (2016) but has difficulties to update noise fields far away from the trajectory of the landmark. This demonstrates that this cost function has many local minima and its gradient with respect to the noise is sometimes too small for the scheme to be able to update the parameters of these noise fields. The genetic algorithm is able almost to recover the correct parameters of the noise, provided that we adjust the parameters of the genetic algorithm.

2.6.4 Summary and outlook

In this section, we used an important property of infinite dimensional integrable systems which is that they contain solitons. The Camassa-Holm equation has the additional property that its solitons, called peakon, can be entirely described by a finite dimensional Hamiltonian system. The reduction from the full infinite dimensional PDE to the landmark or peakon equation is important for many applications and in particular for medical imaging. It heavily relies on the existence of a singular momentum map which maps the

u field of the PDE to the position and momentum of the peakons. We did not investigate this aspect of the coadjoint motion in the previous section, but the noise which we used to deform the Euler-Poincaré equation preserves this momentum map. In fact, this map is intimately connected to the geometrical structure of the equation, which is preserved by the noise. Remarkably, the existence of such a solution is related to the complete integrability of the original system. In the case of the Euler-Poincaré equation on the diffeomorphism group which encompasses the Camassa-Holm equation, these solutions exist even without complete integrability. We can modify the shape of the solutions from a peaked profile to a Gaussian profile and consider the dynamics of these landmarks on the plane, not just on the real line as for the CH equation.

As expected, the Fokker-Planck equation is also of the double bracket from, but instead of a Lie-Poisson bracket, it is formed out of the canonical Poisson bracket. The explicit form of the Fokker-Planck equation is too involved to study it analytically and is too high dimensional to study it numerically. We thus took a different approach and approximated it with a finite dimensional system of equations describing the time evolution of its first moments. For this, we were inspired by a cluster-expansion method often used in quantum mechanics for the description of the interaction of many quantum particles. Here, the interactions correspond to the noise and the double bracket term in the Fokker-Planck equation. The aim of the derivation of this approximate equations is not only to study the effect of the noise on the error in the propagation of landmarks but also to recover the amount of noise needed to reproduce a given final error in observations. The finite dimensional approximation allows us to have a fast and precise algorithm for such studies. We tested it for simple test cases and will apply it to real data in future works; some test were already done in Arnaudon et al. (2017d).

Several open problems and extensions of this theory exist. For the approximation of the Fokker-Planck equation, we can use a higher order approximation or change the shape of the kernels. It is also possible to implement the reduction from the infinite dimensional equation to other data structures, such as curves, surfaces or vector fields. This noise will work for these

structures as well, but other methods must be implemented to derive the noise amplitudes as we did here. Finally, we only considered exact matching but inexact matching can also be done and the noise will preserve the corresponding momentum map, see Arnaudon et al. (2017d) for more details.

3

Infinite dimensional systems

3.1 Introduction

Most of the previous chapter was devoted to the stochastic deformations of finite dimensional dynamical systems with symmetries. This chapter has more variations around the theme of deformations of integrable systems even if its core consists of the Sobolev deformation of hierarchies of integrable systems. It was known for a long time that integrable partial differential equations come in hierarchies and that this structure is the basis for understanding the various properties of integrability. Indeed, this hierarchy structure provides us with infinitely many conserved quantities and an isospectral problem used to derive exact soliton solutions.

As in the first chapter, our aim is to deform these equations while preserving their geometric structure. We will thus in a first step, describe integrable hierarchies in a way that suits this goal, that is by interpreting the entire hierarchies as reduced dynamical systems in the Lagrangian sense. This procedure is an attempt to merge two different descriptions of integrable partial differential equations. The first interpretation is the one we just presented, tied to the notion of integrability and the inverse scattering transform. The

second is the reduction by symmetry on the diffeomorphism group, as we already mentioned in section 2.6 in the context of image matching. This description means that the integrable equations are geodesic motions on a particular group of diffeomorphisms. The equation of motion is thus the corresponding Euler-Poincaré equation obtained after reduction by symmetry. This derivation works perfectly for equations such as KdV or the Camassa-Holm equation but does not work for the NLS equation for example. By merging these two approaches, we mean that we will derive each hierarchy as an Euler-Poincaré system on a particular infinite dimensional Lie group. This derivation is inspired by numerous works, including Flaschka et al. (1983), Semenov-Tyan-Shanskiĭ (1983) and the book Newell (1985) where a Hamiltonian derivation was investigated. We will derive the Lagrangian counterpart of this derivation which will be used to implement a Sobolev deformation. The new feature of this derivation is to be able to directly obtain the Sobolev deformation of well-known equations such as the KdV or the modified KdV. This deformation recovers equations such that the CH or modified CH equation as well as a new deformation of the NLS equation which we will call the CH-NLS equation and will study in some detail. Unfortunately, this equation seems not to be integrable and raises more questions than we can answer.

This Lagrangian interpretation of integrable hierarchies will also be used to implement a stochastic deformation as we did in the first chapter for finite dimensional systems. We will see that to retain a notion of stochastic integrability we will need a rather specific type of noise resulting in stochastic linear combinations of flows of an integrable hierarchy. For these equations, it is possible to use the inverse scattering transform to find exact stochastic solutions. More investigations are possible regarding their statistical properties or the notion of stochastic stability as we did in section 2.4 with the energy-Casimir method.

PLAN OF THIS CHAPTER This second chapter is structured as follows. In section 3.2, we will begin by the Lagrangian reduction of the integrable hierarchies and implement the deformation with the Sobolev norm in subsection 3.2.5. The deformation of the NLS equation will be studied in section 3.3

with the help of numerical simulations where we will mainly focus on the behaviour of its solitary waves. In the next section 3.4, we will use the Lagrangian reduction to introduce the stochastic deformation and study the stochastic inverse scattering transform. Finally, we will end this chapter in section 3.6 where we will apply a different type of deformation on the AKNS hierarchy. Indeed, we will consider the first negative flow of the AKNS hierarchy and move the poles of the spectral problem to produce new integrable systems depending on a deformation parameter. We will study these systems in some details, and show that two of them have applications in physics.

3.2 Lagrangian reduction and Sobolev deformation of integrable hierarchies

To implement various geometric deformations in the theory of infinite dimensional integrable systems, we need a geometric description of them that will allow us to deform entire hierarchies at once. The geometry of integrable equations such as the KdV equation or the CH equation of Camassa and Holm (1993) can be understood as geodesic motion equations on some particular infinite dimensional groups. They are the groups of diffeomorphisms endowed with various metrics and group actions. Although this geometrical viewpoint led to significant geometrical insights about these equations, it is not directly related to the concept of complete integrability and will therefore not be used in this section. This approach based on a diffeomorphism group uses the Euler-Poincaré equation whereas the standard AKNS scheme does not use any such geometrical derivation. One of the aims of this Lagrangian reduction for integrable hierarchies is to merge these two approaches, that is to use the Euler-Poincaré equation to derive integrable hierarchies. We will briefly use this Euler-Poincaré equation on the diffeomorphism group later in section 3.5 to introduce noise compatible with this structure without demanding integrability. We refer here to Khesin and Misioek (2003), Khesin and Wendt (2008), Marsden et al. (2007), Marsden and Ratiu (1999), Misiołek (1998) for more detail on this approach. In this setting of Euler-Poincaré equations on diffeomorphism groups, introducing the Sobolev H^1 norm for the geodesic flow of the KdV equation yields the CH equation as shown in Misiołek (1998) but nothing is known about the other equations in the AKNS hierarchy.

To obtain a more general deformation scheme of this type for integrable hierarchies, we need another geometrical interpretation of integrable equations taking into account their complete integrability. Such a derivation will be performed by using loop groups. This group is an infinite dimensional group different from the groups of diffeomorphisms and will be used to implement a theorem of reduction by symmetry, both in the Hamiltonian and Lagrangian settings. The Hamiltonian derivation was known for a long time

and first derived by Reyman and Semenov-Tian-Shansky (1988), Semenov-Tyan-Shanskiĭ (1983) with the R-matrix. We will here avoid the explicit use of R-matrices to apply the interpretation of Flaschka et al. (1983). This classical Hamiltonian description of integrable systems is the most logical as the conserved quantities and the bi-Hamiltonian structure of Magri (1978) are immediately visible. We will present here an alternative Lagrangian derivation which will be more convenient for implementing deformations with the Sobolev norm. This will be done later in section 3.2.5 and stochastic deformations will be introduced in section 3.4 using this same geometrical description.

This section and the next are based on Arnaudon (2016b) and are structured as follows. We will first describe the mathematical tools in 3.2.1 such as loop groups and central extensions with cocycles. In the second subsection 3.2.2, we will give and prove the reduction theorems in both the Hamiltonian and Lagrangian context. Finally, the last subsection 3.2.3 will be about the derivation of the AKNS hierarchy.

3.2.1 Preliminaries

In this subsection, we will review various mathematical objects needed later in this chapter.

Loop groups

First, recall that a definition of an integrable system is that we can find exact solutions for it. In the case of infinite dimensional integrable systems, one of the best methods for computing solutions is the inverse scattering transform (IST). This method relies on a spectral problem depending on a spectral parameter $\lambda \in \mathbb{C}$ which is a linearization of the nonlinear equation, simpler to solve and still equivalent to the integrable system. We refer to Ablowitz and Segur (1981), Ablowitz et al. (1973b), Newell (1985) and references therein for more detail on this method. We will not describe the IST method in detail here, but we will extensively use the spectral parameter λ in the context of loop groups. We will use basic tools of the IST in later sections such as 3.4

and 3.6. We refer to the monographs Khesin and Wendt (2008), Pressley and Segal (1986) for a complete account of the theory of loop groups as we will only discuss the necessary definitions and results for this work.

Definition 3.2.1 (Loop group). *Given a Lie group G, the corresponding loop group is*

$$LG := \mathrm{Map}(S^1, G)\,,$$

where $S^1 \in \mathbb{C}$ is the unit circle.

This definition is rather general, and no assumption is made on the regularity of this map. We will here only consider that these maps are Laurent polynomials in λ with positive and negative powers of λ. Recall that we can view $S^1 \subset \mathbb{C}$, hence elements of a polynomial loop group can be seen as Laurent polynomials taking values in a Lie group G.

Definition 3.2.2 (Polynomial loop group). *There is a dense subset of LG called a polynomial loop group and denoted $\widetilde{LG}$ with elements of the form*

$$g(\lambda) = \sum_{i=a}^{b} \lambda^i g_i \in G, \qquad g_i \in G,\ \forall i \quad \text{and} \quad a \in]-\infty, \infty[,\ b > a\,. \quad (3.2.1)$$

Notice that a cannot be $-\infty$ to ensure that the inverse of $g(\lambda)$ exists. These Laurent polynomials have a pole at 0 in the complex plane but loop group elements can have poles at other positions. We will stick to the simple case here and in the last section 3.6 we will study another equation called the reduced Maxwell-Bloch equation which will require poles at other positions.

Recall that given a Lie group G, one can derive its Lie algebra $\mathfrak{g}$ by looking at the tangent space of G at the identity element $e \in G$ that is $\mathfrak{g} = T_eG$. In the case of loop groups, the identity element is the constant $e(\lambda) = e \in G$ and the Lie algebra is a loop algebra.

Definition 3.2.3 (Loop algebra). *Given the Lie algebra $\mathfrak{g}$ of the Lie group*

G, the loop algebra corresponding to the loop group LG is

$$L\mathfrak{g} := \mathrm{Map}(S^1, \mathfrak{g})\,.$$

The corresponding polynomial loop algebra $\widetilde{L\mathfrak{g}}$ has elements of the form

$$\xi(\lambda) = \sum_{i=a}^{b} \lambda^i \xi_i \in \mathfrak{g}, \qquad \xi_i \in \mathfrak{g},\ \forall i \quad \text{and} \quad a \in]-\infty, \infty[,\ b > a\,. \tag{3.2.2}$$

To obtain dynamical equations from the theory of reduction, we need to identify the dual of the loop algebra. For simplicity we will assume that G is semisimple, that is, the Lie algebra admits a non-degenerate bilinear pairing, often called the Killing form and denoted $\langle \cdot, \cdot \rangle : \mathfrak{g} \times \mathfrak{g}^* \to \mathbb{R}$. This pairing is extended to the loop algebra as

$$\begin{aligned} \sum_{i,j} \langle \xi_i \lambda^i, \eta_j \lambda^j \rangle_{\widetilde{\mathfrak{g}}} &:= \sum_{i,j} \frac{1}{2\pi i} \int_{S^1} \langle \xi_i, \eta_j \rangle \lambda^{i+j} d\lambda \\ &= \sum_{i,j} \langle \xi_i, \eta_j \rangle \delta_{i,-1-j}\,, \end{aligned} \tag{3.2.3}$$

where we used Cauchy residue theorem to simplify the form of this pairing.

Central extension and cocycle

We already mentioned a possible extension of the Euler-Poincaré equation to include a semi-direct product describing symmetry breaking. In this section, we will use another extension of the Euler-Poincaré equation which is called a central extension. A central extension is another type of geometric structure for systems with symmetries which relies on a so-called cocycle to describe the group action in its centre. Similar extensions have applications in physical systems such as complex fluids used to describe internal dynamics of the fluid molecules. We refer to Gay-Balmaz and Ratiu (2009) for a thorough treatment of this theory for complex fluids and Khesin and Wendt (2008), Marsden et al. (2007) for the general mathematical theory. In the present case, the type of central extension that we will be using is different

from the extension used for complex fluid and is, in fact, simpler. Indeed, for complex fluids, the cocycle term is connected to a semi-direct product structure whereas it is only a central extension here.

A central extension consists of extending the Lie group G by a vector space V with a direct product. In general, V can be any vector space, but we will use here $V = \mathbb{R}$. We will denote the new group by $G^c := G \times V$ and we need to define its group multiplication. For this purpose, we introduce a symmetric map $B(g,h) : G \times G \to \mathbb{R}$ called a group 2-cocycle if it satisfies the following cocycle identity

$$B(f,g) + B(fg,h) = B(f,gh) + B(g,h), \quad \forall f,g,h \in G\,. \tag{3.2.4}$$

This identity ensures that the following group multiplication is associative

$$(g,a)\cdot(h,b) = (gh, a+b+B(g,h))\,. \tag{3.2.5}$$

We will also need the formula for the inverse of the group element given by

$$(g,0)^{-1} = (g^{-1}, -B(g^{-1},g))\,, \tag{3.2.6}$$

and its action on tangent vectors

$$(g,0)^{-1}(\dot{g},0) = (g^{-1}\dot{g}, -D_2B(g^{-1},\dot{g}))\,, \tag{3.2.7}$$

where D_2B stands for the in the second slot of the group cocycle B. We now need to compute the corresponding structures on the level of the Lie algebra.

The underlying vector space of the Lie algebra $\mathfrak{g}^c$ of G^c is $\mathfrak{g} \times \mathbb{R}$, and the adjoint action is

$$\mathrm{ad}_{(\xi,a)}(\eta,b) = [(\xi,a),(\eta,b)] = ([\xi,\eta], c(\xi,\eta)), \quad \forall \xi,\eta \in \mathfrak{g}, a,b \in \mathbb{R}\,, \tag{3.2.8}$$

where $c : \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}$ is the corresponding Lie algebra cocycle computed from differentiating the group cocycle B in both entries. a and b do not appear on the right-hand side because the vector space $\mathbb{R}$ is an abelian group with a commutative Lie algebra. Using the extending dual pairing of $\mathfrak{g}^c$, that is

$\langle(\xi, a), (\eta, b)\rangle_{\mathfrak{g}^c} = \langle\xi, \eta\rangle_{\mathfrak{g}} + ab$, the coadjoint action is

$$\mathrm{ad}^*_{(\xi,a)}(\mu, m) = (\mathrm{ad}^*_\xi \mu + mc(\xi, \cdot), 0), \qquad \forall (\mu, m) \in (\mathfrak{g}^c)^* \,, \tag{3.2.9}$$

where $c(\xi, \cdot) \in \mathfrak{g}^*$ is viewed as a linear functional on the Lie algebra $\mathfrak{g}$, thus in its dual. The fact that that the coadjoint action does not affect the centre of the dual of the Lie algebra is a direct consequence of the centrality of the group extension.

So far we have not specified the form of the cocycle, and it is time to do so. First, recall that the main topic of this section is about integrable 1+1 partial differential equations, so we need to introduce the space variable $x \in \mathbb{R}$. We do this by letting the Lie group of Lie algebra elements depend on x. The cocycle that we will need must involve the space variable, and it does with the help of partial derivatives in space. At the group level, the cocycle is

$$B(g, h) := \int \langle g, \partial_x h\rangle_G \, dx \,, \tag{3.2.10}$$

and, for the Lie algebra it descends to

$$c(\xi, \eta) = \int_{\mathbb{R}} \langle \xi, \partial_x \eta\rangle_{\mathfrak{g}} \, dx \,, \tag{3.2.11}$$

which satisfies the corresponding cocycle identity

$$c([\xi, \eta], \zeta) + c([\eta, \zeta], \xi) + c([\zeta, \xi], \eta) = 0 \,, \quad \forall \xi, \eta, \zeta \in \mathfrak{g} \,.$$

The pairing used for the group cocycle is not defined and is rather involved, see Khesin and Wendt (2008) for the details. This formula is in fact not entirely correct but will do the job in getting the equations of motion and understanding its derivation. A couple of other remarks are needed here. First, we will not address the regularity in the space variable and only consider the equations formally. Second, all the integrals and boundary conditions are such that integration by part can be done without boundary terms. This can be achieved by either assuming vanishing boundary conditions at infinity or by considering a periodic domain.

3.2.2 Reduction theorems

In this setting of centrally extended Lie groups, it is possible to apply the theory of reduction by symmetry in both the Hamiltonian and the Lagrangian framework. We first state the Hamiltonian theorem of reduction, see Marsden et al. (2007) for more detail. We will denote by $L(x,t) \in \mathfrak{g}^*$ the momentum variable which will be time and space dependent.

Theorem 3.2.1 (Lie-Poisson). *Let $\mathfrak{g}^c$ be the central extension of the Lie algebra $\mathfrak{g}$ with cocycle $c : \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}$ and $h : (\mathfrak{g}^c)^* \to \mathbb{R}$ be the Hamiltonian function. The Lie-Poisson bracket is*

$$\{F,G\}((L,a)) = \left\langle L, \left[\frac{\delta F}{\delta L}, \frac{\delta G}{\delta L}\right]\right\rangle + a\, c\left(\frac{\delta F}{\delta L}, \frac{\delta G}{\delta L}\right), \tag{3.2.12}$$

and the corresponding Lie-Poisson equation is

$$\partial_t L = \mathrm{ad}^*_{\frac{\delta h}{\delta L}} L + a\, c\left(\frac{\delta h}{\delta L}, \cdot\right). \tag{3.2.13}$$

Proof. The proof of these formulae is straightforward from the definition of the Lie-Poisson bracket and the adjoint and coadjoint action of the centrally extended Lie algebra defined in (3.2.8) and (3.2.9). ■

The Lagrangian counterpart of this theorem is the derivation of the Euler-Poincaré equation from a variational principle. This gives the following theorem.

Theorem 3.2.2 (Euler-Poincaré). *The following statements are equivalent:*

(1) Hamilton's variational principle with G-invariant Lagrangian $\mathscr{L} = \int \widetilde{\mathscr{L}} dx : G \to \mathbb{R}$ holds on G^c for arbitrary variations δg vanishing at the endpoints

$$\delta \int \mathscr{L}(g(t,x), \dot{g}(t,x))dt = \delta \iint \widetilde{\mathscr{L}}(g(t,x), \dot{g}(t,x))dxdt = 0\,;$$

(2) $g(t)$ satisfies the Euler-Lagrange equations on G;

(3) the constrained variational principle

$$\delta \int l(M(t,x))dt = \delta \iint \ell(M(t,x))dxdt = 0$$

holds on $\mathfrak{g}^c$*, using variations of the form*

$$\delta(M,0) = (\dot{\eta} + [M,\eta], c(\eta, M)) \ , \tag{3.2.14}$$

for arbitrary η *vanishing at the endpoints;*

(4) the Euler-Poincaré equation with central extension holds

$$\frac{\partial}{\partial t}\frac{\delta \ell}{\delta M} = \mathrm{ad}^*_M \frac{\delta \ell}{\delta M} + c(M, \cdot) \, . \tag{3.2.15}$$

Proof. The proof that (1) and (2) are equivalent is standard and comes from the general theory of Hamilton's principle. For (3), the form of the reduced Lagrangian l is straightforward to obtain from the invariance of L under the group action, so only the reduced variations must be computed. With the formula (3.2.7), we have a left trivialised generic element as

$$(M,0) = (g,0)^{-1}(\dot{g},0) = \left(g^{-1}\dot{g}, -\int g^{-1}\partial_x \dot{g} dx \right) ,$$

and its variation is

$$\delta(M,0) = \delta(g^{-1}(\dot{g},0)) = \left(\delta(g^{-1}\dot{g}), -\int \delta(g^{-1}\partial_x \dot{g}) dx \right) .$$

The first slot will give the standard constraint variations $\delta M = \dot{\eta} + [M, \eta]$ for arbitrary elements $\eta = g^{-1}\delta g$. We then compute the second slot

$$\begin{aligned} -\int \delta(g^{-1}\partial_x \dot{g}) dx &= \int \left[-g^{-1}\delta g \partial_x (g^{-1}\dot{g}) + g^{-1}\delta g g^{-1}\partial_x g g^{-1}\dot{g} - g^{-1}\partial_x g g^{-1}\delta \dot{g} \right] dx \\ &= \int \left[\eta \partial_x M + \eta g^{-1}\partial_x g M - g^{-1}\partial_x g \dot{\eta} - g^{-1}\partial_x g M \eta \right] dx \, . \end{aligned}$$

By noticing that

$$0 = \frac{1}{2}\int \partial_x(g^{-1}g\dot{\eta})dx = \frac{1}{2}\int \partial_x g^{-1}g\dot{\eta} + g^{-1}\partial_x g\dot{\eta} + \partial_x\dot{\eta}dx = \int g^{-1}\partial_x g\dot{\eta} + \partial_x\dot{\eta}\,,$$

the second term in the previous calculation vanishes. The last two terms also vanish by using the relations

$$\begin{aligned} 0 &= \int \partial_x(\eta g^{-1}gM)dx = 2\eta g^{-1}\partial_x gM + \partial_x(\eta M) \\ \text{and} \qquad 0 &= \int \partial_x(g^{-1}gM\eta)dx = 2g^{-1}\partial_x gM\eta + \partial_x(M\eta)\,. \end{aligned}$$

This proves (3).

We can now derive the Euler-Poincaré equation with central extension by using these constraint variations in the variational principle to prove (4). We compute to obtain

$$\begin{aligned} \delta\int \ell(M)dt &= \iint \left(\left(\frac{\delta\ell}{\delta M}, 1\right), (\dot{\eta} + [M,\eta], c(\eta, M))\right) dt\,dx\,d\lambda \\ &= \iint \frac{\delta\ell}{\delta M}(\dot{\eta} + [M,\eta])\,dt\,dx + \int c(\eta, M)\,dt \\ &= \iint \left((-\partial_t + \mathrm{ad}^*_M)\frac{\delta\ell}{\delta M} + \partial_x M\right)\eta\,dt\,dx\,. \end{aligned}$$

We implicitly used the freedom in the form of the Lagrangian on the centre of $\mathfrak{g}^c$ and chose the kinetic term $\frac{1}{2}a^2$ with $a \in \mathbb{R}$. Then, because the coadjoint action does not affect the centre, its dynamics is trivial, that is $\dot{a} = 0$, and can select $a = 1$ to recover the Euler-Poincaré equations (3.2.15). ■

3.2.3 Reduction for hierarchies

In this section, we will show how to use the previous theory of reduction to derive integrable hierarchies such as the AKNS hierarchy. We will introduce the correct notion of space-time and in particular the proper loop algebra fields such that the Lie-Poisson or Euler-Poincaré equation with the central extension will become equivalent to an entire integrable hierarchy. We begin

by describing the notion of multi-times needed to associate a time variable for each flow of the hierarchy following the ideas of Date et al. (1983), see also Flaschka et al. (1983), Newell (1985).

We consider the multi-time space to be $N := \lim_{n\to\infty} \mathbb{R}^n$, that is the Euclidian space with countably infinite dimensions and with coordinates $\mathbf{a} = (a_1, a_2, \ldots)$. We are voluntarily using the variable a as we do not know at this point if the a are time or space variables. We also need the notation $N_{ij} = \text{span}(a_i, a_j)$ to denote a two dimensional sub-manifold of N, or a slice in the (i, j) direction. With this space-time manifold it is in principle possible to construct a covariant field theory in the sense of Castrillón López et al. (2001), Gotay et al. (1997) (see also Arnaudon et al. (2017f)). This is a way to write dynamical equations of Euler-Poincaré type for multiple space-time variables in a coordinate free geometrical setting. The biggest problem with this theory here is that the space-time manifold is infinite dimensional, and the dynamical equations have divergence operator with respect to these multiple time variables. This simple observation implies that it does not make sense to try to derive the equations of motion from the Euler-Poincaré equation. We will instead use another property of the fields which is necessary to implement the reduction by symmetry from $T\widetilde{G}$ to $\widetilde{\mathfrak{g}}$. It is the zero curvature relation of integrable systems. To show this property, we need to introduce more geometrical objects. In covariant field theories, a tangent vector is not only an element of a tangent bundle of a manifold but is an element of the so-called jet bundle. This structure encodes the space-time direction used to compute the tangent vector to say $\widetilde{G}$. An element of the jet bundle is a map $j^1 g : N \to T^*N \otimes T\widetilde{G}$ where $\otimes$ is the tensor product, $T\widetilde{G}$ is the usual tangent bundle of $\widetilde{G}$. More explicitly, we have $j^1 g = \partial_1 g da_1 + \partial_2 g da_2 + \ldots \cdots$ where $\partial_i g \in T\widetilde{G}$ for all i. The similar construction is done at the Lie algebra level, that is $j^1 \xi = \xi_1 da_1 + \xi_2 da_2 + \ldots$. The theory of reduction by symmetry says that the reduced fields satisfy a covariance Euler-Poincaré equation as well as a certain compatibility condition if the unreduced fields satisfy a covariant Euler-Lagrange equation. This compatibility condition only appears in field theory and must be imposed to be able to reconstruct the complete

dynamics. This condition is, for a map $\mathsf{M} : N \to T^* \otimes \widetilde{\mathfrak{g}}$, given by

$$d^{\mathsf{M}}\mathsf{M} := dM + [\mathsf{M}, \mathsf{M}] = 0\,. \tag{3.2.16}$$

The operation d^{M} is called the covariant exterior derivative with respect to M and is expressed in term of the usual exterior derivative d and a commutator.

For now, we pick a generic loop algebra element

$$\mathsf{M} = \sum_{i=1}^{\infty} M^{(i)} da_i\,, \tag{3.2.17}$$

where $M^{(i)} \in \widetilde{\mathfrak{g}}$ are N dependent loop algebra valued fields. The vanishing curvature condition (3.2.16) applied to (3.2.17) simplifies to

$$d^{\mathsf{M}}\mathsf{M} = \sum_{i,j=1}^{\infty} (\partial_{a_j} M^{(i)} - \partial_{a_i} M^{(j)} + [M^{(i)}, M^{(j)}]) da_i \wedge da_j = 0\,, \tag{3.2.18}$$

where ∂_{a_i} denotes the partial derivative with respect to a_i. For this equation to vanish, each coefficient of the two-form on N must vanish independently. This gives the infinite set of ZCRs

$$\partial_{a_j} M^{(i)} - \partial_{a_i} M^{(j)} + [M^{(i)}, M^{(j)}] = 0\,.$$

As it is now these equations are useless as they contain to much freedom. Recall that each $M^{(j)}$ is a polynomial in λ, with possibly an infinite number of powers, each of them with an associated Lie algebra $\mathfrak{g}$ element. We need to find a natural way to reduce the number of available fields to proceed further and obtain integrable hierarchies from the relation (3.2.18). The key observation is that the only special choice that we can make is $i \to \infty$ (or $j \to \infty$). By special choice, we mean one which is different from all the others and would give us an extra constraint on the fields. We thus begin with $i \to \infty$ and define its associated loop algebra fields

$$M^{(\infty)} := \sum_{i=1}^{\infty} M_i(\mathbf{a}) \lambda^{-i}\,, \tag{3.2.19}$$

It contains an infinite number of negative powers of λ, each of them associated to a field $M_i : N \to \mathfrak{g}$. From this field, we can define all the other fields $M^{(i)}$ as

$$M^{(i)} := P_+(\lambda^i M^{(\infty)}) \,, \tag{3.2.20}$$

where P_+ is the projection onto positive powers of λ. This construction is close to the R-matrix construction of Semenov-Tyan-Shanskiĭ (1983) or the construction of Flaschka et al. (1983). In fact, this operator is a projection to a subalgebra of the loop algebra, the other subalgebra being the strictly negative powers of λ. We have now expressed all the $M^{(i)}$ in term of only $M^{(\infty)}$, but they still contain an infinite number of $\mathfrak{g}$-valued fields. To constrain further this construction we use the ZCR (3.2.18) on the slice $N_{i\infty}$

$$\partial_{a_i} M^{(\infty)} + [M^{(\infty)}, M^{(i)}] = 0 \,. \tag{3.2.21}$$

This formula gives an extra relation between $M^{(\infty)}$ and some of its fields M_i encoded in $M^{(i)}$. We will use it to reduce the freedom on the M_i and obtain integrable hierarchies. We first prove this formula as a particular case of the next proposition.

Proposition 3.2.1. *Equation* (3.2.18) *holds on every slice and in particular on $N_{i\infty}$, where $a_\infty := \lim_{j\to\infty} a_j$. This is encoded in the formulae*

$$\lim_{j\to\infty} d^{\mathsf{M}}\mathsf{M}\left(w, \frac{\partial}{\partial a_j}\right) = 0 \,, \tag{3.2.22}$$

where $d^{\mathsf{M}}\mathsf{M} : TN \times TN \to \widetilde{\mathfrak{g}}$ and any vector $w \in T_{\mathbf{a}}N$. In the case $w = \frac{\partial}{\partial a_i}$ one obtains (3.2.21).

Proof. We rewrite the ZCR using the definition of $M^{(j)}$ in the case of $w = \frac{\partial}{\partial a_i}$ to obtain

$$\frac{\partial}{\partial a_j} P_+(\lambda^i M^{(\infty)}) - \frac{\partial}{\partial a_i} P_+(\lambda^j M^{(\infty)}) + [P_+(\lambda^i M^{(\infty)}), P_+(\lambda^j M^{(\infty)})] = 0 \,. \tag{3.2.23}$$

Then, by noticing that

$$\lim_{j\to\infty} \lambda^{-j} P_+(\lambda^j M^{(\infty)}) = M^{(\infty)}\,,$$

multiplying (3.2.23) by λ^{-i}, taking the limit $j \to \infty$ and using

$$\lambda^{-j} P_+(\lambda^i M^{(\infty)}) \to 0\,,$$

we obtain the ZCR (3.2.21). ■

We are now ready to connect the ZCR (3.2.21) to the dynamical equations obtained after reduction by symmetry in the previous section. We first need to choose a particular integer k and compute $M^{(j)}(M^{(k)})$ for all $j \neq k$ using (3.2.21). This is possible as all the $M^{(j)}$ are contained in $M^{(\infty)}$, but in practice only the first few $M^{(j)}$ fields should be computed explicitly. For a given j, we can write the ZCR (3.2.19) on the slice N_{kj} and be left with two choices for the interpretation of the variables a_k and a_j. Depending on which one will be chosen as the time or the space variable, we will obtain different yet equivalent formulations of the integrable hierarchy.

Theorem 3.2.3. *Let the connection* M *satisfy the ZCR* (3.2.18). *For each slice* N_{ij}, *the ZCR* (3.2.21) *has a Hamiltonian or a Lagrangian formulation.*

1. *Hamiltonian formulation.*
   The space variable is $x := a_i$. *We define* $L := M^{(i)}$ *and use the ZCR* (3.2.21) *on the slice* $N_{i\infty}$ *to express every* M_j *as a function of the fields in* L. *For the time variable* $t := a_j$, *the ZCR* (3.2.21) *is a Lie-Poisson equation in the sense of theorem 3.2.1 with Hamiltonian*

$$h_j(L) = \frac{1}{2}\int \langle M^{(\infty)}(L), \lambda^{j-i} M^{(\infty)}(L)\rangle\, dx\,. \tag{3.2.24}$$

2. *Lagrangian formulation.*
   The time variable is $t := a_j$. *We define* $M := M^{(j)}$ *and use the ZCR* (3.2.21) *on the slice* $N_{j\infty}$ *to express every* M_i *as a function of the fields*

in M. For the space variable $x := a_i$, the ZCR (3.2.21) is an Euler-Poincaré equation in the sense of theorem 3.2.2 with Lagrangian

$$l_j(L) = \frac{1}{2}\int \langle M^{(\infty)}(M), \lambda^{j-i} M^{(\infty)}(M)\rangle \, dx \,. \tag{3.2.25}$$

Furthermore, these two interpretations are equivalent when written for the same space and time variable.

Proof. Using the formula (3.2.20) which relates $L = M^{(i)}$ with $M^{(\infty)}$, we directly obtain for the Hamiltonian interpretation

$$\begin{aligned}\frac{\delta h^j(M^{(\infty)})}{\delta M^{(i)}} &= P_+\left(\lambda^i \frac{\delta h^j(M^{(\infty)})}{\delta M^{(\infty)}}\right) \\ &= P_+\left(\lambda^{i+j} M^{(\infty)}\right) \\ &= M^{(i+j)} \,.\end{aligned}$$

The Lagrangian interpretation is the same, but with M instead of L. The equivalence between the two formulations is difficult to see at this level, but the Euler-Poincaré equation contains higher order terms in λ than the Hamiltonian equation (if $j > i$). These terms can be used to express $M(L)$ and obtain an equation for L only, as in the Hamiltonian derivation. ■

The scheme to derive the equations from the Lagrangian formalism is as follows. We first compute the Euler-Poincaré equation, where M contains the independent fields, and use the Legendre transform to rewrite this equation in term of a smaller number of independent fields. We want to mention that this derivation is different from the one of Flaschka et al. (1983) even if they have a similar Hamiltonian function. They did not use (3.2.21) before writing down the Hamiltonian and the system remains at an even more abstract level. Here, by using (3.2.21) we can not only disentangle the Lagrangian and the Hamiltonian formulation but also obtain a Lagrangian and Hamiltonian containing more information about the hierarchy than they have. Indeed, the Hamiltonians h_i can be evaluated explicitly and produce the conserved quantities of the integrable hierarchies for each value of i. The Lagrangians

are not supposed to correspond to conserved quantities, but upon computing the Legendre transform, or the relation $M(L)$, they can be obtained.

We now briefly show how to derive the AKNS hierarchy using the previous general derivation to then introduce the Sobolev deformation in section 3.2.5 and the stochastic deformation in section 3.4.

3.2.4 The AKNS hierarchy

The AKNS hierarchy is the most studied example of integrable hierarchies and is based on the Lie algebra $\mathfrak{sl}(2)$. It contains the well-known KdV and NLS equations in the slices with $x := a_1$. For instance, the flow with $t := a_2$ can be reduced to the NLS equation and the flow for $t := a_3$ to the KdV or modified KdV equation. We begin the derivation of this hierarchy by fixing our notation for the Lie algebra $\mathfrak{sl}(2)$. The standard basis is given by

$$e = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \qquad f = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \qquad h = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \tag{3.2.26}$$

with commutations relations

$$[e, f] = h, \qquad [e, h] = 2e, \qquad [f, h] = -2f\,. \tag{3.2.27}$$

For simplicity, we will use the notation $\xi^{\parallel}$ for the component of ξ in the Cartan subalgebra (along h) and $\xi^{\parallel}$ for its complement, that is

$$\xi^{\perp} := \xi - \langle \xi, h \rangle \frac{h}{\langle h, h \rangle}\,.$$

Hamiltonian derivation

We first derive the equations from the standard Hamiltonian formalism by following the scheme outlined in theorem 3.2.3. We choose the space variable

to be $x := a_1$ and set the L operator

$$L = \lambda M_0 + M_1 \,, \tag{3.2.28}$$

where $M_0, M_1 \in \mathfrak{sl}(2)$ are two arbitrary fields. We then use the ZCR (3.2.21) in the slice $N_{1\infty}$ to express the first few M_i and solve the following set of recursive relations

$$\begin{aligned}
\lambda^0 &: \partial_x M_0 = 0 \\
\lambda^{-1} &: \partial_x M_1 + [M_2, M_0] = 0 \\
\lambda^{-2} &: \partial_x M_2 + [M_2, M_1] + [M_3, M_0] = 0 \\
&\vdots \\
\lambda^{-i} &: \partial_x M_i + [M_i, M_1] + [M_{i+1}, M_0] = 0 \,.
\end{aligned} \tag{3.2.29}$$

From the equation proportional to λ^0, we deduce that M_0 is a constant. There are two possible choices for the form of M_0, either an element proportional to the Cartan subalgebra h or another element. We will use the first choice as it is the most common in the literature, but both are possible and give equivalent results. We refer to Newell (1985) for more detail on this related to the gradings of Kac-Moody algebras. We will denote $M_0 := A$ for $A = ih$ and proceed with the second equation $\partial_x M_1 + [A, M_2] = 0$. From the commutations relations of $\mathfrak{sl}(2)$, we deduce that M_1 has no component in the Cartan subalgebra and $M_2^\perp$ can then be computed using the relation $[A, [A, \xi]] = -4\xi^\perp$, written for an arbitrary ξ. We obtain

$$M_2^\perp = -\frac{1}{4}[A, \partial_x M_1^\perp] \,. \tag{3.2.30}$$

The parallel part of this field M_2 is found by projecting out the perpendicular part of the third equation to get

$$\partial_x M_2^\parallel = -[M_2^\perp, M_1^\perp] = \frac{1}{4}[[A, \partial_x M_1^\perp], M_1^\perp] = -\frac{1}{8}\partial_x [M_1^\perp, [A, M_1^\perp]] \,,$$

where we used the Jacobi identity. Integrating with vanishing boundary conditions yields

$$M_2^{\parallel} = -\frac{1}{8}[M_1^{\perp}, [A, M_1^{\perp}]]\,. \tag{3.2.31}$$

We proceed to compute the next field M_3 which is needed for the KdV equation. Using again the equation proportional to λ^{-2}, we can compute $M_3^{\perp}$ to get

$$M_3^{\perp} = -\frac{1}{4}[A, \partial_x M_2^{\perp}] - \frac{1}{4}[A, [M_2^{\parallel}, M_1^{\perp}]]\,.$$

This will be enough for our purposes and by denoting $U := M_1$, the ZCR on the slice N_{13} is the Lie-Poisson equation

$$\partial_t U^{\perp} + \frac{1}{4}\partial_x^3 U^{\perp} - \frac{1}{32}\partial_x[A, [U^{\perp}, [U, [A, U]]]] + \frac{1}{4}[U^{\perp}, [U^{\perp}, \partial_x U^{\perp}]] = 0\,. \tag{3.2.32}$$

It is not clear from this formulation but the last two terms are of the same form and can be combined. This equation demonstrates the rather unusual appearance of a Lie-Poisson system when written on a complicated symmetry group.

We now briefly show the various reductions of this system which yield well-known integrable equations. After rescaling time as $t \to 4t$, we have the following choices of matrix U and their corresponding integrable equations

$$U_{KdV} = \begin{bmatrix} 0 & u \\ 1 & 0 \end{bmatrix} \qquad \Rightarrow \qquad u_t + 6uu_x + u_{xxx} = 0\,, \tag{3.2.33}$$

$$U_{mKdV} = \begin{bmatrix} 0 & u \\ \sigma u & 0 \end{bmatrix} \qquad \Rightarrow \qquad u_t + 6\sigma u^2 u_x + u_{xxx} = 0\,, \tag{3.2.34}$$

$$\tag{3.2.35}$$

where $\sigma = \pm 1$ gives the focusing or defocusing mKdV. Another choice with real u and v off-diagonal fields would give a coupled modified KdV equation. The Lie-Poisson equation in the slice N_{12} yields the NLS equation for the

choice

$$U_{NLS} = \begin{bmatrix} 0 & u \\ \sigma\overline{u} & 0 \end{bmatrix} \qquad \Rightarrow \qquad iu_t + u_{xx} + \sigma u|u|^2 = 0\,. \tag{3.2.36}$$

This is, for this flow, the only non-trivial reduction of U as a non-complex reduction will make the right hand side of the Lie-Poisson equation zero. This is due to the required complex numbers introduced via the element $A = ih$ to allow even flows to exist.

Lagrangian derivation

We will give more detail on the computations for the Lagrangian derivation in the next section 3.2.5 and only explain here the principal differences from the previous derivation. We follow the scheme of theorem 3.2.3 and begin by fixing $t := a_3$ and the corresponding M operator

$$M = \lambda^3 A + \lambda^2 U + \lambda V + W\,. \tag{3.2.37}$$

In this case we have more independent fields than before and we need to solve the ZCR (3.2.21) to find $L := M_1$. Because the space variable has a smaller index than the time variable, it is just given by

$$L = \lambda A + U^{\perp}\,, \tag{3.2.38}$$

where only the perpendicular part of U appears. We can then write the Euler-Poincaré equation, or ZCR in the slice N_{13} explicitly in term of these fields to obtain three equations for the $\perp$ projection

$$\begin{aligned} \partial_t U^{\perp} - \partial_x W^{\perp} + [U^{\perp}, W^{\parallel}] &= 0 \\ -\partial_x V^{\perp} + [U^{\perp}, V^{\parallel}] + [A, W^{\perp}] &= 0 \\ -\partial_x U^{\perp} + [A, V^{\perp}] &= 0\,, \end{aligned} \tag{3.2.39}$$

and two equations for the remaining $\|$ components

$$\begin{aligned} -\partial_x W^{\|} + [U^{\perp}, W^{\perp}] &= 0 \\ -\partial_x V^{\|} + [U^{\perp}, V^{\perp}] &= 0\,. \end{aligned} \tag{3.2.40}$$

To see that this set of equations is equivalent to the equation (3.2.32) for U only, one just expresses V and W in terms of U and A and substitute them in the first equation. This computation is in effect the same as what did in the Hamiltonian derivation to obtain the equation (3.2.32) but different from this abstract point of view. Indeed, the computation of expressing V and W in term of U is a Legendre transformation from the Lagrangian variables M (the independent fields U, V, W) and the Hamiltonian variables L (with U as an independent field). This is the main point of this derivation which will allow us in the next subsection to use a Sobolev norm to deform the original integrable hierarchy.

3.2.5 The Sobolev deformation

We are now ready to introduce the first type of deformations of integrable hierarchies which consists of replacing the L^2 norm of the Lagrangian in theorem 3.2.3 by a Sobolev H^1 norm. This operation introduces a length scale parameter $\alpha \in \mathbb{R}$ which when set to 0, reduces the Sobolev norm to the L^2 norm. We will use the following Lagrangian defined on a loop algebra for the slice N_{ij} for $i < j$

$$\begin{aligned} l_{ij}^{H^1}(M^{(j)}) &:= \int \langle M^{(j)}, M^{(i)}(M^{(j)})\rangle + \alpha^2 \langle \partial_x M^{(j)}, \partial_x M^{(i)}(M^{(j)})\rangle dx \\ &= \int \langle M^{(j)}, (1 - \alpha^2 \partial_x^2) M^{(i)}(M^{(j)})\rangle dx\,. \end{aligned} \tag{3.2.41}$$

We will use symbol $\Lambda := 1 - \alpha^2 \partial_x^2$ to simplify the notation. The important point here is that the dependent variable is $M^{(j)}$ and not $L := M^i$, as for the Hamiltonian case. This has the effect of modifying the momentum L by

ΛM, or, in term of $M^{(i)}$ we have the following Euler-Poincaré equation

$$\partial_t \Lambda M^{(i)} = [M^{(j)}, \Lambda M^{(i)}] + \partial_x M^{(j)} \,. \tag{3.2.42}$$

We will later use this equation to derive the deformation of various flows of the AKNS hierarchy and obtain equations such as the well-known CH equation or the new CH-NLS equation studied in more detailed in section 3.3.1. Before that, we want to understand how this deformation affects the integrability of the original hierarchy while being compatible with its geometrical structure. It turns out that some of the equations obtained from (3.2.42) when expanded in term of the independent fields in $M^{(j)}$ cannot be solved. In fact, these same equations when $\alpha = 0$ are automatically satisfied and are not even used for the computation of the Legendre transformation. The only way we found for the Euler-Poincaré equation (3.2.42) to be equivalent to a ZCR and then to various integrable PDE is to introduce a particular projection operator.

Definition 3.2.4. *Given A, the element in the Cartan subalgebra of $\mathfrak{g}$ and k associated with a slice N_{1k}, we define the projection operator $\mathcal{P}^A_{1,k} : \widetilde{\mathfrak{g}} \to \widetilde{\mathfrak{g}}$ acting on polynomial loop algebras*

$$\mathcal{P}^A_{1,k}(Z) = Z - \lambda^k P_+(\lambda^{-k} Z) - \lambda^{k-1} P_+ \left(\lambda^{1-k} A \frac{\langle Z, A \rangle}{\langle A, A \rangle} \right) , \tag{3.2.43}$$

for an arbitrary $Z = \sum_{i=-\infty}^{k} Z_i \in \widetilde{\mathfrak{g}}$ and where P_+ stands for the projection on positive powers of λ.

Definition 3.2.5. *The Lie algebra valued two-form*

$$(\partial_t M^{(1)} - \partial_x M^{(k)} + [M^{(1)}, M^{(k)}]) ds \wedge dt$$

is said to be a weak ZCR if its projection under $\mathcal{P}^A_{1,k}$ vanishes, that is

$$\mathcal{P}^A_{1,k}(\partial_t M^{(1)} - \partial_x M^{(k)} +, [M^{(1)}, M^{(k)}]) = 0 \,, \tag{3.2.44}$$

or equivalently

$$\partial_t M^{(1)} - \partial_x M^{(k)} + \mathcal{P}^A_{1,k}([M^{(1)}, M^{(k)}]) = 0 \,. \tag{3.2.45}$$

If the equation (3.2.44) *is equivalent to a PDE, then this PDE is said to be weakly integrable.*

For the flows with a higher space variable, another projection operator should be introduced, but we will not investigate this further here. We could not obtain any interesting result regarding the interpretation of the weak ZCR, apart from being able to rewrite it in a form similar to the Manakov triple used for certain $2+1$ integrable equations, see the discussion in Arnaudon (2016b).

3.2.6 Deformed AKNS hierarchy

We now proceed with the derivation of the dynamical equation from the Euler-Poincaré equation (3.2.44). With the Sobolev operator, we can only do the derivation in the Lagrangian formalism. This will also illustrate how to derive the equation even with $\alpha = 0$. We begin with the KdV flow where $t := a_3$ and the corresponding M operator

$$M = \lambda^3 A + \lambda^2 U^\perp + \lambda V + W\,, \tag{3.2.46}$$

where A is a constant and $U^\perp, V, W$ are the independent fields. The condition on A and $U^\perp$ comes from the ZCR used to find $L = M^{(1)}$. We then apply the operator $\Lambda := 1 - \alpha^2\partial_x^2$ to L and define

$$\widehat{L} := \Lambda L = \lambda A + \Lambda U^\perp\,. \tag{3.2.47}$$

The Euler-Poincaré equation (3.2.44) expands to four equations for the $\perp$ part

$$\begin{aligned}
&\lambda^0 : \partial_t \Lambda U^\perp - \partial_x W^\perp + [\Lambda U^\perp, W^\parallel] = 0\\
&\lambda^1 : -\partial_x V^\perp + [\Lambda U^\perp, V^\parallel] + [A, W^\perp] = 0\\
&\lambda^2 : -\partial_x U^\perp + [A, V^\perp] = 0\\
&\lambda^3 : [A, \Lambda U^\perp] + [U^\perp, A] = 0\,,
\end{aligned} \tag{3.2.48}$$

and three for the $\|$ part

$$\begin{aligned}
&\lambda^0 : -\partial_x W^{\|} + [\Lambda U^{\perp}, W^{\perp}] = 0 \\
&\lambda^1 : -\partial_x V^{\|} + [\Lambda U^{\perp}, V^{\perp}] = 0 \\
&\lambda^2 : [\Lambda U^{\perp}, U^{\perp}] = 0\,.
\end{aligned} \tag{3.2.49}$$

The main difficulty with these equations is that the last equation of both systems are not solvable, unless $U_{xx} = 0$. This condition is too strong but can be removed by using the projection $\mathcal{P}^A_{13}$ defined by the equation (3.2.43). Such a projection exists only because the unsolvable terms are separated from the other equations when we write them for each power of λ. In a sense, this projection does not take much information out from the original ZCR, thus the name weak ZCR. These terms are nevertheless the main obstruction for the proof of complete integrability using this form of the spectral problem.

We now use the other equations (except the first) to express V and W in term of $U^{\perp}$. We find, using the λ^2 equation of (3.2.48) that

$$V^{\perp} = -\frac{1}{4}[A, \partial_x U^{\perp}]\,,$$

and using the λ equation of (3.2.49) that

$$\partial_x V^{\|} = [\Lambda U^{\perp}, V^{\perp}] = -\frac{1}{4}[\Lambda U^{\perp}, [A, \partial_x U^{\perp}]]\,.$$

This last equation is solved as

$$V^{\|} = -\frac{1}{4}\partial_x^{-1}([\Lambda U^{\perp}, [A, \partial_x U^{\perp}]])\,,$$

and similarly we obtain $W^{\perp}$ as

$$\begin{aligned}
W^{\perp} &= -\frac{1}{4}[A, \partial_x V^{\perp}] - \frac{1}{4}[A, [V^{\|}, \Lambda U^{\perp}]] \\
&= -\frac{1}{4}\partial_x^2 U^{\perp} + \frac{1}{16}[A, [\partial_x^{-1}([\Lambda U^{\perp}, [A, \partial_x U^{\perp}]]), \Lambda U]]\,.
\end{aligned}$$

The parallel part of W is

$$\begin{aligned}\partial_x W^{\parallel} &= [\Lambda U^{\perp}, W^{\perp}] \\ &= -\frac{1}{4}[\Lambda U^{\perp}, \partial_x^2 U^{\perp}] - \frac{1}{16}[\Lambda U^{\perp}, [A, [\partial_x^{-1}([\Lambda U^{\perp}, [A, \partial_x U^{\perp}]]), \Lambda U]]]\,.\end{aligned}$$

The last step is to use these relations in the equation of motion of $U^{\perp}$ to obtain a single dynamical equation for this system. Although these equations are non-local as we needed to invert the differential operator ∂_x, their reductions can be integrated as in the standard $\alpha = 0$ case. Similar computations can in principle be done up to every flows of the hierarchy, that is for loop algebra elements M of any power in λ. It turns out that only W remains local when we compute the Legendre transform as we just did. Other operators Λ could also be used, such as the H^2 operator $1 - \alpha^2\partial_x^2 + \beta^4\partial_x^4$, but the equation cannot be solved anymore. The H^1 Sobolev operator has this particular feature that other higher order Sobolev norms do not have which is that the first flow of the hierarchy can be deformed to yield interesting equations.

CH, mCH and CH-NLS equations

We proceed as before to derive the explicit form of the deformed equations for various reductions of the AKNS scheme. As before, we rescale the time $t \to 4t$. For the KdV reduction U_{KdV} of (3.2.33), we obtain the dispersive CH equation of Camassa and Holm (1993), Dullin et al. (2004), that is

$$m_t + 2mu_x + m_x u + u_{xxx} = 0\,. \tag{3.2.50}$$

Obtaining the CH equation is a good indication that the deformation of the hierarchy does what is meant to do as the CH equation becomes the KdV equation when $\alpha \to 0$. The difference with the standard proof of integrability of the CH equation is in the ZCR, which is now the weak ZCR with

$$L = \begin{bmatrix} \lambda & m \\ 1 & -\lambda \end{bmatrix} \quad \text{and} \quad M = \begin{bmatrix} \lambda^3 - \frac{1}{2}\lambda u - \frac{1}{4}u_x & \lambda^2\frac{1}{2}u_x - \frac{1}{2}mu + \frac{1}{4}u_{xx} \\ \lambda^2 - \frac{1}{2}u & -\lambda^3 + \frac{1}{2}\lambda u + \frac{1}{4}u_x \end{bmatrix}.$$

Notice that for this weak ZCR to correspond to the CH equation (3.2.50) we need to project out the term

$$\lambda^2([A,\Lambda U]+[U,A])+\lambda^3[\Lambda U,U]=\alpha^2 u_{xx}\begin{bmatrix}-\lambda^2 & 2\lambda^3\\ 0 & \lambda^2\end{bmatrix}.$$

For the mKdV reduction with U_{mKdV} defined in (3.2.34), we obtain the dispersive mCH equation

$$m_t+2\sigma[m(u^2-\alpha^2u_x^2)]_x+u_{xxx}=0\,. \tag{3.2.51}$$

The corresponding M operator is

$$M=\begin{bmatrix}\lambda^3-\frac{1}{2}\lambda(u^2-\alpha^2u_x^2) & \lambda^2u+\lambda\frac{1}{2}u_x-\frac{1}{2}m(u^2-\alpha^2u_x^2)+\frac{1}{4}u_{xx}\\ \lambda^2u-\frac{1}{2}\lambda u_x-\frac{1}{2}m(u^2-\alpha^2u_x^2)+\frac{1}{4}u_{xx} & -\lambda^3+\frac{1}{2}\lambda(u^2-\alpha^2u_x^2)\end{bmatrix}.$$

Similarly to the CH equation, the term that we need to project out to recover the dispersive mCH equation (3.2.51) from the previous weak ZCR is

$$\lambda^2([A,\Lambda U]+[U,A])+\lambda^3[\Lambda U,U]=\alpha^2 u_{xx}\begin{bmatrix}-\lambda^2 & 2\lambda^3\\ -2\lambda^3 & \lambda^2\end{bmatrix}.$$

As for the CH equation, the mCH equation is an integrable equation with another ZCR obtained by Fokas (1995), Fokas and Liu (1996), Olver and Rosenau (1996), Qiao (2007) for a linear first order dispersion. The third order dispersion can be related to the linear dispersion by a change of variable of the form $x\to x+ct$.

Finally, for the more general form of U containing two independent fields u and v we obtain the coupled mCH equations of Xia and Qiao (2015)

$$\begin{aligned} m_t+2[m(uv-\alpha^2u_xv_x)]_x-2m(uv_x-u_xv)+u_{xxx}&=0\\ n_t+2[n(uv-\alpha^2u_xv_x)]_x+2n(uv_x-u_xv)+v_{xxx}&=0\,. \end{aligned} \tag{3.2.52}$$

The weak ZCR can be computed as well but will not be displayed here.

So far, the deformed equations were all proven to be completely integrable with a proper ZCR, different from the weak ZCR. We will not investigate

these equations further as they are now well-studied by the community of integrable systems, in particular for their dispersionless limit which contains various type of peaked solutions. One of our interest in this work is to understand the limits of this scheme for deriving new integrable equations. One of them is when we use the flow with $t = a_2$, corresponding to the NLS equation. We will obtain the so-called CH-NLS equation which never appeared before seems not to be integrable. We follow the usual calculations, starting from $M = \lambda^2 A + \lambda U^{\perp} + V$ and the NLS form of U, that is

$$\Lambda U_{NLS} = \begin{bmatrix} 0 & m \\ \sigma\overline{m} & 0 \end{bmatrix} . \tag{3.2.53}$$

Here, u is complex and $\sigma = \pm 1$ selecting between the focusing or defocusing case. We compute the Legendre transform to get the CH-NLS equation on the slice N_{12}

$$im_t + u_{xx} + 2\sigma m(|u|^2 - \alpha^2|u_x|^2) = 0 . \tag{3.2.54}$$

In term of u only it is given by

$$iu_t - i\alpha^2 u_{xxt} + u_{xx} + 2\sigma u|u|^2 - 2\sigma\alpha^2 u|u_x|^2 - 2\sigma\alpha^2 u_{xx}|u|^2 + 2\sigma\alpha^4 u_{xx}|u_x|^2 = 0 . \tag{3.2.55}$$

This equation is obtained from the weak ZCR

$$M = \begin{bmatrix} i\lambda^2 + \frac{i}{2}(|u|^2 - \alpha^2|u_x|^2) & \lambda u - \frac{i}{2}u_x \\ \lambda\overline{u} + \frac{i}{2}\overline{u}_x & -i\lambda^2 - \frac{i}{2}(|u|^2 - \alpha^2|u_x|^2) \end{bmatrix} ,$$

which needs the projection $\mathcal{P}_{12}^{A}$ to be equivalent to the CH-NLS equation (3.2.54). This ZCR cannot guarantee any integrability, and as we will see in the next section 3.3, bi-Hamiltonian structures do not work either. We will not comment further on this equation here, as we will show a more detailed analysis of it in the next section 3.3.

3.2.7 Summary and outlook

Although we broke this section down into several subsections, it contains two main parts. The first part introduces a different way of understanding integrable hierarchies such as the AKNS hierarchy (other examples may include hierarchies using $SO(3)$, as in Arnaudon (2016b), or $SL(3)$, etc...). For this, we used a rather abstract description using loop groups and cocycles to be able to disentangle the Hamiltonian and the Lagrangian derivations of integrable hierarchies. The first derivation is similar to the standard derivation of integrable hierarchies while the second is different and more useful for us as our aim is to deform the hierarchy in a geometrically consistent way. Both derivations should deserve more investigations from an abstract point of view as they provide a different link between the dynamical equations and the Lie group structures. This link was partially investigated in Flaschka et al. (1983) in a slightly different context and shown to be connected to the theory of τ functions and vertex operator (see also Newell (1985)). We did not take this path here as our main interest is in the deformation of integrable systems rather than in a general abstract theory which may surely overlap with the numerous works in this direction.

One of these geometric deformations is the Sobolev deformation which is the second theme of this section. The modification of the L^2 norm with spatial derivatives to obtain an H^1 norm is only possible in the Lagrangian as the Hamiltonian does not define a metric as it is written on a cotangent space, but a co-metric. This deformation produces known integrable equations such as the CH or the modified CH equation as well as the new CH-NLS equation. We will study it in more detail in the next section 3.3 and only mention that its apparent lack of integrability is one of the main open problems regarding this equation. In fact, the deformation of integrable hierarchies seems to destroy the complete integrability viewed from the usual isospectral problem. Indeed, we had to define a projection operator to retain the equivalence between the abstract deformed Euler-Poincaré equation and the deformed dynamical equations such as the CH equation. All the obtained equations can be shown to be completely integrable using a different ZCR, except for the CH-NLS. In some sense, the CH-NLS equation is at the boundary

between integrable deformation and non-integrable deformations, but we do not know the exact mathematical description for this fact. We will see in the next section that the CH-NLS has many integrable features and others non-standard behaviours, especially for the collision of solitons with high momentum where high-frequency waves will be observed.

3.3 The CH-NLS equation

We now study the CH-NLS equation derived in the previous section. Most of the results of this section are available in Arnaudon (2016a). We first recall the CH-NLS equation for a complex field u and with $m = u - \alpha^2 u_{xx}$

$$im_t + u_{xx} + 2\sigma m \left(|u|^2 - \alpha^2 |u_x|^2\right) = 0\,, \qquad m = u - \alpha^2 u_{xx}\,, \tag{3.3.1}$$

where $\sigma = \pm 1$ selects between the focusing and the defocusing equation. Although these two equations have different properties, we will only consider the focusing case when $\sigma = 1$. We make this choice as the focusing equation is the most common case for the NLS equation. In section 3.3.1, we will briefly review the main properties of the CH-NLS and discuss its integrability. We will then derive peaked and smooth standing wave solutions in section 3.3.2. They both appear in a one-parameter family of solutions. Although rare in dispersive equations, such transitions between smooth and peaked standing waves have been studied in a generalised NLS equation by Amiranashvili et al. (2011, 2014). We then numerically solve the initial value problem to isolate and characterise some solitary waves in section 3.3.2 and their collision in section 3.3.2. The existence and the stability of solitary waves in non-integrable equations has been an active area of research parallel to integrable systems, see for example Tao (2009) for a recent review. In the case of the CH-NLS, we observed interesting phenomena in the collision of solitary waves such as high amplitude waves and the emergence of sharp structures, almost stationary. Indeed, recall that $u = K * m$ with $K = \frac{1}{2\alpha} e^{-|x|/\alpha}$, thus the linear dispersion relation is

$$\omega(\kappa) = \frac{\kappa^2}{1 + \alpha^2 \kappa^2}\,. \tag{3.3.2}$$

The group velocity thus vanishes at high wavenumbers. These features have already been investigated,for example by Colin and Lannes (2009) in the context of short pulses in optical fibres. These extreme behaviours in the collisions are also present in the modulational instability regime as shown in section 3.3.3.

3.3.1 Symmetries and non-integrability

We review here the main features of the CH-NLS equation and in particular the differences with the NLS equation. The first common property with the NLS equation is that the CH-NLS equation is a non-canonical Hamiltonian system with Hamiltonian

$$P = \frac{i}{2}\int (\overline{m}u_x - m\overline{u}_x)dx\,, \tag{3.3.3}$$

and Hamiltonian structure

$$K = \begin{bmatrix} -2m\partial_x^{-1}m & \partial_x + 2m\partial_x^{-1}\overline{m} \\ \partial_x + 2\overline{m}\partial_x^{-1}m & -2\overline{m}\partial_x^{-1}\overline{m} \end{bmatrix}, \tag{3.3.4}$$

such that the CH-NLS equation (3.3.1) is given by

$$m_t = K\frac{\delta P}{\delta m}\,. \tag{3.3.5}$$

This Hamiltonian structure is the same as the second NLS Hamiltonian structure but the associated energy is modified to contain α-terms. This Hamiltonian is, in fact, the total momentum of the solution and is associated with the space translation invariance of the CH-NLS equation. Another common similarity with the NLS equations is the invariance under phase shifts related to the conservation of the total mass

$$M = \int |m|^2 dx\,. \tag{3.3.6}$$

Notice that the mass depends on m and not u, as for the NLS equation.

The CH-NLS equation is also invariant under the transformation

$$t \to \lambda^{-2}t, \qquad x \to \lambda^{-1}x\,, \qquad u(t,x) \to \lambda u(\lambda^2 t, \lambda x) \quad \text{and} \quad \alpha \to \lambda^{-1}\alpha\,,$$

which could be seen as a scaling transformation, but is not as the constant α is modified. This transformation shows that α introduces a particular length-scale in the equation as in the α-model of turbulence introduced in

Holm et al. (1998b).

The last relevant symmetry of the NLS equation is the Galilean invariance but is absent from the CH-NLS equation. This symmetry can be used to produce new solutions out of known one by performing a Galilean boost. Unfortunately, the length-scale α destroys this symmetry. To see this, we compute the time derivative of the centre of mass $C := \int x|m|^2 dx$ to obtain

$$\frac{d}{dt}C = i\int(\overline{u}u_x - u\overline{u}_x)dx\,, \tag{3.3.7}$$

where the right-hand side is not a conserved quantity. For the NLS equation, the right-hand side is be proportional to the total momentum P and the centre of mass thus moves with a constant speed.

We already saw in the previous section that the CH-NLS equation emerges from the deformation of the AKNS hierarchy and that its complete integrability seems to be lost. We want here to expose other arguments in favour of its non-integrability, in particular using the Hamiltonian structures of the NLS equation. A possible second Hamiltonian structure which could give a bi-Hamiltonian formulation for the CH-NLS equation would be

$$J = \begin{bmatrix} 0 & -i(1-\alpha^2\partial_x^2) \\ i(1-\alpha^2\partial_x^2) & 0 \end{bmatrix}, \tag{3.3.8}$$

which together with P and M produce the travelling wave equation $m_t = m_x$. If this Hamiltonian structure were compatible with the non-canonical structure (3.3.4), the theorem of Magri (1978) would apply and give the result of integrability of this equation, but J and K are not compatible. The failure arises in the Jacobi identity of $J+K$ and is not possible to fix by any modifications in the Hamiltonian or the Hamiltonian structure. This reflects the fact that it is also not possible to find a Hamiltonian for the CH-NLS equation that would converge to the NLS Hamiltonian when $\alpha \to 0$. Indeed, finding h such that $\frac{\delta h}{\delta u}$ or $\frac{\delta h}{\delta m}$ for the right-hand side of the CH-NLS equation (3.3.1) seems impossible.

3.3.2 Solitary wave solutions

We study here the solutions of the CH-NLS equation (3.3.1), starting with the standing waves and going to numerical experiments for travelling waves and their collisions.

Standing wave solutions

We begin the study of the solution of the CH-NLS equation (3.3.1) by trying to derive them analytically. As the weak ZCR of the CH-NLS cannot directly be used to compute solutions via the inverse scattering transform, we will try to find a solution using a travelling wave ansatz. It turns out that the computation for a travelling wave is difficult but possible for the standing wave solutions. Recall that the CH-NLS is not Galilean invariant, hence a standing wave solution will not correspond to a family of travelling waves.

The standing wave solutions will be computed for the two real functions $\phi(t)$ and $a(x)$, such that

$$u = a(x)e^{i\phi(t)}\,, \tag{3.3.9}$$

which in turn gives

$$m = (a(x) - \alpha^2 a_{xx}(x))e^{i\phi(t)} \quad \text{and} \quad u_t = i\phi_t u\,.$$

After inserting the decomposition (3.3.9) into the CH-NLS equation (3.3.1), we obtain

$$\phi_t = \frac{1}{a - \alpha^2 a_{xx}} \left(\frac{1}{2}\sigma a_{xx} + (a - \alpha^2 a_{xx})(a^2 - \alpha^2 a_x^2) \right)\,.$$

Using separation of variables, we can set $\phi = ct$, multiply by a_x and integrate the resulting equation to obtain

$$\alpha^4 a_x^4 + (\sigma + 2\alpha^2(c - a^2))a_x^2 - a^2(2c - a^2) = 0\,.$$

We then choose the integration constant to be 0 as we will only consider solu-

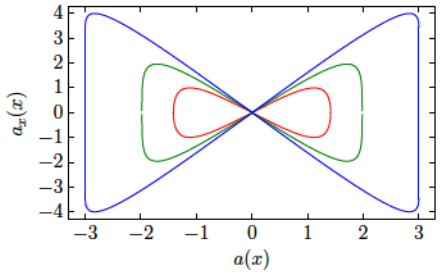
tions vanishing at $\pm\infty$. The travelling wave equation can then be calculated from the roots of the polynomial in a_x and is

$$a_x^2 = \frac{1}{2\alpha^4}\left[-(\sigma + 2\alpha^2(c - a^2)) \pm \sqrt{(\sigma + 2\alpha^2 c)^2 - 4\alpha^2 a^2}\right] . \tag{3.3.10}$$

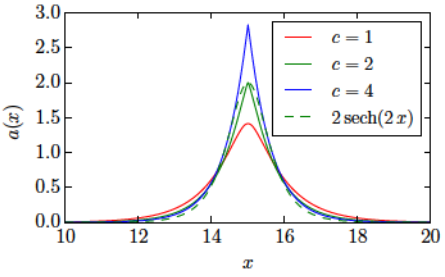
We can compute the height of the standing wave as a function of the parameter c by setting $a_x = 0$ in (3.3.10). It is given by $a_{\max} = \sqrt{2c}$ and does not depend on the parameter α. These considerations suggest that the soliton solutions are similar to the NLS solutions as they have the same height. In this calculation, we did not take into account the fact that the argument of the square root in equation (3.3.10) cannot be negative. This gives another maximum of the standing wave, with

$$a^*_{\max} = \frac{1 + 2\alpha^2 c}{2\alpha}, \quad \text{and} \quad a^*_{x,\max} = \sqrt{c^2 - \frac{1}{4\alpha^4}} . \tag{3.3.11}$$

If the maximum amplitude satisfies $a_{\max} < a^*_{\max}$, the solution will be smooth otherwise it will be peaked. Indeed, the condition $\frac{\partial a_x}{\partial a} = 0$ is equivalent to having a jump in a_x, thus a peaked profile. Finally, we can find a critical value of the parameter c giving a peaked solution in the weak sense. It is $c_{\text{crit}} = \frac{1}{2\alpha^2}$. If $\alpha \to 0$, then $c_{\text{crit}} \to \infty$ and the solution is compatible with the non-peaked NLS standing waves. Figure 3.1(a) displays the phase portrait



(a) Phase portraits of standing waves.



(b) Standing waves

Figure 3.1: The left figure displays the phase portrait of the standing wave solution given by (3.3.10) for different values of c and $\alpha = 0.5$. The legend is on the right panel, and one can see that the critical value of c in the transition between smooth and peaked solutions is $c_{\text{crit}} = 2$, displayed in magenta. The right figure shows the corresponding standing wave profiles for the same values of c and compares them with a standard hyperbolic secant profile plotted in dashed line.

of the standing wave solutions and figure 3.1(b) the corresponding standing wave solutions. We fixed $\alpha = 0.5$ and plotted the solutions for three different values of c with a smooth, a peaked and the limit solution for crit $= 2$. These solutions have also been tested numerically and remain stable even in the case of a peaked profile. This type of standing waves is rather unusual except for the CH equation and its peakon solutions but has recently been discovered and investigated by Amiranashvili et al. (2011, 2014) in the context of short pulses in optical fibres.

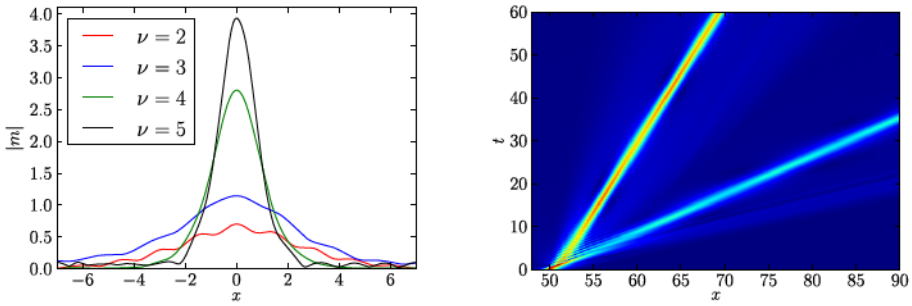
Solitary waves

After the standing waves, we need to study the travelling waves, but as it is too difficult to derive analytically, we will use numerical simulations. Recall that the travelling wave solution of the NLS equation is the soliton described by

$$u(x, t = 0) = e^{i\nu x} A \operatorname{sech}(Ax)\,. \tag{3.3.12}$$

Upon applying a Galilean boost, it can be transformed to give the full time dependant solution, which is not possible for the CH-NLS equation. Here is a possible explanation of why it is hard to find an explicit solution of the CH-NLS equation. In the case of the NLS equation, A and ν are two independent parameters, but for the CH-NLS equation they should not be independent anymore for (3.3.12) to be a solution. Indeed, the Helmholtz operator will couple the phase and the amplitude of the solution in a non-trivial manner as already seen from the previous computation of the standing wave.

To study the travelling waves of the CH-NLS equation, we performed numerical experiments with initial conditions given by the NLS solitons (3.3.12). We display in figure 3.2(a) various profiles of solitary waves which emerged from this initial condition with parameters $A = 1$ and various $\nu = 1, 2, 3, 4, 5$, at time $t = 150$. We also fixed $\alpha = 0.3$ and made sure that the numerical simulation converged for these solutions. In figure 3.2(b) we show a space-time plot of the fission of a solitary wave in two travelling waves and radiation.



(a) Solitary wave profiles extracted from fissions of initial pulses, see the right panel.

(b) Space-time plot from initial condition (3.3.12) with $A = 1.5$ and $\nu = 1$.

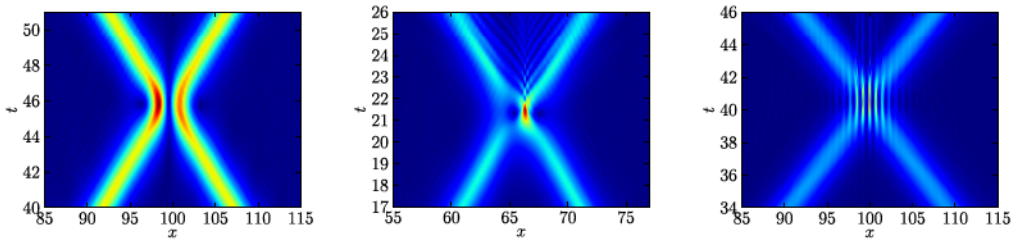
Figure 3.2: We display several numerical solutions of the CH-NLS equation with solitary wave profiles on the left and a fission of an initial pulse in the right panel. In the left panel, we can observe oscillations on top of the solitary wave, which periodically oscillate in time without affecting the propagation of the solitary wave. The fission of the right panel produces two different solitary waves and some radiation.

To understand the results of these numerical simulations, we need to have a closer look at the CH-NLS equation itself. The crucial difference with the NLS equation is the length scale α, which also gives a scale in wavelength with $1/\alpha$. Therefore, depending on the peak frequency κ_0 and the spectral width $\Delta\kappa$ of the initial pulse, we expect different behaviours in the evolution of this pulse. If $|\kappa_0 \pm \Delta\kappa| \ll 1/\alpha$, we expect the solution to behave like an NLS soliton. Indeed, the α-terms of the CH-NLS equation will be small and can be neglected at first order. If instead the solution contains high frequencies with respect to the scale $1/\alpha$, new phenomena are to be expected. This is indeed observed on figure 3.2(b) where the initial pulse with $A = 1.5$ and $\nu = 1$ has frequencies large enough compared to $1/\alpha \approx 3.3$ to trigger a fission. Two pulses emerge and have similar widths but different speeds and amplitudes. This is explained by dispersion relation (3.3.2) where high frequency waves travel slower than low-frequency waves. The interesting thing is that the two pulses have rather different properties. The fast pulse, with low-frequencies, is very close to being a soliton of the NLS equation, whereas the slow pulse shows a breathing behaviour, with high-frequency phase variations. By breathing behaviour, we mean higher frequency components which propagate on top of the main pulse and periodically oscillate in time. Notice

that the dissipative waves that emerged from the fission are to be expected and are also present in the NLS equation when the initial conditions are not perfect solitons. On the left panel 3.2(a), we display a few examples of this new type of solitary waves with breather behaviours that emerged from similar fissions but with various initial frequencies of the phase oscillations. The larger the initial frequency, the sharper the peak, and the slower the solitary wave.

Collision of solitary waves

The next step in the study of solitons is to study their collisions, and in particular what comes out after head-on collisions. The process of the collision itself is rather complicated and nonlinear, but for integrable PDEs, one can show that the outcome of the collision is again two solitons with possibly different amplitudes and phase shifts. For the CH-NLS equation, we expect to lose this property partially. We will study this effect by running initial value



(a) Slow and large solitary wave.

(b) Slow and narrow solitary wave.

(c) Fast and large solitary wave.

Figure 3.3: In these three figures, we display three different head-on collisions of solitary waves extracted from the simulations of figure 3.2(a). The left panel shows a typical NLS collision while the two others have higher amplitude waves during the collisions. The very high-frequency waves that emerge from the collision may not be physical as the numerics do not converge, see the discussion in the appendix B.

problems with two solitary waves extracted from the simulations of the previous section. We display in figures 3.3(a), 3.3(b) and 3.3(c) these collisions featuring different properties. In particular, the first panel of figure 3.3(a) is similar to an NLS collision as the solitary waves have a low-frequency spectrum. The height difference observed in the peak amplitude near the collision

is due to a small loss of symmetry in the simulation and should not be seen in principle.

The other two collisions in panels 3.3(b) and 3.3(c) have solitary waves with higher frequencies, either because they are sharper, as in 3.3(b), or faster as in 3.3(c). In both cases, these frequencies are close enough to the scale $1/\alpha$ such that the nonlinear dispersive effects of the α-terms begin to appear. These effects are observed during the collision where high amplitude structures are created. These peaks are numerically difficult to resolve as they are rather sharp. This explains the high-frequency radiations which do not move after the shock as high-frequency waves have small group velocities. We will discuss the numerical issues in the later appendix B but first focus on the appearance of these high amplitude waves present only for a short time. This experiment is an interesting integrability test despite the lack of numerical convergence. If we collide higher energy or frequencies pulses, we would observe an even more dramatic collision where the energy would be localised in a region of space comparable with or smaller than the spatial resolution. Most of the information of the previous state of the system is then lost in the collision, and no solitary waves would emerge after it. This suggests that the lack of integrability of the CH-NLS equation is seen here as allowing the formation of extremely high density and localised regions during the collision. As we will see, the lack of numerical convergence even for moderate energy collision indicates that these energy peaks can be as high as we want. If they eventually become infinite for a very short time, the information of the previous state will be lost during the collision. A similar phenomenon also appear in the CH equation for the peakon anti-peakon collision where all the energy of the solitons is localised at a single point, making numerical simulations extremely challenging. This analysis thus demonstrates that even if the integrability of the CH-NLS equation is not clear, it shows behaviours similar to a known integrable equation of the same type, the CH equation of Camassa and Holm (1993).

Interestingly, these extreme events with large amplitudes compared to the mean wave amplitudes are part of an active area of study and are called rogue waves. They are very rare and sometimes observed in the deep ocean where

the NLS equation can be used to describe the deep water waves. In fact, the natural place where these waves appear is in the modulational instability regime, which we will now study.

3.3.3 Modulational instability

The modulational instability, also called Benjamin-Feir instability, was first discovered by Benjamin and Feir (1967) in the context of deep water waves in the NLS equation. It is a narrow band instability that can develop on top of plane wave solutions which only exists for the focusing equations such as the focusing NLS equation. We refer to the monographs Agrawal (2007), Hasegawa and Kodama (1995) for more detail about this mechanism. We will show here that this phenomenon also exists in the focusing CH-NLS equation and expose some of the differences with the NLS equation.

The derivation of the condition for the modulational instabilities goes as follows. First, we write a plane wave solution of the CH-NLS equation as

$$u(x) = \beta e^{i2\beta^2 x},$$

for a given real parameter β. We then perturb this solution with a small function $\epsilon(x,t)$ such that $|\epsilon| \ll |\beta|$, that is we define a perturbed solution

$$u_\epsilon(x,t) := (\beta + \epsilon(x,t))e^{2i\beta^2 x}.$$

After a direct substitution into the CH-NLS equation (3.3.1), we obtain an approximate equation for the perturbation ϵ which is now a linear partial differential equation of the form

$$i\epsilon_t - i\alpha^2\epsilon_{xxt} + \epsilon_{xx} + 2\beta^2(\epsilon + \bar{\epsilon}) = 0. \tag{3.3.13}$$

We are then looking at the growth rate of this small perturbation. We do so by analysing the plane wave solution of this equation, that is we choose

$$\epsilon = \epsilon_+ e^{-i\omega t + i\kappa x} + \epsilon_- e^{i\omega t - i\kappa x},$$

where ϵ_+, ϵ_- and ω are free parameters but κ is not (or ω, depending if κ is taken to be free). The relationship between κ and ω can be directly computed from the linear equation (3.3.13). The imaginary part of κ produces either an exponential damping or an exponential growth of the perturbation ϵ. The imaginary part of κ, that we call the gain $g(\kappa)$, is given by

$$g(\kappa) = \mathrm{Im}\left(\frac{\kappa\sqrt{2\beta^2 - \kappa^2}}{1 + \alpha^2\kappa^2}\right), \tag{3.3.14}$$

and corresponds to the regime of modulational instability when its right-hand side has the correct sign.

Already by looking at the equation (3.3.13) for ϵ, one notices that the only α-dependent term is the time derivative of ϵ. Indeed, the other terms disappeared after the linearization. We obtained a rather simple modification of the growth rate (3.3.14) as it only becomes smaller for larger values of α, but the instability window persists. This effect is only due to the inverse Helmholtz operator of the linear dispersion.

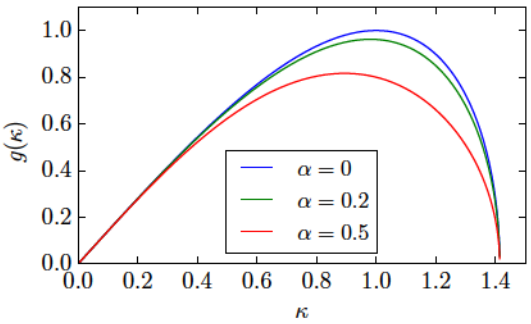


Figure 3.4: We plotted the growth rate of the modulational instability given by equation 3.3.14 as a function of the perturbation wavelength κ. With different values of α, we illustrate the effect of the nonlocality of the CH-NLS equation.

We illustrate this by plotting on figure 3.4 the gain g as a function of κ for different values of α and with $\beta = 1$. The CH-NLS has a similar modulational instability to the NLS equation, at least during the linear stage of growth of the perturbation ϵ. After some time, the nonlinear effect will become non-negligible, and the additional α-terms will further modify the behaviour of

the solution compared to the NLS regime. This nonlinear regime is hard to analyse analytically but has seen recent progress in the form of integrable turbulence, see for example Agafontsev and Zakharov (2015), Zakharov (2009). Indeed, the main feature of the non-linear regime is the creation of many pulses from the instabilities which interact and produce a chaotic, or turbulent regime. We will only study this regime numerically and extract the rogue wave events of the turbulent nonlinear regime. We refer to Dudley et al. (2014) for a recent numerical study on the statistics of the rogue wave events in the NLS equation in the context of optical fibres. The initial condition of the simulations is a plane wave of amplitude 1 and two oscillations in a periodic domain of length 50 units. The instabilities are triggered by the numerical errors of the spatial discretization. We display some of the

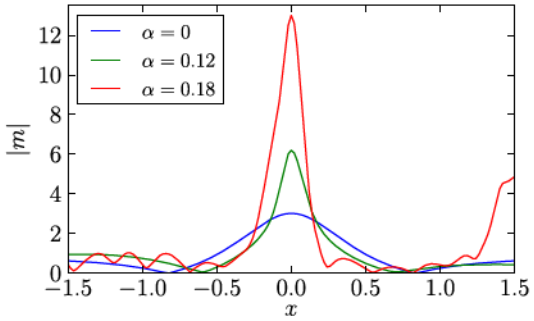


Figure 3.5: In this figure, we display three different rogue wave events of the CH-NLS equation, extracted from three simulations of modulational instabilities obtained with various values of the parameter α. The blue line with $\alpha = 0$ corresponds to the standard Peregrine solution of the NLS equation and has a peak amplitude of three times the background, so an amplitude of 1. The Peregrine solution of the CH-NLS equation shows a much larger amplitude, up to 12 times the background.

Peregrine solutions in figure 3.5 which were obtained from simulations with different values for α. The result is in accordance with the previous section where we studied the nonlinear interactions of solitary waves. Indeed, the extreme events of the NLS equation are becoming more extreme for larger values of α. For Peregrine solutions, the peak amplitude is three times larger than the background field, which is 1 here. In the two other simulations where the value of α remains rather small, we observe events with up to 12 times the height of the background. In these simulations, the wave produces

high-frequency radiation similar to that in the collision of solitary waves.

3.3.4 Summary and outlook

Even after investigating the properties of the CH-NLS equation, many mysteries remain. The principal one is about its non-integrability and the existence of another Hamiltonian function. The latter could provide a bi-Hamiltonian structure, and a path to integrability but no such Hamiltonian had been found. One of the properties of this equation indicating a near integrability is the existence of solitary waves and near elastic collisions. Even here, standard numerical schemes are not precise enough to fully resolve the collisions and provide us with a way to exactly understand this process. We can nevertheless say that for large momenta the collisions produce high-frequency waves which almost do not move. These waves are the problem in the numerics and would require a higher order scheme with an excellent spatial resolution. Nevertheless, we observed the emergence of large amplitude Peregrine solutions from the modulational instability regime, waves which are known to be excellent candidates for rogue wave events. Similar equations can be derived with the Helmholtz operator for a nonlocal dispersion, but the particular form of the nonlinear term seems to have never appeared before in the physics literature. The closest systems are the ones which take into account a small delay in the effect of the nonlinearities. This effect is captured by a convolution operator in the nonlinear term but does not seem to be reducible to the CH-NLS nonlinearities.

Other paths for investigating the physical nature of the CH-NLS equation is to study its higher dimensional extension

$$i\frac{\partial}{\partial t}\Psi + \Delta\Phi + 2\sigma\Psi\left(|\Phi|^2 - \alpha^2|\nabla\Phi|^2\right) = 0\,, \tag{3.3.15}$$

for a field $\Psi(\mathbf{x},t)$, where $(\mathbf{x},t) \in \mathbb{R}^n \times \mathbb{R}$, $\Psi = \Phi - \alpha^2\Delta\Phi \in \mathbb{C}$ and $\sigma = \pm 1$. This equation is closer to the Gross-Pitaevskii equation used in the theory of Bose-Einstein condensates (see section 3.6), but here again, the nonlinear term that is often used does not correspond to these nonlinearities.

Nevertheless, it would be interesting to study this equation in two dimensions, and in particular, the possible effect of the α-terms on the blow-up of localised solutions, see for example the monograph Sulem and Sulem (2007).

3.4 Stochastic hierarchies

In this section, we will use the geometric interpretation of the integrable hierarchies and, in particular, the AKNS hierarchy to introduce stochastic perturbations such that the notion of stochastic integrability defined in the first chapter in section 2.2.1 is retained. By adding noise in a geometrically consistent way, we will show that arbitrary linear combinations of integrable flows with stochastic parameters will remain integrable. From this result, we conjecture that no other stochastic PDEs can be completely integrable. Indeed, there are no other cases of integrable stochastic PDEs in the literature than the original work of Wadati (1983) on the KdV equation and subsequent papers. As we will see, these works are contained in the present theory.

We will start by introducing noise into the Lagrangian derivation of the AKNS hierarchy in section 3.4.1, then derive various equations of the stochastic AKNS hierarchy and develop a stochastic extension of the inverse scattering transform in section 3.4.2 to end by the study of some of the stochastic equations in section 3.4.3. This section is contained in Arnaudon (2015).

3.4.1 Stochastic constrained variational principle

In the first chapter, we saw that for finite dimensional systems, the noise is compatible with the reduction by symmetry if we choose the stochastic potential to be linear in the momentum. In this case, we will slightly relax this condition to obtain a more interesting type of noise while having the reduction by symmetry available together with the complete integrability. To write this stochastic deformation into a variational principle, we will use the Clebsch derivation described in section 2.2.1. Notice that the Hamilton-Pontryagin approach of section 2.2.1 would also work.

We will use the following constrained variational principle for a time t_i and

associated velocity $M := M^{(i)}$

$$S = \int \left(l(M) + \left\langle (p,\mu), \partial_{t_i} (q,a) + \mathrm{ad}_{(M,\xi)}(q,a) \right\rangle \right) dt_i + \sum_l \int \Phi_l(q,p) \circ dW_{t_i}^l \,. \tag{3.4.1}$$

where $(p,\mu) \in T^*\widetilde{LG}$ are the so-called Clebsch variables used to enforce the constraint that the variables $(q,a) \in T\widetilde{LG}$ evolve advected by $(M,\xi) \in \widetilde{L\mathfrak{g}}$. We recall that $\widetilde{LG}^c$ is the polynomial loop group with a central extension. The element in the centre denoted a is also a polynomial in λ with the same order as q. The function Φ is the stochastic potential and introduces noise into this variational principle. At this point it is arbitrary, but it does not depend on the variable in the centre of the loop algebra. We first compute the variations of (3.4.1) to find the following set of equations

$$\begin{aligned} \frac{\delta l}{\delta M} &= \mathrm{ad}_q^* p + \mu c(q,\cdot) =: L \\ d(p,\mu) &= \mathrm{ad}^*_{(M,\xi)}(p,\mu)\, dt_i + \sum_l \left(\frac{\delta \Phi_l}{\delta q}, 0 \right) \circ dW_{t_i}^l \\ d(q,a) &= -\mathrm{ad}_{(M,\xi)}(q,a)\, dt_i - \sum_l \left(\frac{\delta \Phi_l}{\delta p}, 0 \right) \circ dW_{t_i}^l \,. \end{aligned} \tag{3.4.2}$$

The first equation is only for the loop algebra and has no component in the centre so that the coadjoint action is the standard coadjoint action. The μ variable is in the centre of the Lie algebra but does not have any dynamics as seen from the second equation. It can thus be taken as a constant. The last equation has an adjoint action and produces some dynamics for the a variable in the centre. This variable will not affect the dynamics for the momentum L as it does not appear in the first equation.

To proceed with the derivation of the equation for L, we need to assume that $\Phi_l(q,p) = \Phi_l(L)$. However, we do not restrict Φ_l to be linear in the L variable, so it can be a quadratic form such that we can define the conjugate

variable $\Xi_l = \frac{\delta \Phi_l}{\delta L}$ and

$$dX_i := M dt_i + \sum_l \Xi_l \circ dW_{t_i}^l \,. \tag{3.4.3}$$

We then skip the derivation of the equation for L which is standard and uses the Jacobi identity. We obtain the stochastic Euler-Poincaré equation with cocycle

$$d\frac{\delta l(M)}{\delta M} = \mathrm{ad}^*_{dX_i} \frac{\delta l(M)}{\delta M} + c(dX_i, \cdot) \,. \tag{3.4.4}$$

Explicitly, using the derivative cocycle and the pairing of the semi-simple Lie algebra $\mathfrak{g}$ this equation simplifies to the stochastic ZCR

$$dL - \partial_x dX_i + [L, dX_i] = 0 \,. \tag{3.4.5}$$

By writing the ZCR even in term of Ξ, we obtain

$$dL = (\partial_x M^{(i)} - [L, M^{(i)}]) dt_i + \sum_l (\partial_x \Xi_l - [L, \Xi_l]) \circ dW_{t_i}^l \,, \tag{3.4.6}$$

and if Ξ_l is taken to be constant, as dictated by the reduction by symmetry, the space derivative will only be multiplied by a commutator of a constant with L. By choosing Ξ_l such that this term does not vanish, it is straightforward to see that the noise amplitude is proportional, for example for KdV, to $u \circ dW_t$. As we seek here a more nonlinear noise amplitude, we proceed by finding Lie algebra elements $\Xi_l(L)$ which would retain the complete integrability.

3.4.2 The stochastic AKNS hierarchy

From the previous derivation of the equations on the loop algebra, we can easily find the possible form for Ξ_l. If we follow the approach of the deterministic case, we will need to compute the Legendre transform, that is to find the relation $M(L)$, and this $\Xi_k(L)$. What helps us here is that the noise amplitude has the same form as the drift, obtained by setting $M = M^{(i)}$ for

(i,k)	Reduction	Stochastic PDE
(2,1)	$q = u$ $r = \overline{u}$	$du = i\left(u_{xx} + 2u\lvert u\rvert^2\right) dt + u_x \circ dW_t$
(2,2)	$q = u$ $r = \overline{u}$	$du = i\left(u_{xx} + 2u\lvert u\rvert^2\right)\left(dt + \circ dW_t\right)$
(2,3)	$q = u$ $r = \overline{u}$	$du = i\left(u_{xx} + 2u\lvert u\rvert^2\right) dt + \left(u_{xxx} + 6\lvert u\rvert^2 u_x\right) \circ dW_t$
(3,1)	$q = u$ $r = u^n,\ n = 0, 1$	$du = \left(u_{xxx} + 6u^{n+1}u_x\right) dt + u_x \circ dW_t$
(3,2)	$q = u$ $r = \overline{u}$	$du = \left(u_{xxx} + 6\lvert u\rvert^2 u_x\right) dt + i\left(u_{xx} + 2u\lvert u\rvert^2\right) \circ dW_t$
(3,3)	$q = u$ $r = u^n,\ n = 0, 1$	$du = \left(u_{xxx} + 6u^{n+1}u_x\right)\left(dt + \circ dW_t\right)$

Table 3.1: In this table we present the first few flows of the stochastic AKNS hierarchy time flow t_i and a single stochastic time t_k. The second column displays the reduction we applied on the fields L and M to arrive at the stochastic integrable equations of the last column.

a given i. We can thus set $\Xi_l = \sigma_l M^{(k_l)}$ for a possibly different index k_l for each noise index l. The real number σ_l represent the overall strength of the noise. Of course, if $k_l = k_{l'}$ for two different l and l', then both noises will equally contribute and give a single noise with overall strength $\sigma_l + \sigma_{l'}$. From now on, we will call k_l the stochastic times.

The stochastic hierarchy will thus have two indices for space i, and time j, as well as several indices for the stochastic perturbations denoted k_l. The integrable flows thus live on the slice N_{ij} but are stochastically deformed in the slices N_{ik_l}. For the AKNS hierarchy, the computations are straightforward from those of section 3.2. We collected the first few stochastic flows in table 3.1 and we will study some of them in more details in the next section 3.4.3. Before that, we investigate the integrability of these equations via the bi-Hamiltonian formulation and the inverse scattering transform.

Stochastic bi-Hamiltonian structures

One of the easiest proof of integrability is to show the existence of a bi-Hamiltonian structure. This method is due to Magri (1978). So far we only mentioned it for the CH-NLS equation in section 3.3 but all of the standard integrable 1+1 PDEs are bi-Hamiltonian. The stochastic deformation presented here preserves the integrability, and thus the bi-Hamiltonian structure. Indeed, one can directly check that for two Hamiltonian structures J and K and a sequence of Hamiltonians h_i generated by these two structures, we can write a flow with time i and stochastic time k as

$$\begin{aligned} du &= J\frac{\delta h_i}{\delta u}dt + J\frac{\delta h_k}{\delta u} \circ dW_t \\ &= K\frac{\delta h_{i+1}}{\delta u}dt + K\frac{\delta h_{k+1}}{\delta u} \circ dW_t \,. \end{aligned} \tag{3.4.7}$$

It is important to note that the two Wiener processes which appear in this formula have the same realisation of the noise despite having a different so-called stochastic time. Indeed, if they were to have different realisations of the noise, no equality could hold.

Stochastic inverse scattering transform

Recall that all of these stochastic PDEs were derived from the stochastic ZCR (3.4.5) which is can be interpreted as a compatibility condition for a stochastic isospectral problem. The compatibility condition is $\partial_x d\Psi = d\partial_x \Psi$, where $\Psi(x,t,\lambda)$ is the wavefunction evolving in time and space according to

$$d\Psi = M\Psi dt + \sum_l \Xi_l \Psi \circ dW_t^l \tag{3.4.8}$$

$$\partial_x \Psi = L\Psi \,. \tag{3.4.9}$$

As expected, only the time evolution of the wavefunction is stochastic and the spectral problem with L remains the same as in the deterministic case. We now want to briefly go through the theory of IST following the classic exposition of Ablowitz and Segur (1981), Ablowitz et al. (2004) and show

how to modify this method to include the stochastic terms. For the clarity of this derivation, we will also restrict ourselves to the simple case of NLS with a single noise.

We first need to introduce the scattering data associated with the scattering problem (3.4.9). Recall that L is linear in λ and $u, \overline{u}$ vanishes at $\pm\infty$. The four boundary conditions for the eigenfunction Ψ can thus be written

$$\begin{aligned} \phi(x,\lambda) &\sim \begin{pmatrix} 1 \\ 0 \end{pmatrix} e^{-i\lambda x}, \quad \overline{\phi}(x,\lambda) \sim \begin{pmatrix} 0 \\ 1 \end{pmatrix} e^{i\lambda x} \quad \text{as } x \to +\infty \\ \psi(x,\lambda) &\sim \begin{pmatrix} 0 \\ 1 \end{pmatrix} e^{i\lambda x}, \quad \overline{\psi}(x,\lambda) \sim \begin{pmatrix} 1 \\ 0 \end{pmatrix} e^{-i\lambda x} \quad \text{as } x \to -\infty\,. \end{aligned} \tag{3.4.10}$$

Having fixed the boundary conditions, we introduce the transitions coefficients $a(\lambda)$ and $b(\lambda)$ which relate the boundary conditions between $-\infty$ and ∞. They are defined by the relations

$$\begin{aligned} \phi(x,\lambda) &= b(\lambda)\psi(x,\lambda) + a(\lambda)\overline{\psi}(x,\lambda) \\ \overline{\phi}(x,\lambda) &= \overline{a}(\lambda)\psi(x,\lambda) + \overline{b}(\lambda)\overline{\psi}(x,\lambda)\,. \end{aligned} \tag{3.4.11}$$

At this point, it is possible to solve the spectral problem and compute the shape of soliton solutions for example. As this computation does not involve noise and is standard, we will skip it and assume that we know its solution. The important part is that it will depend on the scattering data λ and a, b, which are arbitrary at this point. The problem is isospectral, so λ will not depend on time, but a and b will. To find the time evolution of these parameters, we need the stochastic evolution of Ψ by M and Ξ.

We have their values at infinity, denoted by $M_\infty = M(\infty)$ and $\Xi_\infty = \Xi(\infty)$. They are matrices depending only on λ and of the form $M_\infty = \operatorname{diag}(m_\infty, -m_\infty)$ and $\Xi_\infty = \operatorname{diag}(\xi_\infty, -\xi_\infty)$. A solution of (3.4.8) at $x = \pm\infty$ can now be derived in term of these quantities and read

$$v^+ = \begin{pmatrix} e^{m_\infty t + \xi_\infty W_t} \\ 0 \end{pmatrix} \quad \text{and} \quad v^- = \begin{pmatrix} 0 \\ e^{-m_\infty t - \xi_\infty W_t} \end{pmatrix}. \tag{3.4.12}$$

Notice that they are not compatible with the fixed boundary conditions for ϕ and ψ which we previously defined in (3.4.10) to solve the spectral problem. As in the deterministic case, by using intermediate functions one arrives at the following equation for the time evolution of the eigenfunctions ϕ and Ψ

$$\begin{aligned} d\phi &= (M - M_\infty)\phi\, dt + (\Xi - \Xi_\infty)\phi \circ dW_t \\ d\overline{\phi} &= (M + M_\infty)\overline{\phi}\, dt + (\Xi + \Xi_\infty)\overline{\phi} \circ dW_t\,. \end{aligned} \tag{3.4.13}$$

Using the formula (3.4.11) we can express these equation in term of a and b only to get

$$\begin{aligned} da &= 0 \\ db &= -2m_\infty b\, dt - 2\xi_\infty b \circ dW_t\,. \end{aligned}$$

The transition coefficient a is thus a constant and b satisfies a stochastic process which can directly be integrated. Its solution is

$$b(\lambda, t) = e^{-2\Theta(t,\lambda)} b(\lambda, 0)\,, \tag{3.4.14}$$

for an initial condition $b(\lambda, 0)$ where we defined

$$\Theta := m_\infty t + \xi_\infty W_t\,. \tag{3.4.15}$$

We can now use these results to find the missing parameters of the soliton solutions which depend on time. For the NLS reduction $q = u, r = -\overline{u}$, the standard theory of IST gives the one-soliton solution

$$u(x,t) = -2i\frac{\overline{b}(t)}{|b(t)|}\eta e^{-2i\xi x}\,\mathrm{sech}(2\eta x - 2\phi)\,, \tag{3.4.16}$$

where ϕ is a phase such that $|b(0)|/(2\eta) = e^{2\phi}$ and the complex eigenvalue decomposes as $\lambda = \eta + i\xi$ to give two real valued parameters η and ξ. Notice that the values ξ_∞ and m_∞ will also depend on λ, thus on ξ and η. After substituting the explicit form of $b(t)$ in the solution (3.4.16) we can rewrite

it as

$$u(x,t) = 2\eta e^{-2i\xi x} e^{2i\mathrm{Im}\,\Theta(t)} \operatorname{sech}(2\eta x + 2\mathrm{Re}\,\Theta(t))\,. \tag{3.4.17}$$

This formula will be use in the next section to derive the soliton solutions of some of the integrable stochastic PDE of the stochastic AKNS hierarchy.

3.4.3 Stochastic integrable equations

We now study several integrable stochastic equations of the stochastic AKNS hierarchy displayed in table 3.1. We will simply survey some interesting features of them and derive their one-soliton solutions. We will denote by the type of the stochastic equation the stochastic time, that is for example type III is with $k = 3$.

Stochastic KdV and NLS equations of type I

The simplest stochastic deformation of the KdV or the NLS equation is with $k = 1$. Recall that the flow with $k = 1$ is a travelling wave, so the stochastic equations are

$$du = i\left(u_{xx} + 2u|u|^2\right) dt + u_x \circ dW_t \quad \text{and} \tag{3.4.18}$$

$$du = \left(u_{xxx} + 6u^{n+1}u_x\right) dt + u_x \circ dW_t \quad n = 0, 1\,. \tag{3.4.19}$$

Their Itô formulation gives a simple second order dissipative term

$$du = i\left(u_{xx} + 2u|u|^2\right) dt + \frac{1}{2}u_{xx}dt + u_x\, dW_t \quad \text{and} \tag{3.4.20}$$

$$du = \left(u_{xxx} + 6u^{n+1}u_x\right) dt + \frac{1}{2}u_{xx}dt + u_x\, dW_t \quad n = 0, 1\,. \tag{3.4.21}$$

From the fact that the noise depends on the space variable only through the gradient of the fields u, the spatial transformation

$$x \to x - W_t \tag{3.4.22}$$

maps these stochastic equations into their deterministic version. Any solutions of both equations will then be of the form $a(t, x - W_t)$ where $a(t, x)$ is the soliton of the deterministic equation. The simplicity of this equation comes from the fact that the stochastic time is the same as the space variable when viewed from the general theory studied before. In turn, a simple linear stochastic change of variable in the slice N_{1i} for a given i removes the noise of the PDE.

The one-soliton solution only requires the computation of Θ defined in (3.4.15)

$$\Theta(t) = i\lambda^2 t + i\lambda W_t = i(\xi^2 - \eta^2 + 2i\eta\xi)t + i(\xi + i\eta)W_t\,, \tag{3.4.23}$$

and the solution (3.4.17) is

$$u(x,t) = 2\eta e^{-2i\xi x} e^{2i(\xi^2-\eta^2)t+2i\xi W_t} \mathrm{sech}[2\eta(x - 2\xi t - W_t)]\,. \tag{3.4.24}$$

This solution corresponds to the deterministic solution with a stochastic spatial shift as we expected from the previous discussion.

The stochastic KdV equation (3.4.19) corresponds to one of the first instances of a stochastic integrable equation discovered and studied by Wadati (1983). Wadati studied the stochastic KdV equation $u_t = (u_{xxx} + 6uu_x)dt + Z_t$ for a Brownian motion Z_t. After the change of variable $u \to u - Z_t$, we obtain the equation (3.4.19) with $W_t = \int Z_t dt$. This shows that Wadati's equation is equivalent to the equation (3.4.19) and in fact, Wadati used this change of variables to prove the complete integrability of this stochastic KdV equation. Interestingly, this change of variables only works for the KdV equation. For example, this modification for the stochastic NLS equation modifies the form of the equation.

Finally, there is a simple physical interpretation of these equations where the noise corresponds to a small constant linear dispersion fluctuating around zero. Such systems were studied for example by de Bouard and Debussche (2010) with a fluctuating second order dispersion. The fact that a change a variable can remove this stochastic first order dispersion makes this equation less interesting than the one of de Bouard and Debussche (2010). Never-

theless, an extension to a space-time white noise $dW_t(x)$ would make things more physical but out of the scope of this work. Such a stochastic linear dispersion parameter can be a measurable and meaningful quantity in the study of optical fibres, see for example the monograph Agrawal (2007).

Stochastic NLS equation of type II and KdV/mKdV equations of type III

The next stochastic deformation of the NLS or KdV equations is when the time and the stochastic time are the same. This corresponds to $k = 2$ (type II) for NLS and $k = 3$ (type III) for KdV. They can be written as

$$du = i\left(u_{xx} + 2u|u|^2\right)(dt + \circ dW_t) \quad \text{and} \tag{3.4.25}$$

$$du = \left(u_{xxx} + 6u^{n+1}u_x\right)(dt + \circ dW_t)\,, \quad n = 0, 1. \tag{3.4.26}$$

From this formulation and the fact that the stochastic time is the same as the deterministic time, it is clear that the change of variable $t \to t + W_t$ will transform these equations into their deterministic version. Notice that their Itô correction contains rather involved dissipative terms. For example, for NLS we obtain the following correction

$$-2u_{xx}|u|^2 + u^2\overline{u}_{xx} - 2u|u|^4 - \frac{1}{2}u_{xxxx} - 2(u|u|^2)_{xx}\,.$$

We proceed as before and derive the one-soliton solution by computing $\Theta(t)$ to get

$$\Theta(t) = i\lambda^2 t + i\lambda^2 W_t = i(\xi^2 - \eta^2 + i2\eta\xi)(t + W_t)\,. \tag{3.4.27}$$

The speed of the soliton is thus stochastic and follow the process $dc(t) = 2\xi(dt + dW_t)$. Notice that this solution corresponds to the previous change of variables $t \to t + W_t$. In this equation, the noise amplitude is proportional to the speed of the deterministic soliton parametrized by ξ, the imaginary part of the eigenvalue of the spectral parameter λ. This noise thus influences the speed of the soliton and its phase shift. For higher order solitons or breathers,

the speed of the oscillations or its shape will also become stochastic. We illustrate this effect in figure 3.6 where we present a numerical simulation of the collision of a first order and a second order soliton of the NLS equation of type II.

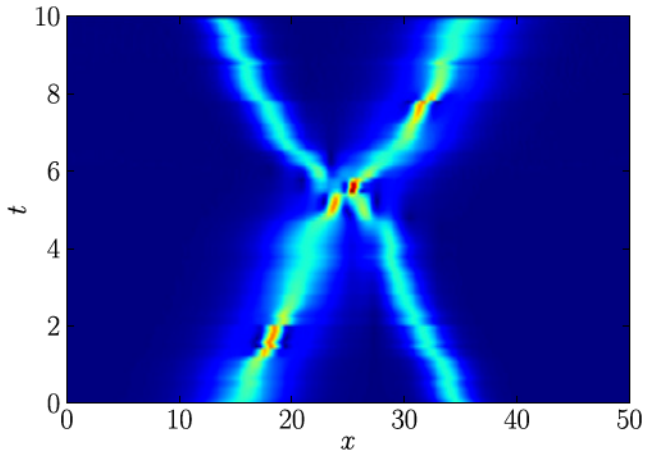


Figure 3.6: This figure shows the amplitude $|u|$ of a numerical solution of the stochastic NLS of type II with parameters $\sigma = 0.3$, $dt = 0.5 \cdot 10^{-4}$ and $dx = 0.1$. The initial condition is created from two solitons, one with parameters $\eta = 0.5$ and $\xi = -0.5$ and the other, a second order soliton (the amplitude is multiplied by 2 compared to the first order soliton), with parameters $\eta = 0.25$ and $\xi = 0.5$. The solitons keep their identity even after a collision and the phase shift of the solitons after the collision depends on the realisation of the noise. The higher order soliton sees its periodic motion affected by the noise but remains cyclic, with a random periodicity.

From the change of variables of the stochastic NLS I and II, we can find other solutions such as breather solutions. We study here the classic solution of Peregrine (1983), given by the formula

$$u(x,t) = \left(1 - \frac{4(1+2it)}{1+4x^2+4t^2}\right) e^{it}. \tag{3.4.28}$$

For the stochastic NLS, the solution given is the same but with the change of variables $u(x + W_t, t)$ for the type I and $u(x, t + W_t)$ for the type II. We display in figure 3.7 the deterministic Peregrine solution and the two stochastic solutions. For comparison, the realisation of the noise is the same in the two simulations, but it influences them differently. The type I noise produces a random horizontal shear whereas the type II noise has a vertical

random stretching effect on the solution.

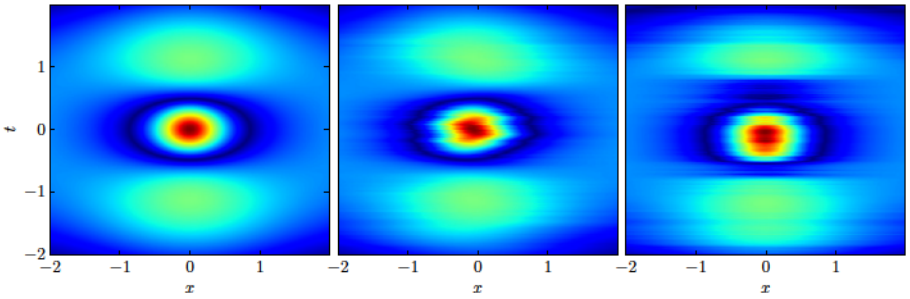


Figure 3.7: A deterministic and two stochastic Peregrine solutions are displayed in this figure. The left panel shows the deterministic solution, the middle panel shows the type I noise and the right panel the type II noise. The standard deviation of the noise is $\sigma = 0.01$, and the noise realisation is the same for both simulations.

Stochastic complex mKdV of type II

There is only one equation when $k = 2$ and the flow time follows the time t_3. The reduction is the NLS reduction as it is the only reduction for the time t_2. The equation becomes a complex mKdV equation or CmKdV of type II. The deterministic part is the so-called Hirota equation introduced by Hirota (1973). The noise is proportional to the NLS equation, and the corresponding stochastic PDE is

$$du = \left(u_{xxx} + 6|u|^2 u_x\right) dt + i\left(u_{xx} + 2u|u|^2\right) \circ dW_t\,. \tag{3.4.29}$$

In this case, it is not possible to find a change of variables which would remove the noise as we did in the previous cases and we only rely on the solutions given by stochastic IST.

We compute Θ in (3.4.15) to get

$$\begin{aligned}\Theta(t) &= i\lambda^3 t + i\lambda^2 W_t \\ &= \left[i((\xi^2-\eta^2)\xi - 2\eta^2\xi) - \eta(3\xi^2-\eta^2)\right] t + \left[-2\eta\xi + i(\xi^2-\eta^2)\right] W_t\,.\end{aligned}$$

The speed of the soliton follows the stochastic process $dc(t) = (3\xi^2 - \eta^2)\, dt +$

$2\xi dW_t$. The solitons with different values of ξ will have a different amplitude of the noise affecting its speed.

Stochastic NLS equation of type III

If we exchange the time and the stochastic time we arrive at the stochastic NLS equation of type III

$$du = i\left(u_{xx} + 2u|u|^2\right)dt + \left(u_{xxx} + 6u_x|u|^2\right) \circ dW_t\,. \tag{3.4.30}$$

In this case the noise is proportional to the CmKdV equation and the deterministic part is still the NLS equation.

We compute the solution by evaluating Θ in (3.4.15) and get

$$\begin{aligned}\Theta(t) &= i\lambda^2 t + i\lambda^3 W_t \\ &= \left[i(\xi^2 - \eta^2) - 2\eta\xi\right] t + \left[-\eta(3\xi^2 - \eta^2) + i((\xi^2 - \eta^2)\xi - 2\eta^2\xi)\right] W_t\,.\end{aligned}$$

As before, no change of variables can remove the noise from this equation. The speed of the soliton is given by $dc(t) = 2\xi\, dt + [3\xi^2 - \eta^2]dW_t$. Compared to the previous case, the noise is proportional to ξ^2 and vanishes for $\eta = \sqrt{3}\xi$ and the solution still has a stochastic phase. This noise could model randomly small higher order physical effects for example in optical fibres (see Agrawal (2007)).

The stochastic CH equations

In this last section, we will apply both types of deformations of the AKNS hierarchy which are the Sobolev and the stochastic deformations to obtain stochastic CH equations. First, the stochastic CH equation of type I is the same as the stochastic KdV of type I as it has a stochastic first order dispersion. This equation is a reduced version of the stochastic CH equation of Holm and Tyranowski (2016) in the case when their cylindrical noise is constant in space, see the next section 3.5 for more detail on this. Second, the stochastic CH equation of type II is more interesting and also reduces to

a stochastic peakon dynamics.

The stochastic CH equation of type II reads

$$dm = (2mu_x + m_x u)\,(dt + \circ dW)\,, \quad m = u - \alpha^2 u_{xx}\,. \tag{3.4.31}$$

Using the peakon ansatz $u(x,t) = \sum_i p_i e^{-|x-q_i(t)|/\alpha}$ in this equation, we obtain the following stochastic Hamiltonian equations for the variables $q_i(t)$ and $p_i(t)$

$$\begin{aligned} dq_i &= u(q_i,t)(dt + \circ dW_t) \\ dp_i &= -p_i \frac{\partial u(q_i,t)}{\partial q_i}(dt + \circ dW_t)\,. \end{aligned} \tag{3.4.32}$$

This is a finite dimensional Hamiltonian stochastic process in the sense of Bismut (1982), with Hamiltonian

$$h(p,q) = \frac{1}{2}\sum_{i,j} p_i p_j e^{-|q_i - q_j|/\alpha}\,, \tag{3.4.33}$$

and stochastic potential $\Phi = h$.

These two stochastic CH equations seem to be the only integrable stochastic cases reducing to a stochastic peakon system. Indeed, the higher order flows of CH are non-local and cannot be reduced to such a system. For the type II equation, we can ask the question of the crossing of peakons. This was numerically studied in Holm and Tyranowski (2016) for a generalisation of this noise introduced directly in the peakon system but which break the link with the stochastic CH equation. Following the original proof of Camassa and Holm (1993), we can prove that they will not cross for this stochastic CH equation. If we consider a simple two peakons system, there are two conserved quantities in the deterministic case, the total momentum $P = p_1 + p_2$ and the Hamiltonian $H = \frac{1}{2}(p_1^2 + p_2^2) + p_1 p_2 e^{|q_1 - q_2|/\alpha}$. The condition for the two peakons to cross is simply $q_1 = q_2$, and the Hamiltonian is in this case $H = \frac{1}{2}p_1^2 + p_2^2$. The Hamiltonian can therefore not have the same value as when the peakons are separated, thus contradicting the conservation of energy. This proof works only if the Hamiltonian H is a conserved quantity of the

stochastic equation which is the case here but not in Holm and Tyranowski (2016).

Finally, the type III noise for the CH equation only works when the deterministic equation is the complex mCH equation of Xia et al. (2015) and when the noise is proportional to the CH-NLS equation. The integrability of the CH-NLS equation is not clear as well as the physical interpretation of both equations, hence we leave the study of this stochastic equation for future works.

3.4.4 Summary and outlook

In this section, we did the obvious thing to do, that is to use the Lagrangian interpretation of integrable systems of section 3.2 to introduce noise into the integrable hierarchies. We saw that for the noise to be interesting, we need to slightly generalise the finite dimensional theory and allow non-linear stochastic potentials. A simple linear potential would only give a small set of the possible integrable stochastic equations. They consist of stochastic linear combinations of flows of the same hierarchy and with the same reduction from the general Lie algebra element. With the stochastic Euler-Poincaré equation, or equivalently the stochastic ZCR associated to a stochastic isospectral problem, we were able to derive exact soliton solutions via a stochastic inverse scattering transform. The solution of some simple stochastic flows could be found by a stochastic change of variable, but equations with higher order stochastic flows need the stochastic IST method to compute the time-related parameters of the solutions. Indeed, it is well-known that the shape of the solitons remains the same for each stochastic equation of the same hierarchy because the L operator of the spectral problem is the same for all. We only look at the one-soliton solution of these equations, but the stochastic IST could be used to derive all the other standard N-solitons solutions or breathers for example. We leave the investigation of the effect of the noise on these solutions for future works, as well as the numerical simulations of high-order stochastic flows. Indeed, the standard pseudo-spectral scheme, which we used to solve the stochastic equation, appeared to be unstable

when the stochastic flow cannot be absorbed by a direct change of x or t variable. This also suggests that studying the stability of the stochastic solitons would be an interesting and necessary next step. Methods similar to the energy-Casimir method for the stability of solitons could be extended to this stochastic context and provide information about their stochastic stability, similar to the results of section 2.4. The methods we have in mind originated in the early works Benjamin (1972), Bona (1975), Bona et al. (1987), Grillakis et al. (1987) and use conserved quantities to assess the orbital stability of solitons.

3.5 Stochastic PDE with stochastic transport

There is another way of introducing noise into PDEs similar to what we did in the first chapter for finite dimensional systems. This is the original idea of Holm (2015) for stochastic fluid dynamics which we will apply to 1+1 integrable PDEs. It uses the Hamiltonian structure of the equation associated with the reduction by symmetries of the diffeomorphism group. This geometrical interpretation of these equations is closer to the physical meaning of these equations in term of fluid dynamics. As we will see, the noise will correspond to having a stochastic transport velocity field. We will not explore the resulting equations in details, but we will just derive some of them and explore some of their properties in section 3.5.1. One interesting result of this section is the application of this method to the nonlinear Schrödinger equation in 3.5.2. Indeed, we need another way to introduce noise so that we obtain a non-trivial deformation. We will achieve that by using the Madelung transformation viewed as a momentum map to transfer the noise from fluid equations to the NLS equation.

3.5.1 Classical transport noise

We will not describe how to introduce noise in PDEs in details here as it is the same as we already did for finite dimensional systems in section 2.2 or for the AKNS hierarchy in section 3.4.2. Recall that the introduction of noise relies on the geometrical structure of the equation. Instead of using the Lagrangian structure related to the loop algebras to describe equations such as KdV or NLS as we did in section 3.4.2, we will use the standard Hamiltonian structures.

The idea behind this type of stochastic deformations is to modify the velocity variable to a stochastic velocity, or in other terms to append a stochastic element to the Hamiltonian. The stochastic Hamiltonian PDEs that we will consider will thus be of the form

$$du = J\frac{\delta h}{\delta u}dt + J\frac{\delta \Phi_i}{\delta u} \circ dW_t^i, \tag{3.5.1}$$

where J is the Hamiltonian structure, h the Hamiltonian and Φ_i the stochastic potential, usually taken to be linear in u to match the theory of reduction by symmetry. The important remark here is that the Hamiltonian structure J is the same for both the drift and the noise of the stochastic PDE (3.5.1).

In the case of KdV, we have two possible Hamiltonian structures, one being $J_1 = u\partial_x + \partial_x u + \partial_x^3$ with Hamiltonian $h = \frac{1}{2}\int u^2 dx$. The resulting stochastic KdV is thus

$$du = (3uu_x + u_{xxx})dt + \sum_i (2\sigma_i u_x + \sigma_{i,x}u + \sigma_{i,xxx}) \circ dW_t^i \,. \tag{3.5.2}$$

If instead we use the other structure $J_2 = \partial_x$, the noise will be an additive noise of the form $\sum_i \sigma_{i,x} \circ dW_t^i$. This is a well-studied equation, and we refer to de Bouard and Debussche (1998), de Bouard et al. (1999), Debussche and Printems (1999) and references therein for interesting works on the existence, well-posedness and numerical simulation of this equation. Similar investigations could be done with the stochastic KdV (3.5.2) instead, but we leave this for future works.

In the case of the NLS equation, the Hamiltonian structure is the canonical one, that is $J = \frac{i}{2}\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$, and the Hamiltonian is $h = \int (|u_x|^2 + |u|^4)dx$. The corresponding stochastic process is then, for $\Phi_i = \sigma_i(u + \overline{u})$,

$$idu = (u_{xx} + u|u|^2)dt + \sum_i \sigma_i \circ dW_t^i \,. \tag{3.5.3}$$

The choice of the stochastic potential is to ensure that it is a real valued function. This stochastic deformation produces an additive cylindrical noise studied by for example by de Bouard and Debussche (2003), Debussche and Di Menza (2002). Following the KdV example, we could use the other Hamiltonian structure of the NLS equation

$$K = \begin{bmatrix} -2u\partial_x^{-1}u & \partial_x + 2u\partial_x^{-1}\overline{u} \\ \partial_x + 2\overline{u}\partial_x^{-1}u & -2\overline{u}\partial_x^{-1}\overline{u} \end{bmatrix} . \tag{3.5.4}$$

This structure will give a non-trivial multiplicative noise with a nonlocal

amplitude in term of u. For this last reason, we will not investigate this stochastic process any further and rather consider another way of introducing a stochastic deformation to obtain a more tractable and interesting noise.

3.5.2 Stochastic NLS via the Madelung variables

The idea is the following. We want to add noise into the NLS equation such that we retain certain of its geometrical properties. As we saw in the previous sections, retaining either complete integrability or Hamiltonian structures does not give any new form of noise in the NLS equation. In this section, will explore the introduction of noise via another geometrical structure of the NLS equation, coming from the existence of the Madelung transformation. We first describe what is this transformation which maps the NLS equation to a particular type of fluid equations.

The Madelung transformation is due to Madelung (1927) and is a well-studied and well-used transformation that relates the NLS equation to a set of quantum hydrodynamical equations. We will not explain this connection with quantum hydrodynamics here, but rather describe the geometrical nature of this transformation. We will follow the recent derivation of the Madelung transformation by Fusca (2016) seen as a momentum map. We first recall the Madelung transformation is obtained by introducing the two real variables ρ and ϕ such that

$$u = \sqrt{\rho} e^{i\phi} \,. \tag{3.5.5}$$

ρ corresponds to the amplitude of the solution and ϕ its phase. We will also use the notation

$$v := \phi_x \quad \text{and} \quad \mu := \rho v = \mathrm{Im}(u_x \overline{u}) \,,$$

which can be seen as a current, or momentum variable. This change of

variables yields the relations

$$\begin{aligned} iu_t &= \left(i\frac{\rho_t}{2\sqrt{\rho}} - \sqrt{\rho}\phi_t \right) e^{i\phi} \\ u_x &= \left((\sqrt{\rho})_x + i\sqrt{\rho}v \right) e^{i\phi} \\ u_{xx} &= \left((\sqrt{\rho})_{xx} + i\sqrt{\rho}v_x + iv\frac{\rho_x}{\sqrt{\rho}} - \sqrt{\rho}v^2 \right) e^{i\phi} \\ 2u|u|^2 &= 2\rho\sqrt{\rho}e^{i\phi} \,. \end{aligned}$$

By considering separately the real and imaginary parts of the NLS equation, we obtain the equations for ρ and v as

$$\begin{aligned} \rho_t + (2\rho v)_x &= 0 \\ v_t + 2vv_x &= 2\rho_x + \left(\frac{(\sqrt{\rho})_{xx}}{\sqrt{\rho}} \right)_x . \end{aligned} \tag{3.5.6}$$

The first is a continuity equation for the density ρ and the second is the Euler equation for the fluid velocity v, with a particular forcing term on the right-hand side related to the theory of quantum hydrodynamics.

In the rest of this section, we will follow Fusca (2016) where it is explained in detail why the Madelung transformation is a momentum map. We will not give the proofs that are in this work and just use the results to introduce noise. Interestingly, the structure linked to this momentum map is the Clebsch construction with the Clebsch variable belonging to a symplectic space with a canonical symplectic form. This is the space for the NLS equation which has a canonical Poisson structure. We already know the canonical structure of the NLS equation, and we need the Hamiltonian structure of the equations (3.5.6). For this, we write this equation in term of the momentum $\mu := \rho v$ only

$$\partial_t \mu = -\left(\frac{1}{\rho}\mu^2 \right)_x - \rho \left(\frac{(\sqrt{\rho})_{xx}}{\sqrt{\rho}} \right)_x . \tag{3.5.7}$$

This equation together with the ρ equation is, in fact, a Lie-Poisson equation

on the semi-direct product space of compressible fluids, with Poisson bracket

$$\begin{aligned}\{F,G\}(\mu,\rho) = &\int \mu\left(\frac{\delta G}{\delta\mu}\partial_x\frac{\delta F}{\delta\mu} - \frac{\delta F}{\delta\mu}\partial_x\frac{\delta G}{\delta\mu}\right)dx \\ &+ \int \rho\left(\frac{\delta G}{\delta\mu}\partial_x\frac{\delta F}{\delta\rho} - \frac{\delta F}{\delta\mu}\partial_x\frac{\delta G}{\delta\rho}\right)dx\,.\end{aligned} \tag{3.5.8}$$

One can directly check that together with the Hamiltonian

$$H = \int\left(\frac{1}{2}\frac{\mu^2}{\rho} + \frac{1}{8}\frac{\rho_x^2}{\rho} + \frac{1}{2}\rho^2\right)dx\,, \tag{3.5.9}$$

the Lie-Poisson equation with the bracket (3.5.8) gives the equation of motion (3.5.7). The Madelung transformation maps this Lie-Poisson system to the Hamiltonian NLS equation.

We introduce some noise into the equation (3.5.6) as before by adding a stochastic element to the Hamiltonian (3.5.9). We use a linear stochastic potential of the form $\Phi_i = \mu\sigma_i$, where $\sigma_i(x)$ are real functions of the position x. We directly obtain the stochastic processes for ρ and μ

$$\begin{aligned}d\rho &= -(2\rho v)_x\,dt + \sum_i(\rho\sigma_i)_x \circ dW_t^i \\ d\mu &= -\left(\left(\frac{1}{\rho}\mu^2\right)_x + \rho\left(\frac{(\sqrt{\rho})_{xx}}{\sqrt{\rho}}\right)_x\right)dt + \sum_i(\mu_x\sigma_i + 2\mu\sigma_{i,x}) \circ dW_t^i\,,\end{aligned} \tag{3.5.10}$$

where the drift is from equation (3.5.6) and (3.5.7).

The key ingredient of Fusca (2016) is to interpret u as a complex half-density, denoted Ψ the variables (μ,ρ) as the momentum variable of compressible fluid, that is $\mu \in \mathfrak{X}^*(\mathbb{R})$ and $\rho \in C(\mathbb{R})$. In this case, the infinitesimal generator of the action of $\mathrm{Diff}(\mathbb{R})\circledS C(\mathbb{R})$ on half-densities is given by

$$\xi_\Psi(u) = -\frac{1}{2}uv_x - u_xv - i\alpha u\,, \tag{3.5.11}$$

where $\xi = (v,\alpha)$. With this action, the Madelung transform was shown to be a momentum map in Fusca (2016). We are now going to use this action to

introduce noise in the system NLS/Madelung, and see how the noise affects both the NLS and its equivalent quantum hydrodynamic equation (3.5.10).

Proposition 3.5.1. *The stochastic NLS equation, for* $\alpha = 0$

$$\begin{aligned} idu &= J_{NLS}\delta H dt + \frac{1}{2} i(u\sigma_x + 2u_x\sigma) \circ dW_t^i \\ &= J_{NLS}\delta H dt + -\frac{1}{2} i\xi_\Psi(u) \circ dW_t^i \\ &= J_{NLS}\delta H dt + J_{NLS}\xi_\Psi(u) \circ dW_t^i \,, \end{aligned} \tag{3.5.12}$$

where $J_{NLS} = -\frac{i}{2}\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$, *is equivalent to the stochastic fluid equation* (3.5.10).

Proof. The abstract proof is straightforward from the properties of the momentum map and the above construction. Nevertheless, we directly check it by computing

$$\frac{1}{2} i(u\sigma_x + 2u_x\sigma) = \frac{1}{2}\left(i\sqrt{\rho}\sigma_x + i2\sigma(\sqrt{\rho})_x - 2\sqrt{\rho}v\sigma\right) e^{i\phi} \,,$$

which then gives

$$\begin{aligned} d\rho &= (\cdot)dt + (\rho\sigma)_x \circ dW_t^i \\ dv &= (\cdot)dt + (v\sigma)_x \,, \end{aligned}$$

where we did not display the drift term for simplicity. In term of μ, we have

$$d\mu = (\cdot)dt + \mu_x\sigma + 2\mu\sigma_x \,,$$

which is the same as the stochastic fluid equation (3.5.10). ■

It turns out that this is the only noise term compatible with both the fluid and the NLS equation via the Madelung transform. We originally derived this noise term by hand by looking for such a compatible noise term, and we were able to obtain only this noise amplitude. Notice that it conserves the total mass of the NLS equation, as the overall phase is still a symmetry of the stochastic equation.

3.5.3 Summary and outlook

In this short section, we discussed the effect of the noise when directly added to the coadjoint action corresponding to the diffeomorphism group. The case of Camassa-Holm was used for the stochastic landmark dynamics in section 2.6. In the case of dispersive equations, there is an extra stochastic term coming from the central extension of the diffeomorphism group, or Virasoro algebra. This extension is similar to the extension used for the derivation of integrable hierarchies but has a different cocycle to produce the dispersive terms of for example the KdV equation. Compared to the stochastic hierarchy, the freedom in the choice of the noise amplitude is bigger as the noise is described by functions of x. Conditions on them for well-posedness and other existence results could be derived but are out of the scope of this work. From this freedom in the choice of the noise, we do not expect to retain any integrability, unless the noise fields are constant in space. In this case, we would recover the simple integrable equations of the stochastic hierarchies.

The second part of this section was devoted to the derivation a more interesting type of noise in the NLS equation. We used the Madelung transformation to transfer the advection noise of the associated fluid equations to the NLS equation to obtain a non-trivial multiplicative noise. From there, many research directions are open, but the question of integrability remain the same as for the KdV equation that we just discussed. Another way to introduce a non-trivial noise into the NLS equation would be to use the change of variables between the NLS equation and the Landau-Lifshitz equation for which we know how to introduce the stochastic advection, that is

$$d\mathbf{S}(x,t) = \mathbf{S}(x,t) \times \mathbf{S}_{xx}(x,t)\, dt + \sum_i \mathbf{S}(x,t) \times \sigma(x) \circ dW_t^i\,,$$

for a spin $\mathbf{S}(x,t) \in \mathbb{R}^3$, see for example the review Lakshmanan (2011). The change of variable from this equation to the NLS equation is more involved than the Madelung variables and we leave this for future investigations.

3.6 The reduced Maxwell-Bloch equations

In this very last section, we will present the results of Arnaudon and Gibbon (2017) on the reduced Maxwell-Bloch equations (RMB). These equations are completely integrable and appear as negative flows of the AKNS hierarchy but with non-zero poles in the M operator of the isospectral problem. One of them was derived and studied by Caudrey et al. (1974), Dodd et al. (1982), Eilbeck et al. (1973), Gibbon et al. (1973) in the early days of integrable systems but has not attracted much attention from the community since then, apart from isolated paper such as Aiyer (1983) where the Hamiltonian and recursion operators were derived. From our perspective of deformations of integrable systems, these equations provide us with another type of deformations preserving complete integrability. As we will see, the original RMB equation depends on a parameter ω_0 and reduces to the sine-Gordon equation in the limit $\omega_0 \to 0$. This parameter can thus be interpreted as characterising the amount of deformation with respect to the sine-Gordon equation. The first main result of this section is to obtain a larger class of such deformations of integrable systems. We will obtain four non-equivalent integrable equations in section 3.6.1 which reduce either to the sine-Gordon or the sinh-Gordon equation. They differ only by a few signs and we will justify the name hyperbolic RMB equations for the two equations which reduce to the sinh-Gordon equation. We will then study an interesting finite dimensional reduction of these equation, related to the Lorenz 63 model in section 3.6.2 and used it to analyse the stability of the constant solutions of the full PDE in section 3.6.3. In the last part 3.6.4, we will show that one of the new equations which we will call the hRMB equation can be obtained from the theory of strongly correlated Bose-Einstein condensates.

3.6.1 The four integrable reduced Maxwell-Bloch equations

In this section, we will first present the four reduced Maxwell-Bloch equations and show how to derive them as negative flows of the AKNS hierarchy.

The four equations

The RMB equations are written for four independent space-time real scalar fields E, N, P and Q for a one-dimensional space $x \in \mathbb{R}$. These four integrable equations have only sign differences between then and can be written together as

$$\begin{aligned} E_x &= P \\ P_t &= EN + \sigma_2 \omega_0 Q \\ N_t &= -\sigma_1 EP \\ Q_t &= -\omega_0 P\,, \end{aligned} \tag{3.6.1}$$

where $\sigma_{1,2} = \pm 1$ selects between the four equations and ω_0 is a real parameter. The four equations can be listed such that

- for $\sigma_1 = \sigma_2 = 1$,we have the original RMB equation,
- for $\sigma_1 = -1, \sigma_2 = 1$, we have the new hyperbolic RMB equation for strongly correlated Bose-Einstein condensate,
- and for $\sigma_1 = \pm 1, \sigma_2 = -1$, we have two new equations with a different M operator for the zero curvature relation but which do not have a clear physical interpretation yet.

The RMB equations can also be seen as a combination of two rotations, one in the (P, Q) variables with rotation speed proportional to ω_0, and the other in the (P, N) variables with a rotation speed proportional to the E variable. In this view, the parameters $\sigma_{1,2}$ change these elliptic motions to hyperbolic dynamics.

As we already mentioned the parameter ω_0 can be seen as a deformation parameter and in the limit $\omega_0 \to 0$ we recover the sine-Gordon equation for the case $\sigma_1 = 1$ and the sinh-Gordon for $\sigma_1 = -1$. In this limit, the Q variable can be discarded as well as the σ_2 parameter. For the case $\sigma_1 = 1$, the change of variables $E = \phi_x$, $P = \sin(\phi)$ and $N = \cos(\phi)$ where we introduced another real scalar field ϕ gives the sin-Gordon equation of the

form

$$\phi_{xt} = 2\sin(\phi)\,. \tag{3.6.2}$$

For $\sigma_1 = -1$, the change of variable is $E = \phi_x$, $P = \sinh(\phi)$ and $N = \cosh(\phi)$ and the sinh-Gordon equation is

$$\phi_{xt} = 2\sinh(\phi)\,. \tag{3.6.3}$$

These two equations are important in physics and in mathematics and were at the origin of the developments of integrable systems. Indeed, soon after that the KdV and the NLS equations were solved by Zakharov and Shabat (1972), the Sine-Gordon equation was shown to be integrable with kink solitons by several authors, including Ablowitz et al. (1973a,b), Caudrey et al. (1973).

We need to mention the existence of an important conserved quantity, especially in physics. On can check that the RMB equations (3.6.1) conserve the quantity

$$\eta := P^2 + \sigma_2 Q^2 + \sigma_1 N^2\,. \tag{3.6.4}$$

In the classic RMB case, this corresponds to the well-known Bloch sphere of optics or quantum mechanics. In the other cases, this sphere becomes a hyperboloid with various orientations.

The complete integrability of these equations is obtained by writing the zero curvature relation associated with the spectral problem

$$\begin{aligned} \Psi_x &= L_{\sigma_1}\Psi \\ \Psi_t &= M_{\sigma_1,\sigma_2}\Psi\,, \end{aligned} \tag{3.6.5}$$

where $\Psi = (\psi_1, \psi_2)^T$ is the scattering wavefunction for the spectral parameter λ. The L operator is the usual Zakharov-Shabat spectral operator which we

recall here

$$L_{\sigma_1} = \lambda \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} + \begin{bmatrix} 0 & E \\ -\sigma_1 E & 0 \end{bmatrix} . \tag{3.6.6}$$

For the RMB equations the field E is real, thus this spectral problem corresponds to the mKdV reduction of the AKNS hierarchy rather than the NLS reduction. The M operator has two different forms depending on the value of σ_2. For $\sigma_2 = 1$, we have

$$M_{\sigma_1,+} = \frac{1}{2(\lambda^2 - \omega_0^2)} \left(\lambda \begin{bmatrix} -iN & P \\ \sigma_1 P & iN \end{bmatrix} - \omega_0 \begin{bmatrix} 0 & Q \\ -\sigma_1 Q & 0 \end{bmatrix} \right) , \tag{3.6.7}$$

where the poles in the complex plane of the spectral parameter is at $\pm\omega_0$. For $\sigma_2 = -1$, the M operator has a double pole at ω_0 and is given by

$$\begin{aligned} M_{\sigma_1,-} = \frac{1}{2(\lambda^2 - \omega_0^2)(\lambda - \omega_0)} \Bigg(& i\lambda^2 \begin{bmatrix} 0 & P \\ -\sigma_1 P & 0 \end{bmatrix} - \\ & - \lambda\omega_0 \begin{bmatrix} 0 & Q \\ \sigma_1 Q & 0 \end{bmatrix} + i\omega_0^2 \begin{bmatrix} N & 0 \\ 0 & -N \end{bmatrix} \Bigg) . \end{aligned} \tag{3.6.8}$$

One can check that these operators satisfy the following ZCR

$$\partial_t L_{\sigma_1} - \partial_x M_{\sigma_1,\sigma_2} + [L_{\sigma_1}, M_{\sigma_1,\sigma_2}] = 0 , \tag{3.6.9}$$

if the fields satisfy the corresponding RMB equation. The derivation of these equations from the negative flow of the AKNS hierarchy is in Appendix C. We obtain these equations from a generic form for the negative flow and successively apply reductions to obtain the non-equivalent RMB equations.

Soliton solutions and inverse scattering

We begin by solving the RMB equations by hand using the travelling wave ansatz $E(x,t) = E(t - c^{-1}x)$, where c is the speed of the wave. With a non-vanishing boundary condition only for N, that is $N(\pm\infty) = N_\infty$, we can

reduce the RMB equations to the following second order nonlinear ODE

$$E_{xx} = -E\left(\frac{1}{2}\sigma_1 E^2 + cN_\infty + \sigma_2\omega_0^2\right). \tag{3.6.10}$$

This ODE is close to the usual travelling wave equation for KdV or NLS, with the extra feature of the ω_0 term and the possible sign modifications. The sign in front of the E^3 term will select between a sech-profile and a tanh-profile. In $\sigma_1 = 1$, that is the classical RMB equations, we obtain by setting $N_\infty = -1$ the soliton solution of Gibbon et al. (1973)

$$E(x,t) = E_0 \operatorname{sech}\left(\frac{1}{2}E_0\left(t - \frac{4}{E_0^2 + 4\sigma_2\omega_0^2}x\right)\right), \tag{3.6.11}$$

where E_0 is the amplitude parameter of the soliton.

A direct integration of the travelling wave ODE (3.6.10) would give the kink soliton of the form

$$E(x,t) = \pm E_\infty \tanh\left(2E_\infty\left(t - \frac{N_\infty}{E_\infty^{-2}/2 - \sigma_2\omega_0^2}x\right)\right), \tag{3.6.12}$$

but this is not a valid solution of the HRMB equation, due to unusual boundary conditions. In the hyperbolic case, we have $Q(\pm\infty) = \frac{E(\pm\infty)N_0}{\sigma_2\omega_0}$, which takes opposite values at $\pm\infty$. Using this boundary condition, the travelling wave ODE is

$$E_{xx} = E\left(\frac{1}{2}E^2 - cN_\infty - \sigma_2\omega_0^2\right) + \operatorname{sgn}(x)cE_\infty N_0\,, \tag{3.6.13}$$

where the last constant is positive for $x > 0$, and negative for $x < 0$.

Explicit solutions of this equation are difficult to obtain hence we compare in Figure 3.8 the tanh solution with a numerical solution of (3.6.13). They can be obtained by considering an idealised model for (3.6.13)

$$E''(x) = \nabla V(E(x))\,, \tag{3.6.14}$$

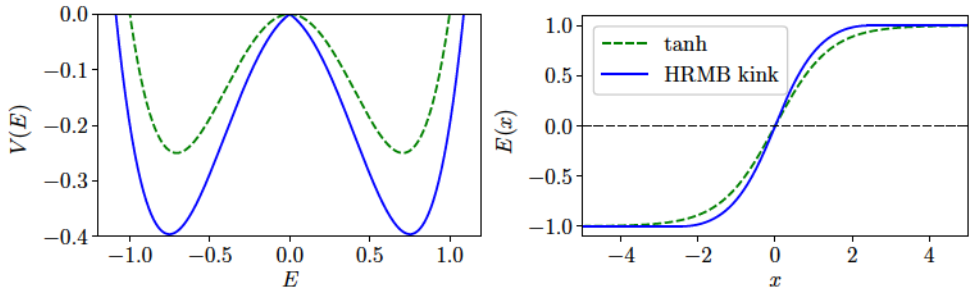


Figure 3.8: In the top panel we display the potential function $V(E)$ in (3.6.15) describing the travelling wave solutions of the HRMB equation in blue, and compare with the potential corresponding to the tanh solution in dashed green. In the bottom panel, we compare the solutions, numerically integrated from (3.6.14).

with the potential function

$$V(E(x)) = E(x)^4 - E(x)^2 - 0.2\mathrm{sgn}(x)E(x)\,. \tag{3.6.15}$$

In contrast with the classical case, if $\sigma_1 = -1$ and $\sigma_2 = 1$ the speed of the kink soliton can propagate in either directions and is standing if $\omega_0^2 E_\infty^2 = \frac{1}{2}$.

In principle, these solutions can be derived from the theory of inverse scattering transform using the M and L operators. In the case $\sigma_1 = 1$, the M operator is anti-Hermitian with real valued E. The spectrum can thus contain discrete eigenvalues either purely imaginary, or in complex pairs, i.e. two eigenvalues λ_1 and λ_2 such that $\lambda_2 = -\overline{\lambda}_1$. For $\sigma_1 = -1$, there is no discrete eigenvalue unless the boundary conditions at $\pm\infty$ do not vanish. In this case, the theory of inverse scattering is more involved and gives dark solitons or kinks. We will not consider this case here and only study the two cases where $\sigma_1 = 1$.

As for the stochastic IST in section 3.4, we follow the exposition of Ablowitz and Segur (1981). The scattering problem is the standard Zakharov-Shabat problem, and we recall the one-soliton solution

$$u(x,t) = 2\eta e^{2i\mathrm{Im}(m_\infty)t}\,\mathrm{sech}(2\eta x + 2\mathrm{Re}(m_\infty)t)\,, \tag{3.6.16}$$

where $\lambda = i\eta$ and m_∞ is the diagonal value of the M operator evaluated at

$x = \infty$. For $\sigma_2 = 1$, we have

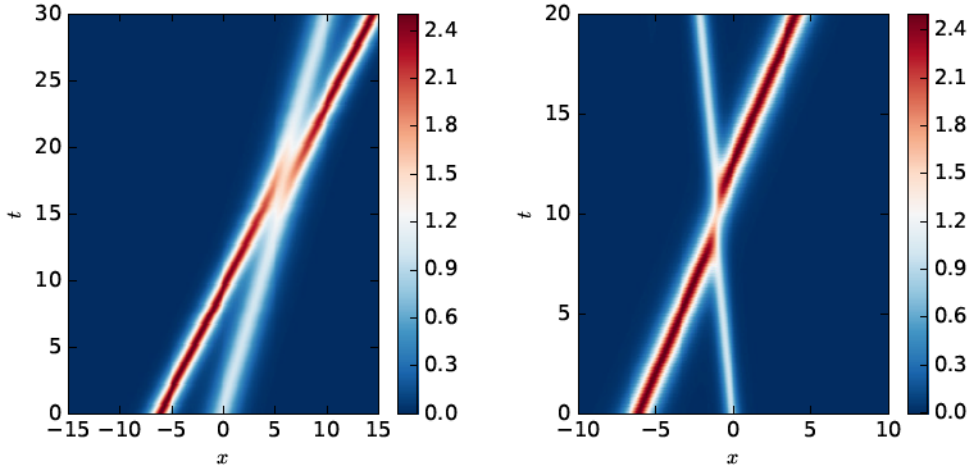
$$m_\infty(i\eta) = \frac{\eta}{2(\eta^2 + \omega_0^2)}\,, \tag{3.6.17}$$

which is real and gives the solution (3.6.11). For $\sigma_2 = -1$, we instead have

$$m_\infty(i\eta) = \frac{-\eta\omega_0^2 + i\omega_0^3}{2(\eta^2 + \omega_0^2)^3)}\,, \tag{3.6.18}$$

which contains an imaginary part. This result is in contradiction with the fact that E is real so it is not a valid solution of the RMB equation with $\sigma_2 = -1$. We will not investigate this further here as it is out of the scope of this work and involves a deeper analysis of this system with the IST. One possible solution to this problem could be to use two eigenvalues in conjugate pairs instead of a single real eigenvalue. In this case, the solutions would correspond to breathers which are one-soliton modulated by a sine function, and typical of the sine-Gordon equation. It may be that the solution (3.6.11) for $\sigma_2 = -1$ is a particular case of a breather which behaves as a one-soliton.

To illustrate the soliton solutions of the RMB equation we display in figure 3.9 two collisions of RMB solitons with the same initial conditions but $\sigma_2 = \pm 1$. We used $A = 1$ and $A = 2.5$ with $\omega_0 = 0.6$. Notice that the initial conditions are not the same as they depend on σ_2, see equation (3.6.11). The standard RMB equation with $\sigma_2 = 1$ has only right going solitons which overtake each other and exchange their momenta. For $\sigma_2 = -1$, one of the solitons has a negative velocity, and we observe a head-on collision which would never be a solution of the standard RMB equation. These simulations were obtained by directly solving the RMB equation where the equation for E is replaced by the forced travelling wave equation $E_x + E_t = P$. We used the Python package Dedalus (Burns et al. (2016)) to perform these simulations. It is a spectral scheme with a Runge-Kutta time integrator.



(a) E (b) E

Figure 3.9: The collision of two RMB solitons for $\sigma_1 = 1$ and $\sigma_2 = 1$ is displayed in the left and for $\sigma_2 = -1$ on the right. It shows a standard overtaking for the RMB case and produces a head-on collision in the new integrable equation.

3.6.2 Finite dimensional reduction

Before looking at the modulational stability or instability of the RMB equation in the next subsection, we study a particular finite dimensional reduction of these equations. For $c \neq 0$, we assume that the fields are constant in space, which amounts in replacing the first equation of (3.6.1) with $cE_t = P$. By setting the constant of integration to 0 for simplicity (having it would complicate the results but would be more general), we can set $E = -\frac{1}{c\omega_0}Q$ and obtain the finite dimensional system

$$\begin{aligned} P_t &= -\frac{1}{c\omega_0}QN + \sigma_2\omega_0 Q \\ N_t &= \frac{\sigma_1}{c\omega_0}QP \\ Q_t &= -\omega_0 P\,. \end{aligned} \tag{3.6.19}$$

This system has two independent conserved quantities

$$H = \frac{\sigma_1}{2c\omega_0} Q^2 + \omega_0 N \tag{3.6.20}$$

$$C = \frac{1}{2} P^2 + \frac{\sigma_1}{2} (N - \sigma_2 c \omega_0^2)^2 \,, \tag{3.6.21}$$

as well as the Bloch sphere, or the Bloch hyperboloid, but is dependent of H nor C. It turns out that the ODE (3.6.19) is the non-dissipative Lorenz 63 model of Lorenz (1963), which we will call here the ODE-RMB. This equation has not attracted much attention in the literature, but we can mention the monograph Sparrow (1982), and in particular, the Appendix K where the non-dissipative Lorenz system is studied in the limit $\omega_0 \to 0$, that is the sin/sinh-Gordon reduction. A similar reduction to an ODE, but from the self-induced transparency equation to the complex Lorenz equation, was derived and investigated by Fowler et al. (1982, 1983), Gibbon and McGuinness (1980, 1982).

3.6.3 Modulational instability

We study here the stability of the constant solution of the four RMB equations. The constant solution depends on two constants Q_0 and N_0 and reads

$$\begin{aligned} &Q(x,t) = Q_0, \quad N(x,t) = N_0 \\ &P(x,t) = 0 \quad \text{and} \quad E(x,t) = -\sigma_2 \frac{Q_0}{N_0} \,. \end{aligned} \tag{3.6.22}$$

As we did for the CH-NLS equation, we will study the linear stability of this family of solutions by linearising the RMB equation around (3.6.22). We use perturbations of the form $N = N_0 + \delta N$, and the same for all the fields to obtain the linearised equation

$$\begin{pmatrix} c\partial_t + \partial_x & \alpha & 0 & 0 \\ N_0 & \partial_t & -\sigma_2 \frac{Q_0}{N_0} & \sigma_2 \omega_0 \\ 0 & \sigma_1 \sigma_2 \frac{Q_0}{N_0} & \partial_t & 0 \\ 0 & -\omega_0 & 0 & \partial_t \end{pmatrix} \begin{pmatrix} \delta E \\ \delta P \\ \delta N \\ \delta Q \end{pmatrix} = 0 \,. \tag{3.6.23}$$

By assuming the perturbations to be plane waves of the form $\delta N \propto \cos(\kappa x + \omega t)$ for all the fields and by a direct substitution into the linearised equations, the system (3.6.23) has a solution only if the following relation holds

$$\begin{aligned} -c\omega^3 - \kappa\omega^2 + \left(c\omega_0 \frac{\sigma_2 N_0^2 + \sigma_1 Q_0^2}{N_0^2} - \alpha N_0 \right) \omega \\ + \kappa\omega_0 \frac{\sigma_2 N_0^2 + \sigma_1 Q_0^2}{N_0^2} = 0 \,. \end{aligned} \tag{3.6.24}$$

This implicit linear dispersion relation is too involved for an analytical treatment. We can nevertheless see that with $\sigma_2 = -\sigma_1$, the choice $N_0^2 = Q_0^2$ corresponds to a transition between two different regimes. We thus plot the dispersion relation $\omega(\kappa)$ for several values of the parameters σ_1, σ_2 an Q_0 in Figure 3.10 computed numerically. For these computations, we used the values $N_0 = -\sigma_1$, $c = -1$, $\omega_0 = 0.5$ and $\alpha = 1$.

When $\sigma_1 = -\sigma_2$, there is a transition at $Q_0 = N_0$ from a stable to an unstable regime. For $\sigma_1 = 1$, the stable regime has a band of modulational instability for low wavenumbers, and for $\sigma_1 = -1$ the solution is stable for all wavenumbers. In more details, we have the following analysis of different regimes.

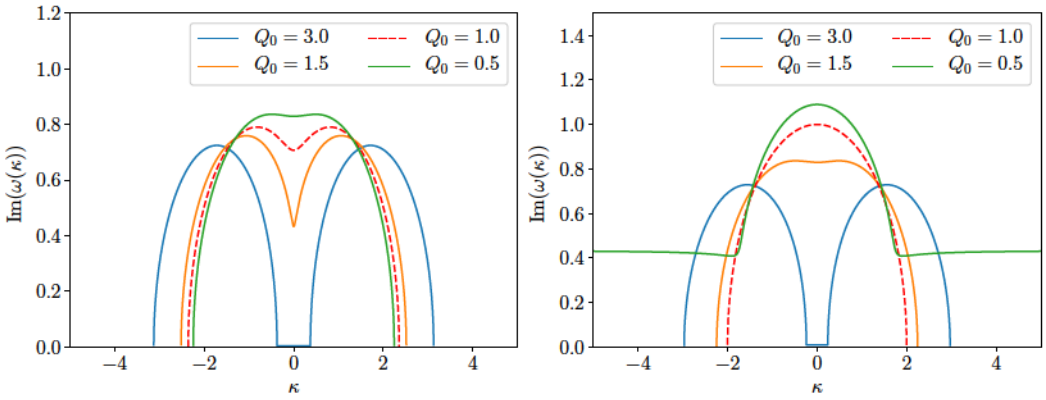
(a) $\sigma_1 = 1, \sigma_2 = 1$: Instabilities for low wavenumbers, corresponding to a modulational instability (MI). Depending on the value of the parameters, and in this case for large Q_0, a stable low-frequency region can exist. Increasing c would have a similar effect.

(b) $\sigma_1 = 1, \sigma_2 = -1$: There are two cases:

- if $Q_0^2 \leq N_0^2$, MI regime, similar to the previous case $\sigma_1 = \sigma_2 = 1$;
- if $Q_0^2 > N_0^2$, MI regime together with a smaller amplitude instability for all frequencies.

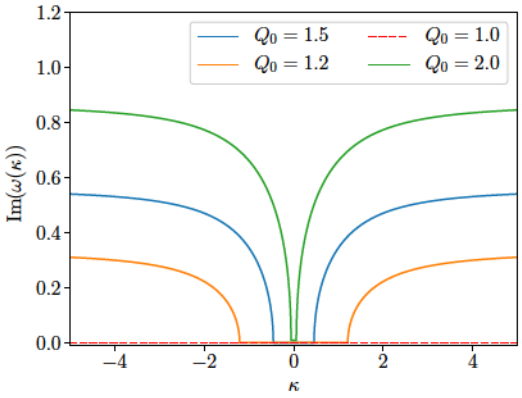
(c) $\sigma_1 = -1, \sigma_2 = 1$: There are two cases:

- if $Q_0^2 \leq N_0^2$, the constant solution is stable for all frequencies;

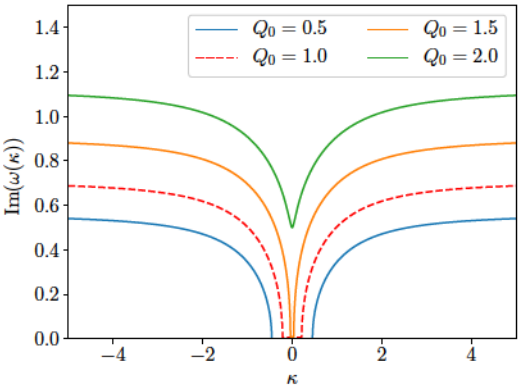


(a) $\sigma_1 = 1, \sigma_2 = 1$: MI regime

(b) $\sigma_1 = 1, \sigma_2 = -1$: MI regime or unstable



(c) $\sigma_1 = -1, \sigma_2 = 1$: Stable or unstable



(d) $\sigma_1 = -1, \sigma_2 = -1$: Unstable

Figure 3.10: In these four panels, we display the instability gain of the four RMB equations, for some values of Q_0, $N_0 = -\sigma_1$, $c = 1$, $\omega = 0.5$ and $\alpha = 1$. We observe two regimes depending on the relative value of Q_0 and N_0 if $\sigma_1 = -\sigma_2$. In particular for $\sigma_2 = 1$, there is a modulational instability regime for $\sigma_1 = 1$, and a stable regime for $\sigma_1 = -1$.

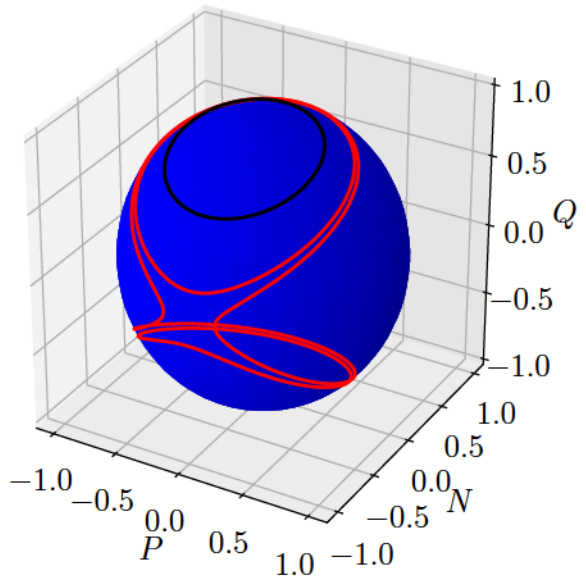
- if $Q_0^2 > N_0^2$, the constant solution is unstable for all frequencies, except for a small region of low frequencies.

(d) $\sigma_1 = -1, \sigma_2 = -1$: the solution is unstable for almost all frequencies and for any values of Q_0. As in the previous case, there is a small stable band at low frequencies, for small values of Q_0.

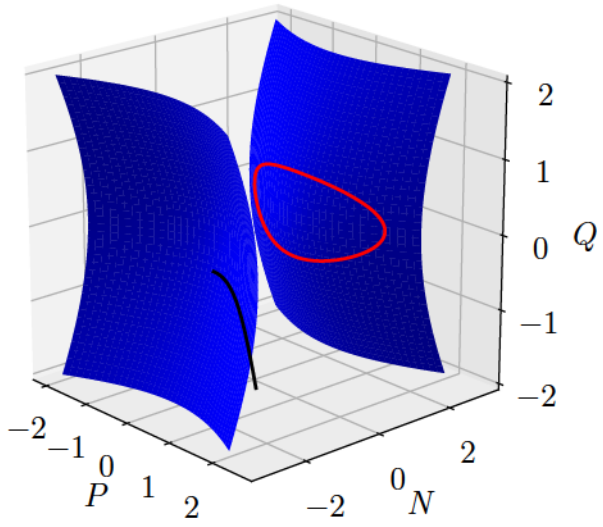
Interestingly, the finite dimensional reduction to the ODE-RMB equation (3.6.19) derived before can be used to shed light on these regimes of stability or instability. First, for $\sigma_1 = \sigma_2 = 1$, we know that the dynamics is restricted to the Bloch sphere (3.6.25) intersected by a cylinder in the Q directions (given by the quantity C in (3.6.21)). Notice that this shape can also be seen as the Bloch sphere intersecting the parabolic sheet in the P direction (given by H in (3.6.20)). The physical fix point corresponds in the model to the position $Q = 0$, as we set the constant of integration in the finite dimensional reduction to 0. This fix point is hyperbolic and hence corresponds to the MI regime as the dynamics leave this unstable position, make a long trajectory to eventually pass nearby it several times. The other fix points of the finite dimensional reduction are $P = 0, \quad Q = \pm Q_0 \quad$ and $\quad N = \sigma_2 c\omega_0^2$. They are stable fix points but not physical as we have a constraint on the value of the fields dictated by physics, see section 3.6.4. To illustrate this discussion, we plot some stable and unstable orbits in Figure 3.11(a).

For the hyperbolic RMB, that is for $\sigma_1 = -1$ and $\sigma_2 = 1$, the dynamic takes place along the intersection of a hyperboloid in the N direction given by (3.6.25) with a parabolic sheet in the same direction (given by H in (3.6.20)). The hyperboloid can have one or two sheets, depending on the relative value of Q and N. In the two sheets case (that is for $Q_0^2 \leq N_0^2$), if we pick $N_0 = 1$ the dynamics is restricted to the stable sheet where the parabola is bounded from below. This regime corresponds to the stable regime of the full HRMB. If the hyperboloid is of one sheet, or if $N_0 = -1$, the trajectory can reach the unbounded region of the parabola, which corresponds to the unstable regime of the HRMB equation. In Figure 3.11(b), we illustrate this discussion by showing a stable and unstable orbit of the ODE-RMB equation (3.6.19).

For the RMB equation, this regime of modulational instability was known for



(a) Bloch sphere with $\sigma_1 = 1, \sigma_2 = 1$



(b) Bloch hyperboloid with $\sigma_1 = -1, \sigma_2 = -1$

Figure 3.11: In this figure, we display typical orbits of ODE-RMB equation (3.6.19) on the Bloch sphere (left) and Bloch hyperboloid (right), obtained as a finite dimensional reduction of the RMB equations. On the left panel, the standard RMB equation with $\sigma_1 = \sigma_2 = 1$ of quantum optics has solutions restricted to the Bloch sphere, with stable equilibrium in the center of the black orbit, and unstable saddle point near the center of the red orbit. This position corresponds to the MI regime of the RMB equation. On the right panel, we show two orbits of the reduction of the HRMB equation of strongly correlated Bose-Einstein condensates, which are, in this case, restricted to a two-sheet hyperboloid. The stationary solution at the center of the red orbit is stable, but unstable near the black orbit. These two scenarios correspond to the stability and instability of the full HRMB equation.

a long time and corresponds to the focusing nonlinear Schrödinger equation instability. In the nonlinear regime of the modulational instability, when the linear approximation of the integrable equation is not valid anymore, the nonlinearities prevent the solution to blow up and form a train of interacting pulses. This nonlinear solution can be understood as a regime of integrable turbulence and can produce high amplitude waves, called rogue waves. We only refer to Zakharov (2009) for the notion of integrable turbulence and to Agafontsev and Zakharov (2015) for the formation of rogue waves for the nonlinear Schrödinger equation.

3.6.4 Application to strongly correlated Bose-Einstein condensates

The original RMB equation was derived in the context of quantum optics and in particular within the theory of self-induced transparency by Caudrey et al. (1974), Dodd et al. (1982), Eilbeck et al. (1973), Gibbon et al. (1973). It turns out that a similar derivation holds for the hyperbolic RMB equation (with $\sigma_1 = -1$ and $\sigma_2 = 1$) but in the context of Bose-Einstein condensates (BEC). In this section, we will show how to obtain the hRMB equation by applying several approximations of the complete physical model. We first describe the hyperbolic Bloch equation which we will then couple to the Gross-Pitaevskii equation to obtain the hRMB equation.

The hyperbolic Bloch equations

The RMB and hRMB equations contain two distinct parts. The equation for the variables (P, Q, N) is an ODE coupled with E. This coupling makes the whole system a $1+1$ PDE. In the case of the RMB, this ODE is the Bloch equation, which preserves the Bloch sphere

$$P^2 + \sigma_2 Q^2 + \sigma_1 N^2 = \eta\,, \tag{3.6.25}$$

and for the hRMB, it is the hyperbolic Bloch equation which instead preserves the hyperboloid $P^2 + Q^2 - N^2$. This finite dimensional hyperbolic Bloch

equation was recently derived in the context of strongly interacting Bose-Einstein condensates (BEC) by Kira (2015b). We will only summarise this derivation here, based on the method of cluster-expansion that we used in our study of the Fokker-Planck equation for stochastic landmark matching in section 2.6.

The hyperbolic Bloch equation (HBE) describes the dynamics of bosons evaporated from a BEC and forming a gas surrounding it. This gas is still interacting with the BEC as some of its atoms can be captured back into the BEC while others can evaporate from it. There is, therefore, a time-dependent process of exchanging atoms between the BEC and this surrounding gas. This effect only appears in strongly correlated BEC, where the internal correlations are significant enough to evaporate parts of its constituents. This interaction is characterised by a length scale a_{scatt} which in the case of strongly correlated BEC can be considered as $a_{\text{scatt}} \to \infty$. In the Gross-Pitaevski description of the BEC, this coefficient multiplies the nonlinear cubic term which represents the nonlinear interactions of the atoms in the BEC.

In quantum mechanics, the description of a Bose gas is written in terms of bosonic operators B_k for wavenumber k with the commutation relations

$$[B_k, B_{k'}^\dagger] = \delta_{k,k'}, \quad [B_k, B_{k'}] = [B_k^\dagger, B_{k'}^\dagger] = 0\,,$$

where $B^\dagger$ is the conjugate transpose of B. Each operator has a momentum $\hbar k$ and an energy

$$E_k = \frac{\hbar^2 k^2}{2m}\,,$$

where m is the mass of the bosons and $\hbar$ the Planck constant. The case $k = 0$ corresponds to the lowest energy level and describes the condensed atoms in the BEC. All the other states with $k \neq 0$ are the evaporated atoms. This approximation allows us to consider a single wavenumber and to only use four bosonic operators $B_+, B_-, B_+^\dagger$ and $B_-^\dagger$. The Hamiltonian is a simplification

of the original one derived in Kira (2015a,b) and is given by

$$\hat{H} = \omega_0 B_+^\dagger B_+ + \Omega(B_- B_+ + B_-^\dagger B_+^\dagger)\,, \tag{3.6.26}$$

where ω_0 is the transition energy and

$$\Omega = V N_c = 4\pi\hbar^2 m^{-1} a_{\text{scatt}} N_c \tag{3.6.27}$$

is the quantum-depletion source. This term is proportional to the number of condensed atoms

$$N_c = N_{\text{tot}} - f\,. \tag{3.6.28}$$

We denote by N_{tot} the total number of atoms in the system. Notice that the first term is the usual quantum harmonic oscillator, and the second term is a source and depletion term for the pair of atoms with opposite momenta.

Now recall that these operators would not correspond to any observable quantities and we needs to instead compute the time evolution of their expectation, following the dynamics prescribed by the Hamiltonian (3.6.26). This is given by an equation similar to the Fokker-Planck equation, but where the dissipative term related to the noise is replaced by the particle interactions. As for the landmark dynamics, the issue is that this equation is a PDE on a high dimensional space and must be approximated by a finite set of ODEs to be solved, even numerically. This can be achieved by using the cluster expansion approach. We will use standard notation for expectations, that is

$$\left\langle B_k B_{k'}^\dagger \right\rangle = \langle B_k \rangle \left\langle B_{k'}^\dagger \right\rangle + \Delta \left\langle B_k B_{k'}^\dagger \right\rangle .$$

There is one trick that one needs in order to be able to derive an approximate equation. It consists in a change of representation of these operators such that the BEC becomes the ground state of the system, needed as the BEC is a fully correlated system which would require all orders in the cluster expansion. This is done by choosing a particular basis for the operation of taking expectations, see Kira (2014) for all the details. In more dilute BEC, a standard truncation suffices, but not here. The result of this operation is

that the expectation simplifies as

$$\begin{aligned}\left\langle B_k^\dagger B_k \right\rangle &= \Delta \left\langle B_k^\dagger B_k \right\rangle := f_k \\ \langle B_k B_{-k} \rangle &= \Delta \left\langle B_k^\dagger B_{-k} \right\rangle := p_k \,,\end{aligned}$$

where we defined the two fields (f_k, p_k). The real valued quantity f_k counts the number of atoms in the state k, and the complex quantity p_k counts the number of pairs of atoms ejected from the BEC state B_0 to two evaporated states with opposite momenta. Its complex conjugate is the reverse process. Higher order processes can be considered and will correspond to higher order correlations in the cluster expansion method. We will not show how to derive the equations for f_k and p_k but only refer to the original derivation of Kira (2015b) as it is lengthy and not relevant for the present study. We will not consider any higher order interactions other than f_k and p_k, and we will approximate the interaction potential between bosons by a contact potential, that is a Dirac delta function. This is equivalent to expressing the Fourier transform of the contact potential as

$$V_k = 4\pi\hbar^2 m^{-1} a_{\text{scatt}} \,,$$

and removing the wavenumber dependence of f and s, as this assumption removes the terms coupling the dynamics of atoms at different wavenumbers.

The simplified HBE equation is then, from the Hamiltonian (3.6.26), given by

$$\begin{aligned} i\dot{p} &= \omega_0 s + (2f+1)\Omega \\ \dot{f} &= 2\,\mathrm{Im}\left(\Omega p^*\right) \,, \end{aligned} \tag{3.6.29}$$

This equation is very close to the usual Bloch equation except for a sign flip in the f equation, giving the hyperbolic character of the equation which preserves the hyperboloid $\eta = (f + \frac{1}{2})^2 - |p|^2$ instead of the Bloch sphere.

Coupling with the Gross-Pitaevskii equation

The second step needed to obtain the hRMB equation is to couple the HBE equation (3.6.29) with the dynamical equation describing the BEC itself. It makes sense to couple the HBE with the Gross-Pitaevski (GP) equation as both are equations for the expected value of some quantum operators representing the dynamics of the BEC. We will not derive the GP equation here but just use it and include a standard source term to model the flux of atoms between the BEC and the surrounding gas.

We will use a simple version of the GP equation obtained using the standard local-density approximation (LDA) which consists of studying the BEC dynamics locally, thus neglecting the exterior trapping potential. We will also consider a one-dimensional condensate, which could still be physically relevant in appropriate experiments, see for example Burger et al. (1999). We obtain the GP equation coupled to the field s as

$$i\hbar\psi_t + \alpha\psi_{xx} + \beta\psi|\psi|^2 = i\beta\mathrm{Im}(p^*)\psi\,, \tag{3.6.30}$$

where ψ is a complex-valued wavefunction describing the BEC, $\alpha = \hbar^2/2m$ and $\beta = 8\pi a_{\mathrm{scatt}}\alpha$. We now write the GP equation in the amplitude phase variables using the Madelung transformation

$$\psi(x,t) = \sqrt{n(x,t)}e^{i\phi(x,t)}\,,$$

for the amplitude $n(x,t)$ and the phase $\phi(x,t)$. This transformation which we already used in the previous section gives the equations

$$n_t + 2(n\phi_x)_x = \beta\mathrm{Im}(p^*)n \tag{3.6.31}$$

$$\phi_t = \alpha\left(\frac{(\sqrt{n})_{xx}}{\sqrt{n}} - \phi_x^2\right) + 2\beta n\,. \tag{3.6.32}$$

Using the LDA again, we decompose the amplitude n as $n(x,t) = n_0 + n_1(x,t)$, for a constant background n_0 and a small amplitude perturbation $n_1(x,t)$, i.e. $|n_1| \ll n_0$. The steady solution of this equation is obtain by setting $n_1 = 0$ whereas a time independent phase $\phi_0(x)$ in the decomposition

$\phi(x,t) = \phi_0(x) + \phi_1(x,t)$ is found by solving $\alpha\phi_{0,x}^2 = 2\beta n_0$. We directly find that $\phi_0(x) = \kappa x$ where $\kappa = 4\sqrt{\pi a_{\text{scatt}} n_0}$. Recall that the strong interaction gives $a_{\text{scatt}} \to \infty$ and corresponds to κ being a large quantity. The phase $\phi_0(x)$ is thus highly oscillating and ϕ_1 is a slowly varying phase. We then approximate the equation (3.6.31) at first order in κ to obtain the following forced wave equation

$$n_{1,t} + 8\sqrt{\pi a_{\text{scatt}} n_0} n_{1,x} = \hbar m^{-1} 4\pi a_{\text{scatt}} n_0 \text{Im}(p^*)\,. \tag{3.6.33}$$

To summarized we obtained a simple wave equation to describe the local dynamics of the BEC affected by the evaporation process of its atoms. This derivation is only formal and should be better physically justified.

By coupling this equation with the HBE equation (3.6.29), and using the variables

$$\begin{aligned} &Q = \text{Re}(p), \quad P = -\text{Im}(p), \\ &N = 2\pi\frac{\hbar}{m}(2f+1) \quad \text{and} \quad E = n_0 + n_1\,, \end{aligned} \tag{3.6.34}$$

one can check that we obtain the hRMB equation (3.6) with $\sigma_1 = -1$ and $\sigma_1 = 1$. In the case of the RMB with $\sigma_1 = 1$, a similar coupling is done with the Maxwell wave equation which describes the light propagation.

3.6.5 Summary and outlook

In this final section, we discussed the RMB equations which are not such well-known equations but can nevertheless be seen as members of the AKNS hierarchy. Usually, the sine-Gordon or the sinh-Gordon equations are derived from the first negative flow of the AKNS hierarchy. The fact that moving the poles of the spectral problem produces non-equivalent integrable equations seems to have not been fully explored in the past. We addressed the question of finding all the possible non-equivalent first negative flows of the AKNS hierarchy by moving the poles away from 0. This introduces parameters which can be interpreted as deformation parameters of the sine-Gordon or the sinh-Gordon. We found two possible deformations for each equation,

so we have 4 non-equivalent deformed negative flows of the AKNS hierarchy. All these equations are completely integrable via the inverse scattering transform, and we can in principle find their soliton solutions. We derived some of them by hand to show that the hyperbolic equations have kink solutions corresponding to the usual soliton of the original RMB equation. We did not investigate much further the use of IST, as the hyperbolicity of the equation makes the spectral problem difficult to solve and out of the scope of this work. Interestingly, we cannot solve the second RMB equation as easily as the original RMB equation. The reason for this seems to be hidden in the M operator, but we could not yet fully understand it.

We also investigated the link between a finite dimensional reduction of the RMB equations and the non-dissipative Lorenz 63 equation and how we could use the dynamics of the ODE to understand better the stability of constant solutions of the full PDE equations. We found several regime of stability and instability, with a modulational instability for the original RMB equation, and a region of stability for the most interesting new hRMB equation, which means that solutions can persists in a physical system. In fact, we showed that this hRMB equation can be derived from a particular approximation of physical equations in strongly correlated Bose-Einstein condensates. This derivation is rather formal and deserves to be better physically motivated to be able to predict for example the existence of solitary wave in these physical systems. We did not investigate this direction further as we do not have enough physical intuition about this new system to adequately justify our approximations. Indeed, the observation of strongly correlated Bose-Einstein condensates goes back to only a few years, and its mathematical description with the hyperbolic Bloch equation was developed in 2015 by Kira (2015b).

4

Conclusion

4.1 Geometric deformations of integrable systems ...

This work has been a mathematical exploration through several topics, with emphasis on integrability and stochasticity. The common thread of the studies undertaken here is the geometric deformation of integrable systems, stochasticity being one of the principal types of deformation. At first sight, noise does not seem compatible with the notion of integrability, but we saw that by using a particular multiplicative noise consistent with the geometrical structure of the original equation, integrability could be preserved. We systematically introduced noise following this idea of modifying the Hamiltonian with a stochastic potential, thus preserving the geometrical structure of the original equation. The freedom of the choice of the noise is thus limited to the choice of the number and form of the stochastic potentials.

In finite dimensions, the stochastic deformation of the rigid body motion with or without gravity remains integrable only if the body has a cylindrical symmetry along the axis of the noise. In this particular case, the noise appears in the equation for the angle variables in the action-angle decomposition and thus leaves the action variables preserved by the stochastic flow.

This observation was crucial in the definition of stochastic finite-dimensional integrable systems which we defined as such a system with a noise appearing only in the angle equation, in the sense of Liouville integrability.

We then looked at another classical integrable system, the Toda lattice. With the noise arising from the geometrical structure of the Flaschka variables, we found that the Toda lattice can admit a stochastic isospectral problem only if the noise does not preserve the coadjoint orbit, that the tri-diagonal structure of matrices and the connection to the physical lattice. This is because the noise does not couple two or more flows of the Toda hierarchy but has a noise attached to each particle.

Going to infinite dimensions, we can include noise in an integrable hierarchy while maintaining integrability only if the noise does not depend on space and if it multiplies a flow of the same integrable hierarchy as the deterministic equation. If we introduce noise as for finite dimensional systems using the geometric structure of the diffeomorphism group, we cannot obtain non-trivial integrable stochastic deformations of the original equations. Hence, we did not investigate this type of deformation in details and only showed that we could use other properties of an equation to include a consistent noise, as we did for the NLS equation with the noise via the Madelung change of variables. These examples are good illustrations of why using the structure of an equation to introduce systematic perturbations is a valid and fruitful approach. For example, in the case of the NLS equation the total mass of the solution is preserved, and in the case of the rigid body motion, the coadjoint orbit is preserved. These are properties necessary for integrability but also important physically as they often correspond to observed conservation laws, whereas the additional constants of motion required for integrability rarely exist in physical systems. In some sense, the noise represents uncertainties in the dynamics which may arise from various sources but which are never predictable in practice.

Apart from stochastic perturbations, we also considered other types of geometric deformations, such as a deformation depending on a parameter such that when the deformation parameter vanishes, we should recover known equations. We did that for the reduced Maxwell-Bloch equation, a member

of the negative AKNS hierarchy, where we could derive three new equations, with one of them appearing in the theory of strongly correlated Bose-Einstein condensates. It turns out that finding exact solutions via the inverse scattering transform is more difficult than expected because of the non-standard boundary conditions of this equation. We showed some numerical solutions and a study of the stability of the constant solutions as a first investigation of these equations.

The last type of deformation, and the most important after the stochastic perturbation is the insertion of the Sobolev H^1 norm directly into integrable hierarchies. We were able to implement this deformation in a systematic way by not using the diffeomorphism group structure of the equation, but by using the more abstract loop group structure aimed at better describing complete integrability. Interestingly, not all the equations obtained from this deformation are integrable. With this scheme, we were able to recover known integrable equations such as the Camassa-Holm equation, or its cubic counterpart, the modified Camassa-Holm equation, or FORQ equation. Nevertheless, some equations such as the new deformation of the NLS equation does not seem to be integrable. In this sense, the NLS equation plays a particular role, and its deformed equation, the so-called CH-NLS equation is at the boundary between integrable and non-integrable PDEs. We could still show that many features of integrable PDEs persist for the CH-NLS equation such as solitary waves, breathers or rogue waves. The place where we can observe an obstruction to its complete integrability is in the collision of solitary waves with high momenta. Indeed, the collisions are difficult to resolve numerically and seem to leave residual high frequencies waves behind. These observations could not be confirmed as the numerics do not converge well enough in this case, but it may be a clue for understanding the non-integrability of the CH-NLS equation.

To sum up, the question of robustness of integrable systems under geometric deformations has only been touched on this work and deserves more thoughts. In particular, we can mention a few open problems which arose from this work, such as the integrability of the Sobolev deformation of integrable hierarchies, the properties of stochastic solitons of the stochastic hierarchies,

the application of these deformations to other integrable hierarchies, understanding the integrability of the stochastic Toda lattice, applying the inverse scattering transform to the new RMB equations, etc...

4.2 ... AND BEYOND

In most of our studies, we began with an integrable equation and applied a geometric deformation to obtain new equations with similar geometrical properties. Apart from testing integrability, we were able to use this remaining geometric structure for other means.

For the stochastic deformation of finite dimensional systems such as rigid bodies, we included a nonlinear dissipation preserving the geometric structure, or in this case coadjoint orbits. We then studied the interaction of noise and dissipation on these usually non-flat coadjoint orbits. In our most important example, the free rigid body written on $SO(3)$, the coadjoint orbit is compact, which makes the investigation of the Fokker-Planck equation and its stationary distribution more transparent and possible to observe with numerical simulations. This system turned out to have a profound connection with statistical physics as the stationary distribution is a Gibbs distribution which appears in the description of physical systems coupled to an external heat bath of a given fixed temperature. In our case, the inverse temperature is the ratio between the noise and the dissipation amplitudes where a larger noise means a larger temperature. To obtain a notion of temperature in the statistical sense, we needed dissipation to control the energy input from the noise and obtain a statistically stable regime with a fixed temperature. We investigated this connection by showing that this distribution arises as a maximum entropy distribution, given a fixed average energy.

By following this idea of interaction between noise and dissipation, we took the point of view of random dynamical systems and tried to reveal the possible attractors of these systems. With noise, an attractor has a different meaning because the system is non-autonomous. The attractors are observed at time $t = 0$ for each noise realization by letting the initial conditions be pulled back to $-\infty$. Without noise, this definition coincides with the usual notion

of an attractor; with noise, it is called a pullback attractor. The topic of pullback attractors is an active area of research in dynamical systems providing a better understanding of the long-time dynamics of stochastic systems. In our case, we have even more interesting systems as we could numerically show that the motion on their pullback attractor can be chaotic, obtained by an estimation of the top Lyapunov exponent. This is a remarkable feature, difficult to prove analytically and subject to a vast amount of work, mainly devoted to understanding which class of dynamical systems admits it, why, and its properties. In most cases, these so-called random attractors arise from perturbations of non-zero dimensional deterministic attractors, such as limit cycles for example. Here, they arise from singular deterministic attractors which correspond to the minimum energy position of the free rigid body. From our investigations of this system, several other research directions are left open. In particular, obtaining a better, or even analytical description of these random attractors would surely be interesting as well as trying to extend these result from the rigid body motion to other mechanical systems with symmetries.

For this purpose, we applied a similar stochastic deformation and a geometric dissipation to another integrable system, the Toda lattice. Unfortunately, we did not have time to study the possible existence of random attractors in this system. The Toda lattice can be a high dimensional system, but the simple 3-particle lattice could be of interest for such an investigation. Another interesting point about the Toda lattice is that its continuum limit is the Burger's, or the KdV equation if higher order terms are taken into account. In our stochastic dissipative regime, we obtained an interesting stochastic dissipative PDE in the limit where the dissipation is highly nonlinear. This equation should deserve more work which we leave for later. The interesting point is that this new type of nonlinear dissipation is obtained from a modified double bracket type dissipation at the level of Flaschka variables, and is not a usual viscous damping of the particles in the lattice. We may thus expect interesting results regarding this equation.

Another particle system admitting a continuum limit is the landmark dynamics obtained from a reduction of the CH equation using singular solutions.

These solutions behave as interacting particles, and in the limit of a large number of particles, many solutions of the original PDE can be described. For this system, we can reconstruct a solution of the original PDE even from a single landmark. The interesting feature here is that the inclusion of noise preserves the relation between the stochastic landmark dynamics and the stochastic PDE. This is a crucial point for using these landmarks in the field of image matching or computational anatomy, as a finite dimensional representation of an image is easier to handle in practice. In this context, the noise is meant to model uncertainties and variabilities of observed images as it is inserted in the reconstruction equation from the velocity field to the solution on the diffeomorphism group. We then studied the effect of the noise in the evolution of the error of each landmark via a finite dimensional approximation of the Fokker-Planck equation. Using this scheme, we showed that we could reliably estimate the covariance around each landmark by including the effect of the noise combined with the landmark interaction. We can also estimate the noise correlation from a set of initial and final images, thus solving a difficult inverse problem. This is only possible because the noise preserves the original Hamiltonian dynamics, equivalently or the Euler-Poincaré structure of the original PDE. In short, we applied two successive finite dimensional reductions of the original stochastic PDE to obtain an estimate of the effect of the noise to the local error in the standard algorithms of large deformation matching.

A

Moment equation for stochastic landmarks

This appendix contains the computation of the moment equation for the stochastic landmark dynamics of section 2.6. It is based on the cluster expansion method applied to the first two moments of the Fokker-Planck equation. The method is described in subsection 2.6.2 and the next computations only work for Gaussian noise kernels.

A.1 First moments

Recall the backward Kolmogorov operator on q_i^α,

$$\mathcal{L}q_i^\alpha = \frac{\partial h}{\partial p_i^\alpha} + \frac{1}{2}\sum_{l,\gamma,\delta} \frac{\partial \sigma_l^\alpha(\mathbf{q}_i)}{\partial q_i^\gamma}\sigma_l^\gamma(\mathbf{q}_i)\,, \tag{A.1.1}$$

which is used to compute the time evolution of the singlet

$$\frac{d}{dt}\langle q_i^\alpha\rangle = \langle p_j^\alpha\rangle\, K(\langle q_i\rangle - \langle q_j\rangle) - \sum_{l,\gamma,\delta}\frac{1}{2\sigma_l^2}\sigma_l^\alpha(\langle \mathbf{q}_i\rangle)(\langle q_i^\gamma\rangle - \delta_l^\gamma)\sigma_l^\gamma(\langle \mathbf{q}_i\rangle)\,. \tag{A.1.2}$$

In this case, the equation only depends on the singlet of the momentum variable. We thus compute

$$\mathcal{L}p_i^\alpha = -\frac{\partial h}{\partial q_i^\alpha} + \frac{1}{2}\sum_{l,\gamma,\delta} p_j^\gamma \frac{\partial\sigma_l^\gamma(\mathbf{q}_i)}{\partial q_i^\delta}\frac{\partial\sigma_l^\delta(\mathbf{q}_i)}{\partial q_i^\alpha} - \frac{1}{2}\sum_{l,\gamma,\delta} p_i^\gamma \frac{\partial^2\sigma_l^\gamma(\mathbf{q}_i)}{\partial q_i^\alpha \partial q_i^\delta}\sigma_l^\delta(\mathbf{q}_i)\,. \tag{A.1.3}$$

which similarly gives the time evolution of the momentum singlet in two terms as

$$\frac{d}{dt}\langle p_i^\alpha\rangle = A_p + B_p\,,$$

where

$$\begin{aligned} A_p &= \frac{1}{\alpha^2}\sum_{j,\gamma} K(\langle q_i\rangle - \langle q_j\rangle)\,\langle p_i^\gamma p_j^\gamma (q_i^\alpha - q_j^\alpha)\rangle \\ B_p &= \sum_{l,\gamma}\frac{1}{2\sigma_l^2}\,\langle p_i^\gamma\rangle\,\sigma_l^\gamma(\langle \mathbf{q}_i\rangle)\sigma_l^\alpha(\langle \mathbf{q}_i\rangle)\,. \end{aligned}$$

We then expand A_p further using the cluster-expansion method on the triplet to get

$$\begin{aligned} A_p &= \frac{1}{\alpha^2}\sum_{j,\gamma} K(\langle \mathbf{q}_i\rangle - \langle \mathbf{q}_j\rangle)(\langle p_i^\gamma p_j^\gamma q_i^\alpha\rangle - \langle p_i^\gamma p_j^\gamma q_j^\alpha\rangle) \\ &\approx \frac{1}{\alpha^2}\sum_{j,\gamma} K(\langle \mathbf{q}_i\rangle - \langle \mathbf{q}_j\rangle)\Big(\,\langle p_i^\gamma\rangle\,\langle p_j^\gamma\rangle\,\langle q_i^\alpha\rangle + \Delta_2\,\langle p_i^\gamma p_j^\gamma\rangle\,\langle q_i^\alpha\rangle + \langle p_i^\gamma\rangle\,\Delta_2\,\langle p_j^\gamma q_i^\alpha\rangle \\ &\qquad + \Delta_2\,\langle p_i^\gamma q_i^\alpha\rangle\,\langle p_j^\gamma\rangle - \langle p_i^\gamma\rangle\,\langle p_j^\gamma\rangle\,\langle q_j^\alpha\rangle - \Delta_2\,\langle p_i^\gamma p_j^\gamma\rangle\,\langle q_j^\alpha\rangle \\ &\qquad - \langle p_i^\gamma\rangle\,\Delta_2\,\langle p_j^\gamma q_j^\alpha\rangle - \Delta_2\,\langle p_i^\gamma q_j^\alpha\rangle\,\langle p_j^\gamma\rangle\,\Big)\,. \end{aligned}$$

Already this term depends on the mixed correlations which we will compute shortly, but we first compute the position correlation.

A.2 $\langle qq \rangle$ CORRELATION

Recall the formula of the Kolmogorov operator applied to $q_i^\alpha q_j^\beta$,

$$\mathcal{L}(q_i^\alpha q_j^\beta) = q_i^\alpha \frac{\partial h}{\partial p_j^\beta} + \sum_l \sigma_l^\alpha(\mathbf{q}_i)\sigma_l^\beta(\mathbf{q}_j) + \frac{1}{2}\sum_l q_i^\alpha \sigma_l^\gamma(\mathbf{q}_j)\frac{\partial \sigma_l^\beta(\mathbf{q}_j)}{\partial q_j^\gamma} + t.\,, \tag{A.2.1}$$

which together with (A.1.1) gives the time evolution of the position correlation in the form

$$\frac{d}{dt}\Delta_2 \langle q_i^\alpha q_j^\beta \rangle = A_{qq} + B_{qq} + C_{qq}\,,$$

where

$$\begin{aligned}
A_{qq} &= \left\langle q_i^\alpha \frac{\partial h}{\partial p_j^\beta} \right\rangle - \langle q_i^\alpha \rangle \left\langle \frac{\partial h}{\partial p_j^\beta} \right\rangle + t. \\
B_{qq} &= \sum_l \left\langle \sigma_l^\alpha(\mathbf{q}_i)\sigma_l^\beta(\mathbf{q}_j) \right\rangle \\
C_{qq} &= \frac{1}{2}\sum_l \left\langle q_i^\alpha \sigma_l^\gamma(\mathbf{q}_j)\frac{\partial \sigma_l^\beta(\mathbf{q}_j)}{\partial q_j^\gamma} \right\rangle - \frac{1}{2}\sum_l \langle q_i^\alpha \rangle \left\langle \sigma_l^\gamma(\mathbf{q}_j)\frac{\partial \sigma_l^\beta(\mathbf{q}_j)}{\partial q_j^\gamma} \right\rangle + t..
\end{aligned}$$

Throughout we will denote by A the terms corresponding to the drift, by B the terms which are not present in the first moment equation and by C the other terms which only depend on the noise and the derivative of the noise fields. We proceed by first approximating the expectation of the kernels to get

$$\begin{aligned}
B_{qq} &\approx \sum_l \sigma_l^\alpha(\langle \mathbf{q}_i \rangle)\sigma_l^\beta(\langle \mathbf{q}_j \rangle) \\
C_{qq} &\approx -\frac{1}{2\alpha_l^2}\sum_{l,\gamma} \Delta_2 \langle q_i^\alpha q_j^\gamma \rangle \sigma_l^\gamma(\langle \mathbf{q}_j \rangle)\sigma_l^\beta(\langle \mathbf{q}_j \rangle) + t..
\end{aligned}$$

where we also used the explicit form of σ_l as a Gaussian and use its form of the derivative. We will now approximate the A_{qq} term to get

$$\begin{aligned} A_{qq} &= \sum_k \left\langle q_i^\alpha p_k^\beta K(\mathbf{q}_j - \mathbf{q}_k) \right\rangle - \sum_k \langle q_i^\alpha \rangle \left\langle p_k^\beta K(\mathbf{q}_j - \mathbf{q}_k) \right\rangle + t. \\ &\approx \sum_k \Delta_2 \left\langle q_i^\alpha p_k^\beta \right\rangle K(\langle \mathbf{q}_j \rangle - \langle \mathbf{q}_k \rangle) + t.. \end{aligned}$$

It is now clear that the B term will linearly increase the position correlation, which will then exponentially increase by the C term and be affected by the momentum-position correlation by the A term. We now proceed by computing the momentum correlation.

A.3 $\langle pp \rangle$ CORRELATION

We compute the Kolmogorov operator on $p_o^\alpha p_j^\beta$ to get

$$\begin{aligned} \mathcal{L}(p_i^\alpha p_j^\beta) = &-p_i^\alpha \frac{\partial h}{\partial q_j^\beta} - \frac{1}{2} \sum_{l,\gamma,\delta} p_i^\alpha p_j^\gamma \sigma_l^\delta(\mathbf{q}_j) \frac{\partial^2 \sigma_l^\gamma(\mathbf{q}_j)}{\partial q_j^\beta \partial q_j^\delta} + t. \\ &+ \frac{1}{2} \sum_{l,\gamma,\delta} p_i^\delta p_j^\gamma \frac{\partial \sigma_l^\delta(\mathbf{q}_i)}{\partial q_i^\alpha} \frac{\partial \sigma_l^\gamma(\mathbf{q}_j)}{\partial q_j^\beta} + \frac{1}{2} \sum_{l,\gamma,\delta} p_i^\alpha p_j^\delta \frac{\partial \sigma_l^\delta(\mathbf{q}_j)}{\partial q_j^\gamma} \frac{\partial \sigma_l^\gamma(\mathbf{q}_j)}{\partial q_j^\beta} + t., \end{aligned}$$

and using (A.1.3), we obtain the time evolution of the correlation in three terms as

$$\frac{d}{dt} \Delta_2 \langle p_i^\alpha p_j^\beta \rangle = A_{pp} + B_{pp} + C_{pp},$$

where

$$A_{pp} = -\left\langle p_i^\alpha \frac{\partial h}{\partial q_j^\beta} \right\rangle + \langle p_i^\alpha \rangle \left\langle \frac{\partial h}{\partial q_j^\beta} \right\rangle + t.$$

$$B_{pp} = \frac{1}{2} \sum_{l,\gamma,\delta} \left\langle p_i^\delta p_j^\gamma \frac{\partial \sigma_l^\delta(\mathbf{q}_i)}{\partial q_i^\alpha} \frac{\partial \sigma_l^\gamma(\mathbf{q}_j)}{\partial q_j^\beta} \right\rangle + t.$$

$$C_{pp} = -\frac{1}{2} \sum_{l,\gamma,\delta} \left\langle p_i^\alpha p_j^\gamma \sigma_l^\delta(\mathbf{q}_j) \frac{\partial^2 \sigma_l^\gamma(\mathbf{q}_j)}{\partial q_j^\beta \partial q_j^\delta} \right\rangle + \frac{1}{2} \sum_{l,\gamma,\delta} \langle p_i^\alpha \rangle \left\langle p_j^\gamma \sigma_l^\delta(\mathbf{q}_j) \frac{\partial^2 \sigma_l^\gamma(\mathbf{q}_j)}{\partial q_j^\beta \partial q_j^\delta} \right\rangle$$
$$+ \frac{1}{2} \sum_{l,\gamma,\delta} \left\langle p_i^\alpha p_j^\delta \frac{\partial \sigma_l^\delta(\mathbf{q}_j)}{\partial q_j^\gamma} \frac{\partial \sigma_l^\gamma(\mathbf{q}_j)}{\partial q_j^\beta} \right\rangle - \frac{1}{2} \sum_{l,\gamma,\delta} \langle p_i^\alpha \rangle \left\langle p_j^\delta \frac{\partial \sigma_l^\delta(\mathbf{q}_j)}{\partial q_j^\gamma} \frac{\partial \sigma_l^\gamma(\mathbf{q}_j)}{\partial q_j^\beta} \right\rangle + t..$$

We first approximate

$$C_{pp} \approx \frac{1}{2} \sum_{l,\gamma,\delta} \frac{1}{\alpha_l^2} \sigma_l^\beta(q_j) \sigma_l^\gamma(q_j) \Delta_2 \left\langle p_i^\alpha p_j^\gamma \right\rangle$$
$$- \frac{1}{2} \sum_{l,\gamma,\delta} \sigma_l^\delta(q_j) \sigma_l^\gamma(q_j) \left(\left\langle p_i^\alpha p_j^\gamma (q_j^\beta - \delta_l^\beta)(q_j^\delta - \delta_l^\delta) \right\rangle - \langle p_i^\alpha \rangle \left\langle p_j^\gamma (q_j^\beta - \delta_l^\beta)(q_j^\delta - \delta_l^\delta) \right\rangle \right)$$
$$+ \frac{1}{2} \sum_{l,\gamma,\delta} \sigma_l^\delta(q_j) \sigma_l^\gamma(q_j) \left(\left\langle p_i^\alpha p_j^\delta (q_j^\gamma - \delta_l^\gamma)(q_j^\beta - \delta_l^\beta) \right\rangle + \langle p_i^\alpha \rangle \left\langle p_j^\delta (q_j^\gamma - \delta_l^\gamma)(q_j^\beta - \delta_l^\beta) \right\rangle \right).$$

The last two terms cancel as they are symmetric under the transpose operation because of the sum on the free indices, thus the C term is

$$C_{pp} \approx \frac{1}{2} \sum_{l,\gamma,\delta} \frac{1}{\alpha_l^2} \sigma_l^\beta(q_j) \sigma_l^\gamma(q_j) \Delta_2 \left\langle p_i^\alpha p_j^\gamma \right\rangle + t.. \tag{A.3.1}$$

We proceed with the B_{pp} term, which is also symmetric under the transpose

operation, thus giving the approximation

$$
\begin{aligned}
B_{pp} &= \sum_{l,\gamma,\delta} \frac{1}{\alpha_l^4} \sigma_l^\delta(q_i)\sigma_l^\gamma(q_j) \left\langle p_i^\delta p_j^\gamma (q_i^\alpha - \delta_l^\alpha)(q_j^\beta - \delta_l^\beta) \right\rangle \\
&\approx \sum_{l,\gamma,\delta} \frac{1}{\alpha_l^4} \sigma_l^\delta(q_i)\sigma_l^\gamma(q_j) \Big(\left\langle p_i^\delta \right\rangle \left\langle p_j^\gamma \right\rangle \left\langle q_i^\alpha \right\rangle \left\langle q_j^\beta \right\rangle + \Delta_2 \left\langle p_i^\delta p_j^\gamma \right\rangle \left\langle q_i^\alpha \right\rangle \left\langle q_j^\beta \right\rangle \\
&\qquad + \left\langle p_i^\delta \right\rangle \Delta_2 \left\langle p_j^\gamma q_i^\alpha \right\rangle \left\langle q_j^\beta \right\rangle + \left\langle p_i^\delta \right\rangle \left\langle p_j^\gamma \right\rangle \Delta_2 \left\langle q_i^\alpha q_j^\beta \right\rangle \\
&\qquad + \Delta_2 \left\langle p_i^\delta q_i^\alpha \right\rangle \left\langle p_j^\gamma \right\rangle \left\langle q_j^\beta \right\rangle + \Delta_2 \left\langle p_i^\delta q_j^\beta \right\rangle \left\langle p_j^\gamma \right\rangle \left\langle q_i^\alpha \right\rangle \\
&\qquad + \left\langle p_i^\delta \right\rangle \left\langle q_i^\alpha \right\rangle \Delta_2 \left\langle p_j^\gamma q_j^\beta \right\rangle + \Delta_2 \left\langle p_i^\delta p_j^\gamma \right\rangle \Delta_2 \left\langle q_i^\alpha q_j^\beta \right\rangle \\
&\qquad + \Delta_2 \left\langle p_i^\delta q_i^\alpha \right\rangle \Delta_2 \left\langle p_j^\gamma q_j^\beta \right\rangle + \Delta_2 \left\langle p_i^\delta q_j^\beta \right\rangle \Delta_2 \left\langle p_j^\gamma q_i^\alpha \right\rangle \\
&\qquad - \left\langle p_i^\delta \right\rangle \left\langle p_j^\gamma \right\rangle \left\langle q_i^\alpha \right\rangle \delta_l^\beta - \Delta_2 \left\langle p_i^\delta p_j^\gamma \right\rangle \left\langle q_i^\alpha \right\rangle \delta_l^\beta \\
&\qquad - \left\langle p_i^\delta \right\rangle \Delta_2 \left\langle p_j^\gamma q_i^\alpha \right\rangle \delta_l^\beta - \Delta_2 \left\langle p_i^\delta q_i^\alpha \right\rangle \left\langle p_j^\gamma \right\rangle \delta_l^\beta \\
&\qquad - \left\langle p_i^\delta \right\rangle \left\langle p_j^\gamma \right\rangle \left\langle q_j^\beta \right\rangle \delta_l^\alpha - \Delta_2 \left\langle p_i^\delta p_j^\gamma \right\rangle \left\langle q_j^\beta \right\rangle \delta_l^\alpha \\
&\qquad - \left\langle p_i^\delta \right\rangle \Delta_2 \left\langle p_j^\gamma q_j^\beta \right\rangle \delta_l^\alpha - \Delta_2 \left\langle p_i^\delta q_j^\beta \right\rangle \left\langle p_j^\gamma \right\rangle \delta_l^\alpha \\
&\qquad + \left\langle p_i^\delta \right\rangle \left\langle p_j^\gamma \right\rangle \delta_l^\alpha \delta_l^\beta + \Delta_2 \left\langle p_i^\delta p_j^\gamma \right\rangle \delta_l^\alpha \delta_l^\beta \Big) .
\end{aligned}
$$

we treat the two Hamiltonian terms separately by first writing them explicitly as

$$
\begin{aligned}
A_{pp} &= \sum_{k,\gamma} \frac{1}{\alpha^2} \left\langle p_i^\alpha p_j^\gamma p_k^\gamma (q_j^\beta - q_k^\beta) K(q_j - q_k) \right\rangle \\
&\quad - \frac{1}{\alpha^2} \sum_{k,\gamma} \left\langle p_i^\alpha \right\rangle \left\langle p_j^\gamma p_k^\gamma (q_j^\beta - q_k^\beta) K(q_j - q_k) \right\rangle + t. \\
&=: \sum_{k,\gamma} \frac{1}{\alpha^2} (A_{pp}^1 - A_{pp}^2) + t. .
\end{aligned}
$$

We expand the first term to arrive at

$$
\begin{aligned}
A^1_{pp} &\approx K(\langle q_j\rangle - \langle q_k\rangle)\left(\left\langle p_i^\alpha p_j^\gamma p_k^\gamma q_j^\beta\right\rangle - \left\langle p_i^\alpha p_j^\gamma p_k^\gamma q_k^\beta\right\rangle\right)\\
&\approx K(\langle q_j\rangle - \langle q_k\rangle)\Big(\langle p_i^\alpha\rangle\langle p_j^\gamma\rangle\langle p_k^\gamma\rangle\langle q_j^\beta\rangle + \Delta_2\langle p_i^\alpha p_j^\gamma\rangle\langle p_k^\gamma\rangle\langle q_j^\beta\rangle\\
&\qquad + \langle p_i^\alpha\rangle\Delta_2\langle p_j^\gamma p_k^\gamma\rangle\langle q_j^\beta\rangle + \langle p_i^\alpha\rangle\langle p_j^\gamma\rangle\Delta_2\langle p_k^\gamma q_j^\beta\rangle\\
&\qquad + \Delta_2\langle p_i^\alpha p_k^\gamma\rangle\langle p_j^\gamma\rangle\langle q_j^\beta\rangle + \Delta_2\left\langle p_i^\alpha q_j^\beta\right\rangle\langle p_j^\gamma\rangle\langle p_k^\gamma\rangle\\
&\qquad + \langle p_i^\alpha\rangle\langle p_k^\gamma\rangle\Delta_2\langle p_j^\gamma q_j^\beta\rangle + \Delta_2\langle p_i^\alpha p_j^\gamma\rangle\Delta_2\langle p_k^\gamma q_j^\beta\rangle\\
&\qquad + \Delta_2\langle p_i^\alpha p_k^\gamma\rangle\Delta_2\langle p_j^\gamma q_j^\beta\rangle + \Delta_2\left\langle p_i^\alpha q_j^\beta\right\rangle\Delta_2\langle p_k^\gamma p_j^\gamma\rangle\\
&\qquad - \langle p_i^\alpha\rangle\langle p_j^\gamma\rangle\langle p_k^\gamma\rangle\langle q_k^\beta\rangle - \Delta_2\langle p_i^\alpha p_j^\gamma\rangle\langle p_k^\gamma\rangle\langle q_k^\beta\rangle\\
&\qquad - \langle p_i^\alpha\rangle\Delta_2\langle p_j^\gamma p_k^\gamma\rangle\langle q_k^\beta\rangle - \langle p_i^\alpha\rangle\langle p_j^\gamma\rangle\Delta_2\langle p_k^\gamma q_k^\beta\rangle\\
&\qquad - \Delta_2\langle p_i^\alpha p_k^\gamma\rangle\langle p_j^\gamma\rangle\langle q_k^\beta\rangle - \Delta_2\left\langle p_i^\alpha q_k^\beta\right\rangle\langle p_k^\gamma\rangle\langle p_k^\gamma\rangle\\
&\qquad - \Delta_2\left\langle p_k^\gamma q_k^\beta\right\rangle\langle p_i^\alpha\rangle\langle p_k^\gamma\rangle - \Delta_2\langle p_i^\alpha p_j^\gamma\rangle\Delta_2\langle p_k^\gamma q_k^\beta\rangle\\
&\qquad - \Delta_2\langle p_i^\alpha p_k^\gamma\rangle\Delta_2\langle p_j^\gamma q_k^\beta\rangle - \Delta_2\left\langle p_i^\alpha q_k^\beta\right\rangle\Delta_2\langle p_k^\gamma p_j^\gamma\rangle\Big),
\end{aligned}
$$

and the second term

$$
\begin{aligned}
A^2_{pp} &\approx K(\langle q_j\rangle - \langle q_k\rangle)\left(\langle p_i^\alpha\rangle\left\langle p_j^\gamma p_k^\gamma q_j^\beta\right\rangle - \langle p_i^\alpha\rangle\left\langle p_j^\gamma p_k^\gamma q_k^\beta\right\rangle\right)\\
&\approx K(\langle q_j\rangle - \langle q_k\rangle)\Big(\langle p_i^\alpha\rangle\langle p_j^\gamma\rangle\langle p_k^\gamma\rangle\langle q_j^\beta\rangle + \langle p_i^\alpha\rangle\Delta_2\langle p_j^\gamma p_k^\gamma\rangle\langle q_j^\beta\rangle\\
&\qquad + \langle p_i^\alpha\rangle\Delta_2\langle p_j^\gamma p_j^\beta\rangle\langle q_k^\gamma\rangle\langle p_i^\alpha\rangle\langle p_j^\gamma\rangle\Delta_2\langle p_k^\gamma q_j^\beta\rangle\\
&\qquad - \langle p_i^\alpha\rangle\langle p_j^\gamma\rangle\langle p_k^\gamma\rangle\langle q_k^\beta\rangle - \langle p_i^\alpha\rangle\Delta_2\langle p_j^\gamma p_k^\gamma\rangle\langle q_k^\beta\rangle\\
&\qquad - \langle p_i^\alpha\rangle\langle p_j^\gamma\rangle\Delta_2\langle p_k^\gamma q_k^\beta\rangle - \langle p_i^\alpha\rangle\langle p_k^\gamma\rangle\Delta_2\langle p_j^\gamma q_k^\beta\rangle\Big).
\end{aligned}
$$

This term cancels the terms of the A^1_{pp} proportional to $\langle p_i^\alpha\rangle$ to give the approximation

$$
\begin{aligned}
A_{pp} \approx \frac{1}{\alpha^2}\sum_{k,\gamma} K(\langle q_j\rangle - \langle q_k\rangle)\Big(&\Delta_2\langle p_i^\alpha p_j^\gamma\rangle\langle p_k^\gamma\rangle\langle q_j^\beta\rangle + \Delta_2\langle p_i^\alpha p_k^\gamma\rangle\langle p_j^\gamma\rangle\langle q_j^\beta\rangle\\
&+ \Delta_2\left\langle p_i^\alpha q_j^\beta\right\rangle\langle p_j^\gamma\rangle\langle p_k^\gamma\rangle + \Delta_2\langle p_i^\alpha p_j^\gamma\rangle\Delta_2\langle p_k^\gamma q_j^\beta\rangle\\
&+ \Delta_2\langle p_i^\alpha p_k^\gamma\rangle\Delta_2\langle p_j^\gamma q_j^\beta\rangle + \Delta_2\left\langle p_i^\alpha q_j^\beta\right\rangle\Delta_2\langle p_k^\gamma p_j^\gamma\rangle\\
&- \Delta_2\langle p_i^\alpha p_j^\gamma\rangle\langle p_k^\gamma\rangle\langle q_k^\beta\rangle - \Delta_2\langle p_i^\alpha p_k^\gamma\rangle\langle p_j^\gamma\rangle\langle q_k^\beta\rangle\\
&- \Delta_2\left\langle p_i^\alpha q_k^\beta\right\rangle\langle p_j^\gamma\rangle\langle p_k^\gamma\rangle - \Delta_2\langle p_i^\alpha p_j^\gamma\rangle\Delta_2\langle p_k^\gamma q_k^\beta\rangle\\
&- \Delta_2\langle p_i^\alpha p_k^\gamma\rangle\Delta_2\langle p_j^\gamma q_k^\beta\rangle - \Delta_2\left\langle p_i^\alpha q_k^\beta\right\rangle\Delta_2\langle p_k^\gamma p_j^\gamma\rangle\Big) + t..
\end{aligned}
$$

We end this computation by approximating the dynamics of the mixed

correlation.

A.4 $\langle pq \rangle$ CORRELATION

We compute

$$\begin{aligned}\mathcal{L}(p_i^\alpha q_j^\beta) = &-q_j^\beta \frac{\partial h}{\partial q_i^\alpha} + p_i^\alpha \frac{\partial h}{\partial p_j^\beta} \\ &+ \frac{1}{2}\sum_{l,\gamma} p_i^\alpha \sigma_l^\gamma(\mathbf{q}_j)\frac{\partial \sigma_l^\beta(\mathbf{q}_j)}{\partial q_j^\gamma} - \sum_{l,\gamma} p_i^\gamma \sigma_l^\beta(\mathbf{q}_j)\frac{\partial \sigma_l^\gamma(\mathbf{q}_i)}{\partial q_i^\alpha} \\ &- \frac{1}{2}\sum_{l,\gamma,\delta} q_j^\beta p_i^\gamma \frac{\partial^2 \sigma_l^\gamma(\mathbf{q}_i)}{\partial q_i^\alpha \partial q_i^\delta}\sigma_l^\delta(\mathbf{q}_i) + \frac{1}{2}\sum_{l,\gamma,\delta} p_i^\delta q_j^\beta \frac{\partial \sigma_l^\gamma(\mathbf{q}_i)}{\partial q_i^\alpha}\frac{\partial \sigma_l^\delta(\mathbf{q}_i)}{\partial q_i^\gamma}\,.\end{aligned}$$

Then, using (A.1.1) and (A.1.3) we obtain the time evolution of $\Delta\langle p_i^\alpha q_j^\beta\rangle$ as

$$\frac{d}{dt}\Delta\langle p_i^\alpha q_j^\beta\rangle = A_{pq} + B_{pq} + C_{pq}\,,$$

where

$$\begin{aligned}A_{pq} &= -\left\langle q_j^\beta \frac{\partial h}{\partial q_i^\alpha}\right\rangle + \left\langle q_j^\beta\right\rangle\left\langle \frac{\partial h}{\partial q_i^\alpha}\right\rangle + \left\langle p_i^\alpha \frac{\partial h}{\partial p_j^\beta}\right\rangle - \langle p_i^\alpha\rangle\left\langle \frac{\partial h}{\partial p_j^\beta}\right\rangle \\ B_{pq} &= -\sum_{l,\gamma}\left\langle p_i^\gamma \sigma_l^\beta(\mathbf{q}_j)\frac{\partial \sigma_l^\gamma(\mathbf{q}_i)}{\partial q_i^\alpha}\right\rangle \\ C_{pq} &= \frac{1}{2}\sum_{l,\gamma}\left\langle p_i^\alpha \sigma_l^\gamma(\mathbf{q}_j)\frac{\partial \sigma_l^\beta(\mathbf{q}_j)}{\partial q_j^\gamma}\right\rangle - \frac{1}{2}\sum_{l,\gamma}\langle p_i^\alpha\rangle\left\langle \sigma_l^\gamma(\mathbf{q}_j)\frac{\partial \sigma_l^\beta(\mathbf{q}_j)}{\partial q_j^\gamma}\right\rangle \\ &\quad - \frac{1}{2}\sum_{l,\gamma,\delta}\left\langle q_j^\beta p_i^\gamma \frac{\partial^2 \sigma_l^\gamma(\mathbf{q}_i)}{\partial q_i^\alpha \partial q_i^\delta}\sigma_l^\delta(\mathbf{q}_i)\right\rangle + \frac{1}{2}\sum_{l,\gamma,\delta}\left\langle q_j^\beta\right\rangle\left\langle p_i^\gamma \frac{\partial^2 \sigma_l^\gamma(\mathbf{q}_i)}{\partial q_i^\alpha \partial q_i^\delta}\sigma_l^\delta(\mathbf{q}_i)\right\rangle \\ &\quad + \frac{1}{2}\sum_{l,\gamma,\delta}\left\langle p_i^\delta q_j^\beta \frac{\partial \sigma_l^\gamma(\mathbf{q}_i)}{\partial q_i^\alpha}\frac{\partial \sigma_l^\delta(\mathbf{q}_i)}{\partial q_i^\gamma}\right\rangle - \frac{1}{2}\sum_{l,\gamma,\delta}\langle q_j^\beta\rangle\left\langle p_i^\delta \frac{\partial \sigma_l^\gamma(\mathbf{q}_i)}{\partial q_i^\alpha}\frac{\partial \sigma_l^\delta(\mathbf{q}_i)}{\partial q_i^\gamma}\right\rangle\,.\end{aligned}$$

We first approximate

$$B_{pq} \approx \frac{1}{\alpha_l^2} \sum_{l,\gamma} \sigma_l^\beta(\mathbf{q}_j)\sigma_l^\gamma(\mathbf{q}_i)\left(\langle p_i^\gamma q_i^\alpha\rangle - \langle p_i^\gamma\rangle \delta_l^\alpha\right)$$

$$C_{pq} \approx -\sum_{l,\gamma} \frac{1}{2\alpha_l^2}\sigma_l^\beta(\mathbf{q}_j)\sigma_l^\gamma(\mathbf{q}_j)\Delta_2\left\langle p_i^\alpha q_j^\gamma\right\rangle + \sum_{l,\gamma}\frac{1}{2\alpha_l^2}\Delta_2\left\langle q_j^\beta p_i^\gamma\right\rangle \sigma_l^\gamma(\langle \mathbf{q}_i\rangle)\sigma_l^\alpha(\langle \mathbf{q}_i\rangle)\,.$$

For the Hamiltonian term we obtain

$$\begin{aligned}
A_{pq} &\approx \frac{1}{\alpha^2}\sum_{k,\gamma} K(\langle \mathbf{q}_i\rangle - \langle \mathbf{q}_k\rangle)\Big(\left\langle q_j^\beta p_i^\gamma p_k^\gamma q_i^\alpha\right\rangle - \left\langle q_j^\beta\right\rangle \langle p_i^\gamma p_k^\gamma q_i^\alpha\rangle \\
&\quad - \left\langle q_j^\beta p_i^\gamma p_k^\gamma q_k^\alpha\right\rangle + \left\langle q_j^\beta\right\rangle \langle p_i^\gamma p_k^\gamma q_k^\alpha\rangle\Big) + \sum_{k,\gamma} K(\langle \mathbf{q}_j\rangle - \langle \mathbf{q}_k\rangle)\Delta_2\left\langle p_i^\alpha p_k^\beta\right\rangle \\
&\approx \frac{1}{\alpha^2}\sum_{k,\gamma} K(\langle \mathbf{q}_i\rangle - \langle \mathbf{q}_k\rangle)\Big(\Delta_2\left\langle p_i^\gamma q_j^\beta\right\rangle \langle p_k^\gamma\rangle \langle q_i^\alpha\rangle + \Delta_2\left\langle p_k^\gamma q_j^\beta\right\rangle \langle p_i^\gamma\rangle \langle q_i^\alpha\rangle \\
&\qquad + \Delta_2\left\langle q_i^\alpha q_j^\beta\right\rangle \langle p_i^\gamma\rangle \langle p_k^\gamma\rangle + \Delta_2\left\langle p_i^\gamma q_j^\beta\right\rangle \Delta_2 \langle p_k^\gamma q_i^\alpha\rangle \\
&\qquad + \Delta_2\left\langle p_k^\gamma q_j^\beta\right\rangle \Delta_2 \langle p_i^\gamma q_i^\alpha\rangle + \Delta_2\left\langle q_i^\alpha q_j^\beta\right\rangle \Delta_2 \langle p_k^\gamma p_i^\gamma\rangle \\
&\qquad - \Delta_2\left\langle p_i^\gamma q_j^\beta\right\rangle \langle p_k^\gamma\rangle \langle q_k^\alpha\rangle - \Delta_2\left\langle p_k^\gamma q_j^\beta\right\rangle \langle p_i^\gamma\rangle \langle q_k^\alpha\rangle \\
&\qquad - \Delta_2\left\langle q_j^\beta q_k^\alpha\right\rangle \langle p_i^\gamma\rangle \langle p_k^\gamma\rangle - \Delta_2\left\langle p_i^\gamma q_j^\beta\right\rangle \Delta_2 \langle p_k^\gamma q_k^\alpha\rangle \\
&\qquad - \Delta_2\left\langle p_k^\gamma q_j^\beta\right\rangle \Delta_2 \langle p_i^\gamma q_k^\alpha\rangle - \Delta_2\left\langle q_j^\beta q_k^\alpha\right\rangle \Delta_2 \langle p_k^\gamma p_i^\gamma\rangle\Big) \\
&\quad + \sum_{k,\gamma} K(\langle \mathbf{q}_j\rangle - \langle \mathbf{q}_k\rangle)\Delta_2\left\langle p_i^\alpha p_k^\beta\right\rangle\,.
\end{aligned}$$

B

Numerical scheme for the CH-NLS equation

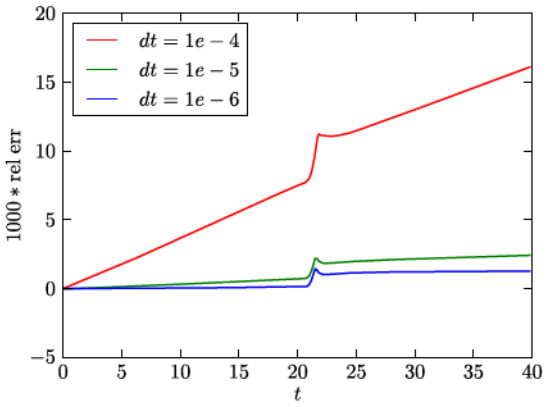
In this short section, we want to quickly explain the numerical scheme that we used for the previous simulation of the CH-NLS equation. This is important also to show that the radiation coming out of the collisions are difficult to resolve numerically and may not be physical.

We used a pseudo-spectral numerical scheme with a Fourier basis in space. We also needed a second order centred finite difference approximation for the space derivatives of the nonlinear terms. The time integration is a second order implicit scheme for the linear part and a first order explicit scheme for the nonlinear term. The scheme is given by

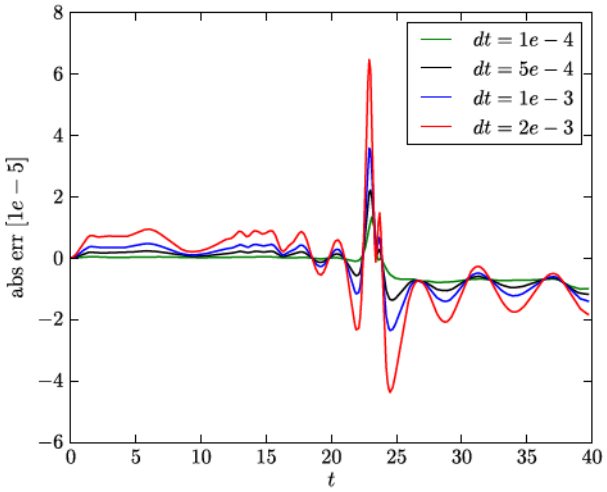
$$\frac{\widehat{u}_{i+1}-\widehat{u}_i}{\Delta t}=\frac{i}{1+\alpha^2\kappa^2}\left[-\kappa^2\frac{\widehat{u}_{i+1}+\widehat{u}_i}{2}+\widehat{\left\{m(|u|^2-\alpha^2|u_x|^2)\right\}}_i\right], \quad \text{(B.0.1)}$$

where $\widehat{u}(\kappa)$ is the Fourier transform of $u(x)$ and i indexes the timesteps. The integration of the linear term is not dissipative, but the integration of the nonlinear term will introduce dissipation. For each simulation, we verify that the space resolution is small enough to capture most of the spectrum of the

solution and such that the solution can converge for simple soliton emergence experiments. For the collision of the modulational instability simulations, the required spectrum band is too large for realistic simulations. This will prevent us from having a good convergence of the scheme. In figure B.1, we display the errors on the conservation of mass and momentum of the collision solution of figure 3.3(b), for three different timesteps. This plot illustrates the first order of convergence of our scheme before the collision as well as the systematic error during the collision. We also did this simulation again with Dedalus Burns et al. (2016) which has a better implementation of spectral methods, but we obtain similar results. The main issue is the creation of a extremely narrow peak of high amplitude at the moment of collision which cannot be captured by these schemes, even by considerably increasing the spatial and temporal resolution.



(a) Numerical convergence of mass M.



(b) Numerical convergence of the momentum P.

Figure B.1: Both panels of this figure show the errors for the collision displayed in 3.3(b). The error for the mass is relative whereas the error for the momentum is absolute because the value of the momentum is close to 0. As expected, the convergence is of order 1, but the collision is never completely resolved, as seen from the jumps in the error of both the momentum and mass.

C

Derivation of the RMB equations

We will show here how to derive the L and M operators of the RMB equation of section 3.6 in the context of the AKNS hierarchy by introducing the deformation parameter ω_0 in its first negative flow.

For the derivation of the M and L operator we will use the same notation as in section 3.2 for the elements of the Lie algebra $\mathfrak{sl}(2)$, that is we denoted its Chevalley basis by h, e and f. We start with the standard L operator but with an M operator which contains negative powers of the spectral parameter instead of positive powers for the positive AKNS hierarchy. In contrast with the positive flows, a modification of the roots of the polynomial in λ is important. Indeed, for the positive flows, the roots correspond to zeros and can be moved in the complex plane without affecting the integrable equations. The extra term resulting from this procedure can be absorbed by a suitable gauge transformation. For negative flows the roots become poles, and their position modifies the equations significantly. This effect can be seen by the fact that if the spectral parameter takes the value of the pole, the M operator will blow up and ZCR cannot be used anymore. The position

of the pole is given by the parameter ω_0 which produces the extra terms of the RMB equation.

C.1 Standard RMB equations

We start by deriving the classical RMB equation, with the following form for L and M

$$L = \lambda A + B, \quad \text{and} \quad M = \Delta^{-1}(\lambda C + \omega_0 D)\,, \tag{C.1.1}$$

where $\Delta = (\lambda - \omega_0)(\lambda + \omega_0)$ and A, B, C and D are generic $\mathfrak{sl}(2)$ matrices. We expand the commutator in powers of λ

$$[L, M] = \Delta^{-1}(\lambda^2[A, C] + \lambda\omega_0[A, D] + \lambda[B, C] + \omega_0[B, D])\,,$$

and use the relations

$$\begin{aligned}
\Delta^{-1} &= \frac{1}{2\omega_0(\lambda - \omega_0)} - \frac{1}{2\omega_0(\lambda + \omega_0)} := \frac{1}{2\omega_0}(\Delta_-^{-1} - \Delta_+^{-1}) \\
\lambda\Delta^{-1} &= \frac{1}{2(\lambda - \omega_0)} + \frac{1}{2(\lambda + \omega_0)} = \frac{1}{2}(\Delta_-^{-1} + \Delta_+^{-1}) \\
\lambda^2\Delta^{-1} &= \frac{\omega_0}{2(\lambda - \omega_0)} - \frac{\omega_0}{2(\lambda + \omega_0)} + 1 = 1 + \frac{\omega_0}{2}(\Delta_-^{-1} - \Delta_+^{-1})\,,
\end{aligned} \tag{C.1.2}$$

which define Δ_- and Δ_+. We expand this commutator further into three parts as

$$\begin{aligned}
\Delta_-^{-1} &: \frac{1}{2}(\omega_0[A, D] + [B, D] + [B, C] + \omega_0[A, C]) \\
\Delta_+^{-1} &: \frac{1}{2}(\omega_0[A, D] - [B, D] + [B, C] - \omega_0[A, C]) \\
\lambda^0 &: [A, C]\,.
\end{aligned}$$

We now compute the space derivative of the M operator and expand it similarly and obtain

$$\Delta_{-}^{-1} : \frac{1}{2}(D_x + C_x)$$
$$\Delta_{+}^{-1} : \frac{1}{2}(-D_x + C_x)\,.$$

From here, we will have three independent equations which need to be solved. They can first be rearranged as

$$\begin{aligned} B_t &= [C, A] \\ C_x &= \omega_0[D, A] + [C, B] \\ D_x &= \omega_0[C, A] + [D, B]\,. \end{aligned} \tag{C.1.3}$$

By direct computation one can see that only the two cases where $\sigma_2 = 1$ will solve this system of matrix equations.

C.2 Hyperbolic RMB equations

We will show how to this computation for the other case $\sigma_2 = -1$ which contains a higher order pole in the λ complex plane. We use the following form of the L and M operator

$$L = \lambda A + B, \quad \text{and} \quad M = \Delta^{-1}(\lambda^2 C + \widetilde{\omega}_1 \lambda D + \widetilde{\omega}_0 E)\,, \tag{C.2.1}$$

where now $\Delta = (\lambda - \omega_1)(\lambda - \omega_0)(\lambda + \omega_1)$. At this stage, we have ω_0, ω_1 and $\widetilde{\omega}_0, \widetilde{\omega}_1$ as free real parameters but we will see that to obtain complete integrability we will only have a single parameter. We will need relations

different than (C.1.2) involving Δ given by

$$\begin{aligned}
\Delta^{-1} &= \frac{1}{2\omega_1(\omega_1+\omega_0)}\frac{1}{\lambda+\omega_1} + \frac{1}{2\omega_1(\omega_1-\omega_0)}\frac{1}{\lambda-\omega_1} \\
&\qquad - \frac{1}{(\omega_1-\omega_0)(\omega_1+\omega_0)}\frac{1}{\lambda-\omega_0} \\
&:= \alpha_+\Delta_{+1}^{-1} + \alpha_-\Delta_{-1}^{-1} - \alpha_0\Delta_0^{-1} \\
\lambda\Delta^{-1} &= -\frac{1}{2(\omega_1+\omega_0)}\frac{1}{\lambda+\omega_1} + \frac{1}{2(\omega_1-\omega_0)}\frac{1}{\lambda-\omega_1} \\
&\qquad - \frac{\omega_0}{(\omega_1-\omega_0)(\omega_1+\omega_0)}\frac{1}{\lambda-\omega_0} \\
&:= -\omega_1\alpha_+\Delta_{+1}^{-1} + \omega_1\alpha_-\Delta_{-1}^{-1} - \omega_0\alpha_0\Delta_0^{-1} \\
\lambda^2\Delta^{-1} &= \frac{\omega_1}{2(\omega_1+\omega_0)}\frac{1}{\lambda+\omega_1} + \frac{\omega_1}{2(\omega_1-\omega_0)}\frac{1}{\lambda-\omega_1} \\
&\qquad - \frac{\omega_0^2}{(\omega_1-\omega_0)(\omega_1+\omega_0)}\frac{1}{\lambda-\omega_0} \\
&:= \omega_1^2\alpha_+\Delta_{+1}^{-1} + \omega_1^2\alpha_-\Delta_{-1}^{-1} - \omega_0^2\alpha_0\Delta_0^{-1} \\
\lambda^3\Delta^{-1} &= -\frac{\omega_1^2}{2(\omega_1+\omega_0)}\frac{1}{\lambda+\omega_1} + \frac{\omega_1^2}{2(\omega_1-\omega_0)}\frac{1}{\lambda-\omega_1} \\
&\qquad - \frac{\omega_0^3}{(\omega_1-\omega_0)(\omega_1+\omega_0)}\frac{1}{\lambda-\omega_0} + 1 \\
&:= -\omega_1^3\alpha_+\Delta_{+1}^{-1} + \omega_1^3\alpha_-\Delta_{-1}^{-1} - \omega_0^3\alpha_0\Delta_0^{-1} + 1\,,
\end{aligned}$$

where we have defined the quantities

$$\Delta_0 = \lambda - \omega_0, \qquad \Delta_{+1} = \lambda + \omega_1 \quad \text{and} \quad \Delta_{-1} = \lambda - \omega_1\,,$$

as well as α_+, α_- and α_-. We proceed as before and expand the commutator of the ZCR using (C.2.1) to get

$$\begin{aligned}
&\Delta^{-1}\lambda^3 : [A, C] \\
&\Delta^{-1}\lambda^2 : \widetilde{\omega}_1[A, D] + [B, C] \\
&\Delta^{-1}\lambda^1 : \widetilde{\omega}_0[A, E] + \widetilde{\omega}_1[B, D] \\
&\Delta^{-1}\lambda^0 : \widetilde{\omega}_0[B, E]\,.
\end{aligned}$$

Then, using the previous relations for the inverse of Δ, we arrive at

$$\begin{aligned}
\lambda^0 &: [A,C] \\
\Delta_0^{-1} &: -\omega_0^3[A,C] - \omega_0^2(\widetilde{\omega}_1[A,D]+[B,C]) - \omega_0(\widetilde{\omega}_0[A,E]+\widetilde{\omega}_1[B,D]) - \widetilde{\omega}_0[B,E] \\
\Delta_{-1}^{-1} &: \omega_1^3[A,C] + \omega_1^2(\widetilde{\omega}_1[A,D]+[B,C]) + \omega_1(\widetilde{\omega}_0[A,E]+\widetilde{\omega}_1[B,D]) + \widetilde{\omega}_0[B,E] \\
\Delta_{+1}^{-1} &: -\omega_1^3[A,C] + \omega_1^2(\widetilde{\omega}_1[A,D]+[B,C]) - \omega_1(\widetilde{\omega}_0[A,E]+\widetilde{\omega}_1[B,D]) + \widetilde{\omega}_0[B,E]\,,
\end{aligned}$$

where we discarded the α coefficients. The space derivative of M similarly expands to

$$\begin{aligned}
\Delta_0^{-1} &: -\omega_0^2 C_x - \widetilde{\omega}_1\omega_0 D_x - \widetilde{\omega}_0 E_x \\
\Delta_{-1}^{-1} &: \omega_1^2 C_x + \widetilde{\omega}_1\omega_1 D_x + \widetilde{\omega}_0 E_x \\
\Delta_{+1}^{-1} &: \omega_1^2 C_x - \widetilde{\omega}_1\omega_1 D_x + \widetilde{\omega}_0 E_x\,.
\end{aligned}$$

As before, we have a set of dependent equations which we need to solve. One can see that we need to impose some relationships between the ω parameters. The most general that we can do is to set $\omega_1 = \omega_0 = \widetilde{\omega}_0 = \omega$ and $\widetilde{\omega}_1 = \omega_0$ so that the terms coming from the commutator simplify to

$$\begin{aligned}
\lambda^0 &: [A,C] \\
\Delta_0^{-1} &: -\omega^3[A,C] - \omega^2(\omega_0[A,D]+[B,C]) - \omega(\omega[A,E]+\omega_0[B,D]) - \omega[B,E] \\
\Delta_{-1}^{-1} &: \omega^3[A,C] + \omega^2(\omega_0[A,D]+[B,C]) + \omega(\omega[A,E]+\omega_0[B,D]) + \omega[B,E] \\
\Delta_{+1}^{-1} &: -\omega^3[A,C] + \omega^2(\omega_0[A,D]+[B,C]) - \omega(\omega[A,E]+\omega_0[B,D]) + \omega[B,E]\,,
\end{aligned}$$

and the terms from M_x are

$$\begin{aligned}
\Delta_0^{-1} &: -\omega^2 C_x - \omega_0\omega D_x - \omega E_x \\
\Delta_{-1}^{-1} &: \omega^2 C_x + \omega_0\omega D_x + \omega E_x \\
\Delta_{+1}^{-1} &: \omega^2 C_x - \omega_0\omega D_x + \omega E_x\,.
\end{aligned}$$

After combining all the terms from the ZCR, we arrive at

$$\begin{aligned}
&B_t^\perp = -[A^\parallel, C^\perp], \quad E_x^\parallel = \omega[B^\perp, C^\perp], \quad D_x^\parallel = [B^\perp, D^\perp], \\
&C_x^\perp = \omega_0[A^\parallel, D^\perp] + \frac{1}{\omega}[B^\perp, E^\parallel], \quad D_x^\perp = \frac{\omega^2}{\omega_0}[A^\parallel, C^\perp] + [B^\perp, D^\parallel]\,,
\end{aligned} \tag{C.2.2}$$

where we recall that the notation $\perp$ means the e, f components and $\parallel$ the h, or diagonal component of the matrix. We are now going to change variables

to make the rest of the computation simpler. We set

$$\begin{aligned} E &:= B^e = B^f \\ iP &:= C^e = -C^f \\ Q &:= D^e = D^f \\ \text{and} \quad iN &:= E^h \,, \end{aligned} \tag{C.2.3}$$

where the upper index notation corresponds to the Lie algebra basis e, f, h of $\mathfrak{sl}(2)$. By a direct substitution in (C.2.2), we obtain the following equations of motion for E, P, Q and N

$$\begin{aligned} E_t &= 2P \\ P_x &= -2\omega_0 Q + \frac{2}{\omega} NE \\ Q_x &= -2\frac{\omega^2}{\omega_0} P \\ N_x &= 2\omega EP \,. \end{aligned} \tag{C.2.4}$$

Without changing the equation we rescale $N \to \omega N$, absorbe a factor of 2, use $\omega_0 = \omega$ and change $x \leftrightarrow t$ to obtain the final hyperbolic RMB equations (3.6.1) with $\sigma_2 = -1$.

Index

AKNS hierarchy, 142, 171

Bi-Hamiltonian structure, 157, 173

Central extension, 131, 191
CH-NLS, 150, 155
Clebsch, 20, 169

Double bracket dissipation, 43, 98

Euler-Poincaré reduction, 23

Fokker-Planck equation, 31, 40, 48, 82, 112

Gibbs distribution, 48, 52

Hamilton-Pontryagin principle, 26, 46

Isospectral problem, 173, 194

Landmarks dynamics, 108
Lyapunov exponents, 60, 67

Madelung variable, 187, 208
Modulational instability, 164, 200
Moment dynamics, 113

Random attractors, 54, 69
Reconstruction relation, 23, 110
RMB equations, 192

Semi-direct product, 27
Sobolev deformation, 146
Soliton, 158
Solitons, 175, 195
SRB measure, 59
Stationary distribution, 35
Stochastic Euler-Poincaré equation, 22, 24, 110, 171
Stochastic free rigid body, 36, 67, 85
Stochastic integrability, 30
Stochastic integrable equations, 176
Stochastic nonlinear stability, 80

References

Ablowitz, M. and Segur, H. (1981). *Solitons and the Inverse Scattering Transform*. Society for Industrial and Applied Mathematics.

Ablowitz, M. J., Kaup, D. J., Newell, A. C., and Segur, H. (1973a). Method for solving the sine-gordon equation. *Physical Review Letters*, 30(25):1262.

Ablowitz, M. J., Kaup, D. J., Newell, A. C., and Segur, H. (1973b). Nonlinear-evolution equations of physical significance. *Physical Review Letters*, 31(2):125.

Ablowitz, M. J., Prinari, B., and Trubatch, A. D. (2004). *Discrete and continuous nonlinear Schrödinger systems*, volume 302. Cambridge University Press.

Abraham, R., Marsden, J. E., and Marsden, J. E. (1978). *Foundations of mechanics*. Benjamin/Cummings Publishing Company Reading, Massachusetts.

Agafontsev, D. and Zakharov, V. E. (2015). Integrable turbulence and formation of rogue waves. *Nonlinearity*, 28(8):2791.

Agrawal, G. P. (2007). *Nonlinear fiber optics*. Academic press.

Aiyer, R. N. (1983). Hamiltonian and recursion operator for the reduced maxwell-bloch equations. *Journal of Physics A: Mathematical and General*, 16(8):1809.

Amiranashvili, S., Bandelow, U., and Akhmediev, N. (2011). Dispersion of nonlinear group velocity determines shortest envelope solitons. *Physical Review A*, 84(4).

Amiranashvili, S., Bandelow, U., and Akhmediev, N. (2014). Ultrashort optical solitons in transparent nonlinear media with arbitrary dispersion. *Optical and Quantum Electronics*, 46(10):1233–1238.

Arnaudon, A. (2015). The integrable stochastic AKNS hierarchy. *arXiv preprint arXiv:1511.07080*.

Arnaudon, A. (2016a). On a deformation of the nonlinear Schrödinger equation. *Journal of Physics A: Mathematical and Theoretical*, 49(12):125202.

Arnaudon, A. (2016b). On a lagrangian reduction and a deformation of completely integrable systems. *Journal of Nonlinear Science*, 26(5):1133–1160.

Arnaudon, A., Castrillón López, M., and Holm, D. D. (2015). Covariant unreduction for curve matching. *Mathematical Foundation of Computational Anatomy 2015, arXiv:1508.05325*.

Arnaudon, A., De Castro, A. L., and Holm, D. D. (2016). Noise and dissipation in rigid body motion. In *Stochastic Geometric Mechanics, Springer*.

Arnaudon, A., De Castro, A. L., and Holm, D. D. (2017a). Noise and dissipation on coadjoint orbits. *To appear in: Journal of Nonlinear Science*.

Arnaudon, A., Ganaba, N., and Holm, D. D. (2017b). The stochastic energy-Casimir method. *arXiv:1702.03899*.

Arnaudon, A. and Gibbon, J. D. (2017). Integrability of the hyperbolic reduced Maxwell-Bloch equations for strongly correlated Bose-Einstein condensates. *Physical Review A*, 96(1):013610.

Arnaudon, A., Holm, D. D., and Ivanov, R. I. (2017c). G-strands on symmetric spaces. *Proc. R. Soc. A*, 473(2199):20160795.

Arnaudon, A., Holm, D. D., Pai, A., and Sommer, S. (2017d). A stochastic large deformation model for computational anatomy. In *Information Processing in Medical Imaging*, pages 571–582.

Arnaudon, A., Holm, D. D., and Sommer, S. (2017e). A geometric framework for stochastic shape analysis. *arXiv preprint arXiv:1703.09971*.

Arnaudon, A., Lopez, M. C., and Holm, D. D. (2017f). Un-reduction in field theory. *To appear in: Letters in Mathematical Physics*.

Arnaudon, M., Chen, X., and Cruzeiro, A. B. (2014). Stochastic Euler-Poincaré reduction. *Journal of Mathematical Physics*, 55(8):081507.

Arnold, L. (1995). *Random dynamical systems*. Springer.

Arnold, V. (1965). Conditions for non-linear stability of plane steady curvilinear flows of an ideal fluid. *Doklady Akademii Nauk SSSR*, 162(5):773–777.

Arnol'd, V. (1966a). Sur la géométrie différentielle des groupes de Lie de dimension infinie et ses applications à l'hydrodynamique des fluides parfaits. *Annales de l'institut Fourier*, 16(1):319–361.

Arnol'd, V. I. (1966b). An a priori estimate in the theory of hydrodynamic stability. *Izvestiya Vysshikh Uchebnykh Zavedenii. Matematika*, (5):3–5.

Arnol'd, V. I. (1989). *Mathematical Methods of Classical Mechanics*, volume 60 of *Graduate Texts in Mathematics*. Springer-Verlag, New York.

Baxendale, P. H. (1991). *Spatial Stochastic Processes: A Festschrift in Honor of Ted Harris on his Seventieth Birthday*, chapter Statistical Equilibrium and Two-Point Motion for a Stochastic Flow of Diffeomorphisms, pages 189–218. Birkhäuser Boston, Boston, MA.

Benettin, G., Christodoulidi, H., and Ponno, A. (2013). The Fermi-Pasta-Ulam problem and its underlying integrable dynamics. *Journal of Statistical Physics*, 152(2):195–212.

Benettin, G. and Ponno, A. (2011). Time-scales to equipartition in the Fermi–Pasta–Ulam problem: finite-size effects and thermodynamic limit. *Journal of Statistical Physics*, 144(4):793–812.

Benjamin, T. B. (1972). The stability of solitary waves. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 328(1573):153–183.

Benjamin, T. B. and Feir, J. (1967). The disintegration of wave trains on deep water part 1. theory. *Journal of Fluid Mechanics*, 27(03):417–430.

Bismut, J.-M. (1982). Mécanique aléatoire. In *Tenth Saint Flour Probability Summer School—1980 (Saint Flour, 1980)*, volume 929 of *Lecture Notes in Math.*, pages 1–100. Springer, Berlin-New York.

Bloch, A., Baillieul, J., Crouch, P., Marsden, J. E., Krishnaprasad, P. S., Murray, R., and Zenkov, D. (2003). *Nonholonomic mechanics and control*, volume 24. Springer.

Bloch, A., Krishnaprasad, P., Marsden, J. E., and Ratiu, T. S. (1996). The Euler-Poincaré equations and double bracket dissipation. *Communications in Mathematical Physics*, 175(1):1–42.

Bloch, A. M., Brockett, R. W., and Ratiu, T. S. (1992). Completely integrable gradient flows. *Communications in Mathematical Physics*, 147(1):57–74.

Bona, J. (1975). On the stability theory of solitary waves. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 344(1638):363–374.

Bona, J. L., Souganidis, P. E., and Strauss, W. A. (1987). Stability and instability of solitary waves of Korteweg-de Vries type. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 411(1841):395–412.

Brockett, R. (1991). Dynamical systems that sort lists, diagonalize matrices, and solve linear programming problems. *Linear Algebra and its Applications*, 146:79–91.

Burger, S., Bongs, K., Dettmer, S., Ertmer, W., Sengstock, K., Sanpera, A., Shlyapnikov, G. V., and Lewenstein, M. (1999). Dark solitons in Bose-Einstein condensates. *Physical Review Letters*, 83:5198–5201.

Burns, K., Brown, B., Lecoanet, D., Oishi, J., and Vasil, G. (2016). Dedalus: Flexible framework for spectrally solving differential equations. *Astrophysics Source Code Library*.

Camassa, R. and Holm, D. D. (1993). An integrable shallow water equation with peaked solitons. *Physical Review Letters*, 71(11):1661–1664.

Castrillón López, M., García Pérez, P. L., and Ratiu, T. S. (2001). Euler-Poincaré reduction on principal bundles. *Letters in Mathematical Physics*, 58(2):167–180.

Caudrey, P. J., Eilbeck, J. C., and Gibbon, J. D. (1974). Exact multisoliton solution of the reduced maxwell-bloch equations of non-linear optics. *IMA Journal of Applied Mathematics*, 14(3):375–386.

Caudrey, P. J., Eilbeck, J. C., Gibbon, J. D., and Bullough, R. K. (1973). Exact multisoliton solutions of the self-induced transparency and sine-gordon equations. *Physical Review Letters*, 30:273– 278.

Cendra, H., Marsden, J. E., and Ratiu, T. S. (2001). *Lagrangian reduction by stages*. Number 722. American Mathematical Society.

Chandler, D. (1987). *Introduction to Modern Statistical Mechanics*. Oxford University Press.

Chen, X., Cruzeiro, A. B., and Ratiu, T. S. (2015). Constrained and stochastic variational principles for dissipative equations with advected quantities. *arXiv preprint arXiv:1506.05024*.

Chirikjian, G. S. (2009). *Stochastic models, information theory, and Lie groups. Vol. 1*. Applied and Numerical Harmonic Analysis. Birkhäuser Boston, Inc., Boston, MA. Classical results and geometric methods.

Chirikjian, G. S. (2012). *Stochastic models, information theory, and Lie groups. Volume 2*. Applied and Numerical Harmonic Analysis. Birkhäuser/Springer, New York. Analytic methods and modern applications.

Colin, M. and Lannes, D. (2009). Short pulses approximations in dispersive media. *SIAM Journal on Mathematical Analysis*, 41(2):708–732.

Crauel, H. (1991). Markov measures for random dynamical systems. *Stochastics: An International Journal of Probability and Stochastic Processes*, 37(3):153–173.

Crauel, H. and Flandoli, F. (1994). Attractors for random dynamical systems. *Probability Theory and Related Fields*, 100(3):365–393.

Crauel, H. and Flandoli, F. (1998). Additive noise destroys a pitchfork bifurcation. *Journal of Dynamics and Differential Equations*, 10(2):259–274.

Cruzeiro, A. B., Holm, D. D., and Ratiu, T. S. (2016). Momentum maps and stochastic clebsch action principles. *arXiv preprint arXiv:1604.04554*.

Da Prato, G. and Zabczyk, J. (2014). *Stochastic equations in infinite dimensions*. Cambridge university press.

Date, E., Kashiwara, M., Jimbo, M., and Miwa, T. (1983). *Transformation groups for soliton equations*. World Scientific Publishing, Singapore.

Dauxois, T. (2008). Fermi, pasta, ulam, and a mysterious lady. *Physics today*, 61(1):55–57.

de Bouard, A. and Debussche, A. (1998). On the stochastic Korteweg-de Vries equation. *Journal of Functional Analysis*, 154(1):215 – 251.

de Bouard, A. and Debussche, A. (2003). The stochastic nonlinear Schrödinger equation in H1. *Stochastic analysis and applications*, 21(1):97–126.

de Bouard, A. and Debussche, A. (2010). The nonlinear schrdinger equation with white noise dispersion. *Journal of Functional Analysis*, 259(5):1300 – 1321.

de Bouard, A., Debussche, A., and Tsutsumi, Y. (1999). White noise driven Korteweg-de Vries equation. *Journal of Functional Analysis*, 169(2):532 – 558.

Debussche, A. and Di Menza, L. (2002). Numerical simulation of focusing stochastic nonlinear Schrödinger equations. *Physica D: Nonlinear Phenomena*, 162(3):131–154.

Debussche, A. and Printems, J. (1999). Numerical simulation of the stochastic Korteweg-de Vries equation. *Physica D: Nonlinear Phenomena*, 134(2):200 – 226.

Deift, P., Li, L., Nanda, T., and Tomei, C. (1986). The Toda flow on a generic orbit is integrable. *Communications on pure and applied mathematics*, 39(2):183–232.

Deift, P., Li, L.-C., and Tomei, C. (1989). Matrix factorizations and integrable systems. *Communications on pure and applied mathematics*, 42(4):443–521.

Deift, P. and McLaughlin, K. T. (1998). *A continuum limit of the Toda lattice.* Number 624. American Mathematical Society.

Dodd, R. K., Eilbeck, J. C., Gibbon, J. D., and Morris, H. C. (1982). *Solitons and nonlinear wave equations.* Academic Press, Inc., London-New York, 1982.

Dudley, J. M., Dias, F., Erkintalo, M., and Genty, G. (2014). Instabilities, breathers and rogue waves in optics. *Nature Photonics*, 8(10):755–764.

Dullin, H. R., Gottwald, G. A., and Holm, D. D. (2004). On asymptotically equivalent shallow water wave equations. *Physica D. Nonlinear Phenomena*, 190(1-2):1–14.

Ebin, D. G. and Marsden, J. (1970). Groups of diffeomorphisms and the motion of an incompressible fluid. *Annals of Mathematics*, pages 102–163.

Eckmann, J.-P. and Ruelle, D. (1985). Ergodic theory of chaos and strange attractors. *Reviews of modern physics*, 57(3):617.

Eilbeck, J. C., Gibbon, J. D., Caudrey, P. J., and Bullough, R. K. (1973). Solitons in nonlinear optics. i. a more accurate description of the 2 pulse in self-induced transparency. *Journal of Physics A: Mathematical, Nuclear and General*, 6(9):1337.

Einstein, A. (1905). Über die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen. *Annalen der physik*, 322(8):549–560.

Elworthy, K. D. (1982). *Stochastic differential equations on manifolds*, volume 70. Cambridge University Press.

Engel, M., Lamb, J. S., and Rasmussen, M. (2016). Bifurcation analysis of a stochastically driven limit cycle. *arXiv preprint arXiv:1606.01137*.

Fermi, E., Pasta, J., and Ulam, S. (1955). Studies of nonlinear problems. *Los Alamos Report LA-1940*, 978.

Flandoli, F., Gess, B., and Scheutzow, M. (2016). Synchronization by noise. *Probability Theory and Related Fields*, pages 1–46.

Flaschka, H. (1974a). On the Toda lattice. ii inverse-scattering solution. *Progress of Theoretical Physics*, 51(3):703–716.

Flaschka, H. (1974b). The Toda lattice. ii. existence of integrals. *Physical Review B*, 9(4):1924.

Flaschka, H., Newell, A. C., and Ratiu, T. (1983). Kac-Moody Lie algebras and soliton equations. II. Lax equations associated with $A_1^{(1)}$. *Physica D. Nonlinear Phenomena*, 9(3):300–323.

Fokas, A. (1995). On a class of physically important integrable equations. *Physica D: Nonlinear Phenomena*, 87(1):145–150.

Fokas, A. S. and Liu, Q. M. (1996). Asymptotic integrability of water waves. *Physical Review Letters*, 77(12):2347–2351.

Fokker, A. D. (1914). Die mittlere energie rotierender elektrischer dipole im strahlungsfeld. *Annalen der Physik*, 348(5):810–820.

Fowler, A., Gibbon, J., and McGuinness, M. (1982). The complex Lorenz equations. *Physica D: Nonlinear Phenomena*, 4(2):139–163.

Fowler, A., Gibbon, J., and McGuinness, M. (1983). The real and complex Lorenz equations and their relevance to physical systems. *Physica D: Nonlinear Phenomena*, 7(1-3):126–134.

Fuchssteiner, B. and Fokas, A. S. (1981). Symplectic structures, their bäcklund transformations and hereditary symmetries. *Physica D: Nonlinear Phenomena*, 4(1):47–66.

Fusca, D. (2016). The Madelung transform as a momentum map. *arXiv preprint arXiv:1512.04611*.

Gardiner, C. (2009). *Stochastic Methods*. Springer-Verlag, Berlin.

Gardner, C. S., Greene, J. M., Kruskal, M. D., and Miura, R. M. (1967). Method for solving the Korteweg-de Vries equation. *Physical Review Letters*, 19(19):1095.

Gay-Balmaz, F. and Holm, D. D. (2013). Selective decay by Casimir dissipation in inviscid fluids. *Nonlinearity*, 26(2):495.

Gay-Balmaz, F. and Holm, D. D. (2014). A geometric theory of selective decay with applications in MHD. *Nonlinearity*, 27(8):1747.

Gay-Balmaz, F. and Ratiu, T. S. (2009). The geometric structure of complex fluids. *Advances in Applied Mathematics*, 42(2):176–275.

Gay-Balmaz, F. and Ratiu, T. S. (2011). Clebsch optimal control formulation in mechanics. *Journal of Geometric Mechanics*, 3(1):41–79.

Gibbon, J. and McGuinness, M. (1980). A derivation of the Lorenz equations for some unstable dispersive physical systems. *Physics Letters A*, 77(5):295–299.

Gibbon, J. and McGuinness, M. (1982). The real and complex Lorenz equations in rotating fluids and lasers. *Physica D: Nonlinear Phenomena*, 5(1):108–122.

Gibbon, J. D., Caudrey, P. J., Bullough, R. K., and Eilbeck, J. C. (1973). An N-soliton solution of a nonlinear optics equation derived by a general inverse method. *Lettere Al Nuovo Cimento (1971–1985)*, 8(13):775–779.

Gotay, M. J., Isenberg, J., and Marsden, J. E. (1997). Momentum maps and classical relativistic fields. Part 1: Covariant Field Theory. *arXiv preprint arXiv:9801019*.

Grillakis, M., Shatah, J., and Strauss, W. (1987). Stability theory of solitary waves in the presence of symmetry, i. *Journal of Functional Analysis*, 74(1):160–197.

Hairer, M. (2011). On Malliavin's proof of Hörmanders theorem. *Bulletin des sciences mathematiques*, 135(6):650–666.

Hasegawa, A. and Kodama, Y. (1995). *Solitons in optical communications*. Number 7. Oxford University Press, USA.

Henón, M. (1974). Integrals of the Toda lattice. *Physical Review B*, 9(4):1921.

Hirota, R. (1973). Exact envelope-soliton solutions of a nonlinear wave equation. *Journal of Mathematical Physics*, 14(7):805–809.

Holm, D. D. (2008). *Geometric mechanics. Part II.* Imperial College Press, London. Rotating, Translating and Rolling.

Holm, D. D. (2015). Variational principles for stochastic fluid dynamics. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 471(2176).

Holm, D. D. and Kupershmidt, B. A. (1983). Poisson brackets and clebsch representations for magnetohydrodynamics, multifluid plasmas, and elasticity. *Physica D: Nonlinear Phenomena*, 6(3):347–363.

Holm, D. D. and Marsden, J. E. (2005). *Momentum maps and measure-valued solutions (peakons, filaments, and sheets) for the EPDiff equation*, volume 232 of *Progr. Math.* Birkhäuser Boston, Boston, MA.

Holm, D. D., Marsden, J. E., Ratiu, T., and Weinstein, A. (1985). Nonlinear stability of fluid and plasma equilibria. *Physics Reports*, 123(1):1 – 116.

Holm, D. D., Marsden, J. E., and Ratiu, T. S. (1998a). The Euler–Poincaré equations and semidirect products with applications to continuum theories. *Advances in Mathematics*, 137(1):1–81.

Holm, D. D., Marsden, J. E., and Ratiu, T. S. (1998b). Euler-Poincaré models of ideal fluids with nonlinear dispersion. *Physical Review Letters*, 80:4173–4176.

Holm, D. D., Putkaradze, V., and Tronci, C. (2007). Geometric dissipation in kinetic equations. *Comptes Rendus Mathematique*, 345(5):297 – 302.

Holm, D. D. and Tyranowski, T. M. (2016). Variational principles for stochastic soliton dynamics. *Proc. R. Soc. A*, 472(2187):20150827.

Hörmander, L. (1967). Hypoelliptic second order differential equations. *Acta Mathematica*, 119(1):147–171.

Ikeda, N. and Watanabe, S. (1981). *Stochastic differential equations and diffusion processes*. North-Holland, Kodansha.

Itô, K. (1944). Stochastic integral. *Proceedings of the Imperial Academy*, 20(8):519–524.

Khasminskii, R. (2011). *Stochastic stability of differential equations*, volume 66. Springer Science & Business Media.

Khesin, B. and Misioek, G. (2003). Euler equations on homogeneous spaces and virasoro orbits. *Advances in Mathematics*, 176(1):116 – 144.

Khesin, B. and Wendt, R. (2008). *The geometry of infinite-dimensional groups*, volume 51. Springer Science & Business Media.

Kira, M. (2014). Excitation picture of an interacting Bose gas. *Ann. Phys.*, 351:200–249.

Kira, M. (2015a). Coherent quantum depletion of an interacting atom condensate. *Nature communications*, 6.

Kira, M. (2015b). Hyperbolic bloch equations: atom-cluster kinetics of an interacting Bose gas. *Annals of Physics*, 356:185–243.

Kira, M. and Koch, S. W. (2011). *Semiconductor quantum optics*. Cambridge University Press.

Kloeden, P. E. and Rasmussen, M. (2011). *Nonautonomous dynamical systems*. Number 176. American Mathematical Society.

Kolmogoroff, A. (1931). Über die analytischen methoden in der wahrscheinlichkeitsrechnung. *Mathematische Annalen*, 104(1):415–458.

Korteweg, D. J. and De Vries, G. (1895). On the change of form of long waves advancing in a rectangular canal, and on a new type of long stationary waves. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 39(240):422–443.

Kostant, B. (1979). The solution to a generalized Toda lattice and representation theory. *Advances in Mathematics*, 34(3):195–338.

Kubo, R. (1966). The fluctuation-dissipation theorem. *Reports on progress in physics*, 29(1):255.

Kubo, R., Toda, M., and Hashitsume, N. (1991). *Statistical physics. II*, volume 31 of *Springer Series in Solid-State Sciences*. Springer-Verlag, Berlin, second edition. Nonequilibrium statistical mechanics.

Lakshmanan, M. (2011). The fascinating world of the Landau-Lifshitz-Gilbert equation: an overview. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 369(1939):1280–1300.

Landau, L. and Lifshitz, E. (1980). Statistical physics, vol. 5. *Course of theoretical physics*, 30.

Lax, P. D. (1968). Integrals of nonlinear equations of evolution and solitary waves. *Communications on pure and applied mathematics*, 21(5):467–490.

Lázaro-Camí, J.-A. and Ortega, J.-P. (2008). Stochastic Hamiltonian dynamical systems. *Reports on Mathematical Physics*, 61(1):65–122.

Le Jan, Y. (1987). Équilibre statistique pour les produits de difféomorphismes aléatoires indépendants. *Annales de l'IHP Probabilités et statistiques*, 23(1):111–120.

Ledrappier, F. and Young, L.-S. (1988). Entropy formula for random transformations. *Probability Theory and Related Fields*, 80(2):217–240.

Li, X.-M. (2008). An averaging principle for a completely integrable stochastic hamiltonian system. *Nonlinearity*, 21(4):803.

Lin, K. K. and Young, L.-S. (2008). Shear-induced chaos. *Nonlinearity*, 21(5):899.

Longstaff, F. A. and Schwartz, E. S. (1992). Interest rate volatility and the term structure: A two-factor general equilibrium model. *The Journal of Finance*, 47(4):1259–1282.

Lorenz, E. N. (1963). Deterministic nonperiodic flow. *Journal of the atmospheric sciences*, 20(2):130–141.

Lyons, T. J. (1998). Differential equations driven by rough signals. *Revista Matemática Iberoamericana*, 14(2):215–310.

Madelung, E. (1927). Quantentheorie in hydrodynamischer form. *Zeitschrift für Physik A Hadrons and Nuclei*, 40(3):322–326.

Magri, F. (1978). A simple model of the integrable hamiltonian equation. *Journal of Mathematical Physics*, 19:1156.

Marsden, J. and Weinstein, A. (1974). Reduction of symplectic manifolds with symmetry. *Reports on mathematical physics*, 5(1):121–130.

Marsden, J. and Weinstein, A. (1983). Coadjoint orbits, vortices, and clebsch variables for incompressible fluids. *Physica D: Nonlinear Phenomena*, 7(1):305–323.

Marsden, J. E., Misiołek, G., Ortega, J.-P., Perlmutter, M., and Ratiu, T. S. (2007). *Hamiltonian reduction by stages*, volume 1913 of *Lecture Notes in Mathematics*. Springer, Berlin.

Marsden, J. E., Raţiu, T., and Weinstein, A. (1984). Semidirect products and reduction in mechanics. *Transactions of the American Mathematical Society*, 281(1):147–177.

Marsden, J. E. and Ratiu, T. S. (1999). *Introduction to mechanics and symmetry*, volume 17 of *Texts in Applied Mathematics*. Springer-Verlag, New York, second edition. A basic exposition of classical mechanical systems.

Marsland, S. and Shardlow, T. (2017). Langevin equations for landmark image registration with uncertainty. *SIAM Journal on Imaging Sciences*, 10(2):782–807.

Misiołek, G. (1998). A shallow water equation as a geodesic flow on the Bott-Virasoro group. *Journal of Geometry and Physics*, 24(3):203–208.

Newell, A. C. (1985). *Solitons in mathematics and physics*, volume 48 of *CBMS-NSF Regional Conference series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA.

Øksendal, B. (2003). *Stochastic differential equations*. Springer.

Olver, P. J. and Rosenau, P. (1996). Tri-hamiltonian duality between solitons and solitary-wave solutions having compact support. *Physical Review E*, 53(2):1900.

Oseledec, V. I. (1968). A multiplicative ergodic theorem. Lyapunov characteristic numbers for dynamical systems. *Trans. Moscow Math. Soc*, 19(2):197–231.

Pavliotis, G. A. (2014). *Stochastic processes and applications*, volume 60 of *Texts in Applied Mathematics*. Springer, New York.

Peregrine, D. (1983). Water waves, nonlinear Schrödinger equations and their solutions. *The Journal of the Australian Mathematical Society. Series B. Applied Mathematics*, 25(01):16–43.

Planck, M. (1917). *Über einen Satz der statistischen Dynamik und seine Erweiterung in der Quantentheorie*. Reimer.

Poincaré, H. (1901). Sur une forme nouvelle des équations de la mécanique. *Comptes rendus de l'Acadmie des sciences*, 132:369–371.

Pressley, A. and Segal, G. (1986). *Loop groups*. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York.

Qiao, Z. (2007). New integrable hierarchy, its parametric solutions, cuspons, one-peak solitons, and M/W-shape peak solitons. *Journal of Mathematical Physics*, 48(8):082701, 20.

Reyman, A. and Semenov-Tian-Shansky, M. (1988). Compatible poisson structures for lax equations: An r-matrix approach. *Physics Letters A*, 130(8–9):456 – 460.

Ruelle, D. (1979). Ergodic theory of differentiable dynamical systems. *Publications Mathématiques de l'Institut des Hautes Études Scientifiques*, 50(1):27–58.

Russell, J. S. (1845). *Report on Waves: Made to the Meetings of the British Association in 1842-43*. R. and J. Taylor.

Schenk-Hoppé, K. R. (1998). Random attractors-general properties, existence and applications to stochastic bifurcation theory. *Discrete and Continuous Dynamical Systems*, 4:99–130.

Semenov-Tyan-Shanskiĭ, M. A. (1983). What is a classical r-matrix? *Functional analysis and its application*, 17(4):17–33.

Simo, J. C., Lewis, D., and Marsden, J. E. (1991). Stability of relative equilibria. part i: The reduced energy-momentum method. *Archive for Rational Mechanics and Analysis*, 115(1):15–59.

Sparrow, C. (1982). The Lorenz equations: Bifurcations, chaos and strange attractors. *Applied Mathematical Sciences, New York: Springer, 1982*.

Storn, R. and Price, K. (1997). Differential evolution–a simple and efficient heuristic for global optimization over continuous spaces. *Journal of global optim.*, 11(4):341–359.

Stratonovich, R. (1966). A new representation for stochastic integrals and equations. *SIAM Journal on Control*, 4(2):362–371.

Sulem, C. and Sulem, P.-L. (2007). *The nonlinear Schrödinger equation: self-focusing and wave collapse*, volume 139. Springer Science & Business Media.

Tao, T. (2009). Why are solitons stable? *Bulletin of the American Mathematical Society*, 46(1):1–33.

Theano Development Team (2016). Theano: A Python framework for fast computation of mathematical expressions. *arXiv e-prints*, abs/1605.02688.

Toda, M. (1967). Vibration of a chain with nonlinear interaction. *Journal of the Physical Society of Japan*, 22(2):431–436.

Toda, M. (1989). *Theory of Nonlinear Lattices*, volume 20 of *Springer Series in Solid-State Sciences*. Springer.

Trouvé, A. and Vialard, F.-X. (2012). Shape splines and stochastic shape evolutions: a second order point of view. *Quarterly of Applied Mathematics*, 70(2):219–251.

Vialard, F.-X. (2013). Extension to infinite dimensions of a stochastic second-order model associated with shape splines. *Stochastic Processes and their Applications*, 123(6):2110–2157.

Viana, M. (2014). *Lectures on Lyapunov exponents*, volume 145. Cambridge University Press.

Wadati, M. (1983). Stochastic Korteweg-de Vries equation. *Journal of the Physical Society of Japan*, 52(8):2642–2648.

Wang, Q. and Young, L.-S. (2003). Strange attractors in periodically-kicked limit cycles and hopf bifurcations. *Communications in Mathematical Physics*, 240(3):509–529.

Xia, B. and Qiao, Z. (2015). A new two-component integrable system with peakon solutions. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 471(2175).

Xia, B., Qiao, Z., and Zhou, R. (2015). A synthetical two-component model with peakon solutions. *Studies in Applied Mathematics*, 135(3):248–276.

Younes, L. (2010). *Shapes and Diffeomorphisms*. Springer.

Young, L.-S. (2002). What are SRB measures, and which dynamical systems have them? *Journal of Statistical Physics*, 108(5):733–754.

Young, L.-S. (2013). Mathematical theory of Lyapunov exponents. *Journal of Physics A: Mathematical and Theoretical*, 46(25):254001.

Zakharov, V. E. (2009). Turbulence in integrable systems. *Studies in Applied Mathematics*, 122(3):219–234.

Zakharov, V. E. and Shabat, A. v. (1972). Exact theory of two-dimensional self-focussing and one-dimensional self-modulating waves in nonlinear media. *Sov. Phys.-JETP (Engl. Transl.)*, 34.

Zambrini, J.-C. (2009). On the geometry of the Hamilton-Jacobi-Bellman equation. *Journal of Geometric Mechanics*, 1(3):369–387.

Zambrini, J.-C. (2015). The research program of stochastic deformation (with a view toward geometric mechanics). In *Stochastic Analysis: A Series of Lectures*, pages 359–393. Springer.