Non-linear reduced order models for Hamiltonian systems

Internship report

M2 Calcul Scientifique et Mathématiques de l'Information

July 21, 2023

Presented by
Claire Schnoebelen

Supervised by
Emmanuel Franck
Laurent Navoret
Emmanuel Opshtein

Contents

Introduction

This document present the work I have made during my internship, as part of my second year of Masters. The internship took place at INRIA between 13 February and 11 August. I was under the supervision of Emmanuel Franck, Emmanuel Opshtein and Laurent Navoret, researchers at INRIA and IRMA.

INRIA (for Institut National de Recherche en Informatique et Automatique) is a public establishment created in 1967 within the framework of the "plan calcul", a governmental plan intended to develop French knowledge in the field of digital technology and to ensure the digital sovereignty of the country. Today, it has 200 teams spread over 10 research centers, bringing together a total of 3,900 researchers and engineers in mathematics and computer science [12]. The institute works by "équipes-projets", groups of about twenty people working on the same project and for the most part in collaboration with companies [12]. The Nancy research centre was founded in 1986 and today has 20 teams bringing together over 400 people. It has a branch at the University of Strasbourg, where researchers from the TONUS team (for TOkamaks and NUmerical Simulations) work, including one of my supervisor, Emmanuel Franck.

IRMA (for Institut de Recherche en Mathématique Avancée) is a research center in mathematics under the administrative supervision of the University of Strasbourg and the CNRS (for Centre National de la Recherche Scientifique) [13]. It has been created in 1966 as the first research centre associated to the CNRS, a public establishment itself created in 1939 in order to structure and dynamize the French public research [7]. IRMA counts about 130 members distributed in 7 research teams including the Geometry team, to which Emmanuel Opshtein belongs, and the MoCo (for Modélisation et Contrôle) team, to which Laurent Navoret belongs.

This internship is in line with last year internship, where we studied linear reduction methods for Hamiltonian systems. The aim of this year internship was to extend and improve those methods in the non-linear case. Hamiltonian systems are systems of partial differential equations whose flows have the particularity to preserve the energy of the underlying physical systems. In geometry, those problems are studied in the field of symplectic geometry. Given a partial differential equation, the objective of a reduced order model is to find a family of functions which alone can explain a large part of the behaviour of the equation solutions. The interest of reduced order modelling is that it allows a faster computation of solutions at any time and for any value of the equation parameters in the interval considered during reduction. As we observed during my last year internship, linear methods such as Proper Symplectic Decomposition (PSD) gives good results for linear Hamiltonian equations but fail for non-linear one.

The main task of my internship was to think of ways to improve the reduction we obtain with PSD. We have worked on two approaches of this problem. In the first one, we try to correct the application that sends the solutions on the reduced space given by the PSD. Here, we used quadratic corrections and mainly tried to adapt to the Hamiltonian case a method proposed in [10]. This task then involved references readings, mathematical computations to set the problem and programmation (in Python). After fruitless tests for several variations of this idea, we put it aside and start working on the second approach. The second idea is to directly build a Hamiltonian dynamic in the reduced space given by the PSD. Here, we used a control-type method with an explicit computation of the gradient to achieve it. This involved mathematical computations using the adjoint method to find an expression of the gradient we were looking for, reflexion on the resolution of the associated optimization problem and coding. The tests we have made so far have given promising results.

The second part of my internship was devoted to being familiar with some geometrical tools that we expect to use in future works to build models for Hamiltonian problems. I fulfilled this task in parallel with the first one. As the first task, it was divided into two parts that I carried out one after another. In a first time, I was asked to being familiar with generating functions, a kind a application very useful to build symplectic maps. As the central idea of Hamiltonian reduction is to find a good symplectic reduction, this tool may prove to be useful in future works on the subject. In the field of learning Hamiltonian dynamics with neural networks, this idea has already been exploited in [5]. Then, the objective was to get familiar with techniques based on what we call h-principle. Those techniques are used in symplectic geometry to prove the existence of symplectic embeddings between two manifolds. In our case, we expect to use them to prove results which would give geometrical guaranties to the reduction methods we are interested to build. To complete this task, I read [15] and [1] to learn about generating functions and [8] to learn about h-principle.

What follows is divided into two parts that correspond to the two main task I worked on, starting by the numerical part and finishing by the geometrical one. We insist more on the control approach and on the h-principle than on quadratic corrections and generating functions since it is the parts that are expected to give the better results in short time after the internship. You will find a summary of the main notations at the beginning of this report.

Notations

- $2n$: the dimension of the high dimensional space,
- (M^{2n}, ω^{2n}) or (M, ω) : the high dimensional symplectic manifold on which lie the trajectories of the PDE we are interested in (in fact $\mathbf{R}^{2n}$),
- $x = (q, p)$: a point of M (note that we use the order we see on theoretical papers and not the one we see on those dealing with numerical applications),
- k : the dimension of the low dimensional space we are looking for in the reduction context when we do not suppose that it is symplectic,
- $2k$: the dimension of the low dimensional space we are looking for in the reduction context when we suppose that it is symplectic,
- Σ^k, Σ^{2k} or Σ : the low dimensional submanifold of M on which lie the trajectories of the PDE we are interested in,
- $\hat{x}$: a point of the submanifold Σ^k, also denoted by $(\hat{p}, \hat{q})$ when Σ is supposed to be symplectic,
- $D^k : U \subset \mathbf{R}^k \to \Sigma^k$: a (global) parametrisation for Σ^k, called *decoder* in the context of reduction of dimension,
- $E : \Sigma^k \to U \subset \mathbf{R}^k$: a (global) chart of the submanifold, also called *encoder* in the context of reduction of dimension,
- N : the number of trajectories we consider to build the reduction D,
- m : the number of time intervals in each trajectories,
- Nm : the dimension of the sample,
- $H : \mathbf{R}^{2n} \to \mathbf{R}$: a Hamiltonian function,
- $\hat{H} : \mathbf{R}^{2k} \to \mathbf{R}$: its reduction,
- $d \in \mathbf{N}$: when the Hamiltonian in high dimension is parametrized, the dimension of the space where lives the parameter,
- $g \in \mathbf{R}^d$: the parameter itself (whenever it exists),

- $K \in \mathbf{N}$: dimension of the space where we look for an optimal reduced Hamiltonian function when performing optimal control approach,
- $\theta \in \mathbf{R}^K$: parameter of the reduced Hamiltonian when build with hyperreduction based on an optimal control approach,

Part I

Numerical part

1 Context

1.1 Hamiltonian systems

In what follows, we consider a symplectic manifold (M, ω) and we study Hamiltonian systems, i.e. systems of the form

$$\begin{cases} \dot{x} = X_H(x), \\ x(0) = x_0, \end{cases}$$

with $X_H \in \Gamma(M)$ the Hamiltonian vector field defined by

$$\omega(X_H, \cdot) = dH,$$

for $H : M \to \mathbf{R}$ a Hamiltonian function.

If we also consider, in addition to the symplectic structure induced by ω on M, a Riemannian structure induced by a metric g on this same space, the 2-form ω can be formulated in terms of g, like any bilinear form: for all $x \in M$, it exists A_{ω_x} such that for all $u, v \in T_xM$

$$\omega_x(u, v) = g(A_{\omega_x} u, v).$$

The matrix A must be skew-symmetric and non-degenerate. Note that g introduces another structure on M that is not necessary for our purpose. Nevertheless, when we work on $\mathbf{R}^{2n}$, using the Euclidian stucture simplifies computations.

We first consider $M = \mathbf{R}^{2n}$ with the standard symplectic form $\omega = d\mathbf{q} \wedge d\mathbf{p}$ and the Euclidean structure induced by the standard scalar product $\langle \cdot, \cdot \rangle$. In this case, for all $u, v \in \mathbf{R}^{2n}$,

$$\omega(u, v) = \langle \mathbf{J}_{2n}^T u, v \rangle,$$

where

$$\mathbf{J}_{2n} = \begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix}.$$

The previous system is then rewritten as

$$\dot{x} = \mathbf{J}_{2n} \nabla_x H(x). \tag{1.1}$$

If we write $x(t) = (p(t), q(t))$, this is equivalent to

$$\begin{cases} q_t = \dfrac{\partial H}{\partial p}(q, p), \\ p_t = -\dfrac{\partial H}{\partial q}(q, p). \end{cases}$$

1.2 Hamiltonian reduction

1.2.1 Goals

We wish to build a reduced model for the equation

$$\dot{x} = X_H(x),$$

with $x \in M$, which is also in a Hamiltonian form:

$$\dot{\hat{x}} = X_{\tilde{H}}(\hat{x}),$$

that is

$$\dot{\hat{x}} = \mathbf{J}_{2k} \nabla_{\hat{x}} \hat{H}(\hat{x}),$$

with $\hat{x}$ in a $2k$-dimensional manifold, for a certain Hamiltonian function $\hat{H}$ of Σ^k and $k \ll n$.

For that, we start by computing some solutions of the original problem in high dimension. From these data, we then look for a projection to a low dimensional space and a Hamiltonian function in this space. In short, building a reduced order model consists in extracting a low dimensional dynamic from a data set.

We in fact assume that it exists a low dimensional manifold Σ^k on which lies the solutions of the previous problem. The set of solutions can be given by the trajectories for a single Hamiltonian function starting at different points or trajectories for a parametrized Hamiltonian function starting at the same point or even a mix of this two cases. In all cases, the space of solutions is parametrized by a certain parameter g, which takes its value in a certain space G. In this section we omit the mention of this parameter and when, in the following sections, it will be necessary to mention it, we will consider that it is the Hamiltonian function which is parametrized. The case where it is the initial condition will only slighly differ in notations. Note also that if G is of dimension k, we expect that the space of solutions is also of dimension k. However, we still need to compute them in high dimension, which is the space where work our numerical solvers. All the point with the reduction is to explicit the dependance of the solutions in g in a way trajectories can easily been computed without mention of the high dimensional space.

Remark 1.2.1. The reason why we look for another Hamiltonian system in low dimension is that previous works in the field of reduction for Hamiltonian systems showed that the induced reduction usually gives better results in term of stability and accuracy (see []). Here is a completely informal discussion about why this could not be surprinsing. In fact, it seems reasonable since we then compute trajectories in low dimension using the same kind of rules than in high dimension. In particular, we know that the Hamiltonian is conserved along trajectories, and this property is physically important since the Hamiltonian usually represents the energy. If the Hamiltonian in low dimension is well chosen, its conservation along trajectories in low dimension may be a guaranty that the energy actually does not vary in the high dimensional space. By conserving the geometrical structure on the space of solutions, the tools available to describe them are still at our disposal and then the numerical solvers specially created to be stable while integrating the solutions in high dimension has good chances to work also in the low dimensional space. However, we have to note that we may not catch the actual dimension of the solution manifold, which can even be odd. We therefore look for a symplectic manifold Σ^k of very low dimension on which the original problem has a Hamiltonian formulation.

For the reduction, we need a proper low dimensional symplectic manifold (Σ^k, η). In the methods we will present here, we always first start by building Σ^k, before looking for a proper $\hat{H}$, and we always build it as a submanifold of (M, ω). In particular, the symplectic form η will be the restriction to Σ^k of ω. This implies that Σ^k is a symplectic submanifold and so that the inclusion $i : \Sigma^k \to M$ is symplectic.

From now on, we assume that Σ^k is symplectically parametrized by $\mathbf{R}^{2k}$. We call *decoder* the map $D : \mathbf{R}^{2k} \to M$ which associates to the coordinates of a point $\hat{x}$ in Σ^k the point $i(x)$ in M and *encoder* the map E which associates its coordinates to a point of Σ^k as sbmanifold embedded in M. In what follows, (M, ω) will always be $\mathbf{R}^{2n}$ endowed with the usual symplectic form. In practice, we choose the submanifold Σ^k by the mean of D, which is the object that we actually build.

1.2.2 The reduced model

The assumptions we have made on Σ^k implie that D is isosymplectic and so that $D^*\omega_{2n} = \omega_{2k}$, which is equivalent to

$$\omega_{2n}(d_{\hat{x}}D(u), d_{\hat{x}}D(u)) = \omega_{2k}(u, v)$$

for all $\hat{x} \in \mathbf{R}^{2k}$ and all $u, v \in \mathbf{R}^{2k}$.

Using the expression of ω_{2n} and ω_{2k} in terms of the scalar products on $\mathbf{R}^{2n}$ and $\mathbf{R}^{2k}$, we immediately find that a necessary and sufficient condition for $d_{\hat{x}}D$ to preserve the Hamiltonian structure is given by

$$^t d_{\hat{x}}D \mathbf{J}_{2n} d_{\hat{x}}D = \mathbf{J}_{2k} \quad \forall \hat{x} \in \mathbf{R}^{2k}. \tag{1.2}$$

Equation 1.1 can be rewritten as

$$\nabla D(\hat{x})\dot{\hat{x}} = \mathcal{J}\nabla H(D(\hat{x})).$$

From the definition of the gradient in $\mathbf{R}^{2n}$, we get $\nabla(H \circ D)(\hat{x}) = {}^t\nabla D(\hat{x})\nabla H(D(\hat{x}))$ for all $\hat{x} \in \mathbf{R}^{2k}$, where $\nabla D(\hat{x})$ represents the Jacobian matrix of D at $\hat{x} \in \mathbf{R}^{2k}$. Then, if we multiply the previous equation by ${}^t\mathbf{J}_{2k}{}^t\nabla D(\hat{x})\mathbf{J}_{2n}$, the condition on $d_{\hat{x}}D$ gives

$$\begin{aligned}
&{}^t\mathbf{J}_{2k}{}^t\nabla D(\hat{x})\mathbf{J}_{2n}\nabla D(\hat{x})\dot{\hat{x}} = {}^t\mathbf{J}_{2k}{}^t\nabla D(\hat{x})\mathbf{J}_{2n}\mathcal{J}_{2n}\nabla H(D(\hat{x})),\\
\Longleftrightarrow\ &{}^t\mathbf{J}_{2k}\mathbf{J}_{2k}\dot{\hat{x}} = {}^t\mathbf{J}_{2k}{}^t\nabla D(\hat{x})(-I_{2n})\nabla H(D(\hat{x})),\\
\Longleftrightarrow\ &\dot{\hat{x}} = \mathbf{J}_{2k}{}^t\nabla(H \circ D)(\hat{x}).
\end{aligned}$$

The original problem thus takes a Hamiltonian form in the low dimensional space :

$$\dot{\hat{x}} = X_{H\circ D}(\hat{x}).$$

When we have D, and the Σ^k, we have not finished the reduction. If we want to compute the solutions in the low dimensional space without coming back at each step to the high dimensional one, we need to find an expression, or at least an approximation, of $H \circ D$. This last step is called the *hyperreduction*.

1.2.3 Linear reduction

First, let us assume that the equation depends linearly on the parameters and on time. We then look for a linear reduction, in other words, for a matrix $A \in \mathcal{M}_{2n,2k}(\mathbf{R})$ which preserves the Hamiltonian structure. In this case, the symplecticity condition (1.2) becomes

$$^tA\mathbf{J}_{2n}A = \mathbf{J}_{2k}.$$

Symplecticity conditions for linear maps

A matrix $A \in \mathcal{M}_{2n,2k}(\mathbf{R})$ is said to be *symplectic* if it satisfies the above condition. If we decompose A into four submatrices A_1, A_2, A_3, A_4 in $\mathcal{M}_{n,k}(\mathbf{R})$ such that

$$A = \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix},$$

we have

$$\begin{aligned} {}^tA\mathbf{J}_{2n}A &= \begin{pmatrix} {}^tA_1 & {}^tA_3 \\ {}^tA_2 & {}^tA_4 \end{pmatrix} \begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix} \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix} = \begin{pmatrix} {}^tA_1 & {}^tA_3 \\ {}^tA_2 & {}^tA_4 \end{pmatrix} \begin{pmatrix} -A_3 & -A_4 \\ A_1 & A_2 \end{pmatrix} \\ &= \begin{pmatrix} {}^tA_3A_1 - {}^tA_1A_3 & {}^tA_3A_2 - {}^tA_1A_4 \\ {}^tA_4A_1 - {}^tA_2A_3 & {}^tA_4A_2 - {}^tA_2A_4 \end{pmatrix}. \end{aligned}$$

Then, A is symplectic if and only if ${}^tA_3A_1$ and ${}^tA_4A_2$ are symmetric and ${}^tA_4A_1 - {}^tA_2A_3 = I_k$.

Encoder and symplectic inverse

When a linear map links two spaces that do not have the same dimensions, it is hopeless to try to inverse it. However, when it has full rank, it admits a left or right inverse which takes a simple form in some cases. For example, it is the transposed for orthogonal matrices. It happens that in the symplectic case, we also have a simple expression for the left or right inverse.

We define the *symplectic inverse* of a matrix $A \in \mathcal{M}_{2n,2k}$ as

$$A^+ := {}^t\mathbf{J}_{2k}{}^tA\mathbf{J}_{2n}.$$

It is easy to check that if A is symplectic, then $A^+A = I_{2k}$ and ${}^t(A^+)$ is symplectic: if A is symplectic, then using the fact that $\mathbf{J}_{2k}$ is orthogonal,

$$A^+A = {}^t\mathbf{J}_{2k}{}^tA\mathbf{J}_{2n}A = {}^t\mathbf{J}_{2k}\mathbf{J}_{2k} = I_{2k}.$$

Similarly,

$$A^+\mathbf{J}_{2n}{}^tA^+ = ({}^t\mathbf{J}_{2k}{}^tA\mathbf{J}_{2n})\mathbf{J}_{2n}({}^t\mathbf{J}_{2n}A\mathbf{J}_{2k}) = {}^t\mathbf{J}_{2k}({}^tA\mathbf{J}_{2n}A)\mathbf{J}_{2k} = {}^t\mathbf{J}_{2k}\mathbf{J}_{2k}\mathbf{J}_{2k} = \mathbf{J}_{2k}.$$

Since it is a left inverse of A and since it is symplectic if A is symplectic, then it is a reasonnable choice for the encoder if A has been chosen to be the decoder.

Proper Symplectic Decomposition

To perform a linear reduction, the idea is to find a symplectic matrix $A \in \mathcal{M}_{2n,2k}(\mathbf{R})$ such that $A\hat{x}$ is as close as possible to x, where $\hat{x}$ is the solution of the reduced problem induced by A and x is the solution of the initial problem.

To do this, we first compute the solution of the initial problem for some values of the parameters and time and we evaluate the difference between these solutions and their images after encoding and decoding, i.e.

$$\|S - AA^+S\|_F. \tag{1.3}$$

We would like to minimize this loss. However, as an examination of the symplecticity conditions shows, the set of symplectic matrices is not bounded. Therefore, this optimization problem does not admit an explicit solution. Several methods have been proposed to find an optimal A under additional constraints. We present here the Proper Symplectic Decomposition, also known as cotangent lift or complex SVD, which is an adaptation of the Proper Orthogonal Decomposition in the symplectic case.

We restrict the space where we look for the minimum of (1.3). The idea is to look for an operator A of the form

$$\begin{pmatrix} \phi & 0 \\ 0 & \phi \end{pmatrix},$$

with $\phi \in \mathcal{O}(n)$. As shown in [17], this is in fact equivalent to choose the k columns of ϕ among the $\{q_1, ..., q_m, p_1, ..., p_m\}$ using classical POD.

1.3 Application to a piano vibrating string

In this work, we test the reduction on a set of equations modelling a piano string vibration, proposed in [4].

1.3.1 The model

We consider the following problem:

$$\begin{cases} \partial_{tt}^2 U(z,t) = \partial_z \Big[\nabla V(\partial_z U(z,t)) \Big] & \forall (z,t) \in \Omega \times \mathbf{R}_+ \\ U(z,0) = U_0(z) & \forall z \in \Omega, \\ \partial_t U(z,0) = U_1(z) & \forall z \in \Omega, \\ U(z,t) = 0 & \forall (z,t) \in \partial\Omega \times \mathbf{R}_+. \end{cases}$$

In what follows, $U(z,t) = (v(z,t), u(z,t))$ represents the longitudinal and transverse variations of the position of the point z in a piano string on the oscillation plane. The domain Ω is the interval $[0,1]$.

Let $q = (u,v)$ and $p = (\partial_t u, \partial_t v)$. The previous equation is rewritten

$$\begin{cases} \dfrac{\partial q}{\partial t} = p, \\ \dfrac{\partial p}{\partial t} = \partial_z \Big[\nabla V(\partial_z q) \Big]. \end{cases}$$

It has a Hamiltonian formulation with the energy function

$$H(p,q,t) = \int_\Omega \frac{1}{2}|p|^2 + V(\partial_z q) dz.$$

Indeed, on the one hand

$$H(p+p',q,t) = \int_\Omega \frac{1}{2}|p|^2 + V(\partial_z q) dz + \int_\Omega p \cdot p' dz + \int_\Omega \frac{1}{2}|p'|^2 dz = H(p,q,t) + \langle p, p' \rangle + o(|p'|),$$

from which

$$\nabla_p H(p,q,t) = p.$$

On the other hand,

$$\begin{aligned} H(p, q+q', t) &= \int_\Omega \frac{1}{2}|p|^2 + V(\partial_z q)dz + \int_\Omega \nabla V(\partial_z q)\cdot \partial_z q' dz + \int_\Omega o(|\partial_z q'|) \\ &= H(p,q,t) + \int_\Omega \nabla V(\partial_z q)\cdot \partial_z q' dz + o(|q'|). \end{aligned}$$

After an integration by parts, since by hypothesis q' is zero on $\partial\Omega$, we find

$$\int_\Omega \nabla V(\partial_z q)\cdot \partial_z q' dz = -\int_\Omega \partial_z \nabla V(\partial_z q)\cdot q' dz,$$

from which

$$\nabla_q H(p,q,t) = -\partial_z \nabla V(\partial_z q).$$

The Hamiltonian is thus separated. The term in p represents the kinetic energy and the one in q the potential one. We study different expressions for V, all given in [4].

To solve this problem in high dimension, we use Störmer-Verlet symplectic solver, given in [11].

1.3.2 Application of the PSD

One of the potential energy proposed in [4] induces a linear model. As we have seen in a previous work (see my M1 internship report), the PSD works well in this case. As we see on Figures 1.1 and 1.2, this is not the case for a choice of V which leads to a non-linear model. This work aims to find reduction methods that improves the results we get with the PSD. We therefore not consider the linear model and focus on the non-linear one, that we present now.

We consider $V(u,v) = \frac{1-\alpha}{2}u^2 + \frac{1}{2}v^2 + \frac{\alpha}{2}(u^2 v + \frac{1}{4}u^4)$, which yields the following system :

$$\begin{cases} \partial_{tt}^2 u = \partial_z\Big[(1-\alpha)\partial_z u + \alpha(\partial_z u \partial_z v + \frac{1}{2}(\partial_z u)^3)\Big], \\ \partial_{tt}^2 v = \partial_z\Big[\partial_z v + \frac{\alpha}{2}(\partial_z u)^2\Big]. \end{cases}$$

Numerically, we approach first and second derivatives with finite differences :

$$\partial_z u \approx \frac{u(z^{i+1}) - u(z^i)}{\Delta z}$$

and

$$\partial_z^2 u \approx \frac{u(z^{i+1}) - 2u(z^i) + u(z^{i-1})}{\Delta z^2}.$$

As we see on Figure 1.1, reduction using PSD gives inaccurate reduced models for $k = 5$. The "bumps" are too sharp on the solution computed with PSD and one can see some oscillations. Looking at H^1 errors and energy, the model built with PSD produces an unstable solution.

Figure 1.1: ...

Figure 1.2: ...

When we take $k = 10$, these problems are still visible. To obtain a satisfactoring solution with the reduced model, one should increase the reduced space dimension and take $k = 20$. This in fact means that the reduction failed because we did not succeed in capturing the low dimensional structure of the problem. We can deduce that the problem we want to solve here is too far from being linear to admit a linear reduction. If we still want to use PSD, we thus have to look further to improve the reduction.

2 Hyperreduction with an optimal control approach

In this chapter, we suppose that we have already performed a Proper Symplectic Decomposition or any other linear reduction method that gives a decoding map $D : \mathbf{R}^{2k} \to \mathbf{R}^{2n}$ from the low dimensional space, where we want to compute the solutions of the studied Hamiltonian system, to the high dimensional space, where we originally compute them. We also suppose that the same method gives us an encoding map $E : \mathbf{R}^{2n} \to \mathbf{R}^{2k}$ to compress solutions from the high dimensional space to the low dimensional one.

In the original PSD approach, we then compute solutions in low dimension using the Hamiltonian obtained by composition of the Hamiltonian H of the original, and high-dimensional, problem with the decoder. Of course, computing solutions this way is as costly as computing them in the classical high-dimensional way. This is why we usually add a hyperreduction step to the reduction, which consists in finding a Hamiltonian function in low dimension which interpolates $H \circ D$.

Here, we take another approach. Instead of interpolating $H \circ D$, we directly want to find the Hamiltonian in the low dimensional space that gives the more accurate trajectories, that is trajectories which, when decompressed in the high dimensional space, are the closest to the those that are computed in high dimension. This is achieved using an optimal control approach. In the field of learning Hamiltonian dynamics with neurol networks, this kind of methods have been used in [6, 18, 14] for instance. Here, we use a method of type Sparse Identification of Non-linear Dynamics (SINDy), proposed in [2]. It briefly consists in taking the target function in the space spanned by a set of given non-linearities. Coefficients of the target in this space are chosen such that the image of the source data, which here is the set of the decompressed low dimensional trajectories, corresponds as much as possible to the target, which here is the set of high-dimensional trajectories, while keeping a lot of coefficients exactly equal to zero. We detail the application of this method to our problem in the following section.

2.1 Problem setting

Let $F = \{f_i : \mathbf{R}^{2k} \to \mathbf{R}\}_i \in [\![1, K]\!]$ with $K \in \mathbf{N}$ a family of non-linear functions of class $\mathcal{C}^2$ and V_F the finite dimensional Hilbert subspace spanned by F. For $\theta \in \mathbf{R}^K$, denote by $\hat{H}_\theta$ the element of V_F such that $\hat{H}_\theta = \sum_{i=1}^K \theta_i f_i$. In the following, we are looking for the value of the parameter θ which minimizes the loss

$$\mathcal{L}(\theta) = \int_{g \in G} \int_{t \in [0,1]} \|D\hat{x}_{\theta,g}(t) - x_g(t)\|_{2k}^2 dg dt,$$

where $\hat{x}_{\theta,g} : [0,1] \to \mathbf{R}^{2k}$ denotes the solution of the problem

$$\begin{cases} x'(t) = X_{\hat{H}_\theta}(x(t)) & \forall t \in [0,1], \\ x(0) = Ex_0, \end{cases} \tag{2.1}$$

where $X_{\hat{H}_\theta}$ is the Hamiltonian flow of $\hat{H}_\theta$, thought of as the reduce version of H_g. Recall that G is the set of admissible parameter for the high dimensional Hamiltonian H_g and $x_g : [0,1] \to \mathbf{R}^{2n}$ is the trajectory of x_0 along the flow of H_g. We also denote by E and D the linear encoder and decoder built at the previous reduction step using the PSD.

A classical method to find numerically a value close to an optimal value, is a gradient descent. To implement it, however, one needs the gradient of the loss $\mathcal{L}$ with respect to the variable θ. The following computations aims to find it.

To simplify further computations, we introduce the loss function

$$\mathcal{L}_g(\theta) := \int_{t\in[0,1]} \|D\hat{x}_{\theta,g}(t) - x_g(t)\|_{2k}^2 dt.$$

Remark 2.1.1. Note that if we have the gradient of $\mathcal{L}_g$ for all $g \in G$, then we simply obtain the gradient of $\mathcal{L}$ by integration over G. Indeed, for all $\theta \in \mathbf{R}^K$, we have $\mathcal{L}(\theta) = \int_{g\in G} \mathcal{L}_g(\theta)dg$ so for all $h \in \mathbf{R}^K$ small enough, it comes

$$\begin{aligned}\mathcal{L}(\theta + h) &= \int_{g\in G} \mathcal{L}_g(\theta) + \nabla\mathcal{L}_g(\theta)\cdot h + o(\|h\|_K)dg = \mathcal{L}(\theta) + \int_{g\in G} \sum_{i=1}^K \partial_i\mathcal{L}_g(\theta)h_i dg + o(\|h\|_K) \\ &= \mathcal{L}(\theta) + \sum_{i=1}^K h_i \int_{g\in G}^K \partial_i\mathcal{L}_g(\theta)dg + o(\|h\|_K) = \mathcal{L}(\theta) + \left\langle \int_{g\in G} \nabla\mathcal{L}_g(\theta)dg, h \right\rangle_K + o(\|h\|_K).\end{aligned}$$

We deduce that $\nabla\mathcal{L}(\theta) = \int_{g\in G} \nabla\mathcal{L}_g(\theta)dg$. In what follows, we will therefore restrict our study to the case where $G = \{g_0\}$ and skip mentions of g in our notations.

2.2 In-out application

Denote by $\mathcal{F} : \mathbf{R}^K \to H^1([0,1])^{2k}$ the application which sends a value of the parameter θ to the trajectory $\hat{x}_\theta$ generated by $\hat{H}_\theta$ in low dimension. In the following, it will be called the *in-out application*. Let also $f : H^1([0,1])^{2k} \to \mathbf{R}$ be such that $f(\hat{x}) = \|D\hat{x} - x\|_{L^2([0,1])^{2k}}$. We have $\mathcal{L} = f \circ \mathcal{F}$ so, provided that both f and $\mathcal{F}$ are differentiable, the chain rule gives

$$d_\theta\mathcal{L} = d_{\mathcal{F}(\theta)}f \circ d_\theta\mathcal{F}.$$

If the variationnal method gives immediately the differential of f, it does not work with $\mathcal{F}$, which is implicitely defined. The strategy that we develop here is a classical method of optimal control. Briefly, it consists in finding what we call an *adjoint state*, build to satisfy a well-chosen ordinary differential equation. Inserted into the expression of $d_\theta\mathcal{L}$, it allows to neutralize the problematic terms.

2.2.1 Regularity of $\mathcal{F}$

In the first step, we find a differential equation followed by $d_\theta\mathcal{F}$. Before that, we have to prove that $\mathcal{F}$ is at least differentiable. We use here the implicit function theorem to achieve this. By translating $\mathcal{F}$ by the constant function x_0, we can reduce the problem where $x_0 = 0$.

Consider the application

$$G : \begin{cases} \mathbf{R}^K \times V \to L^2([0,1])^{2k} \\ \quad (\theta, x) \mapsto x' - X_{\hat{H}_\theta}(x) \end{cases}$$

where $V := \{f \in H^1([0,1])^{2k} \quad | \quad f(0) = 0\}$. It is a closed subspace of H^1 so it is a Hilbert space too for the H^1 scalar product.

It is of class $\mathcal{C}^1$. To prove it, we first need the following result concerning $\theta \mapsto X_{\hat{H}_\theta}$.

Lemma 2.2.1. *The application $\theta \to X_{\hat{H}_\theta}$ is linear and continuous. More precisely, we have $X_{\hat{H}_h} = {}^t\mathbf{X}h$ with*

$$\mathbf{X} := [X_{f_i}]_{i=1,\dots,K} \in \mathcal{C}^1(\mathbf{R}^{2k}, \mathbf{R}^{2k})^K.$$

Proof. Let us first see that $\theta \to X_{\hat{H}_\theta}$ is linear. Clearly, this is the case of $\theta \to \hat{H}_\theta$. Thanks to the properties of the symplectic form ω, this is also the case of $\hat{H} \to X_{\hat{H}}$: the bilinearity makes that

$$\omega(X_{\hat{H}_1+\lambda\hat{H}_2}, \cdot) = d_x(\hat{H}_1 + \lambda\hat{H}_2) = d_x\hat{H}_1 + \lambda d_x\hat{H}_2 = \omega(X_{\hat{H}_1}, \cdot) + \lambda\omega(X_{\hat{H}_2}, \cdot) = \omega(X_{\hat{H}_1} + \lambda X_{\hat{H}_2}, \cdot)$$

and the non-degeneracy then insures that $X_{\hat{H}_1+\lambda\hat{H}_2} = X_{\hat{H}_1} + \lambda X_{\hat{H}_2}$.

Then, since $X_{\hat{H}_\theta} = \sum_{i=1}^K \theta_i X_{f_i}$, if we introduce $\mathbf{X}$ as in the statement of the lemma, we have $X_{\hat{H}_\theta} = {}^t\mathbf{X}\theta$.

Finally,

$$\|X_{\hat{H}_h}\|_{L^2} = \int_{x\in\mathbf{R}^{2k}} \|{}^t\mathbf{X}(x) \cdot h\|_{2k} dt \leqslant \|h\|_K \int_{x\in\mathbf{R}^{2k}} |||\mathbf{X}(x)||| dt.$$

As $\mathbf{X}(x)$ lies in $\mathcal{M}_{2k,K}(\mathbf{R})$, which is a finite dimensional space, the operator norm $|||\cdot|||$ and the Froebenius norm are equivalent so

$$\|X_{\hat{H}_h}\|_{L^2} \leqslant cst\ \|h\|_K \int_{x\in\mathbf{R}^{2k}} \|\mathbf{X}(x)\|_F dt = cst\ \|h\|_K \sum_{i=1}^{2k}\sum_{j=1}^{K} \|\mathbf{X}_{ij}\|_{L^2(\mathbf{R}^{2k},\mathbf{R})}.$$

All the coefficients of $\mathbf{X}$ are continuous functions so the double sum takes a finite value. $\square$

Let us go back to G. Let θ be a point in $\mathbf{R}^K$ and x a function in $V \subset H^1([0,1])^{2k}$. Let h and y be two small elements of the same spaces. For all $t \in [0,1]$, we have

$$\begin{aligned} G(\theta + h, x + y)(t) - G(\theta, x)(t) = &\left(y'(t) - d_{x(t)}X_{\hat{H}_\theta}(y(t)) - X_{\hat{H}_h}(x(t))\right) \\ &- \left(\epsilon_1(y)(t) + d_{x(t)}X_{\hat{H}_h}(y)(t) + \epsilon_2(y)(t)\right), \end{aligned}$$

where $\epsilon_1(y)$ and $\epsilon_2(h)$ are respectively in $o(\|y\|_{H^1})$ and $o(\|h\|_K)$.

From the regularity assumption on the elements of V_F and from Lemma 2.2.1, we know that the terms in the first parentheses are linear and continuous in (h, y). In the second one, ϵ_1 and ϵ_2 are the remaining terms in the first oder Taylor's expansion of $X_{\hat{H}_\theta}$ and $X_{\hat{H}_h}$ so they are in $o(\|y\|_{H^1})$ and *a fortiori* in $o(\|(h,y)\|_{\mathbf{R}^K\times H^1})$. The last term, $d_{x(\cdot)}X_{\hat{H}_h}(y)$, is bilinear and continuous in h and y so is also in $o(\|(h,y)\|_{\mathbf{R}^K\times H^1})$. This proves that G is differentiable ad that it's differential is given by

$$d_{\theta,x}G(h,y) : t \mapsto y'(t) - d_{x(t)}X_{\hat{H}_\theta}(y(t)) - X_{\hat{H}_h}(x(t)).$$

We now have to see that $dG : \mathbf{R}^K \times V \to \mathcal{L}_c\left(V, L^2([0,1])^{2k}\right)$ is continuous. The first term does not depend on x nor θ so it is continuous in this variables. The last one is continuous too for the regularity we assumed on the functions in V_F.

Let $\bar{y}$ and $\bar{h}$ be two small elements of $H^1([0,1])$ and $\mathbf{R}$. We have

$$|||d_{x+\bar{y}}X_{\hat{H}_{\theta+\bar{h}}} - d_x X_{\hat{H}_\theta}||| \leqslant |||d_{x+\bar{y}}X_{\hat{H}_\theta} - d_x X_{\hat{H}_\theta}||| + |||d_{x+\bar{y}}X_{\hat{H}_{\bar{h}}}|||,$$

where $|||\cdot|||$ denotes the operator norm for linear continuous functions between $H^1([0,1])^{2k}$ and $L^2([0,1])^{2k}$.

The second term in the last expression tends to zero as $(\bar{h}, \bar{y}) \to 0$ since

$$|||d_{x+\bar{y}}X_{\hat{H}_{\bar{h}}}||| \leqslant \sum_{i=1}^{K} \bar{h}_i \; |||d_{x+\bar{y}}X_{f_i}|||$$

with $|||d_{x+\bar{y}}X_{f_i}||| \to |||d_x X_{f_i}|||$ from Lemma 2.2.1.

For the first term, we have that $t \mapsto d_{x(t)}X_{\hat{H}_\theta}$ is continuous on the compact $[0,1]$ as composition of such functions so is uniformly continuous. Then,

$$\forall \epsilon > 0,\ \exists \delta_\epsilon \quad | \quad \|\bar{y}(t)\|_{2k} < \delta_\epsilon \implies \sup_{z \in \mathbf{R}^{2k}} \left(\frac{\|d_{x(t)+\bar{y}(t)}X_{\hat{H}_\theta}(z) - d_{x(t)}X_{\hat{H}_\theta}(z)\|_{2k}}{\|z\|_{2k}} \right) \leqslant \epsilon.$$

In particular,

$$\|d_{x(t)+\bar{y}(t)}X_{\hat{H}_\theta}(y(t)) - d_{x(t)}X_{\hat{H}_\theta}(y(t))\|_{2k} \leqslant \epsilon \|y(t)\|_{2k}$$

for all $y \in V$. Passing to the L^2 norm, we get

$$\|d_{x+\bar{y}}X_{\hat{H}_\theta}(y) - d_x X_{\hat{H}_\theta}(y)\|_{L^2} \leqslant \epsilon \|y\|_{L^2} \leqslant \epsilon \|y\|_{H^1}$$

for all $y \in V$, which gives

$$\sup_{y \in H^1([0,1])^{2k}} \left(\frac{\|d_{x+\bar{y}}X_{\hat{H}_\theta}(y) - d_x X_{\hat{H}_\theta}(y)\|_{L^2}}{\|y\|_{H^1}} \right) \leqslant \epsilon$$

provided that $\|\bar{y}\|_\infty \leqslant \delta_\epsilon$. As $H^1([0,1])^{2k}$ is continuously injected in $\mathcal{C}^0([0,1])^{2k}$, this condition is satisfied when $\|\bar{y}\|_{H^1}$ is small enough. Since ϵ can be chosen arbitrarily small, this proves that dG is continuous and so that G is of class $\mathcal{C}^1$.

Now, the differential of G in the direction x at $\hat{x}_\theta$, which is given by

$$y \mapsto y' - d_x X_{\hat{H}_\theta}(y)$$

is bijective. In fact, for all $g \in L^2([0,1])^{2k}$, the system

$$\begin{cases} y' = d_x X_{\hat{H}_\theta}(y) + g, \\ y(0) = 0 \end{cases}$$

has an unique solution by Cauchy-Lipschiz's theorem. Then, applying the implicit function theorem, it exists an open neighbourhood $\Omega = \Omega_1 \times \Omega_2$ of any $(\theta, \hat{x}_\theta)$ in $\mathbf{R}^K \times V$ and a $\mathcal{C}^1$ function $\phi : \mathbf{R}^K \to V$ such that

$$((h,y) \in \Omega \wedge G(h,y) = 0) \quad \iff \quad (\theta \in \Omega_1 \wedge y = \phi(h)).$$

But if $G(h,y) = 0$, it means that $\mathcal{F}(h) = y$ so $\mathcal{F} = \phi$ on Ω_1. It means that $\mathcal{F}$ is $\mathcal{C}^1$ around θ and since θ has been arbitrarily chosen, it also means that $\mathcal{F}$ is $\mathcal{C}^1$ on $\mathbf{R}^K$.

2.2.2 Caracterization of $d_\theta \mathcal{F}$

Now that we have seen that $d_\theta \mathcal{F}$ is well defined and continuous, we characterize it with a differential problem in $[0,1]$.

Proposition 2.2.1. *The differential of $\mathcal{F}$ at θ in the direction h is solution of the following Cauchy's system*

$$\begin{cases} \dot{z}(t) = d_{\mathcal{F}(\theta)(t)} X_{\hat{H}_\theta}(z(t)) + X_{\hat{H}_h}(\mathcal{F}(\theta)(t)) \quad \forall t \in [0,1], \\ z(0) = 0. \end{cases} \tag{2.2}$$

Proof. Consider the equation

$$\mathcal{F}(\theta + h) - \mathcal{F}(\theta) = d_\theta \mathcal{F}(h) + \epsilon_1(h)$$

for some infinitesimal h in $\mathbf{R}^K$ and derive it with respect to time. This leads to

$$X_{\hat{H}_{(\theta+h)}}(\mathcal{F}(\theta+h)(t)) - X_{\hat{H}_\theta}(\mathcal{F}(\theta)(t)) = \frac{d}{dt} d_\theta \mathcal{F}(h)(t) + \frac{d}{dt}\epsilon_1(h)(t)$$

$$\iff X_{\hat{H}_\theta}(\mathcal{F}(\theta+h)(t)) + X_{\hat{H}_h}(\mathcal{F}(\theta+h)(t)) - X_{\hat{H}_\theta}(\mathcal{F}(\theta)(t)) = \frac{d}{dt} d_\theta \mathcal{F}(h)(t) + \frac{d}{dt}\epsilon_1(h)(t) \tag{2.3}$$

for all t in $[0,1]$.

Now, developing $X_{\hat{H}_\theta}$, $X_{\hat{H}_h}$ and $\mathcal{F}$ at first order, the left hand side becomes

$$\begin{aligned} & X_{\hat{H}_\theta}\Big(\mathcal{F}(\theta)(t)\Big) + d_{\mathcal{F}(\theta)(t)} X_{\hat{H}_\theta}\Big(d_\theta \mathcal{F}(h)(t)\Big) + d_{\mathcal{F}(\theta)(t)} X_{\hat{H}_\theta}\Big(\epsilon_1(h)(t)\Big) + \epsilon_2\Big(d_\theta \mathcal{F}(h)(t) + \epsilon_1(h)(t)\Big) \\ & + X_{\hat{H}_h}\Big(\mathcal{F}(\theta)(t)\Big) + d_{\mathcal{F}(\theta)(t)} X_{\hat{H}_h}\Big(d_\theta \mathcal{F}(h)(t) + \epsilon_1(h)(t)\Big) + \epsilon_3\Big(d_\theta \mathcal{F}(h)(t) + \epsilon_1(h)(t)\Big) \\ & - X_{\hat{H}_\theta}\Big(\mathcal{F}(\theta)(t)\Big) \end{aligned}$$

for all $t \in [0,1]$.

Rearranging the terms, we get

$$\begin{aligned} & \left[d_{\mathcal{F}(\theta)(t)} X_{\hat{H}_\theta}\Big(d_\theta \mathcal{F}(h)(t)\Big) + X_{\hat{H}_h}\Big(\mathcal{F}(\theta)(t)\Big) \right] + \left[d_{\mathcal{F}(\theta)(t)} X_{\hat{H}_\theta}\Big(\epsilon_1(h)(t)\Big) \right. \\ & \left. + \epsilon_2\Big(d_\theta \mathcal{F}(h)(t) + \epsilon_1(h)(t)\Big) + d_{\mathcal{F}(\theta)(t)} X_{\hat{H}_h}\Big(d_\theta \mathcal{F}(h)(t) + \epsilon_1(h)(t)\Big) + \epsilon_3\Big(d_\theta \mathcal{F}(h)(t) + \epsilon_1(h)(t)\Big) \right]. \end{aligned}$$

The function

$$m : h \in \mathbf{R}^K \mapsto \Big(t \mapsto d_{\mathcal{F}(\theta)(t)} X_{\hat{H}_\theta}\Big(d_\theta \mathcal{F}(h)(t)\Big) + X_{\hat{H}_h}\Big(\mathcal{F}(\theta)(t)\Big)\Big) \in L^2([0,1])^{2k}$$

is linear and continuous as sum and compositions of such functions :

$$\begin{aligned} \|m(h)\|_{L^2}^2 &\leqslant \int_0^1 \|d_{x_\theta(t)} X_{\hat{H}_\theta}(d_\theta \mathcal{F}(h)(t))\|_{2k}^2 dt + \int_0^1 \|{}^t\mathbf{X}(\mathcal{F}(\theta)(t))h\|_{2k}^2 dt \\ &\leqslant \int_0^1 |\!|\!| d_{x_\theta(t)} X_{\hat{H}_\theta} |\!|\!|^2 \ |\!|\!| d_\theta \mathcal{F}(t) |\!|\!|^2 \ \|h\|_K^2 dt + \int_0^1 |\!|\!| \mathbf{X}(\mathcal{F}(\theta)(t)) |\!|\!|^2 \|h\|_K^2 dt \\ &= cst \ \|h\|_K^2. \end{aligned}$$

Lemma 2.2.2. *The other terms, that is*

$$t \mapsto d_{\mathcal{F}(\theta)(t)} X_{\hat{H}_\theta}\Big(\epsilon_1(h)(t)\Big) + \epsilon_{2+3}\Big(d_\theta \mathcal{F}(h)(t) + \epsilon_1(h)(t)\Big) + d_{\mathcal{F}(\theta)(t)} X_{\hat{H}_h}\Big(d_\theta \mathcal{F}(h)(t) + \epsilon_1(h)(t)\Big),$$

with $\epsilon_{2+3} = \epsilon_2 + \epsilon_3$, are in $o(\|h\|_K)$ for the L^2 norm.

Proof. By continuity of $d_{\mathcal{F}(\theta)(t)} X_{\hat{H}_\theta}$ for all $t \in [0,1]$

$$\Big\| d_{\mathcal{F}(\theta)(\cdot)} X_{\hat{H}_\theta}\Big(\epsilon_1(h)(\cdot)\Big)\Big\|_{L^2}^2 = \int_0^1 \| d_{\mathcal{F}(\theta)(t)} X_{\hat{H}_\theta}\Big(\epsilon_1(h)(t)\Big)\|_{2k}^2 dt \leqslant \int_0^1 |\!|\!| d_{\mathcal{F}(\theta)(t)} X_{\hat{H}_\theta} |\!|\!| \ \|\epsilon_1(h)(t)\|_{2k}^2 dt.$$

As $\hat{H}_\theta$ is supposed to be of class $\mathcal{C}^2$, $x \mapsto d_x X_{\hat{H}_\theta}$ is continuous. Since $t \mapsto \mathcal{F}(\theta)(t)$ is continuous, we finally have that $t \mapsto |\!|\!| d_{\mathcal{F}(\theta)(t)} X_{\hat{H}_\theta} |\!|\!|$ is continuous on $[0,1]$ and therefore bounded. This gives

$$\Big\| d_{\mathcal{F}(\theta)(\cdot)} X_{\hat{H}_\theta}\Big(\epsilon(h)(\cdot)\Big)\Big\|_{L^2}^2 \leqslant cste \int_0^1 \|\epsilon_1(h)(t)\|_{2k}^2 dt = cste\ \|\epsilon_1(h)\|_{L^2}^2.$$

As ϵ_1 is in $o(\|h\|_K)$ for the L^2 norm, so is $t \mapsto d_{\mathcal{F}(\theta)(t)} X_{\hat{H}_\theta}\Big(\epsilon_1(h)(t)\Big)$.

In the second non-linear term, the fact that $\epsilon_{2+3}(x)$ is in $o(\|x\|_{2k})$ when x tends to zero can be written as

$$\forall \epsilon > 0,\ \exists \delta_\epsilon > 0 \mid \forall x \in \mathbf{R}^{2k},\ \|x\| < \delta_\epsilon \implies \|\epsilon_{2+3}(x)\|_{2k} < \epsilon \|x\|_{2k}. \tag{2.4}$$

When h tends to zero in $\mathbf{R}^K$, $d_\theta \mathcal{F} \cdot h + \epsilon_1(h)$ tends to zero in $H^1([0,1])$, which is continuously injected in $\mathcal{C}^0([0,1])$. Then,

$$\forall \delta > 0,\ \exists \eta_\delta > 0 \mid \forall h \in \mathbf{R}^K,\ \|h\| < \eta_\delta \implies \big(\forall t \in [0,1],\ \|d_\theta \mathcal{F}(t) \cdot h + \epsilon_1(h)(t)\|_{2k} < \delta\big). \tag{2.5}$$

Now take $\epsilon > 0$, consider the δ_ϵ given in (2.4) and the $\eta_\epsilon := \eta_\delta$ given in (2.5) for $\delta = \delta_\epsilon$. If $\|h\|_K < \eta_\epsilon$, then by (2.5),

$$\|d_\theta \mathcal{F}(t) \cdot h + \epsilon_1(h)(t)\|_{2k} < \delta_\epsilon$$

so by (2.4),

$$\|\epsilon_{2+3}\big(d_\theta \mathcal{F}(t) \cdot h + \epsilon_1(h)(t)\big)\|_{2k} < \epsilon \|d_\theta \mathcal{F}(t) \cdot h + \epsilon_1(h)(t)\|_{2k}$$

for all $t \in [0,1]$. If we take the L^2 norm of the previous inequality, we get

$$\int_0^1 \|\epsilon_{2+3}\big(d_\theta \mathcal{F}(t) \cdot h + \epsilon_1(h)(t)\big)\|_{2k}^2 dt < \epsilon^2 \int_0^1 \|d_\theta \mathcal{F}(t) \cdot h + \epsilon_1(h)(t)\|_{2k}^2 dt.$$

Since $d_\theta \mathcal{F} \cdot h + \epsilon_1(h)$ is in $\mathcal{O}(\|h\|_K)$, we have

$$\|\epsilon_{2+3}\big(d_\theta \mathcal{F}(t) \cdot h + \epsilon_1(h)(t)\big)\|_{L^2} < \ cste\ \epsilon \|h\|_K.$$

As ϵ is arbitrarily chosen, we have proven that the second non-linear term is also in $o(\|h\|_K)$.

For the remaining non-linear term, we have

$$\begin{aligned}
\|d_{\mathcal{F}(\theta)(\cdot)} X_{\hat{H}_h}\Big(d_\theta \mathcal{F}(h)(\cdot) + \epsilon_1(h)(\cdot)\Big)\|_{L^2}^2 &= \int_0^1 \|d_{\mathcal{F}(\theta)(t)} X_{\hat{H}_h}\Big(d_\theta \mathcal{F}(h)(t) + \epsilon_1(h)(t)\Big)\|_{2k}^2 dt \\
&\leqslant \int_0^1 |\!|\!| d_{\mathcal{F}(\theta)(t)} X_{\hat{H}_h} |\!|\!|^2 \ \|d_\theta \mathcal{F}(h)(t) + \epsilon_1(h)(t)\|_{2k}^2 dt \\
&\leqslant \int_0^1 cste\ |\!|\!| \sum_{i=1}^K h_i\ d_{\mathcal{F}(\theta)(t)} X_{f_i} |\!|\!|^2 \ \|h\|_K^2 dt \\
&\leqslant \|h\|_K^4 \int_0^1 cste\ \max_{1 \leqslant i \leqslant K} |\!|\!| d_{\mathcal{F}(\theta)(t)} X_{f_i} |\!|\!|^2 dt
\end{aligned}$$

so this last term is also in $o(\|h\|_K)$. $\square$

To finish the proof of the proposition, see that the first term in (2.3) is linear and continuous with respect to h while the second term is in $o(\|h\|)$ in L^2. Identifying the linear parts in both sides of (2.3), we finally find that $d_\theta\mathcal{F}(\theta)\cdot h$ is solution of the affine ordinary differential equation

$$\dot{z}(t) = d_{\mathcal{F}(\theta)(t)}X_{\hat{H}_\theta}(z(t)) + X_{\hat{H}_h}(\mathcal{F}(\theta)(t)) \quad \forall t \in [0,1].$$

Moreover, as $\mathcal{F}(\theta)(0) = Ex_0$ for all θ, $d_\theta\mathcal{F}(\theta)\cdot h$ also satisfies

$$z(0) = 0_{\mathbf{R}^{2k}}.$$

□

2.3 Gradient of $\mathcal{L}$

We now express the differential of $\mathcal{L}$ in terms of $d_\theta\mathcal{F}$. Fix $g \in G$. Let θ be any point of $\mathbf{R}^K$ and h be an infinitesimal displacement in this same space. Consider $\mathcal{F}$ first order Taylor's expansion

$$\mathcal{F}(\theta + h) = \mathcal{F}(\theta) + d_\theta\mathcal{F}(h) + \epsilon_1(h),$$

where $\epsilon_1 : \mathbf{R}^K \to H^1([0,1])^{2k}$ satisfies $\|\epsilon_1(h)\|_{H^1} = o(\|h\|_K)$.

The developpement of $\mathcal{L}$ gives

$$\begin{aligned}\mathcal{L}(\theta + h) &= \|D\mathcal{F}(\theta + h) - x\|_{L^2}^2 \\ &= \|D\mathcal{F}(\theta) - x\|_{L^2} + 2\langle Dd_\theta\mathcal{F}(h), D\mathcal{F}(\theta) - x\rangle_{L^2} \\ &\quad + 2\langle D\epsilon_1(h), D\mathcal{F}(\theta) - x\rangle_{L^2} + \|Dd_\theta\mathcal{F}(h) + D\epsilon_1(h)\|_{L^2}^2.\end{aligned}$$

Since D is orthogonal, we have for each pair x, y in $L^2([0,1])^{2k}$ that

$$\langle Dx, Dy\rangle_{L^2} = \int_0^1 Dx \cdot Dy = \int_0^1 x \cdot {}^tDDy = \int_0^1 x \cdot y = \langle x, y\rangle_{L^2}.$$

Thus,

$$\mathcal{L}(\theta + h) = \mathcal{L}(\theta) + 2\langle d_\theta\mathcal{F}(h), \mathcal{F}(\theta) - {}^tDx\rangle_{L^2} + 2\langle \epsilon_1(h), \mathcal{F}(\theta) - {}^tDx\rangle_{L^2} + \|d_\theta\mathcal{F}(h) + \epsilon_1(h)\|_{L^2}^2.$$

The second term of the previous sum is linear and continuous as composition of such applications. By Cauchy's inequality in $L^2([0,1])^{2k}$, the third term is bounded by

$$\|\epsilon_1(h)\|_{L^2}\|\mathcal{F}(\theta) - {}^tDx\|_{L^2}$$

and since $\|\epsilon_1(h)\|_{H^1}$, and a fortiori $\|\epsilon_1(h)\|_{L^2}$, are supposed to be in $o(\|h\|_K)$, so is the third term. Applying the triangular inequality to the last term and invoking the assymptotic behaviours of the linear continuous $d_\theta\mathcal{F}$ and the 1-order neglectible ϵ_1, we prove that the last term is also in $o(\|h\|_K)$. Therefore,

$$d_\theta\mathcal{L} = 2\langle d_\theta\mathcal{F}(h), \mathcal{F}(\theta) - {}^tDx\rangle_{L^2} \tag{2.6}$$

Remark 2.3.1. Unless what usally occurs in optimal control problem, where the parameter θ depends on time, it is here a fixed point of $\mathbf{R}^K$. This allows to rewrite quite immediately the differential of $d_\theta \mathcal{L}$ in the finite dimensional space :

$$
\begin{aligned}
\langle d_\theta \mathcal{F}, \mathcal{F} - Dx \rangle_{L^2} &= \int_0^1 \sum_{i=1}^{K} \big(d_\theta \mathcal{F}^i(t) \cdot h\big) \times \big(\mathcal{F}^i(\theta)(t) - D^i x(t)\big)\, dt \\
&= \int_0^1 \sum_{i=1}^{2k} \left(\sum_{j=1}^{K} \partial_j \mathcal{F}^i(\theta)(t) \times h_j \right) \times \big(\mathcal{F}^i(\theta)(t) - D^i x(t)\big)\, dt \\
&= \sum_{j=1}^{K} h_j \times \left(\int_0^1 \sum_{i=1}^{2k} \partial_j \mathcal{F}^i(\theta)(t) \times \big(\mathcal{F}^i(\theta)(t) - D^i x(t)\big)\, dt \right) \\
&= h \cdot \int_0^1 \big(\mathcal{F}^i(\theta)(t) - D^i x(t)\big) \cdot \nabla \mathcal{F}(\theta)(t) dt.
\end{aligned}
$$

Then, we also have $\nabla \mathcal{L}(\theta) = \int_0^1 \big(\mathcal{F}^i(\theta)(t) - D^i x(t)\big) \cdot \nabla \mathcal{F}(\theta)(t) dt$. Now that we have caracterized $\nabla \mathcal{F}(\theta)$, we can compute $\nabla \mathcal{L}(\theta)$. However, the computation of $\nabla \mathcal{F}(\theta)$ involves K differential systems in $\mathbf{R}^{2k}$, with K possibly very large.

In fact, it is possible to write $\nabla \mathcal{L}(\theta)$ in another way which only involves one differential system in $\mathbf{R}^{2k}$. To show that, we use a classical method in optimal control, which makes use of a well-chosen adjoint function $a : [0,1] \to \mathbf{R}^{2k}$.

More precisely, we set a as the unique solution of the adjoint Cauchy's problem

$$
\begin{cases}
\dot{a}(t) = d_{\mathcal{F}(\theta)(t)} X_{\hat{H}_\theta}(a(t)) + \Big(\mathcal{F}(\theta)(t) - {}^t D x(t)\Big), \forall t \in [0,1], \\
a(T) = 0.
\end{cases}
\tag{2.7}
$$

Theorem 2.3.1. *For all θ in $\mathbf{R}^{2k}$,*

$$
\nabla \mathcal{L}(\theta) = -2 \int_0^1 \mathbf{X}\Big(\mathcal{F}(\theta)(t)\Big) a(t) dt,
\tag{2.8}
$$

with a and $\mathbf{X}$ as previously defined.

Proof. We have

$$
\dot{a}(t) - d_{\mathcal{F}(\theta)(t)} X_{\hat{H}_\theta}(a(t)) = \Big(\mathcal{F}(\theta)(t) - {}^t D x(t)\Big)
$$

for all t in $[0,1]$. Inserting this equality in (2.6), we get

$$
d_\theta \mathcal{L}(h) = 2 \int_0^1 \langle \dot{a}(t) - d_{\mathcal{F}(\theta)(t)} X_{\hat{H}_\theta} \cdot a(t), d_\theta \mathcal{F}(h)(t) \rangle_{\mathbf{R}^{2k}} dt.
$$

After an integration by part of the first term, we obtain

$$
\begin{aligned}
d_\theta \mathcal{L}(h) = 2 \int_0^1 & -\langle a(t), d_\theta \dot{\mathcal{F}}(h)(t) \rangle_{\mathbf{R}^{2k}} - \langle d_{\mathcal{F}(\theta)(t)} X_{\hat{H}_\theta} \cdot a(t), d_\theta \mathcal{F}(h)(t) \rangle_{\mathbf{R}^{2k}} dt \\
& + a(T) d_\theta \mathcal{F}(h)(T) - a(0) d_\theta \mathcal{F}(h)(0).
\end{aligned}
$$

Thanks to the initial and final condition on a and $d_\theta \mathcal{F}(h)$, the last two terms vanish. Using the equation satisfied by $d_\theta \mathcal{F}(h)$, we have

$$
\begin{aligned}
d_\theta \mathcal{L}(h) = 2 \int_0^1 & -\langle a(t), d_{\mathcal{F}(\theta)(t)} X_{\hat{H}_\theta} \cdot d_\theta \mathcal{F}(h)(t) \rangle_{\mathbf{R}^{2k}} - \langle a(t), X_{\hat{H}_h}(\mathcal{F}(\theta)(t)) \rangle_{\mathbf{R}^{2k}} \\
& - \langle d_{\mathcal{F}(\theta)(t)} X_{\hat{H}_\theta} \cdot a(t), d_\theta \mathcal{F}(h)(t) \rangle_{\mathbf{R}^{2k}} dt.
\end{aligned}
$$

Now, it suffices to notice that $d_{\mathcal{F}(\theta)(t)}X_{\hat{H}_\theta}$ is skew-symmetric (it is $\mathbf{J}\mathrm{Hess}_{\hat{H}_\theta}(\mathcal{F}(\theta)(t))$) to finally get

$$\begin{aligned}
d_\theta\mathcal{L}(h) &= -2\int_0^1 \langle a(t), X_{\hat{H}_h}\big(\mathcal{F}(\theta)(t)\big)\rangle_{2k} dt \\
&= -2\int_0^1 \langle a(t), {}^t\mathbf{X}\Big(\mathcal{F}(\theta)(t)\Big)h\rangle_{2k} dt \\
&= -2\int_0^1 \langle \mathbf{X}\Big(\mathcal{F}(\theta)(t)\Big)a(t), h\rangle_K dt \\
&= \langle -2\int_0^1 \mathbf{X}\Big(\mathcal{F}(\theta)(t)\Big)a(t)dt, h\rangle_K.
\end{aligned}$$

We conclude that

$$\nabla_\theta\mathcal{L} = -2\int_0^1 \mathbf{X}\Big(\mathcal{F}(\theta)(t)\Big)a(t)dt.$$

□

To compute this gradient, we have to compute the solution of an ODE in dimension $2k$, where k is supposed to be small, evaluate $2k \times K$ functions on the points of the trajectory $\hat{x}_\theta$ and compute a matricial multiplication along the small dimension of $\mathbf{X}$. This is far less expensive than solving $2k \times K$ ODE.

Remark 2.3.2. The equation satisfied by the adjoint can be recovered from a formal resolution of the optimization problem

$$\inf_{x_\theta} J(x)$$

under the constraint that $x_\theta \in H^1([0,1])^{2k}$ is solution of the Hamiltonian problem associated to $\hat{H}_\theta$. with $J(x_\theta) = \mathcal{L}(\theta)$.

The Lagrangian operator of this problem is given by

$$L(\theta,\lambda) = J(x_\theta) + \int_0^1 \lambda \cdot \big(x_\theta' - X_{\hat{H}_\theta}(x_\theta)\big)dt.$$

The point $(\theta,\lambda) \in \mathbf{R}^K \times H^1([0,1])^{2k}$ is a critical point if

$$\begin{cases} \partial_\theta L(\theta,\lambda) = 0, \\ \partial_\lambda L(\theta,\lambda) = x_\theta' - X_{\hat{H}_\theta}(x_\theta) = 0. \end{cases}$$

As all the functions are at stake are at least L^2, we can switch the integral and the derivative in the first equation, which gives for all $h \in \mathbf{R}^K$ small enough

$$\partial_x J(x_\theta)\big(d_\theta\mathcal{F}(h)\big) + \int_0^1 \lambda \cdot \left(\frac{d}{dt}d_\theta\mathcal{F}(h) - d_{x_\theta}X_{\hat{H}_\theta}\big(d_\theta\mathcal{F}(h)\big)\right) dt = 0.$$

After an integration by parts, the first term in the integral becomes

$$-\int_0^1 \lambda' \cdot d_\theta\mathcal{F}(h)dt + \lambda(1)\cdot d_\theta\mathcal{F}(h)(1) - \lambda(0)\cdot d_\theta\mathcal{F}(h)(0),$$

where the last term vanishes thanks to the initial condition on $d_\theta\mathcal{F}$.

We get

$$\int_0^1 \left[(x_\theta - {}^tDx) - \lambda' - \lambda \cdot d_{x_\theta} X_{\hat{H}_\theta}\right] \cdot d_\theta \mathcal{F}(h)dt + \lambda(1) \cdot d_\theta \mathcal{F}(h)(1) = 0.$$

Note that the jacobian of $X_{\hat{H}_\theta}$ is skew-symmetric. The previous equality is then verified if

$$\begin{cases} \lambda'(t) = d_{x_\theta} X_{\hat{H}_\theta}(\lambda(t)) + (x_\theta(t) - {}^tDx(t)), & \forall t \in [0,1] \\ \lambda(1) = 0, \end{cases}$$

which is exactly the adjoint problem.

2.4 Penalisation

In practice, the number K can be chosen very large. When this is the case, it slow down the evaluation of $X_{\hat{H}_\theta}$ and then reduce the efficiency of the reduction. To remedy this problem, we can add a penalization term at the loss function and minimize

$$\mathcal{L}(\theta) = \int_{g\in G} \int_{t\in[0,1]} \|D\hat{x}_{\theta,g}(t) - x_g(t)\|^2_{\mathbf{R}^{2k}} dgdt + \alpha\|\theta\|_1,$$

where α is a positive real number. The additional term forces a lot of coefficients of α to be exactly set to zero.

As it is given, this penalisation term is not differentiable. This is why we replace it in practice by $\theta \mapsto \sqrt{\theta^2 + \epsilon}$ for an epsilon chosen very small, 10^{-6} for example. Then, we just have to add $\frac{\theta}{\sqrt{\theta^2+\epsilon}}$ at the gradient we have found at the previous section.

2.5 Algorithms

Now, that we have a satifying expression for $\nabla\mathcal{L}(\theta)$, we can perform a numerical optimization using Algorithm 2.5.

Recall that we suppose that we have computed a set of solution $\{x_{g_i}\}_{i=1,\ldots,m}$ for some values $\{g_i\}_i \subset G$ of the parameter g. It can be a parameter of the primal equation or the initial condition. When g parameters the primal equation, a way to take this dependance into account in the reduction is to take functions f_i which involves this parameter. By doing this, the optimization will give a family $\{H_{\theta*,g}\}_g$, where θ^* is the optimal value of the parameter θ over all the trajectories x_{g_i}. When g only appears in the inital condition, we are only looking for one Hamiltonian function $H_{\theta*}$ so we have to take the f_i independant of g.

Of course, this basic algorithm can be sophisticated by adding momentum or by using Adam descent instead of the basic gradient descent. We will present comparisons for some cases in the following section.

During tests, it happens that taking into account only a small portion of the studied interval $[0,1]$ at each step can highly improve the descent efficiency. Below are exposed two precise algorithms which make use of this idea.

In algorithm 2.5, we in fact use a different gradient at each step, which is actually the gradient of $\mathcal{L}_g$ when the integration interval is $[b\Delta t, c\Delta t]$ instead of $[0,1]$. This way, we introduce

Algorithm 1 Simple gradient descent

Require: $\theta_0 \in \mathbf{R}^K$, α, $\eta > 0$, $\rho > 0$, $G = \{g_1, ..., g_m\}$, $x_G := \{x_{g_1}, ..., x_{g_m}\}$, and $\mathbf{X}$

$\theta \leftarrow \theta_0$

while $\frac{\|\nabla\mathcal{L}(\theta)\|}{\|\nabla\mathcal{L}(\theta_0)\|} > \eta$ **do**

  $\nabla \leftarrow \frac{\alpha\theta}{\sqrt{\theta^2+\epsilon}}$

  for all $g \in G$ **do**

    compute the solution x_θ of (2.1) for g and current θ

    compute a_θ of (2.7) for current θ, g, x_θ and x_g.

    compute $\nabla\mathcal{L}_g(\theta)$ from (2.8) with a_θ and x_θ

    $\nabla \leftarrow \nabla + \nabla\mathcal{L}_g(\theta)$

  end for

  $\theta \leftarrow \theta - \rho\nabla$

end while

Algorithm 2 Progressive gradient descent 1

Require: $\theta_0 \in \mathbf{R}^K$, α, $\eta > 0$, $\rho > 0$, $G = \{g_1, ..., g_N\}$, $x_G := \{x_{g_1}, ..., x_{g_N}\}$, $w \in [\![1, m]\!]$, Δt, $\mathbf{X}$

$\theta \leftarrow \theta_0$

while $\frac{\|\nabla\mathcal{L}(\theta)\|}{\|\nabla\mathcal{L}(\theta_0)\|} > \eta$ **do**

  $\nabla \leftarrow \frac{\alpha\theta}{\sqrt{\theta^2+\epsilon}}$

  for $i = 0, ..., \lfloor\frac{m}{w}\rfloor$ **do**

    $b \leftarrow iw$

    $c \leftarrow b + w$

    for all $g \in G$ **do**

      compute the solution $x_{\theta,g}$ of (2.1) starting at $x_0 = x_g(b\Delta t)$ on the interval $[b\Delta t, c\Delta t]$.

      compute $a_{\theta,g}$ of (2.7) from $x_{\theta,g}$ ending at $a(c\Delta t) = 0$ on the interval $[b\Delta t, c\Delta t]$.

      compute $\nabla\mathcal{L}_g(\theta)$ from (2.8) with a_θ and x_θ on the interval $[b\Delta t, c\Delta t]$.

      $\nabla \leftarrow \nabla + \nabla\mathcal{L}_g(\theta)$

    end for

    $\theta \leftarrow \theta - \rho\nabla$

  end for

end while

"stochastic-like" effect in the descent. A variation of this method can be to shift the considered window of one time step Δt instead of w time steps. Another one would consist in taking randomly the starting points b at each step. One can also imagine taking at each step several intervals instead of one. We will discuss all these variations in the following section.

Algorithm 3 Progressive gradient descent 2

Require: $\theta_0 \in \mathbf{R}^K$, α, $\eta > 0$, $\rho > 0$, $G = \{g_1, ..., g_N\}$, $x_G := \{x_{g_1}, ..., x_{g_N}\}$, $w \in [\![1, m]\!]$, Δt and $\mathbf{X}$

$\theta \leftarrow \theta_0$

while $\frac{\|\nabla \mathcal{L}(\theta)\|}{\|\nabla \mathcal{L}(\theta_0)\|} > \eta$ **do**

  $\nabla \leftarrow \frac{\alpha\theta}{\sqrt{\theta^2+\epsilon}}$

  for all $g \in G$ **do**

    compute the solution $x_{\theta,g}$ of (2.1) starting at $x_0 = x_g(0)$ on the interval $[0, 1]$.

  end for

  for $i = 0, ..., \lfloor \frac{m}{w} \rfloor$ **do**

    $b \leftarrow iw$

    $c \leftarrow b + w$

    for all $g \in G$ **do**

      compute $a_{\theta,g}$ of (2.7) from $x_{\theta,g}$ ending at $a(c\Delta t) = 0$ on the interval $[b\Delta t, c\Delta t]$.

      compute $\nabla \mathcal{L}_g(\theta)$ from (2.8) with a_θ and x_θ on the interval $[b\Delta t, c\Delta t]$.

      $\nabla \leftarrow \nabla + \nabla \mathcal{L}_g(\theta)$

    end for

    $\theta \leftarrow \theta - \rho\nabla$

  end for

end while

Algorithm 2.5 presents another way to optimize on subintervals which is less interpretable but which gives in some cases better results than the previous one. There is a single change in this new version : to compute the adjoint state, we use the solution of the primal problem computed with an inital condition at $t = 0$ whatever is b. Looking back to the proof of (2.8), we see that doing this way, the inital term after the integration by parts does not vanish. Therefore, the quantity we obtain is not longer the gradient of $\mathcal{L}_g$ on any interval. We do not even know if it is a gradient at all. We mentionned this method because in some cases, it gives better results than the two previous ones. We will present and discuss these results in the following section.

2.6 Tests

2.6.1 First tests on known target Hamiltonians

2.6.2 Tests on the reduction problem

3 Quadratic corrections for PSD

In this chapter, we present a different approach to learn the Hamiltonian in low dimension. We worked on it at the beginning of the internship but it does not give satisfying results. We present our work all the same because it is a lead that we could follow in future works.

As for the previous hyperreduction approach, we supposed that the PSD, or any other linear reduction method, has already given us a decoder $D : \mathbf{R}^{2k} \to \mathbf{R}^{2n}$ and an encoder $E : \mathbf{R}^{2n} \to \mathbf{R}^{2k}$ between the high and the low dimensional spaces. As we have seen for the non-linear piano string model, the induced reduction give bad results when the equations are too non-linear. This suggest that the submanifold Σ^k, on which lie the solutions of the PDE we are looking at, has too strong non-linearities to be embedded in a low dimensional vector subspace of $\mathbf{R}^{2n}$. Therefore, it seems that we have to look for non-linear decoders D if we want to obtain better results.

Here, we do not care about hyperreduction, which is viewed as the following step in the reduction process. We actually look if it is possible to improve the passage from the low dimensional space to the high dimensional one by adding a non-linear part to the decoder. More precisely, we want to adapt the method proposed in [10] in the symplectic case and add to the decoder a quadratic term which reduces the compression-decompression error. By doing so, we have to make sure that the new decoder remains symplectic. In the following, we present different variations of this idea and the results we obtained.

3.1 Shears

Before presenting the methods we have explored, we introduce the notion of *shear* that will turn out to be useful to build families of symplectic quadratic maps. We say that a map from $\mathbf{R}^m$ to $\mathbf{R}^r$ is quadratic if all of its r coordinate functions are polynomial of degree inferior or equal to 2.

Definition 3.1.1 (Shear, [16])**.** *A* shear transformation *is a map*

$$\sigma_V : \begin{cases} \mathbf{R}^{2m} \to \mathbf{R}^{2m} \\ (q,p) \mapsto (Q,P) \end{cases}$$

such that $Q_i = q_i$ and $P_i = p_i + \partial_{q_i} V(q)$ with $V : \mathbf{R}^n \to \mathbf{R}$ a cubic potential, that is a polynomial function of degree 3*.*

Proposition 3.1.1. *In the symplectic space $\mathbf{R}^{2m}$ endowed with the usual symplectic form, a shear is a symplectic quadratic map.*

Proof. To prove this assertion, it is sufficient to see that σ_V is the transformation induced by the generating function

$$S : (q, P) \mapsto Pq - V(q).$$

In fact, the transformation $(q, p) \mapsto (Q, P)$ induced by S verifies

$$\begin{cases} p_i = \frac{\partial S}{\partial q_i}(q, P) = P_i - \partial_{q_i} V(q), \\ Q_i = \frac{\partial S}{\partial P_i}(q, P) = q_i \end{cases}$$

for all i between 1 and m.

Moreover, as V is supposed to be cubic, σ_V is quadratic. □

Note that the shears form an Abelian group for the composition and that the correspondance $V \mapsto \sigma_V$ is a homomorphism between the additive group of the cubic functions and the shears.

Let us now see what is the physical meaning of a shear. Consider the Hamiltonian function $H : (q, p) \mapsto \frac{1}{2}|p| + V(q)$. It represents an energy function which is obtained by the sum of a kinetic energy and a potential one. The corresponding system is given by

$$\begin{cases} \dot{p} = -\frac{\partial H}{\partial q} = -V'(q), \\ \dot{q} = \frac{\partial H}{\partial p} = p \end{cases}$$

Denote by (q^t, p^t) the state of the system at a time t. When ϵ comes close to zero, we have

$$\begin{cases} p^{t+\epsilon} = p^t - \epsilon V'(q), \\ q^{t+\epsilon} = q^t + \epsilon p^t \end{cases}$$

We can see the transformation $(q^t, p^t) \mapsto (q^{t+\epsilon}, p^{t+\epsilon})$, induced by the flow of the equation, as the composition of the transformation $(q^t, p^t) \mapsto (q^t + \epsilon p^t, p^t)$, induced by the Hamiltonian function without potential energy, and the shear $\sigma_{\epsilon V}$. Therefore, we can link a shear with the effect of a potential on a system. With that point of view, it seems to be reasonable to try correct a bad reduction which is particularly wrong on the speed, as it is our case, with a shear.

The following result, proved in [16], gives a normal form, involving shears, for any quadratic symplectic map.

Theorem 3.1.1 (Normal form for symplectic quadratic maps, [16])**.** *Any quadratic symplectic map ϕ on $\mathbf{R}^{2m}$ can be decomposed as the composition of a symplectic linear map l, a shear σ and a symplectic affine function a of $\mathbf{R}^{2m}$, that is*

$$\phi = a \circ \sigma \circ l. \tag{3.1}$$

Moreover, a and l are linked by the formula

$$a = d\phi(0) \cdot l^{-1} + \phi(0).$$

This results implies that any quadratic symplectomorphism is invertible and has another quadratic symplectomorphism for inverse map.

Unfortunately, the proof of this result can not be adapted to the case of symplectic quadratic maps between $\mathbf{R}^{2m}$ and $\mathbf{R}^{2l}$ with $l < m$.

3.2 Quadratic correction with shears in low dimension

In this section, we look for a decoder D_{corr} of the form $D \circ \phi_\lambda$, where $\phi_\lambda : \mathbf{R}^{2k} \to \mathbf{R}^{2k}$ is a quadratic symplectic map.

3.2.1 Expression of the optimization problem

Expression of the corrected decoder

We first take for the family $\left(\phi_\lambda\right)_\lambda$ the group of the shears. The parameter λ is then the coefficients of V, which we write

$$V(y) = \sum_{1\leqslant i\leqslant k} \lambda_i y_i + \sum_{1\leqslant i\leqslant j\leqslant k} \lambda_{ij} y_i y_j + \sum_{1\leqslant i\leqslant j\leqslant l\leqslant k} \lambda_{ijl} y_i y_j y_l.$$

In the following, R will represent the number of coefficients. We count k terms of order 1 and $\frac{k(k+1)}{2}$ terms of order 2 in the previous expression for V. In the same way, there are

$$\begin{aligned}
\sum_{i=1}^{k}\sum_{j=i}^{k}\sum_{l=j}^{k} 1 = \sum_{i=1}^{k}\sum_{j=i}^{k}(k-j+1) = \sum_{i=1}^{k}\sum_{m=1}^{k-i+1} m = \sum_{i=1}^{k} \frac{(k-i+1)(k-i+2)}{2} \\
= \sum_{i=1}^{k}\Big(\frac{(k+1)(k+2)}{2} - i\frac{2k+3}{2} + \frac{i^2}{2}\Big) \\
= \frac{k(k+1)(k+2)}{2} - \frac{2k+3}{2}\frac{k(k+1)}{2} + \frac{k(k+1)(2k+1)}{12} \\
= \frac{k(k+1)(k+2)}{6}
\end{aligned}$$

terms of order 3 and

$$R = \frac{k(k+1)(k+5)}{6} + k.$$

According with the normal form of symplectic quadratic maps, if we would like to cover the whole space of quadratic symplectomorphism of $\mathbf{R}^{2k}$ in $(\phi_\lambda)_\lambda$, we would have to compose the shears with affines and linear symplectomorphism as in 3.1. This complicates a lot the problem so we limit ourselves to taking $a = l = id$.

Objective function

We are looking for the $\lambda \in \mathbf{R}^R$ such that

$$\|X - D\phi_\lambda \hat{X}\|_{F,2n,N}^2$$

is minimal, where $\|\cdot\|_{F,2n,N}$ denotes the Froebenius norm in $\mathcal{M}_{2n,N}(\mathbf{R})$, X the matrix of the N sample in $\mathbf{R}^{2n}$ and $\hat{X}$ the matrix of their PSD reduction in $\mathbf{R}^{2k}$. It happens that

$$\begin{aligned}
\|X - D\phi_\lambda \hat{X}\|_{F,2n,N}^2 &= \|(Q,P) - D(\hat{Q}, \hat{P} + \nabla V(\hat{Q}))\|_{F,2n,N}^2 \\
&= \|Q - A_{psd}\hat{Q}\|_{F,n,N}^2 + \|P - \hat{P} - \nabla V(\hat{Q}))\|_{F,n,N}^2
\end{aligned}$$

where A_{psd} is the submatrix of D such that $D = \begin{pmatrix} A_{psd} & 0 \\ 0 & A_{psd} \end{pmatrix}$. The previous problem is then equivalent to solving

$$\underset{V\in\mathcal{P}_3(\mathbf{R}^k)}{\operatorname{argmin}} \|P - \hat{P} - \nabla V(\hat{Q}))\|_{F,n,N}^2, \tag{3.2}$$

where $\mathcal{P}_3(\mathbf{R}^k)$ denotes the ring of polynomial functions of degree at most 3 on $\mathbf{R}^k$.

3.2.2 Least-square formulation

Clearly, $\lambda \mapsto \phi_\lambda$ is a linear map. We can therefore rewrite the problem in a way such that it becomes a least-square one. Set

$$\Lambda = {}^t\begin{pmatrix}\lambda_1 & \dots & \lambda_k & \lambda_{11} & \lambda_{12} & \dots & \lambda_{kk} & \lambda_{111} & \dots & \lambda_{kkk}\end{pmatrix} \in \mathbf{R}^R.$$

In what follows, we use the lexico-graphic order when we work with the elements $\{i,j,l\}$ of $[\![1,k]\!]^3$, starting by the smallest indice of the set and finishing by the largest. To simplify further notations, we introduce the function $P : [\![1,k]\!]^3 \to [\![1,R]\!]$ which associates to the set $\{i,j,l\}$ its position in the ordered list of all the 3-uplets (i_1,i_2,i_3) verifying $i_1 \leqslant i_2 \leqslant i_3$. For $i \leqslant j \leqslant l$, we have that

$$\begin{aligned}
P(i,i,i) &= \sum_{r=1}^{i-1}\sum_{s=r}^{k}\sum_{t=s}^{k} 1 = \sum_{r=1}^{i-1} \frac{(k-r+1)(k-r+2)}{2} \\
&= \frac{i-1}{2}\left(k(k+3) + i(k+\frac{3}{2}+\frac{i(2i-1)}{6})\right) + i, \\
P(i,j,j) &= P(i,i,i) + \sum_{s=i}^{j-1}\sum_{t=s}^{k} 1 = P(i,i,i) + \frac{(j-i)(2k-i-j+3)}{2}, \\
P(i,j,l) &= P(i,j,j) + l - j.
\end{aligned}$$

For $1 \leqslant l \leqslant k$, let $S = 1 + k + \frac{k(k+1)}{2}$. Let also $F_l \in \mathcal{M}_{S,1}(\mathbf{R})$ be the matrix of the map $y \in \mathbf{R}^k \mapsto \partial_l V(y)$, verifying $\partial_l V(y) = F_l \bar{Y}$ with $\bar{Y} = \begin{pmatrix}1 & y_1 & \dots & y_k & y_1y_1 & \dots & y_ky_k\end{pmatrix}$. For all $1 \leqslant l \leqslant k$, we have that

$$\partial_l V(y) = \lambda_l + \sum_{i=1 i\neq j}^{k} \lambda_{il}x_i + 2\lambda_{ll}x_l + \sum_{1\leqslant i\leqslant j\leqslant k i,j\neq l} \lambda_{ijl}x_ix_j + \sum_{i=1 i\neq l}^{k} 2\lambda_{ill}x_ix_l + 3\lambda_{lll}x_l^2$$

so $F_l = G^l\Lambda$ with G^l in $\mathcal{M}_{S,R}(\mathbf{R})$ such that

$$\begin{cases}
G^l_{1,l} = G^l_{i+1,P(1,i,l)+k} = G^l_{P(1,i,j)+k+1,P(i,j,l)+K+k} = 1, \\
G^l_{l+1,P(1,l,l)+k} = G^l_{P(1,i,l)+k+1,P(i,l,l)+K+k} = 2, \\
G^l_{P(1,l,l)+k+1,P(l,l,l)+K+k} = 3
\end{cases}$$

for all i,j in $[\![1,k]\!]$ different from l with $K = \frac{k(k+1)}{2}$.

If we set

$$G = \begin{pmatrix} G^1 \\ \hline \dots \\ \hline G^k \end{pmatrix} \qquad \text{and} \qquad \bar{\bar{Y}} = \begin{pmatrix} {}^t\bar{Y} & & 0 \\ & \ddots & \\ 0 & & {}^t\bar{Y} \end{pmatrix},$$

we then have that $\nabla V(y) = \bar{\bar{Y}}G\Lambda$. Now, if we want to apply this to the matrix $\hat{Q}$ of the snapshots and if we take

$$\bar{\bar{\hat{Q}}} = \begin{pmatrix} {}^t\bar{\hat{Q}} & & 0 \\ & \ddots & \\ 0 & & {}^t\bar{\hat{Q}} \end{pmatrix},$$

we have to multiply the result by the permutation matrix P with sends the $(ik + j)$-th line of $\bar{\bar{\hat{G}}}G\Lambda$ on the $(jN + i)$-th one, where N is the number of solutions recorded in X.

This finally gives the least-sqare formulation of the problem 3.2 :

$$\underset{V \in \mathcal{P}_3(\mathbf{R}^k)}{\operatorname{argmin}} \|X - P\bar{\bar{\hat{Q}}}G\Lambda\|^2_{F,n,N},$$

whose solution is given by

$$\left({}^t(P\bar{\bar{\hat{Q}}}G)(P\bar{\bar{\hat{Q}}}G) \right)^{-1} {}^t(P\bar{\bar{\hat{Q}}}G)X.$$

3.2.3 Corrected reduced model

The new decoder is written $D_{corr} = D \circ \sigma_V$. We have

$$D(q,p) = (D^q_{corr}(q,p), D^p_{corr}(q,p)) = (A_{psd}q, A_{psd}p + \nabla V(q)$$

and therefore

$$\begin{cases} \nabla_q D^q_{corr}(q,p) = A_{psd}, \\ \nabla_p D^q_{corr}(q,p) = 0, \end{cases} \qquad \begin{cases} \nabla_q D^p_{corr}(q,p) = \nabla^2 V(q), \\ \nabla_p D^p_{corr}(q,p) = A_{psd}. \end{cases}$$

Then, the components of the new reduced Hamiltonian are given by

$$\begin{cases} \nabla_q \hat{H}(q,p) = {}^t\nabla_q D^q_{corr}(q,p) \cdot \nabla_q H(D_{corr}(q,p)) + {}^t\nabla_q D^p_{corr}(q,p) \cdot \nabla_p H(D_{corr}(q,p)) \\ \qquad\qquad\quad = {}^tA_{psd} \cdot \nabla_q H(D_{corr}(q,p)) + {}^t\nabla^2 V(q) \cdot \nabla_p H(D_{corr}(q,p)) \\ \nabla_p \hat{H}(q,p) = {}^t\nabla_p D^q_{corr}(q,p) \cdot \nabla_q H(D_{corr}(q,p)) + {}^t\nabla_p D^p_{corr}(q,p) \cdot \nabla_p H(D_{corr}(q,p)) \\ \qquad\qquad\quad = {}^tA_{psd} \cdot \nabla_p H(D_{corr}(q,p)). \end{cases}$$

3.2.4 Results

We have implemented this correction and we obtained the results presented Figures **??** and **??**. We see that the correction do not improve the solution in low dimension.

This can be explained by the fact that we do not change the image of the decoder in $\mathbf{R}^{2n}$. More precisely, the PSD gives a linear subspace of $\mathbf{R}^{2n}$ but the manifold on which lies the soltuions may not be linear at all. The best that we can obtain with the PSD is therefore a linear subspace in which this target manifold is included. In the cases where this manifold is highly non-linear, this subspace can be of high dimension and this is why the linear reduction fails.

Now, when we correct the decoder produced by the PSD, we change the reduced Hamiltonian but we do not change the space in which it takes its values. If the dimension of the linear subspace obtained with the PSD is too high or if it does not include the true manifold, then the quadratic correction has few chance to really improve the resolution in low dimension.

3.3 Additive quadratic correction

Following an idea found in [10], we now look for a decoder of the form $D_{corr} = D + \phi_\lambda$.

3.3.1 Tentative 1

Symplecticity condition

We look for ϕ_λ of the form $y \mapsto \bar{A}\tilde{Y}$ where $\tilde{Y} = (y_1y_1, y_2y_2, ..., y_{2k}y_{2k})$ and $\bar{A}$ is in $\mathcal{M}_{2n,2k}(\mathbf{R})$. In this case, the jacobian of the new decoder at point $y \in \mathbf{R}^{2k}$ is given by $D + 2\bar{A}Y$, with $Y = \begin{pmatrix} Y_q & 0 \\ 0 & Y_p \end{pmatrix}$, where Y_q and Y_p are the $n \times k$ diagonal matrices with diagonal coefficients $y_1, ..., y_k$ and $y_{k+1}, ..., y_{2k}$.

Decompose $\bar{A}$ into four submatrices a, b, c and d of $\mathcal{M}_{2n,2k}(\mathbf{R})$:

$$\bar{A} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

This gives

$$\mathcal{J}D_{corr}(y) = \begin{pmatrix} A_{psd} + 2AY_q & 2BY_p \\ 2CY_q & A_{psd} + 2DY_p \end{pmatrix}.$$

According to previous computations, D_{corr} is symplectic if and only if

$$\begin{cases} 2Y_q^t C(A_{psd} + 2AY_q) = 2^t(A_{psd} + 2AY_q)CY_q, \\ 2^t(A_{psd} + 2DY_p)BY_p = 2Y_p^t B(A_{psd} + 2DY_p), \\ {}^t(A_{psd} + 2DY_p)(A_{psd} + 2AY_q) - 4Y_p^t BCY_q = I_k \end{cases}$$

for all Y_p and Y_q diagonal matrices in $\mathcal{M}_{n,k}(\mathbf{R})$.

Rearranging the terms and taking into account that ${}^tA_{psd}A_{psd} = I_k$ by construction, we obtain

$$\begin{cases} 2Y_q({}^tAC - {}^tCA)Y_q + {}^tA_{psd}CY_q - Y_q{}^tCA_{psd} = 0, \\ 2Y_p({}^tDB - {}^tBD)Y_p + {}^tA_{psd}BY_p - Y_p{}^tBA_{psd} = 0, \\ 2Y_p({}^tBC - {}^tDA)Y_q - {}^tA_{psd}AY_q - Y_p{}^tDA_{psd} = 0. \end{cases} \tag{3.3}$$

for all Y_p and Y_q.

When we set $(Y_q, Y_p) = (I_k, 0)$ and $(Y_q, Y_p) = (0, I_k)$ in the third equation, we respectively get ${}^tA_{psd}A = 0$ and ${}^tA_{psd}D = 0$. Inserting these results in the third equation with $(Y_q, Y_p) = (I_k, I_k)$, we have ${}^tBC - {}^tDA = 0$. Conversely, if we have ${}^tA_{psd}A = {}^tA_{psd}D = 0$ and ${}^tBC - {}^tDA = 0$, the third equation is true for all Y_q, Y_p.

On the other hand, when we take $Y_q = I_k$ and $-I_k$ in the first equation, we obtain $2({}^tAC - {}^tCA) + {}^tA_{psd}C - {}^tCA_{psd} = 0$ and $-2({}^tAC - {}^tCA) + {}^tA_{psd}C - {}^tCA_{psd} = 0$. Adding the two equation gives ${}^tA_{psd}C - {}^tCA_{psd} = 0$, substrating them ${}^tAC - {}^tCA = 0$. Now, if we introduce the last expression in the third equation of 3.3, we get ${}^tA_{psd}CY_q - Y_q{}^tCA_{psd} = 0$ for all Y_q. Let us see that the symmetry of ${}^tA_{psd}C$ that we have just shown implies then that ${}^tA_{psd}C = 0$. In fact, if for all indices i, j in $[\![1, n]\!]$,

$$\sum_{l=1}^{n}(A_{psd})_{li}C_{lj} = \left[{}^tA_{psd}C\right]_{ij} = \left[{}^tCA_{psd}\right]_{ij} = \sum_{l=1}^{n}(A_{psd})_{lj}C_{li},$$

then

$$\sum_{l=1}^{n}(A_{psd})_{li}C_{lj}(Y_q)_{jj} = \left[{}^tA_{psd}CY_q\right]_{ij} = \left[Y_q{}^tCA_{psd}\right]_{ij} = \sum_{l=1}^{n}(A_{psd})_{lj}C_{li}(Y_q)_{ii} = \sum_{l=1}^{n}(A_{psd})_{li}C_{lj}(Y_q)_{ii}$$

so

$$((Y_q)_{jj} - (Y_q)_{ii})\left[{}^tA_{psd}C\right]_{ij} = ((Y_q)_{jj} - (Y_q)_{ii})\sum_{l=1}^{n}(A_{psd})_{li}C_{lj} = 0$$

for all indices i, j and for all values of $(Y_q)_{jj}$ and $(Y_q)_{ii}$. This implies that ${}^tA_{psd}C = 0$. Conversely, if we have ${}^tA_{psd}C = 0$ and ${}^tAC - {}^tCA = 0$, then the first equation of 3.3 in true for all value of Y_q. Exactly in the same way, we find that the second equation of 3.3 is equivalent to ${}^tA_{psd}B = 0$ and ${}^tDB - {}^tBD = 0$.

Optimization problem

Consider the loss function

$$\mathcal{L} : \begin{cases} \mathcal{M}_{n,k}(\mathbf{R})^4 \to \mathbf{R} \\ (A, B, C, D) \mapsto \|\bar{X} - \bar{A}\tilde{\hat{X}}\|^2_{F,2n,N}, \end{cases}$$

where $\bar{X}$ denotes the compression-decompression error made by the PSD on the samples, that is $X - D\hat{X}$.

According to the preceeding section, D_{corr} is symplectic if and only if $\bar{A}$ is a zero of the functions

$$\begin{aligned} g_1 &: (A, B, C, D) \mapsto \|{}^tA_{psd}A\|^2_{F,k,k}, \\ g_2 &: (A, B, C, D) \mapsto \|{}^tA_{psd}B\|^2_{F,k,k}, \\ g_3 &: (A, B, C, D) \mapsto \|{}^tA_{psd}C\|^2_{F,k,k}, \\ g_4 &: (A, B, C, D) \mapsto \|{}^tA_{psd}D\|^2_{F,k,k}, \\ g_5 &: (A, B, C, D) \mapsto \|{}^tAC - {}^tCA\|^2_{F,k,k}, \\ g_6 &: (A, B, C, D) \mapsto \|{}^tDB - {}^tBD\|^2_{F,k,k}, \\ g_7 &: (A, B, C, D) \mapsto \|{}^tBC - {}^tDA\|^2_{F,k,k}. \end{aligned}$$

Whatever the dimension of the matrices we are looking at, the Froebenius norm is Euclidian for the scalar product $(\cdot, \cdot) : (A, B) \mapsto Tr({}^tAB)$. On $\mathcal{M}_{n,k}(\mathbf{R})^4$, we use the scalar product induced by the Cartesian product

$$\langle (A, B, C, D); (E, F, G, H) \rangle = (A, E) + (B, F) + (C, G) + (D, H)$$

and the associated norm $||| \cdot |||$.

Note

$$K_i = \{(A, B, C, D) \in \mathcal{M}_{n,k}(\mathbf{R})^4 \mid g_i(A, B, C, D) = 0\}$$

and $K = \bigcap_{i=1}^7 K_i$. We want to solve

$$\underset{(A,B,C,D)\in K}{min} \mathcal{L}(A, B, C, D).$$

Consider Taylor's expansion of $\mathcal{L}$ at (A, B, C, D) in the direction of (h_1, h_2, h_3, h_4):

$$\begin{aligned} &\mathcal{L}(A + h_1, B + h_2, C + h_3, D + h_4) \\ &= \|\bar{Q} - (A + h_1)\tilde{\hat{Q}} - (B + h_2)\tilde{\hat{P}}\|^2_{F,n,k} + \|\bar{P} - (C + h_3)\tilde{\hat{Q}} - (D + h_4)\tilde{\hat{P}}\|^2_{F,n,k} \\ &= \|\bar{Q} - A\tilde{\hat{Q}} - B\tilde{\hat{P}}\|^2_F - 2\left(\bar{Q} - A\tilde{\hat{Q}} - B\tilde{\hat{P}}, h_1\tilde{\hat{Q}} + h_2\tilde{\hat{P}}\right) + \|h_1\tilde{\hat{Q}} + h_2\tilde{\hat{P}}\|^2_F \\ &\quad + \|\bar{P} - C\tilde{\hat{Q}} - D\tilde{\hat{P}}\|^2_F - 2\left(\bar{P} - C\tilde{\hat{Q}} - D\tilde{\hat{P}}, h_3\tilde{\hat{Q}} + h_4\tilde{\hat{P}}\right) + \|h_3\tilde{\hat{Q}} + h_4\tilde{\hat{P}}\|^2_F. \end{aligned}$$

Thanks to the properties of the trace, we have

$$\left(\bar{Q} - A\tilde{\hat{Q}} - B\tilde{\hat{P}}, h_1\tilde{\hat{Q}}\right) = Tr\left({}^t(\bar{Q} - A\tilde{\hat{Q}} - B\tilde{\hat{P}})h_1\tilde{\hat{Q}}\right) = Tr\left(\tilde{\hat{Q}}^t(\bar{Q} - A\tilde{\hat{Q}} - B\tilde{\hat{P}})h_1\right)$$
$$= \left((\bar{Q} - A\tilde{\hat{Q}} - B\tilde{\hat{P}})^t\tilde{\hat{Q}}, h_1\right).$$

Equivalent expressions holds for the other terms involving one of the h_i, which gives

$$\begin{aligned}
&\mathcal{L}(A + h_1, B + h_2, C + h_3, D + h_4) \\
&= \mathcal{L}(A, B, C, D) - 2\left((\bar{Q} - A\tilde{\hat{Q}} - B\tilde{\hat{P}})^t\tilde{\hat{Q}}, h_1\right)) - 2\left((\bar{Q} - A\tilde{\hat{Q}} - B\tilde{\hat{P}})^t\tilde{\hat{P}}, h_2\right) \\
&\quad - 2\left((\bar{P} - C\tilde{\hat{Q}} - D\tilde{\hat{P}})^t\tilde{\hat{Q}}, h_3\right) - 2\left((\bar{P} - C\tilde{\hat{Q}} - D\tilde{\hat{P}})^t\tilde{\hat{P}}, h_4\right) + \mathcal{O}(\|\|h_1, h_2, h_3, h_4\|\|^2)
\end{aligned}$$

so the gradient of $\mathcal{L}$ for our scalar product is

$$-2\left((\bar{Q} - A\tilde{\hat{Q}} - B\tilde{\hat{P}})^t\tilde{\hat{Q}}, (\bar{Q} - A\tilde{\hat{Q}} - B\tilde{\hat{P}})^t\tilde{\hat{P}}, (\bar{P} - C\tilde{\hat{Q}} - D\tilde{\hat{P}})^t\tilde{\hat{Q}}, (\bar{P} - C\tilde{\hat{Q}} - D\tilde{\hat{P}})^t\tilde{\hat{P}}\right).$$

Using the same arguments, we find that

$$\nabla g_1(A, B, C, D) = (2a^tA_{psd}A, 0, 0, 0)$$

and that ∇g_2, ∇g_3 and ∇g_4 have similar expressions. Unfortunately, we also find that

$$\nabla g_5(A, B, C, D) = 4(C^tCA - C^tAC, 0, A^tAC - A^tCA, 0)$$

which equals to zero when $g_5(A, B, C, D) = 0$. Similar results hold for g_6 and g_7 so we can't use usual theoretical tools to caracterize local minima.

Numerically, we will use a gradient descent to find a value of $\bar{A}$ which achieve a small value of the loss.

Results

3.3.2 Tentative 2

We now want to add crossed terms y_iy_j for $i \neq j$ in the quadratic map. We then look for ϕ_λ of the form $y \mapsto \bar{A}\tilde{Y}$ where $\tilde{Y} = (y_1y_1, y_1y_2, ..., y_{2k}y_{2k})$ and $\bar{A}$ is in $\mathcal{M}_{2n,S}(\mathbf{R})$. Recall from a previous section that $S = 1 + k + \frac{k(k+1)}{2}$ is the dimension of $\tilde{Y}$.

Optimization problem

Decompose $\bar{A}$ in $\binom{A}{B}$ with A and B in $\mathcal{M}_{n,S}$.

Let G the matrix in $\mathcal{M}_{S,4k^2}(\mathbf{R})$ whose coefficients are 0 except for : the $G_{P(1,i,i);2ki+i}$ which are 2 for $i \in [\![1, 2k]\!]$ and for the $G_{P(1,i,j),2ki+j}$ and the $G_{P(1,i,j),2kj+i}$, which are 1 for all $i < j$ between 1 and $2k$. The Jacobian matrix of $f : y \mapsto \tilde{Y}$ at y is given by $\mathcal{J}f(y) = G\check{Y}$, where

$$\check{Y} = \begin{pmatrix} Y & 0 & \dots & 0 \\ 0 & Y & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & Y \end{pmatrix} \in \mathcal{M}_{4k^2,2k}(\mathbf{R}).$$

Then, for $D_{corr} : x = (q,p) \mapsto (A_{psd}q + Af(x), A_{psd}p + Bf(x))$, we have

$$\begin{cases} \nabla_q D^q_{corr}(\hat{x}) = A_{psd} + AG\check{\hat{Q}}, \\ \nabla_q D^p_{corr}(\hat{x}) = BG\check{\hat{Q}}, \end{cases} \qquad \begin{cases} \nabla_p D^q_{corr}(\hat{x}) = AG\check{\hat{P}}, \\ \nabla_p D^p_{corr}(\hat{x}) = A_{psd} + BG\check{\hat{P}}, \end{cases}$$

where $\check{\hat{Q}}$ and $\check{\hat{P}}$ in $\mathcal{M}_{4k^2,k}(\mathbf{R})$ are such that $\check{\hat{X}} = \left(\check{\hat{Q}} \mid \check{\hat{P}} \right)$.

Therefore, the symplecticity conditions for D_{corr} are, for all $\check{\hat{Q}}$ and $\check{\hat{P}}$

$$\begin{cases} {}^tA_{psd}BG\check{\hat{P}} + {}^t\check{\hat{Q}}{}^tG({}^tAB - {}^tBA)G\check{\hat{P}} + {}^t\check{\hat{Q}}{}^tG^tAA_{psd} = 0, \\ {}^t(A_{psd} + AG\check{\hat{Q}})BG\check{\hat{Q}} - {}^t\check{\hat{Q}}{}^tG^tB(A_{psd} + AG\check{\hat{Q}}) = 0, \\ {}^t\check{\hat{P}}{}^tG^tA(A_{psd} + BG\check{\hat{P}}) - {}^t(A_{psd} + BG\check{\hat{P}})AG\check{\hat{P}} = 0. \end{cases}$$

If we take $\check{\hat{Q}} = 0$ in the first equation, we have ${}^tA_{psd}BG\check{\hat{P}} = 0$ for all $\check{\hat{P}}$. We can choose $\check{\hat{P}}$ with all but one coefficient equal to zero. This leads to ${}^tA_{psd}BG_i = 0$ for all column G_i of G, which means that ${}^tA_{psd}BG = 0$. The same argument with $\check{\hat{P}} = 0$ shows that ${}^tA_{psd}AG = 0$. Now, taking $\check{\hat{P}}$ and $\check{\hat{Q}}$ with all but one coefficients equal to zero, we see that all the coefficients of ${}^tG({}^tAB - {}^tBA)G$ are zero. Conversely, if ${}^tA_{psd}AG = {}^tA_{psd}BG = {}^tG({}^tAB - {}^tBA)G = 0$, then the three equation above are satisfied for all $\check{\hat{Q}}$ and $\check{\hat{P}}$ and D_{corr} is symplectic.

The optimization problem that we want to solve is therefore

$$\underset{(A,B)\in K}{min} \mathcal{L}(A,B,C,D).$$

with the loss function

$$\mathcal{L} : \begin{cases} \mathcal{M}_{n,S}(\mathbf{R})^2 \to \mathbf{R} \\ (A,B) \mapsto \|\bar{X} - \bar{A}\tilde{\hat{X}}\|^2_{F,2n,N}, \end{cases}$$

where $\bar{X}$ denotes the compression-decompression error made by the PSD on the samples, that is $X - D\hat{X}$ and $K = \{(A,B) \in \mathcal{M}_{n,S}(\mathbf{R})^2 \quad | \quad {}^tA_{psd}AG = {}^tA_{psd}BG = {}^tG({}^tAB - {}^tBA)G = 0\}$.

Results

Part II

Geometric part

4 Homotopy principle

4.1 Goals

The aim of the geometrical part of this work is to give theoretical justifications to the methods we develop in the numerical part. In particular, we would like to know if there is no geometrical obstacle to learning the manifold Σ^k. More precisely, we would like to prove the following conjecture :

Conjecture 4.1.1. *Let $n, k \in \mathbf{N}$ such that $k \ll n$. Consider two manifolds Σ^k and $\tilde{\Sigma}^k$ embedded in $\mathbf{R}^{2n}$ endowed with its usual symplectic structure. Denote by $i : \Sigma^k \to \mathbf{R}^{2n}$ and $\tilde{i} : \tilde{\Sigma}^k \to \mathbf{R}^{2n}$ the corresponding inclusions.*

If k is sufficiently small in front of n, then it exists a symplectic homeomorphism $h : \mathbf{R}^{2n} \to \mathbf{R}^{2n}$ such that $h(\Sigma^k) = \tilde{\Sigma}^k$.

Moreover, if i and $\tilde{i}$ are $\mathcal{C}^0$-close, then h is $\mathcal{C}^0$-close from the identity.

If this result is true, then for all Hamiltonian function $H : \mathbf{R}^{2n} \to \mathbf{R}$ whose flow preserves Σ^k, the flow of the composition $\tilde{H} = H \circ h^{-1}$ preserves $\tilde{\Sigma}^k$. This is immediate since $\phi^t_{\tilde{H}} = h \circ \phi^t_H \circ h^{-1}$. In the case Σ^k and $\tilde{\Sigma}^k$ are $\mathcal{C}^0$-close, then H and $\tilde{H}$ are also $\mathcal{C}^0$-close and the restriction of their flows on borned intervals of $\mathbf{R}$ too.

In other words, if we make a small error when learning the solution manifold, which is highly probable since we interpolate it with a finite number of points, then the part of the errors on solutions induced by errors we made on Σ^k remains small. If this result is true, then we can hope to learn the dynamic on the interpolated manifold as we try to do.

This result has already been proved in the particular case of isotropic submanifolds in [3] using methods based on the Gromov's h-principle. The h-principle, an abbreviation for *homotopy principle*, is a principle or a caracteristic of some spaces which, if it holds, guaranty the existence of solutions for differential problems. It involves a new point of view on differential equalities and inequalities, involving for example notions of jets and differential relations. When working with the h-principle, one usually want to establish it and there is some particular techniques to achieve this.

During this internship, we worked to understand the methods that were used to prove the results in [3]. Eventually, we will use them to extend the proof of 4.1.1 to the general case. To become familiar with the h-principle, we read the parts of [8] devoted to the h-principle and its proofs using holonomic approximation theorem. In this chapter, we present the main notions that one needs to understand what is the h-principle and some of its applications. We present all this notions in a way that the chapter leads to the proof of the following theorem :

Theorem 4.1.1. *Let (V, ω_V) and (W, ω_W) be two symplectic manifolds of respective dimensions*

$n = 2l$ and $q = 2m$. Let $f_0 : V \to W$ be an embedding such that $f_0^*[\omega_W] = [\omega_V]$. Suppose also that $F_0 = df_0$ is homotopic to an isosymplectic homomorphism F_1 via $F_t \subset \mathcal{R}_{imm}$ such that $bsF_t = f_0$ for all $t \in [0,1]$.

Then, if V is open and $m < l$, then it exists an isotopy $f_t : V \to W$ such that f_1 is isosymplectic and df_1 is homotopic to F_1 in the isosymplectic homomorphisms.

Moreover, if K is a core of V, we can choose f_t arbitrarily $\mathcal{C}^0$-close to f_0 near K.

All the definitions and the results we present here are taken from [8]. The proofs are taken from the same reference but we had details in most of them.

In what follows we are sometimes required to consider a metric on some manifolds. Everytime we talk about $\mathcal{C}^0$-closeness of applications, we imply the existence of a distance on the space where those functions take their values. We also need it to build normal neighbourhoods in some proofs. A simple way to define a distance on manifolds is to consider a Riemannian metric on this manifold and we know that this is always possible (see [9] for a proof of this assertion). When needed, we therefore consider a Riemannian structure on the considered manifold. This additionnal structure is only useful to properly define normal neighbourhoods or $\mathcal{C}^0$-closeness, it does not change anything to the geometry of the problems we consider.

4.2 Definitions

4.2.1 Jets

Jet spaces

For our purpose, we are interested on derivatives of functions. The notion of jets allows us to designate and manipulate the functions along with their derivatives. Let us first see how we define jets in the simple case of functions of $\mathbf{R}^l$.

Definition 4.2.1 (Space of r-jets on Euclidian spaces [8])**.** *Let $r \in \mathbf{N}^*$. The* space of r-jets *of functions $\mathbf{R}^n \to \mathbf{R}^q$ is the space of all* a priori *possible values of a function $f : \mathbf{R}^n \to \mathbf{R}^q$ and its derivatives of order at most r at a point of $\mathbf{R}^n$, that is*

$$\mathbf{R}^n \times \mathbf{R}^q \times \mathbf{R}^{qd(n,1)} \times ... \times \mathbf{R}^{qd(n,r)},$$

where $d(m,l)$ is the number of partial derivatives of order l for a function $f : \mathbf{R}^m \to \mathbf{R}$. We note this space $J^r(\mathbf{R}^n, \mathbf{R}^q)$.

The space of all possible values that can take a function and its derivative at a point $v \in \mathbf{R}^n$ is really the product we have mentionned : for any point P in this product space such that $\pi_1(P) = v$, it exists a polynom of degree r in $\mathbf{R}^n$ whose r-order derivatives at v agree with P.

Remark 4.2.1. We have $d(n,r) = \frac{(n+r-1)!}{(n-1)!r!}$. This can be proved by induction on n. The formula is clearly true for $n = 1$. Suppose now that the formula is true for all $i \leqslant n$. We have

$$d(n,r) = \sum_{i_1=0}^{n} \sum_{i_2=0}^{i_1} ... \sum_{i_r=0}^{i_{r-1}} 1$$

so $d(n+1,r) = d(n,r) + d(n+1,r-1)$. An immediate induction on r gives then $d(n+1,r) = \sum_{i=2}^{r} d(n,i) + d(n+1,1)$. Since $d(n,1) = n$ and $d(n,0) = 1$ for any n, this can be rewritten as $d(n+1,r) = \sum_{i=0}^{r} d(n,i)$. By assumption, this is equivalent to $d(n+1,r) = \sum_{i=0} \binom{n-1+i}{n-1}$. Using the "hockey cross" formula,

$$\sum_{j=0}^{m-k} \binom{j+k}{k} = \binom{m+1}{m-k},$$

which is true for all $m, k \in \mathbf{N}$ such that $k < m$, we get $d(n+1,r) = \binom{n+r}{r} = \frac{(n+r)!}{n!r!}$. We then have proved that the formula is also true for $n+1$. By the induction principle, we deduce that the formula is true for all $n > 1$.

We define now jets on manifolds. For this purpose, we use the definition we have given on Euclidian spaces. We simply need to adapt it to make it invariant by change of coordinates.

Definition 4.2.2 (Space of r-jets in the general case [8])**.** *Let V and W two manifolds of respective dimensions n and q. Let $v \in V$ and $U \subset V$ be an open neighbourhood of v in V on which is defined a coordinate system $\phi : U \to \mathbf{R}^n$. We say that two functions f and g from U to W are* r-tangent *at v if they agree at v and if the r-order derivatives of $\phi_* f$ and $\phi_* g$ agree at $\phi(v)$.*

Tangency at v gives rise to an equivalence relation : two functions are in the same class if and only if they are r-tangent. The space of r-jets *$J^r(V,W)$ is defined as the space of all r-tangency classes at any point of V.*

When $V = \mathbf{R}^n$ and $W = \mathbf{R}^q$, the previous definition is equivalent to the first one.

With the chain rule, we verify that this definition is indeed invariant under a change of coordinates. Let $\psi : U \to \mathbf{R}^n$. We have

$$d_{\phi(v)}\phi_* f = d_{\phi(v)}(f \circ \phi^{-1}) = d_{\phi\circ\psi^{-1}\circ\psi(v)}(f \circ \psi^{-1} \circ \psi \circ \phi^{-1}) = d_{\psi(v)}(f \circ \psi^{-1}) \circ d_{\phi(v)}(\psi \circ \phi^{-1}).$$

Since $d_{\phi(v)}(\psi \circ \phi^{-1})$ is invertible,

$$d_{\phi(v)}\phi_* f = d_{\phi(v)}\phi_* g \quad \Longleftrightarrow \quad d_{\psi(v)}\psi_* f = d_{\psi(v)}\psi_* g.$$

If we replace f and g by partial derivatives of order inferior to r, we obtain the invariance of the notion of r-tangency for $r > 1$.

Note that for $l < m$, the projection $p_l^m : J^m(V,W) \to J^l(V,W)$ which sends a class of m-tangency to a class of l-tangency by "forgetting" the derivatives of order superior to l is invariantly defined. In fact, if two functions are m-tangent at a point v, they are also l-tangent at the same point for all $l \leqslant m$.

On the contrary, the inclusions $J^l(V,W) \subset J^m(V,W)$ are not invariant under a change of coordinates : if we want to set an inclusion function $i : J^l(V,W) \to J^m(V,W)$, we can not simply "complete" the l-jets with zeros. To see that, take for example $n = 1$, $q = 2$ and the function whose expresison in polar coordinates is $f : x \mapsto (x,x)_{r,\theta}$. In cartesian coordinates, we have $f : x \mapsto (x\cos(x), x\sin(x))_{u,v}$. Thus, $f_r''(x) = f_\theta''(x) = 0$ while $f_u''(x) = -2\sin(x) - \cos(x)$ and $f_v''(x) = 2\cos(x) - x\sin(x)$ for all x in $\mathbf{R}$. This illustrates the fact that prolongating a jet by 0 does not give the same thing depending on the coordinates we have chosen on W. This is also true for any choice that we could make to prolongate the l-jets into m-jets. Therefore, those inclusions are not invariant by a change of coordinates on W.

Jet extensions of functions

In the following, we will consider functions $f : V \to W$ as sections of the trivial fibration $p : V \times W \to V$. We will always suppose that it is $\mathcal{C}^\infty$ sections.

The notion of r-jets at points v of V gives rise to an another fibration, $p^r : V \to J^r(V, W)$, where $p^r := p \circ p_0^r$ associates to a r-jet the point of V in which it is defined. This fibration is endowed with a structure of smooth fibration thanks to the extensions of an atlas on the product manifold $V \times W$. In fact, as V and W are both endowed with a smooth atlas, so is the cartesian product $V \times W$. Then, we define an atlas on $J^r(V, W)$ by extending each coordinate chart $\phi : U \to \mathbf{R}^n\mathbf{R}^q$ on a coordinate chart $\phi^r : (p_0^r)^{-1}(U) \to J^r(\mathbf{R}^n, \mathbf{R}^q) \sim \mathbf{R}^{n+q(1+d(n,1)+...+d(n,r))}$, where ϕ^r associate a r-jet, or class of r-tangency, to its image by ϕ in $J^r(\mathbf{R}^n, \mathbf{R}^q)$. As we have defined a class of r-tangency as the inverse image of classes in $J^r(\mathbf{R}^n, \mathbf{R}^q)$ by a coordinate chart, it is clear that we build this way a smooth structure on the jets space.

This allows us to consider regular sections of the jet fibration and from now on, all the ones we consider are supposed to be $\mathcal{C}^\infty$. We note bs$F$ the image by p_0^r of F, in other words the section of $V \times W \to V$ induced by a section F of the jet fibration. Conversely, any section $f : V \to V \times W$ gives rise to a section $J_f^r : V \to J^r(V, W)$, which associates each point v in V to the r-class of tangency of f at v. It is called the *r-jet extension of f*.

All the jet sections are not r-jet extensions. Those which has this property, that is $F : V \to J^r(V, W)$ such that it exists $f : V \to V \times W$ with $F = J_f^r$ are called *holonomic sections.*

Example

Let us illustrate all these notions with an example. Take $n = 1$, $q = 2$, $V = \mathbf{R}$ and $W = \mathbf{R}^2$. The space of r-jets is $J^r(\mathbf{R}, \mathbf{R}^2) = \mathbf{R} \times \mathbf{R}^2 \times \mathbf{R}^{r\times 2\times 1}$. For any element $d = (v, a_1, a_2, b_1, b_2, c_1, c_2)$ of $J^2(\mathbf{R}, \mathbf{R}^2)$, the polynom $P : x \mapsto (a_1 + b_1(x - v) + \frac{1}{2}c_1(x - v)^2, a_2 + b_2(x - v) + \frac{1}{2}c_2(x - v)^2)$ around v represents the 2-tangency class of d.

For a function $f : \mathbf{R} \to \mathbf{R}^2$, we have $F := J_f^r(v) = (v, f(v), f'(v), ..., f^{(r)}(v))$. The section F is then a holonomic function, such that bs$F = f$. For other functions g, h, the map $v \mapsto (v, f(v), g(v), h(v))$ is also a section of $J^2(\mathbf{R}, \mathbf{R}^2)$ but is *a priori* not holonomic.

4.2.2 Differential relations

A lot of categories of functions that we use to manipulate in geometry are defined using differential equations, that is conditions on the derivatives of order 1 or more. This is for example the case of immersions, submersions, diffeomorphisms or symplectomorphisms. Using the notion of jets, we can give another view on these categories.

Definition 4.2.3 (Differential relation [8])**.** *A* differential relation of order r *between V and W is a subset of the r-jets space $J^r(V, W)$.*

Let us illustrate this notion with some example.

Example 4.2.1.

- **Immersions :** the differential relation $\mathcal{R}_{imm}$ associated to the notion of immersion is a subset of $J^1(V, W)$ since it only involves derivatives of order 1. It is exactly the set of monomorphisms $T_vV \to T_wW$ above each pair $(v, w) \in V \times W$.

- **Submersions :** the differential relation $\mathcal{R}_{sub}$ associated to the notion of submersion is also a subset of $J^1(V,W)$. It is exactly the set of epimorphisms $T_vV \to T_wW$ above each pair $(v,w) \in V \times W$.

- **Isoymplectomorphism :** isosymplectomorphisms between two symplectic manifolds (V,ω_V) and (W,ω_W) are functions $f : V \to W$ verifying $f^*\omega_W = \omega_V$. This equation involves the first order derivatives of f so it defines a subset of $J^1(V,W)$. We note the differential relation defined this way $\mathcal{R}_{isosymp}$.

  With $V = W = \mathbf{R}^{2n}$ endowed with the canonical symplectic structure, a section $f : V \to V \times W$ is a symplectomorphism if and only if ${}^t\partial_p f_q \partial_p f_p$ and ${}^t\partial_q f_q \partial_q f_p$ are symmetric and ${}^t\partial_q f_q \partial_p f_p - {}^t\partial_q f_p \partial_p f_q = I_k$. Therefore,

  $$\mathcal{R}_{isosymp} = \{(v,w,A) \in \mathbf{R}^n \times \mathbf{R}^n \times \mathbf{R}^{n^2} \quad | \quad {}^tA_3A_1, {}^tA_4A_2 \in S_n(\mathbf{R}) \wedge {}^tA_4A_1 - {}^tA_2A_3 = I_n\},$$

  where $A = \begin{pmatrix} A_1 & A_2 \\ A_3 & A_4 \end{pmatrix}$ with $A_1, A_2, A_3, A_4 \in \mathcal{M}_{l,l}(\mathbf{R})$.

- **Symplectomorphisms :** symplectomorphisms from V to a symplectic manifold (W,ω_W) are maps f such that $f^*\omega_W$ defines a symplectic form on V. Since ω_W is a symplectic form, it is closed and we have $d(f^*\omega_W) = f^*d\omega_W = f^*0 = 0$. The condition on f is then reduced to the fact that $f^*\omega_W$ is non-degenerate.

  In $V = W = \mathbf{R}^{2n}$ endowed with the canonical symplectic structure is symplectic if and only if ${}^t\nabla f(v)\mathbf{J}_{2n}\nabla f(v)$ is non-degenerate for all $v \in V$ so

  $$\mathcal{R}_{symp} = \left\{(v,w,A) \in \mathbf{R}^n \times \mathbf{R}^n \times \mathbf{R}^{n^2} \quad | \quad \det \begin{pmatrix} {}^tA_3A_1 - {}^tA_1A_3 & {}^tA_3A_2 - {}^tA_1A_4 \\ {}^tA_4A_1 - {}^tA_2A_3 & {}^tA_4A_2 - {}^tA_2A_4 \end{pmatrix} \neq 0 \right\}.$$

- More generally, any differential equation of order r induce a differnatial relation of order r.

Together with this definition comes the notion of *open* and *close* relations, which correspond to open and close subsets of the jet space. Immersion and submersions relations are open since they are defined as the complement of the close set composed of morphisms with at least a minor equal to zero. The relation associated with isosymplectomorphisms is closed since it is defined with an equality. On the contrary, the relation associated to symplectomorphisms is open as it is the complement of a closed subset. As usual, relations defined with equalities or large inequalities are closed while relations defined as complement of singularities or with strict inequalities are open.

Of course, smooth solutions of a differential equation, or inequality, of order r are such that their r-jet extension sends V in the induced differential relation. We now extend the notion of solution to all sections of the jet space.

Definition 4.2.4 (Formal and genuine solutions [8])**.** *A* formal solution *of a given differential relation $\mathcal{R}$ of order r is a section of the r-jet space which takes its values in $\mathcal{R}$, that is $F : V \to \mathcal{R}$. We denote by Sec $\mathcal{R}$ the subset of sections of the r-jet space composed of formal solutions of $\mathcal{R}$.*

A genuine solution *of $\mathcal{R}$ is a section $f : V \to V \times W$ whose r-jet extension is a formal solution. We denote by Hol $\mathcal{R}$ the subset of Sec $\mathcal{R}$ composed of holonomic formal solutions of $\mathcal{R}$.*

Example 4.2.2. For $V = \mathbf{R}^n$ and $W = \mathbf{R}^q$ with $n = q = 2l$ and endow these two spaces with the canonical symplectic structure. Consider the relation $\mathcal{R}_{isosymp}$ defined in the previous example and take any section $f : V \to V \times W$. The 1-jet section $F : v \mapsto (v, f(v), id) \in \mathbf{R}^n \times \mathbf{R}^n \times \mathbf{R}^{n^2}$ is a formal solution of $\mathcal{R}_{isosymp}$. Since it is not holonomic, $f = bsF$ is *a priori* not a genuine solution.

When we study differential equations, we look for genuine solutions. In some cases, it can be useful to first see if it exists formal solutions to the considered problem. If this is not the case, it is useless to search for genuine solutions. In the following section, we intoduce the homotopy-principle, which, if it holds, insure the existence of genuine solution from the existence of formal solutions.

4.2.3 Homotopy-principle

Definition 4.2.5 (Homotopy-principle [8]). *A differential relation $\mathcal{R}$ satisfies the homotopy-principle (or h-principle) if all formal solutions of $\mathcal{R}$ are homotopic in $\mathcal{R}$ to a holonomic formal solution.*

In other words, the h-principle holds for a relation $\mathcal{R}$ if any formal solution of the relation can be deformed in Sec $\mathcal{R}$ to the jet extension of a section $f : V \to V \times W$, which is therefore a genuine solution of $\mathcal{R}$.

There is different variations of this principle. Below is a list of some of them.

- **one parameter h-principle :** a differential relation $\mathcal{R}$ satisfies the one parameter h-principle if all homotopy F_t in Sec $\mathcal{R}$ joining two holonomic sections can be smoothly deformed in a homotopy in Hol $\mathcal{R}$ keeping F_0 and F_1 fixed.
- **multi-parameter h-principle :** a differential relation $\mathcal{R}$ satisfies the multi-parameter h-principle if all smooth family $F_T \subset$ Sec $\mathcal{R}$ such that $F_T \in$ Hol$\mathcal{R}$ for $T \in \partial I^k$ can be smoothly deformed in a family in Hol $\mathcal{R}$ keeping F_T fixed for all $T \in \partial I^k$. Here we have noted $I^k = [0, 1]^k$.
- **local h-principle :** let $A \subset V$ and $Op_V(A)$ an open neighbourhood of A in V. A differential relation $\mathcal{R}$ satisfies the local h-principle around A if all formal solutions of $\mathcal{R}$ defined above $Op_V(A)$ in V are homotopic in the sections of $\mathcal{R}$ defined above $Op_V(A)$ to a holonomic formal solution. In other words, the deformation should not change too much the space of definition of the original section but the holonomic section that we obtain are only defined above $Op_V(A)$.
- **relative h-principle :** let $B \subset V$ and $Op_V(B)$ an open neighbourhood of B in V. A differential relation $\mathcal{R}$ satisfies the relative h-principle around B if all formal solutions of $\mathcal{R}$ holonomic above $Op_V(B)$ in V are homotopic in the sections of $\mathcal{R}$ fixed on $Op_V(B)$ to a holonomic formal solution. In other words, we start from a formal solution which is already holonomic above an open set and we deforme it to obtain a holonomic solution on the whole V without changing the part which is already holonomic.
- **$\mathcal{C}^0$-dense h-principle :** a differential relation $\mathcal{R}$ satisfies the $\mathcal{C}^0$-dense h-principle if any formal solution F_0 of $\mathcal{R}$ is homotopic in $\mathcal{R}$ to a holonomic formal solution F_1 such that bsF_0 and bsF_1 are $\mathcal{C}^0$-close.

It is also possible to work with combinations of this versions such as the $\mathcal{C}^0$-close one parameter h-principle, where we ask that the deformation of the homotopy is $\mathcal{C}^0$-small, or the relative one parameter h-principle, where we ask that the deformation of the homotopy is fixed on B.

Proving that the h-principle holds for a given relation $\mathcal{R}$ can be sometimes difficult. In the following section, we present some tools that we can use to achieve it.

4.3 Proving the h-principle : tools and examples

4.3.1 Holonomic approximation

Below is the theorem on which is based all the results we will present in the following. The proof of this result can be found in [8].

Theorem 4.3.1 (Holonomic approximation [8]). *Let $A \in V$ a polyedron of codimension > 0 and $F : Op_V(A) \to J^r(V, W)$ a section of the jet space.*

For all $\delta, \epsilon > 0$, it exists a diffeotopy $(h^\tau)_{\tau\in[0,1]} : V \to V$ δ-small in the $\mathcal{C}^0$ sense and a holonomic section $\tilde{F} : Op_V(h^1(A)) \to J^r(V, W)$ such that $d(\tilde{F}(v), F(v)) < \epsilon$ for all $v \in Op_V(h^1(A))$.

This result also holds in its parametric and relative forms :

Theorem 4.3.2 (Parametric holonomic approximation [8]). *Let $A \in V$ a polyedron of codimension > 0 and $F_z : Op_V(A) \to J^r(V, W)$ a family of sections parametrized by $z \in I^k := [0, 1]^k$ with F_z holonomic for $z \in \partial I^k$.*

For all $\delta, \epsilon > 0$, it exists a family of diffeotopies $(h_z^\tau)_{\tau\in[0,1]} : V \to V$ δ-small in the $\mathcal{C}^0$ sense and a family of holonomic sections $\tilde{F}_z : Op_V(h_z^1(A)) \to J^r(V, W)$ such that $d(\tilde{F}_z(v), F_z(v)) < \epsilon$ for all $v \in Op_V(h_z^1(1))$ and all $z \in I^k$ and such that $h_z^\tau = id_V$ and $\tilde{F}_z = F_z$ for $z \in \partial I^k$.

Theorem 4.3.3 (Relative holonomic approximation [8]). *Let $A \in V$ a polyedron of codimension > 0 and $F : Op_V(A) \to J^r(V, W)$ a section of the jet space holonomic in a neighbourhood $Op_V\partial A$.*

For all $\delta, \epsilon > 0$, it exists a δ-small diffeotopy $(h^\tau)_{\tau\in[0,1]} : V \to V$ fixed on $Op_V\partial A$ and a holonomic section $\tilde{F} : Op_V(h^1(A)) \to J^r(V, W)$ such that $d(\tilde{F}(v), F(v)) < \epsilon$ for all $v \in Op_V(h^1(A))$ and $\tilde{F}(v) = F(v)$ on $Op_V(h^1(\partial A)) = Op_V(\partial A)$.

The relative form of the theorem is particularly useful when one wants to prove the existence of a holonomic section on a whole space divided into pieces that are treated one after another. If it is possible to prove the existence on a piece and to build the section on neighbouring pieces while sticking to what has already being done, then we can prove the existence on the whole space.

Here is a first interesting result which can be prove using holonomic approxiation.

Corollary 4.3.1 (Approximation of differential forms by closed forms). *Let V be an open manifold, A a polyedron of positive codimension, $a \in H^p(V)$ a cohomology class. Near A, we can approach in the $\mathcal{C}^0$ sense any p-form ω by a closed p-form $\tilde{\omega}$ in a.*

Moreover, given $\Omega \in a$ and a $(p-1)$-form α, we can chose $\tilde{\omega}$ of the form $d\tilde{\alpha} + \Omega$ for $\tilde{\alpha}$ $\mathcal{C}^0$-close to α.

Proof. (from [8])

Consider $X = \Lambda^p V$. It exists a map $D : (\Lambda^{p-1}V)^{(1)} \to \Lambda^p V$ which sends the formal derivatives of the coordinate functions of a $(p-1)$-form to a p-form. Consider its extension $\tilde{D} : \mathrm{Sec}(\Lambda^{p-1}V)^{(1)} \to \mathrm{Sec}\Lambda^p V$. Since we can choose what we want on $\mathrm{Sec}(\Lambda^{p-1}V)^{(1)}$, this map is surjective. Let ω be a p-form. It exists a section F_ω of $(\Lambda^{p-1}V)^{(1)}$ such that $\tilde{D} \circ F_\omega = \omega$. Since $\tilde{D}$ only cares about derivatives, we can choose F_ω such that $\mathrm{bs}F_\omega = \alpha$ for any $(p-1)$-form α.

For A a polyedron of positive codimension, we can apply Theorem 4.3.1 which gives us the existence of a diffeotopy $h^\tau : V \to V$ as small as we want and a holonomic section $\tilde{F}_\omega = J^1_{\tilde{\alpha}} : Op_V(h^1(A)) \to (\Lambda^{p-1}V)^{(1)}$ $\mathcal{C}^0$-close to F_ω. In particular, $\tilde{\alpha}$ is $\mathcal{C}^0$-close to α and, since D is continuous, $\tilde{w} := \tilde{D} \circ \tilde{F}_\omega$ is $\mathcal{C}^0$-close to ω. We also have $d\tilde{\alpha} = \tilde{D}(J^1_{\tilde{\alpha}}) = \tilde{\omega}$ so $\tilde{\omega}$ is exact.

Let now a be an arbitrarily cohomology class and $\Omega \in a$. Apply previous argument to $\theta = \omega - \Omega$ and take $\tilde{\omega} = \tilde{\theta} + \Omega$. It is $\mathcal{C}^0$-close to ω and can be written as $d\tilde{\alpha} + \Omega$. This shows that we can approach any p-form by a closed form of any cohomology class on $Op_V(h^1(A))$. $\square$

Note that the parametric version of this proposition is also true : we just have to apply Theorem 4.3.2 instead of 4.3.1.

4.3.2 Open Diff$_V$-invariant relations

We are now interested to a special category of differential relation, that we call *Diff$_V$-invariant*. Let $p : X \to V$ a fibration. All what we have done for $X = V \times W$ can immediately be extend to any X. In particular, we will be required to use $X = \Lambda^p V$ in some following propositions.

Denote by $\mathrm{Diff}_V X$ the group of diffeomorphisms of X which preserve the fibers, that is $h_X : X \to X$ such that it exists $h_V : V \to V$ satisfying $p \circ h_X = h_V \circ p$. If such a h_V exists, then it is obviously unique : for h^1_V and h^2_V satisfying the previous equation, we have $h^1_V \circ p = h^2_V \circ p$ in X. Since p is surjective, this means that $h^1_V = h^2_V$. Note that all of this can be extended to any fibration $p : E \to F$.

Definition 4.3.1 (Natural fibration [8])**.** *Let $p : E \to F$ a fibration and $\pi : Diff_F E \to Diff_F$ the homomorphism which associate a diffeomorphism $h_E \in Diff_F E$ to the unique diffeomorphism $h_F \in Diff_F$ such that $p \circ h_E = h_F \circ p$. If π can be inverted, that is if it exists $j : Diff_F \to Diff_F E$ such that $\pi \circ j = id_{Diff_F}$, then we say that the fibration $p : E \to F$ is* natural.

Note that j is not necessarily unique.

Example 4.3.1. The fibrations we are working with in this chapter are natural.

- the trivial fibration $p : V \times W \to V$ with $j : h \mapsto (h, id_W)$,
- The tangent bundle $p : TV \to V$ with $j : h \mapsto dh$,
- and the fibration of p-forms $p : \Lambda^p V \to V$ with $j : h \mapsto d^p h$, where $d^p h : (v, \omega) \in V \times \Lambda^p_v V \mapsto (h(v), \omega_h : (a_1, ..., a_p) \mapsto \omega(d_v h^{-1}(a_1), ..., d_v h^{-1}(a_p)))$.

The naturality of those fibrations can be extend to the jet fibration. In fact, we also have that $p^r_0 : X^{(r)} \to X$ is natural. Let $g : X \to X$ be a diffeomorphism and take

$$g^r : s \in X^{(r)} \to J^r_{g \circ \bar{s}}(p \circ g \circ p^r_0(s)),$$

where $\bar{s}$ is a local section of $p : X \to V$ whose r-jet coincide with s at $p^r(s)$. We immediately have that $p \circ p_0^r \circ g^r(s) = p \circ g \circ p_0^r(s)$ so $p_0^r \circ g^r(s)$ and $g \circ p_0^r(s)$ are in the same fiber of $X \to V$ for all $s \in X^{(r)}$. It is also obvious that $g^r(s)$ and $g \circ p_0^r(s)$ are in the same class of 0-tangency since $\bar{s}$ and $p_0^r(s)$, and so $g \circ \bar{s}$ and $g \circ p_0^r(s)$, are 0-tangent by construction of $\bar{s}$.

Then, if $p : X \to V$ is natural, we also have that $p^r : X^{(r)} \to V$ is natural. Under the assumption on p, for any diffeomorphism $h : V \to V$, it exists a fiber preserving diffeomorphism $j(h) = h_X : X \to X$. Thanks to the previous argument, we know that it exists $h^r : X^{(r)} \to X^{(r)}$ which preserves the fibers of $p_0^r : X^{(r)} \to X$. Therefore,

$$h \circ p^r = h \circ p \circ p_0^r = p \circ h_X \circ p_0^r = p \circ p_0^r \circ h^r = p^r \circ h^r.$$

Finally, let us see that $j^r(h) = (j(h))^r$ so

$$j^r(h)(s) = J^r_{j(h)\circ\bar{s}}(p \circ j(h) \circ p_0^r(s)) = J^r_{j(h)\circ\bar{s}}(h \circ p^r(s)).$$

In particular, $j^r(h)$ preserves holonomy : $j^r(h)(J^r_f) = J^r_{j(h)\circ f\circ h^{-1}}$.

Definition 4.3.2 (Diff$_V$-invariant differential relation [8])**.** *A differential relation is said to be* Diff$_V$-invariant *if it is invariant under the action $s \in X^{(r)} \to h_* s := j^r(h)(s)$ for all $h \in$ Diff$_V$.*

In other terms, a differential relation is Diff$_V$-invariant if it is invariant under coordinate changes. A lot of the relations that we will use in the following have this property.

Example 4.3.2.

- The relations $\mathcal{R}_{imm}$ and $\mathcal{R}_{subm}$ are Diff$_V$-invariant. More generally, any relation in $J^1(V, W)$ which imposes a condition on the rank of the differential is Diff$_V$-invariant. In fact, for a diffeomorphism $h : V \to V$,

$$h_*(x, y, A) = J^1_{(h\times id)(x,\bar{y},A)}(h(x)) = (h(x), y, A \circ d_x h).$$

  Since h is a diffeomorphism, its differential at any point is invertible so its composition with the homomorphism A has the same rank as A.

- The same formula shows that all differential relation which would impose conditions on the image of the differential is also Diff$_V$-invariant. This will be useful when we will talk about Grassmanians.

The notion of Diff$_V$-invariance is particularly interesting thanks to the following theorem.

Theorem 4.3.4. *Let V, W be two manifolds and $\mathcal{R} \subset J^1(V, W)$ an open Diff$_V$-invariant differential relation. All the local forms of the h-principle holds for $\mathcal{R}$.*

Proof. (from [8])

We first prove the theorem in the 1-parameter case. Let F_0 and F_1 two holonomic solutions of $\mathcal{R}$ and F_t a homotopy between them in $\mathcal{R}$.

Let $A \subset V$ be a polyedron of positive codimension. For all $\delta, \epsilon > 0$, Theorem 4.3.2 insures the existence of a family of δ-small diffeotopies $h_t^\tau : V \to V$ and a family of holonomic sections $\tilde{F}_t : Op_V(h_t^1(A)) \to J^r(V, W)$ such that $\tilde{F}_t$ and F_t are $\mathcal{C}^0$-close for all $t \in [0, 1]$, $\tilde{F}_0 = F_0$, $\tilde{F}_1 = F_1$ on $Op_V(h_t^1(A))$ and $h_0^\tau = h_1^\tau = h_t^0 = id_V$ for all $\tau, t \in [0, 1]$.

Since F_t is a homotopy in the open set $\mathcal{R}$, we can choose ϵ such that $\tilde{F}_t$, which is ϵ-close to F_t, is also in $\mathcal{R}$ for all $t \in [0,1]$ on $Op_V(h_t^1(A))$ and the linear homotopy $\hat{F}_t^\nu = \nu\tilde{F}_t + (1-\nu)F_t$ as well. Now, since $\mathcal{R}$ is Diff$_V$-invariant, $\bar{F}_t^{\nu,\tau} := (h_t^\tau)_*^{-1}\hat{F}_t^\nu$ also has its image on $\mathcal{R}$. Its satisfies $\bar{F}_t^{1,1} = (h_t^1)_*^{-1}\tilde{F}_t$, $\bar{F}_t^{0,0} = F_t$, $\bar{F}_0^{\nu,\tau} = F_0$ and $\bar{F}_1^{\nu,\tau} = F_1$ for all $t, \tau, \nu \in [0,1]$. Then, the family of homotopies given by $G_t^\eta = \bar{F}_t^{2\eta,0}$ on $\eta \in [0, \frac{1}{2}]$ and $G_t^\eta = \bar{F}_t^{1,2\eta-1}$ on $\eta \in [\frac{1}{2}, 1]$ goes from F_t to $G_t^1 = (h_t^1)_*^{-1}\tilde{F}_t$ while staying fixed at $t = 0, 1$. It is defined on $Op_V(A)$ and takes its values in $\mathcal{R}$. Since the action of $(h_t^1)^{-1}$ preserves the holonomy, G_1 is holonomic. Morevover, since h_t^τ can be chosen arbitrarily small and $\tilde{F}_t$ arbitrarily close to F_t, we have that bsG_t^1 is arbitrarily $\mathcal{C}^0$-close to bsF_t.

Therefore, we have established the local $\mathcal{C}^0$-dense 1-parameter h-principle. Note that the above argument also works when t is multivalued, one just has to change the notations. Applying Theorem 4.3.1 instead of 4.3.2 and skiping mentions of t, we also get the result for the simple local h-principle. In the same way, all what we have done is still true in the relative case, after the application of Theorem 4.3.3 instead of 4.3.2 : if the diffeotopies h_t^τ fixes $Op(\partial A)$ and the homotopy $\tilde{F}_t$ coincides with F_t on $Op(\partial A)$, then G_t^1 also coincides with F_t on $Op(\partial A)$. ☐

We now extend this result to the global h-principle. For that, we compress the whole space V to the open set on which are defined the objects we are interested in after application of the previous theorem. To make that this compression is possible, we have to assume that V is an open manifold.

Theorem 4.3.5. *Let V be an open manifold and $\mathcal{R} \subset J^1(V,W)$ an open Diff$_V$-invariant differential relation. All the global forms of the h-principle holds for $\mathcal{R}$, except the $\mathcal{C}^0$-dense and the relative one.*

Nevertheless, if V can be retracted into a polyedron of positive codimension K, then the $\mathcal{C}^0$-closeness is still true in a neighbourhood of K and the relative version of the h-principle holds by respect to K.

Proof. (from [8])

As V is open, it exists a polyedron K of positive codimension such that V can be retracted in an arbitrarily small neighbourhood $Op_V(K)$ via a diffeotopy h^τ such that $h^0 = id_V$, $h^1(V) \subset Op_V(K)$ and $h^\tau = id_V$ on K. The polyedron K is then called a *core* of V. For a proof of this result, see for instance [8].

We start by the 1-parameter version of the h-principle. Let F_0 and F_1 two holonomic section of $\mathcal{R}$ and F_t a homotopy joining them in $\mathcal{R}$. From Theorem 4.3.4, it exists a family of homotopy $\tilde{F}_t^\tau : Op_V(K) \to \mathcal{R}$ such that $\tilde{F}_t^1$ is holonomic, $\tilde{F}_t^0 = F_t$, $\tilde{F}_0^\tau = F_0$ and $\tilde{F}_1^\tau = F_1$ for all $\tau, t \in [0,1]$ on $Op_V(K)$. Moreover, we can choose it in such a way that $\tilde{F}_t^\tau$ is ϵ-close to F_t for all $\tau, t \in [0,1]$.

Define the family of homotopies G_t^τ such that $G_t^\tau = (h^{2\tau})_*^{-1}F_t$ for $\tau \in [0, \frac{1}{2}]$ and $G_t^\tau = (h^1)_*^{-1}\tilde{F}_t^{2\tau-1}$ for $\tau \in [0, \frac{1}{2}]$. It is defined on V and since $\mathcal{R}$ is Diff$_V$-invariant, it takes its values in $\mathcal{R}$.

As in Theorem 4.3.4, changing $t \in [0,1]$ for a multi-valued parameter only changes notations. For the simple h-principle, one just has to use the simple version of Theorem 4.3.4 and forget the indice t in the passage from local to global.

For the relative version, if B is a core of V, then the retraction h^τ can be chosen fixed on $Op_V(B)$. Since the use of the relative version of Theorem 4.3.4 gives homotopies fixed on $Op_V(B)$, the resulting homotopies are fixed on $Op_V(B)$ too.

The decompression of $Op_V(K)$ into V makes that the $\mathcal{C}^0$-closeness does not hold anymore in the global case. However, since this decompression is fixed on K, we still have that G_t^1 is $\mathcal{C}^0$-close to F_t in a small neighbourhood of K. $\square$

Here is a basic example where having proved the h-principle gives the existence of homotopies which are really difficult to visualize.

Example 4.3.3. Let V be the annulus $\{(x, y) \in \mathbf{R}^2 \mid \epsilon < x^2 + y^2 < a\}$ and $W = \mathbf{R}$. The 1-jet space is $J^1(V, W) = V \times \mathbf{R} \times \mathbf{R}^2$. Let $f_0 : (x, y) \mapsto x^2 + y^2$ and $f_1 = -f_0$. The two functions f_0 and f_1 are isotopic.

Since V is open and the relation $\mathcal{R}_{imm}$ is open Diff_V-invariant, by Theorem 4.3.5 it is sufficient to find a formal solution of $\mathcal{R}_{imm}$ linking f_0 and f_1. Identifying $\mathbf{R}^2$ with $\mathbf{C}$, we can take $F_t : (x, y) \mapsto (x, y, tf_1(x, y) + (1 - t)f_0(x, y), e^{i\pi t}\nabla f_0)$.

Using the same proof as for Theorem 4.3.4 followed by 4.3.5 but replacing the invocation of holonomic approximation theorem by Corollary 4.3.1, we obtain :

Proposition 4.3.1. *Let V be an open manifold, $a \in H^p(V)$ a cohomology class and $\mathcal{R} \subset \Lambda^p V$ an open Diff_V-invariant differential relation.*

Any p-form $\omega : V \to \mathcal{R}$ is homotopic in $\mathcal{R}$ to a closed p-form in a.

Any homotopy of p-forms $\omega_t : V \to \mathcal{R}$ between two closed forms ω_0 and ω_1 in a can be deformed in $\mathcal{R}$ to a homotopy of closed forms in a between ω_0 and ω_1 without changing the ends ω_0 and ω_1.

Proof. (from [8])

The proof is exactly the same as the one of the Theorem 4.3.4 followed by the one of 4.3.5, excepting that we replace the invocation of the holonomic approximation theorem by Corollary 4.3.1 and that we also use the fact that the cohomology class of $h^\tau \omega$ is constant for any isotopy h^τ and any p-form ω. $\square$

We present below an application of the h-principle for open Diff_V-invariant relations.

4.3.3 Application of the h-principle to the Grassmanian bundle

Let W be a q-dimensional manifold and V a n-dimensional submanifold embedded in W. For $l \leqslant q$, we note $Gr_l W$ the *Grassmanian* of W, which is defined as the set of all the vectorial spaces of dimension l tangent at W. Denotes by $\pi : Gr_l W \to W$ the Grassmanian bundle of the manifold W. The projection π associates an element of $Gr_l W$ to the point $w \in W$ at which this element is tangent to W.

At each point $v \in V$, the differential of the embedding $dp : TV \to TW$ sends the tangent spaces of v to a vectorial subspace of dimension n tangent to W. We denote $Gdf : V \to Gr_n W$ the map which associates $v \in V$ to $df(T_v V) \subset T_{p(v)} W$. More generally, if $F : TV \to TW$ is a monomorphism, we can define in the same way the map $GF : V \to Gr_n W$.

For $A \subset Gr_n W$, a homomorphism $F : TV \to TW$ is said to be A-*directed* if $GF(V) \subset A$.

Theorem 4.3.6. *Let A be an open subset of $Gr_n W$, V an open manifold, $f_0 : V \to W$ an immersion such that $F_0 := df_0$ is homotopic in the space of monomorphisms from TV to TW to a certain A-directed F_1 with $bsF_1 = f_0$.*

Then, f_0 is isotopic to f_1 such that $Gdf_1(V) \in A$ and df_1 is homotopic to F_1 in the space of A-directed monomorphisms.

Moreover, the isotopy can be chosen as $\mathcal{C}^0$-small as we want on an open neighbourhood of a core K.

Proof. (from [8])

To the open subset A in $Gr_n W$, we associate an open differential relation,

$$\mathcal{R}_A := \{F \in J^1(V,W) \quad | \quad F \text{ monomorphism and } GF(V) \subset A\}.$$

From a previous example, it is open and Diff_V-invariant. Theorem 4.3.5 in its simple version gives the existence of a homotopy $\hat{F}_t : V \to \mathcal{R}_A$ between F_1 and a holonomic A-directed monomorphism F_2 that we can choose such that $\mathrm{bs}F_1 = f_0$ and $f_1 := \mathrm{bs}F_2$ are $\mathcal{C}^0$-close on a small neighbourhood of K in V.

Taking the homotopy between F_0 and F_1 followed by $\hat{F}_t$, we obtain a homotopy $\bar{F}_t$ of monomorphisms between two holonomic sections. Then, Theorem 4.3.5 in its parametric version gives the existence of a holonomic homotopy $\tilde{F}_t : V \to \mathcal{R}_{imm}$ between F_0 and F_2 with $\mathrm{bs}\tilde{F}_t$ $\mathcal{C}^0$-close to $\mathrm{bs}\bar{F}_t$ near K.

In particular, $f_0 = \mathrm{bs}F_0$ is isotopic to $f_1 = \mathrm{bs}F_2$, where f_1 is A-directed and the isotopy small around K, and F_1 is homotopic to $df_1 = F_2$ via a homotopy of A-directed monomorphism. □

4.4 Application to the symplectic relation

In this section, we present an application of what we have presented in the symplectic case. More precisely, we have the following theorem, whose proof is a detailed version of the one proposed in [8].

Theorem 4.4.1. *Let (V, ω_V) and (W, ω_W) be two symplectic manifolds of respective dimensions $n = 2l$ and $q = 2m$. Let $f_0 : V \to W$ be an embedding such that $f_0^*[\omega_W] = [\omega_V]$. Suppose also that $F_0 = df_0$ is homotopic to an isosymplectic homomorphism F_1 via $F_t \subset \mathcal{R}_{imm}$ such that $bsF_t = f_0$ for all $t \in [0,1]$.*

Then, if V is open and $m < l$, then it exists an isotopy $f_t : V \to W$ such that f_1 is isosymplectic and df_1 is homotopic to F_1 in the isosymplectic homomorphisms.

Moreover, if K is a core of V, we can choose f_t arbitrarily $\mathcal{C}^0$-close to f_0 near K.

Proof. (from [8])

The proof is made in three steps.

Step 1 : as the relation associated to the fact of being isosymplectic is not open, we first consider the relation associated to the fact of being symplectic, which is open. Let A_{symp} the associated subset of $Gr_n W$.

By the Theorem 4.3.6, it exists an isotopy $\tilde{f}_t : V \to W$ such that $\tilde{f}_0 = f_0$, $\tilde{f}_1$ is symplectic and $\tilde{f}_t$ is $\mathcal{C}^0$-close to f_0 on K for all $t \in [0,1]$. Moreover, $d\tilde{f}_1$ and F_1 are homotopic via Φ_t such that $G\Phi_t(V) \subset A_{symp}$.

Since the cohomology class is invariant by homotopy, the assumption that $f_0^*[\omega_W] = [\omega_V]$ implies that $\tilde{f}_1^*[\omega_W] = [\omega_V]$. Then, by Theorem 4.3.1 it exists a homotopy of symplectic forms ω_t between $\tilde{f}_1^*\omega_W$ and ω_V such that $[\omega_t]$ is constant on $[0,1]$. This allows us to write $\omega_t = \omega_0 + d\alpha_t$ for $t \in [0,1]$.

The proof of the theorem is now reduced to the proof of the following proposition :

Proposition 4.4.1. *Let V a symplectic manifold of dimension $n = 2m$, (W, ω_W) a symplectic manifold of dimension $q = 2l > n$, $h_0 : V \to W$ a symplectic embedding, $\omega_0 = h_0^*\omega_W$ and $\omega_t = \omega_0 + d\alpha_t$ a homotopy of symplectic forms.*

It exists a symplectic isotopy $h_t : V \to W$ as $\mathcal{C}^0$-small as we want such that $h_1^*\omega_W = \omega_1$.

If we apply this proposition to $h_0 = \tilde{f}_1$, we can take the homotopy f_t given by $\tilde{f}_{2t}$ on $[0, \frac{1}{2}]$ and h_{2t-1} on $[\frac{1}{2}, 1]$. It is an isotopy as its two parts are isotopies and it verifies that $f_1 = h_1$ is isosymplectic. As $\tilde{f}_t$ is $\mathcal{C}^0$-small on K and h_t $\mathcal{C}^0$-small on the whole V, f_t is $\mathcal{C}^0$-small on K. Finally, since the space of isosymplectic homomorphisms is convex, the linear homotopy $F_t = tdf_1 + (1-t)F_1$ realizes the desired homotopy between F_1 and df_1.

Step 2 : we prove the proposition in the case where $w_t = \omega_0 + tdr \wedge ds$ for $r, s : V \to W$ borned.

From the symplectic neighbourhood theorem, it exists $\epsilon > 0$ such that $h_0 : V \to W$ can be extended in an isosymplectic embedding $\hat{h}_0 : (E, \omega_E) \to (W, \omega_W)$, where $E := V \times D_\epsilon^2 \times D_\epsilon^{q-n-2}$ and $\omega_E = \omega_0 \oplus \eta_2 \oplus \eta_{q-n-2}$, for (D_ϵ^k, η_k) the ball of radius ϵ in $\mathbf{R}^k$ endowed with the restriction to D_ϵ^k of the standard symplectic form of $\mathbf{R}^k$.

Consider $\phi = (r, s) : V \to \mathbf{R}^2$. As r and s are supposed to be borned, it exists $R > 0$ such that $\phi(V) \subset D_R^2$. Let $\tau_{R,\epsilon} : D_R^2 \to D_\epsilon^2$ be an area preserving map and set $\psi = \tau_{R,\epsilon} \circ \phi$. We then have that

$$(t\psi)^*\eta_2 = (t\phi)^*\tau_{R,\epsilon}^*\eta_2 = (t\phi)^*\eta_2 = t^2 dr \wedge ds.$$

Consider now $\Phi_t : V \to E$ such that $\Phi_t(v) = (v, \sqrt{t}\psi(v), 0)$. We have

$$\Phi_t^*\omega_E = \Phi_t^*(\omega_0 \oplus \eta_2 \oplus \eta_{q-n-2}) = id^*\omega_0 + \sqrt{t}\psi^*\eta_2 + 0^*\eta_{q-n-2} = \omega_0 + tdr \wedge ds.$$

Now, if we take $h_t : v \mapsto \hat{h}_0(\Phi_t(v))$, we have

$$h_t^*\omega_W = \Phi_y^*\hat{h}_0^*\omega_W = \Phi_t^*\omega_E = \omega_t$$

and $\hat{h}_0(\Phi_0(v)) = \hat{h}_0(v, 0, 0) = h_0(v)$ with $(h_t)_t$ ϵ-small in the $\mathcal{C}^0$ sense.

Step 3 : We now reduce the general case, where $\omega_t = \omega_0 + d\alpha_t$ to the case where $\omega_t = \omega_0 + tdr \wedge ds$, that is, the case where $\alpha_t = trds$. For that, note that we can consider any homotopy α_t provided that it agrees with the original one at $t = 0, 1$ and that the resulting ω_t remains symplectic.

We first consider $\hat{\alpha}_t$ linear by parts, build this way :

$$\begin{cases} \hat{\alpha}_{t_i} = \alpha_{t_i} \text{ for some points } t_i \in [0,1], \\ \hat{\alpha}_{t_i+\tau} = \alpha_{t_i} + \dfrac{\tau}{t_{i+1} - t_i}(\alpha_{t_{i+1}} - \alpha_{t_i}) \text{ for } \tau \in [0, t_{i+1} - t_i], \end{cases}$$

where the finite set $\{t_i\}_i$ contains 0 and 1 and is chosen such that $\hat{\omega}_t = \omega_0 + d\hat{\alpha}_t$ is symplectic for all $t \in [0,1]$.

We can restrict our analysis to an interval of the form $[t_i, t_{i+1}]$, where ω_t is of the form $\omega_0 + t\alpha$. In fact, applying the proposition on each interval, we obtain a finite number of isotopies h_i that we have to take one after another to obtain an isotopy on the whole $[0,1]$.

Suppose now that we have a decomposition of this kind : $\alpha = \beta^1 + ... + \beta^L$ with $L \in \mathbf{N}$ and the β^i of the form $r_i ds_i$, where the r_i and s_i are borned. Consider β_t linear by parts such that

$$\begin{cases} \beta_0 = 0, \\ \beta_{s_j} = \beta^1 + ... + \beta^j \text{ for } s_j = \dfrac{j}{L}, j \in [\![1, L]\!], \\ \beta_{s_j+\tau} = \beta_{s_j} + L\tau\beta^{j+1} \text{ for } \tau \in [0, \dfrac{1}{L}] \end{cases}$$

and $\tilde{\alpha}_t$ such that

$$\begin{cases} \tilde{\alpha}_{t_i} = t_i\alpha \text{ for } t_i = \dfrac{i}{N}, j \in [\![0, N]\!], \\ \tilde{\alpha}_{t_i+\tau} = t_i\alpha + \dfrac{1}{N}\beta_{N\tau} \text{ for } \tau \in [0, \dfrac{1}{N}], \end{cases}$$

where $N \in \mathbf{N}$ can be chosen arbitrarily large. On each subinterval $\left[\frac{i}{N+\frac{j}{NL}}, \frac{i}{N+\frac{j}{NL}}\right]$, $\tilde{\omega} = \omega_0 + d\tilde{\alpha}$ is of the form $\left(\omega_{t_i} + \frac{1}{N}\beta_{s_j}\right) + \tau\beta^{j+1}$ so is linear by parts. With the same argument than previously, we can restrict the problem to the case where ω_t is of the form $\omega_0 + td(rds) = \omega_0 + tdr \wedge ds$ with r and s borned, which is the case where the proposition is already proved.

It now remains to show how we obtain the decomposition of the 1-form α. We suppose that V is compact : if it not the case, we consider a compact extension. Let $(\rho_i)_{i\in I}$ be a partition of the unity subordinate to an atlas $(U_i, \phi_i)_{i\in I}$ of V, that is such that $\sum_{i\in I}\rho_i = 1$ and $\mathrm{supp}\rho_i$ is compact and included in U_i for every i. For every $i \in I$, we also consider χ_i such that $\chi_i = 1$ on $\mathrm{supp}\rho_i$ and $\mathrm{supp}\chi_i \subset U_i$.

Choose a coordinate system $\{x_1, ..., x_n\}$ on $\phi_i(U_i) \subset \mathbf{R}^n$ and write $(\phi_i)_*\alpha|_{U_i} = a_1^i dx_1^i + ... + a_n^i dx_n^i$. Then,

$$\alpha_{U_i} = \phi_i^*(a_1^i dx_1^i + ... + a_n^i dx_n^i) = (a_1^i \circ \phi_i)\phi_i^* dx_1^i + ... = (a_1^i \circ \phi_i) d(x_1^i \circ \phi_i) + ...$$

Now, see that

$$\alpha = (\sum_{i\in I}\rho_i)\alpha = \sum_{i\in I}\rho_i\alpha|_{U_i} = \sum_{i\in I}\rho_i\sum_{j=1}^{n}(a_j^i \circ \phi_i)d(x_j^i \circ \phi_i) = \sum_{i,j}\rho_i(a_j^i \circ \phi_i)d\big(\chi_i(x_j^i \circ \phi_i)\big).$$

If we set $r^{ni+j} = \rho_i(a_j^i \circ \phi_i)$ and $s^{ni+j} = \chi_i(x_j^i \circ \phi_i)$, which are compactly supported, we obtain $\alpha = \sum_l r^l ds^l$, which is the desired form. This achieves at the same time the proofs of the proposition and the theorem. □

Remark 4.4.1. In the proof, we have used the fact that the cohomology class remains constant when we pull back a form by a homotopy. Let us show this. Let ω a closed p-form on W and $f_t : V \to W$ a homotopy. Saying that the cohomology class of $f_t^*\omega$ is constant in $H^k(V) = (H_k(V))^{-1}$ is equivalent to say that $f_t^*\omega$ takes fixed values on a base of $H_k(V)$. Let $\sigma : \Delta_k \to V$ be a cycle. We have that

$$\int_{\Delta_k}\sigma^*(f_t^*\omega) = \int_{\Delta_k}(f_t \circ \sigma)^*\omega.$$

Let $\phi : \Delta_k \times [0,1] \to V$ be such that $\phi(x,t) = (f_t \circ \sigma)(x)$. Then, using Stokes theorem, we have

$$\int_{\Delta_k\times[0,1]} d(\phi^*\omega) = \int_{\Delta\times\{1\}}\phi^*\omega - \int_{\Delta\times\{0\}}\phi^*\omega.$$

Since ω is closed, the left hand side of the equation is equal to zero and $f_0^*\omega$ and $f_1^*\omega$ takes the same value on the cycle σ. As this cycle can be arbitrarily chosen, we have proved that the cohomology class of $f_0^*\omega$ and $f_1^*\omega$ is the same.

5 Generating functions

In this chapter, we present a summary of what we learned about generating functions in [1] and [15]. We expect to use these notions in futur reduction methods for Hamiltonian systems. In fact, as we will see, generating functions are useful tools to build symplectomorphisms $f : V \to V$ on a given symplectic manifold (V, ω). With a generating function $S : V \to \mathbf{R}$, we can caracterize f. When one wants to learn a Hamiltonian dynamic, this caracterization may be interesting. Here, we take $V = \mathbb{R}^{2n}$ and we use coordinates (q, p), with $q, p \in \mathbb{R}^n$. We endow $\mathbb{R}^{2n}$ with the usual smyplectic structure, given by $\omega = d\lambda$ with $\lambda = pdq$.

We are here interested in isosymplectic maps, that is $f : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ such that $f^*\omega = \omega$.

In particular, Hamiltonian flows are isosymplectic transformations :

$$(\phi_H^t)^*\omega = \omega.$$

To show it, first note that this equation is equivalent to $L_{X_H}\omega = 0$. Then, use Cartan's formula :

$$L_z\alpha = \iota_z\alpha + d(\iota_z\alpha),$$

which is true for all p-form α and all vector field z in $\mathbf{R}^{2n}$. In this formula, $L_z\alpha$ represents the Lie derivative of the form α in the direction z and $\iota_z\alpha$ the interior product between z and α. This immediately gives

$$\iota_{X_H}d\omega + d(\iota_{X_H}\omega) = d(dH) = 0.$$

It is obvious that if f is isosymplectic, then the form $\lambda - f^*\lambda$ is closed. In fact, it it even exact: it exists $S : \mathbb{R}^{2n} \to \mathbb{R}$ such that

$$pdq(p, q) + P(p, q)dQ(p, q) = dS(p, q). \tag{5.1}$$

If we assume that the coordinates (q, Q) are independent, we can express S in this coordinate system. Note

$$S_1(q, Q(p, q)) = S(p, q).$$

We have

$$p = \frac{\partial S_1}{\partial q}(q, Q) \quad \text{and} \quad P = -\frac{\partial S_1}{\partial Q}(q, Q). \tag{5.2}$$

Conversely, if a function $S_1 : \mathbb{R}^N \times \mathbb{R}^N \to \mathbb{R}$ verifies $\det \frac{\partial^2 S_1}{\partial Q \partial q} \neq 0$, then the implicite function theorem applied to $\frac{\partial S_1}{\partial q}$ tells that we can express Q in terms of $p := \frac{\partial S_1}{\partial q}$ and q. If we set $P_1(q, Q) = \frac{\partial S_1}{\partial Q}(q, Q)$ and $P(p, q) = P_1(q, Q(p, q))$, we obtain an isosymplectic transformation $g : (p, q) \mapsto (P, Q)$. In fact, it verifies equation (5.1) so if we apply the exterior derivative and use the fact that $ddS = 0$, we have $g^*\omega = \omega$. The map f is such that p and P satisfy (5.2). We then say that S_1 is the *generating function of f*.

Note that we obtained a cananical transformation from a single map from $\mathbb{R}^{2n}$ to $\mathbb{R}$. Moreover, every isosymplectic transformation which verify the independence condition between q and Q can be obtained from a generating function.

It can happen that q and Q are not independent : this is for example the case in the identity function. This do not mean that previous computations can not be done anymore. We can apply the same argument with the coordinates q and P instead of q and Q. We then have

$$p = \frac{\partial S_1}{\partial q}(q, P) \quad \text{and} \quad Q = \frac{\partial S_1}{\partial P}(q, P).$$

For example, a generating function for the identity function is given by $S : (q, P) \mapsto Pq$.

Actually, we can choose any partition $(i_1, ..., i_k), (j_1, ..., j_m)$ de $(1, ..., N)$ such that

$$\det \frac{\partial^2 S_1}{\partial(P_j, Q_i)\partial q} \neq 0.$$

For isosymplectic transformations close to the identity, we can choose generating functions of the form

$$S(P, q) = Pq + \epsilon \bar{S}(P, q, \epsilon).$$

We then have

$$p = P - \epsilon \frac{\partial \bar{S}}{\partial q} \quad \text{and} \quad Q = q + \epsilon \frac{\partial \bar{S}}{\partial P}$$

so if we set $H : (p, q) \mapsto S(p, q, 0)$, we have

$$\left.\frac{dP}{d\epsilon}\right|_{\epsilon=0} = -\frac{\partial H}{\partial q} \quad \text{and} \quad \left.\frac{dQ}{d\epsilon}\right|_{\epsilon=0} = \frac{\partial H}{\partial q}.$$

Conclusion

The internship was in line with last year internship, where we explored different linear reduction methods for Hamiltonian problems, among which the PSD. The aim was to explore different methods to improve the reduction given by the PSD in non-linear cases and established theoretical justifications for the new methods. Two approaches have been taken : quadratic corrections of the decoder and hyperreduction via optimal control. The different variations of the first one which have been tested have given mediocre results, the trajectories computed in low dimension with the corrected decoder did not differ significantly from the trajectories induced by the PSD decoder. On the contrary, tests conducted within hyperreduction via control approach have given promising results. In particular, we have presented a variation of the gradient descent which appeared to be very efficient on simple cases. We are still carrying out additional tests in order to explain it and adapat the method to more complex cases. In the geometrical part of the internship, we have continued to read in order to being familiar with some geometrical tools. We have learned about generating functions, which we hope to use to build new symplectic decoders, and h-principle, which we want to use to justify the symplectic reduction approach.

To conclude, I think that the numerical objectives of internship were partially reached. We have explored the quadratic correction approach but we put it aside due to non-satisfying results. The hyperreduction approach gived results but we still are working on it and testing the methods we presented. On the other hand, the geometrical objectives were reached since we are now ready to start working on the conjecture that we want to prove. I could surely have been expected to code more quickly, especially the quadratic corrections, and this is certainly the reason why the numerical part is less developed than originally planed.

During this internship, I worked on some skills I acquired during the two years of Masters. From the programming point of view, I used Python to implement methods I learned and this gave me the opportunity to practise this language. From the numerical analysis point of view, I enriched my knowledge about reduced order models, that we have seen in class in the case of finite elements and that I have already seen in last year internship. I also used the theoretical tool we learned in the optimal control lesson, namely the adjoint method, that I had to detail in a case a little more difficult than the ones we have seen in class. I also worked on my English skills as almost all the references I had to read were in English. I also developed new skills during this internship : I discovered the h-principle and some techniques of proof very interesting. In the field of "soft" skills, I often had to take a step back and think about the global mechanisms rather than the technical details, but without losing sight of the geometric rigour. Because I had a geometrician and two numerical analysts for supervisors, I sometimes had to switch from a point of view to another to understand what there were saying about the same subject. I found these exercises difficult and I know that I have a lot of room for improvement.

Bibliography

[1] V.I. Arnold. *Mathematical Methods of Classical Mechanics*, chapter 7, 8, 9, 10, pages 163–200. Springer, 1989.

[2] S. L. Brunton, J. L. Proctor, and J. N. Kutz. Discovering governing equations from data by sparse identification of non linear dynamical systems. *Proceedings of the National Academy of Sciences*, 113(5):3932–3937, 2016.

[3] L. Buhovsky and E. Opshtein. Quantitative h-principle in symplectic geometry. *Journal of Fixed Point Theory and Applications*, 24(2):38, 2022.

[4] J. Chabassier and P. Joly. Energy preserving schemes for nonlinear hamiltonian systems of wave equations: Application to the vibrating piano string. *Computer Methods in Applied Mechanics and Engineering*, 199(45-48):2779–2795, 2010.

[5] R. Chen and M. Tao. Data-driven prediction of general hamiltonian dynamics via learning exactly-symplectic maps. *CoRR*, abs/2103.05632, 2021.

[6] T. Q. Chen, Y. Rubanova, J. Bettencourt, and D. Duvenaud. Neural ordinary differential equations. *CoRR*, abs/1806.07366, 2018.

[7] Direction de la communication du CNRS. L'histoire du CNRS. url: https://histoire.cnrs.fr, march 2018. Accessed: 2023-07-20.

[8] Y. Eliashberg and N.M. Mishachev. *Introduction to the h-Principle*, volume 48 of *Graduate studies in mathematics*. American Mathematical Society, 2002.

[9] S. Gallot, D. Hulin, and J. Lafontaine. *Riemannian Geometry*, chapter 1,2. 2004.

[10] R. Geelen, S. Wright, and K. Willcox. Operator inference for non-intrusive model reduction with quadratic manifolds. *Computer Methods in Applied Mechanics and Engineering*, 403, 2023.

[11] E. Hairer, C. Lubich, and G. Wanner. Geometric numerical integration illustrated by the störmer-verlet method. *Acta Numerica*, pages 399–450, 2003.

[12] INRIA. Inria et son écosystème. url: https://www.inria.fr/fr/innovation-numerique-ecosysteme, november 2020. Accessed: 2022-08-20.

[13] IRMA. L'institut - Présentation. url: https://irma.math.unistra.fr/linstitut/presentation.html. Accessed: 2023-07-20.

[14] P. Jin, Z. Zhang, A. Zhu, Y. Tang, and G.E. Karniadakis. Sympnets: Intrinsic structure-preserving symplectic networks for identifying hamiltonian systems, 2020.

[15] D. McDuff and D. Salamon. *Introduction to symplectic topology.* Oxford University Press, 1998.

[16] J. Moser. On quadratic symplectic mappings. *Mathematische Zeitschrift*, 216:417–430, 1994.

[17] L. Peng and K. Mohseni. Symplectic model reduction of hamiltonian systems. Preprint at https://arxiv.org/abs/1407.6118v2, 2015.

[18] Y.D. Zhong, B. Dey, and A. Chakraborty. Symplectic ode-net: Learning hamiltonian dynamics with control. *CoRR*, abs/1909.12077, 2019.

Appendices