UNIVERSIDADE TÉCNICA DE LISBOA

INSTITUTO SUPERIOR TÉCNICO

KAM Theory, Presymplectic Dynamics and Lie algebroids

Hassan Najafi Alishah

Supervisor: Doctor Rui António Loja Fernandes

Thesis approved in public session to obtain the PhD Degree in

Mathematics

Jury final classification: Pass with Merit

Jury

Chairperson: Chairman of the IST Scientific Board

Members of the Committee:

Doctor Nguyen Tien Zung

Doctor Rafael de la Llave

Doctor Rui António Loja Fernandes

Doctor Diogo Luís de Castro Vasconcelos de Aguiar Gomes

Doctor Joana Margarida Mavigné Andrade Alves de Sousa Nunes da Costa

Doctor José António Maciel Natário

2012

UNIVERSIDADE TÉCNICA DE LISBOA

INSTITUTO SUPERIOR TÉCNICO

KAM Theory, Presymplectic Dynamics and Lie algebroids

Hassan Najafi Alishah

Supervisor: Doctor Rui António Loja Fernandes

Thesis approved in public session to obtain the PhD Degree in

Mathematics

Jury final classification: Pass with Merit

Jury

Chairperson: Chairman of the IST Scientific Board

Members of the Committee:

Doctor Nguyen Tien Zung, Professor, Mathematics Institute of Toulouse, France

Doctor Rafael de la Llave, Professor, School of Mathematics, Georgia Institute of Technology, USA

Doctor Rui António Loja Fernandes, Professor Catedrático, Instituto Superior Tecnico, Lisbon

Doctor Diogo Luís de Castro Vasconcelos de Aguiar Gomes, Professor Catedrático, Professor Instituto Superior Tecnico, Lisbon

Doctor Joana Margarida Mavigné Andrade Alves de Sousa Nunes da Costa, Professora Associada (com Agregacão), Science and Technology faculty of the Coimbra University

Doctor José António Maciel Natário, Professor Associado, Instituto Superior Tecnico, Lisbon

Funding Institution: UT-Austin Portugal Program/ Fundacão para a Ciência e a Tecnologia

2012

Abstract

This thesis focus on various aspects of KAM theory, i.e., the study of invariant tori of a dynamical system where the motion is quasi-periodic. We show that the Diophantine condition that appears in such problems, in order to deal with convergence issues related with the appearance of "small divisors", can be interpreted as a cohomological condition. We then consider two concrete distinct classes of KAM problems:

1) We present a KAM theorem for presymplectic dynamical systems. The theorem has an "a posteriori" format: given a Diophantine frequency ω and a family of presymplectic mappings, we show that if for some map in this family we can find an embedded torus which is approximately invariant with rotation ω and satisfies some non-degeneracy condition, then we can find an invariant embedded torus for some map in the family close to the original map. Furthermore, we show that the dimension of the parameter space can be taken smaller if we assume that the presymplectic mappings in the family are exact.

2) We present a new approach to the KAM problem for a general vector field via Lie algebroids. We explain how the problem of persistence of invariant tori can be restated as a problem of stability of leaves of Lie algebroids. Then, we state a conjecture concerning the stability of compact invariant submanifolds of a Lie algebroid. We present some examples supporting this conjecture and we discuss possible approaches to prove this conjecture.

Keywords: KAM Theory, presymplectic map, Lie algebroid, stability of a leaf.

Resumo

Esta tese centra-se em certos aspectos da teoria KAM, i.e., do estudo de toros invariantes de um sistema dinâmico onde o movimento é quase periódico. Mostramos que a condição Diofantina que aparece naturalmente neste tipo de problemas, de forma a controlar-se as questões de convergência relaccionadas com o aparecimento de "divisores pequenos", pode ser interpretada como uma condição cohomológica. De seguida, consideramos dois tipos concretos de problemas KAM:

1) Enunciamos e mostramos um teorema KAM para sistemas dinâmicos presimplécticos. Este resultado tem o seguinte formato "a posteriori": dada uma frequência Diofantina ω e uma família de aplicações presimplécticas, mostramos que se uma aplicação nesta família possui um toro embebido que é aproximadamente invariante com frequência ω e satisfaz uma condição de não degeneresência, então podemos encontrar uma aplicação na família, próxima da aplicação original, que possui um toro invariante. Mostramos ainda que a dimensão do espaço dos parâmetros diminui se as aplicações presimplécticas na família forem exactas.

2) Apresentamos uma nova abordagem ao problema KAM para campos vectoriais arbitrários via teoria dos algebróides de Lie. Explicamos como o problema da persistência de toros invariantes pode ser reformulado na linguagem dos algebróides de Lie como um problema de estabilidade de órbitas. De seguida, enunciamos uma conjectura sobre a estabilidade de subvariedades compactas invariantes de um algebróide de Lie. Damos alguns exemplos que apoiam esta conjectura e discutimos duas vias possíveis para a sua demonstração.

Palavras Chave: Teoria KAM, aplicação presimpléctica, algebróide de Lie, estabilidade de folhas.

Acknowledgements

I would like to express special thanks to my supervisor, Professor Dr. Rui Loja Fernandes, for his guidance and support. This work could not have been done without his extremely useful suggestions and advice during our regular meetings and informal discussions.

I am grateful to Professor Dr. Rafael de la Llave , who introduced me to the subject when I was in Austin. The content of Chapter 3 of this thesis has been done under his guidance and has been published in the journal of Dynamics and Differential Equations.

I have been supported by the Portuguese Science Foundation (FCT) grant through the UT Austin-Portugal CoLab programme.

I would like to thank Portugal and portuguese people for great hospitality.

At the end, I would like to extent my sincerest thanks to my family for their support.

Contents

Chapter 1

Introduction

1.1 KAM theory

KAM theory has its roots in Celestial Mechanics and in the classical works on the planetary motion by astronomers and mathematicians such as Kepler, Newton, Lagrange, Liouville, Delaunary, Weieratrass and others. From a more modern point of view, due mainly to Poincaré, Birkhoff and Siegel, perturbative methods were suggested to handle the problem of stability of solutions of the n-body problem. This method remain problematic, due to the presence of arbitrary small divisors in the perturbative expansion, until **K**olmogorov in 1954, [19], followed by **A**rnold and **M**oser in the early 1960s, made a major breakthrough and succeed in overcoming the formidable technical problems related to the appearance of the small divisors. The main bulk of KAM theory is a set of techniques, based on fast convergent methods of Newton type, which can be used to solve various existence and stability questions about quasi-periodic solutions of Hamiltonian (or generalizations of Hamiltonian) dynamical systems. [1].

[1]We assume that the reader is familiar with basic concepts of symplectic geometry, for an introduction to the subject see, e.g., [7]

Definition 1.1.1. A **Hamiltonian system** is a triple (M, Ω, H) where (M, Ω) is a symplectic manifold and $H : M \to \mathbb{R}$ is a smooth function. The dynamics of the systems is governed by the **Hamiltonian vector field associated to** H:

$$X_H := \Omega^{-1}(\mathrm{d}H), \tag{1.1.1}$$

where Ω is viewed as a map $\Omega : TM \to T^*M$ and inversion is possible due to the non-degeneracy of the symplectic form.

We recall that the Darboux theorem shows that for any neighborhood of a point in a symplectic manifold (M^{2d}, Ω) one can choose local coordinates $(x_1, \ldots, x_n, y_1, \ldots, y_n)$ where the symplectic form takes the canonical form:

$$\Omega = \mathrm{d}x \wedge \mathrm{d}y := \sum_{i=1}^{d} \mathrm{d}x_i \wedge \mathrm{d}y_i.$$

In these canonical coordinates the equations for the trajectories of the Hamiltonian vector field X_H take the standard Hamiltonian form:

$$\begin{cases} \dot{x} = \frac{\partial H}{\partial y} \\ \dot{y} = -\frac{\partial H}{\partial x} \end{cases} \tag{1.1.2}$$

An immediate consequence of the skew-symetry of the symplectic form is that the Hamiltonian H is a **first integral** of the system, i.e., it is constant over trajectories of the vector field X_H, and this justifies using the name *conservative systems*. Moreover, notice that the knowledge of other first integrals of the system helps in understanding the behavior of solutions of the system, since any such solution is constrained to the common level sets of those first integrals.

Actually, the geometry underlying a Hamiltonian dynamical system allows one to explore even further the presence of first integrals. Recall that associated to the symplectic form Ω one has the **Poisson bracket** of two functions:

$$\{f, g\} := \Omega(X_f, X_g).$$

This is a *Lie bracket* on the space of smooth functions $C^\infty(M)$ which, additionally, satisfies the Leibniz identity:

$$\{f, gh\} = \{f, g\}h + g\{f, h\}.$$

Now one sees immediately that a function f is a first integral of X_H if and only if $\{H, f\} = 0$. Moreover, the Jacobi identity also shows that if f_1 and f_2 are first integrals then so is their Poisson bracket $\{f_1, f_2\}$.

Given a first integral f_1 all solutions of X_H will lie entirely on the level sets of f_1, so we are able to "reduce" the dimension by 1. Actually, since $\{H, f_1\} = 0$, one sees that the vector fields X_H and X_{f_1} commute. Hence, the action of $\mathbb{R}$ on M given by the flow of X_{f_1}:

$$(t, x) \mapsto \phi^t_{X_{f_1}}(x),$$

preserves the level sets of f_1 and fixes H, so maps solutions to solutions. Hence, we can further reduce by this action, and so a first integral allows actually to reduce the dimension of the phase space by 2. It turns out that the reduced level sets $M_c := f_1^{-1}(c)/\mathbb{R}$ are again symplectic manifolds and that the reduced dynamics are Hamiltonian: the function H induces a reduced Hamiltonian $H_{red} : M_c \to \mathbb{R}$. If we are given another first integral f_2, this will induce a first integral of the reduced system (M_c, Ω_c, H_{red}) provided that $\{f_1, f_2\} = 0$, and we can then proceed by further reducing using f_2. This leads to the following definition:

Definition 1.1.2. A Hamiltonian system (M, Ω, H) on a $2d$-dimensional manifold is called **completely integrable** if it admits d independent first integrals $f_1 = H, f_2, .., f_d$ which Poisson commute:

$$\{f_i, f_j\} = 0, \quad i, j = 1, ..., d. \tag{1.1.3}$$

Given a completely integrable system, the Arnold-Liouville theorem states that each connected component of the compact level sets of $f = (f_1, ..., f_d)$ is diffeomorphic to a d-dimensional torus $\mathbb{T}^d$ and gives a normal form for the system in a neighborhood of the torus. Note that any such torus will be invariant under the flow of X_H. More precisely ([2]), the theorem states that there exists an open neighborhood U of the connected compact component and a diffeomorphism $(x, y) : U \to \mathbb{T}^d \times \mathbb{R}^d$, such that the symplectic form Ω takes the standard form and H is independent of the angle variables x:

$$\Omega = \mathrm{d}x \wedge \mathrm{d}y := \sum_{i=1}^{d} \mathrm{d}x_i \wedge \mathrm{d}y_i, \qquad H = H(y).$$

This means that we can replace the original first integrals by $y_1, .., y_n$. The coordinates (x, y) are known as *action-angle coordinates* and it follows that the equations for the motion in these coordinates are given by

$$\begin{cases} \dot{x} = \frac{\partial H}{\partial y}(y) \\ \dot{y} = 0 \end{cases}$$

If we set $\omega_i(y) = \frac{\partial H_0}{\partial y_i}(y)$ then the solutions of the system in each torus $\{(x, y) \in \mathbb{T}^d \times \mathbb{R}^d : y = y_0\}$ are quasi-periodic with frequency vector $\omega(y_0)$:

$$x(t) = x_0 + \omega(y_0)t, \quad y(t) = y_0.$$

The map $y \mapsto \omega(y)$ is called the **frequency map**.

Obviously, although many examples of integrable systems are known (see, e.g., [2] or [1, Chapt. 5]), this is far from being a generic situation. However, given an arbitrary Hamiltonian dynamical system (M^{2d}, Ω, H) one may still ask if there exists some invariant torus where the motion is quasi-periodic. In other words, whether there exists a smooth embedding $K : \mathbb{T}^d \to M$ such that:

$$\phi_X^t(K(\theta)) = K(\theta + \omega t), \quad \forall \theta \in \mathbb{T}^d. \tag{1.1.4}$$

For invariant tori with rationally independent frequency vectors $\omega = (\omega_1, ..., \omega_d) \in \mathbb{R}^d$, where rationally independent means:

$$\omega.n := \sum_{i=1}^{d} \omega_i n_i \neq 0 \quad \forall n \in \mathbb{Z}^d \setminus \{0\}$$

one has the following important fact (see Lemma 3.2.5):

Proposition 1.1.3. *Any invariant torus* $K : \mathbb{T}^d \to M$ *of a Hamiltonian system* (M^{2d}, Ω, H) *on an exact symplectic manifold which has rationally independent frequencies* $\omega = (\omega_1, \ldots, \omega_d)$ *is Lagrangian:* $K^*\Omega = 0$.

The Weinstein Lagrangian Neighborhood Theorem states that a neighborhood U of any embedded Lagrangian torus is symplectomorphic to $T^*\mathbb{T}^d \simeq \mathbb{T}^d \times \mathbb{R}^d$ with its canonical symplectic form $\Omega = \mathrm{d}x \wedge \mathrm{d}y$, and so we can choose coordinates in a neighborhood of the torus where the trajectories of X_H are the solutions of Hamilton's equations:

$$\begin{cases} \dot{x} = \frac{\partial H}{\partial y} \\ \dot{y} = -\frac{\partial H}{\partial x} \end{cases}$$

where $\frac{\partial H}{\partial x}(x, 0) = 0$ and $\frac{\partial H}{\partial y}(x, 0) = \omega$ is constant.

It follows from this discussion that in the search for invariant tori of a Hamiltonian system (M^{2d}, Ω, H) where the motion is quasi-periodic one can look first at the case where $M = \mathbb{T}^d \times \mathbb{R}^d$ with its canonical symplectic form $\Omega = \mathrm{d}x \wedge \mathrm{d}y$, and look for Lagrangian embeddings $K : \mathbb{T}^d \to \mathbb{T}^d \times \mathbb{R}^d$ satisfying (1.1.4). Also, it is natural to look at the particular case of **nearly integrable Hamiltonian systems** where

$$H(x, y) = H_0(y) + \epsilon P(x, y).$$

This gives rise to two versions of KAM theory:

(i) In the classical version of KAM theory, one starts with a completely integrable system X_{H_0}, which has plenty of invariant tori, and one looks for invariant tori of small perturbations X_H.

(ii) In modern versions of KAM theory, one starts with a Hamiltonian dynamical system X_H which has an *almost invariant torus* and one looks for close by invariant tori.

Note that (ii) is a more general version that includes (i) as a special case: an invariant torus of X_{H_0} will be an almost invariant torus of X_H provided H is sufficiently close to H_0.

In KAM theory the arithmetic properties of the frequency ω play a major role:

Definition 1.1.4. Given $\gamma > 0$ and $\sigma \geq d$, we will denote by $D(\gamma, \sigma)$ the set of frequency vectors $\omega \in \mathbb{R}^{d+n}$ satisfying the **Diophantine condition**:

$$|k \cdot \omega| \geq \gamma |k|_{\mathbb{Z}}^{-\sigma} \quad \forall\, k \in \mathbb{Z}^{d+n} \backslash \{0\}, \tag{1.1.5}$$

where $|k|_{\mathbb{Z}} = |k_1| + ... + |k_{d+n}|$ and $k.\omega = \sum_{i=1}^{d+n} k_i \omega_i$.

In Chapter 2, we will discuss the nature of this condition.

An invariant torus $K(\theta)$ with frequency $\omega \in D(\gamma, \sigma)$ is called a **KAM torus**. Now the most classical version of the KAM theorem can be stated as follows:

Theorem 1.1.5. *If the frequency map $\omega = \frac{\partial H_0}{\partial y}$ of a real analytic integrable Hamiltonian $H_0(y)$ is a local diffeomorphism (Kolmogorov non-degeneracy condition), then KAM tori persist under small smooth perturbations of H_0.*

Roughly speaking, the proof of this theorem consists of an infinite iteration procedure, where at each step one performs a symplectic change of variables such that the error term decreases quadratically. The change of coordinates is obtained via a generating function, which guarantees that the symplectic structure is preserved under this change. The main technical problem is the appearance of small divisors $\omega.n$ in the process of calculating the generating function.

The loss of regularity is typical in problems involving small divisors, which means that one cannot fix the function space for the iteration. The usual technique to overcome this problem consists of introducing a *scale of Banach spaces* $\{\mathcal{B}_\xi : \xi > 0\}$ with the property $\mathcal{B}_{\xi'} \subset \mathcal{B}_\xi$ when $\xi' < \xi$ and work in $\mathcal{B}_{\xi_j}$ at the step j, for a suitable decreasing sequence ξ_j. Furthermore, one needs to control the small divisors at each step. In order to do so, one uses Kolmogorov's idea of keeping the frequency *fixed* so that one can systematically use the diophantine estimate (1.1.5): the non-degeneracy assumption imposed on the frequency map makes it possible to fix the frequency. The convergence of the procedure is guarantied by the quadratic decrease of the error term, which beats the "divergence" caused by the small divisors. For a detailed example of the above procedure for a concrete class of systems see [38, Chapter 3].

Moser, in 1962, proved a KAM type theorem in the framework of area-preserving twisting maps of the 2-annulus $[0, 1] \times \mathbf{S}^1$. He considered a C^k perturbation of an integrable analytic system. Moser's original set up correspond to the Hamiltonian case considered above, where $d = 2$. He required smoothness of degree $k = 333$. Later, Rüssmann was able to bring down this number to 5 by using a smoothing technique (via convolutions) which re-introduces at each step of the iteration a certain number of derivatives which compensates for the loss caused by the presence of the small divisors.

KAM type theorems have now been proved for other types of dynamical systems beyond the classical set up of Hamiltonian systems on symplectic manifolds: Hamiltonian systems on certain Poisson manifolds, reversible dynamical systems, dissipative systems, etc. KAM theory has also been developed in the infinite dimensional setting to deal with classes of partial differential equation carrying a Hamiltonian structure, such as water wave equations (see, e.g., [40]).

1.2 Outline of the thesis

The remainder of this thesis begins with Chapter 2, where we discuss various aspects of the theme "small divisors". In KAM type problems, small divisors appear when attempting to solve an equation called the "cohomological equation". The presence of small divisors lead to issues about convergence of solutions which can be overcome by imposing certain Diophantine conditions. This is recalled in the first two sections of Chapter 2. However, there is more to say about these conditions: we will eventually show that, at least in some cases, they can be interpreted geometrically as vanishing of certain deformation cohomologies attached to the problem. Hence, one can think of KAM theorems as instances of:

$$\text{infintesimal stability} \qquad \Longrightarrow \qquad \text{stability.}$$

Chapter 3 is devoted to the study of KAM theory in the context of presymplectic geometry. We consider the presymplectic manifold

$$M := T^*\mathbb{T}^d \times \mathbb{T}^n, \tag{1.2.1}$$

with an **exact** presymplectic form Ω of rank $2d$, whose kernel coincides with the $\mathbb{T}^n$-direction. We are interested in investigating invariant tori of presymplectic diffeomorphisms

$$f : M \to M, \quad f^*\Omega = \Omega.$$

One important class are the time 1 flows $f := \phi_X^1$ of a presymplectic vector field X. The approach we follow consists in starting with an *approximate* invariant torus for f, i.e., a torus $K : \mathbb{T}^{d+n} \to M$ for which the right hand side of (1.1.4) is non-zero, but small enough (in some norm that will be made explicit later), and then look for a near by invariant torus. Our main theorem can be stated roughly as follows:

Theorem 1.2.1. *Let $f_\lambda : M \to M$ be an analytic, non-degenerate, $(2d+n)$-parametric family of presymplectic diffeomorphisms of $M = T^*\mathbb{T}^n \times \mathbb{T}^d$, and assume that f_0 has an approximate invariant torus K_0, satisfying a non-degeneracy condition, with frequency ω satisfying a Diophantine condition. Then there exists a diffeomorphism f_{λ_∞} in this family, where λ_∞ is close to 0, which has an invariant torus K_∞ with frequency ω and which is* ***"close"*** *to the initial torus K_0.*

The precise version of the theorem will be stated below in chapter 3, where we formulate precisely the non-degeneracy conditions and we introduce norms to make precise the meaning of "close". In contrast with classical KAM theorem, we will not require the system to be neither nearly integrable, nor to be written in action-angle variables. Indeed, the fact that the dynamics of the system preserve the presymplectic structure implies that the KAM tori are automatically approximately reducible. This leads to an approximate solution of the linearized equations without transformation theory. Moreover, the reducing transformation is given explicitly in terms of the approximately translated torus, which form the basis of an efficient numerical algorithm (an explicit description of this algorithm can be found in [24]).

The proof of our main theorem follows an approach similar to the one developed in [14] for the symplectic case. The presymplectic case however has a few peculiarities due to the degeneracy of the 2-form. The quasi-Newton method used here (and in [14]) is of the type introduced by Moser in [30, 31]. We note however that the approach is not based on transformation theory, which seems problematic in the case of presymplectic mappings since generating functions are not as straightforward as in the symplectic case and the Lie transform method is hampered by the fact that there are several Hamiltonians that give the same vector field (see [8, 9, 17] for the theory of canonical transformations). The approach is based on deriving a parameterization equation and applying corrections additively. The presymplectic geometry leads to cancellations that reduce a Newton step to the constant coefficients cohomology equations of the type one usually finds in KAM theory. We also note that the same cancellations lead to very effective numerical algorithms.

As a byproduct of this approach, we will also prove a flux-type vanishing lemma for exact presymplectic diffeomorphisms. Roughly speaking, we will show that the average of the translation is zero in the directions other than the ones tangent to torus in the basis. Note that in the directions tangent to torus the averaging does not need to vanish. This also shows the need for considering a parametric family of diffeomorphisms, rather than just a single diffeomorphism.

Finally, it should be remarked that the results in Chapter 3 are not applicable to Hamiltonian dynamics in general Poisson manifolds. For regular Poisson structures, which have an underlying regular symplectic foliation, there are cohomological obstructions to find a compatible presymplectic structure, see [39]. Even when these obstructions vanish (e.g., locally around invariant tori), so that one can find a compatible presymplectic structure, Poisson diffeomorphisms do not coincide with presymplectic diffeomorphisms, and these two kinds of diffeomorphisms have quite distinct properties. A KAM theory for Poisson manifolds has been proposed in [33]. In section 3.8, we will compare the Poisson and presymplectic cases.

In Chapter 4, we will present a new approach to the problem of persistence of invariant tori of a vector field. We will consider this problem in the general context of invariant compact submanifolds of a Lie algebroid. We start by recalling the definition of a Lie algebroid and then explain, as a very simple example, how one can associate to a vector field, in a canonical way, a Lie algebroid, so that the leafs of this Lie algebroid are precisely the orbits of the vector filed. The main theorem in [11] states that infinitesimal stability of a compact leaf of a Lie algebroid leads to stability of that leaf. As we will see, when one applies this result to the Lie algebroid of a vector field, one recovers well known results about persistence of fixed points and periodic orbits of the vector field under perturbations.

The natural question to ask is then if one can extend the result of [11] to an invariant submanifold of a Lie algebroid and we conjecture that this is the case. First, one remarks that the condition of infinitesimal stability for orbits, which amounts to the vanishing of a certain cohomology group defined for the restricted Lie algebroid structure over the leaf, actually makes sense for any invariant submanifold. When applied to invariant tori of vector fields, this condition yields the Diophantine condition! We will present two examples giving some evidence for our conjecture, at least in the case of invariant tori. The main difficulty in extending the proof in [11] to the case of invariant submanifolds, is that the complex defining the cohomology group mentioned above, is not, in general, elliptic anymore. This fact makes the analytic part of the proof in [11] fail. We close Chapter 4 presenting two different attempts to circumvent these difficulties.

In the first attempt, we assume that the invariant torus of the unperturbed vector field $X_0 \in \mathfrak{X}(M)$ is a leaf of a certain Lie algebroid structure D_0 on a vector bundle $A \to M$ which contains X_0 in the image of its anchor map. If the map from the space of Lie algebroid structures on $A \to M$ to the space of vector fields $\mathfrak{X}(M)$, which sends D_0 to X_0, was an open map around D_0, then the result for stability of orbits would imply a result for stability of vector fields. We provide a counter example which shows that this map is not open, in general.

As a second attempt, we restrict ourself to Hamiltonian vector fields. A Lagrangian torus is invariant under the flow of a Hamiltonian vector field if and only if the restriction of the Hamiltonian to the torus is constant. On the other hand, the dynamics of the system restricted to the torus is conjugate to a linear one if and only if the normal variation of the Hamiltonian restricted to the torus is also constant. These two conditions lead to the construction of a functional, parametrized by the Hamiltonian, for which the zeros of the functional correspond precisely to the invariant tori where the dynamics are conjugate to linear ones. For the unperturbed Hamiltonian, one has that zero is a strong non-degenerate critical point of the functional, provided that the frequency map is a local diffeomorphism, i.e., provided that the Kolmogorov condition holds. Hence, under this assumption, it follows from an infinite dimensional implicit function theorem that every nearby functional also has a critical point as well. In the case of Lie algebroids one could show that this critical point must be a zero of the functional as well. But here, this does not seem to be the case.

Chapter 2

The Diophantine Condition

2.1 Small divisors

A basic feature in KAM theory is the presence of small divisors and the necessity to impose some kind of Diophantine condition to achieve any meaningful results.

In order to illustrate what are small divisors, consider a harmonic oscillator

$$\ddot{x} + \omega_1^2 x = 0,$$

whose general solution $x(t) = Acos(\omega_1 t + \theta)$ is a periodic motion of frequency ω_1. If we add to the system a periodic perturbation of frequency $\omega_2 \neq \omega_1$ such as:

$$\ddot{x} + \omega_1^2 x + \cos(\omega_2 t) = 0,$$

the new solution has the form

$$x(t) = A\cos(\omega_1 t + \theta) - \frac{\cos(\omega_2 t)}{\omega_1^2 - \omega_2^2}.$$

This function is only periodic when ω_2 is a rational multiple of ω_1. But even when ω_1 and ω_2 are rationally independent, the solutions are nice, bounded, quasi-regular oscillations. However, if ω_2 is close to $\pm\omega_1$, the quotient on the right hand side can become arbitrarily large. To see what happens as ω_2 approaches ω_1, let us focus on the simplest initial conditions $x(0) = \dot{x}(0) = 0$. Then

$$x(t) = \frac{\cos(\omega_1 t) - \cos(\omega_2 t)}{\omega_1^2 - \omega_2^2}$$

and L'Hôpital's rule gives that in the limit, as $\omega_1 \to \omega_2$, the solution approaches

$$x(t) = \frac{-t\sin(\omega_1 t)}{2\omega_1}.$$

This last function is unbounded. The periodic kicks of the perturbation build up, without canceling, to make the solution grow indefinitely. This kind of *resonance phenomena* is of course well-known to engineers. A more sophisticated phenomenon than simple resonance occurs when two (or more) distinct periodic motions of frequencies ω_1 and ω_2 interact with each other. The prototype example is given by the motion of two planets around the sun, which is a nearly integrable systems. This example fits in the more general problem studied by Poincaré of the dynamics generated by the flow of a one-parameter family of Hamiltonians of the form

$$H_0(y) + \epsilon P(x, y, \epsilon), \quad 0 < \epsilon \ll 1,$$

he called this problem *le problème général de la dynamique* to which he dedicated a large part of his monumental *Méthodes Nouvelles de la mécanique Céleste*, see [10].

In our case, the small divisors appear when one attempts to solve the following equation, sometimes known as the *cohomological equation*:

$$\phi(\theta + \omega) - \phi(\theta) = \eta(\theta), \tag{2.1.1}$$

or its continuous version:

$$\sum_{i=1}^{n} \omega_i \, \partial_{\theta_i}\phi(\theta) = \eta(\theta), \tag{2.1.2}$$

where ϕ and η are functions on the torus $\mathbb{T}^n$ and ω is the frequency vector. Equation (2.1.1) will show up in the KAM theory for persymplectic maps while (2.1.2) will appear in the KAM theory for general vector fields.

Before we discuss the appearance of small divisors when one solves this equations, let us justify the use of the term "cohomological equation". Recall that if G is a group with a representation $\rho : G \to GL(V)$ on some vector space V, then one defines the complex $(C^k(G,V), \mathrm{d})$ where $C^k(G,V)$ consist of all maps defined on k-copies of G, $c : G \times \cdots \times G \to V$, and the differential $\mathrm{d} : C^k(G,V) \to C^{k+1}(G,V)$ is given by:

$$\begin{aligned} \mathrm{d}c(g_0, \ldots, g_k) = \rho(g_0)c(g_1, \ldots, g_k) + \\ \sum_{i=0}^{k-1} (-1)^i c(g_0, \ldots, g_i g_{i+1}, \ldots, g_k) + (-1)^k c(g_0, \ldots, g_{k-1}). \end{aligned} \tag{2.1.3}$$

When $k = 0$, $C^0(G,V) = V$ and one defines $\mathrm{d}v(g) = \rho(g)v - v$. The corresponding cohomology groups:

$$H^k(G,V) = \frac{\text{Ker d}}{\text{Im d}},$$

define the *group cohomology* with coefficients in the representation V.

Now consider the vector space $V = C^\infty(M)$ consisting of smooth real valued functions on a manifold M. Any diffeomorphism $f : M \to M$ defines an action of $G = \mathbb{Z}$ on V by setting:

$$(n \cdot \eta)(x) = \eta(f^n(x)).$$

Then, given any $\eta \in V$ one can define $c \in C^1(\mathbb{Z}, V)$ by setting $c(n) = \sum_{i=0}^{n-1} \eta(f^i)$. One checks easily that $\mathrm{d}c = 0$, so c is a 1-cocyle and that the equation for this cocyle to be a coboundary $c = \mathrm{d}\phi$ reduces to:

$$\phi \circ f - \phi = \eta.$$

Therefore, equation (2.1.1) is a cohomological condition when we think of η as defining a cocycle for the transformation $T : \theta \to \theta + \omega$, and one says that η is a *coboundary* if the equation (2.1.1) has a solution. For example, if ω be a Diophantine vector then this equation can be solved for every η with vanishing average and it yields a unique solution with zero average, so $H^1(\mathbb{Z}, V) = 0$, where $V = \{\eta \in C^\infty(\mathbb{T}^d) : \mathrm{avg}(\eta) := \int_{\mathbb{T}^d} \eta \mathrm{d}\theta = 0\}$. The vanishing of this cohomology space has consequences on the rigidity of the dynamical systems and existence of invariant measurers on them (see [3] and references therein).

Let us now turn to the appearance of small divisors in solving equation (2.1.2) (the same considerations apply to (2.1.1) with minor modifications). We use Fourier expansions, so we write:

$$\eta(\theta) = \sum_{k\in\mathbb{Z}^n} \hat{\eta}_k e^{2\pi\mathfrak{i}(k.\theta)}, \quad \hat{\eta}_0 = 0,$$

where $\mathfrak{i} = \sqrt{-1}$. Then equation (2.1.1) has the formal solution:

$$\hat{\phi}_k 2\pi\mathfrak{i}(k.\omega) = \hat{\eta}_k.$$

Hence, for a non-resonant frequency vector ω, i.e., if $(k.\omega) \neq 0$, $\forall k \in \mathbb{Z}^n\backslash\{0\}$, one gets:

$$\hat{\phi}_k = \frac{\hat{\eta}_k}{2\pi\mathfrak{i}(k.\omega)} \tag{2.1.4}$$

so one finds the presence of small divisors.

2.2 The Diophantine Condition

If one does not put any quantitive restriction on how fast $|(k.\omega)|^{-1}$ grows, the solutions given by (2.1.4) may even fail to be distributions. For example, take $\hat{\eta}_k = e^{-|k|_{\mathbb{Z}}}$, then it is not hard to construct $\omega \in \mathbb{R}^n\backslash\mathbb{Q}^n$ such that there are infinitely many k for which $|(k.\omega)|^{-1} \geq e^{e^{|k|_{\mathbb{Z}}}}$, see [24]. This means that $\sum_{k\in\mathbb{Z}^n} \hat{\phi}_k(1 + |k|^2_{\mathbb{Z}})^{s/2}$ is not convergent even for $s < 0$, i.e. $\hat{\phi}_k$, cannot be the Fourier coefficient of neither a function, nor a distribution.

Definition 2.2.1. Given $\gamma > 0$ and $\sigma \geq n$, we will denote by $D(\gamma, \sigma)$ the set of frequency vectors $\omega \in \mathbb{R}^n$ satisfying the **Diophantine condition**:

$$|(k.\omega)| \geq \gamma.|k|_{\mathbb{Z}}^{-\sigma}, \quad \forall k \in \mathbb{Z}^n\backslash\{0\}. \tag{2.2.1}$$

Let U_ρ denote the complex strip of width $\rho > 0$:

$$U_\rho = \{\theta \in \mathbb{C}^n/\mathbb{Z}^n : |\mathrm{Im}(\theta)| \leq \rho\},$$

and for a function η defined on U_ρ set:

$$\|\eta\|_\rho := \sup_{\theta \in U_\rho} |\eta(\theta)|, \tag{2.2.2}$$

and denote by $\|\eta\|_r$ its C^r-norm. If η is analytic, then its Fourier coefficients satisfy the bound

$$|\hat{\eta_k}| \leq e^{-2\pi\rho|k|_{\mathbb{Z}}}\|\eta\|_\rho,$$

while for $\eta \in C^r$ we have the bound:

$$|\hat{\eta_k}| \leq (2\pi)^{-r}|k|_{\mathbb{Z}}^{-r}\|\eta\|_r.$$

Hence, assuming the Diophantine condition (2.2.1), the solution ϕ of (2.1.4) will satisfy:

$$|\hat{\phi_k}| \leq (2\pi)^{-1}\gamma^{-1}|k|_{\mathbb{Z}}^{-\sigma}e^{-2\pi\rho|k|_{\mathbb{Z}}}\|\eta\|_\rho$$

for analytic η, and it will satisfy:

$$|\hat{\phi_k}| \leq (2\pi)^{-r-1}\gamma^{-1}|k|_{\mathbb{Z}}^{-\sigma-r}\|\eta\|_r.$$

for $\eta \in C^r$. These estimates does not allow one to conclude that ϕ belongs to exactly the same function space as η, but rather to a slightly weaker space: the word "weaker" in the analytic case means a loss in the domain, while in the differentiable case it means a loss of degree of differentiability. See [24] for more details.

We have just seen the necessity of the Diophantine condition. What about the existence of frequency vectors satisfying this condition?
The existence of such frequencies is easy to deduce. Observe that $D(\gamma, \sigma)$ is the complement of the open dense set $R_\gamma^\sigma = \bigcup_{0 \neq k \in \mathbb{Z}^n} R_{\gamma,k}^\sigma$, where

$$R_{\gamma,k}^\sigma = \{\omega \in \mathbb{R}^n : |(k.\omega)| < \gamma.|k|_{\mathbb{Z}}^{-\sigma}\}$$

Obviously, for any bounded domain $\Omega \subset \mathbb{R}^n$, its Lebesgue measure satisfies the estimate $m(R^{\sigma}_{\gamma,k} \cap \Omega) = O(\gamma/|k|_{\mathbb{Z}}^{\sigma+1})$, hence:

$$m(R^{\sigma}_{\gamma} \cap \Omega) \leq \sum_{k} m(R^{\sigma}_{\gamma,k} \cap \Omega) = O(\gamma)$$

provided that $\sigma > n-1$. Therefore, $R^{\sigma} = \bigcap_{\gamma>0} R^{\sigma}_{\gamma}$ is a set of measure zero, and its complement

$$D(\sigma) := \bigcup_{\gamma>0} D(\gamma, \sigma)$$

is a set of full measure in $\mathbb{R}^n$, for any $\sigma > n-1$. In other words, almost every ω in $\mathbb{R}^n$ belongs to $D(\sigma)$, $\sigma > n-1$, which is the set of all ω in $\mathbb{R}^n$ satisfying (3.2.2) for some $\gamma > 0$ while σ is fixed.

Remark 2.2.2. In the classical KAM theorem, the non degeneracy assumption of the frequency map guarantees that the set of action values that correspond to Diophantine frequencies has full measure as well. In other word, the set of KAM tori in Theorem (1.1.5), includes almost all tori, so almost all tori of the completely integrable Hamiltonian system survive small perturbations.

2.3 The Diophantine condition in our work

In Chapters 3 and 4, we will discuss two different kinds of KAM theorems, one for presymplectic maps and one for general vector fields. Here we explain how the cohomological equation appears in these two problems and hence how one uses the Diphantine condition in each of these problems.

2.3.1 The case of presymplectic maps

In this case, as we have explained in the introduction, the main idea is to find an invariant torus given an approximate invariant one. One starts with some torus

$K_0(\theta) : \mathbb{T}^{d+n} \to \mathbb{T}^d \times \mathbb{R}^d \times \mathbb{T}^n$ which satisfies

$$f_0(K_0(\theta)) - K_0(\theta + \omega) = e_0(\theta) \tag{2.3.1}$$

where f_λ is a parametric family of presymplectic maps (to be defined precisely later) and $e(\theta)$ is an small enough error term, and the aim is to find $(K_\infty, \lambda_\infty)$ such that

$$f_{\lambda_\infty}(K_\infty(\theta)) - K_\infty(\theta + \omega) = 0.$$

In order to do so, we will perform an iteration procedure: starting with (2.3.1), we look for an approximate solution for the corresponding linearized equation

$$\begin{aligned} &DG(K_0, 0)|_{(\Delta_0(\theta), \varepsilon_0)} := \\ &\quad \left.\frac{\partial f_\lambda(K_0(\theta))}{\partial \lambda}\right|_{\lambda=0} \varepsilon_0 + Df_0(K_0(\theta))\Delta_0(\theta) - \Delta_0(\theta + \omega) = -e_0(\theta). \end{aligned}$$

By an approximate solution we mean one that satisfies the linearized equation up to a quadratic error, i.e., a solution $\Delta_0(\theta)$ such that:

$$\|DG(K_0, 0)|_{(\Delta_0(\theta), \varepsilon_0)} + e_0\|_{\rho_0 - \delta_0} \leq C\|e_0\|^2_{\rho_0}$$

where C is a constant to be determined later. To do this, we remove the terms bounded by a multiple of the quadratic error and then performes a change of variable that transforms the equation to a new one which is a cohomology equation of the form (2.1.1). As we explained, the Diophantine condition is necessary to solve the equation (2.1.1). The approximate solution of the linearized equation will be used to produce a new torus which has smaller error term, so one can proceed with the iteration.

2.3.2 The case of general vector fields

In this case, as we have pointed out in the introduction, our aim is to understand how one can view KAM as part of the *moto*:

$$\text{infintesimal stability} \qquad \Longrightarrow \qquad \text{stability.}$$

Here one starts with the manifold $M = \mathbb{T}^n \times \mathbb{R}^m$ with the vector field

$$X_0 = \sum_i \omega_i \frac{\partial}{\partial x_i} + \sum_{j,k} \Omega_{j,k} y^k \frac{\partial}{\partial y_j},$$

which has the invariant torus $K_0 : \mathbb{T}^n \to M$, $\theta \mapsto (\theta, 0)$ where the motion is quasi-periodic and asks if a nearby vector field X has a nearby invariant torus $K : \mathbb{T} \to M$ where the motion is quasi-periodic. It turns out that to a vector field X one can associated a Lie algebroid A_X and that the geometric meaning of infinitesimal stability is here formulated in terms of a cohomology group attached to this Lie algebroid: one needs the vanishing of the cohomology group:

$$H^1(A_{X_0}|_{K_0}, \nu(K_0)) = 0.$$

This cohomology can be seen as the deformation cohomology attached to changes of the invariant subset K_0 upon small deformations of the Lie algebroid structure.

On the other hand, we will see in Chapter 4 that this group is trivial if and only if one is able to find at least one solution $h \in C^\infty(M, \mathbb{R}^m)$ for the equation:

$$(\sum_i \omega_i \frac{\partial}{\partial x_i} \cdot I_m - \Omega).e(x) = h(x) \tag{2.3.2}$$

for every $h \in C^\infty(M, \mathbb{R}^m)$. Using Fourier analysis, as above, we will see that this is indeed the case provided Ω is diagonalizable, with non zero eigenvalues $\Omega_1, \dots, \Omega_m$, and that the (ω_i, Ω_j) satisfy a Diophantine condition. In conclusion, we have:

(i) The Diophantine condition amounts to vanishing of a cohomology group, and

(ii) The vanishing of this cohomology group amounts to infinitesimal stability.

Chapter 3

Tracing KAM tori in presymplectic Dynamical systems

3.1 Presymplectic dynamical systems

Presymplectic structures (constant rank, closed 2-forms) arise naturally in the study of degenerate Lagrangian and Hamiltonian mechanical systems with constrains, in time dependent Hamiltonian systems and in control theory. (see, e.g., [13, 15, 20, 21, 28, 29, 34]). For other situations where presymplectic dynamics occur see, e.g., [5].

Given a presymplectic form $\Omega \in \Omega^2(M)$, a vector field $X \in \mathfrak{X}(M)$ is said to be a Hamiltonian vector field associated with a function $H \in \mathcal{C}^\infty(M)$ if:

$$i_X\Omega = \mathrm{d}H.$$

Due to the degeneracy of Ω, there can be different functions H associated with X, not differing by a constant. The corresponding flow $\phi_X^t : M \to M$ is a 1-parameter group of presymplectic diffeomorphisms: $(\phi_X^t)^*\Omega = \Omega$. Hence, the dynamics of such systems leave the presymplectic structure invariant. One example to keep in mind could be the three dimensional torus endowed with a presymplectic form $\Omega = d\Psi_1 \wedge d\Psi_2$.

Clearly, the kernel is given by the level sets of Ψ_1, Ψ_2.

A more complicated example on $\mathbb{T}^3$ is $\Omega = d\Psi_1 \wedge \gamma$ where γ is a closed but not exact form. In this case, the kernel can be an irrational foliation.

Another example related to the previous ones is the study of quasi-periodically perturbed Hamiltonian systems $H(x, \omega t)$. These can be made autonomous by adding an extra variable $\theta \in \mathbb{T}^d$ that satisfies $\frac{d}{dt}\theta = \omega$. The phase space is now supplemented by a factor $\mathbb{T}^d$. The symplectic form in the phase space becomes a presymplectic form in the extended phase space having $\mathbb{T}^d$ in the kernel. Even this elementary example was considered as covered by the KAM theory of symplectic systems at the time of writing [25].

The theory of presymplectic manifolds was developed (e.g. in [23]) to give a geometric framework to the Dirac theory of constrained systems, [26, 27]. There are many physically interesting examples of constrained systems to which the present theory applies. Notably, besides the examples in [26, 27], the papers [20, 21] contain a very concrete example of a relativistic system of spinning particles which is close to integrable.

The paper [13] shows how the Pontryagin maximum principle for optimal trajectories can be formulated using presymplectic systems. If we consider a mechanical system with KAM tori and subject it to a control indexed by enough parameters, the results in this chapter give a condition which ensures that the one adjust parameters to maintain the quasi-periodic motion. It would be interesting to study in detail concrete models, specially because the methods we use here, are well suited for numerical implementations.

Note that, in contrast with symplectic manifolds, presymplectic manifolds may be odd dimensional. Hence, it is clear that an extension of the symplectic theory to presymplectic systems will require significant modifications. A general theory of perturbations of quasi-periodic motions independent of geometric structures was undertaken in [32].

Remark 3.1.1. An important well known fact about presymplectic forms Ω is that the kernel of Ω is an integrable distribution.

We recall that the kernel of a form is

$$\mathrm{Ker}(\Omega) = \{X | i_X(\Omega) = 0\} = \{X | \Omega(X, Z) = 0 \ \forall Z\}$$

Note that, for a general 2-form Ω and any three vector fields X, Y, Z, we have:

$$\begin{aligned} \mathrm{d}\Omega(X, Y, Z) = {} & X(\Omega(Y, Z)) - Y(\Omega(X, Z)) + Z(\Omega(X, Y)) \\ & - \Omega([X, Y], Z) + \Omega([X, Z], Y) - \Omega([Y, Z], X) \end{aligned}$$

If $\mathrm{d}\Omega = 0$ and X, Y are in the kernel of Ω, for any Z, we have $\Omega([X, Y], Z) = 0$.

Hence, if $X, Y \in \mathrm{Ker}(\Omega)$, $[X, Y] \in \mathrm{Ker}(\Omega)$. This shows that the distribution given by the kernel can be integrated to a manifold.

Of course, in a torus, it could well happen that the leaves integrating the kernel are not compact (e.g. they could be an irrational foliation).

3.2 Preliminaries and Motivation

3.2.1 Notations and Preliminaries

In this section we will fix some notations and state a few preliminary results. Along the way, we will also justify the assumptions that will appear later in our main result.

As stated in the introduction, we consider $M = \mathbb{T}^d \times \mathbb{R}^d \times \mathbb{T}^n$ equipped with a constant rank exact presymplectic structure, i.e., an exact 2-form $\Omega \in \Omega^2(M)$, such that its kernel is:

$$N := \mathrm{Ker}\, \Omega = \{(u, (0, 0, z)) \in TM \mid u \in M, z \in \mathbb{R}^n\}.$$

The exactness assumption places restrictions on the presymplectic form but for applications to Hamiltonian dynamical systems with constrains or degenerate Lagrangian systems this is not too restrictive. For these systems the phase space is often obtained by restriction to a submanifold where the 2-form is the pullback of the canonical symplectic structure on the cotangent bundle (see [23]) or some other exact symplectic forms (see [28]).

Let $V = \{(u,(x,y,0)) \in TM | u \in M, (x,y) \in \mathbb{R}^{2d}\}$ so that $TM = V \oplus N$, and denote by $\pi : TM \to V$ the canonical projection on V. For each $u \in M$, we have the linear isomorphism $\tilde{J}(u) : T_uM \to T_uM$ defined by:

$$\Omega_u(\xi,\eta) = \langle \xi, \tilde{J}(u)\eta \rangle, \quad \xi,\eta \in T_uM \tag{3.2.1}$$

where

$$\tilde{J}(u) = \begin{bmatrix} J(u) & 0 \\ 0 & 0 \end{bmatrix}$$

and $\langle \cdot,\cdot \rangle$ denotes the standard Euclidean inner product on $\mathbb{R}^{2d+n}$. The skew-symmetry of Ω implies that $J^\intercal = -J$.

We will be using the following norms. If $x = (x_1, ..., x_{d+n}) \in \mathbb{R}^{d+n}$ we set:

$$|x| := \max_{j=1,..,d+n} |x_j|.$$

For an analytic function g on a complex domain $\mathcal{B}$ we denote by $|g|_{C^m,\mathcal{B}}$ its C^m-norm:

$$|g|_{C^m,\mathcal{B}} := \sup_{0\leq |k|_{\mathbb{Z}} \leq m} \sup_{z\in\mathcal{B}} |D^k g(z)|,$$

where $|l|_{\mathbb{Z}} := |l_1| + ... + |l_{d+n}|$.

We will be looking for real analytic invariant tori which extend holomorphically to a small strip in the complex space. More precisely, let U_ρ denote the complex strip of width $\rho > 0$:

$$U_\rho = \{\theta \in \mathbb{C}^{d+n}/\mathbb{Z}^{d+n} : |\mathrm{Im}(\rho)| \leq \rho\},$$

and introduce the following family of maps.

Definition 3.2.1. The space $(\mathcal{P}_\rho, \|.\|_\rho)$ consists of functions $K : U_\rho \to M$ which are one periodic in all their arguments, real analytic on the interior of U_ρ and continuous on the closure of U_ρ. We endow this space with the norm

$$\|K\|_\rho := \sup_{\theta \in U_\rho} |K(\theta)|, \tag{3.2.2}$$

which makes it into a Banach space.

We will also use the same notations for functions taking values in vector spaces or in matrices.

Some well known results about the spaces above are the Cauchy bounds below (a consequence of Cauchy's integral representation of the derivative). For $0 < \delta < \rho$, we have:

$$\|D^j K\|_{\rho-\delta} \leq C_j \delta^{-j} \|K\|_\rho \tag{3.2.3}$$

Like in all other KAM type results we will have to deal with small divisors. For that we set:

Definition 3.2.2. Given $\gamma > 0$ and $\sigma \geq d + n$, we will denote by $D(\gamma, \sigma)$ the set of frequency vectors $\omega \in \mathbb{R}^{d+n}$ satisfying the **Diophantine condition**:

$$|l \cdot \omega - m| \geq \gamma |l|_{\mathbb{Z}}^{-\sigma} \quad \forall l \in \mathbb{Z}^{d+n} \backslash \{0\}, m \in \mathbb{Z} \tag{3.2.4}$$

The aim of this paper is to find invariant tori of a given frequency ω for a m-parametric family of presymplectic diffeomorphism f_λ, defined as follows:

Definition 3.2.3. A m**-parametric family of presymplectic diffeomorphisms** f_λ is a function

$$f : M \times B \to M, \qquad B \subseteq \mathbb{R}^m,$$

such that for each $x \in M$ the map $f(x, \cdot)$ is of class C^2 and for each $\lambda \in B$ the map $f_\lambda := f(\cdot, \lambda)$ is a real analytic presymplectic diffeomorphism.

We will introduce an algorithm to solve the equation

$$f_\lambda(K(\theta)) - K(\theta+\omega) = 0, \quad \omega \in D(\gamma,\sigma), \tag{3.2.5}$$

given that one knows an approximate solution $K_0(\theta)$ for the diffeomorphism f_{λ_0}, where, without loss of generality, we will set $\lambda_0 = 0$. In other words, we know that

$$f_{\lambda_0}(K_0(\theta)) - K_0(\theta+\omega) = e_0(\theta), \tag{3.2.6}$$

where the error term $e_0(\theta)$ has small enough norm. Equation (3.2.5) will be solved by Newton method where at each step we have infinitesimal equations given by

$$Df_{\lambda_i}(K_i(\theta))\Delta_i(\theta) - \Delta_i(\theta+\omega) + \left.\frac{\partial f_\lambda(K_i(\theta))}{\partial\lambda}\right|_{\lambda=\lambda_i} \varepsilon_i = -e_i(\theta). \tag{3.2.7}$$

The approximate invariant tori and the geometry of the problem will lead us to a change of variables that will reduce (3.2.7) to a simpler equation with constant coefficients (cohomological equation) that can be solved by the following result of Rüssmann which will also be useful in some other proofs.

Proposition 3.2.4 ([24, 35]). *Let $\omega \in D(\sigma,\gamma)$ and assume that $\eta : \mathbb{T}^{d+n} \to \mathbb{R}^{2d+n}$ is analytic on U_ρ and has zero average,* $\operatorname{avg}(\eta) = 0$*. Then for all $0 < \delta < \rho$, the difference equation*

$$\phi(\theta) - \phi(\theta+\omega) = \eta(\theta) \tag{3.2.8}$$

has a unique zero average solution $\phi : \mathbb{T}^{d+n} \to \mathbb{R}^{2d+n}$ which is analytic in $U_{\rho-\delta}$. Moreover, this solution satisfies the following estimate:

$$\|\phi\|_{\rho-\delta} \le c_0\gamma^{-1}\delta^{-\sigma}\|\eta\|_\rho, \tag{3.2.9}$$

where c_0 is a constant depending on n and σ.

3.2.2 Lagrangian properties of invariant tori

A first, very important, consequence of the Diophantine condition on ω and the exactness of the presymplectic form is that KAM tori are actually Lagrangian submanifolds:

Lemma 3.2.5. *If $K(\theta) \in \mathcal{P}_\rho$ is a solution of* (3.2.5) *then $K^*\Omega$ is identically zero.*

Proof. Since $K(\theta)$ satisfies (3.2.5) and f is presymplectic we have

$$K^*\Omega = (K \circ T_\omega)^*\Omega,$$

where $T_\omega(\theta) = \theta + \omega$. Moreover, since ω is rationally independent then rotation on the torus is ergodic and this implies that $K^*\Omega$ is constant. If we write $K^*\Omega$ in matrix form, exactly as we did for Ω in (3.2.1) we have

$$K^*\Omega(\xi, \eta) = \langle \xi, L(\theta)\eta \rangle \quad \xi, \eta \in T_\theta(\mathbb{T}^{d+n}) \tag{3.2.10}$$

where $L(\theta)$ is actually constant. It remains to show that $L(\theta) \equiv 0$.

The 2-form Ω is exact, so we can write $\Omega = \mathrm{d}\alpha$ where

$$\alpha(u) = a(u)\mathrm{d}u, \quad a(u) = (a_1(u), ..., a_{2d+n}(u))^\intercal.$$

Then we find that

$$(K^*\alpha) = \sum_{j=1}^{d+n} C_j(\theta)\mathrm{d}\theta^j$$

where the components C_j have the following expression

$$C_j(\theta) = DK(\theta)a(K(\theta))_j.$$

This implies $L(\theta) = DC(\theta)^\intercal - DC(\theta)$. But now:

$$\mathrm{avg}(DC(\theta)) := \int_{\mathbb{T}^{d+n}} DC(\theta)\mathrm{d}\theta = 0, \tag{3.2.11}$$

which shows that:

$$\mathrm{avg}(L(\theta)) = 0.$$

But $L(\theta)$ being constant, we conclude that $L(\theta) = 0$, i.e., $K^*\Omega = 0$. ■

Following simple lemma extends the result of the Lemma 3.2.5 to approximate invariant tori:

Lemma 3.2.6. *Let $f_0 : M \to M$ be a presymplectic analytic diffeomorphism and let $K \in \mathcal{P}_\rho$ be an approximate invariant torus with frequency $\omega \in D(\gamma, \sigma)$:*

$$f_0(K(\theta)) - K(\theta + \omega) = e(\theta). \tag{3.2.12}$$

and assume that f_0 extends holomorphically to some complex neighborhood of the image of U_ρ under K:

$$\mathcal{B}_r = \{z \in \mathbb{C}^{2d+n} : \sup_{\theta \in U_\rho} |z - K(\theta)| < r\}.$$

Then there exist a constant $C > 0$, depending on $n, \sigma, \rho, \|DK\|_\rho, |f_0|_{C^1,\mathcal{B}_r}$ and $|J|_{C^1,\mathcal{B}_r}$, such that for $0 < \delta < \frac{\rho}{2}$

$$\|L\|_{\rho-2\sigma} \leq C\gamma^{-1}\delta^{-(\sigma+1)}\|e\|_\rho \tag{3.2.13}$$

where L is the matrix representing the pullback form $K^*\Omega$ (see 3.2.10).

Proof. Let $g := L - L \circ T_\omega$. Then, we note that g is the expression in coordinates of

$$K^*\Omega - T_\omega^* K^*\Omega = K^* f_0^*\Omega - T_\omega^* K^*\Omega$$

Hence, when K is exactly invariant $g = 0$. One can also easily show that $||g|| \leq ||De||$ See [14] for more details.

Using Proposition 3.2.4, one obtains that:

$$\|L\|_{\rho-2\delta} \leq c_0\gamma^{-1}\delta^{-\sigma}\|g\|_{\rho-\delta}.$$

One can bound the norm of g in exactly the same way as in the symplectic case, which can be found in [14], to obtain the result. ■

3.2.3 Automatic reducibility near invariant tori

In this subsection we will assume that $K(\theta)$ is an invariant torus of f, i.e., a solution of (3.2.5). When one starts instead with an approximate invariant torus $K_0(\theta)$ of f, i.e., a solution of (3.2.6), the results of this subsection do not hold anymore. However, we will see in the next sections that we have versions of these results which hold in the approximate case and which will allow us to perform the Newton method and conclude the existence of an invariant torus. For $K(\theta) \in \mathcal{P}_\rho$ let us decompose its Jacobian in the form

$$DK(\theta) = (X(\theta), Z(\theta)) \tag{3.2.14}$$

where $X(\theta), Z(\theta)$ are the first d and last n columns of $DK(\theta)$. Also, for every vector in $TM = V \oplus N$, we will use the subscripts V and N for the first and second projections in each factor. Assume that $K(\theta)$ solves (3.2.5) and that there exists a $d \times d$-matrix valued function $N(\theta)$ such that

$$N(\theta)(X_V^\intercal(\theta) \cdot X_V(\theta)) = \mathrm{Id}, \tag{3.2.15}$$

where $X(\theta)$ is as in (3.2.14). This non-degeneracy assumption will turn out to be one of the ingredients to solve (3.2.5) approximately. Also, set[1]:

$$Y_V(\theta) := X_V(\theta)N(\theta) \quad \text{and} \quad Y(\theta) := \begin{bmatrix} Y_V(\theta) \\ 0 \end{bmatrix}. \tag{3.2.16}$$

Then the following matrix will provide us the change of variable needed to reduce the linearized equations (3.2.7) to a simple form:

$$M(\theta) := \begin{pmatrix} X_V(\theta) & J^{-1}(K(\theta))Y(\theta) & Z_V(\theta) \\ X_N(\theta) & 0 & Z_N(\theta) \end{pmatrix}, \tag{3.2.17}$$

where X, Z and Y are defined in (3.2.14) and (3.2.16) respectively. The non-degeneracy assumption (3.2.15), together with the fact that $K(\theta)$ is Lagrangian (Lemma 3.2.5), show that:

$$\Omega_{K(\theta)}(X(\theta), J^{-1}(K(\theta))Y(\theta)) = I_d \tag{3.2.18}$$

$$\Omega_{K(\theta)}(X(\theta), X(\theta)) = 0 \tag{3.2.19}$$

[1]We will often abuse notation and will use $Y(\theta)$ to denote both $Y_V(\theta)$ and $Y(\theta)$.

$$\Omega_{K(\theta)}(X(\theta), Z(\theta)) = 0 \tag{3.2.20}$$

Therefore, $X(\theta)$, $J^{-1}(K(\theta))Y(\theta)$ and $Z(\theta)$ do not form a presymplectic basis along the torus $K(\theta)$, the reason is that neither $\Omega_{K(\theta)}(J^{-1}(K(\theta))Y(\theta), J^{-1}(K(\theta))Y(\theta))$ nor $\Omega_{K(\theta)}(J^{-1}(K(\theta))Y(\theta), Z(\theta))$ have to be zero, but they do provide a basis where Ω takes a rather simple form. Moreover, as the following lemma shows, they transform the linearized equations (3.2.7) into a simpler form:

Lemma 3.2.7. *The set $\{X(\theta), J^{-1}(K(\theta))Y(\theta), Z(\theta)\}$ is a basis provided the matrix*

$$V(\theta) = \begin{bmatrix} 0 & I_d & 0 \\ -I_d & -Y^{\intercal}(\theta)J^{-1}(K(\theta))Y(\theta) & (J^{-1}(K(\theta))Y(\theta))^{\intercal}J(K(\theta))Z_V(\theta) \\ X_N(\theta) & 0 & Z_N(\theta) \end{bmatrix} \tag{3.2.21}$$

is invertible. In this case, we have:

$$Df(K(\theta)).(X(\theta), Z(\theta)) = (X(\theta+\omega), Z(\theta+\omega)), \tag{3.2.22}$$

$$\begin{aligned} Df(K(\theta))J^{-1}(K(\theta))Y(\theta) =& X(\theta+\omega)S_1(\theta) + J^{-1}(K(\theta+\omega))Y(\theta+\omega)\,\mathrm{Id} + \\ &+ Z(\theta+\omega)A(\theta), \end{aligned} \tag{3.2.23}$$

where $A(\theta)$ and $S_1(\theta)$ are matrices satisfying: [2]

$$Df(K(\theta)) \cdot M(\theta)) = M(\theta+\omega) \cdot \begin{bmatrix} I_d & S_1(\theta) & 0 \\ 0 & I_d & 0 \\ 0 & A(\theta) & I_n \end{bmatrix}. \tag{3.2.24}$$

Proof. Let

$$Q(\theta) := \begin{bmatrix} X_V^{\intercal}(\theta)J(K(\theta)) & 0 \\ (J^{-1}(K(\theta))Y(\theta))^{\intercal}J(K(\theta)) & 0 \\ 0 & I_n \end{bmatrix}. \tag{3.2.25}$$

[2] We emphasize that identity (3.2.24) holds only when we have an invariant torus. In Corollary 3.4.2, we will prove that for approximately invariant tori (3.2.24) holds up to an error which can be bounded by the error in the invariance equation.

The expression (3.2.17) for M and relations (3.2.18), (3.2.19) and (3.2.20), give:

$$Q(\theta)\cdot M(\theta) = \begin{bmatrix} 0 & I_d & 0 \\ -I_d & -Y^{\intercal}(\theta)J^{-1}(K(\theta))Y(\theta) & (J^{-1}(K(\theta))Y(\theta))^{\intercal}J(K(\theta))Z_V(\theta) \\ X_N(\theta) & 0 & Z_N(\theta) \end{bmatrix}, \tag{3.2.26}$$

which shows that $\{X_V(\theta), J^{-1}(K(\theta))Y(\theta)\}$ is a basis for $V := \pi(TM)$ and $Q(\theta)$ is invertible. Using this fact one can write

$$Z_V(\theta) = a_l^k(\theta)X_V^l(\theta) + b_l^k(\theta)J^{-1}(K(\theta))Y^l, \qquad (l = 1, \dots, d,\ k = 1, \dots, n).$$

Pairing both sides with $X_V^{l_0}(\theta)$ via the presymplectic form Ω, it follows from (3.2.18), (3.2.19) and (3.2.20) that:

$$b_{l_0}^k(\theta) = \Omega(X_V^{l_0}(\theta), Z_V^k(\theta)) - a_l^k(\theta)\Omega(X_V^{l_0}(\theta), X_V^l(\theta)) = 0. \tag{3.2.27}$$

In general, we have no control on $\Omega(J^{-1}(K(\theta))Y(\theta), Z_V(\theta))$, it means we have no control on the $a_l^k(\theta)$, but the assumption that $V(\theta) := Q(\theta) \cdot M(\theta)$ is non-degenerate guarantees that $\{X(\theta), J^{-1}(K(\theta))Y(\theta), Z(\theta)\}$ is a basis.

Assume from now on that $V(\theta)$, and hence $M(\theta)$, is invertible. Since f is presymplectic and $f(K(\theta)) = K(\theta + \omega)$, it follows from (3.2.18), (3.2.19) and (3.2.20) that:

$$\begin{aligned} Df(K(\theta)).(X(\theta), Z(\theta)) =& (X(\theta+\omega), Z(\theta+\omega)), \\ Df(K(\theta))J^{-1}(K(\theta))Y(\theta) =& X(\theta+\omega)S_1(\theta) + J^{-1}(K(\theta+\omega))Y(\theta+\omega)\,\mathrm{Id} + \\ & + Z(\theta+\omega)A(\theta), \end{aligned}$$

for some matrices $S_1(\theta)$ and $A(\theta)$. This shows that relations (3.2.23) hold. Moving the term $J^{-1}(K(\theta+\omega))Y(\theta+\omega)\,\mathrm{Id}$ to the left side of the second equation we obtain that:

$$A(\theta) = T_3(\theta+\omega)\left[Df(K(\theta))J^{-1}(K(\theta))Y(\theta) - J^{-1}(K(\theta+\omega))Y(\theta+\omega)\right], \tag{3.2.28}$$

where $T_3(\theta)$ is the last row in the matrix:

$$M^{-1}(\theta) = \begin{bmatrix} T_1(\theta) \\ T_2(\theta) \\ T_3(\theta) \end{bmatrix}. \tag{3.2.29}$$

Finally, moving the term $Z(\theta+\omega)A(\theta)$ to the left hand side and pairing both sides with $J^{-1}(K(\theta+\omega))Y(\theta+\omega)$, via the presymplectic form Ω, together with (3.2.18), gives:

$$\begin{aligned} S_1(\theta) = [Y_V(\theta+\omega))^\intercal\, 0] \big[& Df(K(\theta))J^{-1}(K(\theta))Y(\theta) - \\ & J^{-1}(K(\theta+\omega))Y(\theta+\omega) - Z(\theta+\omega)A(\theta)\big]\,. \end{aligned} \tag{3.2.30}$$

■

Remark 3.2.8. A straightforward calculation shows that $V^{-1}(\theta)$ takes the following form:

$$\begin{bmatrix} V_{11}^- & V_{12}^- & V_{13}^- \\ I_d & 0 & 0 \\ V_{31}^- & V_{32}^- & V_{33}^- \end{bmatrix}. \tag{3.2.31}$$

We will need this fact later.

3.3 Main results

3.3.1 A KAM theorem for presymplectic dynamical systems

In this section we will give precise statements of our results. The discussion in the previous section motivates introducing the following definitions:

Definition 3.3.1. We will say that $K(\theta) \in \mathcal{P}_\rho$ is a **non-degenerate torus** if

(i) There exists a $d \times d$-matrix valued function $N(\theta)$ such that

$$N(\theta)(X_V(\theta))^\intercal.(X_V(\theta)) = \mathrm{Id}\,. \tag{3.3.1}$$

(ii) The matrix $V(\theta)$, which is defined in (3.2.21), is invertible,

where $X_V(\theta)$ and $V(\theta)$ are defined in (3.2.14) and (3.2.21) respectively.

Remark 3.3.2. In the symplectic case, the matrix $V(\theta)$ is always non-degenerate when Ω is exact. When Ω is not exact, then one also needs to assume that $V(\theta)$ is invertible in order to perform the Newton iteration successfully. In the presymplectic case, even when Ω is exact, we need to assume that $V(\theta)$ is invertible. Also, we do not know how to proceed with the algorithm presented here if one gives up on exactness of Ω. However, one may still be able to proceed with this algorithm in some special problems where the form is non-exact. Dealing with KAM theory for non exact symplectic forms is a deep problem largely unexplored, see [36] for remarks on the problem of non-exact forms.

Definition 3.3.3. A pair $(f_\lambda, K(\theta))$ is **non-degenerate at** $\lambda = \lambda_0$ if f_λ is a $(2d+n)-$parameter family of presymplectic diffeomorphisms, $K(\theta) \in \mathcal{P}_\rho$ is a non-degenerate torus, and the *average* of the $(2d+n) \times (2d+n)$ matrix

$$\Lambda(\theta) := V^{-1}(\theta)Q(\theta) \left(\left. \frac{\partial f_\lambda}{\partial \lambda} \right|_{\lambda=\lambda_0} (K(\theta)) \right) \tag{3.3.2}$$

has rank $2d+n$, where $V(\theta)$ is defined by (3.2.21).

Remark 3.3.4. Note that the condition is an open condition, so that, if the initial error is small enough, the iterative process does not leave the region where Definition 3.3.3 holds.

We can now state the main theorem of this chapter:

Theorem 3.3.5. *Let $\omega \in D(\gamma, \sigma)$, let f_λ be a $2d+n$-parametric family of analytic presymplectic diffeomorphisms and let $K_0 \in \mathcal{P}_{\rho_0}$. Assume that:*

(H1) The pair (f_λ, K_0) is non-degenerate at $\lambda = \lambda_0$.

(H2) The family f_λ can be holomorphically extended to some complex neighborhood of the image of U_ρ under K:

$$\mathcal{B}_r = \{z \in \mathbb{C} : \sup |z - K(\theta)| < r\}$$

such that $|f_\lambda|_{C^2,\mathcal{B}_r} < \infty$.

If

$$e_0(\theta) := f_{\lambda_0}(K_0(\theta)) - K_0(\theta + \omega),$$

then there exists constant $c > 0$, depending on σ, n, d, ρ_0, r, $|f_{\lambda_0}|_{C^2,\mathcal{B}_r}$, $\|DK_0\|_{\rho_0}$, $\|N_0\|_{\rho_0}$, $\|\frac{\partial f_\lambda}{\partial \lambda}|_{\lambda=\lambda_0}(K_0)\|_{\rho_0}$ and $|\operatorname{avg}(\Lambda_0)^{-1}|$ such that if $0 < \delta_0 < \max(1, \frac{\rho_0}{12})$ and

$$\|e_0\|_{\rho_0} < \min\left\{\gamma^4 \delta_0^{4\sigma}, rc\gamma^2\delta_0^{2\sigma}\|e_0\|_{\rho_0}\right\} \tag{3.3.3}$$

then there exists a mapping $K_\infty \in \mathcal{P}_{\rho_0 - 6\sigma_0}$ and a vector $\lambda_\infty \in \mathbb{R}^{2d+n}$ satisfying

$$f_{\lambda_\infty} \circ K_\infty = K_\infty \circ T_\omega \tag{3.3.4}$$

Moreover, the following inequalities hold:

$$\|K_\infty - K_0\|_{\rho_0 - 6\delta_0} < \frac{1}{c}\gamma^2\delta_0^{-2\sigma}\|e_0\|_{\rho_0} \tag{3.3.5}$$

$$|\lambda_\infty| < \frac{1}{c}\gamma^2\delta_0^{-2\sigma}\|e_0\|_{\rho_0} \tag{3.3.6}$$

Sketch of the proof. More details of the proof will be given later, in Sections 3.4, 3.5, 3.6, but it will be useful to start with a brief overview that can serve as a road map.

We will use a modified Newton method of the type introduced by Moser in [30, 31, 41]. The procedure goes as follows. Starting with

$$G(K_0, 0) := f_0(K_0(\theta)) - K_0(\theta + \omega) = e_0(\theta), \tag{3.3.7}$$

we look for an approximate solution for the corresponding linearized equation

$$\begin{aligned} &DG(K_0, 0)|_{(\Delta_0(\theta),\varepsilon_0)} := \\ &\quad \left.\frac{\partial f_\lambda(K_0(\theta))}{\partial \lambda}\right|_{\lambda=0} \varepsilon_0 + Df_0(K_0(\theta))\Delta_0(\theta) - \Delta_0(\theta + \omega) = -e_0(\theta). \end{aligned} \tag{3.3.8}$$

By an approximate solution we mean up to a quadratic error, i.e., a solution $\Delta_0(\theta)$ such that:

$$\|DG(K_0,0)|_{(\Delta_0(\theta),\varepsilon_0)} + e_0\|_{\rho_0-\delta_0} \leq c_0\gamma^{-3}\delta_0^{-(3\sigma+1)}\|e_0\|_{\rho_0}^2$$

where δ_0, c_0 are constants to be determined later.

Having the solution $(\Delta_0(\theta),\varepsilon_0)$ a better approximating torus for the map f_{λ_1}, where $\lambda_1 = \lambda_0 + \varepsilon_0$, is defined as

$$K_1(\theta) = K_0(\theta) + \Delta_0(\theta)$$

and it will be shown that $(K_1(\theta), f_{\lambda_1})$ is a non-degenerate pair. Furthermore, setting

$$e_1(\theta) := f_{\lambda_1}(K_1(\theta)) - K_1(\theta)$$

we find that

$$\|e_1\|_{\rho_0-\delta_0} \leq c_0\gamma^{-4}\delta_0^{-4\sigma}\|e_0\|_{\rho_0}^2.$$

In other words, for the new torus the error has decreased quadratically.

Iterating this procedure, we will see that the sequence

$$(K_0,\lambda_0),(K_1,\lambda_1),\ldots,(K_n,\lambda_n),\ldots$$

of approximate solutions of (3.2.5), obtained by applying the iterative procedure, converges to a solution $(K_\infty,\lambda_\infty)$. One has to be careful with the domain U_ρ which decreases in each iteration (the reason is because we can bound the correction applied at one step only in a domain slightly smaller than the domain of the original function). This loss of domain can be arranged in a way that, in the limit, one does not end up with an empty domain. This choice of decreasing domains so that there is some domain that remains is very standard in KAM theory since the first papers [22, 30, 31]. See [24, 41] for a pedagogical exposition. ■

Remark 3.3.6. There are two possible ways to extend Theorem 3.3.5 from presymplectic maps to presymplectic flows. One is to use the local uniqueness statement in

Theorem 3.3.7 below. The other one is to proceed with automatic reducibility for flows as in the symplectic case (see [14]). Both of them require some technical work which is still in progress and will not be discussed in this thesis.

3.3.2 Local uniqueness

Notice that if K_∞ is a solution of (3.3.4) then for every $\varphi \in \mathbb{T}^d \times \mathbb{T}^n$ the map $K_\infty(\theta+\varphi)$ is also a solution. For this reason, we will consider $K(\theta)$ and $\hat{K}(\theta) := K(\theta+\varphi)$ to be equivalent. By *uniqueness of solutions*, we will mean uniqueness up to this equivalence relation. The following result gives uniqueness of solutions of (3.3.4):

Theorem 3.3.7. *Let $\omega \in D(\gamma, \sigma)$ and assume that K_1 and K_2 are two non-degenerate tori in $\mathcal{P}_\rho$ solving*

$$f_\lambda(K(\theta)) - K(\theta+\omega) = 0, \tag{3.3.9}$$

such that $K_1(U_\rho) \subset \mathcal{B}_r$ and $K_2(U_\rho) \subset \mathcal{B}_r$. Furthermore, assume that the matrix

$$\Theta := \text{avg}\left(\begin{bmatrix} S_1(\theta) \\ A(\theta) \end{bmatrix}\right),$$

where $S_1(\theta), A(\theta)$ are defined by (3.2.30) *and* (3.2.28) *has rank d. Then there exists a constant $\tilde{c} > 0$ depending on σ, n, d, ρ, r, $|f_\lambda|_{C^2,\mathcal{B}_r}$, $\|DK_1\|_\rho$, $\|N_1\|_\rho$ and $|\Theta|$ such that if*

$$\|K_1 - K_2\|_\rho < \tilde{c}\gamma^2\delta^{2\sigma}, \tag{3.3.10}$$

where $\delta = \frac{\rho}{8}$, then there exists an initial phase $\tau \in \mathbb{T}^d \times \mathbb{T}^n$ such that in $U_{\rho/2}$ one has:

$$K_1 \circ T_\tau = K_2$$

The proof of this result is given in Section 3.7

3.3.3 A vanishing lemma

We end this section with one geometric result. Recall that a diffeomorphism $f : M \to M$ is called **exact presymplectic** if at the level of de Rham cohomology one has:

$$[f^*\alpha - \alpha] = 0 \tag{3.3.11}$$

where α is a primitive of the presymplectic form: $\Omega = \mathrm{d}\alpha$. When M is not compact, one must use compactly supported de Rham cohomology. Clearly, the time-1 map of a Hamiltonian vector field is exact. Moreover, using the flux homomorphism (see [4]), one can show that an exact presymplectic diffeomorphism which is close enough to the identity is the time-1 map of a (time-dependent) Hamiltonian vector field.

We now generalize to exact presymplectic diffeomorphisms the Vanishing Lemma of [16], valid for exact symplectic diffeomorphisms, and which allows one to have some control on the size of the parameter λ. Due to the the presence of kernel, our Vanishing Lemma has a slightly different nature (and statement) than [16, Lemma 4.9].

We will assume that we are in the situation described in the statement of Theorem 3.3.5, where f_0 is exact. In order to simplify the notation we write $K(\theta)$ instead of $K_\infty(\theta)$ and λ instead of λ_∞. Let $\tilde{f}_\lambda := f_\lambda - f_0$ and define the average[3]

$$\bar{\mu} := \int_{\mathbb{T}^{d+n}} \tilde{f}_\lambda(K(\theta))\, \mathrm{d}\theta \in \mathbb{R}^{2d+n} \tag{3.3.12}$$

If we express the vector $\bar{\mu}$ in the basis $\{X(\theta), J^{-1}(K(\theta))Y(\theta), Z(\theta)\}$, we obtain the θ-dependent components $(\mu_1(\theta), ..., \mu_{2d+n}(\theta))$, in other word

$$\bar{\mu} = [\mu_1(\theta), ..., \mu_{2d+n}(\theta)] \left[\begin{array}{ccc} X(\theta) & J^{-1}(K(\theta))Y(\theta) & Z(\theta) \end{array} \right].$$

We have

[3] In the sequel, we will not distinguish between a map with values in $\mathbb{T}^{d+n} \times \mathbb{R}^d$ and a lift with values in $\mathbb{R}^{2d+n}$.

Lemma 3.3.8 (Vanishing Lemma). *If $f_0 : M \to M$ is an exact presymplectic diffeomorphism, then*

$$\int_{\mathbb{T}^{d+n}} \mu_k(\theta) \, \mathrm{d}\theta = 0, \qquad (k = d+1, \dots, 2d). \tag{3.3.13}$$

Proof. We fix the following notations

$$\hat{\theta}_i = (\theta_1, ..., \theta_{i-1}, \theta_{i+1}, ..., \theta_{d+n}) \in \mathbb{T}^{d+n-1}$$

$$\hat{\omega}_i = (\omega_1, ..., \omega_{i-1}, \omega_{i+1}, ..., \omega_{d+n}) \in \mathbb{R}^{d+n-1}$$

and we let $\sigma_{i,\theta_i} : \mathbb{T} \to \mathbb{T}^{d+n}$ be the path given by;

$$\sigma_{i,\theta_i}(\eta) = (\theta_1, ..., \theta_{i-1}, \eta, \theta_{i+1}, ..., \theta_{d+n}).$$

Also, we consider the two-cell $B_{i,\hat{\theta}_i} : [0,1] \times \mathbf{S}^1 \to \mathbb{R}^{2d+n}$ defined by:

$$B_{i,\hat{\theta}_i}(\xi, \eta) := K \circ \sigma_{i,\hat{\theta}_i + \hat{\omega}_i}(\eta) - (\bar{\mu}) \circ \sigma_{i,\hat{\theta}_i + \hat{\omega}_i}(\eta)\xi. \tag{3.3.14}$$

We will compute the integral

$$\int_{B_{i,\hat{\theta}_i}} \Omega$$

in two distinct ways:

(1) The boundary of $B_{i,\hat{\theta}_i}$ is the difference between the two paths $K \circ \sigma_{i.\hat{\theta}_i + \hat{\omega}_i}$ and $(K \circ T_\omega - \bar{\mu}) \circ \sigma_{i,\hat{\theta}_i}$, so by Stokes's theorem we conclude

$$\int_{B_{i,\hat{\theta}_i}} \Omega = \int_{(K \circ T_\omega - \bar{\mu}) \circ \sigma_{i,\hat{\theta}_i}} \alpha - \int_{K \circ \sigma_{i,\hat{\theta}_i + \hat{\omega}_i}} \alpha \tag{3.3.15}$$

Since $(\tilde{f}_\lambda(K(\theta)) - \bar{\mu})$ has average zero and satisfies all hypothesizes of Proposition 3.2.4, there exists an analytic function $v : \mathbb{T}^{d+n} \to \mathbb{R}^{2d+n}$ such that

$$v(\theta) - v(\theta + \omega) = \tilde{f}_\lambda \circ K - \bar{\mu}.$$

This, together with the exactness of f_0 implies that:

$$\begin{aligned}\int_{(K\circ T_\omega-\bar\mu)\circ\sigma_{i,\hat\theta_i}}\alpha &= \int_{(f_\lambda\circ K-\bar\mu)\circ\sigma_{i,\hat\theta_i}}\alpha = \int_{(f_0\circ K+\tilde f_\lambda\circ K-\bar\mu)\circ\sigma_{i,\hat\theta_i}}\alpha\\ &= \int_{K\circ\sigma_{i,\hat\theta_i}} f_0^\star\alpha + \int_{v\circ\sigma_{i,\hat\theta_i}-v\circ\sigma_{i,\hat\theta_i+\hat\omega_i}}\alpha\\ &= \int_{K\circ\sigma_{i,\hat\theta_i}}\alpha + \int_{v\circ\sigma_{i,\hat\theta_i}-v\circ\sigma_{i,\hat\theta_i+\hat\omega_i}}\alpha.\end{aligned}$$

Hence, we see that:

$$\int_{B_{i,\hat\theta_i}}\Omega = \int_{(K+v)\circ\sigma_{i,\hat\theta_i}}\alpha - \int_{(K+v)\circ\sigma_{i,\hat\theta_i+\hat\omega_i}}\alpha.$$

By a simple change of variable, we see that if we integrate over the torus $\mathbb{T}^{d+n-1}$ the right-hand side of the previous equation vanishes, so we can conclude that

$$\int_{\mathbb{T}^{d+n-1}}\mathrm{d}\hat\theta_i\int_{B_{i,\hat\theta_i}}\Omega = 0. \tag{3.3.16}$$

(2) Next we compute the integral of Ω over $B_{i,\hat\theta_i}$ explicitly as follows:

$$\int_{B_{i,\hat\theta_i}}\Omega = \int_0^1\int_0^1 \Omega_{B_{i,\hat\theta_i}(\xi,\eta)}(\partial_\xi B_{i,\hat\theta_i}(\xi,\eta),\partial_\eta B_{i,\hat\theta_i}(\xi,\eta))\ \mathrm{d}\xi\mathrm{d}\eta$$

Since $\bar\mu$ is a constant vector, by (3.3.14) and (3.2.14) we have for $i=1,\dots,d$:

$$\begin{aligned}\partial_\eta B_{i,\hat\theta_i} &= \partial_{\theta_i}K\circ\sigma_{i,\hat\theta_i+\hat\omega_i} = X_i\circ\sigma_{i,\hat\theta_i+\hat\omega_i}\\ \partial_\xi B_{i,\hat\theta_i}, &= -(\bar\mu)\circ\sigma_{i,\hat\theta_i+\hat\omega_i}.\end{aligned}$$

So from the partial presymplectic basis relations (3.2.18), (3.2.19) and (3.2.20) we conclude that:

$$\int_{B_{i,\hat\theta}(\xi,\eta)}\Omega = \int_0^1 \mu_{d+i}\circ\sigma_{i,\hat\theta_i+\hat\omega_i}(\eta)\ \mathrm{d}\eta, \qquad (i=1\dots,d). \tag{3.3.17}$$

Now, (3.3.16) and (3.3.17) together show that:

$$\int_{\mathbb{T}^{d+n-1}}\mathrm{d}\hat\theta_i\int_0^1 \mu_{d+i}\circ\sigma_{i,\hat\theta_i+\hat\omega_i}(\eta)\ \mathrm{d}\eta = 0 \qquad (i=1\dots,d),$$

and this yields the result. ∎

Remarks 3.3.9. The following remarks illustrate the relevance of the Vanishing Lemma:

- The Vanishing Lemma concerns invariant tori. It can be extended to the approximate case, as it is done in [16] for the symplectic case, and assuming that the whole family f_λ is exact presymplectic, it leads to a bound on the parameter, which shows that in every step the value of the parameter decreases with the error term. This can be useful in numerical schemes for finding invariant tori.

- In dimension 2, a volume preserving diffeomorphism of $\mathbf{S}^1 \times \mathbb{R}$ is the same as (pre)symplectic diffeomorphism. In this case, as shown by the proof above, the integral (3.3.13) is the oriented area between a circle and its image by the map, as shown in Figure 3.1. This clearly shows that the vanishing of (3.3.13) is an obstruction for the existence of invariant tori (see also [24]).

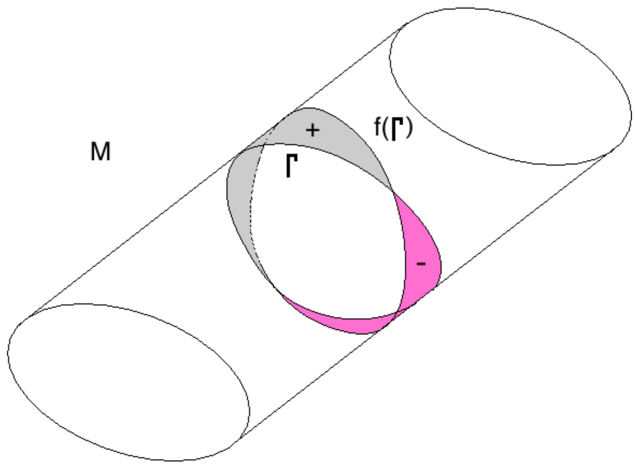

Figure 3.1: Vanishing Lemma

- Recall that we can think of our presymplectic manifold M as $T^*\mathbb{T}^d \times \mathbb{T}^n$. In our Vanishing Lemma we only control the averages in the directions normal to $\mathbb{T}^d$. It is easy to give simple examples of maps satisfying all the assumptions and such that the averages in other directions are non-zero.

3.4 Estimates for the linearized equation

The sketch of the proof of Theorem 3.3.5, given in the previous section, relied on finding an approximate solution of the linearized equation (3.3.8), assuming that one has an approximate solution K_0 of (3.3.7). In this section, we explained how this can be done.

The first claim is that the set $\{X(\theta), J^{-1}(K(\theta))Y(\theta), Z(\theta)\}$ is still a basis for $T_{K_0(\theta)}M$ if the error term is small enough. Note that now, due to the error term, equation (3.2.26) becomes

$$Q(\theta) \cdot M(\theta) = V(\theta) + R(\theta), \tag{3.4.1}$$

where

$$R(\theta) := \begin{bmatrix} X_V^\intercal(\theta)J(\theta)X_V(\theta) & 0 & X_V^\intercal(\theta)J(\theta)Z_V(\theta) \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

If we now use that $K_0(\theta)$ is approximately Lagrangian, i.e., if we apply Lemma 3.2.6, we see that we can control the reminder $R(\theta)$:

Lemma 3.4.1. *Assume the hypotheses of Lemma 3.2.6 hold. Then there exits a constant c_3 depending on d, n, ρ, $|f_\lambda|_{C^1,\mathcal{B}}$, $|J|_{C^1,\mathcal{B}}$, $\|N\|_\rho$, and $\|DK_0\|_\rho$ such that for every $0 < \delta < \frac{\rho}{2}$ we have*

$$\|V^{-1} \cdot R\|_{\rho-2\delta} \leq c_3\gamma^{-1}\delta^{-(\sigma+1)}\|e_0\|_\rho.$$

We conclude that:

Corollary 3.4.2. *Assume the hypotheses of Lemma 3.2.6 hold. If $e_0(\theta)$ satisfies*

$$c_3\gamma^{-1}\delta^{-(\sigma+1)}\|e_0\|_\rho \leq \frac{1}{2}, \tag{3.4.2}$$

then M is invertible and

$$M^{-1}(\theta) = V^{-1}(\theta)Q(\theta) + M_e(\theta),$$

where

$$M_e(\theta) = -[I_{2d+n} + V^{-1}(\theta)R(\theta)]^{-1}V^{-1}(\theta)R(\theta)V(\theta)R(\theta). \tag{3.4.3}$$

Moreover

$$\|M_e\|_{\rho-2\delta} \leq c_4\gamma^{-1}\delta^{-(\sigma+1)}\|e_0\|_\rho, \tag{3.4.4}$$

where c_4 is a constant which depends on the same parameters as c_3.

Proof. A simple application of the Neumman series. See [14]. ∎

We are ready to apply our change of variables. Before that we remark that, since f_{λ_0} is presymplectic, we have

$$Df_{\lambda_0}(K(\theta)) = \begin{bmatrix} F_1(\theta) & 0 \\ F_2(\theta) & F_4(\theta) \end{bmatrix}, \tag{3.4.5}$$

where $F_1(\theta)$ is a symplectic linear map from $V = \pi(T_{K(\theta)}M)$ into itself.

Lemma 3.4.3. *Let $K_0(\theta) \in \mathcal{P}_\rho$ solves*

$$f_{\lambda_0}(K_0(\theta)) - K_0(\theta + \omega) = e_0(\theta)$$

and that $(f_\lambda, K(\theta))$ is non-degenerate at $\lambda = \lambda_0$ in the sense of definition 3.3.3. If $e_0(\theta)$ satisfies (3.4.2)*, then the change of variable $\Delta_0(\theta) = M(\theta)\xi(\theta)$ transforms equation* (3.3.8) *to*

$$\left(\begin{bmatrix} I_d & S(\theta) & 0 \\ 0 & I_d & 0 \\ 0 & A(\theta) & I_n \end{bmatrix} + B(\theta)\right)\xi(\theta) - \xi(\theta+\omega) = \tag{3.4.6}$$

$$- V^{-1}(\theta)Q(\theta)e_0(\theta) - \Lambda(\theta)\varepsilon_0 - M_e(\theta)e_0(\theta) - M_e(\theta)\left(\frac{\partial f_\lambda}{\partial \lambda}\Big|_{\lambda=\lambda_0}\right)\varepsilon_0,$$

where

$$B(\theta) := M^{-1}(\theta+\omega)E(\theta) - \begin{bmatrix} 0 & S_2(\theta) & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\begin{aligned} E(\theta) &:= (D_1 e_0(\theta), E_1(\theta), D_2 e_0(\theta)) \\ E_1(\theta) &:= Df_{\lambda_0}(K_0(\theta))J^{-1}(K_0(\theta))Y(\theta) - X(\theta+\omega)S_1(\theta) + \\ &\qquad - J^{-1}(K_0(\theta))Y(\theta+\omega) - Z(\theta+\omega)A(\theta) \\ S_2(\theta) &:= V_{13}^{-} \cdot \big(F_2(\theta)J^{-1}(K_0(\theta))Y(\theta) - X_N(\theta+\omega)S_1(\theta) - Z_N(\theta+\omega)A(\theta)\big) \\ S(\theta) &:= S_1(\theta) + S_2(\theta), \end{aligned}$$

and $\Lambda(\theta)$, $M_e(\theta)$ *and* $S_1(\theta)$ *are defined by* (3.3.2), (3.4.3) *and* (3.2.30) *respectively. Moreover, we have the estimates:*

$$\|M_e e_0\|_{\rho-2\delta} \le c_4\gamma^{-1}\delta^{-(\sigma+1)}\|e_0\|_\rho^2 \tag{3.4.7}$$

$$\left\| M_e \left.\frac{\partial(f_\lambda \circ K_0)}{\partial\lambda}\right|_{\lambda=0} \varepsilon_0 \right\|_{\rho-2\delta} \le c_4\gamma^{-1}\delta^{-(\sigma+1)} \left\| \left.\frac{\partial(f_\lambda \circ K_0)}{\partial\lambda}\right|_{\lambda=0} \right\|_\rho |\varepsilon_0| \|e_0\|_\rho$$

$$\|B\|_{\rho-2\delta} \le c_5\gamma^{-1}\delta^{-(\sigma+1)}\|e_0\|_\rho \tag{3.4.8}$$

where c_4 *is the same as in* (3.4.4) *and* c_5 *is another constant which depends on the same parameters .*

Proof. The form of the transformed equations follows from substituting the change of variable and elementary computations.

To prove the estimates (3.4.7) and (3.4.8), we note that (3.4.7) follows immediately from (3.4.4), so it only remains to prove (3.4.8). First note that for the first term in the definition of $B(\theta)$ i.e. $M^{-1}(\theta+\omega)E(\theta)$, the Cauchy integral formula provides bounds for $D_1 e_0(\theta)$ and $D_2 e_0(\theta)$ in terms of the error. This enables us to bound $M^{-1}(\theta+\omega)\left(D_1 e_0(\theta), D_2 e_0(\theta)\right)$ by the error term. Calculating bounds for $M^{-1}(\theta+\omega)E_1(\theta)$ is more subtle. By the definition of $A(\theta)$ and the fact that

$$T_3(\theta+\omega)X(\theta+\omega) = 0$$

it follows that

$$T_3(\theta+\omega)E_1(\theta)=0.$$

Therefore:

$$M^{-1}(\theta+\omega)E_1(\theta)=\begin{bmatrix}\begin{bmatrix}T_1(\theta+\omega)\\ T_2(\theta+\omega)\end{bmatrix}E_1(\theta)\\ 0\end{bmatrix}.$$

By the corollary 3.4.2 and the remark 3.2.8 we get:

$$\begin{bmatrix}T_1(\theta)\\ T_2(\theta)\end{bmatrix}=\begin{bmatrix}V_{11}^- & V_{12}^- & V_{13}^-\\ Id & 0 & 0\end{bmatrix}Q(\theta)+\tilde{M}_e(\theta),$$

Where[4]

$$Q(\theta):=\begin{bmatrix}X_V^\intercal(\theta)J(K(\theta)) & 0\\ (J^{-1}(K(\theta))Y(\theta))^\intercal J(K(\theta)) & 0\\ 0\quad 0 & I_n\end{bmatrix},$$

and $\tilde{M}_e(\theta)$ is obtained from $M_e(\theta)$, defined at (3.4.3), by removing the last n rows. So, we have

$$\begin{bmatrix}T_1(\theta+\omega)\\ T_2(\theta+\omega)\end{bmatrix}E_1(\theta)=\overbrace{\begin{bmatrix}\tilde{V}^{-1}(\theta+\omega)\tilde{Q}(\theta+\omega) & \begin{matrix}0\\0\end{matrix}\end{bmatrix}E_1(\theta)}^{(1)}+ \tag{3.4.9}$$

$$+\underbrace{\begin{bmatrix}0 & 0 & V_{13}^-\\ 0 & 0 & 0\end{bmatrix}E_1(\theta)}_{(2)}+\underbrace{\tilde{M}_e(\theta)E_1(\theta)}_{(3)},$$

where we used notations:

$$\tilde{V}^{-1}(\theta):=\begin{bmatrix}V_{11}^- & V_{12}^-\\ I_d & 0\end{bmatrix},$$

$$\tilde{Q}(\theta)=\begin{bmatrix}X_V^\intercal(\theta)J(K(\theta))\\ (J^{-1}(K(\theta))Y(\theta))^\intercal J(K(\theta))\end{bmatrix}.$$

[4] $Q(\theta)$ is defined at (3.2.25), just to make it easier to follow the calculations we restate it again

Note that, by (3.4.4) the term (3) in the right hand side of (3.4.9) is bounded by the error i.e.,

$$\|\tilde{M}_e(\theta)E_1(\theta)\|_{\rho-2\delta} \leq c_6\gamma^{-1}\delta^{-(\sigma+1)}\|e_0\|_\rho,$$

where c_6 depends on c_4 from (3.4.4) and $\|E_1(\theta)\|_\rho$[5]. Considering (3.4.5) and an elementary computation shows that

$$E_1(\theta) = \begin{bmatrix} F_1(\theta)J^{-1}(K(\theta))Y(\theta) - X_V(\theta+\omega)S_1(\theta) - J^{-1}(K(\theta+\omega))Y(\theta+\omega) - Z_V(\theta+\omega)A(\theta) \\ F_2(\theta)J^{-1}(K(\theta))Y(\theta) - X_N(\theta+\omega)S_1(\theta) - Z_N(\theta+\omega)A(\theta) \end{bmatrix}, \tag{3.4.10}$$

substituting (3.4.10) in the term (1) of left hand side of (3.4.9), we get that term (1) is equal to

$$\tilde{V}^{-1}(\theta+\omega)\cdot \begin{bmatrix} X_V^\intercal(\theta+\omega)J(\theta+\omega)E_1^{\mathrm{up}} \\ (J^{-1}(K(\theta+\omega))Y(\theta+\omega))^\intercal J(\theta+\omega)E_1^{\mathrm{up}} \end{bmatrix}, \tag{3.4.11}$$

where E_1^{up} is the upper block of E_1 at (3.4.10). The definition of $S_1(\theta)$, see (3.2.30), and assumption (3.2.15) easily show that the lower block in the equation (3.4.11) is identically zero. The upper block of the equation (3.4.11) is equal to the following term

$$\begin{aligned} \phi(\theta) - \psi(\theta) - X_V^\intercal(\theta+\omega)J(\theta+\omega)X_V(\theta+\omega) + \\ - X_V^\intercal(\theta+\omega)J(\theta+\omega)Z_V(\theta+\omega)A(\theta), \end{aligned} \tag{3.4.12}$$

where

$$\phi(\theta) = (F_1(\theta)X_V(\theta))^\intercal\varphi(\theta)F_1(\theta)J^{-1}(K(\theta))Y(\theta),$$

with $\varphi(\theta) = J(K(\theta+\omega)) - J(f(K(\theta))$ and

$$\psi(\theta) = [F_1(\theta)X_V(\theta) - X_V(\theta)]^\intercal J(\theta+\omega)(F_1(\theta)J^{-1}(K(\theta))Y(\theta)).$$

Both $\varphi(\theta)$ and $F_1(\theta)X_V(\theta) - X_V(\theta)$ are controlled by the error term. This fact and Lemma 3.2.6 show that (3.4.12) is controlled by $\|e_0(\theta)\|_\rho$. Finally the term (2)

[5] As we will see $\|E_1(\theta)\|_\rho$ contains terms that are not bounded by the error, so we do not get quadratic bound by the error as in the symplectic case and the constant depends on $\|E_1(\theta)\|_\rho$ also.

in the left hand side of (3.4.9) is equal to $\begin{bmatrix} S_2(\theta) \\ 0 \end{bmatrix}$ by definition. Since this term is not controlled by the error, we subtract it from $M^{-1}(\theta+\omega)E(\theta)$ to define $B(\theta)$, then we get the bound (3.4.8). We move $S_2(\theta)$ to the coefficients matrix add it to $S_1(\theta)$.

■

Remark 3.4.4. The details to reach expression (3.4.12) are as follows:

$$
\begin{aligned}
&X_V^\intercal(\theta+\omega)J(\theta+\omega)J(\theta+\omega)[F_1(\theta)J^{-1}(K(\theta))Y(\theta) - J^{-1}(K(\theta+\omega))Y(\theta+\omega)] = \\
&= -\underbrace{[F_(\theta)X_V(\theta) - X_V(\theta+\omega)]^\intercal J(\theta+\omega)F_1(\theta)J^{-1}(K(\theta))Y(\theta)}_{\psi} + \\
&\quad + \underbrace{(F_1(\theta)X_V(\theta))^\intercal \overbrace{(J(K(\theta+\omega) - J(f(K(\theta))}^{\varphi} F_1(\theta)J^{-1}(K(\theta))Y(\theta)}_{\phi} + \\
&\quad + \underbrace{(F_1(\theta)X_V(\theta))^\intercal J(f(K(\theta))F_1(\theta)J^{-1}(K(\theta))Y(\theta)}_{(1)} + \\
&\quad - \underbrace{X_V(\theta+\omega)J(\theta+\omega)J^{-1}(K(\theta+\omega))Y(\theta+\omega)}_{(2)}.
\end{aligned}
$$

But we have:

$$
\begin{aligned}
&(1) = \Omega(F_1(\theta)J^{-1}(K(\theta))Y(\theta), F_1(\theta)X_V(\theta)) = \Omega(J^{-1}(K(\theta))Y(\theta), X_V(\theta)) = -I, \\
&(2) = \Omega(J^{-1}(K(\theta+\omega))Y(\theta+\omega), X_V(\theta+\omega)) = -I,
\end{aligned}
$$

so (3.4.12) follows.

We will see that the terms $B(\theta)\xi(\theta)$, $M_e(\theta)e_0(\theta)$ and $M_e(\theta)(\frac{\partial f_\lambda}{\partial \lambda}|_{\lambda=\lambda_0})\varepsilon_0$ have a quadratic dependence on the error $\|e_0(\theta)\|_\rho$, and since we are looking for approximate solution we may omit them. After we omit these terms from (3.4.6) we obtain the linear system:

$$
\begin{bmatrix} I_d & S(\theta) & 0 \\ 0 & I_d & 0 \\ 0 & A(\theta) & I_n \end{bmatrix} \xi(\theta) - \xi(\theta+\omega) = R_0(\theta), \tag{3.4.13}
$$

where

$$R_0(\theta) = -V^{-1}(\theta)Q(\theta)e_0(\theta) - \Lambda(\theta)\varepsilon_0.$$

This linear system can be solved using Proposition 3.2.4, as we show next:

Proposition 3.4.5. *Assume that all hypothesis of Lemma 3.4.3 hold. Then there exists a mapping $\xi(\theta)$, analytic on $U_{\rho-2\delta}$ and a vector $\varepsilon_0 \in \mathbb{R}^{2n}$ such that* (3.4.13) *holds for $\xi(\theta)$ and ε_0. Moreover, there exits c_8 and c_9 depending on n, d, ρ, r, $|f_{\lambda_0}|_{C^2,\mathcal{B}}$, $\|DK_0\|_\rho$, $\|N\|_\rho$, $\left\| \frac{\partial f_\lambda}{\partial \lambda}\Big|_{\lambda=\lambda_0} \right\|_\rho$ such that*

$$\|\xi\|_{\rho-2\delta} \le c_8 \gamma^{-2}\delta^{-2\sigma}\|e_0\|_\rho \tag{3.4.14}$$

$$|\varepsilon_0| \le c_9 |\operatorname{avg}(\Lambda_0)^{-1}| \|e_0\|_\rho \tag{3.4.15}$$

Proof. Since the proof goes through as in the symplectic case, to avoid unnecessary details, we give a short sketch of the proof and refer to [14] for more details. Let

$$R_0(\theta) = \begin{pmatrix} R_x(\theta) \\ R_y(\theta) \\ R_z(\theta) \end{pmatrix}, \quad \xi(\theta) = \begin{pmatrix} \xi_x(\theta) \\ \xi_y(\theta) \\ \xi_z(\theta) \end{pmatrix},$$

so (3.4.13) becomes

$$\begin{cases} \xi_x(\theta) - \xi_x(\theta+\omega) = R_x(\theta) - S(\theta)\xi_y(\theta) \\ \\ \xi_y(\theta) - \xi_y(\theta+\omega) = R_y(\theta) \\ \\ \xi_z(\theta) - \xi_z(\theta+\omega) = R_z(\theta) - A(\theta)\xi_y(\theta) \end{cases} \tag{3.4.16}$$

Using the non-degeneracy of the pair (f_λ, K_λ) at $\lambda = 0$, we can determine $(\varepsilon_0^{d+1}, ..., \varepsilon_0^{2d})$ in such way that $\operatorname{avg}(R_y) = 0$. Then we can apply Proposition 3.2.4 to solve the second equation in (3.4.16) finding a unique zero average solution $\xi_y(\theta)$. After determining $\xi_y(\theta)$ one can choose the remaining components of ε_0 so that

$$\operatorname{avg}(R_x - S\xi_y) = \operatorname{avg}(R_z - A\xi_y) = 0.$$

Applying again Proposition 3.2.4, we solve the first and last equation of (3.4.16) obtaining unique zero average solutions $\xi_x(\theta)$ and $\xi_z(\theta)$. Proposition 3.2.4 shows that these solutions satisfy the following estimates:

$$\|\xi_y\|_{\rho-\delta} \leq c'\gamma^{-1}\delta^{-\sigma}\|R_y\|_\rho$$
$$\|\xi_x\|_{\rho-2\delta} \leq c''\gamma^{-1}\delta^{-\sigma}\|R_x - S\xi_y\|_{\rho-\delta}$$
$$\|\xi_z\|_{\rho-2\delta} \leq c'''\gamma^{-1}\delta^{-\sigma}\|R_z - A\xi_y\|_{\rho-\delta}$$

The proof of the estimates (3.4.15) and (3.4.14) follow just like in the symplectic case (see [14]). ■

Corollary 3.4.6. *Assume all the hypotheses of the proposition* (3.4.5) *hold. then*

$$\|\Delta_0\|_{\rho-2\delta} \leq c\gamma^{-2}\delta^{-2\sigma}\|e_0\|_\rho \tag{3.4.17}$$
$$\|D\Delta_o\|_{\rho-3\delta} \leq c\gamma^{-2}\delta^{-(2\sigma+1)}\|e_0\|_\rho.$$

$$\|DG(K_0,\lambda_0)|_{(\Delta_0(\theta),\varepsilon_0)} + e_0\|_{\rho-2\delta} \leq c_{12}\gamma^{-3}\delta^{-(3\sigma+1)}\|e_0\|_\rho^2, \tag{3.4.18}$$

where $\Delta_0(\theta) = M^{-1}(\theta)\xi(\theta)$.

Proof. The estimates (3.4.17) are immediate consequences of the proposition (3.4.5) and the Cauchy integral formula. Replacing the solution given by Proposition 3.4.5 into the linearized equation (3.3.8) we find that:

$$DG(K_0,\lambda_0)|_{(\Delta_0(\theta),\varepsilon_0)} + e_0(\theta) =$$
$$M(\theta+\omega)\left(B(\theta)\xi(\theta) + M_e(\theta)e_0(\theta) + M_e(\theta)\left.\frac{\partial f_\lambda}{\partial\lambda}\right|_{\lambda=\lambda_0}\varepsilon_0\right),$$

Now (3.4.18) follows from (3.4.7) (3.4.8) (3.4.14) and (3.4.15). This establishes that indeed, we have obtained an approximate solution of the linearized equation (3.3.8). ■

Remark 3.4.7. One of the concerns in the KAM results of the type we are presenting here is how many modifying parameter are needed. A very lucid discussion regarding this matter can be found in [32]. A discussion of the dimension of the space of parameters in the degenerate cases, can be found in [18]. We note that comparing [14, Proposition 8] and Proposition 3.4.5, one sees, that if the family f_λ consists of exact presymplectic diffeomorphisms, then the dimension of parameter space can be reduced by d. Furthermore, if the initial torus satisfies the Kolmogorov [6] non-degeneracy condition [22] i.e. if $\text{avg}(S(\theta))$ is non-singular where $S(\theta)$ is defined in the Lemma 3.4.3, then the dimension of parameter space can be reduced by d again. The reason is that we can choose the averages of the tori as parameters.

In particular, having both families of exact presymplectic mappings and Kolmogorov non-degeneracy condition, it will be enough to consider the parameter space to be n dimensional, see the Vanishing Lemma also.

3.5 Estimates for the improved step

In the previous section, we have shown that the linearized equation (3.3.8) admits approximate solution in a smaller analyticity domains. The estimates blow up if the analyticity loss vanishes. The good point is that they blow up not worse than a power.

The goal of this section is to show that if $\|\Delta_0\|_{\rho-\delta}$ is sufficiently small, the new torus $K_1(\theta) = K_0(\theta)+\Delta_0(\theta)$ has an error in the invariance equation which is quadratically small with respect to the original one (in the smaller domain).

Lemma 3.5.1. *Assume*

$$(K_0 + \Delta_0)(U_{\rho-\delta}, \lambda_0 + \varepsilon_0) \subset Domain(f),$$

where f is defined in (3.2.3)*,then*

$$\|f_{\lambda_0+\varepsilon_0} \circ (K_0 + \Delta_0) - (K_0 + \Delta_0) \circ T_\omega\|_{\rho-\delta} \leq c\gamma^{-2}\delta^{-4\sigma}\|e_0\|_\rho, \tag{3.5.1}$$

[6] It is also known as twist condition

where c now involves $\|f\|_{C^2,\mathcal{B}}$ as well as previous quantities. Furthermore, the pair (f_λ, K_1) is non-degenerate at $\lambda = \lambda_0 + \varepsilon_0$, in the sense of definition 3.3.3.

Note that the linear equation admits estimates for Δ in any domain $U_{\rho-\delta}$ for any $\delta > 0$. If the δ is very small compared with $\|e_0\|_\rho$, the estimates blow up. So that the estimates on the step require some restrictions on the loss of domain δ allowed.

Given the estimates on Δ, ε_0 obtained in Corollary 3.4.2, we see that the requirement on the composition is implied by

$$c\gamma^{-2}\delta^{-(2\sigma+1)}\|e_0\|_\rho \leq \eta \tag{3.5.2}$$

where η is smaller than the distance of $K(U_\rho)$ to the complement of the domain of f.

Proof. This is just a simple consequence of the obvious identity obtained by adding and subtracting some terms:

$$\begin{aligned}
&f_{\lambda_0+\varepsilon_0}(K_0+\Delta_0) - (K_0+\Delta_0)\circ T_\omega = \\
&\quad \underbrace{f_{\lambda_0+\varepsilon_0}(K_0+\Delta_0) - f_{\lambda_0}(K_0) - \frac{\partial f_\lambda}{\partial \lambda}\bigg|_{\lambda=\lambda_0}(K_0)\varepsilon_0 - Df_{\lambda_0}(K_0)\Delta_0}_{(1)} \\
&\quad \underbrace{+ f_{\lambda_0}(K_0) - K_0\circ T_\omega + Df_{\lambda_0}(K_0)\Delta_0 - \Delta_0\circ T_\omega + \frac{\partial f_\lambda}{\partial \lambda}\bigg|_{\lambda=\lambda_0}(K_0)\varepsilon_0}_{(2)}
\end{aligned}$$

The term (1) can be estimated by Taylor theorem, so we have:

$$\begin{aligned}
&\|f_{\lambda_0+\varepsilon_0}(K_0+\Delta_0) - f_{\lambda_0}(K_0) - \frac{\partial f_\lambda}{\partial \lambda}\bigg|_{\lambda=\lambda_0}(K_0)\varepsilon_0 - Df_{\lambda_0}(K_0)\Delta_0\|_{\rho-\delta} \\
&\quad \leq \frac{1}{2}\|f\|_{C^2,\mathcal{B}}(\|\Delta_o\|^2_{\rho-\delta} + |\varepsilon_0|^2) \leq c\frac{1}{2}\|f\|_{C^2,\mathcal{B}}\gamma^{-2}\delta^{-4\sigma}\|e_0\|_\rho
\end{aligned}$$

The term (2) is exactly the left hand side of (3.4.18), so by rearranging the constant we get estimate (3.4.18). Non-degeneracy of the pair (f_λ, K_1) comes from the estimates (3.4.17), (3.4.15) and the fact that non-degeneracy is an open condition. ■

3.6 Iteration of the Newton method and convergence

We shall now perform our modified Newton method, starting with f_{λ_0}, K_0, ω and ρ_0 satisfying the hypotheses of Theorem 3.3.5, and applying at each step the results of Section 3.4. We will see that if we choose $\|e_0\|_{\rho_0}$ small enough we will be able to proceed with the iteration so that the equation

$$f_\lambda(K(\theta)) = K(\theta + \omega) \tag{3.6.1}$$

has a convergent sequence of approximate solutions

$$(K_0, \lambda_0), (K_1, \lambda_1), (K_2, \lambda_2), \ldots$$

defined on domains

$$U_{\rho_0} \supset U_{\rho_1} \supset U_{\rho_2} \supset \cdots$$

with limit an exact solution $(K_\infty, \lambda_\infty)$, defined on a domain U_{ρ_∞}.

Starting with the approximate solution (K_0, λ_0), assume that we have already found the term (K_m, λ_m) in this sequence. The next term will take the form:

$$K_m = K_{m-1} + \Delta_{m-1}(\theta), \quad \lambda_m = \lambda_{m-1} + \varepsilon_{m-1} \quad (m \geq 1),$$

with $(\Delta_{m-1}(\theta), \varepsilon_{m-1})$ an approximate solution of the linear equation

$$DG(K_{m-1}, \lambda_{m-1})|_{(\Delta_{m-1}(\theta), \varepsilon_{m-1})} = -e_{m-1}, \tag{3.6.2}$$

where $e_{m-1} := G(K_{m-1}, \lambda_{m-1})$

The following lemmas are simply restating the lemma 3.5.1 for a general step.

Lemma 3.6.1. *Assume that* (K_{m-1}, λ_{m-1}) *is a non-degenerate (Definition 3.3.3) approximate solution of* (3.6.1) *such that*

$$r_{m-1} := \|K_{m-1} - K_0\|_{\rho_{m-1}} < r. \tag{3.6.3}$$

If $\|e_{m-1}\|_{\rho_{m-1}}$ *is small enough so that Proposition 3.4.5 applies, then for any* $0 < \delta_{m-1} < \rho_{m-1}/3$ *there exist a function* $\Delta_{m-1}(\theta) \in \mathcal{P}_{\rho_{m-1}-3\delta_{m-1}}$ *and* $\varepsilon_{m-1} \in \mathbb{R}^{2d+n}$*, such that*

$$\begin{aligned} &\|\Delta_{m-1}(\theta)\|_{\rho_{m-1}-2\delta_{m-1}} < c_{m-1}\gamma^{-2}\delta_{m-1}^{-2\sigma}\|e_{m-1}\|_{\rho_{m-1}} \\ &\|D\Delta_{m-1}(\theta)\|_{\rho_{m-1}-2\delta_{m-1}} < c_{m-1}\gamma^{-2}\delta_{m-1}^{-2(\sigma+1)}\|e_{m-1}\|_{\rho_{m-1}} \\ &|\varepsilon_{m-1}| \le c_{m-1}|(\mathrm{avg}(\Lambda_{m-1})^{-1}|\|e_{m-1}\|_{\rho_{m-1}} \end{aligned} \tag{3.6.4}$$

where c_{m-1} *is a constant depending on* n*,* d*,* ρ*,* r*,* $|f_{\lambda_{m-1}}|_{C^2,\mathcal{B}_r}$*,* $\|DK_{m-1}\|_\rho$*,* $\|N_{k-1}\|_\rho$ *and* $\left\| \frac{\partial f_\lambda}{\partial \lambda}|_{\lambda=\lambda_{m-1}} \right\|_\rho$*.*

Moreover if

$$r_{m-1} < c_{m-1}\gamma^{-2}\delta_{m-1}^{-2\sigma-1}\|e_{m-1}\|_{\rho_{m-1}} \tag{3.6.5}$$

setting $K_m = K_{m-1} + \Delta_{m-1}$*,* $\lambda_m = \lambda_{m-1} + \varepsilon_{m-1}$*. then,* $e_m(\theta) = G(K_m, \lambda_m)(\theta)$ *the error function of the improved solutions satisfies*

$$\|e_m\|_{\rho_m} \le c_{m-1}\gamma^{-4}\delta_m^{-4\sigma}\|e_{m-1}\|^2_{\rho_{m-1}} \tag{3.6.6}$$

Lemma 3.6.2. *Under the same assumptions as in Lemma 3.6.1, one can improve the constant* c_{m-1} *such that* (3.6.5) *holds and if*

$$c_{m-1}\gamma^{-2}\delta_{m-1}^{-(\sigma+1)}\|e_{m-1}\|_{\rho_{m-1}} \le \frac{1}{2} \tag{3.6.7}$$

then

(i) If $(\pi D_1 K_{m-1})^\intercal \pi D_1 K_{m-1}$ *is invertible with inverse* N_{m-1}*, then the matrix* $(\pi D_1 K_m)^\intercal \pi D_1 K_m$ *is also invertible with inverse* N_m *satisfying*

$$\|N_m\|_{\rho_m} \le \|N_{m-1}\|_{\rho_{m-1}} + c_{m-1}\gamma^{-2}\delta_{m-1}^{-(\sigma+1)}\|e_{m-1}\|_{\rho_{m-1}}; \tag{3.6.8}$$

(ii) If V_{m-1} *is invertible then* V_m *is invertible and the inverse satisfies equation* (3.6.8) *with* N *replaced by* V^{-1}*.*

(iii) If $\operatorname{avg}(\Lambda_{m-1})$ *is invertible then* $\operatorname{avg}(\Lambda_m)$ *is invertible and the inverse satisfies equation* (3.6.8) *with* N *replaced by* $\operatorname{avg}(\Lambda)^{-1}$*.*

(iv) The assumption (3.5.2) *ensurs that the range of* $(K_{m-1} + \Delta_{m-1}, \lambda_{m-1} + \varepsilon_{m-1})$ *is inside of the domain of* f*.*

The most important point is that the constants c_m depend only on n, d, ρ, r, $|f_{\lambda_{m-1}}|_{C^2,\mathcal{B}_r}$, $\|DK_{m-1}\|_\rho$, $\|N_{k-1}\|_\rho$ and $\left\| \frac{\partial f_\lambda}{\partial \lambda}\big|_{\lambda=\lambda_{m-1}} \right\|_\rho$. Hence, when we show that the K does not leave a neighborhood, then, the constants are uniform.

The convergence of the modified Newton method described above is very standard in KAM theory. Indeed, it has sometimes been formulated as an implicit function theorem. Among the many versions of implicit function theorems, the one of [41] is the closest to the problem here. For the sake of completeness, we indicate the main points of the iteration following closely [14, 24] and refer to those papers for more details. One of the main issues to watch out is that the non-degeneracy conditions do not deteriorate much along the iteration and that the assumption (3.5.2), which ensures that we can define the composition, remains valid.

We start by making the choice of the analyticity loss:

$$\rho_m = \rho_{m-1} - 2^{-(m-1)}\delta_0.$$

The most subtle point is to show that the conditions (3.6.3) and (3.6.7) are always satisfied. The first one is to guarantee that the new torus always stays in the domain of the f and the second one is to insure the non-degeneracy condition during the iteration.

The constant c_m depends on the quantities σ, n, d, r, which do not change during the iteration. It also depends on the $\rho_m \le \rho_0$ and the following quantities

$$|f_{\lambda_m}|_{C^2,\mathcal{B}_r}, \|DK_m\|_{\rho_m}, \|N_m\|_{\rho_m}, \|\frac{\partial f_\lambda}{\partial \lambda}\ |_{\lambda=\lambda_m}\ (K_m)\|_{\rho_m}, |\operatorname{avg}(\Lambda_m)^{-1}|.$$

This dependence is polynomial. By similar calculation as follows, can be shown that there exist constant c such that $c_m \le c$ for $m \ge 0$, see [14, Lemma 13]. The main point is that we do not get far away from initial torus. Denote $\epsilon_m = \|e_m\|_{\rho_m}$ with the choice of the domain loss, we obtain :

$$\begin{aligned}\epsilon_m &\leq c\gamma^{-4}(2^{-(m-1)}\delta_0)^{-4\sigma}\epsilon_{m-1}^2 \leq (c\gamma^{-4})^{(1+2)}(2^{-(m-1)-2(m-2)}\delta_0^{(1+2)})^{-4\sigma}\epsilon_{m-2}^4 \quad (3.6.9)\\ &\leq \cdots \leq (c\gamma^{-4}\delta_0^{-4\sigma})^{1+2+\cdots+2^{m-1}}(2^{4\sigma})^{2^0(m-1)+2(m-2)+\cdots+2^{m-2}}\epsilon_0^{2^m}\\ &\leq (c\gamma^{-4}\delta_0^{-4\sigma})^{2^m-1}2^{4\sigma(2^m-m)}\epsilon_0^{2^m} \leq (c\gamma^{-4}\delta_0^{-4\sigma}2^{4\sigma}\epsilon_0)^{2^m-1}2^{-4\sigma(m-1)}\epsilon_0,\end{aligned}$$

where we have used that

$$2^0(m-1)+2(m-2)+\cdots+2^{m-2} = 2^{m-1}\sum_{s=1}^{m-1}s2^{-s} \leq 2^m - m.$$

One sees that if $\|e_0\|_{\rho_0}$ satisfies the assumption (3.3.3), the condition (3.6.7) is always satisfied. It remains to show that (3.6.3) is also satisfied. We denote $\kappa = c\gamma^{-4}\delta_0^{-4\sigma}2^{4\sigma}\epsilon_0$. Now, the first estimate in (3.6.4), estimate (3.6.9) and the definition of r_m gives us:

$$\begin{aligned}r_m &\leq r_{m-1} + c_m\gamma^{-2}\delta_{m-1}^{-2\sigma}\|e_{m-1}\|_{\rho_{m-1}} \leq \cdots \leq c\gamma^{-2}\sigma_0^{-2\sigma}\epsilon_0 + c\gamma^{-2}\sum_{j=1}^{m-1}\delta_j^{-2\sigma}\epsilon_j \quad (3.6.10)\\ &\leq c\gamma^{-2}\sigma_0^{-2\sigma}\epsilon_0 + c\gamma^{-2}\sigma_0^{-2\sigma}\kappa\epsilon_0\sum_{j=1}^{m-1}2^{2j\sigma}2^{-4\sigma(j-1)}\\ &c\gamma^{-2}\sigma_0^{-2\sigma}\epsilon_0\left(1+\kappa 2^{4\sigma}\sum_{j=1}^{\infty}2^{-2j\sigma}\right) = c\gamma^{-2}\sigma_0^{-2\sigma}\epsilon_0\left(1+\kappa\frac{2^{4\sigma}}{2^{2\sigma}-1}\right).\end{aligned}$$

Again having the assumption (3.3.3) and calculations (3.6.10) show that the (3.6.3) is always satisfied.

3.7 Local Uniqueness

The proof of Theorem 3.3.7 follows exactly the same pattern as the proof of uniqueness for the symplectic case given in [14], so we will not reproduce here the same computations. We limit ourselves to some comments and a sketch of the proof, which takes advantage of the fact that in Proposition 3.2.4 two different solutions of (3.2.8) differ by their average. In our situation, one can transfer this difference of averages of two solutions to a difference of the phase between them.

Let K_1 and K_2 be two solutions as in the statement of Theorem 3.3.7. From Taylor's theorem we have:

$$Df_\lambda(K(\theta))(K_2 - K_1) + R(K_1, K_2) = 0 \tag{3.7.1}$$

where

$$\|R(K_1, K_2)\|_\rho \leq c\|K_2 - K_1\|_\rho^2 \tag{3.7.2}$$

Applying the change of variable $M\xi = (K_1 - K_2)$, where M is given by (3.2.17) and replacing K by K_1, the linear equation (3.7.1) is transformed to

$$\xi_x(\theta) - \xi_x(\theta + \omega) = (\tilde{R}(K_1, K_2))_x - S_1(\theta)\xi_y(\theta) \tag{3.7.3}$$

$$\xi_y(\theta) - \xi_y(\theta + \omega) = (\tilde{R}(K_1, K_2))_y \tag{3.7.4}$$

$$\xi_z(\theta) - \xi_z(\theta + \omega) = (\tilde{R}(K_1, K_2))_z - A(\theta)\xi_y(\theta) \tag{3.7.5}$$

where

$$\tilde{R}(K_1, K_2) = -M^{-1}(\theta + \omega)R(K_1, K_2).$$

Using Proposition 3.2.4 and (3.7.2), it follows from (3.7.4) that:

$$\|\xi_y^\perp\|_{\rho-2\delta} \leq c\gamma^{-2}\delta^{-2\sigma}\|\tilde{R}(K_1, K_2)\|_\rho, \tag{3.7.6}$$

where

$$\xi_y^\perp(\theta) = \xi_y(\theta) - \operatorname{avg}(\xi_y)$$

On the other hand, the average of the right hand sides of (3.7.3) and (3.7.5) are zero, so the assumption that Θ (see Theorem 3.3.7) has rank d together with the estimates (3.7.2) and (3.7.6) give us

$$\|(\xi_x, \xi_y, \xi_z) - (\operatorname{avg}(\xi_x), 0, \operatorname{avg}(\xi_z))\|_{\rho-2\delta} \leq c\gamma^{-2}\delta^{-2\sigma}\|K_2 - K_1\|_\rho^2$$

Similarly to [14], this leads to the following lemma.

Lemma 3.7.1. *There exists a constant $\tilde{c}$ depending on d, n, ρ, $|J|_{\mathcal{B}_r}$, $\|K_1\|_{C^2,\rho}$ such that if*

$$\tilde{c}\|K_2 - K_1\|_\rho \leq 1,$$

then one can find $\tau \in \mathbb{R}^{d+n}$ with $|\tau| \leq \|K_2 - K_1\|_\rho$ such that

$$\operatorname{avg}\left(\begin{bmatrix} T_1 \\ T_3 \end{bmatrix} [K_2 \circ T_{\tau_1} - K_1]\right) = 0,$$

where T_1 and T_3 are defined by (3.2.29) *after replacing K by K_1. Therefore, for any $0 < \delta < \rho/2$, we have*

$$\|K_1 \circ T_{\tau_1} - K_2\|_{\rho-2\delta} < \hat{c}\gamma^{-2}\delta^{-2\sigma}\|K_1 - K_2\|_\rho,$$

for a constant $\hat{c}$ depending on the same parameters as $\tilde{c}$ and also on Θ^{-1}.

We then replace K_2 by $K_2 \circ T_{\tau_1}$ and repeat the iteration, which is possible since $K_2 \circ T_{\tau_1}$ is also an invariant torus and the constants $\tilde{c}$ and $\hat{c}$ do not depend on K_2. In this way we produce a convergent sequence of phases $\tau_1, \tau_2, \ldots, \tau_m, \ldots$ such that the limit τ_∞ satisfies:

$$\|K_2 \circ T_{\tau_\infty} - K_1\|_{\rho/2} = 0,$$

therefore completing the proof of Theorem 3.3.7.

A different proof which does not require iteration (but requires setting a normalization condition) appears in [6].

3.8 Comparison between the Poisson and the Presymplectic cases

As we mentioned in the Introduction, our results are not applicable to Hamiltonian dynamical systems on Poisson manifolds. We now justify this statement by presenting a simple example illustrating the difference between Poisson and presymplectic diffeomorphisms.

Consider the presymplectic structure $\Omega = \mathrm{d}x \wedge \mathrm{d}y$ on $M := T^*\mathbb{T} \times \mathbb{T}$ with standard coordinates (x, y, z). This presymplectic structure admits the compatible Poisson structure

$$\Pi = \frac{\partial}{\partial x} \wedge \frac{\partial}{\partial y}.$$

Now a diffeomorphism $f : M \to M$ perserves the Poisson bivector Π if and only if it satisfies the condition:

$$(Df) \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} (Df)^{\intercal} = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \tag{3.8.1}$$

which means that f will have the form

$$f(x, y, z) = (f_1(x, y, z), f_2(x, y, z), f_3(z)),$$

with $\frac{\partial f_1}{\partial x}\frac{\partial f_2}{\partial y} - \frac{\partial f_1}{\partial y}\frac{\partial f_2}{\partial x} = 1$. We remark that (3.8.1) clearly shows that the symplectic leafs of M are invariant under f, a well known fact from Poisson geometry.

On the other hand, a diffeomorphism $f : M \to M$ perserves the presymplectic form Ω if and only if it satisfies the condition:

$$(Df)^{\intercal} \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} (Df) = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

which means that f must have the form

$$f(x, y, z) = (f_1(x, y), f_2(x, y), f_3(x, y, z)), \tag{3.8.2}$$

with $\frac{\partial f_1}{\partial x}\frac{\partial f_2}{\partial y} - \frac{\partial f_1}{\partial y}\frac{\partial f_2}{\partial x} = 1$. In general, there is no canonical way to get a symplectic foliation for a presymplectic manifold. Moreover, even if we consider the symplectic foliation arising from the associated Poisson structure, one can see from (3.8.2) that presymplectic diffeomorphisms, in general, do not preserve symplectic leafs.

Chapter 4

KAM and Lie Algebroids

In this chapter we will point out a different approach to stability problem of the KAM tori in the case of vector fields.

4.1 Lie algebroids and vector fields

Let us start by recalling the definition of a Lie algebroid:

Definition 4.1.1. A **Lie algebroid** over a manifold M consists of a vector bundle A over M together with a Lie algebra bracket $[.,.]$ on the space of $\Gamma(A)$ sections of A and a bundle map $\rho : A \to TM$ (called the **anchor**) satisfying the Leibniz identity

$$[\alpha, f\beta] = f[\alpha, \beta] + \mathcal{L}_{\rho(\alpha)}(f)\beta, \qquad (\alpha, \beta \in \Gamma(A), \quad f \in C^\infty(M)),$$

where we denote by $\mathcal{L}_X$ the Lie derivative along the vector field $X \in \mathfrak{X}(X)$.

Here we are interested in Lie algebroids associated with vector fields, as in the following very simple example:

Example 4.1.2. The Lie algebroid associated to vector field $X \in \mathfrak{X}(M)$ is the line bundle $A := M \times \mathbb{R}$ equipped with the bracket

$$[\alpha, \beta] = \alpha X(\beta) - \beta X(\alpha) \qquad \alpha, \beta \in C^\infty(M, \mathbb{R}) = \Gamma(A),$$

and anchor map

$$\rho : A \to TM$$
$$(m, \lambda) \mapsto \lambda X(m)$$

The anchor map is a Lie algebra homomorphism $(\Gamma(A), [\ ,\ ]_A) \to (\mathfrak{X}(M), [\ ,\ ])$ and its image is an integrable (singular) distribution. Clearly, the leaves of the Lie algebroid associated to a vector field $X \in \mathfrak{X}(M)$ are the orbits of the vector field X. Furthermore, note that for a general Lie algebroid A, given a leaf $L \subset M$ of A, one can restrict the Lie algebroid structure on A to L to obtain a new Lie algebroid $A|_L$, with the bracket determined by:

$$[\alpha_L, \beta_L] = [\alpha, \beta]_L.$$

This Lie algebroid is transitive, i.e., its anchor map is surjective.

A representation of the Lie algebroid A, is a vector bundle E over M endowed with A-derivative operator:

$$\Gamma(A) \otimes \Gamma(E) \to \Gamma(E), \qquad (\alpha, e) \mapsto \nabla_\alpha(e)$$

which satisfy the connection-like identities

$$\nabla_{f\alpha}(e) = f\nabla_\alpha(e), \qquad \nabla_\alpha(fe) = f\nabla_\alpha(e) + \mathcal{L}_{\rho(\alpha)}(f)e, \quad (f \in C^\infty(M))$$

as well as the flatness condition

$$\nabla_{[\alpha,\beta]} = \nabla_\alpha\nabla_\beta - \nabla_\beta\nabla_\alpha.$$

Given a representation E one defines the complex:

$$\Omega^\bullet(A; E) = \Gamma(\wedge^\bullet A^* \otimes E), \tag{4.1.1}$$

with the differential d_A given by the classical Koszul-like formula:

$$\begin{aligned}\mathrm{d}_A\omega(\alpha_1, ..., \alpha_{q+1}) = &\sum_i (-1)^{i+1}\nabla_{\alpha_i}(\omega(\alpha_1, ..., \hat{\alpha}_i, ..., \alpha_{q+1})) \\ &+ \sum_{i<j}(-1)^{i+j}\omega([\alpha_i, \alpha_j], ..., \hat{\alpha}_i, ..., \hat{\alpha}_j, ..., \alpha_{q+1}).\end{aligned} \tag{4.1.2}$$

The resulting cohomology $H^\bullet(A;E)$ is called the A**-de Rham cohomology with coefficients in** E**.** We will be interested on the following example:

Example 4.1.3. If we fix a leaf $L \subset M$ of A, the restricted algebroid $A|_L$ has a canonical representation on the normal bundle $\nu_L = TM/TL$:

$$\nabla_\alpha(X|\text{mod } TL) = [\rho(\alpha), X] \text{ mod } TL. \tag{4.1.3}$$

This representation is known as the **Bott representation** of $A|_L$ and the resulting cohomology $H^\bullet(A_L;\nu_L)$ is the one relevant to the stability of leaves of Lie algebroids.

Remark 4.1.4. Actually, the previous example can be extended to any invariant submanifold $S \subset M$, i.e., any submanifold which is a union of leaves of $A \to M$. Exactly the same formulas define the cohomology $H^\bullet(A_S, \nu_S)$. There is however a fundamental difference between the case of a single leaf L and an arbitrary invariant submanifold S: for a leaf L the Lie algebroid A_L is transitive, so the complex computing the cohomology $H^\bullet(A_L;\nu_L)$ is elliptic and this cohomology is finite dimensional, while this is no long true for an arbitrary submanifold S, where usually $H^\bullet(A_S, \nu_S)$ is infinite dimensional.

4.2 Stability of leaves of Lie algebroids

Let us now turn to the question of stability of leaves of Lie algebroids. In order to state a stability theorem for Lie algebroids, we need to explain first what does it mean to perturb a Lie algebroid, in other words, what is the meaning of "nearby" in the space of Lie algebroids structure on a fixed vector bundle. Fixing coordinates $(U, x_1, ..., x_n)$ on M and a basis of sections $\{e_1, .., e_m\}$ of $A|_U$, we can describe a Lie algebroid $A \to M$ by certain **structure functions** $a_i^\alpha, c_{ij}^k \in C^\infty(U)$ defined by:

$$\rho(e_i) = \sum_\alpha a_i^\alpha \frac{\partial}{\partial x^\alpha}, \qquad [e_i, e_j] = \sum_k c_{ij}^k e_k$$

This makes it possible to compare any two Lie algebroid structure over U by comparing the corresponding structure functions in some C^k-topology. In order to compare them globally, one fixes some connection ∇ on the vector bundle and then define sections $a \in \Gamma(A^* \otimes TM)$ and $c \in \Gamma(\wedge^2 A^* \otimes A)$ by setting:

$$< a, (\alpha, \omega) >=< \rho(\alpha), \omega >, \qquad < c, (\alpha, \beta, \xi) >=< [\alpha, \beta] - \nabla_{\rho(\alpha)}\beta + \nabla_{\rho(\beta)}\alpha, \xi > .$$

The C^k-topology on the sections of a vector bundle now induce a **C^k-topology on the space of Lie algebroid structures** on A. We can now state the stability theorem for Lie algebroids:

Theorem 4.2.1 ([11]). *Let L be a compact n-dimensional leaf of a Lie algebroid A which satisfies $H^1(A_L; \nu_L) = 0$ and let $\kappa > \frac{n}{2}$ be an integer. For any neighborhood V of L there exist a neighborhood $\mathcal{V}$ of the Lie algebroid A in the C^κ-topology such that any Lie algebroid structure in $\mathcal{V}$ has a family of leaves in V, diffeomorphic to L, smoothly parametrized by $H^0(A|_L; \nu_L)$ and depending continuously on the algebroid structure.*

This theorem is a very powerful result which, for example, includes as special cases stability theorems for foliations and for group actions. For more details in this regard and some examples see [11]. Here we will explore only the case of vector fields.

The compact orbits of a vector field are fixed points and periodic orbits. Let us see what Theorem 4.2.1 gives in each of these cases.

Example 4.2.2. Let $X \in \mathfrak{X}(M)$ be some vector field on M and $x_0 \in M$ a fixed point: $X(x_0) = 0$. Then if A is the Lie algebroid associated with X and $L = \{x_0\}$, we have $\nu(L) = T_{x_0}M$ and one checks that:[1]

$$\Omega^0(A_L, \nu(L)) = T_{x_0}M, \quad \Omega^1(A_L, \nu(L)) \simeq T_{x_0}M, \quad \Omega^2(A_L, \nu(L)) = \{0\}.$$

[1]Note that A_L^* is a trivial line bundle.

There is only one non-trivial differential $\mathrm{d} : \Omega^0(A_L, \nu(L) \to \Omega^1(A_L, \nu(L)$, and it follows from Example 4.1.3 that it is given by:

$$\mathrm{d} : T_{x_0}M \to T_{x_0}M, \quad \mathrm{d}v = D_{x_0}X \cdot v,$$

where $D_{x_0}X : T_{x_0}M \to T_{x_0}M$ denotes the linearization of the vector field X at x_0. Hence, we conclude that $H^1(A_L, \nu(L)) = 0$ if and only if this linearization is invertible, in which case we have also that $H^0(A_L, \nu(L)) = 0$. Therefore, in this case, Theorem 4.2.1 states that if a vector field has a non-degenerate fixed point at x_0, then any nearby vector field will also have a nearby zero.

Example 4.2.3. Let $X \in \mathfrak{X}(M)$ be some vector field on M and let $L \subset M$ be a periodic orbit of X. In order to simplify the analysis let us assume that the normal bundle $\nu(L)$ is orientable. Then we parameterize a neighborhood U of $L \simeq \mathbb{S}^1$ so that $U = \mathbb{S}^1 \times \mathbb{R}^{n-1}$ with coordinates $(\theta, x_1, \ldots, x_{n-1})$ and the vector field is given by:

$$X = X_\theta(\theta, x)\frac{\partial}{\partial\theta} + \sum_{i=1}^{n-1} X_i(\theta, x)\frac{\partial}{\partial x_i},$$

with $X_\theta(\theta, 0) = 1$ and $X_i(\theta, 0) = 0$. In these coordinates, the normal bundle $\nu(L) \to L$ can be identified with the trivial bundle $\mathbb{S}^1 \times \mathbb{R}^{n-1} \to \mathbb{S}^1$, and sections of the normal bundle can be identified with vector fields along L of the form $Y(\theta) = \sum_i Y_i(\theta)\frac{\partial}{\partial x_i}$. Then one checks that:

$$\Omega^0(A_L, \nu(L)) = \Gamma(\nu(L)), \quad \Omega^1(A_L, \nu(L)) \simeq \Gamma(\nu(L)),$$
$$\Omega^2(A_L, \nu(L)) = \{0\},$$

and the only non-zero differential $\mathrm{d} : \Omega^0(A_L, \nu(L) \to \Omega^1(A_L, \nu(L)$, is given by:

$$\mathrm{d} : \Gamma(\nu(L)) \to \Gamma(\nu(L)), \quad \mathrm{d}\left(\sum_i Y_i(\theta)\frac{\partial}{\partial x_i}\right) = \sum_i \left(\frac{\partial Y_i}{\partial\theta}(\theta) + \sum_j \frac{\partial X_i}{\partial x_j}(\theta, 0)Y_j(\theta)\right)\frac{\partial}{\partial x_i}$$

Hence, we conclude that $H^1(A_L, \nu(L)) = 0$ if and only if the Floquet exponents are all non-zero, in which case we have also that $H^0(A_L, \nu(L)) = 0$. Therefore, in this case, Theorem 4.2.1 states that if a vector field has a non-degenerate periodic orbit then any nearby vector field will also have a nearby periodic orbit.

4.3 Stability of invariant submanifolds of Lie algebroids

A leaf of a Lie algebroid A is an example of an invariant submanifold, i.e., a manifold which is invariant for all the vector fields in the image of the the anchor map. It is natural to ask wether one has a stability result for invariant submanifolds analogous to Theorem 4.2.1. In fact, we make the following conjecture:

Conjecture 4.3.1. *Let S be a compact n-dimensional invariant submanifold of a Lie algebroid A which satisfies $H^1(A_S; \nu_S) = 0$. Then S is stable: every nearby Lie algebroid structure has a nearby leaf diffeomorphic to S.*

The proof of Theorem 4.2.1 given in [11] does not extend immediately to invariant submanifolds. As we pointed out in Remark 4.1.4, the complex computing the cohomology $H^\bullet(A_S; \nu_S)$ will not be elliptic in general. Hence, in spite of the fact that the formal proof of the theorem still holds for invariant submanifolds, the analysis part of the proof cannot be carried out in the same way, and at this point we do not know if this can be fixed. One possibility would be to use an iteration method of KAM type to prove such a result, but this has not been tried so far, to the best of our knowledge.

Evidence for this conjecture:

1) Families of zeros of a vector field:

Let $M = \mathbb{T} \times \mathbb{R}$, where $\mathbb{T} = \mathbb{S}^1$, and let $X \in \mathfrak{X}(M)$ be a vector field which has a family of fixed points along $S = \mathbb{T} \times \{0\}$. Note that S is an invariant submanifold for the Lie algebroid A associated with our vector field. Let us see what our conjecture predicts in this case.

We choose coordinates (θ, x) in $\mathbb{T} \times \mathbb{R}$ so our vector field is given by:

$$X = X_\theta(\theta, x)\frac{\partial}{\partial \theta} + X_x(\theta, x)\frac{\partial}{\partial x},$$

with $X_\theta(\theta,0)=X_x(\theta,0)=0$. In these coordinates, the normal bundle $\nu(S)\to S$ can be identified with the trivial bundle $\mathbb{S}^1\times\mathbb{R}\to\mathbb{S}^1$, and sections of the normal bundle can be identified with vector fields along L of the form $Y(\theta)=Y(\theta)\frac{\partial}{\partial\theta}$. Then one checks that:

$$\Omega^0(A_L,\nu(L)) = \Gamma(\nu(L)),\quad \Omega^1(A_L,\nu(L))\simeq\Gamma(\nu(L)),$$
$$\Omega^2(A_L,\nu(L))=\{0\},$$

and the only non-zero differential $\mathrm{d}:\Omega^0(A_L,\nu(L)\to\Omega^1(A_L,\nu(L)$, is given by:

$$\mathrm{d}:\Gamma(\nu(L))\to\Gamma(\nu(L)),\quad \mathrm{d}\left(Y(\theta)\frac{\partial}{\partial x}\right)=\frac{\partial X_x}{\partial x}(\theta,0)Y(\theta)\frac{\partial}{\partial x}$$

Hence, we conclude that $H^1(A_L,\nu(L))=0$ if and only if $\frac{\partial X_x}{\partial x}(\theta,0)\neq 0$, in which case we have also that $H^0(A_L,\nu(L))=0$. Therefore, in this case, our conjecture predicts that if a vector field has a non-degenerate circle and the normal derivative of X along the circle is non-zero, then any nearby vector field will also have a nearby invariant circle. This circle can be a family of fixed points, a periodic orbit or some other collection of orbits (e.g., two fixed points and heteroclinic orbits connecting them).

In this case one can check the conjecture directly:

Proposition 4.3.2. *Let $X=X_\theta(\theta,x)\frac{\partial}{\partial\theta}+X_x(\theta,x)\frac{\partial}{\partial x}$ be a vector field on $M=\mathbb{T}\times\mathbb{R}$, such that $X_\theta(\theta,0)=X_x(\theta,0)=0$, so $S_0=\mathbb{T}\times\{0\}$ is a family of fixed points. If $\frac{\partial X_x}{\partial x}(\theta,0)\neq 0$, then any nearby vector field Y has a nearby invariant submanifold S diffeomorphic to S_0.*

We shall not give a proof of this result here. One possible way to proceed is to set up a functional, in a fashion similar to Section 4.4.2, which depends on a vector field $Y=(Y_\theta,Y_x)$, by setting for each $s:\mathbb{T}\to\mathbb{R}$:

$$\Phi_Y(s)=\int_0^{2\pi}\left(Y_x(\theta,s(\theta))-Y_t(\theta,s(\theta))s'(\theta)\right)^2\,\mathrm{d}\theta.$$

Notice that a zero $s(\theta)$ of this functional defines an invariant submanifold of the vector field Y:

$$S=\{(\theta,s(\theta):\theta\in[0,2\pi]\}.$$

When $Y = X$, the condition $\frac{\partial X_x}{\partial x}(\theta, 0) \neq 0$ guarantees that this functional has $s = 0$ as a non-degenerate critical point (a minimum, since the hessian $\mathrm{d}_0\Phi_X$ is positive definite). It follows that for Y close to X, this functional will have a minimum s close to 0. One then shows that this minimum is actually a zero of the functional, so it gives the desired invariant submanifold S of Y close to S_0.

2) Moser's Theorem:

Let $M = \mathbb{T}^n \times \mathbb{R}^m$ with coordinates (x, y) and consider the system of differential equations

$$\begin{cases} \dot{x} = \omega + \varepsilon f(x, y, \varepsilon) \\ \dot{y} = \Omega y + \varepsilon g(x, y, \varepsilon) \end{cases} \tag{4.3.1}$$

where $\omega \in \mathbb{R}^n$ and Ω is a $m \times m$ diagonalizable real matrix with eigenvalues $\mu_1, ..., \mu_n$, we will refer to $\omega_1, ..\omega_n, \mu_1, ..., \mu_m$ as **characteristic numbers**. In the unperturbed case ($\varepsilon = 0$) the $S := \mathbb{T}^d \times 0$ is an invariant torus. We have the following stability result by Moser [32].

Theorem 4.3.3. *Consider the system* (4.3.1) *which possesses $S = \mathbb{T}^d \times 0$ as an invariant torus with characteristic numbers $\omega_1, ..\omega_n, \mu_1, ..., \mu_m$. If these numbers satisfy the following diophantine condition with constants $\gamma, \sigma > 0$,*

$$|(k, \omega) + \Sigma r_l \mu_l| \geq (|k|_{\mathbb{Z}}^{\sigma} + 1)^{-1}\gamma, \tag{4.3.2}$$

for all integer vectors k and all r_l with

$$|\Sigma r_i| \leq 1, \qquad \Sigma |r_l| \leq 2$$

except the finitely many $(k, r) = (0, r)$ for which the left hand side of (4.3.2) *vanishes. Then there exist unique analytic functions $\lambda_1(\varepsilon), \lambda_2(\varepsilon), M(\varepsilon)$ satisfying:*

$$\Omega\lambda_2 = 0, \qquad \Omega M = M\Omega$$

such that the modified system

$$\begin{cases} \dot{x} = \omega + \varepsilon f(x, y, \varepsilon) + \lambda_1 \\ \dot{y} = \Omega y + \varepsilon g(x, y, \varepsilon) + \lambda_2 + M \end{cases} \tag{4.3.3}$$

possesses an invariant torus with the same characteristic numbers as the unperturbed system.

Remark 4.3.4. In the case of nearly integrable systems, the non degeneracy condition imposed on the frequency map makes it possible to find an invariant torus in the perturbed system with a fixed frequency. This is not the case for general vector field. In order to fix this problem one needs to consider modifying parameters. Note that if all the eigenvalues of Ω are non zero then the modifying term λ_2 in Theorem (4.3.3) is zero.

Let us now turn to the Lie algebroid interpretation of this result. The torus $S = \mathbb{T}^n \times 0$ is an invariant submanifold for the Lie algebroid associated to the vector field $X = \sum_i \omega_i \frac{\partial}{\partial x_i} + \Sigma_{j,k} \Omega_{j,k} y^k \frac{\partial}{\partial y_j}$, introduced in Example 4.1.2. A straight forward calculation shows that the Bott representation for the restricted algebroid $A|_S$, defined by (4.1.3), is:

$$\nabla : \Gamma(A|_S) \otimes \Gamma(\nu(S)) \to \Gamma(\nu(S))$$
$$(\alpha(x), e(x)) \mapsto \alpha(x)((\sum_i \omega_i \frac{\partial}{\partial x_i})I_m - \Omega)e(x).$$

Using this representation, we obtain the complex $\Omega^\bullet(A|_S, \nu(S))$:

$$0 \xrightarrow{0} \Omega^0(A|_S, ; \nu(S)) \xrightarrow{d} \Omega^1(A|_S, ; \nu(S)) \xrightarrow{0} 0, \tag{4.3.4}$$

(since $A|_S$ is a line bundle, there are only terms in degree 0 and 1). Furthermore

$$\Omega^0(A|_S; \nu(S)) \simeq \Omega^1(A|_S; \nu(S)) \simeq C^\infty(\mathbb{T}^n, \mathbb{R}^m),$$

and

$$d(e(x)) = (\sum_i \omega_i \frac{\partial}{\partial x_i} \cdot I_m - \Omega)e(x).$$

One can characterize the smooth functions on the torus by their Fourier expansions (see, e.g., [37, chapter 3]): we set

$$s(\mathbb{Z}) = \left\{ u : \mathbb{Z}^n \to \mathbb{R} : \sup_{k \in \mathbb{Z}^n} (1 + |k|^2_{\mathbb{Z}})^N |u(k)|^2 < \infty,\ \forall N \in \mathbb{N} \cup \{0\} \right\},$$

so that the map:

$$\mathcal{F} : C^\infty(\mathbb{T}^n, \mathbb{R}) \to s(\mathbb{Z}),$$
$$g(k) \longmapsto \hat{g}(k) = \frac{1}{(2\pi)^n} \int_{\mathbb{T}^n} g(x) e^{-\mathfrak{i}<k,x>} dx,$$

is both injective and surjective. Theorem (4.3.3) together with following result support our conjecture.

Proposition 4.3.5. *If Ω is diagonalizable with eigenvalues $\Omega_1, ..., \Omega_m$ satisfying the diophantine condition*

$$|\mathfrak{i} < k, \omega > -\Omega_\mu| \geq \gamma(1 + |k|^2_{\mathbb{Z}})^{-\tau}, \quad \forall k \in \mathbb{Z}^n \,,\, \mu = 1, \cdots, m, \tag{4.3.5}$$

then the cohomology groups $H^i(A|_S; \nu(S))$ $i = 0, 1$ vanish.

Proof. Clearly:

- $H^0(A|_S; \nu(S)) = 0$ if and only if $\mathrm{d}(e(x)) = 0$ gives us that $e(x) = 0$
- $H^1(A|_S; \nu(S)) = 0$ if and only if for every $h(x) \in C^\infty(\mathbb{T}^n, \mathbb{R}^m)$ there exists $e(x) \in C^\infty(\mathbb{T}^n, \mathbb{R}^m)$ such that

$$\mathrm{d}(e(x)) = h(x). \tag{4.3.6}$$

By complexifying, we can consider complex valued maps on torus instead of real valued ones. Since Ω is diagonalizable, after a change of coordinates, we can rewrite equation (4.3.6) in the form:

$$\sum_i \omega_i \frac{\partial}{\partial x_i} e_\mu(x) - \Omega_\mu e_\mu(x) = h_\mu(x) \quad \mu = 1, \cdots, m, \tag{4.3.7}$$

where $\Omega_1, \ldots, \Omega_m$ denote the (possible complex) eigenvalues of Ω. Passing to Fourier expansions, we obtain that the Fourier coefficients of the solutions are given by:

$$\hat{e}_\mu(k) = \frac{\hat{h}_\mu(k)}{\mathfrak{i} < k, \omega > -\Omega_\mu} \tag{4.3.8}$$

Now the Diophantine condition (4.3.5) implies that $\mathfrak{i} < k, \omega > -\Omega_\mu \neq 0$ for every k. Hence, the coefficients $\hat{e_\mu}(k)$ are well-defined and (4.3.8) implies immediately that if $\mathrm{d}(e(x)) = 0$ then $e(x) = 0$, i.e., that $H^0(A|_L, \nu(L)) = 0$.

On the other hand, the Diophantine condition (4.3.5) implies also that:

$$\sup_{k\in\mathbb{Z}^n}(1+|k|_{\mathbb{Z}}^2)^N|\hat{e}_\mu(k)|^2 = \sup_{k\in\mathbb{Z}^n}(1+|k|_{\mathbb{Z}}^2)^N\left|\frac{\hat{h}_\mu(k)}{\mathfrak{i} < k,\omega > -\Omega_\mu}\right|^2$$
$$< \sup_k(1+|k|_{\mathbb{Z}}^2)^{N+2\tau}|\hat{h}_\mu(k)|^2\gamma^2 < \infty,$$

which shows that e_μ is smooth. We conclude that (4.3.6) has a smooth solution $e(x)$, for every smooth $h(x)$, i.e., that $H^1(A|_S;\nu(S)) = 0$ as well. ■

Remark 4.3.6. Since we are working over the torus, Proposition 4.3.5 holds also if we replace the space of smooth functions $C^\infty(\mathbb{T}^n)$ by some Sobolev space $H_r(\mathbb{T}^n)$. In fact (see [37]), we have that $H_r(\mathbb{T}^n)$ is isomorphic to the Hilbert space of Fourier series:

$$\ell^{2,r} = \{\{c_k\}, k\in\mathbb{Z}^n|\sum_k|c_k|^2.(1+|k|_{\mathbb{Z}}^2)^r < \infty\},$$

with the weighted version of the usual hermitian inner product:

$$< \{c_k\},\{d_k\} >_r = \sum_k c_k\bar{d}_k.(1+|k|_{\mathbb{Z}}^2)^r.$$

4.4 Lie algebroid stability versus stability for vector fields

One can hope to use the general stability theorem for *orbits* of Lie algebroids to prove stability of invariant submanifolds of vector fields. We describe two possible approaches and explain why they fail.

4.4.1 Stability via the image of the anchor map

Given a vector field X_0 with an invariant submanifold S_0, assume that one can find a Lie algebroid A_0 such that X_0 is in the image of the anchor map and S_0 is a stable orbit of A_0. Then, if any vector field X close to X_0 is also in the image of the anchor of some Lie algebroid A close to A_0, it would follow that S_0 is also stable as an invariant submanifold of X_0: there would exist close by invariant submanifolds S of X, diffeormorphic to S_0. We give here an example that shows that this method in general will fail, since one cannot guarantee that close by vector fields are in the image of a close by Lie algebroid.

Let $A := \mathbb{R}^3 \times \mathfrak{so}(3)$ be the trivial vector bundle over $M := \mathbb{R}^3$ with the action Lie algebroid structure D_0 defined as follows. Sections of A can be considered as functions $\alpha : \mathbb{R}^3 \to \mathfrak{s}o(3)$, the Lie bracket between two sections α, β is:

$$[\alpha, \beta](p) = [\alpha(p), \beta(p)]_{\mathfrak{s}o(3)} + (v_{\beta(p)}.\alpha)(p) - (v_{\alpha(p)}.\beta)(p)$$

where $v_\xi(x) = \frac{\mathrm{d}}{\mathrm{d}t}|_{t=0}(\exp(t\xi)x) \in T_x\mathbb{R}^3$ is the infinitesimal vector field associated with $\xi \in \mathfrak{s}o(3)$ and dot stands for the action of the vector fields on the sections considered as functions of $\mathbb{R}^3$. The anchor map of D_0 is the vector bundle morphism

$$\rho_{D_0} : A \to T\mathbb{R}^3, \quad (\xi, p) \mapsto v_\xi(p).$$

The lie algebra $\mathfrak{s}o(3)$ has the basis

$$J_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \; J_y = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}, \; J_z = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

and we have the bracket relations:

$$[J_x, J_y] = J_z\,,\; [J_y, J_z] = J_x\,,\; [J_z, J_x] = J_y,$$

while for the anchor we obtain:

$$v_{J_x} = (y\partial_z - z\partial_y)\,,\; v_{J_y} = (z\partial_x - x\partial_z)\,,\; v_{J_z} = (x\partial_y - y\partial_x).$$

Proposition 4.4.1. *For $A = \mathbb{R}^3 \times \mathfrak{so}(3)$ with D_0 the action Lie algebroid structure described above, the origin of $\mathbb{R}^3$ is a stable leaf of A. Moreover, one can find a vector field v_ϵ arbitrary close to v_{J_x} and a neighborhood $\mathcal{V}$ of D_0 such that v_{J_x} is not in the image of the anchor of any $D \in \mathcal{V}$.*

Proof. For the leaf $L := \{(0,0,0)\}$, we have $\nu_L = T_L\mathbb{R}^3/TL = T_L\mathbb{R}^3$ and

$$\Omega^i(A|_L;\nu_L) = \Gamma(\wedge^i A^*|_L \otimes \nu_L) = \{(g,r) | g \in \wedge^i \mathfrak{so}^*(3), r \in \mathbb{R}^3\}.$$

The Bott representation defined by (4.1.3) is:

$$\nabla : \Gamma(A|_L) \otimes \Gamma(\nu_L) \to \Gamma(\nu_L), \quad (\alpha(0), X(0)) \mapsto [v_{\alpha(0)}, X](0)$$

where $\alpha : \mathbb{R}^3 \to \mathfrak{so}(3)$ and $X \in \mathfrak{X}(\mathbb{R}^3)$. Using the basis J_x^*, J_y^*, J_z^* for $\mathfrak{so}^*(3)$ and $\{\partial_x, \partial_y, \partial_z\}$ for $T\mathbb{R}^3$, we have

$$\Omega^0(A|_L;\nu_L) = \mathbb{R}^3\,,\ \Omega^1(A|_L;\nu_L) = M(3,\mathbb{R})\,,\ \Omega^2(A|_L;\nu_L) = M(3,\mathbb{R})$$

where

$$(a_1\partial_x + a_2\partial_y + a_3\partial_z) \to \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix},$$

$$(a_{11}J_x^*\otimes\partial_x + a_{12}J_y^*\otimes\partial_x + a_{13}J_y^*\otimes\partial_x + a_{21}J_x^*\otimes\partial_y + \cdots + a_{33}J_z^*\otimes\partial_z) \to \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

$$(b_{11}J_x^* \wedge J_y^* \otimes \partial_x + \cdots + b_{33}J_y^* \wedge J_z^* \otimes \partial_z) \to (b_{ij})\,.$$

By some straight forward calculations one concludes that:

$$\mathrm{d}_A : \Omega^0(A|_L;\nu_L) \to \Omega^1(A|_L;\nu_L)$$

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \mapsto \begin{pmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{pmatrix}$$

and

$$\mathrm{d}_A : \Omega^1(A|_L; \nu_L) \to \Omega^2(A|_L; \nu_L)$$

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \mapsto \begin{pmatrix} a_{31}+a_{13} & -a_{21}-a_{12} & a_{11}-a_{22}-a_{33} \\ a_{32}+a_{23} & a_{11}-a_{22}+a_{33} & a_{12}+a_{21} \\ -a_{11}-a_{22}+a_{33} & -a_{23}-a_{32} & a_{13}+a_{31} \end{pmatrix}.$$

The above formulas show that $H^1(A|_L; \nu_L) = 0$ so, by Theorem 4.2.1, the leaf L is stable.

Now consider the vector field $X_\epsilon = (y\partial_z - z\partial_y + \epsilon\partial_x)$. For arbitrary $\epsilon \neq 0$ the vector field X_ϵ has no fixed point, so there is no Lie algebroid structure D in the neighborhood $\mathcal{V}$ given by Theorem 4.2.1 in which $X_\epsilon \in \mathrm{Im}(\rho_D)$. ■

4.4.2 Stability of tori of Hamiltonian vector fields via Lie algebroids

Let us consider now the more restrictive class of vector fields which define nearly integrable systems. We let $M = \mathbb{T}^n \times \mathbb{R}^n$, with coordinates (x, y), equipped with canonical symplectic form $\Omega = \sum_{i=1}^n \mathrm{d}x \wedge \mathrm{d}y$ and we consider nearly integrable Hamiltonian system over this manifold defined by Hamiltonian vector fields associated with

$$H(x, y) = H_0(y) + \varepsilon H_1(x, y).$$

For any $y_0 \in \mathbb{R}^n$, the torus $\mathbb{T}^n \times \{y_0\}$ is an invariant torus for X_{H_0} with frequency $\omega(y_0) = \frac{\partial H_0}{\partial y}(y_0)$. We fix one such invariant torus, say the one at $y_0 = 0$, and look for nearby invariant tori of X_H, for H close to H_0. These are necessarily of the form:

$$T_s := \{(\theta, s(\theta)) | s : \mathbb{T}^n \to \mathbb{R}^n\}.$$

Furthermore, we will ask that the flow of X_H restricted to T_s is conjugate to a linear one.

Proposition 4.4.2. *A Lagrangian section $s : \mathbb{T}^d \to \mathbb{R}^n$ defines an invariant torus of X_H if and only if*

$$\mathrm{d}H(\theta, s(\theta)) = 0.$$

Proof. The torus T_s is invariant under the flow of $X_H = \sum_{i=1}^n \frac{\partial H}{\partial y^i}\frac{\partial}{\partial x_i} - \frac{\partial H}{\partial x^i}\frac{\partial}{\partial y^i}$ if and only if

$$X_H(\theta, s(\theta)) \in T_{(\theta,s(\theta))}T_s = \left\langle \frac{\partial}{\partial x_i} + \sum_{j=1}^n \frac{\partial s^j}{\partial \theta^i}\frac{\partial}{\partial y^j},\ i = 1, ..., n \right\rangle,$$

a simple substitution yields:

$$\frac{\partial H}{\partial x^j}(\theta, s(\theta)) + \sum_{i=1}^n \frac{\partial H}{\partial y^i}(\theta, s(\theta))\frac{\partial s^i}{\partial \theta^j}(\theta) = 0, \quad (j = 1, .., n),$$

therefore finishing the proof. ∎

Since we are interested in invariant tori where the motion is linear, we will also ask that the frequency vector $\omega_s(\theta) = (\frac{\partial H}{\partial y^1}(\theta, s(\theta)), ..., \frac{\partial H}{\partial y^n}(\theta, s(\theta)))$ is independent of θ, i.e., that $\mathrm{d}\omega_s(\theta) = 0$. In local coordinates, this condition reads:

$$\frac{\partial^2 H}{\partial y^j \partial x^i} + \sum_{k=1}^n \frac{\partial^2 H}{\partial y^j \partial y^k}\frac{\partial s^k}{\partial \theta^i} = 0 \quad i, j = 1, ..., n$$

We conclude that our problem consists in showing that if H is close to H_0 there are "small" sections $s : \mathbb{T}^d \to \mathbb{R}^n$ satisfying:

$$\mathrm{d}H(\theta, s(\theta)) = 0, \quad \mathrm{d}\omega_s(\theta) = 0.$$

One can now try to follow the same method used in Lie algebroid case in the proof of Theorem 4.2.1. One defines the functional:

$$\Phi_H : H_r(\mathbb{T}^n, \mathbb{R}^n) \to \mathbb{R} \quad s \mapsto \|\mathrm{d}H(\theta, s(\theta))\|_r^2 + \|\mathrm{d}\omega_s(\theta)\|_r^2,$$

where $H_r(\mathbb{T}^n, \mathbb{R}^n)$ is the r^{th}-Sobolev space of functions $s : \mathbb{T}^n \to \mathbb{R}^n$ which satisfy the closeness condition:

$$\frac{\partial s^i}{\partial \theta^j}(\theta) = \frac{\partial s^j}{\partial \theta^i}(\theta), \quad (i \neq j).$$

One looks for zeros of the functional Φ_H. The proof of Theorem 4.2.1 suggest breaking the proof into two steps:

(i) Φ_{H_0} has a strongly non-degenerate critical point at $s = 0$, so every functional Φ_H for H close enough to H_0, has a critical point, close to zero.

(ii) Critical points of Φ_H close to zero, where H is close enough to H_0, are also zeros of Φ_H.

For the first step, one has:

Proposition 4.4.3. *Assume that the Hessian matrix* $\frac{\partial^2 H_0}{\partial y^i \partial y^j}(0)$ *is non-degenerate. Then* $s = 0$ *is a strongly non-degenerate critical point of* Φ_{H_0} *for most values of the frequency* ω_0.

Proof. One checks that $\mathrm{d}_0^2\Phi_{H_0} : H_r(\mathbb{T}^n, \mathbb{R}^n) \times H_r(\mathbb{T}^n, \mathbb{R}^n) \to \mathbb{R}$ is the sum of two quadratic forms:

$$(f, g) \longmapsto \left\langle \sum_j \omega_0^j \frac{\partial f}{\partial x^j}(x), \sum_j \omega_0^j \frac{\partial g}{\partial x^j}(x) \right\rangle_r +$$

$$+ \left\langle \sum_j \frac{\partial^2 H_0}{\partial y^i \partial y^j}(0) \frac{\partial f}{\partial x^j}(x), \sum_j \frac{\partial^2 H_0}{\partial y^i \partial y^j}(0) \frac{\partial g}{\partial x^j}(x) \right\rangle_r .$$

The second quadratic form is non-degenerate if $\frac{\partial^2 H_0}{\partial y^i \partial y^j}(0)$ is non-degenerate. Hence, for most values of the frequency ω_0 we will have that $\mathrm{d}_0^2\Phi_{H_0}$ is also non-degenerate. ■

Unfortunately, the method used in the Lie algebroid case to prove step (ii) does not seem to work in this case.

Bibliography

[1] Vladimir I. Arnold, Valery V. Kozlov, and Anatoly I. Neishtadt, *Mathematical aspects of classical and celestial mechanics*, 3rd ed., Encyclopaedia of Mathematical Sciences, vol. 3, Springer-Verlag, Berlin, 2006. [Dynamical systems. III]; Translated from the Russian original by E. Khukhro. MR2269239 (2008a:70001)

[2] Michèle Audin, *Hamiltonian systems and their integrability*, SMF/AMS Texts and Monographs, vol. 15, American Mathematical Society, Providence, RI, 2008. Translated from the 2001 French original by Anna Pierrehumbert; Translation edited by Donald Babbitt. MR2440371 (2009i:37143)

[3] Artur Avila and Alejandro Kocsard, *Cohomological equations and invariant distributions for minimal circle diffeomorphisms*, Duke Math. J. **158** (2011), no. 3, 501–536, DOI 10.1215/00127094-1345662. MR2805066 (2012f:37087)

[4] Augustin Banyaga, *The structure of classical diffeomorphism groups*, Mathematics and its Applications, vol. 400, Kluwer Academic Publishers Group, Dordrecht, 1997. MR1445290 (98h:22024)

[5] G.B. Byrne, F.A. Hagger, and G.R.W. Guispel, *Sufficient conditions for dynamical systems to have presymplectic or pre-implectic structure*, Phys. A **269** (1999), no. 1-2, 99–129, DOI 10.1016/S0378-4371(99)00093-X.

[6] Renato Calleja and Rafael de la Llave, *A numerically accessible criterion for the breakdown of quasi-periodic solutions and its rigorous justification*, Nonlinearity **23** (2010), no. 9, 2029–2058, DOI 10.1088/0951-7715/23/9/001. MR2672635 (2011i:37081)

[7] Ana Cannas da Silva, *Lectures on symplectic geometry*, Lecture Notes in Mathematics, vol. 1764, Springer-Verlag, Berlin, 2001. MR1853077 (2002i:53105)

[8] J. F. Cariñena, L. A. Ibort, J. Gomis, and N. Román-Roy, *Applications of the canonical-transformation theory for presymplectic systems*, Nuovo Cimento B (11) **98** (1987), no. 2, 172–196, DOI 10.1007/BF02721479 (English, with Italian and Russian summaries). MR900784 (88m:58049)

[9] J. F. Cariñena, J. Gomis, L. A. Ibort, and N. Román, *Canonical transformations theory for presymplectic systems*, J. Math. Phys. **26** (1985), no. 8, 1961–1969, DOI 10.1063/1.526864. MR796226 (86j:58041)

[10] Luigi Chierchia, *Kolmogrov-Arnold-Moser Theory*, Encyclopedia of Complexity and Systems Science. Editor-in-chief: Meyers, Robert A (2009), 5064-5091.

[11] Marius Crainic and Rui Loja Fernandes, *Stability of symplectic leaves*, Invent. Math. **180** (2010), no. 3, 481–533, DOI 10.1007/s00222-010-0235-1. MR2609248 (2011i:53130)

[12] Marius Crainic and Ieke Moerdijk, *Deformations of Lie brackets: cohomological aspects*, J. Eur. Math. Soc. (JEMS) **10** (2008), no. 4, 1037–1059, DOI 10.4171/JEMS/139. MR2443928 (2010f:58024)

[13] M. Delgado-Téllez and A. Ibort, *A panorama of geometrical optimal control theory*, Extracta Math. **18** (2003), no. 2, 129–151. MR2002442

[14] R. de la Llave, A. González, À. Jorba, and J. Villanueva, *KAM theory without action-angle variables*, Nonlinearity **18** (2005), no. 2, 855–895, DOI 10.1088/0951-7715/18/2/020. MR2122688 (2005k:37131)

[15] Manuel de León and Paulo R. Rodrigues, *Methods of differential geometry in analytical mechanics*, North-Holland Mathematics Studies, vol. 158, North-Holland Publishing Co., Amsterdam, 1989. MR1021489 (91c:58041)

[16] Ernest Fontich, Rafael de la Llave, and Yannick Sire, *Construction of invariant whiskered tori by a parameterization method. I. Maps and flows in finite dimensions*, J. Differential Equations **246** (2009), no. 8, 3136–3213, DOI 10.1016/j.jde.2009.01.037. MR2507954 (2010c:37139)

[17] J. Gomis, J. Llosa, and N. Román, *Lee Hwa Chung theorem for presymplectic manifolds. Canonical transformations for constrained systems*, J. Math. Phys. **25** (1984), no. 5, 1348–1355, DOI 10.1063/1.526303. MR745746 (85k:58030)

[18] Alejandra González-Enríquez, Àlex Haro, and Rafael de la Llave, *Singularity theory for non–twist KAM tori*, Memoirs of the AMS (2012). MP_ARC #11-83.

[19] A. N. Kolmogorov, *On conservation of conditionally periodic motions for a small change in Hamilton's function*, Dokl. Akad. Nauk SSSR (N.S.) **98** (1954), 527–530 (Russian). MR0068687 (16,924c)

[20] H. P. Künzle, *Dynamics of a rigid test body in curved space-time*, Comm. Math. Phys. **27** (1972), 23–36. MR0303904 (46 #3040)

[21] ———, *Canonical dynamics of spinning particles in gravitational and electromagnetic fields*, J. Mathematical Phys. **13** (1972), 739–744. MR0299139 (45 #8188)

[22] A. N Kolmogorov, *Preservation of Conditionally Periodic Movements with SmallChange in the Hamiltonian Functions*, Akad. Nauk. Dokl. **98** (1954), 527-530.

[23] M.J. Gotay, J.M. Nester, and G. Hinds, *Presymplectic manifolds and the Dirac-Bergman theory of constraints*, J. Math. Phys. **19** (1978), no. 11, 2388–2399, DOI 10.1063/1.524793.

[24] Rafael de la Llave, *A tutorial on KAM theory*, Smooth ergodic theory and its applications (Seattle, WA, 1999), Proc. Sympos. Pure Math., vol. 69, Amer. Math. Soc., Providence, RI, 2001, pp. 175–292. MR1858536 (2002h:37123)

[25] Amadeu Delshams, Rafael de la Llave, and Tere M. Seara, *Orbits of unbounded energy in quasi-periodic perturbations of geodesic flows*, Adv. Math. **202** (2006), no. 1, 64–188, DOI 10.1016/j.aim.2005.03.005. MR2218821 (2007a:37070)

[26] P. A. M. Dirac, *Forms of relativistic dynamics*, Rev. Modern Physics **21** (1949), 392–399. MR0033248 (11,409i)

[27] ———, *Generalized Hamiltonian dynamics*, Canadian J. Math. **2** (1950), 129–148. MR0043724 (13,306b)

[28] Mark J. Gotay and James M. Nester, *Presymplectic Lagrangian systems. I. The constraint algorithm and the equivalence theorem*, Ann. Inst. H. Poincaré Sect. A (N.S.) **30** (1979), no. 2, 129–142. MR535369 (80j:58035)

[29] ———, *Presymplectic Lagrangian systems. II. The second-order equation problem*, Ann. Inst. H. Poincaré Sect. A (N.S.) **32** (1980), no. 1, 1–13. MR574809 (81m:58033)

[30] Jürgen Moser, *A rapidly convergent iteration method and non-linear partial differential equations. I*, Ann. Scuola Norm. Sup. Pisa (3) **20** (1966), 265–315. MR0199523 (33 #7667)

[31] ———, *A rapidly convergent iteration method and non-linear differential equations. II*, Ann. Scuola Norm. Sup. Pisa (3) **20** (1966), 499–535. MR0206461 (34 #6280)

[32] ———, *Convergent series expansions for quasi-periodic motions*, Math. Ann. **169** (1967), 136–176. MR0208078 (34 #7888)

[33] Yong Li and Yingfei Yi, *Persistence of invariant tori in generalized Hamiltonian systems*, Ergodic Theory Dynam. Systems **22** (2002), no. 4, 1233–1261, DOI 10.1017/S0143385702000743. MR1926285 (2003g:37107)

[34] M. C. Muñoz Lecanda and N. Román-Roy, *Lagrangian theory for presymplectic systems*, Ann. Inst. H. Poincaré Phys. Théor. **57** (1992), no. 1, 27–45 (English, with English and French summaries). MR1176356 (93j:58052)

[35] Helmut Rüssmann, *On optimal estimates for the solutions of linear partial differential equations of first order with constant coefficients on the torus*, Dynamical systems, theory and applications (Rencontres, Battelle Res. Inst., Seattle, Wash., 1974), Springer, Berlin, 1975, pp. 598–624.Lecture Notes in Phys., Vol. 38. MR0467824 (57 #7675)

[36] Mikhail B. Sevryuk, *KAM tori: persistence and smoothness*, Nonlinearity **21** (2008), no. 10, T177–T185, DOI 10.1088/0951-7715/21/10/T01. MR2439472 (2009i:37151)

[37] Michael E. Taylor, *Partial differential equations I. Basic theory*, 2nd ed., Applied Mathematical Sciences, vol. 115, Springer, New York, 2011. MR2744150 (2011m:35001)

[38] Walter Thirring, *Classical mathematical physics*, 3rd ed., Springer-Verlag, New York, 1997. Dynamical systems and field theories; Translated from the German by Evans M. Harrell, II. MR1473724 (98g:70002)

[39] Izu Vaisman, *Lectures on the geometry of Poisson manifolds*, Progress in Mathematics, vol. 118, Birkhäuser Verlag, Basel, 1994. MR1269545 (95h:58057)

[40] Xiaoping Yuan, *Recent progress on nonlinear wave equations via KAM theory*, Control theory and related topics (2007), 387–395. MR2391843 (2009d:37131)

[41] E. Zehnder, *Generalized implicit function theorems with applications to some small divisor problems. I*, Comm. Pure Appl. Math. **28** (1975), 91–140. MR0380867 (52 #1764)