Departament de Matemàtica Aplicada I
Universitat Politècnica de Catalunya

Memòria presentada per aspirar al grau
de doctor en Ciències Matemàtiques

Director de tesi : Jordi Villanueva Castelltort

Analytic and numerical tools for the study of quasi-periodic motions in Hamiltonian Systems

Alejandro Luque
<alejandro.luque@upc.edu>

Barcelona, 27th November 2009

Todo el cariño que durante estos años he puesto en esta
tesis está dedicado a nuestras Grandes Mujeres.
Aquellas que, tras vivir una infancia en tiempos muy dificiles,
tuvieron o supieron encontrar el coraje para sacar adelante a sus familias.
Aquellas que, con su buen humor y sentido común,
han hecho mejor la vida de cuántos les rodean.
Aquellas que, volcadas en la educación de hijos y nietos,
han hecho posible (en una o dos generaciones) el salto de la miseria
a la oportunidad de escoger un camino en la vida.
Me refiero a nuestras abuelas (y a todas las Cándidas Villar del mundo).

En concreto, dedico esta tesis a mi yaya **María Jiménez Lozoya**,
ejemplo de buen humor, generosidad y tolerancia,
entre otras virtudes (con algún defecto).

Acknowledgements

I would like to express my sincere acknowledgement to my advisor Jordi Villanueva for so many things that I have learned from him. He introduced me in this research area (actually he introduced me in mathematics a while ago!) and he has strongly motivated and encouraged me during these years. I am in debt with him not only because I have benefitted greatly from his extensive knowledge, but also for his unending patience.

Of course, I would like to acknowledge all people from the academic atmosphere that helped me in any way along these years of learning and researching. Giving the deserved credit to all of them would require an extension larger than any chapter of this thesis.

To the members of *UPC Dynamical Systems Group* for creating an outstanding scientific atmosphere, for their enthusiasm, for all training activities organized, for permanent motivation and exchange of ideas, etc. Amadeu Delshams for giving me valuable support and for his tenacity and patience when solving some bureaucratic problems that eventually occurred; Tere M. Seara for so much mathematics that I have learned from her and also for encouragement and priceless advises; Joaquim Puig for being always aware of my research and for providing me many interesting material; Pau Roldán for teaching and supporting me in the use of the clusters and his patience with users like me (those that only learn from their mistakes); Rafael Ramírez-Ros for introducing and helping me in the PARI-GP software; Tomás Lázaro for the way he cares about students and for his unbounded smile and optimism; Chara Pantazi for always saying yes when I asked for any favor; Primitivo Acosta and David Blázquez for their hospitality in Bogotá and for scientific discussions.

To the members of *UB Dynamical Systems Group* for sharing the collaborative atmosphere to which I referred above. In particular, I want to thank Carles Simó for teaching me on many topics in Dynamical Systems and for working with such enthusiastic and contagious energy.

To Rafael de la Llave for his lessons —always brimming with ideas—, valuable advises and interesting suggestions. Some of them helped to improve certain aspects of this work and some other ones will be object of subsequent research. Indeed, I am going to visit him at *Department of Mathematics* (*University of Texas at Austin*) during next months and I am looking forward to work in the fruitfull research environment that he creates[1].

[1]Addendum: after this staying, I want to thank everybody (the list is too large) that made my days in Austin so pleasant and enjoyable. I want to thank Rafael for always finding time for me, in espite of so many obligations that he had. Also to Nikola Petrov for inviting me one week to Oklahoma, where we had stimulating discussions.

To Florentino Borondo for his hospitality during my staying at *Departamento de Química* (*Universidad Autónoma de Madrid*). I thank him for many formal and informal discussions on Quantum Mechanics that allowed me to “escape” from the world of Classical Mechanics for a while. I appreciate the opportunity to learn and work in different topics from those developed in this thesis. I want to thank also to professors and students there for being so nice to me.

To Jacobo Aguirre and Daniel Peralta-Salas for exchanging with me many valuable ideas and for making our research specially funny and enjoyable. I learned from them many things on electromagnetism and chaotic scattering of open systems.

To the members of the network DANCE for making it possible for Spanish researchers in Dynamical Systems to have a meeting point to exchange knowledge and ideas. I expect to learn many new dances and to dance with as many dancers as possible.

To the people of *Departament de Matemàtica Aplicada I* (*Universitat Politècnica de Catalunya*) for making “MA1” a pleasant place to work. In particular, I would like to mention María Aguareles, Miguel Ángel Barja, Marta Casanellas, Neus Consul and Jesús Fernández. Also I want to thank Gemma Baldrís, Carme Capdevila, Rosa Maria Cuevas, Ester Pineda and Hilda Rodón for attending and/or helping me so many times.

To Yannick Sire for his encouragement and interest on some of the results presented here and for valuable discussions on reference [FdlLS09] before it was published. I appreciate very much his friendship, too.

To my officemates for suffering me during the “days of stress” and most specially to Marina Gonchenko, Marcel Guardia, Gemma Huguet, Martí Lahoz and Abraham de la Rosa. I feel that I have been very lucky to share these years with them (I include also Renato Calleja), not only because all the mathematics that we have learn together and/or discussed about, but also for encouragement, favors, friendship and support.

During this work I have been granted and funded by several fellowships and projects. I acknowledge the Catalan Ph.D. Fellowship 2005FI 00155, the Spanish Ph.D. Fellowship FPU AP2005-2950 and the position co-funded by PTA2008-1693-P. I also want to thank funds from the Catalan CIRIT project num 2009 SGR 859 and the Spanish MCyT/FEDER project MTM2006-00478. Finally, I acknowledge the use of EIXAM, the UPC Applied Math cluster system for research computing.

Por otro lado, quiero expresar mi más profundo agradecimiento a gente fantástica que me ha acompañado a lo largo de estos años y con los que he compartido grandísimos momentos. Gracias por hacer la vida tan fácil y llevadera: Fernando Burgos, Maida Bustreo, Ane Eizaguirre, Verónica Fernández, José Gámez, Ester Garcia, Anna Giménez, Cristóbal Lorente (por estar ahí desde que tengo memoria), Pablo López, Mariluz Luque, Alex Narváez, Luis Nuñez, Šárka Pětivlasová, Ruth Sanchis, “vecinhos” y Yang Wang. Soy afortunado también por la família que tengo y agradezco el cariño de mis padres Fran y Romy y de mi hermano Jorge.

Sara, quiero agradecerte de forma muy especial todas las cosas que compartimos. Por bellas que puedan resultar la Ciencia o las Matemáticas, tu me enseñas que hay cosas mucho más bellas e importantes por las que empezar cada día con más ilusión que el anterior.

Contents

List of Figures

List of Tables

Chapter I

Introduction

It is a general fact that equations derived from the observation of Nature are extremely difficult or impossible to solve. Nevertheless, very often these equations are "close" to some "ideal" ones whose solutions can be obtained and characterized using diverse techniques. The difference between the "real" and the "ideal" problem is known as *perturbation*, and the idea of *perturbation theory* is to use the knowledge of the "ideal" system in order to understand the long term behavior of the "real" problem.

One of the most fruitful areas in perturbation theory is the context of Classical Mechanics, where equations typically have an special structure. The structure that we have in mind is the symplectic structure that appears in Hamiltonian Mechanics, which gathers a wide collection of problems in Celestial Mechanics and other frictionless mechanical systems, molecular dynamics, plasma/beam physics, etc.

The contributions of this thesis are focused in the study of invariant tori carrying quasi-periodic motions in Hamiltonian systems. Invariant tori play an important role in order to understand a given system, since they are part of the "skeleton" that organizes the dynamics, and they are ubiquitous in Hamiltonian Mechanics. Indeed, the characterization of this skeleton in the context described at the beginning has been for long time —and still is— one of the main challenges in Dynamical Systems. Concretely, the aim of our results (let us remark that the methods presented here do not restrict to a perturbative setting) is twofold:

- In order to study quasi-periodic invariant tori, valuable information is obtained from the frequency vector that characterizes the motion. Part of the work in this thesis has been to develop efficient numerical methods for the study of one dimensional quasi-periodic motions in a wide set of contexts. Our methodology is an extension of a recently developed approach to compute rotation numbers of circle maps (see [SV06]) based on suitable averages of iterates of the map. On the one hand, the ideas of [SV06] have been adapted to compute derivatives of the rotation number for parametric families of circle diffeomorphisms, thus obtaining powerful tools for the study of Arnold Tongues and invariant curves for twist maps (this is presented in Chapter II). We want to point out that this variational information cannot be obtained in such a direct way by means of other existing methods to compute rotation numbers (we refer to the works [BS00, Bru92, GMS, dlLP02, LFC92]). On the other hand, further extensions of the previous methods have been developed to deal with more general contexts such as non-twist maps or quasi-periodic signals (the results are contained in Chapter III).

- KAM theory deals with the persistence of quasi-periodic invariant tori in the perturbed problem. The other part of the work, presented in Chapter IV, is to adapt recently developed parameterization methods (we refer to [FdlLS09, JdlLZ99, JS96, dlLGJV05]) to the study of elliptic lower dimensional tori of Hamiltonian system. The method is based in solving iteratively the functional equations that stands for invariance and reducibility. In contrast with classical methods, it is not assumed that the system is close to integrable nor that is written in action-angle variables. It is worth mentioning that the presented approach has a lot of advantages compared with methods which are built in terms of canonical transformations, e.g. it produces constructive proofs that lead to more efficient numerical algorithms for the computation of these objects (see [dlLHS]). Such numerical algorithms are suitable in order to perform computer assisted proofs.

I.1 Background and some previous results

This section pretends to be a bridge for mathematicians which do not work in the field and to motivate the achievements obtained in this thesis. There is no need to remark that this brief introduction does not pretend to be exhaustive nor detailed so we will refer in advance to the standard textbooks: some classical references are [AM78, Arn88, Arn89, Wei79]; see also the modern ones [HZ94, JS98, Laz93] (the second has a physical point of view and presents many worked examples) or the surveys [Bry91, dlL01] (the last one is a very rich source of references and techniques in KAM theory).

Definition I.1.1. *Given a function $h : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$ such that $(q,p) \mapsto h(q,p)$, the following systems of ordinary differential equations*

$$\dot{q}_i = \frac{\partial h}{\partial p_i}(q,p), \qquad \dot{p}_i = -\frac{\partial h}{\partial q_i}(q,p), \qquad i = 1, \ldots, n, \tag{I.1}$$

is called Hamilton's canonical equations of n degrees of freedom *and the function h is called* Hamiltonian. *The variables q_i are called* positions *and p_i are called* momenta.

Let us observe that denoting $x = (q,p)$ it is equivalent to write

$$\dot{x} = J_n \mathrm{grad}\, h(x) = J_n Dh(x)^\top = X_h(x),$$

where the $2n \times 2n$ matrix

$$J_n = \begin{pmatrix} 0 & \mathrm{Id}_n \\ -\mathrm{Id}_n & 0 \end{pmatrix}$$

is skew-symmetric $J_n^\top = -J_n$ and non-singular $\det J_n = 1$.

Given an initial condition $x(0) = x^0$, there exists a unique solution of $\dot{x} = X_h(x)$ that we write as $t \mapsto \phi(t, x^0)$ and the map $(t, x) \mapsto \phi(t, x) = \phi_t(x)$ —whenever defined— is called the *flow associated to* X_h. Moreover, given another function $f : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$, it is natural to ask for how f varies along the flow of the system X_h. This is obtained from the computation

$$\frac{d}{dt} f(\phi(t,x)) = Df(\phi(t,x)) X_h(\phi(t,x)) = \mathrm{grad}\, f(\phi(t,x))^\top J_n \mathrm{grad}\, h(\phi(t,x)). \tag{I.2}$$

Introducing the *Poisson bracket* of two functions f, g as follows

$$\{f, g\} = \sum_{i=1}^{n} \frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q_i}, \tag{I.3}$$

we observe that (I.2) can be written in terms of the Poisson bracket as

$$\frac{d}{dt}(f \circ \phi_t) = \{f, h\} \circ \phi_t. \tag{I.4}$$

Let us also recall that given a differential equation $\dot{x} = X(x)$ and a diffeomorphism $x = \varphi(y)$, we can transform the system as $\dot{y} = \varphi^* X(y) = (D\varphi(y))^{-1} X(\varphi(y))$ (i.e. pull-back by φ). If $X = X_h$ is a Hamiltonian system, then the new vector field in not necessarily Hamiltonian. Diffeomorphisms preserving this property for all Hamiltonians are called *canonical transformations* and they are characterized in the following proposition

Proposition I.1.2. *Given a diffeomorphism $x = \varphi(x)$, the following statements are equivalent:*

(i) The diffeomorphism satisfies $D\varphi(y)^\top J_n D\varphi(y) = J_n$.

(ii) For every function h, we have $\varphi^* X_h = X_{\varphi^* h}$.

(iii) For every pair of functions f, g, we have $\{\varphi^* f, \varphi^* g\} = \varphi^* \{f, g\}$.

Using these properties, and recalling the previous discussion, it is easy to check that for every fixed t the flow $x \mapsto \phi_t(x)$ is a canonical transformation. This fact is very important in order to construct canonical transformations in the context of normal forms and in the classical approach to KAM theory (see Remark I.1.12).

Although all the computations and results that will be presented in this thesis are performed in particular coordinates, it is convenient to recall the corresponding concepts of symplectic geometry. In the following, given a real smooth manifold M, we will denote by $\mathfrak{F}(M)$ the space of functions on M, $\mathfrak{X}(M)$ the space of vector fields on M and $\mathcal{A}^r(M)$ the space of alternating r-forms on M.

Definition I.1.3. *A* symplectic structure *on a smooth even dimensional manifold M is a non-degenerate*[1]*, closed 2-form $\Omega \in \mathcal{A}^2(M)$. The pair (M, Ω) is called a* symplectic manifold*. In addition, a symplectic manifold is called* exact *if Ω is exact, that is, $\Omega = d\alpha$ for some 1-form α.*

If (M, Ω) and (N, Υ) are symplectic manifolds, then a diffeomorphism $\phi : M \to N$ satisfying $\phi^*\Upsilon = \Omega$ is called *symplectomorphism* (or *symplectic map* when it is clear that ϕ is a diffeomorphism), and we write $\mathrm{Sp}(M)$ the group of symplectomorphisms of (M, Ω). In addition, ϕ is an *exact* symplectomorphism if there exist a function $S \in \mathfrak{F}(M)$ such that $\phi^*\alpha = \alpha + dS$, where $\Omega = d\alpha$.

A vector field X on M is said to be *symplectic* if Ω is invariant under the flow of X, this is, if the Lie derivative $L_X\Omega$ vanishes. The space of symplectic vector fields on M is denoted by $\mathfrak{sp}(M)$ and corresponds to the Lie algebra of $\mathrm{Sp}(M)$. Since Ω is closed, it is equivalent to say (use Cartan's identity) that X is symplectic if and only if the form $X \lrcorner \Omega = \Omega(\cdot, X)$ is closed.

Due to the non-degeneracy of Ω, for any vector field X on M, the 1-form $\flat(X) = -X \lrcorner \Omega$ vanishes only where X does. Since TM and T^*M have the same rank, it follows that the map $\flat : \mathfrak{X}(M) \to \mathcal{A}^1(M)$ is and isomorphism. We denote $\sharp : \mathcal{A}^1(M) \to \mathfrak{X}(M)$ its inverse.

With this notation, we can write $\mathfrak{sp}(M) = \sharp(\mathcal{Z}^1(M)))$ where $\mathcal{Z}^1(M)$ denotes the space of closed 1-forms on M. In particular, we can consider the subspace $\mathcal{B}^1(M)$ of exact 1-forms on M giving rise to the following definition.

Definition I.1.4. *For each $f \in \mathfrak{F}(M)$, the vector field $X_f = \sharp(df)$ is called the* Hamiltonian vector field *associated to f. The set of all Hamiltonian vector fields on M is denoted $\mathfrak{h}(M)$.*

Thus, $\mathfrak{h}(M) = \sharp(\mathcal{B}^1(M))$ and for this reason, Hamiltonian vector fields are often called *exact*. Moreover, $\mathfrak{h}(M)$ is an ideal of $\mathfrak{sp}(M)$ and we have the following exact sequence

$$0 \longrightarrow \mathfrak{h}(M) \longrightarrow \mathfrak{sp}(M) \longrightarrow H^1_{dR}(M, \mathbb{R}) \longrightarrow 0,$$

where $H^1_{dR}(M, \mathbb{R})$ is the first deRam cohomology group. Thus, sometimes the elements of $\mathfrak{sp}(M)$ are also called *locally Hamiltonian vector fields*.

Given $X \in \mathfrak{sp}(M)$ we have that the flow ϕ_t of X is a symplectomorphism (i.e., $\phi_t^*\Omega = \Omega$). Moreover, if $X_h \in \mathfrak{h}(M)$ we have that ϕ_t is exact symplectic and that the Hamiltonian h is constant along the flow (i.e., $h \circ \phi_t = h$).

Suppose that (M, Ω) is a $2n$-dimensional symplectic manifold. We say that two tangent vectors $u, v \in T_xM$ are *skew-orthogonal* if $\Omega_x(u, v) = 0$. Then a submanifold L of M is *isotropic* if the tangent space T_xL lies in its skew-orthogonal complement $(T_xL)^\perp$ for every $x \in L$. Note that for any isotropic submanifold L the restriction of Ω to L vanishes and $\dim L \leq n$. A submanifold L is called *Lagrangian* if it is maximal with respect to the property of being isotropic. Note that for Lagrangian manifolds we have $T_xL = (T_xL)^\perp$, $\Omega|_L$ vanishes and $\dim L = n$.

[1]Recall that a 2-form is non-degenerate or nonsingular if for every point $x \in M$ it turns out that if $\Omega_x(u, v) = 0$ for every $u \in T_xM$ then v must be the zero vector.

Definition I.1.5. *Given* $f, g \in \mathfrak{F}(M)$ *we define the* Poisson Bracket *of* f *and* g *as* $\{f, g\} = \Omega(X_f, X_g) = \Omega(\sharp(df), \sharp(dg))$.

The map $f \mapsto \sharp(df)$ is a Lie algebra homomorphism between Hamiltonian functions (with the Poisson bracket) and Hamiltonian vector fields (with the Lie bracket). Concretely, we have $[X_f, X_g] = X_{\{f,g\}}$ for every $f, g \in \mathfrak{F}(M)$.

As a consequence of a celebrated theorem by Darboux, around any point $x \in M$, there exist a canonical chart (U, x) where $x = (q, p)$ such that the 2-form is

$$\Omega|_U = \sum_{j=1}^{n} dq_j \wedge dp_j. \tag{I.5}$$

Furthermore, in these Darboux coordinates the Hamiltonian vector field of a Hamiltonian $h \in \mathfrak{F}(M)$ is written as

$$X_h|_U = \sum_{j=1}^{n} \left(\frac{\partial h}{\partial p_j} \frac{\partial}{\partial q_j} - \frac{\partial h}{\partial q_j} \frac{\partial}{\partial p_j} \right), \quad \text{or} \quad X_h|_U = J \operatorname{grad} h \quad J = \begin{pmatrix} 0 & \mathrm{Id} \\ -\mathrm{Id} & 0 \end{pmatrix}$$

the associated differential equation reads as (I.1), and the Poisson bracket is expressed as (I.3).

Consider again a symplectic $2n$-dimensional manifold (M, Ω). We say that a collection of functions $f_i \in \mathfrak{F}(M)$, $i \in I$, are in *involution* if their Poisson brackets vanish, i.e., $\{f_i, f_j\} = 0$ for any $i, j \in I$. Let us recall that two vector fields $X, Y \in \mathfrak{X}(M)$ are said to *commute* if their Lie Bracket vanishes, $[X, Y] = 0$. Thus, by means of the Poisson bracket, $X_f, X_g \in \mathfrak{h}(M)$ commute if, and only if, their Hamiltonian functions are in involution.

We will say that $f \in \mathfrak{F}(M)$ is a *first integral* of a vector field $X \in \mathfrak{X}(M)$ provided $L_X f = 0$, i.e., if f is invariant under the flow of X. Note that for each first integral we can reduce the evolution of the flow to a hypersurface, so roughly speaking, one can say that the more first integrals the more confined will be the dynamics. In particular the following notion is very important.

Definition I.1.6. *We say that a locally Hamiltonian vector field* X *is* completely integrable *if there exist functions* $f_i \in \mathfrak{F}(M)$*,* $1 \leq i \leq n$*, such that*

(i) The functions f_i *are first integrals of* X*.*

(ii) They are in involution.

(iii) They are functionally independent, except in a set of zero measure.

Now we can present the following well-known result:

Theorem I.1.7 (Arnold-Liouville)**.** *Let* $f_1, \ldots, f_n$ *be functions as in Definition I.1.6 and consider a common level surface of the functions* f_i *given by*

$$N_c = \{x \in M \,:\, f_i(x) = c_i,\ \forall i = 1, \ldots, n\}.$$

Then:

(i) It turns out that N_c is a smooth submanifold of M, invariant under the flow of X_{f_i}.

(ii) If N_c is compact, their connected components are diffeomorphic to $\mathbb{T}^n = \mathbb{R}^n/(2\pi\mathbb{Z})^n$.

(iii) If N_c is non-compact, then their connected components are diffeomorphic to the cylinder $\mathbb{T}^r \times \mathbb{R}^{n-r}$ for some $r \leq n$.

Moreover, the flow of X_{f_i} on each connected component of N_c is of translation type.

Let us precise the definition of a translation on the torus. A translation on the r-dimensional torus $\mathbb{T}^r$ is a map $T_{t,\omega} : \mathbb{T}^r \to \mathbb{T}^r$, that has the following explicit expression on the lift $\widetilde{T}_{t,\omega} : \mathbb{R}^r \to \mathbb{R}^r$, $x \mapsto x + \omega t$ and makes commutative the following diagram

$$\begin{array}{ccc} \mathbb{R}^r & \xrightarrow{\widetilde{T}_{t,\omega}} & \mathbb{R}^r \\ \downarrow{\scriptstyle \pi} & & \downarrow{\scriptstyle \pi} \\ \mathbb{T}^r & \xrightarrow{T_{t,\omega}} & \mathbb{T}^r \end{array}$$

where $\pi : \mathbb{R}^r \to \mathbb{T}^r$ is the canonical quotient projection. Note that, if we write $\theta = \pi(x)$ on $\mathbb{T}^r$, the above flow is generated by the constant vector field

$$X = \omega_1 \frac{\partial}{\partial \theta_1} + \ldots + \omega_r \frac{\partial}{\partial \theta_r}. \tag{I.6}$$

The vector $\omega \in \mathbb{R}^r$ is called the *frequency vector* of the linear flow and determines the dynamical properties of $T_{t,\omega}$. A frequency vector $\omega \in \mathbb{R}^r$ is said to be *non-resonant* if its components $\omega_1, \ldots, \omega_r$ are rationally independent, this is, if

$$\langle \omega, k \rangle = \sum_{j=1}^{r} \omega_j k_j = 0, \quad k_j \in \mathbb{Z}, \quad \forall j = 1, \ldots, r,$$

implies that $k_j = 0$ for every $j = 1, \ldots, r$. If $\omega \in \mathbb{R}^r$ is non-resonant, then the flow of (I.6) on $\mathbb{T}^r$ is *minimal*, i.e. the orbit of any point is dense. In the resonant case, the flow fills densely a sub-torus of dimension $r - \dim M_\omega$, where M_ω is the resonant $\mathbb{Z}$-module

$$M_\omega = \{k \in \mathbb{Z}^r \mid \langle \omega, k \rangle = 0\}.$$

Finally, the following theorem due to Arnold evidences that completely integrable systems exhibit simple dynamics.

Theorem I.1.8. *Under the hypotheses of Theorem I.1.7 and if the manifold N_c is compact, we have that:*

(i) There is a neighborhood of N_c submerged in $\mathbb{R}^{2n}$ and diffeomorphic to the direct product $\mathbb{T}^n \times U$, where U is an open subset of $\mathbb{R}^n$.

(ii) We can choose coordinates (θ, I) in $\mathbb{T}^n \times U$ in such a way that the f_i, $1 \leq i \leq n$ are expressed in terms of the I variables only.

In particular, note that in the above variables the Hamiltonian function turns $h = h(I)$ so Hamilton's equations are written in the following simple form

$$\dot{\theta}_i = \frac{\partial h}{\partial I_i}(I) = \omega_i(I), \qquad \dot{I} = -\frac{\partial h}{\partial \theta_i}(I) = 0, \qquad 1 \leq i \leq n,$$

and thus, the solution in this coordinates turns $I(t) = I_0$, $\theta(t) = \theta_0 + \omega(I_0)t$. Accordingly, the variables I are called *action variables* and θ *angular variables*. Notice that for $I_0 \in U$, the torus $\mathbb{T}^r \times \{I_0\}$ is invariant with dynamics given by a linear flow of frequency vector $\omega(I_0)$. For this reason, $I \mapsto \omega(I)$ is called *frequency map*.

We recall that both symplectic maps and Hamiltonian vector fields are related in a natural way. For example, we already stated that the flow of a Hamiltonian vector field is an exact symplectic map. Another way of obtaining an exact symplectic map from a Hamiltonian h is to consider a $(2n-2)$-dimensional manifold Σ contained on an energy level of the system $\{h(x) = h_0\}$ which is transversal to the flow. Then, the *Poincaré return map* $P : \Sigma \to \Sigma$, if defined, is also a symplectic map. For example, if the Hamiltonian h depends only on the actions and we take

$$\Sigma = \{(\theta, I) \,:\, \theta_n = \theta_n^0, \text{ and } h(I) = h_0\},$$

then P takes the form $P(\theta, I) = (\theta + \widetilde{\omega}(I), I)$, where

$$\widetilde{\omega}(I) = \left(\frac{\omega_1(I)}{\omega_n(I)}, \ldots, \frac{\omega_{n-1}(I)}{\omega_n(I)}\right).$$

It is well-known that integrable systems are rare in Nature. Nevertheless, it turns out that many systems of physical interest are "close to integrable". A classic example is the planetary problem in the Solar System. Despite its complexity, setting the problem of the motion of N-celestial bodies under their mutual gravitational attraction is one of the simplest to formulate in Physics. Assuming the gravitational constant equal to 1, the equations of motion are

$$m_i \ddot{x}_i = \frac{\partial U}{\partial x_i}, \qquad U = \sum_{1 \leq i < j \leq N} \frac{m_i m_j}{\|x_i - x_j\|},$$

where m_i and x_i denote the mass and the position of the ith body, and $\|\cdot\|$ the Euclidean norm. If we order the bodies in the Solar System by decreasing mass then we have that the ratios m_i/m_1, for $i = 2, \ldots, N$, are very small (the value $m_2/m_1 \simeq 10^{-3}$ for Jupiter is the largest one), so the previous system is described approximately by considering the $N-1$ remaining bodies describing trajectories around the Sun (without interacting each other), picture which is mathematically completely understood (this is the well-known Keplerian approximation). Starting from this integrable system, we can then attempt to include the effects of the mutual gravitational interactions.

In a more general situation, given an integrable Hamiltonian h, written in action-angle coordinates (see Theorem I.1.8), we consider the following Hamiltonian function

$$h_\varepsilon(\theta, I) = h(I) + \varepsilon f(\theta, I), \tag{I.7}$$

where f is a periodic function in the angular variables. It seems reasonable that for ε sufficiently small, the solutions of (I.7) must be related to the solutions of the corresponding unperturbed Hamiltonian h. However, this problem is extremely difficult —it was baptized by Poincaré as the *fundamental problem of mechanics*— and actually many posed questions are still unanswered. This problem has been widely studied by many researchers (using rigorous, numeric and heuristic techniques) and nowadays there is a large amount of partial results. The cornerstone of this partial knowledge is the so-called KAM theory —named after A.N. Kolmogorov [Kol54], V.I. Arnold [Arn63] and J.K. Moser [Mos62]— which establishes that many (in the sense of Lebesgue measure) stable solutions of the unperturbed system, which are given by invariant tori, persist. We refer to the survey [dlL01] for a nice introduction to KAM theory (that contains many references together with many hints and interesting remarks), even though we roughly describe here some of the main ideas.

The first result on the persistence of these tori was given by Kolmogorov in [Kol54]. He proved that a single invariant torus of fixed frequency persists provided that some arithmetic conditions of the frequency vector are satisfied.

Theorem I.1.9. *Let h_ε be a real analytic Hamiltonian function like* (I.7) *defined on $\mathbb{T}^n \times U$ and verifying the following hypotheses:*

(i) The unperturbed Hamiltonian is non-degenerate: $\det(\operatorname{hess}_I h(I)) \neq 0$, for all $I \in U$.

(ii) For a fixed $I_0 \in U$, the frequency vector $\omega = \operatorname{grad} h(I_0)$ satisfies Diophantine conditions, i.e., there exist $\gamma > 0$ and $\nu > n - 1$ such that

$$|\langle \omega, k \rangle| \geq \frac{\gamma}{|k|_1^\nu}, \qquad k \in \mathbb{Z}^n \backslash \{0\}, \tag{I.8}$$

where $|k|_1 = |k_1| + \ldots + |k_r|$.

Then, if ε is small enough (indeed, if $\varepsilon < \alpha\gamma^4$ for some constant $\alpha > 0$), there exist an invariant torus of frequency ω which is close to $(\theta(t), I(t)) = (\theta_0 + \omega t, I_0)$. In particular, there is a symplectomorphism $\Phi = \mathrm{id} + \phi$, where ϕ is small in some norm, such that

$$(\theta(t), I(t)) = \Phi(\theta_0 + \omega t, I_0)$$

is a quasi-periodic solution for X_{h_ε}.

Remark I.1.10. *The result is of local character in the following sense: it guaranties the persistence of an isolated torus in the phase space. Thus, close to a Diophantine Lagrangian torus for X_h, there is an invariant Lagrangian one for X_{h_ε} with the same frequency vector ω.*

Remark I.1.11. *Roughly speaking, the non-degeneracy condition allows us to label the tori by their frequency so each torus is identified by $\mathbb{T}^n \times \{I_0\}$ in the unperturbed system.*

Sketch of the proof of Theorem I.1.9. Let us consider the space $\mathcal{K}$ of Kolmogorov's Hamiltonian functions of the form

$$k(\theta, I) = \lambda + \langle \omega, I \rangle + \frac{1}{2}\langle I, A(\theta) I \rangle + O_3(I),$$

where $A : \mathbb{T}^n \to M_{n\times n}(\mathbb{R})$ is such that $A(\theta)^\top = A(\theta)$. Then, Hamilton's equations for $k \in \mathcal{K}$ read as

$$\dot{I}_j = \mathcal{O}(I^2), \quad \dot{\theta}_j = \omega_j + \sum_{i=1}^{n} A_{ji}(\theta) I_i + \mathcal{O}(I^2), \quad j = 1, \dots, n,$$

so it is clear that $(\theta(t), I(t)) = (\omega t, 0)$ is a quasi-periodic solution of the vector field X_k. The proof consist in making changes of coordinates to h_ε in such a way that the Hamiltonian function h_ε looks like k so we can assign to X_{h_ε} a quasi-periodic solution.

In particular, we expand h and f in Taylor series with respect to I

$$h(I) = h(I_0) + \langle \omega, I \rangle + \frac{1}{2}\langle I, \operatorname{hess} h(I_0) I \rangle + \mathcal{O}(I^3),$$

$$f(\theta, I) = f_0(\theta) + \langle f_1(\theta), I \rangle + \frac{1}{2}\langle I, f_2(\theta) I \rangle + \mathcal{O}(I^3).$$

Then we write h_ε as a perturbation of k

$$\begin{aligned} h_\varepsilon(\theta, I) &= h(I_0) + \langle \omega, I \rangle + \frac{1}{2}\langle I, (\operatorname{hess} h(I_0) + \varepsilon f_2(\theta)) I \rangle + \varepsilon f_0(\theta) + \langle \varepsilon f_1(\theta), I \rangle + \mathcal{O}(I^3) \\ &= k(\theta, I) + F(\theta, I), \end{aligned}$$

with $\lambda = h(I_0)$ and $A = \operatorname{hess} h(I_0) + \varepsilon f_2$; and our aim is to construct a symplectomorphism Φ such that $\Phi^* h_\varepsilon \in \mathcal{K}$. This can be done by means of the Lie series method that we recall in Remark I.1.12 below.

Let g be another Hamiltonian function and ϕ^g the time one flow of the vector field X_g. Then

$$\begin{aligned} (\phi^g)^* h_\varepsilon &= h_\varepsilon + \{h_\varepsilon, g\} + \frac{1}{2!}\{\{h_\varepsilon, g\}, g\} + \dots \\ &= k + \underbrace{F + \{k, g\}}_{\mathcal{O}(\varepsilon)} + \underbrace{\{F, g\} + \frac{1}{2!}\{\{h, g\}, g\}}_{\mathcal{O}(\varepsilon^2)} + \dots \end{aligned}$$

Thus, if we find g such that $k+F+\{k,g\}\in\mathcal{K}$ we will obtain a new Hamiltonian which is not in $\mathcal{K}$ due to terms $\mathcal{O}(\varepsilon^2)$. Iterating this process infinitely many times we would obtain a quadratic convergent scheme to overcome the small divisor problems that will appear when solving the previous equation. Moreover, since the condition $k+F+\{k,g\}\in\mathcal{K}$ is very difficult to fulfill, we choose to study a linearized problem. In particular, we restrict g to be of the form

$$g(\theta,I)=c(\theta)+\langle X,\theta\rangle+\langle Y(\theta),I\rangle,$$

where $X\in\mathbb{R}^n$ and $c:\mathbb{T}^n\to\mathbb{R}$, $Y:\mathbb{T}^n\to\mathbb{R}^n$ are 2π-periodic functions. Introducing this Hamiltonian into $k+F+\{k,g\}$ and ignoring terms $O_2(I)$, we obtain the following equation

$$\langle\omega,X\rangle+\varepsilon f_0+L_\omega c+\left\langle L_\omega Y+b+A(X+\operatorname{grad}c)+\varepsilon f_1,I\right\rangle=0,$$

where we have introduced the operator

$$L_\omega=\sum_{j=1}^{n}\omega_j\frac{\partial}{\partial\theta_j},\tag{I.9}$$

which corresponds to the Lie derivative with respect to the vector field (I.6). Then, if we want to eliminate the terms $O(\varepsilon)$ we need

$$L_\omega c=-\varepsilon f_0,\tag{I.10}$$

$$L_\omega Y=-b-A(X+\operatorname{grad}c)-\varepsilon f_1.\tag{I.11}$$

The above scheme will make sense if we can solve the equation $L_\omega\varphi=\eta$ for a given $\eta:\mathbb{T}^n\to\mathbb{R}$. Let us show how the small divisors appear by considering Fourier series

$$\varphi(\theta)=\sum_{k\in\mathbb{Z}}\hat{\varphi}_k\mathrm{e}^{\mathrm{i}\langle\theta,k\rangle},\qquad\eta(\theta)=\sum_{k\in\mathbb{Z}}\hat{\eta}_k\mathrm{e}^{\mathrm{i}\langle\theta,k\rangle}.$$

Then

$$L_\omega\varphi(\theta)=\sum_{j=1}^{n}\omega_j\frac{\partial}{\partial\theta_j}\sum_{k\in\mathbb{Z}}\hat{\varphi}_k\mathrm{e}^{\mathrm{i}\langle\theta,k\rangle}=\sum_{k\in\mathbb{Z}}\mathrm{i}\langle\omega,k\rangle\hat{\varphi}_k\mathrm{e}^{\mathrm{i}\langle\theta,k\rangle}=\eta(\theta),$$

and we obtain the following formal result provided $\hat{\eta}_0=0$

$$\varphi(\theta)=\sum_{k\in\mathbb{Z}\setminus\{0\}}\frac{\hat{\eta}_k}{\mathrm{i}\langle\omega,k\rangle}\mathrm{e}^{\mathrm{i}\langle\theta,k\rangle}.$$

Here is the place where the Diophantine hypothesis plays an important role. This condition is important in order to control the small divisors $\langle\omega,k\rangle$ and ensure analyticity of the solutions of (I.10) and (I.11). We omit the technical details related to the convergence of the whole process. Finally we remark that the non-degeneracy conditions are used in order to guarantee that the averages on the right-hand side of (I.10) and (I.11) vanish.

For complete and well-written proofs we refer to [BGGS84, HI04, Vil08]. See also the remarks and comments in the tutorial [dlL01]. □

Remark I.1.12. *In order to construct a symplectomorphism ϕ we can consider the flow ϕ_t^g of another Hamiltonian vector field X_g to be determined according to our purpose. Concretely, if we take the time one flow, expand $(\phi_1^g)^*h$ and using formula* (I.4)*, we obtain*

$$\begin{aligned}(\phi_1^g)^*h = h \circ \phi_1^g &= h + \frac{d}{dt}(h \circ \phi_t^g)\Big|_{t=0} + \frac{1}{2!}\frac{d^2}{dt^2}(h \circ \phi_t^g)\Big|_{t=0} + \dots \\ &= h + \{h, g\} + \frac{1}{2!}\{\{h, g\}, g\} + \dots \end{aligned} \tag{I.12}$$

This approach is called the Lie series method because we can write $h \circ \phi_1^g = \sum_{m=0}^{\infty} \frac{t^m}{m!} L_g^m h$, where $L_g^0 h = h$ and $L_g^m h = \{L_g^{m-1} h, g\}$, for $m > 1$. We remark that the original proof of Kolmogorov used generating functions rather than Lie series to construct these transformations.

A typical KAM theorem in the spirit of Arnold's paper [Arn63], where the result was extended to deal with a family of tori, is the following:

Theorem I.1.13. *Let $h_\varepsilon : \mathbb{T}^n \times U \to \mathbb{R}$ be a real analytic Hamiltonian function of the form* (I.7) *that satisfies the non-degeneracy assumption*

$$\det(\operatorname{hess}_I h(I)) \neq 0, \qquad \textit{for all } I \in U.$$

Then, if ε is small enough (indeed, if $\varepsilon < \alpha\gamma^2$ for some constant $\alpha > 0$, where γ is the Diophantine constant in (I.8) *of the surviving tori), the majority (in the sense of the Lebesgue measure) of invariant tori of the integrable system persist and are only slightly deformed.*

Sketch of the proof. Basically, this approach consists in performing symplectic transformations that reduce the system to integrable in a (complicated) region of space. We start by considering low order resonances (where $\langle \omega(I), k\rangle$ is small for k not big) reducing the system to a new one with a smaller perturbation. At each iterative step we have to take into account higher order resonances, thus obtaining an integrable Hamiltonian defined on a Cantor set at the end of the procedure.

In particular, starting with the Hamiltonian (I.7), we perform a symplectic transformation ϕ^g (to the determined), which is the time one flow of a vector field X_g. The transformed Hamiltonian is written as

$$\begin{aligned} h_\varepsilon \circ \phi^g &= h + \varepsilon f + \{h + \varepsilon f, g\} + \frac{1}{2}\{\{h + \varepsilon f, g\}, g\} + \dots, \\ &= h + \underbrace{\varepsilon f_{\leq K} + \{h, g\}}_{\Delta h} + \varepsilon f_{>K} + \{\varepsilon f, g\} + \frac{1}{2}\{\{h + \varepsilon f, g\}, g\} + \dots, \end{aligned}$$

where

$$f_{\leq K}(\theta, I) = \sum_{\substack{k \in \mathbb{Z}^n \\ |k|_1 \leq K}} \hat{f}_k(I) \mathrm{e}^{\mathrm{i}\langle \theta, k\rangle}, \qquad f_{>K}(\theta, I) = \sum_{\substack{k \in \mathbb{Z}^n \\ |k|_1 > K}} \hat{f}_k(I) \mathrm{e}^{\mathrm{i}\langle \theta, k\rangle}.$$

Then, we take g in such a way that $\Delta h = \varepsilon f_{\leq K} + \{h, g\}$ depends only on the action variables. Let us observe that $\{h, g\} = -L_\omega g$ so g is obtained by solving a cohomological equation as (I.10). Hence, the domain of g is taken according to some Diophantine-like properties. When iterating this procedure, the term $f_{>K}$ is controlled by taking larger values of K tending to infinity. We refer to [dlL01, DG96a, Gal83] for detailed and instructive proofs. □

Remark I.1.14. *Apart from measure arguments, another difference to remark between the approaches of Kolmogorov and Arnold is that the last establishes a factor γ^2 instead of γ^4. This difference is due to the fact that in Theorem I.1.9 we require to solve two cohomological equations at each iterative step, and only one is required in Theorem I.1.13. Recently, in [Vil08] Kolmogorov theorem was proved under the condition $\varepsilon < \alpha\gamma^2$.*

The hypothesis on the analyticity of the Hamiltonian is not necessary for these results, and there are many works in the literature that deal with smooth and finite differentiable cases [Bos86, Féj04, GdlL08, Laz93, Pös82, Zeh75, Zeh76]. In addition, similar theorems can be proved asking for weaker non-degeneracy conditions. For example, in [Sev95] existence of invariant tori was proved under the so-called Rüssmann's non-degeneracy condition. Nowadays, KAM theory is a vast area of research that involves a large collection of methods and applications to a wide set of contexts: Hamiltonian systems, reversible systems, volume-preserving systems, symplectic maps, PDEs and lattices, just to mention a few. Unfortunately, it is not possible (neither is our goal) to comment all of them. Again, we refer to [dlL01] for a very complete introduction to KAM theory that collect many fundamental aspects and cover a large amount of bibliography.

Apart from Lagrangian tori (i.e., of maximal dimension) it is also important to characterize isotropic tori of lower dimension. In the perturbative context described above, lower dimensional tori appear close to resonant frequencies (see [Nek94, Tre89]). Concretely, let us assume that for the system (I.7), we have certain I_0 such that $\langle \operatorname{grad} h(I_0), k \rangle = 0$ for some $k \in \mathbb{Z}^n$. As we discused, since the frequencies are rationally dependent, the flow on the torus $\mathbb{R}^n \times \{I_0\}$ is not dense. Indeed, if we have r independent frequencies, this torus contains an $(n - r)$-family of r-dimensional invariant tori, and each of these tori is densely filled up by the flow. Here, the natural problem is also to study the persistence of these lower dimensional invariant tori for $\varepsilon \neq 0$. It turns out that KAM theory also extends to this context, and most (in the sense of Lebesgue measure) of these tori survive, having a normal behavior of either elliptic of hyperbolic type. Although a theory for hyperbolic tori has been known for long time (see [Gra74]), first rigorous proofs of existence of elliptic tori of dimension $r < n$ were not given until the work in [Eli88, Kuk88] (the case $r = n - 1$ was studied in [Mos67]).

The main source of difficulty in presence of elliptic normal directions is the so-called *lack of parameters* problem [BHS96, Eli88, JV97a, JV97b, Mos67, Sev99]. Basically, since we have only as many internal parameters ("actions") as the number of basic frequencies of the torus, we cannot control simultaneously the normal ones, so we cannot prevent them from "falling into resonance". This is equivalent to say that, for a given Hamiltonian system, we cannot construct a torus with fixed basic and normal frequencies, because there are not enough internal parameters. The previous fact leads to the exclusion of a small set of these internal parameters in order to avoid resonances involving normal frequencies. To perform the previous exclusion, it is necessary to assume that the normal frequencies "move" as a function of the internal parameters. Another possibility to overcome this problem is to apply the so-called Broer-Huitema-Takens theory (see [BHTB90]). This consists in adding as many (external) parameters as needed to control simultaneously the values of both basic and normal frequencies (this process is referred as *unfolding*). Then, under suitable (generic) hypotheses on the parameter dependence, we can prove that —under small perturbations— there exist invariant tori over a "Cantor set" of large measure in parameter space. Finally, if we are interested in recovering invariant tori for the original system, it is possible to obtain measure estimates for the corresponding set of projected parameters by means of results on *Diophantine approximation on submanifolds* (see [BHS96, Sev99]).

Another difficulty linked to persistence of lower dimensional invariant tori refers to reducibility of the normal variational equations (at least in the elliptic directions) which is usually asked in order to simplify the study of the linearized equations involved. In order to achieve this reducibility, it is typical to consider second order Melnikov conditions [Mel65] to control the small divisors of the cohomological equations appearing in the construction of the reduced matrix (even though after the work in [Bou97, Eli01, XY01] it is known that only first order Melnikov conditions are strictly necessary, as we discuss in Chapter IV). Other approaches for studying persistence of invariant tori in the elliptic context, without second order Melnikov conditions, are discussed in Remark IV.3.10. In Section IV.2.3 we provide a more detailed discussion on reducibility.

To give a more precise exposition of the discussion presented in the previous paragraphs, let us consider an unperturbed Hamiltonian of the form

$$h_0 = e + \sum_{j=1}^{r} \omega_j y_j + \frac{1}{2} \sum_{j=1}^{m} \Omega_j (u_j^2 + v_j^2), \tag{I.13}$$

defined on $(x, y, u, v) \in \mathbb{T}^r \times \mathbb{R}^r \times \mathbb{R}^{n-r} \times \mathbb{R}^{n-r}$ and with the associated symplectic form

$$\sum_{j=1}^{r} dx_j \wedge dy_j + \sum_{j=1}^{n-r} du_j \wedge dv_j.$$

It is straightforward to check that this Hamiltonian has and invariant torus $\mathbb{T}^r \times \{0\} \times \{0\} \times \{0\}$ carrying quasi-periodic dynamics $x(t) = x_0 + \omega t$ with basic frequencies $\omega = (\omega_1, \ldots, \omega_r)$. The normal directions of the torus are given by the variables $(u, v) \in \mathbb{R}^{2(n-r)}$, whose origin is an elliptic fixed point with characteristic frequencies $\Omega = (\Omega_1, \ldots, \Omega_{n-r})$ and $-\Omega$ (the so-called *normal frequencies*). As we said before, normal frequencies cannot be fixed arbitrarily once the basic frequencies $\omega \in \mathbb{R}^r$ have been selected. Typically, these normal frequencies are considered as a function of the basic frequencies of the torus, as it is done in the following classical result (in the spirit of the statement proved in [Pös89]).

Theorem I.1.15. *Let us suppose that we have a Hamiltonian as* (I.13) *defined for* $\omega \in U \subset \mathbb{R}^r$*, whose normal frequencies* $\Omega = \Omega(\omega)$ *are real analytic on a complex neighborhood of* $U \in \mathbb{R}^r$*, and let us suppose that* h *is a Hamiltonian close to* h_0 *in certain norm. Then, there exists a Cantor set* $U_\gamma \subset U$ *of large Lebesgue measure, and a Hamiltonian function* $\hat{h}_0$ *(close to* h_0*) of the form* (I.13)*, with normal frequencies* $\hat{\Omega} = \hat{\Omega}(\omega)$ *defined for* $\omega \subset U_\gamma$*, such that for some* $\nu > r - 1$ *we have*

$$|\langle k, \omega\rangle + \langle l, \hat{\Omega}(\omega)\rangle| \geq \frac{\gamma/2}{|k|_1^\nu}, \qquad |l|_1 \leq 2, |k|_1 \neq 0$$

on U_γ *and there exists also a symplectic transformation* Φ *that is real analytic for each* ω *and smooth in the sense of Whitney with respect to* ω*, such that* $h \circ \Phi = \hat{h}_0$ *(modulo terms of third order in* (y, u, v)*), i.e., it possesses an elliptic invariant torus with frequencies* ω *and* $\hat{\Omega}(\omega)$*, for each frequency vector* $\omega \in U_\gamma$*.*

So far we have presented some classical results on persistence of both Lagrangian and lower dimensional invariant tori, which are based on symplectic transformations performed on the Hamiltonian function. These methods typically deal with a perturbative setting in such a way that the problem is written as a perturbation of an integrable Hamiltonian (in the sense that it has a family of reducible invariant tori), and take advantage of action-angle-like coordinates for the unperturbed Hamiltonian system. As we will clarify in Section IV.1, the classical approach presents some shortcomings, whose origin is that the method restricts to perturbative problems. For example, these procedures presents some difficulties when facing concrete problems[2].

In this thesis we face persistence of lower dimensional tori using so-called parameterization methods for Hamiltonian systems, under non-perturbative setting and without using action-angle variables (this is done in Chapter IV). The idea of parameterization methods is to build an iterative scheme to solve the invariance equation of the torus. Instead of performing canonical transformations, this scheme is performed by adding an small function to the approximation of the torus. This correction is obtained by solving (approximately) the corresponding linearized equations —let us remark that the geometry of the problem plays an important role in the study of these equations. Such approach —also known as KAM theory without action-angle variables— was introduced in [dlLGJV05] for Lagrangian tori and in [FdlLS09] for hyperbolic lower dimensional tori, following long-time developed ideas (relevant work can be found in [JdlLZ99, Rüs76, SZ89, Zeh76] but we refer to [dlL01, dlLGJV05] for a detailed list).

Our aim is to adapt KAM methods without action-angle variables to the normally elliptic context. Concretely, we assume that we have a 1-parameter family of Hamiltonian systems for which we know a 1-parameter family of approximately invariant r-dimensional elliptic tori

[2]Just to illustrate that this task is non-trivial we remark that a rigorous application of KAM theory to the Solar System (asking for arbitrarily small masses) was just given recently in [Féj04]. Moreover, existence of KAM tori was proved in [CC07] for the so-called Restricted Planar Circular Three Body Problem (RPC3BP) with realistic values of the parameter for the system Sun-Jupiter-asteroid (12 Victoria). To carry out these studies it was required to introduce very sophisticated mathematical techniques and some parts in [CC07] required to be computer assisted.

—all of them with the same vector of basic frequencies— and also approximations of the vectors of normal frequencies and the corresponding normal directions associated to these frequencies (i.e., a basis of the normal directions along each torus that approximately reduces the normal variational equations to constant coefficients). Then we show that, under suitable hypothesis of non-resonance and non-degeneracy, for a Cantorian subset of parameters —of large relative Lebesgue measure— there exist true elliptic tori close to the approximate ones, having the same vector of basic frequencies and slightly modified vector of normal frequencies. We have selected this setting —rather than using the basic frequencies as parameters— in order to simplify some (well understood) technical aspects of the problem but all the geometrical ideas linked to parameterizations methods for reducible lower dimensional tori are still present in our approach.

Apart from theoretical results on the existence of quasi-periodic motion, there is also a great interest in the numerical computation and characterization of these objects (from a more practical point of view). As as illustration, let us pay attemption to the following topics:

Frequency Analysis: As we said before, in order to study quasi-periodic invariant tori, valuable information is obtained from the frequency vector that characterizes the motion. For example, the classical *frequency analysis* approach introduced by J. Laskar (see [LFC92]) to obtain an approximation of frequencies is to look for the frequencies as peaks of the modulus of the Discrete Fourier Transform (DFT) of the studied signal. This idea was refined in [GMS] by means of the simultaneous improvement of the frequencies and the amplitudes of the signal. We refer also to the works [TLBB01, RGJ05] for interesting application of Frequency Analysis methods.

Computation of invariant tori: The methods in [CJ00, dlLHS, JO] have been applied efficiently in a wide set of contexts. They require to compute a representation —by means of a trigonometric polynomial— of the curve which solves the invariance equation of the problem, so it is requested to solve large systems of equations —as large as the used number of Fourier modes, say M. One possibility to face this difficulty is to solve these full linear systems, with a cost $\mathcal{O}(M^3)$ in time and $\mathcal{O}(M^2)$ in memory, by means of efficient parallel algorithms as it is proposed in [JO]. Another recent approach presented in [dlLHS], based on the analytic and geometric ideas developed in [dlLGJV05], allows us to reduce the computational effort of the problem to a cost of order $\mathcal{O}(M \log M)$ in time and $\mathcal{O}(M)$ in memory.

In the rest of this section we will focus in the case of 2-dimensional tori for vector fields (we will not assume explicitly that they are Hamiltonian) and we will present a method introduced in [SV06], which is an alternative of the methods discussed above. As we explained before, we can resort to a Poincaré section in order to reduce the study to consider 1-dimensional tori (so-called invariant curves) for maps. For example, let us consider a planar map $F : \mathbb{R}^2 \to \mathbb{R}^2$ having an invariant curve Γ which carries quasi-periodic dynamics of frequency θ (in this context θ is usually called *rotation number* and, if the curve corresponds to a 2-dimensional torus of a flow, it corresponds to the ratio of the frequencies of the torus). Then, we have an embedding $\gamma : \mathbb{T} \to \mathbb{R}$, with $\gamma(\mathbb{T}) = \Gamma$, that makes the following diagram commute

$$\begin{array}{ccc} \Gamma & \xrightarrow{F} & \Gamma \\ {\scriptstyle\gamma}\uparrow & & \uparrow{\scriptstyle\gamma} \\ \mathbb{T} & \xrightarrow{R_\theta} & \mathbb{T} \end{array},$$

that is, we have $f \circ \gamma = \gamma \circ R_\theta$, where $R_\theta(x) = x + \theta$ is a rigid rotation. For this reason, the study of invariant curves has may connections with the study of circle maps. As a matter of motivation, let us assume that F is a map on the real annulus $\mathbb{T} \times I$, where I is a real interval and let $X : \mathbb{T} \times I \to \mathbb{R}$ denote the canonical projection $X(x, y) = x$. For example, if F is a twist map (i.e., if F satisfies that $\partial(X \circ F)/\partial y$ does not vanish), the Birkhoff Graph Theorem (see [Gol01]) ensures that every invariant curve Γ is a graph over its projection on the circle by means of X, and its dynamics induces a circle map by projecting the iterates.

Recently, a new method for computing Diophantine rotation numbers of circle diffeomorphisms with high precision at low computational cost has been introduced in [SV06]. This method is built assuming that the circle map is conjugate to a rigid rotation in a sufficiently smooth way and, basically, it consists in averaging the iterates of the map together with Richardson extrapolation. It has a cost of order $\mathcal{O}(N \log N)$ in terms of the used number of iterates N and free cost in memory, and the construction takes advantage of the geometry and the dynamics of the problem, so it turns out to be very efficient in multiple applications.

Next we briefly summarize the construction of this numerical method (more details will be given in Section II.2). From now on, it is convenient to consider $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ (angles modulo one). Let us denote $\mathrm{Diff}_+^r(\mathbb{T})$, $r \in [0, +\infty) \cup \{\infty, \omega\}$, the group of orientation-preserving homeomorphisms of $\mathbb{T}$ of class $\mathcal{C}^r$. Given $f \in \mathrm{Diff}_+^r(\mathbb{T})$, we identify f with its lift to $\mathbb{R}$ by fixing the normalization condition $f(0) \in [0, 1)$.

Definition I.1.16. *Let f be the lift of an orientation-preserving homeomorphism of the circle such that $f(0) \in [0, 1)$. Then the* rotation number of f *is defined as the limit*

$$\rho(f) := \lim_{|n|\to\infty} \frac{f^n(x_0) - x_0}{n},$$

that exists for all $x_0 \in \mathbb{R}$, is independent of x_0 and satisfies $\rho(f) \in [0, 1)$.

Given $f \in \mathrm{Diff}_+^2(\mathbb{T})$ with $\rho(f) \in \mathbb{R}\backslash\mathbb{Q}$, Denjoy's theorem ensures that f is topologically conjugate to the rigid rotation $R_{\rho(f)}$, i.e., there exist $\eta \in \mathrm{Diff}_+^0(\mathbb{T})$ such that $f \circ \eta = \eta \circ R_{\rho(f)}$, where $R_\theta(x) = x + \theta$. In addition, if we require $\eta(0) = x_0$, for fixed $x_0 \in [0, \infty)$, then η is unique. The theoretical support of the method is provided by the regularity of η, that follows from the next result:

Theorem I.1.17. *Assume that* $f \in \mathrm{Diff}_+^r(\mathbb{T})$ *has Diophantine rotation number* $\theta = \rho(f)$ *of* (C,τ)*-type, i.e., there exist constants* $C > 0$ *and* $\tau \geq 1$ *such that*

$$|1 - \mathrm{e}^{2\pi \mathrm{i} k\theta}|^{-1} \leq C|k|^\tau, \qquad \forall k \in \mathbb{Z}_*.$$

Then, if $\tau + 1 < r$*,* f *is conjugated to* $R_{\rho(f)}$ *by means of a conjugacy* $\eta \in \mathrm{Diff}_+^{r-\tau-\varepsilon}(\mathbb{T})$*, for any* $\varepsilon > 0$*. Note that* $\mathrm{Diff}_+^\omega(\mathbb{T}) = \mathrm{Diff}_+^{\omega-\tau-\varepsilon}(\mathbb{T})$ *while the domain of analyticity is reduced.*

From now on we focus in the analytic case, although finite differentiability is enough. Let us consider $f \in \mathrm{Diff}_+^\omega(\mathbb{T})$ with rotation number $\theta = \rho(f) \in \mathcal{D}$. Notice that we can write $\eta(x) = x + \xi(x)$, being ξ a 1-periodic function normalized in such a way that $\xi(0) = x_0$, for a fixed $x_0 \in [0,1)$. Now, we can write the following expression for the iterates under the lift:

$$f^n(x_0) = f^n(\eta(0)) = \eta(n\theta) = n\theta + \sum_{k \in \mathbb{Z}} \hat{\xi}_k \mathrm{e}^{2\pi \mathrm{i} kn\theta}, \qquad \forall n \in \mathbb{Z}, \tag{I.14}$$

where the sequence $\{\hat{\xi}_k\}_{k\in\mathbb{Z}}$ denotes the Fourier coefficients of ξ. Then, the above expression gives us the following formula

$$\frac{f^n(x_0) - x_0}{n} = \theta + \frac{1}{n} \sum_{k \in \mathbb{Z}_*} \hat{\xi}_k (\mathrm{e}^{2\pi \mathrm{i} kn\theta} - 1),$$

to compute θ modulo terms of order $\mathcal{O}(1/n)$. Unfortunately, this order of convergence is very slow for practical purposes, since it requires a huge number of iterates if we want to compute θ with high precision. Nevertheless, by averaging the iterates $f^n(x_0)$ in a suitable way, we can manage to decrease the order of this quasi-periodic term.

We introduce the following *recursive sums* for $p \in \mathbb{N}$

$$S_N^0(f) := f^N(x_0) - x_0, \qquad S_N^p(f) := \sum_{j=1}^{N} S_j^{p-1}(f). \tag{I.15}$$

Then, the result presented in [SV06] says that, under the previous hypotheses, the following *averaged sums of order* p

$$\widetilde{S}_N^p(f) := \binom{N+p}{p+1}^{-1} S_N^p(f)$$

satisfy the expression (basically, the idea is to use (I.14) and the fact that the Fourier coefficients decay very fast due to the analyticity of η)

$$\widetilde{S}_N^p(f) = \theta + \sum_{l=1}^{p} \frac{A_l^p}{N^l} + E^p(N), \tag{I.16}$$

where the coefficients A_l^p depend on f and p but are independent of N. Furthermore, the remainder $E^p(N)$ is uniformly bounded by an expression of order $\mathcal{O}(1/N^{p+1})$. Formula (I.16) allows us to approximate numerically the rotation number θ by performing the Richardson extrapolation (see Algorithm II.2.5).

Remark I.1.18. *Let us point out that in the averaging procedure there appear factors* $1 - \mathrm{e}^{2\pi \mathrm{i} k\theta}$ *dividing the Fourier coefficients of the involved functions. Diophantine conditions are used to avoid the effect of these small divisors.*

In this thesis we adapt the previous construction in order to compute derivatives of the rotation number in parametric families of circle maps (Chapter II). The obtained methodology leads to a wide set of applications since, using the approximate derivatives of the rotation number, we can continue invariant curves numerically with respect to parameters by means of the Newton method. Furthermore, using the obtained variational information, we are able to compute asymptotic expansions relating parameters and initial conditions that correspond to curves of fixed rotation number.

If we want to apply the previous methodology to a planar map which does not satisfy the twist condition or it is not written in suitable coordinates, its invariant curves are not necessarily graphs over the projection on a circle. In this situation, invariant curves can fold in a very wild way. Hence, another goal of this thesis (discussed in Chapter III) is to propose a numerical method to construct a circle map —preserving the rotation number— from a general invariant curve on the plane. Therefore, we obtain a very efficient toolkit for the study of invariant curves of planar maps and their numerical continuation. Furthermore, since our construction does not use the invariance of the quasi-periodic curve under the map, it can be applied to more general contexts.

I.2 Summary of achieved results

In order to help the reader, in this section we summarize the results achieved in the thesis. We want to remark that Chapters II and III share the same context, but they have been published separately (in [LV08] and [LV09], respectively). For this reason, we have slightly reorganized part of the material presented in these chapters, in order to avoid unnecessary repetitions of concepts. Of course, we have tried that both chapters can still be read independently.

Chapter II: Computation of derivatives of the rotation number for parametric families of circle diffeomorphisms

Let us consider a family $\mu \in I \subset \mathbb{R} \mapsto f_\mu$ of orientation-preserving analytic circle diffeomorphisms depending $\mathcal{C}^d$-smoothly with respect to μ. The rotation numbers (see Definition I.1.16) of the family $\{f_\mu\}_{\mu \in I}$ induce a function $\theta : I \to [0, 1)$ given by $\theta(\mu) = \rho(f_\mu)$. Let us remark that the function θ is continuous but non-smooth: generically, there exist a family of disjoint open intervals of I, with dense union, such that θ takes distinct constant values on these intervals (a so-called Devil's Staircase). However, the derivatives of θ are defined in "many" points (see the discussion in Section II.3 and references given therein).

In order to compute $D^d_\mu\theta$ (the dth derivative of θ) at a point μ_0 where it exists, we define the recursive sums of order p (we omit the notation regarding the fact that the map is evaluated at $\mu = \mu_0$) of the derivatives of the iterates of a given point x_0

$$D^d_\mu S^0_N = D^d_\mu(f^N_\mu(x_0) - x_0), \quad D^d_\mu S^p_N = \sum_{j=0}^{N} D^d_\mu S^{p-1}_j,$$

(here f represents the lift of the circle map) and the corresponding averaged sums

$$D^d_\mu \widetilde{S}^p_N = \binom{N+p}{p+1}^{-1} D^d_\mu S^p_N.$$

Then, if $\theta(\mu_0)$ satisfies Diophantine properties (see Definition II.2.2) and $D^d_\mu\theta(\mu_0)$ exists, we have that

$$D^d_\mu \widetilde{S}^p_N = D^d_\mu\theta + \sum_{l=1}^{p-d} \frac{D^d_\mu \hat{A}^p_l}{N^l} + D^d_\mu \hat{E}^p(N), \tag{I.17}$$

where the remainder $D^d_\mu \hat{E}^p(N)$ is of order $\mathcal{O}(1/N^{p-d+1})$. Equation (I.17) is justified in Section II.3 by adapting the construction in [SV06]. The key idea is to use the fact that the dynamics is conjugate to a rigid rotation (see Theorem I.1.17) in order to express the iterates $f^N_\mu(x_0)$ in terms of the conjugation. By averaging the derivatives of this expression we obtain formula (I.17).

Therefore, according to formula (I.17) we obtain Algorithm II.3.2 to extrapolate the dth derivative of the rotation number. As it is shown in Section II.3.1, when we apply Algorithm II.3.2 for computing the dth derivative of θ, in general, we are forced to select an averaging order $p > d$ and the remainder turns out to be of order $\mathcal{O}(1/N^{p-d+1})$. Nevertheless, if the rotation number is known to be constant as a function of the parameters, we can avoid the previous limitations. Concretely, in this case we can select any averaging order p, independent of d, since the remainder is now of order $\mathcal{O}(1/N^{p+1})$. Of course, if the rotation number is constant, then the derivatives of θ are all zero and the fact that we can obtain them with better precision seems to be irrelevant. Nevertheless, from the computation of these vanishing derivatives, we can obtain information about other involved objects. This is the case of many applications in which this methodology turns out to be very useful

The rest of Chapter II is devoted to illustrate the method through several applications. On the one hand in Section II.4 we study the Arnold family of circle maps, and on the other hand in Section II.5 we focus on the computation and continuation of invariant curves for planar twist maps (we select the conservative Hénon map). Let us describe the performed computations.

- For the Arnold family $x \mapsto f_{\alpha,\varepsilon}(x) = x + 2\pi\alpha + \varepsilon\sin(x)$, we fix $\varepsilon = 0.75$ and compute derivatives of the rotation number along a Devil's Staircase, with a typical error that is less that 10^{-18}. See Figure II.2 for some magnifications of these derivatives.

- The set of parameters (α, ε) for which the rotation number of $f_{\alpha,\varepsilon}$ takes a fixed value is an analytic curve when this value is a Diophantine number. Using the approximate derivatives of the rotation number, we compute these sets (called Arnold Tongues) numerically by means of the Newton method.

- Since an Arnold Tongue is given by the graph of an analytic function $\alpha(\varepsilon)$, for $\varepsilon \in [0, 1)$, we expand α at the origin as

$$\alpha(\varepsilon) = \theta + \frac{\alpha'(0)}{1!}\varepsilon + \frac{\alpha''(0)}{2!}\varepsilon^2 + \cdots + \frac{\alpha^{(d)}(0)}{d!}\varepsilon^d + \mathcal{O}(\varepsilon^{d+1}),$$

and we numerically approximate the terms in this expansion. To this end, we use that the derivative of the rotation number vanishes on the tongue (see details in Section II.4.3). In addition, we use these computations to obtain formulas, depending on θ, for the first coefficients of the Taylor expansion. To make this dependence explicit, we introduce the notation $\alpha_r(\theta) := \alpha^{(2r)}(0)$. In particular, the first two coefficients (that were obtained analytically in [SV09]) are

$$\alpha_1(\theta) = \frac{\cos(\pi\theta)}{2^2\pi\sin(\pi\theta)}, \qquad \alpha_2(\theta) = -\frac{3\cos(4\pi\theta) + 9}{2^5\pi\big(\sin(\pi\theta)\big)^2\sin(2\pi\theta)},$$

and in general, they are given by

$$\alpha_r(\theta) = \frac{P_r(\theta)}{2^{c(r)}\pi\big(\sin(\pi\theta)\big)^{2^{r-1}}\big(\sin(2\pi\theta)\big)^{2^{r-2}}\cdots\big(\sin((r-1)\pi\theta)\big)^2\sin(r\pi\theta)},$$

where $c(r)$ is a natural number and P_r is a trigonometric polynomial of the form

$$P_r(\theta) = \sum_{j=1}^{d_r} a_j\cos(j\pi\theta),$$

with integer coefficients and degree $d_r = 2^{r+1} - r - 2$ that coincides with the degree of the denominator. In addition, the coefficients a_j vanish except for indexes j such that $j \equiv d_r (\bmod 2)$. Concretely, we have found the values $c(3) = 10$, $c(4) = 19$, and $c(5) = 38$. On the other hand, the corresponding polynomials P_r are given in Table II.2. The comparison between these pseudo-analytical coefficients and the values computed numerically for $\theta = (\sqrt{5} - 1)/2$ is shown in column e_2 of Table II.1, obtaining a very good agreement. Let us observe that the coefficients of P_r grow very fast with respect to r, and the same occurs to $c(r)$.

- Let us consider $\alpha : \Lambda \subset \mathbb{R} \mapsto F_\alpha$ a one-parameter family of twist maps on the annulus, that induces a function $(\alpha, y_0) \in U \subset \Lambda \times I \mapsto \rho(f_{\alpha,y_0})$ differentiable in the sense of Whitney. In Section II.5.2 we compute the derivatives of this function (at the points where they exist) and use them to compute numerically invariant curves of F_α by means of the Newton method. As an example we continue the invariant curve of rotation number $\theta = (\sqrt{5}-1)/2$ for the conservative Hénon map with a tolerance of 10^{-23}.

- Finally, if (x_0, y_0^*) is a point on an invariant curve of rotation number θ for a twist map F_{α^*}, then we consider the expansion

$$\alpha(y_0) = \alpha^* + \alpha'(y_0^*)(y_0 - y_0^*) + \frac{\alpha''(y_0^*)}{2!}(y_0 - y_0^*)^2 + \cdots ,$$

  that corresponds to the value of the parameter for which (x_0, y_0) is contained in an invariant curve of $F_{\alpha(y_0)}$ having the same rotation number. Coefficients of this expansion and their estimated error, are given in Table II.3. Finally, in order to verify the results, we compare the truncated expansions of the curve with the numerical approximation computed in Section II.5.2. The deviation is plotted in $\log_{10}$ scale in Figure II.5 (right).

Chapter III: Numerical computation of rotation numbers of quasi-periodic planar curves

We present numerical algorithms to deal with quasi-periodic invariant curves of planar maps by adapting the previous ideas. For convenience, we identify the real plane with the set of complex numbers by defining $z = u + \mathrm{i}v$ for any $(u, v) \in \mathbb{R}^2$. Let $\Gamma \subset \mathbb{C}$ be a quasi-periodic invariant curve for a map $F : U \subset \mathbb{C} \to \mathbb{C}$ of rotation number $\theta \in \mathbb{R}\backslash\mathbb{Q}$. Let us assume, for example, that the curve "rotates" around the origin and that it is a graph of the angular variable. Then, the projection

$$\begin{array}{ccc} \Gamma & \longrightarrow & \mathbb{T} \\ z & \longmapsto & x = \arg(z)/2\pi \end{array} \tag{I.18}$$

generates a circle map induced by the dynamics of $F|_\Gamma$. On the other hand, if Γ is folded, then the projection (I.18) does not provide a circle map, but defines a correspondence on $\mathbb{T}$ that we can "lift" to $\mathbb{R}$. In some cases, for example if the rotation number is large enough as to avoid the folds, we can compute numerically the "lift" of (I.18) using the iterates of an orbit. However, if the invariant curve presents large folds or we cannot identify directly a good point around which the curve is rotating, we cannot compute this "lift" in a systematic way. Then, our aim is to construct another curve, having the same rotation number, by using suitable averages of iterates of the original map. If we manage to eliminate (or at least minimize) the folds in the new curve, then we are able to obtain a circle diffeomorphism (or at least a circle correspondence that we can "lift" numerically).

As Γ is a quasi-periodic invariant curve of rotation number θ, there exists an analytic embedding $\gamma : \mathbb{T} \to \mathbb{C}$ verifying $\Gamma = \gamma(\mathbb{T})$ and

$$F(\gamma(x)) = \gamma(x + \theta).$$

In this situation, we consider the iterates $z_n = F^n(z_0)$ that can be expressed using γ as

$$z_n = F^n(z_0) = F^n(\gamma(0)) = F^{n-1}(\gamma(\theta)) = \gamma(n\theta) = \sum_{k\in\mathbb{Z}} \hat{\gamma}_k \mathrm{e}^{2\pi\mathrm{i}kn\theta}.$$

Then, for any $\theta_0 \in \mathbb{R}$ and $L \in \mathbb{N}$, we introduce the following iterates

$$z_n^{(L,\theta_0)} = \frac{1}{L} \sum_{m=n}^{L+n-1} z_m \mathrm{e}^{2\pi\mathrm{i}(n-m)\theta_0}.$$

These new iterates correspond to a curve that is not necessarily embedded in $\mathbb{C}$ nor invariant under any map (we refer to such objects as *quasi-periodic signals*), but still carries quasi-periodic dynamics with the same rotation number, i.e., there exists a curve $\Gamma^{(L,\theta_0)} = \gamma^{(L,\theta_0)}(\mathbb{T})$ such that

$$z_n^{(L,\theta_0)} = \gamma^{(L,\theta_0)}(n\theta) = \sum_{k\in\mathbb{Z}} \hat{\gamma}_k^{(L,\theta_0)} \mathrm{e}^{2\pi\mathrm{i}kn\theta},$$

and the new Fourier coefficients are given by

$$\hat{\gamma}_k^{(L,\theta_0)} = \frac{\hat{\gamma}_k}{L} \sum_{m=n}^{L+n-1} \mathrm{e}^{2\pi\mathrm{i}(m-n)(k\theta-\theta_0)} = \frac{\hat{\gamma}_k}{L} \frac{1-\mathrm{e}^{2\pi\mathrm{i}L(k\theta-\theta_0)}}{1-\mathrm{e}^{2\pi\mathrm{i}(k\theta-\theta_0)}},$$

In Section III.2.2 we show that, under conditions on regularity and non-resonance, if $\theta_0 = \theta$, then the new curve $\gamma^{(L,\theta)}$ is arbitrarily $\mathcal{C}^1$-close to a circle (see Lemma III.2.6) for L large enough. On the other hand, if $\varepsilon = \theta_0 - \theta$ is small, then we can choose $L = L(\varepsilon)$ such that the new curve is $\mathcal{C}^1$-close to a circle with an error of order $\mathcal{O}(\varepsilon)$ (this is concluded from Proposition III.2.8) so that the projection

$$\begin{array}{rcl} \Gamma^{(L,\theta_0)} \subset \mathbb{C}^* & \longrightarrow & \mathbb{T} \\ z_n^{(L,\theta_0)} & \longmapsto & x_n^{(L,\theta_0)} = \arg(z_n^{(L,\theta_0)})/2\pi, \end{array}$$

provides an orbit of a circle diffeomorphism $f_\Gamma^{(L,\theta_0)}$. Once this circle map has been obtained (as we have discussed, in practice it suffices to obtain a slightly folded curve such that we can compute the “lift” of the circle correspondence defined by the direct projection), we can apply the methodology exposed in Chapter II to compute the rotation number and derivatives with respect to parameters (this is described in Section III.2.5). In order to justify this construction, we require the rotation number to be Diophantine (see Definition II.2.2).

Let us remark that, in order to apply our method, we require an approximation θ_0 of the rotation number. Sometimes this is given by the context of the problem —for example, if we look for invariant curves of fixed rotation number— or it can be obtained by means of any method of frequency analysis (see for example [GMS, LFC92]). Furthermore, our construction gives us a heuristic method for computing an approximation θ_0 of the rotation number θ (see Section III.2.4).

Moreover, in the same way that the method of [SV06] accelerates the convergence of the rotation number from $\mathcal{O}(1/N)$ (suing the definition) to $\mathcal{O}(1/N^{p+1})$, we introduce higher order averages to the iterates $z_n^{(L,\theta_0)}$ to accelerate the convergence of the new curve to a circle. Concretely, by performing averages of order p we improve the rate of convergence from $\mathcal{O}(1/L)$ to $\mathcal{O}(1/L^p)$. The goal of Section III.2.6 is to justify these higher order averages.

In Section III.2.7 we also adapt the previous methodology in order to obtain the Fourier coefficients of a quasi-periodic signal of known rotation number. Let us recall that standard FFT algorithms are based in equidistant samples of points. Since the iterates of a quasi-periodic signal are not distributed in such a way on $\mathbb{T}$, one has to implement a non-equidistant FFT or resort to interpolation of points (see for instance [AdlLP05, dlLP02, OP08]). We avoid this difficulty using the fact that the iterates are equidistant "according with the quasi-periodic dynamics". In particular, we adapt these extrapolation method in order to compute Fourier coefficients with and error of order $\mathcal{O}(1/N^p)$, where N is the used number of iterates.

We illustrate several features of the methods developed in this chapter along Section III.3. Now we describe the computations that we have performed in several contexts:

- We consider invariant curves in the Siegel domain of $F(z) = \lambda(z - \frac{1}{2}z^2)$,with $\lambda = \mathrm{e}^{2\pi \mathrm{i}\theta}$, for several rotation numbers θ (see details in Section III.3.1). Even in this simple example, the direct projection on the angular variable does not always give a diffeomorphism on $\mathbb{T}$ (see Figure III.3). Our aim is to emphasize the difficulty of unfolding an invariant curve depending on its rotation number. To this end, we propose simple criteria to decide if the curve is already unfolded or not.

- We consider a quasi-periodic signal $z_n = \gamma(n\theta)$, with $\theta = (\sqrt{5}-1)/2 \in \mathcal{D}$, given by

$$\gamma(x) = \hat{\gamma}_{-1}\mathrm{e}^{-2\pi \mathrm{i}x} + \hat{\gamma}_0 + \hat{\gamma}_1\mathrm{e}^{2\pi \mathrm{i}x} + \hat{\gamma}_2\mathrm{e}^{4\pi \mathrm{i}x}, \tag{I.19}$$

  where the Fourier coefficients of γ (see Table III.1) have been selected in such a way that the curve $\gamma(\mathbb{T})$ intersects itself. First, we unfold this invariant curve and compute its non-zero Fourier coefficients. We also study the dependence with respect to L of the norms in Lemma III.2.6 and Proposition III.2.8.

- It is very interesting to study the effect of a random error in the evaluation of iterates, trying to simulate that the source of our quasi-periodic signal is experimental data. Concretely, we consider the iterates $z_n = \gamma(n\theta) + \varepsilon x_n$, where γ is (I.19) and the real and the imaginary parts of the noise x_n are normally distributed with zero mean and unit variance. Of course, the new iterates z_n do not belong to a curve but they are distributed in a cloud around the curve γ (see the left plot in Figure III.5). If we compute the iterates $z_n^{(L,\theta_0)}$, then it turns out that we can "unfold" the cloud of points in a similar way. We also focus on the effect of this external noise when computing the rotation number θ of the "circle map" thus obtained.

- Let us consider the family of area preserving non-twist maps given by

$$F_{a,b} : (x, y) \longmapsto (\bar{x}, \bar{y}) = \big(x + a(1 - \bar{y}^2), y - b \sin(2\pi x)\big)\,,$$

  where $(x, y) \in \mathbb{T} \times \mathbb{R}$ are phase space coordinates and a, b parameters. This family is usually called *standard non-twist map* and it is studied as a paradigmatic example of a non-twist family. We study invariant curves for this map close to a reconnection scenario; and we focus on the shearless curve (this concept is introduced at the beginning of Section III.3.3) computing its rotation number by applying the extrapolation method of Algorithm II.2.5 to the circle correspondence obtained by direct projection on the angular variable x (see Table III.2) In addition, we illustrate the methodology of Section III.2 to unfold the shearless invariant curve.

- A further step is to consider the case of the so-called *higher order meanderings* that appear due to reconnections involving periodic orbits in a neighborhood of a meandering curve. The considered example is selected from [Sim98b] and we refer there for a constructive explanation (see the curve in the left plot of Figure III.10). We apply Laskar's method of frequency analysis (implemented as described in Section III.2.4) to obtain a rough approximation of the rotation number in order to unfold the curve. Finally, we compute the rotation number of this invariant curve from the circle correspondence obtained by means of the unfolding procedure.

- We also illustrate the computation of derivatives of the rotation number by applying Algorithm II.3.2 to a circle correspondence that we can "lift" numerically to $\mathbb{R}$. To this end, we use the Hénon family.

- Finally, we consider the case of *labyrinthic curves*, that follow arbitrarily complicated paths in phase space (we refer to [Sim98b]). We first obtain an approximation of the rotation number by means of frequency analysis as is explained in Section III.2.4. Then, we apply Algorithm II.2.5 to compute the rotation number with an estimated extrapolation error of order 10^{-34}, that corresponds with the arithmetic precision of our computations. Furthermore, we use the Newton method to follow the evolution of an invariant curve having a prefixed rotation number.

Chapter IV: A KAM theorem without action-angle variables for elliptic lower dimensional tori

Here we summarize informally the result obtained in Chapter IV. We refer to Section IV.2 for the basic terminology and objects related to the problem.

Let us consider a 1-parameter family of Hamiltonian systems $\mu \in I \subset \mathbb{R} \mapsto h_\mu$ with $h_\mu : U \subset \mathbb{R}^{2n} \to \mathbb{R}$, where I is a finite interval and U is an open set. Let us assume that for these Hamiltonian systems we have a 1-parameter family of approximately (lower dimensional) elliptic invariant tori with vector of basic frequencies $\omega \in \mathbb{R}^r$ and approximated vector of normal frequencies $\lambda_\mu \in \mathbb{R}^{n-r}$, i.e., assume that there exist a 1-parameter family of embeddings

$$\mu \in I \subset \mathbb{R} \longmapsto \tau_\mu, \qquad \tau_\mu : \mathbb{T}^r \longrightarrow \tau_\mu(\mathbb{T}^r) \subset U,$$

matrix functions

$$\mu \in I \subset \mathbb{R} \longmapsto N_\mu, \qquad N_\mu : \mathbb{T}^r \longrightarrow M_{2n\times(n-r)}(\mathbb{C}),$$

and approximated normal eigenvalues $\Lambda_\mu = \mathrm{diag}\,(\mathrm{i}\lambda_\mu)$, satisfying

$$\begin{aligned} L_\omega \tau_\mu(\theta) &= J\mathrm{grad}\, h_\mu(\tau_\mu(\theta)) + e_\mu(\theta), \\ L_\omega N_\mu(\theta) &= J\mathrm{hess}\, h_\mu(\tau_\mu(\theta)) N_\mu(\theta) - N_\mu(\theta)\Lambda_\mu + R_\mu(\theta). \end{aligned}$$

Next, we formulate the conditions on τ_μ, N_μ, ω and λ_μ which allow us to prove, using the Newton method, that there exists a large set of parameters for which we have a true invariant tori close to the corresponding approximate one (see Theorem IV.3.1). We emphasize that we do not assume that the system is given in action-angle-like coordinates nor that the Hamiltonians are close to integrable. Let us assume that the following hypotheses hold.

- The previous objects are analytic and (for certain norms to be described in Section IV.2.1) we can control $\|D\tau\|_{I,\rho}$, $\|N\|_{I,\rho}$, $\|G^{-1}_{D\tau}\|_{I,\rho}$ and $\|G^{-1}_{N,N^*}\|_{I,\rho}$.

- The basic frequencies satisfy Diophantine conditions for $\hat{\gamma} > 0$ and $\nu > r-1$, i.e.

$$|\langle k, \omega\rangle| \geq \frac{\hat{\gamma}}{|k|_1^\nu}, \qquad \forall k \in \mathbb{Z}^r \backslash \{0\}.$$

- We have $\mathrm{diag}\,\left[\Omega_{N_\mu, N_\mu^*}\right]_{\mathbb{T}^r} = \mathrm{i}\mathrm{Id}_{n-r}$ for every $\mu \in I$.

- The family of matrix functions

$$A_{1,\mu}(\theta) := G^{-1}_{D\tau_\mu}(\theta) D\tau_\mu(\theta)^\top (T_{1,\mu}(\theta) + T_{2,\mu}(\theta) + T_{2,\mu}(\theta)^\top) D\tau(\theta) G^{-1}_{D\tau_\mu}(\theta),$$

  where

$$\begin{aligned} T_{1,\mu}(\theta) &:= J^\top \mathrm{hess}\, h_\mu(\tau_\mu(\theta)) J - \mathrm{hess}\, h_\mu(\tau_\mu(\theta)), \\ T_{2,\mu}(\theta) &:= T_{\mu,1}(\theta) J \left[D\tau_\mu(\theta) G_{D\tau_\mu}(\theta)^{-1} D\tau_\mu(\theta)^\top - \mathrm{Id}\right] \mathrm{Re}(\mathrm{i}N_\mu(\theta) N_\mu^*(\theta)^\top), \end{aligned}$$

  is non-singular.

- The objects h_μ, τ_μ, N_μ and λ_μ are at least $\mathcal{C}^1$ with respect to μ, and we can control

$$\left\|\frac{dh}{d\mu}\right\|_{I,\mathcal{C}^3,\mathcal{U}}, \left\|\frac{dD\tau}{d\mu}\right\|_{I,\rho}, \left\|\frac{dN}{d\mu}\right\|_{I,\rho}, \left\|\frac{d\lambda_i}{d\mu}\right\|_{I} < \infty.$$

  for $i = 1, \dots, n-r$. Moreover, the approximated normal frequencies λ_μ satisfy for every $\mu \in I$

$$0 < |\lambda_{i,\mu}| < \infty, \qquad 0 < |\lambda_{i,\mu} \pm \lambda_{j,\mu}|,$$

  for $i, j = 1, \dots, n-r$, with $i \neq j$ and we also have the next separation condition

$$0 < \left|\frac{d}{d\mu}\lambda_{i,\mu}\right|, \qquad 0 < \left|\frac{d}{d\mu}\lambda_{i,\mu} \pm \frac{d}{d\mu}\lambda_{j,\mu}\right|,$$

  for $i, j = 1, \dots, n-r$, with $i \neq j$.

Under these assumptions, if

$$\varepsilon = \frac{16}{\rho}\left(\|e\|_{I,\rho} + \left\|\frac{de}{d\mu}\right\|_{I,\rho}\right) + \|R\|_{I,\rho} + \left\|\frac{dR}{d\mu}\right\|_{I,\rho}$$

is sufficiently small (depending on the previous conditions), then there exist a Cantorian subset $I_{(\infty)} \subset I$ such that $\forall \mu \in I_{(\infty)}$ the Hamiltonian h_μ has an r-dimensional elliptic invariant torus $\mathscr{T}_{\mu,(\infty)}$ with basic frequencies ω and normal frequencies $\lambda_{\mu,(\infty)}$ that satisfy Diophantine conditions

$$|\langle k, \omega\rangle + \lambda_{i,(\infty)}| \geq \frac{C\varepsilon^{1/10}}{|k|_1^\nu}, \qquad |\langle k, \omega\rangle + \lambda_{i,(\infty)} \pm \lambda_{j,(\infty)}| \geq \frac{C\varepsilon^{1/10}}{|k|_1^\nu}, \qquad \forall k \in \mathbb{Z}^r \backslash \{0\}.$$

Moreover, $I_{(\infty)}$ has big relative Lebesgue measure

$$\mathrm{meas}_{\mathbb{R}}(I \backslash I_{(\infty)}) \leq \alpha \varepsilon^\beta$$

where $\beta = \frac{3\nu - 2r + 2}{10\nu}$ and C, α depend on the previous conditions. Finally, we have that the obtained objects $\tau_{(\infty)}$, $N_{(\infty)}$ and $\lambda_{(\infty)}$ are just slightly changed.

As it is customary in parameterization methods, the conditions of this result depend only on the initial approximate solution. This fact is useful in the validation of numerical computations that typically consist in looking for trigonometric functions that satisfy invariance and reducibility equations approximately.

Chapter II

Computation of derivatives of the rotation number for parametric families of circle diffeomorphisms

Abstract of the chapter. *In this chapter we present a numerical method to compute derivatives of the rotation number for parametric families of circle diffeomorphisms with high accuracy. Our methodology is an extension of a recently developed approach to compute rotation numbers based on suitable averages of the iterates of the map and Richardson extrapolation. We focus on analytic circle diffeomorphisms, but the method also works if the maps are differentiable enough. In order to justify the method, we also require the family of maps to be differentiable with respect to the parameters and the rotation number to be Diophantine. In particular, the method turns out to be very efficient for computing Taylor expansions of Arnold Tongues of families of circle maps. Finally, we adapt these ideas to study invariant curves for parametric families of planar twist maps.*

II.1 Introduction

The rotation number, introduced by Poincaré, is an important topological invariant in the study of the dynamics of circle maps and, by extension, invariant curves for maps or two dimensional invariant tori for vector fields. For this reason, several numerical methods for approximating rotation numbers have been developed during the last years. We refer to the works [BS00, Bru92, dlLP02, GJSM01b, GMS, LFC92, Pav95, vV88] as examples of methods of different nature and complexity. This last ranges from the pure definition of the rotation number to sophisticated and involved methods like frequency analysis. The efficiency of these methods varies depending if the approximated rotation number is rational or irrational. Moreover, even though some of them can be very accurate in many cases, they are not adequate for every kind of application, for example due to the violation of their assumptions or due to practical reasons, like the required amount of memory.

Recently, a new method for computing Diophantine rotation numbers of circle diffeomorphisms with high precision at low computational cost has been introduced in [SV06]. This method is built assuming that the circle map is conjugate to a rigid rotation in a sufficiently smooth way and, basically, it consists in averaging the iterates of the map together with Richardson extrapolation. This construction takes advantage of the geometry and the dynamics of the problem, so it results very efficient in multiple applications. The method is specially suited if we are able to compute the iterates of the map with a high precision, for example if we can work with a computer arithmetic having a large number of decimal digits.

The goal of this chapter is to extend the method of [SV06] in order to compute derivatives of the rotation number with respect to parameters in families of circle diffeomorphisms. We follow the same averaging-extrapolation process applied to the derivatives of the iterates of the map. To this end, we require the family to be differentiable with respect to the parameters. Hence, we are able to obtain accurate variational information at the same time that we approximate the rotation number. Consequently, the method allows us to study parametric families of circle maps from a point of view that is not given by any of the previously mentioned methods.

From the practical point of view, circle diffeomorphisms appear in the study of quasi-periodic invariant curves for maps. In particular, for planar twist maps, any such a curve induces a circle diffeomorphism in a direct way just by projecting the iterates on the angular variable. Then, using the approximate derivatives of the rotation number, we can continue these invariant curves numerically with respect to parameters by means of the Newton method. The methodology presented is an alternative to more common approaches based on solving numerically the invariance equation, interpolation of the map or approximation by periodic orbits (see for example [CJ00, dlLGJV05, GJMS93, Sim98a]). Furthermore, using the obtained variational information, we are able to compute asymptotic expansions relating parameters and initial conditions that correspond to curves of fixed rotation number.

Finally, we point out that the method can be formally extended to deal with maps on the torus with Diophantine rotation vector. However, in order to apply the method to the study of quasi-periodic tori for symplectic maps in higher dimension, there is not an analogue of the twist condition to guarantee a well defined projection of the iterates on the standard torus. Then, the immediate interest is focused in the generalization of the method in the case of non-twist maps with folded invariant curves (for example, the so-called meanderings [Sim98b]). These and other extensions will be discussed in Chapter III.

The contents of the chapter are organized as follows. In Section II.2 we recall some fundamental facts about circle maps and we briefly review the method of [SV06]. In Section II.3 we describe the method for the computation of derivatives of the rotation number. The rest of the chapter is devoted to illustrate the method through several applications. Concretely, in Section II.4 we study the Arnold family of circle maps. Finally, in Section II.5 we focus on the computation and continuation of invariant curves for planar twist maps and, in particular, we present some computations for the conservative Hénon map.

II.2 Notation and previous results

All the results presented in this section can be found in the bibliography, but we include them for self-consistence of the text. Concretely, in Section II.2.1 we recall the basic definitions, notations and properties of circle maps that we need in the chapter (we refer to [dMvS93, KH95] for more details and proofs). On the other hand, in Section II.2.2 we review briefly the method of [SV06] for computing rotation numbers of circle diffeomorphisms.

II.2.1 Circle diffeomorphisms

Let $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ be the real circle which inherits both a group structure and a topology by means of the natural projection $\pi : \mathbb{R} \to \mathbb{T}$ (also called the universal cover of $\mathbb{T}$). We denote by $\mathrm{Diff}^r_+(\mathbb{T})$, $r \in [0, +\infty) \cup \{\infty, \omega\}$, the group of orientation-preserving homeomorphisms of $\mathbb{T}$ of class $\mathcal{C}^r$. Concretely, if $r = 0$ it is the group of homeomorphisms of $\mathbb{T}$; if $r \geq 1$, $r \in (0, \infty)\backslash\mathbb{N}$, it is the group of $\mathcal{C}^{\lfloor r \rfloor}$-diffeomorphisms whose $\lfloor r \rfloor$th derivative verifies a Hölder condition with exponent $r - \lfloor r \rfloor$; if $r = \omega$ it is the group of real analytic diffeomorphisms.

Given $f \in \mathrm{Diff}^r_+(\mathbb{T})$, we can lift f to $\mathbb{R}$ by π obtaining a $\mathcal{C}^r$ map $\widetilde{f}$ that makes the following diagram commute

$$\begin{array}{ccc} \mathbb{R} & \xrightarrow{\widetilde{f}} & \mathbb{R} \\ {\scriptstyle\pi}\downarrow & & \downarrow{\scriptstyle\pi} \\ \mathbb{T} & \xrightarrow{f} & \mathbb{T} \end{array} \qquad\qquad \pi \circ \widetilde{f} = f \circ \pi.$$

Moreover, we have $\widetilde{f}(x + 1) - \widetilde{f}(x) = 1$ (since f is orientation-preserving) and the lift is unique if we ask for $\widetilde{f}(0) \in [0, 1)$. Accordingly, from now on we choose the lift with this normalization so we can omit the tilde without any ambiguity.

Definition II.2.1. *Let f be the lift of an orientation-preserving homeomorphism of the circle such that $f(0) \in [0, 1)$. Then the* rotation number *of f is defined as the limit*

$$\rho(f) := \lim_{|n|\to\infty} \frac{f^n(x_0) - x_0}{n}, \tag{II.1}$$

that exists for all $x_0 \in \mathbb{R}$, is independent of x_0 and satisfies $\rho(f) \in [0, 1)$.

Let us remark that the rotation number is invariant under orientation-preserving conjugation, i.e., for every $f, h \in \mathrm{Diff}^0_+(\mathbb{T})$ we have that $\rho(h^{-1} \circ f \circ h) = \rho(f)$. Furthermore, given $f \in \mathrm{Diff}^2_+(\mathbb{T})$ with $\rho(f) \in \mathbb{R}\backslash\mathbb{Q}$, Denjoy's theorem ensures that f is topologically conjugate to the rigid rotation $R_{\rho(f)}$, where $R_\theta(x) = x + \theta$. That is, there exists $\eta \in \mathrm{Diff}^0_+(\mathbb{T})$ making the following diagram commute

$$\begin{array}{ccc} \mathbb{T} & \xrightarrow{f} & \mathbb{T} \\ \eta \uparrow & & \uparrow \eta \\ \mathbb{T} & \xrightarrow{R_{\rho(f)}} & \mathbb{T} \end{array} \qquad f \circ \eta = \eta \circ R_{\rho(f)}. \tag{II.2}$$

In addition, if we require $\eta(0) = x_0$, for fixed x_0, then the conjugacy η is unique.

All the ideas and algorithms described in this chapter make use of the existence of such conjugation and its regularity. Let us remark that, although smooth or even finite differentiability is enough, in this chapter we are concerned with the analytic case. Moreover, it is well known that regularity of the conjugation depends also on the rational approximation properties of $\rho(f)$, so we will focus on Diophantine numbers.

Definition II.2.2. *Given $\theta \in \mathbb{R}$, we say that θ is a* Diophantine number *of (C, τ) type if there exist constants $C > 0$ and $\tau \geq 1$ such that*

$$\left|1 - \mathrm{e}^{2\pi \mathrm{i} k\theta}\right|^{-1} \leq C|k|^{\tau}, \qquad \forall k \in \mathbb{Z}_*.$$

We will denote by $\mathcal{D}(C, \tau)$ the set of such numbers and by $\mathcal{D}$ the set of Diophantine numbers of any type.

Although Diophantine sets are Cantorian (i.e., compact, perfect and nowhere dense) a remarkable property is that $\mathbb{R}\backslash\mathcal{D}$ has zero Lebesgue measure. For this reason, this condition fits very well in practical issues and we do not resort to other weak conditions on small divisors such as the Brjuno condition (see [Yoc84a]).

The first result on the regularity of the conjugacy (II.2) is due to Arnold [Arn61] but we also refer to [Her79, KO89, SK89, Yoc84a] for later contributions. In particular, theoretical support of the methodology presented in this chapter is provided by the following result:

Theorem II.2.3 (Katznelson and Ornstein [KO89])**.** *If $f \in \mathrm{Diff}_+^r(\mathbb{T})$ has Diophantine rotation number $\rho(f) \in \mathcal{D}(C, \tau)$ for $\tau + 1 < r$, then f is conjugated to $R_{\rho(f)}$ by means of a conjugacy $\eta \in \mathrm{Diff}_+^{r-\tau-\varepsilon}(\mathbb{T})$, for any $\varepsilon > 0$. Note that $\mathrm{Diff}_+^{\omega}(\mathbb{T}) = \mathrm{Diff}_+^{\omega-\tau-\varepsilon}(\mathbb{T})$ while the domain of analyticity is reduced.*

II.2.2 Computing rotation numbers by averaging and extrapolation

We review here the method developed in [SV06] for computing Diophantine rotation numbers of analytic circle diffeomorphisms (the $\mathcal{C}^r$ case is similar). This method is highly accurate with low computational cost and it turns out to be very efficient when combined with multiple precision arithmetic routines. The reader is referred there for a detailed discussion and several applications.

Let us consider $f \in \mathrm{Diff}_+^\omega(\mathbb{T})$ with rotation number $\theta = \rho(f) \in \mathcal{D}$. Notice that we can write the conjugacy of Theorem II.2.3 as $\eta(x) = x + \xi(x)$, ξ being a 1-periodic function normalized in such a way that $\xi(0) = x_0$, for a fixed $x_0 \in [0,1)$. Now, by using the fact that η conjugates f to a rigid rotation, we can write the following expression for the iterates under the lift:

$$f^n(x_0) = f^n(\eta(0)) = \eta(n\theta) = n\theta + \sum_{k\in\mathbb{Z}} \hat{\xi}_k \mathrm{e}^{2\pi \mathrm{i} k n\theta}, \qquad \forall n \in \mathbb{Z}, \tag{II.3}$$

where the sequence $\{\hat{\xi}_k\}_{k\in\mathbb{Z}}$ denotes the Fourier coefficients of ξ. Then, the above expression gives us the following formula

$$\frac{f^n(x_0) - x_0}{n} = \theta + \frac{1}{n} \sum_{k\in\mathbb{Z}_*} \hat{\xi}_k (\mathrm{e}^{2\pi \mathrm{i} k n\theta} - 1),$$

to compute θ modulo terms of order $\mathcal{O}(1/n)$. Unfortunately, this order of convergence is very slow for practical purposes, since it requires a huge number of iterates if we want to compute θ with high precision. Nevertheless, by averaging the iterates $f^n(x_0)$ in a suitable way, we can manage to decrease the order of this quasi-periodic term.

As a motivation, let us start by considering the sum of the first N iterates under f, that has the following expression (we use (II.3) to write the iterates)

$$S_N^1(f) := \sum_{n=1}^{N} (f^n(x_0) - x_0) = \frac{N(N+1)}{2}\theta - N \sum_{k\in\mathbb{Z}_*} \hat{\xi}_k + \sum_{k\in\mathbb{Z}_*} \hat{\xi}_k \frac{\mathrm{e}^{2\pi \mathrm{i} k\theta}(1 - \mathrm{e}^{2\pi \mathrm{i} k N\theta})}{1 - \mathrm{e}^{2\pi \mathrm{i} k\theta}},$$

and we observe that the new factor multiplying θ grows quadratically with the number of iterates, while it appears a linear term in N with constant $A_1 = -\sum_{k\in\mathbb{Z}_*} \hat{\xi}_k$. Moreover, the quasi-periodic sum remains uniformly bounded since θ is Diophantine and η is analytic (use Lemma II.2.4 with $p = 1$). Thus, we obtain

$$\frac{2}{N(N+1)} S_N^1(f) = \theta + \frac{2}{N+1} A_1 + \mathcal{O}(1/N^2), \tag{II.4}$$

that allows us to extrapolate the value of θ with an error $\mathcal{O}(1/N^2)$ if, for example, we compute $S_N^1(f)$ and $S_{2N}^1(f)$.

In general, we introduce the following *recursive sums* for $p \in \mathbb{N}$

$$S_N^0(f) := f^N(x_0) - x_0, \qquad S_N^p(f) := \sum_{j=1}^{N} S_j^{p-1}(f). \tag{II.5}$$

Then, the result presented in [SV06] says that under the above hypotheses, the following *averaged sums of order* p

$$\widetilde{S}_N^p(f) := \binom{N+p}{p+1}^{-1} S_N^p(f) \tag{II.6}$$

satisfy the expression

$$\widetilde{S}_N^p(f) = \theta + \sum_{l=1}^{p} \frac{A_l^p}{(N+p-l+1)\cdots(N+p)} + E^p(N), \tag{II.7}$$

where the coefficients A_l^p depend on f and p but are independent of N. Furthermore, we have the following expressions for them

$$A_l^p = (-1)^l (p-l+2)\cdots(p+1) \sum_{k\in\mathbb{Z}_*} \hat{\xi}_k \frac{\mathrm{e}^{2\pi \mathrm{i} k(l-1)\theta}}{(1-\mathrm{e}^{2\pi\mathrm{i}k\theta})^{l-1}},$$

$$E^p(N) = (-1)^{p+1} \frac{(p+1)!}{N\cdots(N+p)} \sum_{k\in\mathbb{Z}_*} \hat{\xi}_k \frac{\mathrm{e}^{2\pi\mathrm{i}kp\theta}(1-\mathrm{e}^{2\pi\mathrm{i}kN\theta})}{(1-\mathrm{e}^{2\pi\mathrm{i}k\theta})^p}.$$

Finally, the remainder $E^p(N)$ is uniformly bounded by an expression of order $\mathcal{O}(1/N^{p+1})$. This follows immediately from the next standard lemma on small divisors.

Lemma II.2.4. *Let $\xi \in \mathrm{Diff}_+^\omega(\mathbb{T})$ be a circle map that can be extended analytically to a complex strip $B_\Delta = \{z \in \mathbb{C} : |\mathrm{Im}(z)| < \Delta\}$, with $|\xi(z)| \leq M$ up to the boundary of the strip. If we denote $\{\hat{\xi}_k\}_{k\in\mathbb{Z}}$ the Fourier coefficients of ξ and consider $\theta \in \mathcal{D}(C,\tau)$, then for any fixed $p \in \mathbb{N}$ we have*

$$\left| \sum_{k\in\mathbb{Z}_*} \hat{\xi}_k \frac{\mathrm{e}^{2\pi\mathrm{i}kp\theta}(1-\mathrm{e}^{2\pi\mathrm{i}kN\theta})}{(1-\mathrm{e}^{2\pi\mathrm{i}k\theta})^p} \right| \leq \frac{\mathrm{e}^{-\pi\Delta}}{1-\mathrm{e}^{-\pi\Delta}} 4MC^p \left(\frac{\tau p}{\pi\Delta\mathrm{e}}\right)^{\tau p}.$$

To conclude this survey, we describe the implementation of the method and discuss the expected behavior of the extrapolation error. In order to make Richardson extrapolation we assume, for simplicity, that the total number of iterates is a power of two. Concretely, we select an averaging order $p \in \mathbb{N}$, a maximum number of iterates $N = 2^q$, for some $q \geq p$, and compute the averaged sums $\{\widetilde{S}_{N_j}^p(f)\}_{j=0,\ldots,p}$ with $N_j = 2^{q-p+j}$. Then, we can use formula (II.7) to obtain θ by neglecting the remainders $E^p(N_j)$ and solving the resulting linear set of equations for the unknowns $\theta, A_1^p, \ldots, A_p^p$.

However, let us point out that, due to the denominators $(N_j+p-l+1)\cdots(N_j+p)$, the matrix of this linear system depends on q, and this is inconvenient if we want to repeat the computations using different number of iterates. Nevertheless, we note that expression (II.7) can be written alternatively as

$$\widetilde{S}_N^p(f) = \theta + \sum_{l=1}^{p} \frac{\hat{A}_l^p}{N^l} + \hat{E}^p(N), \tag{II.8}$$

for certain $\{\hat{A}_l^p\}_{l=1,\ldots,p}$, also independent of N, and with a new remainder $\hat{E}^p(N)$ that differs from $E^p(N)$ only by terms of order $\mathcal{O}(1/N^{p+1})$. Then, by neglecting the remainder $\hat{E}^p(N)$

in (II.8), we can obtain θ by solving a new $(p+1)$-dimensional system of equations, independent of q, for the unknowns $\theta, \hat{A}_1^p/2^{1(q-p)}, \ldots, \hat{A}_p^p/2^{p(q-p)}$. Therefore, the rotation number can be computed as follows

$$\theta = \Theta_{q,p}(f) + \mathcal{O}(2^{-(p+1)q}), \tag{II.9}$$

where $\Theta_{q,p}$ is an *extrapolation operator,* that is given by

$$\Theta_{q,p}(f) := \sum_{j=0}^{p} c_j^p \widetilde{S}^p_{2^{q-p+j}}(f), \tag{II.10}$$

and the coefficients $\{c_j^p\}_{j=0,\ldots,p}$ are

$$c_l^p = (-1)^{p-l} \frac{2^{l(l+1)/2}}{\delta(l)\delta(p-l)}, \tag{II.11}$$

where we define $\delta(n) := (2^n-1)(2^{n-1}-1)\cdots(2^1-1)$ for $n \geq 1$ and $\delta(0) := 1$. The previous discussion is summarized in the following algorithm.

Algorithm II.2.5. *Once an averaging order p is selected, we take $N = 2^q$ iterates of the map, for some $q > p$, and compute the sums $\{\widetilde{S}^p_{N_j}\}_{j=0,\ldots,p}$ with $N_j = 2^{q-p+j}$. We approximate the rotation number using the formula*

$$\theta = \Theta_{q,p}(f) + \mathcal{O}(2^{-(p+1)q}), \quad \Theta_{q,p}(f) = \sum_{j=0}^{p} c_j^{(p)} \widetilde{S}^p_{2^{q-p+j}}, \tag{II.12}$$

where the coefficients $c_j^{(p)}$ are given by (II.11)*. The operator $\Theta_{q,p}$ corresponds to the Richardson extrapolation of order p of equation* (II.8)*.*

Remark II.2.6. *Note that the dimension of this linear system and the asymptotic behavior of the error only depend on the averaging order p. For this reason, in [SV06] p is called the extrapolation order. However, this is not always the case when computing derivatives of the rotation number. As we discuss in Section II.3, the extrapolation order is in general less than the averaging order.*

As far as the behavior of the error is concerned, if we fix the extrapolation order p, using (II.9) we have that

$$|\theta - \Theta_{q,p}(f)| \leq c/2^{q(p+1)},$$

for certain (unknown) constant c independent of q, that we estimate heuristically as follows. To estimate c, we compute $\Theta_{q-1,p}$ and consider the expression $|\theta - \Theta_{q-1,p}| \leq c/2^{(q-1)(p+1)}$. Then, we replace in this inequality the exact value of θ by $\Theta_{q,p}$, as we expect $\Theta_{q,p}$ to be closer to θ than $\Theta_{q-1,p}$. After that, we estimate c by

$$c \simeq 2^{(q-1)(p+1)}|\Theta_{q,p}(f) - \Theta_{q-1,p}(f)|.$$

Then, we obtain the expression

$$|\theta - \Theta_{q,p}(f)| \leq \frac{\nu}{2^{p+1}} |\Theta_{q,p}(f) - \Theta_{q-1,p}(f)|, \tag{II.13}$$

where ν is a "safety parameter" whose role is to prevent the oscillations in the error as a function of q due to the quasi-periodic part. In every numerical computation we take $\nu = 10$. For more details on the behavior of the error we refer to [SV06].

Now, we comment two sources of error to take into account in the implementation of the method:

- The sums $S^p_{N_j}(f)$ are evaluated using the lift rather than the map itself. Of course, this makes the sums $S^p_{N_j}(f)$ to increase (actually they are of order $\mathcal{O}(N^{p+1})$) and is recommended to store their integer and decimal parts separately in order to keep the desired precision.

- If the required number of iterates increases, we have to be aware of the round-off errors in the evaluation of the iterates. For this reason, when implementing the above scheme in a computer, we multiple-precision arithmetics. The computations presented in this chapter have been performed using a C++ compiler and the multiple arithmetic has been provided by the routines *double-double and quad-double package* of [HLB05], which include a *double-double* data type of approximately 32 decimal digits and a *quadruple-double* data type of approximately 64 digits.

Along this section we have required the rotation number to be Diophantine. Of course, if $\theta \in \mathbb{Q}$ equation (II.7) is not valid since, in general, the dynamics of f is not conjugate to a rigid rotation. Anyway, we can compute the sums $S^p_N(f)$ and it turns out that the method works as well as for Diophantine numbers. We can justify this behavior from the known fact that, for any circle homeomorphism of rational rotation number, every orbit is either periodic or its iterates converge to a periodic orbit (see [dMvS93, KH95]). Then, the iterates of the map tend toward periodic points, and for such points, one can see that the averaged sums $\widetilde{S}^p_N(f)$ also satisfy an expression like (II.8) with an error of the same order, and this is all we need to perform the extrapolation. In fact, the worst situation appears when computing irrational rotation numbers that are "close" to rational ones (see also the discussion in Section II.4.1).

II.3 Derivatives of the rotation number with respect to parameters

Now we adapt the method already described in Section II.2 in order to compute derivatives of the rotation number with respect to parameters (assuming that they exist). For the sake of simplicity, we introduce the method for one-parameter families of circle diffeomorphisms, albeit the construction can be adapted to deal with multiple parameters (we discuss this situation in Remark II.3.4). Thus, consider $\mu \in I \subset \mathbb{R} \mapsto f_\mu \in \mathrm{Diff}_+^\omega(\mathbb{T})$ depending on μ in a regular way. The rotation numbers of the family $\{f_\mu\}_{\mu\in I}$ induce a function $\theta : I \to [0,1)$ given by $\theta(\mu) = \rho(f_\mu)$. Then, our goal is to approximate numerically the derivatives of θ at a given point μ_0.

Let us remark that the function θ is only continuous in the $\mathcal{C}^0$-topology and, actually, the rotation number depends on μ in a very non-smooth way: generically, there exist a family of disjoint open intervals of I, with dense union, such that θ takes constant rational values on these intervals (a so-called Devil's Staircase). However, $\theta'(\mu)$ is defined for any μ such that $\theta(\mu) \notin \mathbb{Q}$ (see [Her77]).

Higher order derivatives are defined in "many" points in the sense of Whitney. Concretely, let $J \subset I$ be the subset of parameters such that $\theta(\mu) \in \mathcal{D}$ (typically a Cantor set). Then, from Theorem II.2.3, there exists a family of conjugacies $\mu \in J \mapsto \eta_\mu \in \mathrm{Diff}_+^\omega(\mathbb{T})$, satisfying $f_\mu \circ \eta_\mu = \eta_\mu \circ R_{\theta(\mu)}$, that is unique if we fix $\eta_\mu(0) = x_0$. Then, if f_μ is $\mathcal{C}^d$ with respect to μ, the Whitney derivatives $D_\mu^j \eta_\mu$ and $D_\mu^j \theta$, for $j = 1, \ldots, d$, can be computed by taking formal derivatives with respect to μ on the conjugacy equation and solving small divisors equations thus obtained. Actually, we know that, if we define $J(C,\tau)$ as the subset of J such that $\theta(\mu) \in \mathcal{D}(C,\tau)$, for certain $C > 0$ and $\tau \geq 1$, then the maps $\mu \in J(C,\tau) \mapsto \eta_\mu$ and $\mu \in J(C,\tau) \mapsto \theta$ can be extended to $\mathcal{C}^s$ functions on I, where s depends on d and τ, provided that d is big enough (see [Van02]).

As it is shown in Section II.3.1, when we extend the method for computing the dth derivative of θ, in general, we are forced to select an averaging order $p > d$ and the remainder turns out to be of order $\mathcal{O}(1/N^{p-d+1})$. Nevertheless, if the rotation number is known to be constant as a function of the parameters, we can avoid previous limitations. Concretely, in this case we can select any averaging order p, independent of d, since the remainder is now of order $\mathcal{O}(1/N^{p+1})$. Of course, if the rotation number is constant, then the derivatives of θ are all zero and the fact that we can obtain them with better precision seems to be irrelevant. However, from the computation of these vanishing derivatives, we can obtain information about other involved objects. This is the case of many applications in which this methodology turns out to be very useful (two examples are worked out in Sections II.4.3 and II.5.3).

II.3.1 Computation of the first derivative

We start by explaining how to compute the first derivative of θ. For notational convenience, from now on we fix μ_0 such that $\theta(\mu_0) \in \mathcal{D}$ and we omit the dependence on μ as a subscript in families of circle maps. In addition, let us recall that we can write any conjugation as $\eta(x) = x + \xi(x)$ and denote by $\{\hat{\xi}_k\}_{k\in\mathbb{Z}}$ the Fourier coefficients of ξ. Finally, we denote the first derivatives as $\theta' = D_\mu \theta$ and $\hat{\xi}_k' = D_\mu \hat{\xi}_k$.

As we did in Section II.2.2, we begin by computing the first averages (of the derivatives of the iterates) in order to illustrate the idea of the method. Thus, we proceed by formally taking derivatives with respect to μ at both sides of equation (II.3)

$$D_\mu f^n(x_0) = n\theta' + \sum_{k\in\mathbb{Z}} \hat{\xi}_k' \mathrm{e}^{2\pi \mathrm{i} k n\theta} + 2\pi \mathrm{i} n\theta' \sum_{k\in\mathbb{Z}} k\hat{\xi}_k \mathrm{e}^{2\pi \mathrm{i} k n\theta}, \qquad \forall n \in \mathbb{Z}.$$

Then, notice that a factor n appears multiplying the second quasi-periodic sum. However, if we perform recursive sums, we can still manage to control the growth of this term due to the quasi-periodic part. Let us compute the sum

$$\begin{aligned} D_\mu S_N^1(f) &:= \sum_{n=1}^{N} D_\mu(f^n(x_0) - x_0) \\ &= \frac{N(N+1)}{2}\theta' - N\sum_{k\in\mathbb{Z}_*} \hat{\xi}_k' + \sum_{k\in\mathbb{Z}_*} \hat{\xi}_k' \frac{\mathrm{e}^{2\pi \mathrm{i} k\theta}(1-\mathrm{e}^{2\pi \mathrm{i} kN\theta})}{1-\mathrm{e}^{2\pi \mathrm{i} k\theta}} \\ &\quad +2\pi \mathrm{i}\theta' \sum_{k\in\mathbb{Z}_*} k\hat{\xi}_k \frac{N\mathrm{e}^{2\pi \mathrm{i} k(N+2)\theta} - (N+1)\mathrm{e}^{2\pi \mathrm{i} k(N+1)\theta} + \mathrm{e}^{2\pi \mathrm{i} k\theta}}{(1-\mathrm{e}^{2\pi \mathrm{i} k\theta})^2}. \end{aligned}$$

Hence, we observe that the method is still valid, even though for $\theta' \neq 0$ the quasi-periodic sum is bigger than expected a priori. Indeed, we obtain the following formula

$$\frac{2}{N(N+1)} D_\mu S_N^1(f) = \theta' + \mathcal{O}(1/N), \tag{II.14}$$

that is similar to equation (II.4), but notice that the term $2A_1/(N+1)$ has been included in the remainder since there are oscillatory terms of the same order. Proceeding as in Section II.2.2, we introduce *recursive sums* for the derivatives of the iterates

$$D_\mu S_N^p(f) := D_\mu(f^N(x_0) - x_0), \qquad D_\mu S_N^p(f) := \sum_{j=1}^{N} D_\mu S_j^{p-1}(f),$$

and their corresponding *averaged sums of order* p

$$D_\mu \widetilde{S}_N^p(f) := \binom{N+p}{p+1}^{-1} D_\mu S_N^p(f).$$

Finding an expression like (II.14) for $p > 1$ is quite cumbersome to do directly, since the computations are very involved. However, the computation is straightforward if we take formal derivatives at both sides of equation (II.7). The resulting expression reads as

$$D_\mu \widetilde{S}_N^p(f) = \theta' + \sum_{l=1}^{p} \frac{D_\mu A_l^p}{(N+p-l+1)\cdots(N+p)} + D_\mu E^p(N),$$

where the new coefficients are $D_\mu A_l^p = (-1)^l (p-l+2)\cdots(p+1) D_\mu A_l$ with

$$D_\mu A_l = \sum_{k\in\mathbb{Z}_*} \frac{\mathrm{e}^{2\pi \mathrm{i} k(l-1)\theta}}{(1-\mathrm{e}^{2\pi\mathrm{i}k\theta})^{l-1}} \left(\hat{\xi}_k' + \frac{2\pi\mathrm{i}k(l-1)\hat{\xi}_k\theta'}{1-\mathrm{e}^{2\pi\mathrm{i}k\theta}} \right),$$

and the new remainder is

$$\begin{aligned} D_\mu E^p(N) &= (-1)^{p+1} \frac{(p+1)!}{N\cdots(N+p)} \sum_{k\in\mathbb{Z}_*} \frac{\mathrm{e}^{2\pi\mathrm{i}kp\theta}}{(1-\mathrm{e}^{2\pi\mathrm{i}k\theta})^p} \left\{ \hat{\xi}_k'(1-\mathrm{e}^{2\pi\mathrm{i}kN\theta}) \right. \\ & \left. +2\pi\mathrm{i}k\hat{\xi}_k\theta' \left(p\frac{1-\mathrm{e}^{2\pi\mathrm{i}kN\theta}}{1-\mathrm{e}^{2\pi\mathrm{i}k\theta}} - N\mathrm{e}^{2\pi\mathrm{i}kpN\theta} \right) \right\}. \end{aligned}$$

Assuming that $\theta(\mu_0) \in \mathcal{D}$ and $\theta'(\mu_0) \neq 0$, we can obtain analogous bounds as those of Lemma II.2.4 and conclude that the remainder satisfies $D_\mu E^p(N) = \mathcal{O}(1/N^p)$. Moreover, we observe that the coefficient $D_\mu A_p^p$ corresponds to a term of the same order, so we have to redefine the remainder in order to include this term. Hence, as we did in equation (II.8), we can arrange the unknown terms and obtain

$$D_\mu \widetilde{S}_N^p(f) = \theta' + \sum_{l=1}^{p-1} \frac{D_\mu \hat{A}_l^p}{N^l} + \mathcal{O}(1/N^p),$$

where the coefficients $\{D_\mu \hat{A}_l^p\}_{l=1,\ldots,p-1}$ are the derivatives of $\{\hat{A}_l^p\}_{l=1,\ldots,p-1}$ that appear in equation (II.8).

Finally, we can extrapolate an approximation to θ' using Richardson's method of order $p-1$ as in Section II.2.2. Concretely, if we compute $N = 2^q$ iterates, we can approximate the derivative of the rotation number by means of the following formula

$$\theta' = \sum_{j=0}^{p-1} c_j^{p-1} D_\mu \widetilde{S}_{2^{q-p+1+j}}^p(f) + \mathcal{O}(2^{-pq}), \tag{II.15}$$

where the coefficients $\{c_j^{p-1}\}_{j=0,\ldots,p-1}$ are given by (II.11).

II.3.2 Computation of higher order derivatives

The goal of this section is to generalize formula (II.15) to approximate $D_\mu^d\theta$ for any d, when they exists. Then, we assume that the family $\mu \mapsto f \in \mathrm{Diff}_+^\omega(\mathbb{T})$ depends $\mathcal{C}^d$-smoothly with respect to the parameter. As usual, we define the recursive sums for the d-derivative and their averages of order p as

$$D_\mu^d S_N^0(f) := D_\mu^d(f^n(x_0)-x_0), \quad D_\mu^d S_N^p(f) := \sum_{j=0}^{N} D_\mu^d S_j^{p-1}(f),$$

and

$$D^d_\mu \widetilde{S}^p_N(f) := \binom{N+p}{p+1}^{-1} D^d_\mu S^p_N(f),$$

respectively. As before, we relate these sums to $D^d_\mu\theta$ by taking formal derivatives in equation (II.7), thus obtaining

$$D^d_\mu \widetilde{S}^p_N(f) = D^d_\mu\theta + \sum_{l=1}^{p} \frac{D^d_\mu A^p_l}{(N+p-l+1)\cdots(N+p)} + D^d_\mu E^p(N). \qquad \text{(II.16)}$$

Is immediate to check that, if $\theta(\mu_0) \in \mathcal{D}$ and $D^d_\mu\theta(\mu_0) \neq 0$, the remainder $D^d_\mu E^p(N)$ is of order $\mathcal{O}(1/N^{p-d+1})$, so this expression makes sense if the averaging order satisfies $p > d$.

Remark II.3.1. *Notice that, in order to work with reasonable computational time and round-off errors, p cannot be taken arbitrarily big. Consequently, there is a (practical) limitation in the computation of high order derivatives.*

In addition, as it was done for the first derivative, the remainder $D^d_\mu E^p(N)$ must be redefined in order to include the terms corresponding to $l \geq p-d+1$ in equation (II.16). Then we can extrapolate $D^d_\mu\theta$ by computing $N = 2^q$ iterates and solving the linear $(p-d+1)$-dimensional system associated to the following rearranged equation

$$D^d_\mu \widetilde{S}^p_N(f) = D^d_\mu\theta + \sum_{l=1}^{p-d} \frac{D^d_\mu \hat{A}^p_l}{N^l} + \mathcal{O}(1/N^{p-d+1}). \qquad \text{(II.17)}$$

Since the averaging order p and the extrapolation order $p-d$ do not coincide, let us define the *extrapolation operator of order m for the d-derivative* as

$$\Theta^d_{q,p,m}(f) := \sum_{j=0}^{m} c^m_j D^d_\mu \widetilde{S}^p_{2^{q-m+j}}(f), \qquad \text{(II.18)}$$

where the coefficients $\{c^m_j\}_{j=0,\ldots,m}$ are given by (II.11). Therefore, according to formula (II.17), we can approximate the dth derivative of the rotation number as

$$D^d_\mu\theta = \Theta^d_{q,p,p-d}(f) + \mathcal{O}(2^{-(p-d+1)q}).$$

As before, it is convenient to summarize the previous computations in the following algorithm.

Algorithm II.3.2. *Once an averaging order p is selected, we take $N = 2^q$ iterates of the map, for some $q > p$, and compute the sums $\{D^d_\mu \widetilde{S}^p_{N_j}\}_{j=0,\ldots,p-d}$ with $N_j = 2^{q-p+j+d}$. We approximate the dth derivative of the rotation number using the formula*

$$D^d_\mu\theta = \Theta^d_{q,p,p-d} + \mathcal{O}(2^{-(p-d+1)q}), \quad \Theta^d_{q,p,m} = \sum_{j=0}^{m} c^{(m)}_j D^d_\mu \widetilde{S}^p_{2^{q-m+j}},$$

where the coefficients $c^{(m)}_j$ are given by equation (II.11). *The operator $\Theta^d_{q,p,p-d}$ corresponds to the Richardson extrapolation of order $p-d$ of equation* (II.17).

Furthermore, as explained in Section II.2.2, by comparing the approximations that correspond to 2^{q-1} and 2^q iterates, we obtain the following heuristic formula for the extrapolation error:

$$|D_\mu^d \theta - \Theta_{q,p,p-d}^d(f)| \leq \frac{\nu}{2^{p-d+1}} |\Theta_{q,p,p-d}^d(f) - \Theta_{q-1,p,p-d}^d(f)|, \tag{II.19}$$

where, once again, ν is a "safety parameter" that we take as $\nu = 10$.

Remark II.3.3. *Up to this point we have assumed that $D_\mu^d \theta \neq 0$ at the computed point. However, if we know a priori that $D_\mu^r \theta = 0$ for $r = 1, \ldots, d$, then equation (II.16) holds with the following expression for the remainder:*

$$D_\mu^d E^p(N) = (-1)^{p+1} \frac{(p+1)!}{N \cdots (N+p)} \sum_{k \in \mathbb{Z}_*} D_\mu^d \hat{\xi}_k \frac{e^{2\pi i k p \theta}(1 - e^{2\pi i k N \theta})}{(1 - e^{2\pi i k \theta})^p},$$

which now is of order $\mathcal{O}(1/N^{p+1})$. As in Section II.2, this allows us to approximate $D_\mu^d \theta$ with the same extrapolation order as the averaging order p. Indeed, we obtain

$$0 = D_\mu^d \theta = \Theta_{q,p,p}^d(f) + \mathcal{O}(2^{-(p+1)q}),$$

and we observe that the order d is not limited by p anymore.

The case remarked above is very interesting since we know that many applications can be modeled as a family of circle diffeomorphisms of fixed rotation number. The possibilities of this approach are illustrated by computing the Taylor expansion of Arnold Tongues (Section II.4.3) and the continuation of invariant curves for the Hénon map (Section II.5.3).

II.3.3 Scheme for evaluating the derivatives of the averaged sums

Let us introduce a recursive way for computing the sums $D_\mu^d \widetilde{S}_N^p(f)$ required to evaluate the extrapolation operator (II.18). First of all, notice that by linearity it suffices to compute $D_\mu^d(f^n(x_0))$ for any $n \in \mathbb{N}$.

To compute the derivatives of $f^n = f \circ \overset{(n)}{\cdots} \circ f$, we proceed inductively with respect to n and d. Thus, let us assume that the derivatives $D_\mu^r(f^{n-1}(x_0))$ are known for a given $n \geq 1$ and for any $r \leq d$. Then, if we denote $z := f^{n-1}(x_0)$, our goal is to compute $D_\mu^r(f(z))$ for $r \leq d$ by using the known derivatives of z.

For $d = 1$, a recursive formula appears directly by applying the chain rule

$$D_\mu(f(z)) = \partial_\mu f(z) + \partial_x f(z) D_\mu(z). \tag{II.20}$$

This formula can be implemented provided the partial derivatives $\partial_\mu f$ and $\partial_x f$ can be numerically evaluated at the point z.

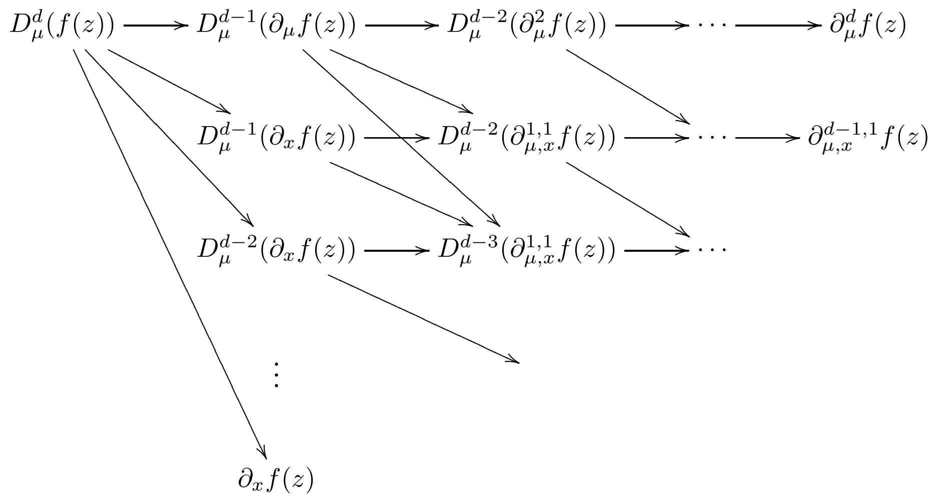

Figure II.1: Schematic representation of the recurrent computations performed to evaluate $D_\mu^d(f(z))$.

In general, we can perform higher order derivatives and obtain the following expression

$$\begin{aligned} D_\mu^d(f(z)) &= D_\mu^{d-1}\Big(\partial_\mu f(z) + \partial_x f(z) D_\mu(z)\Big) \\ &= D_\mu^{d-1}(\partial_\mu f(z)) + \sum_{r=0}^{d-1} \binom{d-1}{r} D_\mu^r(\partial_x f(z)) D_\mu^{d-r}(z). \end{aligned}$$

This motivates the extension of recurrence (II.20), since for evaluating the previous formula we require to know the derivatives $D_\mu^r(\partial_x f(z))$ for $r < d$ and $D_\mu^{d-1}(\partial_\mu f(z))$. We note that these derivatives can also be computed recursively using similar expressions for the maps $\partial_x f$ and $\partial_\mu f$, respectively. Concretely, assuming that we can evaluate $\partial_{\mu,x}^{i,j} f(z)$ for any $(i,j) \in \mathbb{Z}_+^2$ such that $i + j \leq d$, we can use the following recurrences

$$D_\mu^r(\partial_{\mu,x}^{i,j} f(z)) = D_\mu^{r-1}(\partial_{\mu,x}^{i+1,j} f(z)) + \sum_{s=0}^{r-1} \binom{r-1}{s} D_\mu^s(\partial_{\mu,x}^{i,j+1} f(z)) D_\mu^{r-s}(z),$$

to compute in a tree-like order the corresponding derivatives. To prevent redundant computations in the implementation of the method, we store (in memory) the value of the "intermediate" derivatives $D_\mu^r(\partial_{\mu,x}^{i,j} f(z))$ so they only have to be computed once. For this reason, this scheme turns out to be more efficient than evaluating explicit expressions such as Faà di Bruno formulas (see for example [KP02]). Figure II.1 summarizes the recursive computations required and the convenience of storing these intermediate computations.

Remark II.3.4. *The above scheme can be generalized immediately to the case of several parameters. For example, consider a two-parameter family* $(\mu_1, \mu_2) \mapsto f_{\mu_1,\mu_2} \in \mathrm{Diff}_+^\omega(\mathbb{T})$ *whose* *rotation number induces a map* $(\mu_1, \mu_2) \mapsto \theta(\mu_1, \mu_2)$. *Then, if* $\theta(\mu_1^0, \mu_2^0) \in \mathcal{D}$, *we can obtain a similar scheme to approximate* $D^{d_1,d_2}_{\mu_1,\mu_2}\theta(\mu_1^0, \mu_2^0)$. *In this context, note that the operator* $\Theta^{d_1,d_2}_{q,p,p-d_1-d_2}$ *can be defined as (II.18), but averaging the derivatives* $D^{d_1,d_2}_{\mu_1,\mu_2}(f^n(x_0))$. *Finally, if we write* $z := f^{n-1}(x_0)$, *we can compute inductively the derivatives* $D^{m,l}_{\mu_1,\mu_2}(f(z))$, *for* $m \leq d_1$ *and* $l \leq d_2$, *using the following recurrences*

$$\begin{aligned} D^{m,l}_{\mu_1,\mu_2}(\partial^{i,j,k}_{\mu_1,\mu_2,x} f(z)) \;=\;& D^{m-1,l}_{\mu_1,\mu_2}(\partial^{i+1,j,k}_{\mu_1,\mu_2,x} f(z)) \\ &+ \sum_{s=0}^{m-1}\sum_{r=0}^{l} \binom{m-1}{s}\binom{l}{r} D^{s,r}_{\mu_1,\mu_2}(\partial^{i,j,k+1}_{\mu_1,\mu_2,x} f(z)) D^{m-s,l-r}_{\mu_1,\mu_2}(z), \end{aligned}$$

if $m \neq 0$ *and*

$$D^{0,l}_{\mu_1,\mu_2}(\partial^{i,j,k}_{\mu_1,\mu_2,x} f(z)) = D^{0,l-1}_{\mu_1,\mu_2}(\partial^{i,j+1,k}_{\mu_1,\mu_2,x} f(z)) + \sum_{r=0}^{l-1}\binom{l-1}{r} D^{0,r}_{\mu_1,\mu_2}(\partial^{i,j,k+1}_{\mu_1,\mu_2,x} f(z)) D^{0,l-r}_{\mu_1,\mu_2}(z),$$

if $l \neq 0$. *Of course,* $D^{0,0}_{\mu_1,\mu_2}(\partial^{i,j,k}_{\mu_1,\mu_2,x} f(z)) = \partial^{i,j,k}_{\mu_1,\mu_2,x} f(z)$ *corresponds to the evaluation of the partial derivative of the map.*

II.4 Application to the Arnold family

As a first example, let us consider the Arnold family of circle maps, given by

$$\begin{aligned} f_{\alpha,\varepsilon} : \mathbb{S} &\longrightarrow \mathbb{S} \\ x &\longmapsto x + 2\pi\alpha + \varepsilon \sin(x), \end{aligned} \tag{II.21}$$

where $(\alpha, \varepsilon) \in [0,1) \times [0,1)$ are parameters and $\mathbb{S} = \mathbb{R}/(2\pi\mathbb{Z})$. Notice that this family satisfies $f_{\alpha,\varepsilon} \in \text{Diff}^{\omega}_{+}(\mathbb{S})$ for any value of the parameters. Let us remark that (II.21) allows us to illustrate the method in a direct way, since there are explicit formulas for the partial derivatives $\partial^{i,j,k}_{\alpha,\varepsilon,x} f(x)$ of the map, for any $(i,j,k) \in \mathbb{Z}^3_+$. In Section II.5 we will consider another interesting application in which the studied family is not given explicitly.

For this family of maps, it is convenient to take the angles modulo 2π just for avoiding the loss of significant digits due to the factors $(2\pi)^{d-1}$ that would appear in the d-derivative of the map.

The contents of this section are organized as follows. First, in Section II.4.1 we compute the derivative of a Devil's Staircase, that corresponds to the variation of the rotation number of (II.21) with respect to α for a fixed ε. In Section II.4.2 we use the computation of derivatives of the rotation number to approximate the Arnold Tongues of the family (II.21) by means of the Newton method. Furthermore, we compute the asymptotic expansion of these tongues and obtain pseudo-analytical expressions for the first coefficients, as a function of the rotation number.

II.4.1 Stepping up to a Devil's Staircase

Let us fix the value of $\varepsilon \in [0, 1)$ and consider the one-parameter family $\{f_\alpha\}_{\alpha\in[0,1)}$ given by equation (II.21), i.e. $f_\alpha := f_{\alpha,\varepsilon}$. Let us recall that we can establish an ordering in this family since the normalized lifts satisfy $f_{\alpha_1}(x) < f_{\alpha_2}(x)$ for all $x \in \mathbb{R}$ if and only if $\alpha_1 < \alpha_2$. Then, we conclude that the function $\alpha \mapsto \rho(f_\alpha)$ is monotone increasing. In particular, for $\alpha_1 < \alpha_2$ such that $\rho(f_{\alpha_1}) \in \mathbb{R}\backslash\mathbb{Q}$ we have $\rho(f_{\alpha_1}) < \rho(f_{\alpha_2})$. On the other hand, if $\rho(f_{\alpha_1}) \in \mathbb{Q}$, there is an interval containing α_1 giving the same rotation number. As the values of α for which f_α has rational rotation number are dense in $[0, 1)$ (the complement is a Cantor set), there are infinitely many intervals where $\rho(f_\alpha)$ is locally constant. Therefore, the map $\alpha \mapsto \rho(f_\alpha)$ gives rise to a "staircase" with a dense number of stairs, that is usually called a Devil's Staircase (we refer to [dMvS93, KH95] for more details).

To illustrate the behavior of the method we have computed the above staircase for $\varepsilon = 0.75$. The computations have been performed by taking 10^4 points of $\alpha \in [0, 1)$, using 32-digit arithmetics (*double-double* data type from [HLB05]), and a fixed averaging order $p = 8$. In addition, we estimate the error in the approximation of $\rho(f_\alpha)$ and $D_\alpha\rho(f_\alpha)$ using formulas (II.13) and (II.19), respectively. Then, we stop the computations for a tolerance of 10^{-26} and 10^{-24}, respectively, using at most $2^{22} = 4194304$ iterates.

Let us discuss the results obtained. First, we point out that only 11.4 % of the selected points have not reached the previous tolerances for 2^{22} iterates. Moreover, we observe that the rotation number for 98.8 % of the points has been obtained with an error less that 10^{-20}, while the estimated error in the derivatives is less than 10^{-18} for 97.7 % of the points. Let us focus in $\alpha = 0.3377$, that is one of the "bad" points. The estimated errors for the rotation number and the derivative at this point are of order 10^{-18} and 10^{-9}, respectively. We observe that, even though this rotation number is irrational (the derivative does not vanish), it is very close to the rational $105/317$, since $|317 \cdot \Theta_{22,9}(f_{0.3377}) - 105| \simeq 4.2 \cdot 10^{-6}$.

In Figure II.2 we show $\alpha \mapsto \rho(f_\alpha)$ and its derivative $\alpha \mapsto D_\alpha\rho(f_\alpha)$ for those points that satisfy that the estimated error is less than 10^{-18} and 10^{-16}, respectively. We recall that the rational values of the rotation number correspond to the constant intervals in the top-left plot, and note that by looking at the derivative (top-right plot) we can visualize the density of the stairs better than looking at the staircase itself. We remark that both these rational rotation numbers and their vanishing derivatives have been computed as well as in the Diophantine case.

Moreover, at the bottom of the same figure, we plot some magnifications of the derivative to illustrate the non-smoothness of a Devil's Staircase. Concretely, the plot in the bottom-left corresponds to 10^5 values of $\alpha \in [0.2, 0.3]$ using the same implementation parameters as before. Once again, if the estimated error is bigger than 10^{-16} the point is not plotted. Finally, on the right plot we give another magnification for 10^6 values of $\alpha \in [0.282, 0.292]$ that are computed with $p = 7$, and allowing at most $2^{21} = 2097152$ iterates. In this case, points that correspond to the branch in the left (i.e. close to $\alpha = 0.2825$), are typically computed with an error 10^{-10}.

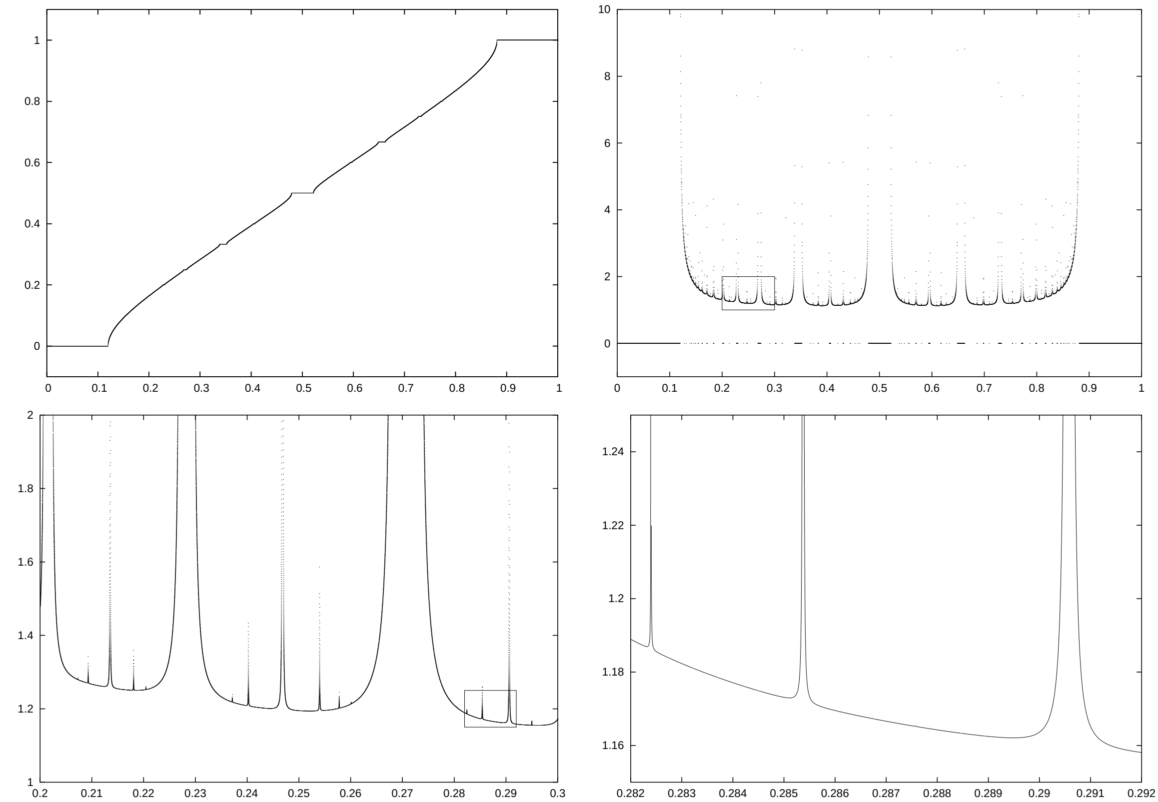

Figure II.2: Devil's Staircase $\alpha \mapsto \rho(f_\alpha)$ (top-left) and its derivative (top-right) for the Arnold family with $\varepsilon = 0.75$. The plots in the bottom correspond to some magnifications of the top-right one.

II.4.2 Newton method for computing the Arnold Tongues

Since $f_{\alpha,\varepsilon} \in \mathrm{Diff}_+^\omega(\mathbb{S})$, we obtain a function $(\alpha, \varepsilon) \mapsto \rho(\alpha, \varepsilon) := \rho(f_{\alpha,\varepsilon})$ given by the rotation number. Then, the Arnold Tongues of (II.21) are defined as the sets $T_\theta = \{(\alpha, \varepsilon) \,:\, \rho(\alpha, \varepsilon) = \theta\}$, for any $\theta \in [0, 1)$. It is well known that if $\theta \in \mathbb{Q}$, then T_θ is a set with interior; otherwise, T_θ is a continuous curve which is the graph of a function $\varepsilon \mapsto \alpha(\varepsilon)$, with $\alpha(0) = \theta$. In addition, if $\theta \in \mathcal{D}$, the corresponding tongue is given by an analytic curve (see [Ris99]).

Using the method described in Section II.2.2, some Arnold Tongues T_θ of Diophantine rotation number were approximated in [SV06] by means of the secant method. Now, since we can compute derivatives of the rotation number, we are able to repeat the computations using a Newton method. To do that, we fix $\theta \in \mathcal{D}$ and solve the equation $\rho(\alpha, \varepsilon) - \theta = 0$ by continuing the known solution $(\theta, 0)$ with respect to ε. Indeed, we fix a partition $\{\varepsilon_j\}_{j=0,\ldots,K}$ of $[0, 1)$, and compute a numerical approximation α_j^* for every $\alpha(\varepsilon_j)$.

To this end, assume that we have a good approximation α_{j-1}^* to $\alpha(\varepsilon_{j-1})$ and let us first compute an initial approximation for $\alpha(\varepsilon_j)$. Taking derivative in the equation $\rho(\alpha(\varepsilon), \varepsilon) - \theta = 0$

we obtain

$$D_\alpha\rho(\alpha(\varepsilon),\varepsilon)\alpha'(\varepsilon) + D_\varepsilon\rho(\alpha(\varepsilon),\varepsilon) = 0. \tag{II.22}$$

Thus, we can approximate $\alpha'(\varepsilon_{j-1})$ by computing numerically the derivatives $D_\alpha\rho$ and $D_\varepsilon\rho$ at $(\alpha^*_{j-1},\varepsilon_{j-1})$. Hence, we obtain an approximated value $\alpha_j^{(0)} = \alpha^*_{j-1} + \alpha'(\varepsilon_{j-1})(\varepsilon_j - \varepsilon_{j-1})$ for $\alpha(\varepsilon_j)$. Next, we apply the Newton method

$$\alpha_j^{(n+1)} = \alpha_j^{(n)} - \frac{\rho(\alpha_j^{(n)},\varepsilon_j) - \theta}{D_\alpha\rho(\alpha_j^{(n)},\varepsilon_j)},$$

and stop when we converge to a value α_j^* that approximates $\alpha(\varepsilon_j)$.

Computations are performed using 64 digits (*quadruple-double* data type from [HLB05]) and, in order to compare with the results obtained in [SV06], we select the same parameters in the implementation. In particular, we take a partition $\varepsilon_j = j/K$ with $K = 100$ of the interval $[0,1)$, we select an averaging order $p = 9$ and allow at most $2^{23} = 8388608$ iterates of the map. The required tolerances are taken as 10^{-32} for the computation of the rotation number (we use (II.13) to estimate the error) and 10^{-30} for the convergence of the Newton method. Let us remark that the computations are done without any prescribed tolerance for the computation of the derivatives $D_\alpha\rho$ and $D_\varepsilon\rho$, even though we check, using (II.19), that the extrapolation is done correctly.

Let us discuss the results obtained for $\theta = (\sqrt{5}-1)/2$. As expected, the number of iterates of the Newton method is less than the ones required by the secant method. Concretely, we perform from 2 to 3 corrections as we approach the critical value $\varepsilon = 1$, while using the secant method we need at least 4 steps to converge. However, we observe that the computation of the derivatives $D_\alpha\rho$ and $D_\varepsilon\rho$ fails if we take $\varepsilon = 1$, even though the secant method converges after 18 iterations. This is totally consistent since we know that $f_{\alpha,1} \in \mathrm{Diff}^0_+(\mathbb{T})$ but is still an analytic map, and that the conjugation to a rigid rotation is only Hölder continuous (see [dlLP02, Yoc84b]).

In Figure II.3 (left) we plot the derivatives $\varepsilon \mapsto D_\alpha\rho(\alpha(\varepsilon),\varepsilon)$ and $\varepsilon \mapsto D_\varepsilon\rho(\alpha(\varepsilon),\varepsilon)$ evaluated on the previous tongue. We observe that the derivatives have been normalized in order to fit together in the same plot. On the other hand, in the right plot we show the estimated error in the computation of these derivatives (obtained from equation (II.19)). In the worst case, $\varepsilon = 0.99$, we obtain errors of order 10^{-27} and 10^{-29} for $D_\alpha\rho$ and $D_\varepsilon\rho$, respectively.

II.4.3 Computation of the Taylor expansion of the Arnold Tongues

As we have mentioned in Section II.4.2, if $\theta \in \mathcal{D}$ then the Arnold Tongue T_θ of (II.21) is given by the graph of an analytic function $\alpha(\varepsilon)$, for $\varepsilon \in [0,1)$. Then, we can expand α at the origin as

$$\alpha(\varepsilon) = \theta + \frac{\alpha'(0)}{1!}\varepsilon + \frac{\alpha''(0)}{2!}\varepsilon^2 + \cdots + \frac{\alpha^{(d)}(0)}{d!}\varepsilon^d + \mathcal{O}(\varepsilon^{d+1}), \tag{II.23}$$

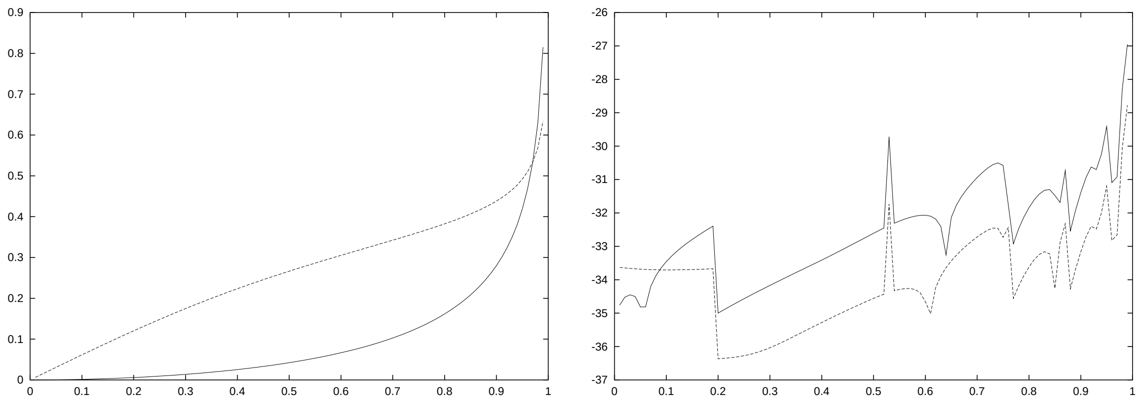

Figure II.3: Left: Graph of the derivatives $\varepsilon \mapsto D_\alpha \rho(\alpha(\varepsilon), \varepsilon)$ and $\varepsilon \mapsto D_\varepsilon \rho(\alpha(\varepsilon), \varepsilon)$ along T_θ, for the fixed rotation number $\theta = (\sqrt{5} - 1)/2$. The solid curve corresponds to $(D_\alpha \rho - 1)$ and the dashed one to $(20 \cdot D_\varepsilon \rho)$. Right: error (estimated using (II.13)) in $\log_{10}$ scale in the computation of these derivatives.

and the goal is now to numerically approximate the terms in this expansion. We know that every odd derivative in this expansion vanishes, so the Taylor expansion can be written in terms of powers in ε^2 (see [SV09] for details). However, we do not use this symmetry, but instead we verify the accuracy of our computations according to this information (see the results presented in Table II.1).

First of all, we want to emphasize that the direct extension of the computations performed in the previous section is hopeless. Concretely, as we did for approximating $\alpha'(\varepsilon)$, we could take higher order derivatives with respect to ε at equation (II.22) and, after evaluating at the point $(\alpha, \varepsilon) = (\theta, 0)$, isolate the derivatives $\alpha^{(r)}(0)$, $1 \leq r \leq d$. For example, once we know $\alpha'(0)$, the computation of $\alpha''(0)$ would follow from the expression

$$D_\alpha \rho(\alpha(\varepsilon), \varepsilon)\alpha''(\varepsilon) + \left(D^2_\alpha \rho(\alpha(\varepsilon), \varepsilon)\alpha'(\varepsilon) + 2D_{\alpha,\varepsilon}\rho(\alpha(\varepsilon), \varepsilon) \right)\alpha'(\varepsilon) + D^2_\varepsilon \rho(\alpha(\varepsilon), \varepsilon) = 0, \tag{II.24}$$

that requires to compute the second order partial derivatives of the rotation number (see Remark II.3.4). Then, by induction, we would obtain recurrent formulas to compute the expansion (II.23) up to order d. However, this approach is highly inefficient due to the following reasons:

- As discussed in Section II.3.2, using this approach we are limited to compute $\alpha^{(r)}(0)$ up to order $p-1$, where p is the selected averaging order. Of course, the precision for $\alpha^{(r)}(0)$ decreases dramatically when r increases to p.

- Note that, for the Arnold family, we can write explicitly $D_\alpha(f^n_{\alpha,\varepsilon}(x_0))|_{(\theta,0)} = n$. Then, if we look at the formulas in Remark II.3.4, we expect the terms $D^{m,l}_{\alpha,\varepsilon}(f^n_{\alpha,\varepsilon}(x_0))|_{(\theta,0)}$ to grow very fast, since they contain factors of the previous expression. Actually, we find that these quantities depend polynomially on n, with a power that increases with the order of the derivative. On the other hand, we expect the sums $D^{m,l}_{\alpha,\varepsilon}\widetilde{S}^p_N(f_{\alpha,\varepsilon})$ to converge, and therefore many cancellations are taking place in the computations. Consequently, when implementing this approach we unnecessarily lose a high amount of significant digits.

- Even if we could compute $D^{m,l}_{\alpha,\varepsilon}\rho(\theta,0)$ up to any order, it turns out that generalization of equation (II.24) for computing $\alpha^{(r)}(0)$ is badly conditioned. Concretely, derivatives of the rotation number increase with the order, giving rise to a large propagation of errors. Actually, the round-off errors increase so fast that, in practice, we cannot go beyond order 5 in the computation of (II.23) with the above methodology.

Therefore, we have to approach the problem in a different way. Concretely, our idea is to use the fact that the rotation number is constant on the tongue combined with Remark II.3.3. To this end, we consider the one-parameter family $\{f_{\alpha(\varepsilon),\varepsilon}\}_{\varepsilon\in[0,1)}$ of circle diffeomorphisms, where the graph of α parameterizes the tongue T_θ. For this family, we have $\rho(f_{\alpha(\varepsilon),\varepsilon}) = \theta$ for any $\varepsilon \in [0,1)$, and hence, from Remark II.3.3 we read the expression

$$0 = \Theta^d_{q,p,p}(f_{\alpha(\varepsilon),\varepsilon}) + \mathcal{O}(2^{-(p+1)q}), \tag{II.25}$$

where p is the averaging order, we use 2^q iterates and $\Theta^d_{q,p,p}$ is the extrapolation operator (II.18) that, in this case, depends on the derivatives of $\alpha(\varepsilon)$ up to order d. With this idea in mind, the aim of the next paragraphs is to show how we can isolate inductively these derivatives at $\varepsilon = 0$ from the previous equation.

Let us start by describing how to approximate the first derivative $\alpha'(0)$. As mentioned above, we have to write $\Theta^1_{q,p,p}(f_{\alpha(\varepsilon),\varepsilon})|_{\varepsilon=0}$ in terms of $\alpha'(0)$ and we note that, by linearity, it suffices to work with the expression $D_\varepsilon(f^n_{\alpha(\varepsilon),\varepsilon}(x_0))|_{\varepsilon=0}$. To do that, we write

$$f(x) = 2\pi\alpha(\varepsilon) + g(x), \qquad g(x) = x + \varepsilon\sin(x),$$

in order to uncouple the dependence on α in the circle map. Observe that, as usual, we omit dependence on the parameter in the maps. Using this notation, we have:

$$\begin{aligned} D_\varepsilon(f(x_0)) &= 2\pi\alpha'(\varepsilon) + \partial_\varepsilon g(x_0), \\ D_\varepsilon(f^2(x_0)) &= 2\pi\alpha'(\varepsilon) + \partial_\varepsilon g(f(x_0)) + \partial_x g(f(x_0)) D_\varepsilon(f(x_0)) \\ &= 2\pi\alpha'(\varepsilon)\left\{1 + \partial_x g(f(x_0))\right\} + \partial_\varepsilon g(f(x_0)) + \partial_x g(f(x_0))\partial_\varepsilon g(x_0). \end{aligned}$$

Similarly, we can proceed inductively and split the derivative of the nth iterate, $D_\varepsilon(f^n(x_0))$, in two parts, one of them having a factor $2\pi\alpha'(\varepsilon)$. Moreover, if we set $\varepsilon = 0$ in $D_\varepsilon(f^n(x_0))$, then it is clear that, with the exception of the previous factor, the resulting expression does not depend on $\alpha'(0)$ but only on $\alpha(0) = \theta$.

Now, we generalize the above argument to higher order derivatives. Let us assume that the values $\alpha'(0), \ldots, \alpha^{(d-1)}(0)$ are known, and isolate the derivative $\alpha^{(d)}(0)$ from $D^d_\varepsilon(f^n(x_0))|_{\varepsilon=0}$. We claim that the following formula holds

$$D^d_\varepsilon(f^n(x_0))|_{\varepsilon=0} = 2\pi n \alpha^{(d)}(0) + g^d_n, \tag{II.26}$$

where the factor $2\pi n$ comes from the fact that $\partial_x g|_{\varepsilon=0} = 1$, and $g^d := \{g^d_n\}_{n\in\mathbb{N}}$ is a sequence that only requires the known derivatives $\alpha^{(r)}(0)$, for $r < d$. Concretely, let us obtain the term g^d_n of the sequence by induction with respect to n. Once again, it is straightforward to write

$$\begin{aligned} D^d_\varepsilon(f^n(x_0)) &= D^{d-1}_\varepsilon\Big(2\pi\alpha'(\varepsilon) + \partial_\varepsilon g(f^{n-1}(x_0)) + \partial_x g(f^{n-1}(x_0)) D_\varepsilon(f^{n-1}(x_0))\Big) \\ &= 2\pi\alpha^{(d)}(\varepsilon) + D^{d-1}_\varepsilon(\partial_\varepsilon g(f^{n-1}(x_0)) \\ &\quad + \sum_{r=0}^{d-1} \binom{d-1}{r} D^r_\varepsilon(\partial_x g(f^{n-1}(x_0))) D^{d-r}_\varepsilon(f^{n-1}(x_0)). \end{aligned}$$

We note that the term $r = 0$ in this expression contains $D^d_\varepsilon(f^{n-1}(x_0))$. Then, if we set $\varepsilon = 0$ and replace inductively the previous term by equation (II.26), we find that

$$\begin{aligned} g^d_n &= D^{d-1}_\varepsilon(\partial_\varepsilon g(f^{n-1}(x_0))|_{\varepsilon=0} \\ &\quad + \sum_{r=1}^{d-1} \binom{d-1}{r} D^r_\varepsilon(\partial_x g(f^{n-1}(x_0))) D^{d-r}_\varepsilon(f^{n-1}(x_0))|_{\varepsilon=0} + g^d_{n-1} \end{aligned}$$

and let us remark that, as mentioned, this expression is independent of $\alpha^{(d)}(0)$.

We conclude the explanation of the method by describing the extrapolation process that allows us to approximate these derivatives. To this end, we introduce an extrapolation operator as (II.10) for the sequence g^d. Indeed, we extend the recursive sums (II.5) and the averaged sums (II.6) for this sequence, thus obtaining

$$\Theta_{q,p}(g^d) := \sum_{j=0}^{p} c^p_j \widetilde{S}^p_{2^{q-p+j}}(g^d).$$

Recalling that $D^d_\varepsilon \theta$ vanishes, we obtain from equation (II.25) that

$$\Theta^d_{q,p,p}(f)|_{\varepsilon=0} = 2\pi\alpha^{(d)}(0) + \Theta_{q,p}(g^d) = \mathcal{O}(2^{-(p+1)q}).$$

Therefore, the Taylor expansion (II.23) follows from sequential computation of $\alpha^{(d)}(0)$ by means of the expression

$$\alpha^{(d)}(0) = -\frac{1}{2\pi}\Theta_{q,p}(g^d) + \mathcal{O}(2^{-(p+1)q}).$$

d	$2\pi\alpha^{(d)}(0)$	e_1	e_2
0	3.883222077450933154693731259925391915269339787692096599014776434	-	-
1	5.289596087298835974306750728481413682115174017433159533705768026 $\cdot 10^{-54}$	$2 \cdot 10^{-50}$	$5 \cdot 10^{-54}$
2	-1.944003667801032197325141712953470682792841985057545477738933600 $\cdot 10^{-1}$	$7 \cdot 10^{-50}$	$2 \cdot 10^{-53}$
3	6.353866339253870417285870622952031667026712174414003758743809499 $\cdot 10^{-52}$	$3 \cdot 10^{-48}$	$6 \cdot 10^{-52}$
4	9.865443989835495993231949890783720243438883460505483297079900562 $\cdot 10^{-1}$	$2 \cdot 10^{-47}$	$5 \cdot 10^{-51}$
5	4.733853534850495777271526084574485398105534790325269345544052633 $\cdot 10^{-49}$	$2 \cdot 10^{-45}$	$5 \cdot 10^{-49}$
6	-1.451874181864020963416053802229271731186248529989217665545212404 $\cdot 10^{1}$	$6 \cdot 10^{-45}$	$1 \cdot 10^{-48}$
7	-1.986768674642925514096249083525472601734104441662711304098209993 $\cdot 10^{-47}$	$7 \cdot 10^{-44}$	$2 \cdot 10^{-47}$
8	1.673363822376717001078781931538386967523434046199355922539083323 $\cdot 10^{1}$	$8 \cdot 10^{-42}$	$2 \cdot 10^{-45}$
9	-5.559060362825539878039137008326038842079877436013501651866007318 $\cdot 10^{-44}$	$2 \cdot 10^{-40}$	$6 \cdot 10^{-44}$
10	1.974679484744669888248485084754876332689468886829840384314732615 $\cdot 10^{4}$	$2 \cdot 10^{-39}$	$4 \cdot 10^{-43}$
11	4.019718902900154426125206309959051888079502318143227318836414835 $\cdot 10^{-42}$	$1 \cdot 10^{-38}$	$4 \cdot 10^{-42}$
12	3.594891944526889578314748272295019294147597687816868847742850594 $\cdot 10^{5}$	$6 \cdot 10^{-37}$	-
13	-4.123166034989923032518732576715313341946051550138603536248010821 $\cdot 10^{-39}$	$2 \cdot 10^{-35}$	$4 \cdot 10^{-39}$
14	2.198602821435568153883567054383394767567371744732559263055644337 $\cdot 10^{6}$	$3 \cdot 10^{-33}$	-
15	1.307318024754974551233761145122558811543944190022138837513637182 $\cdot 10^{-35}$	$6 \cdot 10^{-32}$	$1 \cdot 10^{-35}$
16	-4.009257214040427899940043656551946700300230713255210114705187412 $\cdot 10^{10}$	$4 \cdot 10^{-31}$	-
17	-6.641638995605492204184114438636683272452899190211080822408603857 $\cdot 10^{-33}$	$4 \cdot 10^{-29}$	$7 \cdot 10^{-33}$
18	-2.582559893723659427522610275977697024396910000154382754643273110 $\cdot 10^{12}$	$1 \cdot 10^{-27}$	-
19	-4.366235264281358239242428788236090577328510850575386329987344515 $\cdot 10^{-30}$	$2 \cdot 10^{-26}$	$4 \cdot 10^{-30}$

Table II.1: Derivatives of $2\pi\alpha(\varepsilon)$ at the origin for $\theta = (\sqrt{5}-1)/2$. The column e_1 corresponds to the estimated error using (II.13). The column e_2 is the real error, that for even derivatives is computed comparing with the analytic expressions (II.27) and (II.28) using the coefficients from table II.2.

Let us discuss some obtained results. The following computations are performed using 64 digits (*quadruple-double* data type from [HLB05]). Implementation parameters are selected as $p = 11$, $q = 23$ and any tolerance is required in the extrapolation error (which is estimated by means of (II.13)).

In Table II.1 we show the computations of $2\pi\alpha^{(d)}(0)$, for $0 \leq d \leq 19$, that correspond to the Arnold Tongue associated to $\theta = (\sqrt{5}-1)/2$.

In addition, we use the above computations to obtain formulas, depending on θ, for the first coefficients of (II.23). To make this dependence explicit, we introduce the notation $\alpha_r(\theta) := \alpha^{(2r)}(0)$, where $(\varepsilon, \alpha(\varepsilon))$ parameterizes the Arnold Tongue T_θ. Analytic expressions for these coefficients can be found, for example, by solving the conjugation equation of diagram (II.2) using Lindstedt series. However, the complexity of the symbolic manipulations required for carrying the above computations is very large. In particular, the first two coefficients, whose computation is detailed in [SV09], are

$$\alpha_1(\theta) = \frac{\cos(\pi\theta)}{2^2\pi\sin(\pi\theta)}, \qquad \alpha_2(\theta) = -\frac{3\cos(4\pi\theta)+9}{2^5\pi\big(\sin(\pi\theta)\big)^2\sin(2\pi\theta)}. \tag{II.27}$$

From these formulas and a heuristic analysis of the small divisors equations to be solved for computing the remaining coefficients, we make the following guess for $\alpha_r(\theta)$:

$$\alpha_r(\theta) = \frac{P_r(\theta)}{2^{c(r)}\pi\big(\sin(\pi\theta)\big)^{2^{r-1}}\big(\sin(2\pi\theta)\big)^{2^{r-2}}\cdots\big(\sin((r-1)\pi\theta)\big)^2\sin(r\pi\theta)}, \tag{II.28}$$

P_3		P_4			
j	a_j	j	a_j	j	a_j
1	-105	0	-360150	14	-177625
3	825	2	40950	16	-14770
5	-465	4	469630	18	34755
7	-315	6	91140	20	49735
9	120	8	-378700	22	-53235
11	-60	10	67165	24	18900
		12	215355	26	-3150

P_5					
j	a_j	j	a_j	j	a_j
1	33992959770	21	46136915685	41	6059661930
3	-96457394880	23	-28888862310	43	-4422651975
5	107920471050	25	23182141500	45	1217211030
7	-47792873520	27	-24695086815	47	651686490
9	-1102024980	29	7313756940	49	-826836885
11	3276815850	31	14354738685	51	404729640
13	-38366469540	33	-20342636055	53	-112651560
15	97991931555	35	13721635620	55	17781120
17	-74144022120	37	-4249642635	57	-1270080
19	-11687638410	39	-3152375100		

Table II.2: Coefficients for the trigonometric polynomials P_3, P_4 and P_5.

where $c(r)$ is a natural number and P_r is a trigonometric polynomial of the form

$$P_r(\theta) = \sum_{j=1}^{d_r} a_j \cos(j\pi\theta),$$

with integer coefficients and degree $d_r = 2^{r+1} - r - 2$ that coincides with the degree of the denominator. In addition, the coefficients a_j vanish except for indexes j such that $j \equiv d_r (\mathrm{mod}\, 2)$.

In order to obtain the coefficients of P_r, we have computed the Taylor expansions of the Arnold Tongues for 120 different rotation numbers. Concretely, we have selected the quadratic irrationals $\theta_{a,b} = (\sqrt{b^2 + 4b/a} - b)/2$, for $1 \leq a \leq b \leq 5$, that have periodic continued fraction given by $\theta_{a,b} = [0; \widehat{a, b}]$. Then, we fix the value of $c(r)$ and perform minimum square fit for the coefficients a_j. We validate the computations if the solution corresponds to integer numbers, or we increase $c(r)$ otherwise. In order to detect if $a_j \in \mathbb{Z}$, we require an arithmetic precision higher than 64 digits. Then, these computations have been implemented in PARI-GP (available at [Par]) using 100-digit arithmetics.

Following the above idea, we have obtained expressions for the next three coefficients. Concretely, we find the values $c(3) = 10$, $c(4) = 19$, and $c(5) = 38$. On the other hand, the corresponding polynomials P_r are given in Table II.2. The comparison between these pseudo-analytical coefficients and the values computed numerically for $\theta = (\sqrt{5} - 1)/2$ is shown in column e_2 of Table II.1, obtaining a very good agreement. Let us observe that the coefficients

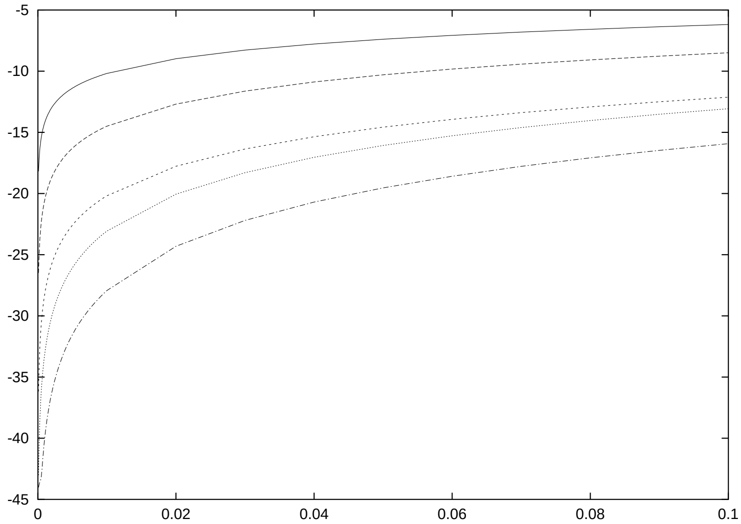

Figure II.4: Comparison between the numerical expressions of $\alpha(\varepsilon)$ for the Arnold Tongue T_θ, with $\theta = (\sqrt{5}-1)/2$, obtained using the Newton method and the truncated Taylor expansion (II.23) up to order d. Concretely we plot, as a function of ε, the difference in $\log_{10}$ scale between these quantities. The curves from top to bottom correspond, respectively, to $d = 2, 4, 6, 8$ and 10.

of P_r grow very fast with respect to r, and the same occurs to $c(r)$. Indeed, the values that correspond to $r = 6$ are too large to be computed with the selected precision, due to the loss of significant digits.

Finally, we also compare the truncated Taylor expansions with the numerical approximation of the Arnold Tongue for $\theta = (\sqrt{5} - 1)/2$, computed using Newton method. To this end, we perform the computation of Section II.4.2 for $\varepsilon \in [0, 0.1]$, using *quadruple-double* precision, an averaging order $p = 9$ and requiring tolerances of 10^{-42} for the computation of the rotation number, and 10^{-40} for the convergence of the Newton method. In all the computations, we allow at most 2^{23} iterates of the map. Then, in Figure II.4 we compare the approximated tongue with the Taylor expansions truncated at orders $2, 4, 6, 8$ and 10.

II.5 Study of invariant curves for planar twist maps

In this last section we deal with a classical problem in dynamical systems that arise in many applications: the study of quasi-periodic invariant curves for planar maps. Concretely, we focus on the context of so-called twist maps, because in this case we can easily make a link with circle diffeomorphisms. First of all, in Section II.5.1 we formalize the problem and fix some notation. Then, in Section II.5.2 we adapt our methodology to compute invariant curves and their evolution with respect to parameters by means of the Newton method. Finally, in Section II.5.3 we follow the ideas of Section II.4.3 and compute asymptotic expansions relating initial conditions and parameters that correspond to invariant curves of fixed rotation number. As an example, we study the neighborhood of the elliptic fixed point for the Hénon map, which appears generically in the study of area-preserving maps.

II.5.1 Description of the problem

Let $\mathcal{A} = \mathbb{T} \times I$ be the real annulus, where I is any real interval, that can be lifted to the strip $A = \mathbb{R} \times I$ using the universal cover $\pi : A \to \mathcal{A}$. Let also $X : A \to \mathbb{R}$ and $Y : A \to I$ denote the canonical projections $X(x, y) = x$ and $Y(x, y) = y$.

In this section, we consider diffeomorphisms $F : \mathcal{A} \to \mathcal{A}$ and their lifts $\widetilde{F} : A \to A$ given by $F \circ \pi = \pi \circ \widetilde{F}$. Note that the lift is unique if we require $X(\widetilde{F}(0, y_0)) \in [0, 1)$ for certain $y_0 \in I$, so we omit the tilde in the lift. In addition, we restrict to maps satisfying that $\partial(X \circ F)/\partial y$ does not vanish, a condition that is called *twist*.

Assume that $F : A \to A$ is a twist map having an invariant curve Γ, homotopic to the circle $\mathbb{T} \times \{0\}$, of rotation number $\theta \in \mathbb{R}\backslash\mathbb{Q}$. Concretely, there exists an embedding $\gamma : \mathbb{R} \to A$, such that $\Gamma = \gamma(\mathbb{R})$, satisfying $\gamma(x + 1) = \gamma(x) + (1, 0)$ for all $x \in \mathbb{R}$, and making the following diagram commute

$$\begin{array}{ccc} \Gamma \subset A & \xrightarrow{F} & \Gamma \subset A \\ \uparrow{\scriptstyle \gamma} & & \uparrow{\scriptstyle \gamma} \\ \mathbb{R} & \xrightarrow{R_\theta} & \mathbb{R} \end{array} \qquad F(\gamma(x)) = \gamma(x + \theta). \tag{II.29}$$

Since F is a twist map, the Birkhoff Graph Theorem (see [Gol01]) ensures that Γ is a Lipschitz graph over its projection on the circle $\mathbb{T} \times \{0\}$, and hence the dynamics on Γ induces a circle homeomorphism f_Γ simply by projecting the iterates, i.e., $f_\Gamma(X(\gamma(x))) = X(F(\gamma(x))$. We observe that, if F and γ are $\mathcal{C}^r$-diffeomorphisms, then $f_\Gamma \in \mathrm{Diff}^r_+(\mathbb{T})$.

From now on, we fix an angle $x_0 \in \mathbb{T}$ and identify invariant curves with points $y_0 \in I$. Then, if (x_0, y_0) belongs to an invariant curve Γ, we also denote the previous circle map as f_{y_0} instead of f_Γ. Of course, the parameterization γ is unknown in general, so we do not have an expression for f_{y_0}. But we can evaluate the orbit $(x_n, y_n) = F^n(x_0, y_0)$ and consider $x_n = f^n_{y_0}(x_0)$. We recall that this is the only that we need to compute numerically the rotation number θ using the method of [SV06] (reviewed in Section II.2.2).

Remark II.5.1. *If the map F does not satisfy the twist condition, their invariant curves are not necessarily graphs over the circle $\mathbb{T} \times \{0\}$. Of course, if Γ is an invariant curve of F, its dynamics still induces a circle diffeomorphism, even though its construction is not so obvious. Since the non-twist case presents another kind of difficulties and has its own interest, we plan to adapt the method to consider the general situation in a subsequent work [LV08].*

If F is a $\mathcal{C}^r$-integrable twist map, then there is a $\mathcal{C}^r$-family of invariant curves of F satisfying (II.29), and $y_0 \mapsto f_{y_0}$ is a one-parameter family in $\mathrm{Diff}^r_+(\mathbb{T})$. In this case, we obtain a $\mathcal{C}^r$-function $y_0 \in I \mapsto \rho(f_{y_0})$. Of course, this is not the general situation and, actually, we do not expect this function to be defined for every $y_0 \in I$. Nevertheless, in many problems we have a family of invariant curves defined on a Cantor subset $J \subset I$ having positive Lebesgue measure and we still have differentiability of $\rho(f_{y_0})$ in the sense of Whitney. For example, if the map F is a perturbation of an integrable twist map that is symplectic or satisfies the intersection condition, KAM theory establishes (under other general assumptions) the existence of such a Cantor family of invariant curves (we refer to [dlL01, Mos62]).

For practical purposes, even if a point $(x_0, y_0) \in A$ does not belong to a quasi-periodic invariant curve, we can compute the orbit $x_n = f_{y_0}^n(x_0) = X(F^n(x_0, y_0))$, even though f_{y_0} is not a circle diffeomorphism. Then, we can also compute the averaged sums $S_N^p(f_{y_0})$ of these iterates but we cannot guarantee in general that $\Theta_{q,p}(f_{y_0})$ converges when $q \to \infty$. Nevertheless, if (x_0, y_0) is an initial condition close enough to an invariant curve of Diophantine rotation number θ, we expect $\Theta_{q,p}(f_{y_0})$ to converge to a number close to θ, due to the existence of neighboring invariant curves for a set of big relative measure (that is called *condensation* phenomena in KAM theory). On the other hand, if (x_0, y_0) belongs to a periodic island, then we expect $\Theta_{q,p}(f_{y_0})$ to converge to the winding number of the "central" periodic orbit. Finally, we recall that the Aubry-Mather theorem (we refer to [Gol01]) states that F has orbits of all rotation numbers, so it can occur that the method converges if (x_0, y_0) corresponds to a periodic orbit or to a ghost curve (Cantori).

On the other hand, in order to approximate the derivatives of the rotation number by means of $\Theta_{q,p,p-d}^d(f_{y_0})$, we have to compute derivatives of the iterates x_n. However, as we do not have an explicit formula for the induced map f_{y_0}, the scheme for computing derivatives of the iterates is slightly different from the one presented in Section II.3.3. Modified recurrences are detailed in the moment that they are required.

II.5.2 Numerical continuation of invariant curves

Let us consider $\alpha : \Lambda \subset \mathbb{R} \mapsto F_\alpha$ a one-parameter family of twist maps on $\mathcal{A}$, that induces a function $(\alpha, y_0) \in U \subset \Lambda \times I \mapsto \rho(f_{\alpha,y_0})$ differentiable in the sense of Whitney. In this situation, we can compute the derivatives of this function (at the points where they exist) by using the method of Section II.3. Our goal now is to use these derivatives to compute numerically invariant curves of F_α by means of the Newton method, similarly as we did in Section II.4.2 for computing the Arnold Tongues.

Concretely, let Γ_{α_0} be an invariant curve of rotation number $\theta \in \mathcal{D}$ for the map F_{α_0}. Then, given any α close to α_0, we want to compute the curve Γ_α, invariant under F_α, having the same rotation number. Once we have fixed an angle $x_0 \in \mathbb{T}$, we identify the invariant curve Γ_α by the point $(x_0, y(\alpha)) \in \Gamma_\alpha$. Then, our purpose is to solve, with respect to y, the equation $\rho(f_{\alpha,y}) = \theta$ by continuing the known solution $(\alpha_0, y(\alpha_0)) \in \Lambda \times I$. We just remark that, when solving this equation by means of the Newton method, we have to prevent us from falling into a resonant island, where the rotation number is locally constant around this point.

Now, in order to approximate numerically $D_\alpha\rho$ and $D_{y_0}\rho$, we have to discuss the computation of the derivatives of the iterates, i.e., $D_\alpha(x_n)$ and $D_{y_0}(x_n)$, where $x_n = f_{\alpha,y_0}^n(x_0)$. Omitting the dependence on the parameter α in the family of twist maps, we denote $F_1 = X \circ F$ and $F_2 = Y \circ F$, and we obtain the recurrent expression

$$D_{y_0}(x_n) = \partial_x F_1(z_{n-1}) D_{y_0}(x_{n-1}) + \partial_y F_1(z_{n-1}) D_{y_0}(y_{n-1}), \tag{II.30}$$

where $z_n := (x_n, y_n)$. Furthermore, $D_{y_0}(y_n)$ follows from a similar expression replacing F_1 by F_2. According to our convention of fixing $x_0 \in \mathbb{T}$, the computations have to be initialized by $D_{y_0}(x_0) = 0$ and $D_{y_0}(y_0) = 1$. Analogous formulas hold for $D_\alpha(x_n)$:

$$D_\alpha(x_n) = \partial_\alpha F_1(z_{n-1}) + \partial_x F_1(z_{n-1}) D_\alpha(x_{n-1}) + \partial_y F_1(z_{n-1}) D_\alpha(y_{n-1}),$$

and similarly for $D_\alpha(y_n)$ using F_2. The recursive computations are now initialized by $D_\alpha(x_0) = 0$ and $D_\alpha(y_0) = 0$.

Let us illustrate the above ideas studying the well known Hénon family, that is a paradigmatic example since it appears generically in the study of a saddle-node bifurcation. In Cartesian coordinates, the family can be written as

$$H_\alpha : \begin{pmatrix} u \\ v \end{pmatrix} \longmapsto \begin{pmatrix} \cos(2\pi\alpha) & -\sin(2\pi\alpha) \\ \sin(2\pi\alpha) & \cos(2\pi\alpha) \end{pmatrix} \begin{pmatrix} u \\ v - u^2 \end{pmatrix}. \tag{II.31}$$

Note that the origin is an elliptic fixed point that corresponds to a "singular" invariant curve. We can blow-up the origin if, for example, we bring the map to the annulus by means of polar coordinates $x = \arctan(v/u)$ and $y = \sqrt{u^2 + v^2}$, thus obtaining a family $\alpha \in \Lambda = [0, 1) \mapsto F_\alpha$ of maps $F_\alpha : \mathbb{S} \times I \mapsto \mathbb{S} \times I$, given by

$$X \circ F_\alpha = \arctan \frac{\sin(x + 2\pi\alpha) - \cos(2\pi\alpha) y (\cos(x))^2}{\cos(x + 2\pi\alpha) + \sin(2\pi\alpha) y (\cos(x))^2}, \tag{II.32}$$

$$Y \circ F_\alpha = y \sqrt{1 - 2y(\cos(x))^2 \sin(x) + y^2 (\cos(x))^4}. \tag{II.33}$$

We remark that, analogously as we did in Section II.4, in this application we consider angles in $\mathbb{S} = \mathbb{R}/2\pi\mathbb{Z}$ in order to avoid factors 2π that would appear in the derivatives (specially in Section II.5.3 when we consider higher order ones).

Albeit it is not difficult to check that the twist condition $\partial(X \circ F)/\partial y \neq 0$ is not fulfilled in these polar coordinates, we can perform a close to the identity change of variables to guarantee the twist condition except for the values $\alpha = 1/3, 2/3$. Then, it turns out that there exist invariant curves of F_α in a neighborhood of $\mathbb{S} \times \{0\}$, whose rotation number tend to α and are "close to the identity" to graphs over $\mathbb{S} \times \{0\}$. However, for values of α close to $1/3$ and $2/3$, meandering phenomena arises (we refer to [DMS00, Sim98b]), i.e., there are folded invariant curves (see Remark II.5.1).

As an example, we study the invariant curves of rotation number $\theta = (\sqrt{5} - 1)/2$ by continuing the initial values $\alpha_0 = \theta$ and $y_0 = 0$, i.e, the curve $\mathbb{S} \times \{0\}$. The computations have been performed by using the *double-double* data type, a fixed averaging order $p = 8$ and up

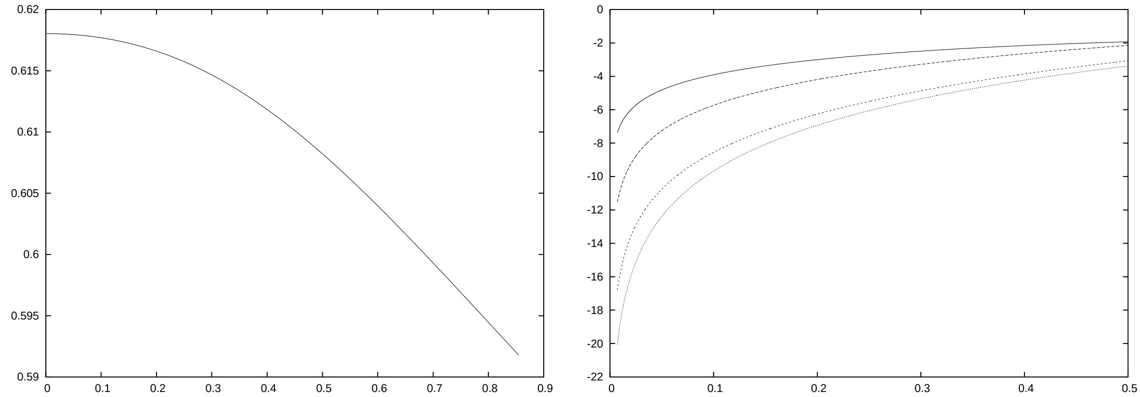

Figure II.5: Left: Numerical continuation of y_0 (horizontal axis) with respect to α (vertical axis) of the invariant curve of rotation number $\theta = (\sqrt{5}-1)/2$ for the Hénon map (II.31). Right: Difference in $\log_{10}$ scale between $\alpha(y_0)$ in the left plot and its truncated Taylor expansion (II.34) up to order d (see Table II.3). The curves from top to bottom correspond, respectively, to $d = 2, 4, 6$ and 8.

to 2^{23} iterates of the map, at most. As usual, we estimate the error in the rotation number by using (II.13), and we validate the computation when the error is smaller than 10^{-26}. For the Newton method, we require a tolerance smaller than 10^{-23} when comparing two successive computations. Finally, we do not require a prescribed tolerance in the computation of the derivatives $D_\alpha \rho$ and $D_{y_0}\rho$, but the largest error in their computation is less that 10^{-21}.

The resulting curve in the space $I \times \Lambda$ is shown in Figure II.5 (left). During the continuation, the step in α is typically taken between 10^{-4} and 10^{-3}, but falls to 10^{-10} when we compute the last point $(\alpha, y(\alpha)) = (0.5917905628, 0.8545569509)$. In Figure II.6 we plot the graph corresponding to this invariant curve and its derivative. We observe that, even though the curve is still a graph, this parameterization is close to have a vertical tangency, so our approach is not suitable for continuing the curve. However, since the fractalization of the curve has not occurred, we expect that it still exists beyond this point.

II.5.3 Computing expansions with respect to parameters

In the same situation of Section II.5.2, our aim now is to use the variational information of the rotation number to compute the Taylor expansion at the origin of Figure II.5 (left). Notice that in the selected example $\alpha'(0) = 0$, so we work with the expansion of the function $\alpha(y_0)$ rather than $y_0(\alpha)$.

In general, if (x_0, y_0^*) is a point on an invariant curve of rotation number θ for a twist map F_{α^*}, then we consider the expansion

$$\alpha(y_0) = \alpha^* + \alpha'(y_0^*)(y_0 - y_0^*) + \frac{\alpha''(y_0^*)}{2!}(y_0 - y_0^*)^2 + \cdots , \tag{II.34}$$

that corresponds to the value of the parameter for which (x_0, y_0) is contained in an invariant curve of $F_{\alpha(y_0)}$ having the same rotation number. We know that if $\theta \in \mathcal{D}$ and the family F_α is analytic, then (II.34) is an analytic function around y_0^*. Once again, during the rest of the section, we omit the dependence on the parameter α in the family of twist maps, and we denote $F_1 = X \circ F$ and $F_2 = Y \circ F$.

As in Section II.4.3, we use that the family $y_0 \mapsto f_{\alpha(y_0),y_0} \in \mathrm{Diff}_+^\omega(\mathbb{S})$ induced by $y_0 \mapsto F_{\alpha(y_0)}$ has constant rotation number, together with Remark II.3.3. Concretely, for any integer $d \geq 1$ we have

$$0 = \Theta_{q,p,p}^d(f_{\alpha(y_0),y_0}) + \mathcal{O}(2^{-(p+1)q}), \qquad \text{(II.35)}$$

where $\Theta_{q,p,p}^d$ is the extrapolation operator (II.18). We observe that the value of $\Theta_{q,p,p}^d(f_{y_0})$ at the point y_0^* only depends on the derivatives $\alpha^{(r)}(y_0^*)$ up to $r \leq d$. We use this fact to compute inductively these derivatives from equation (II.35). To achieve this, we have to isolate them from $D_{y_0}^d(x_n)|_{y_0=y_0^*}$ for any $d \geq 1$, as we discuss through the next paragraphs.

The following formula generalizes (II.30):

$$\begin{aligned} D_{y_0}^d(x_n) &= \sum_{j=0}^{d-1} \binom{d-1}{j} \Big\{ D_{y_0}^j(\partial_\alpha F_1(z_{n-1}))\, \alpha^{(d-j)}(y_0) \\ &\quad + D_{y_0}^j(\partial_x F_1(z_{n-1})) D_{y_0}^{d-j}(x_{n-1}) + D_{y_0}^j(\partial_y F_1(z_{n-1})) D_{y_0}^{d-j}(y_{n-1}) \Big\}, \end{aligned} \qquad \text{(II.36)}$$

while a similar equation holds for $D_{y_0}^d(y_n)$ replacing F_1 by F_2. Moreover, as in Section II.3.3, we compute the derivatives $D_{y_0}^r$ of $\partial_\alpha F_1(z_{n-1})$, $\partial_x F_1(z_{n-1})$ and $\partial_y F_1(z_{n-1})$ by means of the following recurrent expression

$$\begin{aligned} D_{y_0}^r(\partial_{\alpha,x,y}^{k,l,m} F_i(z_{n-1})) &= \sum_{j=0}^{r-1} \binom{r-1}{j} \Big\{ D_{y_0}^j(\partial_{\alpha,x,y}^{k+1,l,m} F_i(z_{n-1}))\, \alpha^{(r-j)}(y_0) \\ &\quad + D_{y_0}^j(\partial_{\alpha,x,y}^{k,l+1,m} F_i(z_{n-1})) D_{y_0}^{r-j}(x_{n-1}) \\ &\quad + D_{y_0}^j(\partial_{\alpha,x,y}^{k,l,m+1} F_i(z_{n-1})) D_{y_0}^{r-j}(y_{n-1}) \Big\}, \end{aligned}$$

which only requires to evaluate the partial derivatives of F_1 and F_2 with respect to α, x and y.

Using the above expressions, we reproduce the inductive argument of Section II.4.3. Let us assume that the values $\alpha'(y_0^*), \ldots, \alpha^{(d-1)}(y_0^*)$ are known. Then, we observe that if we set $y_0 = y_0^*$ and $\alpha = \alpha^*$ in equation (II.36), the only term containing the derivative $\alpha^{(d)}(y_0^*)$ is the one corresponding to $j = 0$. By induction, it is easy to find that

$$D_{y_0}^d(x_n)|_{y_0=y_0^*} = \mathcal{X}_n^d\, \alpha^{(d)}(y_0^*) + \hat{\mathcal{X}}_n^d, \qquad D_{y_0}^d(y_n)|_{y_0=y_0^*} = \mathcal{Y}_n^d\, \alpha^{(d)}(y_0^*) + \hat{\mathcal{Y}}_n^d,$$

where the coefficients $\mathcal{X}_n^d$, $\hat{\mathcal{X}}_n^d$, $\mathcal{Y}_n^d$ and $\hat{\mathcal{Y}}_n^d$ are obtained recursively and only depend on the

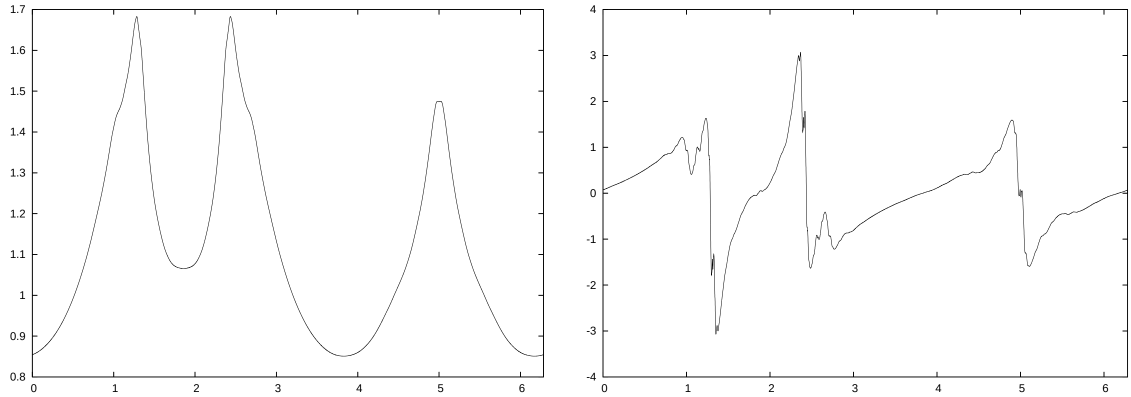

Figure II.6: Left: Invariant curve of (II.31) of rotation number $\theta = (\sqrt{5}-1)/2$ corresponding to the last computed point in Figure II.5 (see text) expressed as a graph $x \mapsto y$ on the annulus $\mathbb{S} \times I$. Right: Derivative of the left plot computed using finite differences.

derivatives $\alpha^{(r)}(y_0^*)$, with $r < d$. Concretely, $\mathcal{X}_n^d$ and $\hat{\mathcal{X}}_n^d$ satisfy

$$\begin{aligned} \mathcal{X}_n^d &= \left(\partial_\alpha F_1(z_{n-1}) + \partial_x F_1(z_{n-1})\mathcal{X}_{n-1}^d + \partial_y F_1(z_{n-1})\mathcal{Y}_{n-1}^d\right)|_{y_0=y_0^*}, \\ \hat{\mathcal{X}}_n^d &= \Bigg(\partial_x F_1(z_{n-1})\hat{\mathcal{X}}_{n-1}^d + \partial_y F_1(z_{n-1})\hat{\mathcal{Y}}_{n-1}^d + \sum_{j=1}^{d-1}\binom{d-1}{j}\Big\{D_{y_0}^j(\partial_\alpha F_1(z_{n-1}))\alpha^{(d-j)}(y_0) \\ &\quad + D_{y_0}^j(\partial_x F_1(z_{n-1}))D_{y_0}^{d-j}(x_{n-1}) + D_{y_0}^j(\partial_y F_1(z_{n-1}))D_{y_0}^{d-j}(y_{n-1})\Big\}\Bigg)\Bigg|_{y_0=y_0^*}, \end{aligned}$$

and similar equations hold for $\mathcal{Y}_n^d$ and $\hat{\mathcal{Y}}_n^d$ replacing F_1 by F_2. These sequences are initialized as

$$\mathcal{X}_0^1 := \hat{\mathcal{X}}_0^1 := \mathcal{Y}_0^1 := 0,\ \hat{\mathcal{Y}}_0^1 := 1, \quad \text{and} \quad \mathcal{X}_0^d := \hat{\mathcal{X}}_0^d := \mathcal{Y}_0^d := \hat{\mathcal{Y}}_0^d := 0, \quad \text{for } d > 1.$$

Finally, if we evaluate the extrapolation operator $\Theta_{q,p}$ for the sequences $\mathcal{X}^d = \{\mathcal{X}_n^d\}_{n=1,\ldots,N}$ and $\hat{\mathcal{X}}^d = \{\hat{\mathcal{X}}_n^d\}_{n=1,\ldots,N}$, then we obtain from (II.35) the following expression

$$\alpha^{(d)}(y_0^*) = -\frac{\Theta_{q,p}(\hat{\mathcal{X}}^d)}{\Theta_{q,p}(\mathcal{X}^d)} + \mathcal{O}(2^{-(p+1)q}).$$

Now, we apply this methodology to the Hénon family $\alpha \in \Lambda = [0,1) \mapsto F_\alpha$ given by (II.32) and (II.33). In particular, we fix $x_0 = 0$ and compute the expansion (II.34) at $y_0^* = 0$ corresponding to invariant curves of rotation number $\alpha^* = \theta = (\sqrt{5}-1)/2$.

Observe that the derivatives of this map are hard to compute explicitly, so we have to introduce another recursive scheme for them. Moreover, in order reduce the amount of computations, we use that the iterates of $(0,0)$ are $x_n = 2\pi n\theta$ and $y_n = 0$.

d	$2\pi\alpha^{(d)}(0)$	e_1
0	3.8832220774509331546937312599254	-
1	$2.9215929940647956972904287221575 \cdot 10^{-29}$	$6 \cdot 10^{-27}$
2	$-3.9914536995187621201317645570286 \cdot 10^{-1}$	$6 \cdot 10^{-28}$
3	$-7.2312013917244657534375078612123 \cdot 10^{-1}$	$5 \cdot 10^{-27}$
4	$-1.4570409862191278806067261207843 \cdot 10^{0}$	$9 \cdot 10^{-27}$
5	$2.0167847130561842764416032369501 \cdot 10^{1}$	$4 \cdot 10^{-26}$
6	$1.2357011948811946999538300791232 \cdot 10^{2}$	$1 \cdot 10^{-25}$
7	$-9.1717201199029959021691212417954 \cdot 10^{1}$	$2 \cdot 10^{-25}$
8	$-3.0832824868383111456060167381447 \cdot 10^{3}$	$4 \cdot 10^{-24}$
9	$-7.2541251340271326844826925983923 \cdot 10^{4}$	$2 \cdot 10^{-22}$

Table II.3: Derivatives of $2\pi\alpha(y_0)$ at the origin for $\theta = (\sqrt{5}-1)/2$. The column e_1 corresponds to the estimated error using (II.13).

We detail the computations of $\partial^{k,l,m}_{\alpha,x,y}(Y \circ F_\alpha)$ at the point $(\alpha, x, y) = (\theta, x_n, 0)$, while the derivatives of $X \circ F_\alpha$ satisfy completely analogous expressions. Let us introduce the function

$$g(x,y) = 1 - 2y(\cos(x))^2 \sin(x) + y^2(\cos(x))^4,$$

so we can write $Y \circ F_\alpha(x,y) = y\sqrt{g(x,y)}$. First, we observe that for any $s \in \mathbb{Q}$ we have $\partial^{k,l,m}_{\alpha,x,y}(yg^s)|_{y=0} = 0$ provided $k \neq 0$ or $m = 0$. Otherwise, the required derivatives can be computed by means of the following recurrent expressions

$$\partial^{l,m}_{x,y}(yg^s) = \partial^{l,m-1}_{x,y}(g^s) + s\sum_{i=0}^{l}\sum_{j=0}^{m-1}\binom{l}{i}\binom{m-1}{j}\partial^{i,j}_{x,y}(yg^{s-1})\partial^{l-i,m-j}_{x,y}(g)$$

and

$$\partial^{l,m}_{x,y}(g^s) = s\sum_{i=0}^{l}\sum_{j=0}^{m-1}\binom{l}{i}\binom{m-1}{j}\partial^{i,j}_{x,y}(yg^{s-1})\partial^{l-i,m-j}_{x,y}(g).$$

Finally, we observe that the derivatives $\partial^{l-i,m-j}_{x,y}(g)$ can be computed easily by expanding the function as a trigonometric polynomial

$$g(x,y) = 1 - \frac{y}{2}\Bigg(\sin(3x) + \sin(x)\Bigg) + \frac{y^2}{2}\left(\frac{3}{4} + \cos(2x) + \frac{1}{4}\cos(4x)\right).$$

The computations are performed by using *double-double* data type, $p = 7$ and 2^{21} iterates, at most. We stop the computations if the estimated error is less than 10^{-25}. Derivatives of the expansion (II.34) and their estimated error, are given in Table II.3. Finally, in order to verify the results, we compare the truncated expansions of the curve with the numerical approximation computed in Section II.5.2. The deviation is plotted in $\log_{10}$ scale in Figure II.5 (right).

Chapter III

Numerical computation of rotation numbers of quasi-periodic planar curves

Abstract of the chapter. *Recently, a new numerical method has been proposed to compute rotation numbers of analytic circle diffeomorphisms, as well as derivatives with respect to parameters, that takes advantage of the existence of an analytic conjugation to a rigid rotation. This method can be directly applied to the study of invariant curves of planar twist maps by simply projecting the iterates of the curve onto a circle. In this work we extend the methodology to deal with general planar maps. Our approach consists in computing suitable averages of the iterates of the map that allow us to obtain a new curve for which the direct projection onto a circle is well posed. Furthermore, since our construction does not use the invariance of the quasi-periodic curve under the map, it can be applied to more general contexts. We illustrate the method with several examples.*

III.1 Introduction

In this chapter we present numerical algorithms to deal with quasi-periodic invariant curves of planar maps by adapting a method presented in [SV06] to compute rotation numbers of analytic circle diffeomorphisms. The developed ideas do not require the curve to be invariant under any map; so they can be applied to more general objects that we refer to as *quasi-periodic signals* (see Definition III.2.2).

The method of [SV06] is built assuming that the circle map is analytically[1] conjugate to a rigid rotation and, basically, it consists in computing suitable averages of the iterates of the map followed by Richardson extrapolation. Since this construction takes advantage of the geometry and the dynamics of the problem, the method turns out to be highly accurate and very efficient in multiple applications. In a few words, if we compute N iterates of the map, then we can

[1]The methods of [LV08, SV06] also work in the class of $\mathcal{C}^r$ circle diffeomorphisms, r being sufficiently large, but we restrict the discussion to the analytic case in order to simplify the exposition.

approximate the rotation number with an error of order $\mathcal{O}(1/N^{p+1})$ where p is the selected order of averaging (compared with $\mathcal{O}(1/N)$ obtained using the definition). This methodology has been extended in Chapter II to deal with derivatives of the rotation number with respect to parameters. In this case, it is required to compute and average the corresponding derivatives of the iterates of the circle map. We want to point out that this variational information cannot be obtained in such a direct way by means of other existing methods to compute rotation numbers (we refer to the works [BS00, Bru92, dlLP02, GMS, LFC92, vV88]).

As a matter of motivation, let us assume first that F is a map on the real annulus $\mathbb{T} \times I$, where I is a real interval and $\mathbb{T} = \mathbb{R}/\mathbb{Z}$, and let $X : \mathbb{T} \times I \to \mathbb{R}$ denote the canonical projection $X(x, y) = x$. If F is a twist map, the Birkhoff Graph Theorem (see [Gol01]) ensures that every invariant curve Γ is a graph over its projection on the circle by means of X, and its dynamics induces a circle map by projecting the iterates. Hence, it is straightforward to apply the method of [SV06] in order to approximate the rotation number of Γ, since for any $(x_0, y_0) \in \Gamma$ we can compute the orbit $x_n = X(F^n(x_0, y_0))$ —this is the only data that the method requires. Furthermore, if F has a differentiable family of invariant curves or a Cantorian family differentiable in the sense of Whitney, we can approximate derivatives of the rotation number with respect to initial conditions and parameters. This allows us to implement a Newton scheme for the computation and continuation of invariant curves of twist maps (as it is discussed in detail in Section II.5).

If the map does not satisfy the twist condition or it is not written in suitable coordinates, its invariant curves are not necessarily graphs over the projection on a circle. In this situation, invariant curves can fold in a very wild way (see Section III.3.3 and references given therein for examples of such curves). Nevertheless, if we can select a suitable circle so that the folded curve "rotates" around it, then the projection of the iterates of the map does not define a circle map but a "circle correspondence" and we can compute the rotation number of the curve from the "lift" of this correspondence to the real line —see Section III.2.1 for details. Moreover, albeit we do not have a justification of this fact, we realize that the extrapolation methods work quite well when applied to the iterates of this "lift".

In some cases —for example, if the rotation number is large compared with the size of the folds— we can compute numerically this "lift" from the iterates of the map. However, if the curve is extremely folded additional work is required in order to face the problem in a systematic way. Hence, we propose a numerical method to construct a circle map —preserving the rotation number— from a general invariant curve on the plane. The method consists in averaging the iterates of an orbit of the curve in such a way that the new iterates belong to another curve, no longer invariant under the map, but having the same rotation number. Concretely, if we know an approximation of the rotation number with error ε, we construct a sequence of (averaged) curves that approaches a circle up to terms of order $\mathcal{O}(\varepsilon)$. We refer to this construction as *unfolding of the curve* since if ε is small enough, then this construction provides us with a circle map. Taking into account the discussion in the previous paragraph, in order to apply the methods of Chapter II it is not necessary to unfold completely the curve, but only to "kill" its main folds so we can compute the "lift" of the correspondence generated by the projection of the iterates of the new (less-folded) curve. In order to justify this unfolding procedure we require the curve to be analytic (or at least differentiable enough) and the rotation number to be Diophantine. Sometimes the requested approximation of the rotation number is given by the context of the problem —for example, if we look for invariant curves of fixed rotation number— or it can be obtained by means of any method of frequency analysis (see for example [GMS, LFC92]). Therefore, we obtain a very efficient toolkit for the study of invariant curves of planar maps and their numerical continuation.

Let us remark that due to the importance, both theoretical and applied, of invariant curves of maps or 2-dimensional tori of flows (for example, they play a fundamental role in the design of space missions [GJSM01a, GJSM01b] and also in the study of models in Celestial Mechanics [SM71], Molecular Dynamics [PCU08, TLBB01] or Plasma-Beam Physics [Mic95], just to mention a few), several approaches to deal with these objects have been developed in the literature. For example, the methods in [CJ00, dlLHS, JO] have been applied efficiently in a wide set of contexts. However, they require to compute a representation —by means of a trigonometric polynomial— of the curve which solves the invariance equation of the problem, so it is required to solve large systems of equations —as large as the used number of Fourier modes, say M. One possibility to face this difficulty is to solve these full linear systems, with a cost $\mathcal{O}(M^3)$ in time and $\mathcal{O}(M^2)$ in memory, by means of efficient parallel algorithms as is proposed in [JO]. Another recent approach presented in [dlLHS], based on the analytic and geometric ideas developed in [dlLGJV05], allows us to reduce the computational effort of the problem to a cost of order $\mathcal{O}(M \log M)$ in time and $\mathcal{O}(M)$ in memory. On the other hand, we can compute the invariant curve by looking for a point so that the corresponding orbit has a prefixed rotation number. Then, rather than approximating explicitly the parameterization of the curve, we reduce the problem to finding a zero of a function. This approach can be implemented using interpolation methods as in [Sim98a] or also using the extrapolation methods discussed in Chapter II. These extrapolation methods, that are also the cornerstone of this chapter, have a cost of order $\mathcal{O}(N \log N)$ in terms of the used number of iterates N and are free in memory. Once we know a point on the curve and its rotation number, we can compute a trigonometric approximation of the curve "a posteriori", using Fourier Transform (FT) on the iterates of the curve. In addition, in Section III.2.7 we develop a method for performing this FT based also on averaging-extrapolation ideas.

Given a numerical method for the continuation of invariant curves, it is specially interesting to verify if the method is valid up to the *breakdown threshold* corresponding to the *critical invariant curve* (see [dlLO06, Gre79, OS87]). These critical curves are specially important objects that organize the long-term behavior of a given dynamical system, because of their role as "last barriers" or "bottlenecks" to chaos (see [Gol01]). Actually, the critical value for the breakdown of the golden curve for the Chirikov standard map was estimated by means of extrapolation methods in [SV06] obtaining a good agreement with the value predicted by means of the classical Greene criterion in [Gre79]. For the non-twist case, we refer to computations in [AdlLP05, dCNGM96] as examples of breakdown studies in non-twist maps. It is worth mentioning that the methods presented in this chapter can be applied also in this context.

Since our construction does not use the invariance of the curve under the map, it can be applied to the study of quasi-periodic curves that are not necessarily embedded (that we call quasi-periodic signals). This context is very interesting since it allows us to analyze sets of data obtained from real experiments or observed natural phenomena. Actually, in order to check that the methods are robust when facing experimental data, we consider the effect of Gaussian error in the evaluation of iterates of a known quasi-periodic function.

We want to point out that our approach can be also understood as a method for the refinement of the frequency analysis of [LFC92]. Actually, an efficient refinement of these methods, based in the simultaneous improvement of the frequencies and the amplitudes of the signal, is given in [GMS]. Once again, the main advantage of our approach is that we do not have to compute Fourier coefficients of the curve. This fact reduces the computational effort of solving big linear systems of equations required to refine the representation of the signal. In addition, the accuracy in the computation of the rotation number is not limited by the truncation error in the representation of the signal.

Finally, we notice that the methodology of Chapter II also works for dealing with maps of the d-dimensional torus that admit an analytic conjugation to a rigid rotation having a Diophantine rotation vector. Our aim is to explore the extension of the ideas presented in this chapter to deal with invariant tori and quasi-periodic signals of arbitrary number of frequencies.

The contents are organized as follows. In the first part, contained in Section III.2, we develop and justify different results, methods and algorithms to study quasi-periodic invariant curves (or quasi-periodic signals). In the second part, presented in Section III.3, we consider several examples in order to illustrate different features of the presented methodology. These examples have been selected in order to sustain the presentation of the methods and to highlight both some of the possibilities and limitations of our approach.

III.2 Exposition of methods

As we said in the introduction, we approach the study of quasi-periodic signals by computing the rotation number of a circle map (or a circle correspondence) induced by the curve. The main definitions and notation, together with a brief overview of the problem, are given in Section III.2.1. After that, we present and justify a method to unfold a quasi-periodic signal. We first assume in Section III.2.2 that the rotation number is known exactly in order to highlight the involved ideas. Basically, we construct a sequence of curves that converges to a circle whose dynamics corresponds to a rigid rotation. In Section III.2.3 we assume that we only have an approximation of the rotation number and we show that the previous construction allows us to obtain a curve that is $\mathcal{C}^1$-close to be a circle —the proximity being of the same order as the error in the initial guess of the rotation number.

In order to obtain the required approximation, a possibility is to resort to frequency analysis methods. In Section III.2.4 we review Laskar's frequency analysis method in terms of the language presented in this chapter, just to stress that the same algorithms derived to unfold the curve can be adapted to obtain the required approximation of the rotation number as alternative of the classical methods.

For the sake of completeness we include in Section III.2.5 some details about the application of the methods of Chapter II to the obtained circle correspondence. Indeed, it is convenient to have Chapter II in mind to understand the higher order method that we develop in Section III.2.6 to improve the unfolding of curves. During the exposition it will be clear that the ideas used in the unfolding are related to FT. This fact is exploited in Section III.2.7 in order to extrapolate Fourier coefficients once the rotation number is known.

III.2.1 Setting of the problem

For convenience, we identify the real plane with the set of complex numbers by defining $z = u + \mathrm{i}v$ for any $(u, v) \in \mathbb{R}^2$. Let $\Gamma \subset \mathbb{C}$ be a quasi-periodic invariant curve for a map $F : U \subset \mathbb{C} \to \mathbb{C}$ of rotation number $\theta \in \mathbb{R}\backslash\mathbb{Q}$. Let us assume, for example, that the curve "rotates" around the origin and that it is a graph of the angular variable. Then, the projection

$$\begin{array}{ccc} \Gamma & \longrightarrow & \mathbb{T} \\ z & \longmapsto & x = \arg(z)/2\pi \end{array} \tag{III.1}$$

generates a circle map induced by the dynamics of $F|_\Gamma$. On the other hand, if Γ is folded, then the projection (III.1) does not provide a circle map, but defines a correspondence on $\mathbb{T}$ that we can "lift" to $\mathbb{R}$. For example, in the left plot of Figure III.1 we show a "folded" invariant curve on the complex plane for an example considered in Section III.3.3. In the right plot of Figure III.1 we show the "lift" of the correspondence on $\mathbb{T}$ given by (III.1). Since the rotation number of the curve is no more than the averaged number of revolutions per iterate, it is not surprising that we can compute it as $\lim_{n\to\infty}(x_n - x_0)/n$, where x_n are the iterates under the "lift" to $\mathbb{R}$ of the circle correspondence. In this situation, we have observed that the methods of Chapter II can be applied to such a "lift" (see examples in Section III.3.3), even though we do not have a justification of this fact.

In some cases, for example if the rotation number is large enough as to avoid the folds, we can compute numerically the "lift" of (III.1) using the iterates of an orbit. However, if the invariant curve presents large folds or we cannot identify directly a good point around which the curve is rotating, we cannot compute this "lift" in a systematic way. Then, our aim is to construct another curve, having the same rotation number, by using suitable averages of iterates of the original map. If we manage to eliminate (or at least minimize) the folds in the new curve, then we are able to obtain a circle diffeomorphism (or at least a circle correspondence that we can "lift" numerically).

As Γ is a quasi-periodic invariant curve of rotation number θ, there exists an analytic embedding $\gamma : \mathbb{T} \to \mathbb{C}$ verifying $\Gamma = \gamma(\mathbb{T})$ and

$$F(\gamma(x)) = \gamma(x + \theta).$$

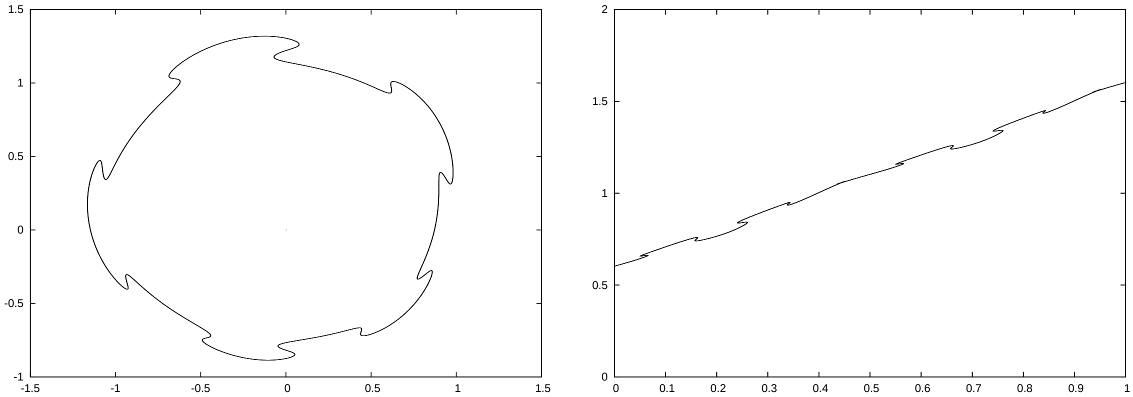

Figure III.1: Left: Folded invariant curve with quasi-periodic dynamics that rotates around the origin in the complex plane (this curve corresponds to an example discussed in Section III.3.3). Right: "Lift" of the associated circle correspondence given by (III.1).

In this situation, since the parameterization γ is periodic, we can use the Fourier series

$$\gamma(x) = \sum_{k\in\mathbb{Z}} \hat{\gamma}_k \mathrm{e}^{2\pi \mathrm{i} k x},$$

and, moreover, for a given $z_0 \in \Gamma$ we can ask for $\gamma(0) = z_0$. Then, the iterates of z_0 under F can be expressed using γ as

$$z_n = F^n(z_0) = F^n(\gamma(0)) = F^{n-1}(\gamma(\theta)) = \gamma(n\theta) = \sum_{k\in\mathbb{Z}} \hat{\gamma}_k \mathrm{e}^{2\pi \mathrm{i} k n\theta}. \tag{III.2}$$

As we will see, our method does not use the invariance of Γ under F but only the expression (III.2) for the iterates. Furthermore, even if we start with an invariant curve of a map, the intermediate stages of our construction may produce curves that are not embedded in $\mathbb{C}$. Using this fact as a motivation, we state the following definitions:

Definition III.2.1. *We say that a complex sequence $\{z_n\}_{n\in\mathbb{Z}}$ is a* quasi-periodic signal *of rotation number θ if there exists a periodic function $\gamma : \mathbb{T} \to \mathbb{C}$ such that $z_n = \gamma(n\theta)$. We also call $\Gamma = \gamma(\mathbb{T})$ a quasi-periodic curve.*

Definition III.2.2. *Under the above conditions, let $\{z_n\}_{n\in\mathbb{Z}}$ be a quasi-periodic signal. Then, for any $\theta_0 \in \mathbb{R}$ and $L \in \mathbb{N}$, we define the following iterates*

$$z_n^{(L,\theta_0)} = \frac{1}{L} \sum_{m=n}^{L+n-1} z_m \mathrm{e}^{2\pi \mathrm{i}(n-m)\theta_0}. \tag{III.3}$$

It is clear that $\{z_n^{(L,\theta_0)}\}_{n\in\mathbb{Z}}$ is a quasi-periodic signal on another curve $\Gamma^{(L,\theta_0)} = \gamma^{(L,\theta_0)}(\mathbb{T})$ of the same rotation number, i.e.,

$$z_n^{(L,\theta_0)} = \gamma^{(L,\theta_0)}(n\theta) = \sum_{k\in\mathbb{Z}} \hat{\gamma}_k^{(L,\theta_0)} \mathrm{e}^{2\pi \mathrm{i} kn\theta}, \tag{III.4}$$

and the new Fourier coefficients are given by

$$\hat{\gamma}_k^{(L,\theta_0)} = \frac{\hat{\gamma}_k}{L} \sum_{m=n}^{L+n-1} \mathrm{e}^{2\pi\mathrm{i}(m-n)(k\theta-\theta_0)} = \frac{\hat{\gamma}_k}{L} \frac{1-\mathrm{e}^{2\pi\mathrm{i}L(k\theta-\theta_0)}}{1-\mathrm{e}^{2\pi\mathrm{i}(k\theta-\theta_0)}}, \tag{III.5}$$

In Section III.2.2 we show that, under conditions on regularity and non-resonance, if $\theta_0 = \theta$, then the new curve $\gamma^{(L,\theta)}$ is arbitrarily $\mathcal{C}^1$-close to a circle (see Lemma III.2.6) for L large enough. On the other hand, if $\varepsilon = \theta_0 - \theta$ is small, then we can choose $L = L(\varepsilon)$ such that the new curve is $\mathcal{C}^1$-close to a circle with an error of order $\mathcal{O}(\varepsilon)$ (this is concluded from Proposition III.2.8) so that the projection

$$\begin{array}{rcl} \Gamma^{(L,\theta_0)} \subset \mathbb{C}^* & \longrightarrow & \mathbb{T} \\ z_n^{(L,\theta_0)} & \longmapsto & x_n^{(L,\theta_0)} = \arg(z_n^{(L,\theta_0)})/2\pi, \end{array} \tag{III.6}$$

provides an orbit of a circle diffeomorphism $f_\Gamma^{(L,\theta_0)}$. Once this circle map has been obtained (as we have discussed, in practice it suffices to obtain a slightly folded curve such that we can compute the "lift" of the circle correspondence defined by the direct projection), we can apply Algorithms II.2.5 and II.3.2 to compute the rotation number and derivatives with respect to parameters (this is described in Section III.2.5). In order to justify this construction, we require the rotation number to be Diophantine.

Definition III.2.3. *Given $\theta \in \mathbb{R}$, we say that θ is a* Diophantine number *of (C,τ) type if there exist constants $C > 0$ and $\tau \geq 1$ such that*

$$|k\theta - l|^{-1} \leq C|k|^\tau, \qquad \forall (l,k) \in \mathbb{Z}\times\mathbb{Z}_*. \tag{III.7}$$

We will denote $\mathcal{D}(C,\tau)$ the set of such numbers and $\mathcal{D}$ the set of Diophantine numbers of any type.

In the aim of KAM theory, we know that the hypothesis of Diophantine rotation number for the dynamics on the curve is consistent with its own existence. Although Diophantine sets are Cantorian —i.e., compact, perfect and nowhere dense— a remarkable property is that $\mathbb{R}\backslash\mathcal{D}$ has zero Lebesgue measure. For this reason, this condition fits very well in practical issues and we do not resort to other weaker conditions on small divisors such as the Brjuno condition (see [Yoc95]). It is worth mentioning that if θ is a "bad" Diophantine rotation number, i.e., having a large constant C in (III.7), then the methods presented in this chapter turn out to be less efficient as we discuss in Section III.3.

III.2.2 Unfolding a curve of known rotation number

Let us consider the previous setting and suppose that the Fourier coefficients of γ are $\hat{\gamma}_1 \neq 0$ and $\hat{\gamma}_k = 0$ for $k \in \mathbb{Z}\backslash\{1\}$. In this case, $\Gamma = \{z \in \mathbb{C} \mid |z| = |\hat{\gamma}_1|\}$ and the circle map obtained by the projection (III.1) is a rigid rotation $R_\theta(x) = x + \theta$.

Assume now that $\hat{\gamma}_1 \neq 0$ and $|\hat{\gamma}_k|$ is small (compared with $|\hat{\gamma}_1|$) for every $k \in \mathbb{Z}\backslash\{1\}$ in such a way that the curve γ is also $\mathcal{C}^1$-close to a circle. In this case, the projection $x = \arg(z)/2\pi$ also makes sense and defines a circle diffeomorphism.

In the general case this projection does not provide a circle map. However, it turns out that the projection of the iterates $z_n^{(L,\theta_0)}$ is well posed if we take $\theta_0 = \theta$ and L large enough. More quantitatively, we assert that the curve $\gamma^{(L,\theta)}$ differs from a circle by an amount of order $\mathcal{O}(1/L)$. In Lemma III.2.6 bellow we make precise the above arguments.

Definition III.2.4. *Given an analytic function $\gamma : \mathbb{T} \to \mathbb{C}$ and its Fourier coefficients $\{\hat{\gamma}_k\}_{k\in\mathbb{Z}}$, we consider the norm $||\gamma|| = \sum_{k\in\mathbb{Z}} |\hat{\gamma}_k|$.*

Definition III.2.5. *Given $k \in \mathbb{Z}\backslash\{0\}$ and $r \in \mathbb{C}$, we define the map $\gamma_k[r] : \mathbb{T} \to \mathbb{C}$ as*

$$\gamma_k[r](x) = r\, \mathrm{e}^{2\pi \mathrm{i} k x}.$$

Then we state the following result:

Lemma III.2.6. *Let us consider a quasi-periodic signal $z_n = \gamma(n\theta)$ of rotation number $\theta \in \mathcal{D}(C,\tau)$. Assume that $\gamma : \mathbb{T} \to \mathbb{C}$ is analytic in the complex strip $B_\Delta = \{z \in \mathbb{C} \,:\, |\mathrm{Im}(z)| < \Delta\}$ and bounded in the closure, with $M = \sup_{z\in B_\Delta} |\gamma(z)|$. Then, if $\hat{\gamma}_1 \neq 0$, the curve $\gamma^{(L,\theta)} : \mathbb{T} \to \mathbb{C}$ given by* (III.4) *and* (III.5) *satisfies*

$$||\gamma^{(L,\theta)} - \gamma_1[\hat{\gamma}_1]|| \leq \frac{A}{L},$$

where A is a constant depending on M, C, τ and Δ.

Proof. First, let us observe that the Fourier coefficients of the new curve are given by

$$\hat{\gamma}_1^{(L,\theta)} = \hat{\gamma}_1, \qquad \hat{\gamma}_k^{(L,\theta)} = \frac{\hat{\gamma}_k}{L} \frac{1 - \mathrm{e}^{2\pi \mathrm{i}(k-1)\theta L}}{1 - \mathrm{e}^{2\pi \mathrm{i}(k-1)\theta}}, \qquad k \in \mathbb{Z}\backslash\{1\}.$$

Then, we have to bound the expression

$$||\gamma^{(L,\theta_0)} - \gamma_1[\hat{\gamma}_1]|| \leq \frac{1}{L} \sum_{k\in\mathbb{Z}\backslash\{1\}} \left| \hat{\gamma}_k \frac{1 - \mathrm{e}^{2\pi \mathrm{i}(k-1)\theta L}}{1 - \mathrm{e}^{2\pi \mathrm{i}(k-1)\theta}} \right| \leq \frac{1}{L} \sum_{k\in\mathbb{Z}\backslash\{1\}} \frac{2|\hat{\gamma}_k|}{|1 - \mathrm{e}^{2\pi \mathrm{i}(k-1)\theta}|}.$$

We observe that the Fourier coefficients of γ satisfy $|\hat{\gamma}_k| \leq M\mathrm{e}^{-2\pi\Delta|k|}$ and we use (III.7) to control the small divisors. Concretely, standard manipulations show that (see for example [AP90])

$$|1 - \mathrm{e}^{2\pi \mathrm{i}(k-1)\theta}|^{-1} \leq \frac{C}{4} |k-1|^\tau.$$

Introducing these expressions in the previous sum we obtain that

$$||\gamma^{(L,\theta_0)} - \gamma_1[\hat{\gamma}_1]|| \leq \frac{CM}{2L} \sum_{k\in\mathbb{Z}\backslash\{1\}} |k-1|^\tau \mathrm{e}^{-2\pi\Delta|k|} \leq \frac{CM}{L} \sup_{x\geq 0}\{\mathrm{e}^{-\pi\Delta x}(x+1)^\tau\} \sum_{k=0}^{\infty} \mathrm{e}^{-\pi\Delta k}.$$

Moreover, we observe that

$$\sup_{x\geq 0}\{\mathrm{e}^{-sx}(x+1)^m\} = \left\{ \begin{array}{cl} 1 & \text{if } s \geq m, \\ (m/(s\mathrm{e}))^m \mathrm{e}^s & \text{if } s < m. \end{array} \right.$$

Finally, taking $A = \frac{MC}{1-\mathrm{e}^{-\pi\Delta}}(1 + (\frac{\tau}{\pi\Delta})^\tau)$ the stated bound follows immediately. □

Remark III.2.7. *Note that in order to guarantee that the projection* (III.6) *is well posed we also need to control the derivative* $(\gamma^{(L,\theta)})'(x)$. *Of course, this can be done modifying slightly the proof of Lemma III.2.6.*

III.2.3 Unfolding a curve of unknown rotation number

Since we are concerned with the computation of θ, the construction presented in the previous section seems useless. Next we show that the method still works —with certain restrictions— if the rotation number θ is unknown, but we have an approximation θ_0.

Proposition III.2.8. *Let us consider a quasi-periodic signal* $z_n = \gamma(n\theta)$ *of rotation number* $\theta \in \mathcal{D}(C,\tau)$. *Assume that* $\gamma : \mathbb{T} \to \mathbb{C}$ *is analytic in the complex strip* B_Δ *and bounded in the closure, with* $M = \sup_{z\in B_\Delta} |\gamma(z)|$. *Suppose that* θ_0 *is an approximation of* θ *and let us denote* $\varepsilon = \theta_0 - \theta$ *and* $K_\varepsilon = \lfloor (2C|\varepsilon|)^{-1/\tau} \rfloor$. *Then, if* $\hat{\gamma}_1 \neq 0$ *and* $K_\varepsilon \geq 1$, *for every* $L \in \mathbb{N}$ *the following estimate holds*

$$\frac{||\gamma^{(L,\theta_0)} - \gamma_1[\hat{\gamma}_1^{(L,\theta_0)}]||}{|\hat{\gamma}_1^{(L,\theta_0)}|} \leq \left| \frac{\sin(\pi\varepsilon)}{\sin(\pi\varepsilon L)} \right| \left(\frac{A}{|\hat{\gamma}_1|} + \frac{2ML}{|\hat{\gamma}_1|} \frac{\mathrm{e}^{-2\pi\Delta(K_\varepsilon - 1)}}{1-\mathrm{e}^{-2\pi\Delta}} \right), \qquad \text{(III.8)}$$

where A *is a constant depending on* M, C, τ *and* Δ.

Proof. Let us consider the sets

$$\begin{aligned} K(\varepsilon) &= \{k \in \mathbb{Z}\backslash\{1\} \,:\, |k-1| \leq K_\varepsilon\}, \\ K^*(\varepsilon) &= \mathbb{Z}\backslash(K(\varepsilon) \cup \{1\}). \end{aligned}$$

Then, if $k \in K(\varepsilon)$ the following bound is satisfied $\forall l \in \mathbb{Z}$

$$|k\theta - \theta_0 - l| \geq |(k-1)\theta - l| - |\varepsilon| \geq \frac{1}{C|k-1|^\tau} - |\varepsilon| \geq \frac{1}{2C|k-1|^\tau}$$

allowing us to control the small divisors

$$|1-\mathrm{e}^{2\pi\mathrm{i}(k\theta-\theta_0)}| \geq \frac{2}{C|k-1|^\tau} \qquad \forall k\in K(\varepsilon). \tag{III.9}$$

Then, from formula (III.5) and recalling that the Fourier coefficients satisfy $|\hat{\gamma}_k| \leq M\mathrm{e}^{-2\pi\Delta|k|}$, we obtain estimates for $\hat{\gamma}_k^{(L,\theta_0)}$. If $k\in K(\varepsilon)$, we use the expression (III.9) to obtain $|\hat{\gamma}_k^{(L,\theta_0)}| \leq MCL^{-1}|k-1|^\tau \mathrm{e}^{-2\pi\Delta|k|}$. On the other hand, for indexes $k\in K^*(\varepsilon)$ we use that $|\hat{\gamma}_k^{(L,\theta_0)}| \leq |\hat{\gamma}_k|$. Therefore, we have to consider the following sums

$$||\gamma^{(L,\theta_0)} - \gamma_1[\hat{\gamma}_1^{(L,\theta_0)}]|| \leq \frac{MC}{L}\sum_{k\in K(\varepsilon)} |k-1|^\tau \mathrm{e}^{-2\pi\Delta|k|} + M\sum_{k\in K^*(\varepsilon)} \mathrm{e}^{-2\pi\Delta|k|}.$$

Now, the sum for $k\in K(\varepsilon)$ is controlled by splitting it into the sets $[-K_\varepsilon+1,0]\cap\mathbb{Z}$ and $[2,K_\varepsilon+1]\cap\mathbb{Z}$. Then, we proceed as in the proof of Lemma III.2.6 obtaining the constant $A=\frac{2MC}{1-\mathrm{e}^{-\pi\Delta}}\left(1+\left(\frac{\tau}{\pi\Delta}\right)^\tau\right)$. Finally, we compute the first Fourier coefficient

$$|\hat{\gamma}_1^{(L,\theta_0)}| = \left|\frac{\hat{\gamma}_1}{L}\right| \left|\frac{1-\mathrm{e}^{-2\pi\mathrm{i}\varepsilon L}}{1-\mathrm{e}^{-2\pi\mathrm{i}\varepsilon}}\right| = \left|\frac{\hat{\gamma}_1}{L}\right| \left|\frac{\sin(\pi\varepsilon L)}{\sin(\pi\varepsilon)}\right|,$$

ending up with estimate (III.8). □

Remark III.2.9. *If we restrict Lemma III.2.8 to those values of* θ_0 *such that the Diophantine condition* (III.9) *is valid* $\forall k\in\mathbb{Z}\backslash\{1\}$, *then we obtain the estimate*

$$\frac{||\gamma^{(L,\theta_0)} - \gamma_1[\hat{\gamma}_1^{(L,\theta_0)}]||}{|\hat{\gamma}_1^{(L,\theta_0)}|} \leq \left|\frac{\sin(\pi\varepsilon)}{\sin(\pi\varepsilon L)}\right| \frac{A}{|\hat{\gamma}_1|}. \tag{III.10}$$

Indeed, if we denote $\mathcal{E}\subset\mathbb{R}$ *the set of values of* ε *such that* $\theta_0=\theta+\varepsilon$ *satisfies estimate* (III.9) *for every* $k\in\mathbb{Z}\backslash\{1\}$, *then for every* ε_0 *sufficiently small the measure of the set* $[-\varepsilon_0,\varepsilon_0]\backslash\mathcal{E}$ *is exponentially small in* ε_0.[2]

Observe that for any fixed $|\varepsilon|>0$, estimate (III.10) depends $|\frac{1}{\varepsilon}|$-periodically on L and also does (III.8) modulo exponentially small terms in $|\varepsilon|$. Since we are interested in the minimization of (III.8), we point out that if $L\simeq|\frac{1}{2\varepsilon}|$, then

$$\frac{||\gamma^{(L,\theta_0)} - \gamma_1[\hat{\gamma}_1^{(L,\theta_0)}]||}{|\hat{\gamma}_1^{(L,\theta_0)}|} = \mathcal{O}(\varepsilon).$$

Hence, the new parameterization is closer to a circle for $|\varepsilon|$ sufficiently small and the projection

$$\begin{array}{rcl} \Gamma^{(L,\theta_0)}\subset\mathbb{C}^* & \longrightarrow & \mathbb{T} \\ z_n^{(L,\theta_0)} & \longmapsto & x_n^{(L,\theta_0)} = \arg(z_n^{(L,\theta_0)})/2\pi, \end{array}$$

[2]These two points of view are analogous to different approaches followed in [JRRV97] and [JS96] to study reducibility of quasi-periodic linear equations. See the discussion in Chapter IV.

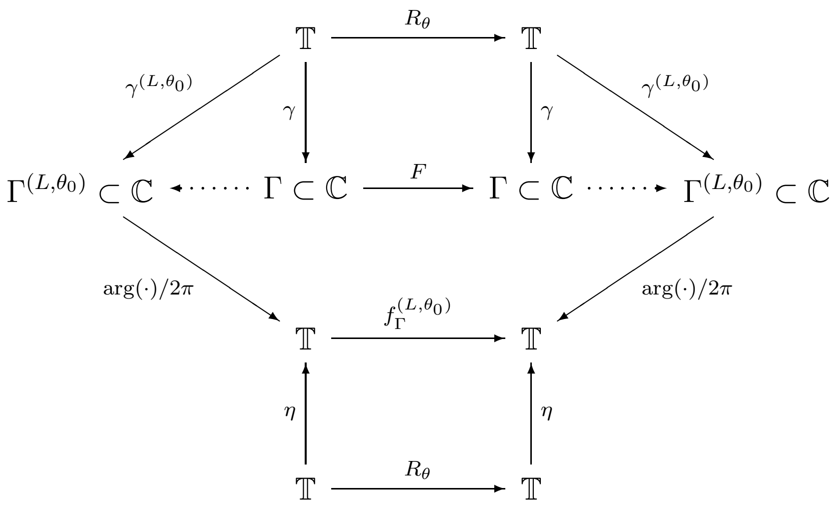

Figure III.2: This diagram summarizes the construction of the analytic circle diffeomorphism $f_\Gamma^{(L,\theta_0)}$ from a folded invariant curve Γ of F of rotation number θ.

induces a well-posed circle diffeomorphism that we denote as $f_\Gamma^{(L,\theta_0)}$. Of course, the regularity of the circle map $f_\Gamma^{(L,\theta_0)}$ follows from the regularity of γ. Hence, we can compute the rotation number θ and derivatives with respect to parameters by applying the methods of Chapter II.

III.2.4 First approximation of the rotation number

The classical *frequency analysis* approach introduced by J. Laskar (see [LFC92]) to obtain an approximation of frequencies of a quasi-periodic signal —here we are considering only one independent frequency— is to look for the frequencies as peaks of the modulus of the Discrete Fourier Transform (DFT) of the studied signal. In this section we translate the elementary ideas used in frequency analysis into the terminology introduced in Section III.2.3.

Let us focus on the iterates $\{z_n^{(L,\theta_0)}\}_{n\in\mathbb{Z}}$ of Definition III.2.2. We observe that they look very similar to the DFT of the signal. We also notice that they can be defined for any θ_0 but the Fourier coefficient $\hat{\gamma}_1^{(L,\theta_0)}$, given explicitly in (III.5), has a local maximum when θ_0 equals θ, and from Proposition III.2.8 we conclude that, for L large enough, the function $\theta_0 \mapsto |z_n^{(L,\theta_0)}|$ has a local maximum for a value of θ_0 close to θ. In general, this phenomena occurs for the tth coefficient if we select θ_0 close to an integer multiple of the rotation number, i.e., $\theta_0 = t\theta + \varepsilon$, $t \in \mathbb{Z}$. The corresponding justification is given by the following proposition (the proof is analogous to that of Proposition III.2.8).

Proposition III.2.10. *Let us consider a quasi-periodic signal $z_n = \gamma(n\theta)$ of rotation number $\theta \in \mathcal{D}(C,\tau)$. Assume that $\gamma : \mathbb{T} \to \mathbb{C}$ is analytic in the complex strip B_Δ and bounded in the closure, with $M = \sup_{z\in B_\Delta} |\gamma(z)|$. Suppose that θ_0 is an approximation of $t\theta$ and let us denote*

$\varepsilon = \theta_0 - t\theta$ and $K_\varepsilon = \lfloor (2C|\varepsilon|)^{-1/\tau} \rfloor$. Then, if $\hat{\gamma}_t \neq 0$ and $K_\varepsilon \geq |t|$, for every $L \in \mathbb{N}$ the following estimate holds

$$\frac{||\gamma^{(L,\theta_0)} - \gamma_t[\hat{\gamma}_t^{(L,\theta_0)}]||}{|\hat{\gamma}_t^{(L,\theta_0)}|} \leq \left| \frac{\sin(\pi\varepsilon)}{\sin(\pi\varepsilon L)} \right| \left(\frac{A}{|\hat{\gamma}_t|} + \frac{2ML}{|\hat{\gamma}_t|} \frac{\mathrm{e}^{-2\pi\Delta(K_\varepsilon - |t|)}}{1 - \mathrm{e}^{-2\pi\Delta}} \right), \tag{III.11}$$

where $A = \frac{2MC}{1-\mathrm{e}^{-\pi\Delta}} \left(|t|^\tau + \left(\frac{\tau}{\pi\Delta}\right)^\tau \right)$.

According with this result and the previous discussion, we summarize the following observations:

- First, we notice that Remark III.2.9 also holds in this context, and so we conclude that this estimate behaves periodically in L for most of the values of θ_0 close to $t\theta$.
- From equation (III.5), we observe that $\hat{\gamma}_t^{(L,\theta_0)} \to \hat{\gamma}_t$ when $\varepsilon \to 0$ and that the modulus $|\hat{\gamma}_t|$ is an upper bound for $|\hat{\gamma}_t^{(L,\theta_0)}|$. Moreover, for L sufficiently large, we obtain a local maximum in the modulus of the iterates $z_n^{(L,\theta_0)}$ for a value of θ_0 close to $t\theta$.
- On the other hand, the estimate (III.11) grows with $|t|$ thus implying that, for fixed L, only low order harmonics can be detected.

The previous discussion gives us a heuristic method for computing an approximation θ_0 of the rotation number θ (and its multiples modulo 1). Basically, we fix L and compute the iterates $z_n^{(L,\theta_0)}$ for different values of θ_0 in order to compute local maxima of the modulus. In particular, if we just study the modulus of the initial iterate $z_0^{(L,\theta_0)}$, we recover the method of [LFC92]. This method is enough for our purposes —we recall that we just look for a rough approximation of the rotation number— but in Remark III.2.12 we explain some refinements that can be performed in this procedure. Thus, from the method of [LFC92] we find a finite number of candidates for the rotation number and we have to decide which one is the generator. Details are given in the next four steps.

Step 1: *Maxima chasing*. First, we fix $L \in \mathbb{N}$ and define the function

$$\begin{array}{ccc} \mathbb{T} & \longrightarrow & \mathbb{R} \\ \theta_0 & \longmapsto & |z_0^{(L,\theta_0)}| = \left|\frac{1}{L}\sum_{m=0}^{L-1} z_m \mathrm{e}^{-2\pi \mathrm{i} m\theta_0}\right| \end{array}. \tag{III.12}$$

We want to obtain values of θ_0 that correspond to maxima of the function (III.12). To this end, let us consider a sample of points $\{\theta_0^i\}_{i=1,\ldots,N}$, where $N \in \mathbb{N}$ and $\theta_0^i \in [0,1]$ (actually, one can reduce the interval if some information about the rotation number is available, say $\theta \in [\theta_{\min}, \theta_{\max}]$). Then, for every pair $\{\theta_0^j, \theta_0^{j+1}\}$, $j = 1, \ldots, N-1$, we compute a local maximum for (III.12) by means of *golden section search* using a tolerance ε_{GSS} (we refer to [PTVF02] for details). Then, we introduce the following terminology:

- $\widetilde{\theta}_0^j$: Maximum obtained from the pair $\{\theta_0^j, \theta_0^{j+1}\}$. Let us observe that this maximum is not necessarily contained in the interval $[\theta_0^j, \theta_0^{j+1}]$.
- n_p: Number of maxima obtained at the end of this step. In order to avoid redundant information, two maxima are considered equivalent if $d_{\mathbb{T}}(\widetilde{\theta}_0^j, \widetilde{\theta}_0^k) \leq 10\varepsilon_{GSS}$ where $d_{\mathbb{T}}$ is the quotient metric induced on the torus.

Step 2: *Maxima selection.* Now we sort the obtained points $\{\widetilde{\theta}_0^i\}_{1,\ldots,n_p}$ according to

$$|z_0^{(L,\widetilde{\theta}_0^i)}| > |z_0^{(L,\widetilde{\theta}_0^j)}| \qquad \text{provided} \qquad i < j$$

At this point, we select the first n_u points just omitting those ones whose corresponding maxima are small when compared with $|z_0^{(L,\widetilde{\theta}_0^1)}|$. In particular, we only take those elements such that

$$|z_0^{(L,\widetilde{\theta}_0^1)}| < \nu |z_0^{(L,\widetilde{\theta}_0^i)}|, \tag{III.13}$$

where $\nu > 1$ is a "selecting factor" (we typically take values of ν between 3 and 6) and we denote by $\{\theta_0^k\}_{k=1,\ldots,n_u}$ the set of numbers thus obtained.

Step 3: *Rotation number selection.* This set $\{\theta_0^k\}_{k=1,\ldots,n_u}$, with $\theta_0^k \in [0, 1]$, corresponds to approximate multiples of the rotation number computed modulo 1. In addition, if $|\hat{\gamma}_1|$ is not too small, there is an element in this set that approximates the rotation number. Notice that for every θ_0^k there exist $m_k, n_k \in \mathbb{Z}$ such that $\theta_0^k \approx m_k\theta + n_k$. This motivates the following definitions

$$k_{ij} = \mathrm{argmin}_{k\in\mathbb{Z}} \{d_{\mathbb{T}}(k\theta_0^i, \theta_0^j)\}, \qquad \kappa_i = \sum_{j=1}^{n_u} |k_{ij}| \tag{III.14}$$

$$d_{ij} = \min_{k\in\mathbb{Z}} \{d_{\mathbb{T}}(k\theta_0^i, \theta_0^j)\}, \qquad \delta_i = \sum_{j=1}^{n_u} d_{ij}. \tag{III.15}$$

Let us observe that if we assume that $\theta_0^i \approx \theta$, then κ_i corresponds to the sum of the order of Fourier terms that allow us to approximate the remaining points θ_0^j. On the other hand, δ_i gives an idea of the error when θ_0^i is selected. If the minimum values of $\{\kappa_i\}_{i=1,\ldots,n_u}$ and $\{\delta_i\}_{i=1,\ldots,n_u}$ correspond to the same index $k \in \{1, \ldots, n_u\}$, then we select $\theta_0 = \theta_0^k$ as an approximation of θ. If they do not coincide but $\delta_k = \min_i\{\delta_i\}$ is small, then we select $\theta_0 = \theta_0^k$. Otherwise we start again from Step 1 using a larger value of L.

Step 4: *Validation and iteration.* Of course, it is recommended to verify that the computations are stable by repeating the process (from Step 1) with larger values of N and L.

Remark III.2.11. *When the frequency of the quasi-periodic signal is not an integer multiple of the "basic frequency"* $1/L$ *associated to the sample interval of the associated DFT in* (III.12)*, there appear in the DFT spurious frequencies, that is, the DFT is different from zero at frequencies not being multiple of the frequency of the function. This is a phenomenon known as* leakage*, that it can be reduced by means of the so-called filter or window functions (see [GMS, LFC92]). Nevertheless, these spurious peaks are smaller than the corresponding harmonic that generates them, so we get rid of them in Step 2 of the described procedure.*

Remark III.2.12. *Finally, we notice that this procedure can be modified in several ways taking into account the ideas introduced in this chapter. For example, when looking for local maxima of function* (III.12) *we can minimize with respect to* θ_0 *the distance of the iterates* $\{z_n^{(L,\theta_0)}\}_{n\in\mathbb{Z}}$ *to be on a circle —see Proposition III.2.10— that can be measured by means of several criteria discussed in Section III.3.1. On the other hand, we can use higher order averages as discussed in Section III.2.6 in order to improve the resolution of the maxima. These refinements may become relevant when dealing with more than one frequency as a possible alternative to the filters mentioned in Remark III.2.11.*

III.2.5 Computation of rotation numbers and derivatives

Let us observe that, in order to approximate derivatives of the rotation number, we require to compute efficiently the quantities $D_\mu^d(f_\mu^n(x))$, i.e., the derivatives with respect to the parameter of the iterates of an orbit. If the family $\mu \mapsto f_\mu$ is known explicitly or it is induced directly by a map on the annulus, several algorithms based on recursive and combinatorial formulas are detailed in Section II.3.3. In the rest of this section we develop recursive formulas to compute these derivatives when the family comes from a general planar map.

Let us consider an analytic map $F : \mathbb{C} \to \mathbb{C}$, with $F = F_1 + \mathrm{i}F_2$, having a Cantor family of invariant curves differentiable in the sense of Whitney, i.e., there exists a family of parameterizations $\mu \in U \mapsto \gamma^\mu$ defined in a Cantor set U such that $\gamma^\mu(\mathbb{T}) = \Gamma^\mu$ and $F(\gamma^\mu(x)) = F(x + \theta(\mu))$, for $\theta(\mu) \in \mathcal{D}$. In the following, we fix a value of the parameter and we omit the dependence on μ in order to simplify the notation and we write $z_n = F^n(z_0)$, for $z_0 \in \Gamma$. As in Section III.2, we consider a curve γ of rotation number θ and we assume that we have an approximation θ_0. Then, we suppose that we can select $L \in \mathbb{N}$ (depending on μ_0 and θ_0) in order to unfold the curve and obtain an orbit of a circle map $f = f^{(L,\theta_0)}$ (or circle correspondence), that has the same rotation number θ, given by

$$x_n^{(L,\theta_0)} = \frac{1}{2\pi} \arctan \frac{\mathrm{Im}\, z_n^{(L,\theta_0)}}{\mathrm{Re}\, z_n^{(L,\theta_0)}}$$

where $z_n^{(L,\theta_0)}$ are given in equation (III.3). The computation of the derivatives of $x_n^{(L,\theta_0)} =$

$f^n(x_0^{(L,\theta_0)})$, that are required to compute $D_\mu S_N^p$, are carried out as

$$D_\mu(x_n^{(L,\theta_0)}) = \frac{1}{2\pi}\frac{\text{Im}\,((D_\mu z_n)^{(L,\theta_0)})\text{Re}\,z_n^{(L)} - \text{Re}\,((D_\mu z_n)^{(L,\theta_0)})\text{Im}\,z_n^{(L,\theta_0)}}{(\text{Re}\,z_n^{(L,\theta_0)})^2 + (\text{Im}\,z_n^{(L,\theta_0)})^2},$$

where

$$(D_\mu z_n)^{(L,\theta_0)} = \frac{1}{L}\sum_{m=n}^{L+n-1} D_\mu z_m \mathrm{e}^{2\pi\mathrm{i}(n-m)\theta_0}.$$

Then, given an averaging order p, we can compute the sums $D_\mu S_N^p$ that allow us to extrapolate $D_\mu\theta$ with an error of order $\mathcal{O}(1/N^p)$. The only point that we need to clarify is the computation of the derivatives $D_\mu z_n$. They are easily obtained by means of the recursive formula

$$\text{Re}(D_\mu z_n) = \frac{\partial F_1}{\partial u}(z_{n-1})\text{Re}(D_\mu z_{n-1}) + \frac{\partial F_1}{\partial v}(z_{n-1})\text{Im}(D_\mu z_{n-1}).$$

and similarly for $\text{Im}(D_\mu z_n)$ replacing F_1 by F_2.

Furthermore, if we consider a family of analytic maps $\alpha \in \Lambda \subset \mathbb{R} \mapsto F_\alpha$ such that for any α we have a family of invariant curves as described before, i.e., there is a parameter μ labeling invariant curves of F_α in a Cantor set U_α. This setting induces a function $(\alpha,\mu) \mapsto \theta(\alpha,\mu)$. Omitting the dependence on (α,μ), let $z_n^{(L,\theta_0)}$ be the unfolded iterates of an orbit that belongs to one of the above curves. Then, we can compute the derivative of θ with respect to α just by averaging the sums of $D_\alpha(x_n^{(L,\theta_0)})$. These iterates are evaluated as explained in the text but using now the recursive formulas

$$\text{Re}(D_\alpha z_n) = \frac{\partial F_1}{\partial \alpha}(z_{n-1}) + \frac{\partial F_1}{\partial u}(z_{n-1})\text{Re}(D_\alpha z_{n-1}) + \frac{\partial F_1}{\partial v}(z_{n-1})\text{Im}(D_\alpha z_{n-1}),$$

and similarly for $\text{Im}(D_\alpha z_n)$ replacing F_1 by F_2.

The generalization of the previous recurrences to compute high order derivatives of the rotation number is straightforward from Leibniz and product rules (as in Remark II.3.4). In addition, expression (II.17) allows us to obtain (pseudo-analytic) asymptotic expansions relating parameters and initial conditions that correspond to curves of prefixed rotation number (see the application to Hénon's map in Section II.5.3).

III.2.6 Higher order unfolding of curves

As it is discussed in Section III.2.3, if we know the rotation number with an error ε small enough, then we can select a number $L \in \mathbb{N}$ (depending on ε) to unfold the curve obtaining a new curve which is a circle with an error of order $\mathcal{O}(\varepsilon) = \mathcal{O}(1/L)$ —we refer to the discussion that follows Proposition III.2.8. Roughly speaking, in the same way that the method of [SV06] accelerates the convergence of the definition in (II.1) to the rotation number from $\mathcal{O}(1/N)$ to

$\mathcal{O}(1/N^{p+1})$, we introduce higher order averages to the iterates $z_n^{(L,\theta_0)}$ to accelerate the convergence of the new curve to a circle. Concretely, by performing averages of order p we improve the rate of convergence from $\mathcal{O}(1/L)$ to $\mathcal{O}(1/L^p)$.

Given $\theta_0 \in \mathbb{R}$, a complex sequence $\{z_n\}_{n\in\mathbb{Z}}$ and a natural number L we introduce the following *recursive sums* of order p

$$\mathcal{S}^1_{L,\theta_0,n} = \sum_{m=n}^{L+n-1} z_m \mathrm{e}^{-2\pi\mathrm{i}m\theta_0}, \qquad \mathcal{S}^p_{L,\theta_0,n} = \sum_{l=1}^{L} \mathcal{S}^{p-1}_{l,\theta_0,n},$$

and the corresponding *averaged sums*

$$\widetilde{\mathcal{S}}^p_{L,\theta_0,n} = \binom{L+p-1}{p}^{-1} \mathcal{S}^p_{L,\theta_0,n}.$$

Definition III.2.13. *Under the above conditions, given $p \in \mathbb{N}$, we define the following iterates for any integer $q \geq p$*

$$z_n^{(2^q,\theta_0,p)} = \left(\sum_{j=0}^{p-1} c_j^{(p-1)} \widetilde{\mathcal{S}}^p_{L_j,\theta_0,n} \right) \mathrm{e}^{2\pi\mathrm{i}n\theta_0}, \tag{III.16}$$

where $L_j = 2^{q-p+j+1}$ and the coefficients $c_j^{(p-1)}$ are given in formula (II.11).

We remark that $z_n^{(2^q,\theta,1)} = z_n^{(2^q,\theta)}$, but that the iterates $z_n^{(L,\theta_0,p)}$ are only defined for L being a power of 2, since they are constructed following the ideas in Algorithms II.2.5 and II.3.2. We see next that if certain non-resonance conditions are fulfilled, these new iterates belong to a quasi-periodic signal such that the corresponding curve approaches a circle improving Proposition III.2.8. For the sake of simplicity, we assume non-resonance conditions as those discussed in Remark III.2.9.

Proposition III.2.14. *Let $\{z_n\}_{n\in\mathbb{Z}}$ be a quasi-periodic signal of rotation number $\theta \in \mathcal{D}$ and averaging order p. Let us consider that $\varepsilon = \theta_0 - \theta$ is small and that θ_0 satisfies*

$$|1 - \mathrm{e}^{2\pi\mathrm{i}(k\theta-\theta_0)}| \geq \frac{2}{C|k-1|^\tau} \qquad \forall k \in \mathbb{Z}\backslash\{1\}, \tag{III.17}$$

for some $C, \tau > 0$. Then, there exists a periodic analytic function $\gamma^{(2^q,\theta_0,p)} : \mathbb{T} \to \mathbb{C}$ such that $z_n^{(2^q,\theta_0,p)} = \gamma^{(2^q,\theta_0,p)}(n\theta)$ and it turns out that

$$||\gamma^{(2^q,\theta_0,p)} - \gamma_1[\hat{\gamma}_1^{(2^q,\theta_0,p)}]|| = \mathcal{O}(2^{-qp}), \tag{III.18}$$

where the function $\gamma_1[\cdot]$ was introduced in Definition III.2.5. Moreover, $\hat{\gamma}_1^{(2^q,\theta_0,p)}$ —the first Fourier coefficient of $\gamma^{(2^q,\theta_0,p)}$— has the following expression

$$\hat{\gamma}_1^{(2^q,\theta_0,p)} = \sum_{j=0}^{p-1} c_j^{(p-1)} \widetilde{\Delta}^p_{L_j,\varepsilon}, \qquad L_j = 2^{q-p+j+1}, \tag{III.19}$$

where $\widetilde{\Delta}^p_{L,\varepsilon}$ *is defined recursively as follows*

$$\Delta^1_{L,\varepsilon} = \hat{\gamma}_1 \frac{1-\mathrm{e}^{-2\pi\mathrm{i}\varepsilon L}}{1-\mathrm{e}^{-2\pi\mathrm{i}\varepsilon}}, \quad \Delta^p_{L,\varepsilon} = \sum_{l=1}^{L} \Delta^{p-1}_{l,\varepsilon}, \quad \widetilde{\Delta}^p_{L,\varepsilon} = \binom{L+p-1}{p}^{-1} \Delta^p_{L,\varepsilon}.$$

In particular, we have that $\lim_{\varepsilon\to 0} \hat{\gamma}_1^{(2^q,\theta_0,p)} = \hat{\gamma}_1$.

Proof. This result is obtained by means of the same arguments used in [SV06]. First, we claim that the following expression follows by induction

$$\widetilde{\mathcal{S}}^p_{L,\theta_0,n} = \widetilde{\Delta}^p_{L,\varepsilon} \mathrm{e}^{-2\pi\mathrm{i}n\varepsilon} + \left(\sum_{l=1}^{p-1} \frac{\mathcal{A}^p_{l,\theta_0}(n\theta)}{(L+p-l)\cdots(L+p-1)} + \mathcal{E}^p_{L,\theta_0}(n\theta) \right) \mathrm{e}^{-2\pi\mathrm{i}n\theta_0}, \qquad \text{(III.20)}$$

where the coefficients $\{\mathcal{A}^p_{l,\theta_0}(n\theta)\}_{l=1,\ldots,p-1}$ are given by

$$\mathcal{A}^p_{l,\theta_0}(n\theta) = (-1)^{l+1}(p-l+1)\cdots p \sum_{k\in\mathbb{Z}\setminus\{1\}} \hat{\gamma}_k \frac{\mathrm{e}^{2\pi\mathrm{i}(l-1)(k\theta-\theta_0)}}{(1-\mathrm{e}^{2\pi\mathrm{i}(k\theta-\theta_0)})^l} \mathrm{e}^{2\pi\mathrm{i}kn\theta}$$

and the remainder is

$$\mathcal{E}^p_{L,\theta_0}(n\theta) = \frac{(-1)^{p+1}p!}{L\cdots(L+p-1)} \sum_{k\in\mathbb{Z}\setminus\{1\}} \hat{\gamma}_k \frac{\mathrm{e}^{2\pi\mathrm{i}(p-1)(k\theta-\theta_0)}(1-\mathrm{e}^{2\pi\mathrm{i}L(k\theta-\theta_0)})}{(1-\mathrm{e}^{2\pi\mathrm{i}(k\theta-\theta_0)})^p} \mathrm{e}^{2\pi\mathrm{i}kn\theta}.$$

For example, we consider the sum for $p=2$

$$\begin{aligned}
\mathcal{S}^2_{L,\theta_0,n} &= \sum_{l=1}^{L} \sum_{m=n}^{l+m-1} z_m \mathrm{e}^{-2\pi\mathrm{i}m\theta_0} = \sum_{l=1}^{L} \sum_{m=n}^{l+m-1} \sum_{k\in\mathbb{Z}} \hat{\gamma}_k \mathrm{e}^{2\pi\mathrm{i}m(k\theta-\theta_0)} = \sum_{l=1}^{L} \sum_{m=n}^{l+m-1} \hat{\gamma}_1 \mathrm{e}^{-2\pi\mathrm{i}m\varepsilon} \\
&+ \sum_{k\in\mathbb{Z}\setminus\{1\}} \hat{\gamma}_k \sum_{l=1}^{L} \sum_{m=n}^{l+m-1} \mathrm{e}^{2\pi\mathrm{i}m(k\theta-\theta_0)} = \Delta^2_{L,\varepsilon} \mathrm{e}^{-2\pi\mathrm{i}n\varepsilon} + L \sum_{k\in\mathbb{Z}\setminus\{1\}} \hat{\gamma}_k \frac{\mathrm{e}^{2\pi\mathrm{i}n(k\theta-\theta_0)}}{1-\mathrm{e}^{2\pi\mathrm{i}(k\theta-\theta_0)}} \\
&- \sum_{k\in\mathbb{Z}\setminus\{1\}} \hat{\gamma}_k \frac{\mathrm{e}^{2\pi\mathrm{i}(n+1)(k\theta-\theta_0)}(1-\mathrm{e}^{2\pi\mathrm{i}L(k\theta-\theta_0)})}{(1-\mathrm{e}^{2\pi\mathrm{i}(k\theta-\theta_0)})^2}.
\end{aligned}$$

Dividing this expression by $L(L+1)/2$, we obtain (III.20) and we proceed inductively to prove the claim. Hence, it is clear that the sequence $z_n^{(2^q,\theta_0,p)}$ in (III.16) corresponds to a quasi-periodic signal since it is a linear combination of quasi-periodic functions.

Using the analyticity assumptions and estimates in (III.17) it turns out that the obtained remainder is of order $\mathcal{E}^p_{L,\theta_0}(n\theta) = \mathcal{O}(1/L^p)$. To extrapolate in this expression using the coefficients (II.11) we require the denominators $(L+p-l)\cdots(L+p-1)$ in (III.20) not to depend on p. To this end, we write

$$\widetilde{\mathcal{S}}^p_{L,\theta_0,n} = \widetilde{\Delta}^p_{L,\varepsilon} \mathrm{e}^{-2\pi\mathrm{i}n\varepsilon} + \left(\sum_{l=1}^{p-1} \frac{\hat{\mathcal{A}}^p_{l,\theta_0}(n\theta)}{L^l} + \hat{\mathcal{E}}^p_{L,\theta_0}(n\theta) \right) \mathrm{e}^{-2\pi\mathrm{i}n\theta_0},$$

by redefining the coefficients $\{\hat{\mathcal{A}}^p_{l,\theta_0}(n\theta)\}_{l=1,\dots,p-1}$, also independent of L, where $\hat{\mathcal{E}}^p_{L,\theta_0}(n\theta)$ differs from $\mathcal{E}^p_{L,\theta_0}(n\theta)$ only by terms of order $\mathcal{O}(1/L^p)$. Hence, we can use Richardson extrapolation using a maximum number of iterates $L = 2^q$ and introduce the corresponding expression into (III.16), thus obtaining

$$z_n^{(2^q,\theta_0,p)} = \left(\sum_{j=0}^{p-1} c_j^{(p-1)} \widetilde{\Delta}^p_{L_j,\varepsilon} \right) \mathrm{e}^{2\pi i n\theta} + \sum_{j=0}^{p-1} c_j^{(p-1)} \hat{\mathcal{E}}^p_{L_j,\theta_0}(n\theta) = \gamma^{(2^q,\theta_0,p)}(n\theta).$$

Therefore, the estimate (III.18) is obtained after observing that the first Fourier coefficient is given by equation (III.19). Finally, the limit $\lim_{\varepsilon\to 0} \hat{\gamma}_1^{(2^q,\theta_0,p)} = \hat{\gamma}_1$ follows from the fact that $\sum_{j=0}^{p-1} c_j^{(p-1)} = 1$. $\square$

III.2.7 Extrapolation of Fourier coefficients

Our goal now is to adapt the previous methodology in order to obtain the Fourier coefficients of a quasi-periodic signal of known rotation number. Let us recall that standard FFT algorithms are based in equidistant samples of points. Since the iterates of a quasi-periodic signal are not distributed in such a way on $\mathbb{T}$, one has to implement a non-equidistant FFT or resort to interpolation of points (see for instance [AdlLP05, dlLP02, OP08]). We avoid this difficulty using the fact that the iterates are equidistant "according with the quasi-periodic dynamics".

We consider a quasi-periodic signal $z_n = \gamma(n\theta)$ of rotation number $\theta \in \mathcal{D}$ as given by Definition III.2.1. Let us observe that we can compute the tth Fourier coefficient, $\hat{\gamma}_t$, as the average of the quasi-periodic signal $z_n \mathrm{e}^{-2\pi i n t\theta}$. For this purpose, we introduce the following recursive sums of order p

$$\mathfrak{S}^0_{N,t} = z_N \mathrm{e}^{-2\pi \mathrm{i} N t\theta}, \qquad \mathfrak{S}^p_{N,t} = \sum_{n=1}^{N} \mathfrak{S}^{p-1}_{n,t}, \tag{III.21}$$

and their corresponding averages

$$\widetilde{\mathfrak{S}}^p_{N,t} = \binom{N+p-1}{p}^{-1} \mathfrak{S}^p_{N,t}.$$

Proposition III.2.15. *For any analytic quasi-periodic signal $z_n = \gamma(n\theta)$ of rotation number $\theta \in \mathcal{D}$ the following expression is satisfied*

$$\widetilde{\mathfrak{S}}^p_{N,t} = \hat{\gamma}_t + \sum_{l=1}^{p-1} \frac{\mathfrak{A}^{p-1}_{t,l}}{N^l} + \mathfrak{E}^p_t(N),$$

where the coefficients $\mathfrak{A}^p_{t,l}$ are independent of N. Furthermore, the remainder $\mathfrak{E}^p_t(N)$ is uniformly bounded by an expression of order $\mathcal{O}(1/N^p)$.

This proposition, that can be proved using analogous arguments as in [SV06], allows us to obtain the following extrapolation scheme to approximate $\hat{\gamma}_t$.

Algorithm III.2.16. *Once an averaging order p is selected, we take $N = 2^q$ iterates of the map, for some $q > p$, and compute the sums $\{\mathfrak{S}^p_{N_j,t}\}_{j=0,\dots,p}$ with $N_j = 2^{q-p+j+1}$. We approximate the tth Fourier coefficient using the formula*

$$\hat{\gamma}_t = \Phi_{q,p,t} + \mathcal{O}(2^{-pq}), \quad \Phi_{q,p,t} = \sum_{j=0}^{p-1} c_j^{(p-1)} \widetilde{\mathfrak{S}}^p_{N_j,t},$$

using the same formula (II.11) *for the coefficients $c_j^{(p-1)}$.*

The extrapolation error of this algorithm can be estimated by means of the following (heuristic) expression

$$|\hat{\gamma}_t - \Phi_{q,p,t}| \leq \frac{\nu}{2^p} |\Phi_{q,p,t} - \Phi_{q-1,p,t}|. \tag{III.22}$$

Remark III.2.17. *As was mentioned in the introduction, the typical approach to compute an invariant curve is to look for it in terms of its Fourier representation. One of the main features of the methodology discussed in this chapter is that we can compute these objects looking for an initial condition on the curve (see Sections II.5 and III.3.3) without computing simultaneously any Fourier expansion or similar approximation. Then, the method discussed above allows us to obtain "a-posteriori" Fourier coefficients of the parameterization from the iterates of the mentioned initial condition.*

Remark III.2.18. *If we want to compute M Fourier coefficients, notice that Algorithm III.2.16 involves a computational cost of order $\mathcal{O}(NM)$ that seems to be deceiving when comparing with FFT methods. Nevertheless, it is clear that the sums* (III.21) *can be also computed as it is standard in FFT since they also satisfy the Danielson-Lanczos Lemma (see for example [PTVF02]), thus obtaining a cost of order $\mathcal{O}(N \log N)$ for computing N coefficients. However, unlike in the other algorithms presented in this chapter, this fast implementation requires to store the iterates of the map.*

Remark III.2.19. *We point out that with Algorithm III.2.16 we can compute isolated coefficients, meanwhile FFT computes simultaneously all the coefficients up to a given order. This can be useful if one is interested only in computing families of coefficients with high precision, as for example it is done in [SV09] —using similar ideas— for coefficients corresponding to Fibonacci numbers.*

III.3 Some numerical illustrations

In this part of the chapter we illustrate several features of the methods discussed in Section III.2. To this end, we have selected three different contexts that we summarize next.

- First, in Section III.3.1, we study invariant curves inside the Siegel domain of a quadratic polynomial. We use this example, where the rotation number is known "a-priori", as a test of the methods. In particular, we show how difficult it is to unfold a given invariant curve as a function of the arithmetic properties of the selected rotation numbers. Furthermore, we introduce two simple criteria to decide if the projection of the iterates of the invariant curve induces a circle map.

- Then, in Section III.3.2, we deal with a toy model obtained by fixing the Fourier coefficients and the rotation number that define a non-embedded quasi-periodic signal. In this example we study the behavior of the unfolded curve $\gamma^{(L,\theta_0)}$ and we check the estimates in Lemma III.2.6 and Proposition III.2.8. Furthermore, in order to simulate uncertainty coming from experimental data we add to this signal a normally distributed random error and we show that the method still provides very accurate results to compute the rotation number of the signal.

- In Section III.3.3 we consider the study of quasi-periodic invariant curves for planar non-twist maps. For the *standard non-twist map*, we apply Algorithm II.2.5 to compute the rotation number in cases where we can compute easily the "lift" of the circle correspondence induced by the direct projection, and also in a very folded curve that we require to unfold. Moreover, we *unfold a shearless* invariant curve comparing the methods in Sections III.2.3 and III.2.6. For Hénon's map, we apply Algorithm II.3.2 to illustrate the computations of derivatives of the rotation number from the "lift" of circle correspondences induced by a family of invariant curves. Finally, we use our methodology to continue numerically a folded (labyrinthic) invariant curve in a more degenerate family of maps.

Let us observe that, since all the recursive sums are evaluated using lifts rather than maps, they turn out to be very large when we increase the order of averaging and the number of iterates. Concretely,

$$S_N^p = \mathcal{O}(N^{p+1}), \qquad D_\mu^d S_N^p = \mathcal{O}(N^{p+1}), \qquad \mathcal{S}_{L,\theta_0,n}^p = \mathcal{O}(L^p), \qquad \mathfrak{S}_N^p = \mathcal{O}(N^p).$$

A natural way to overcome this problem is to do computations by using a representation of real numbers using a computer arithmetic having a large number of decimal digits. Moreover, we have to be very careful with the manipulation of this large numbers to prevent the loss of significant digits (for example, by storing separately integer and decimal parts) and beware not to "saturate" them.

The presented computations have been performed using a C++ compiler and multiple arithmetic (when it is required) has been provided by the routines *double-double* and *quad-double* package of [HLB05], which include a *double-double* data type of approximately 32 decimal digits and a *quadruple-double* data type of approximately 64 digits.

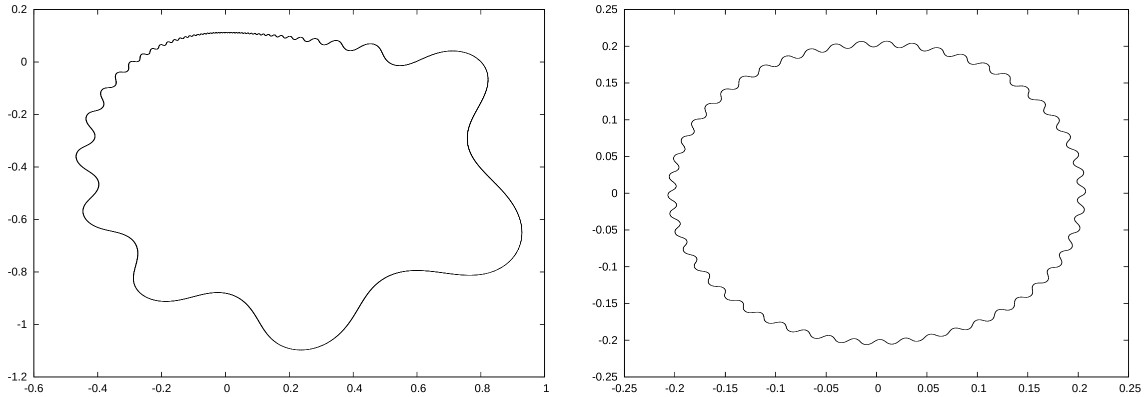

Figure III.3: Left: iterates (in the complex plane) of the point $z_0 = 0.8$ for the quadratic polynomial with $\theta = \theta^{(50)}$. Right: averaged iterates $z_n^{(200,\theta^{(50)})}$ of this curve given by (III.3).

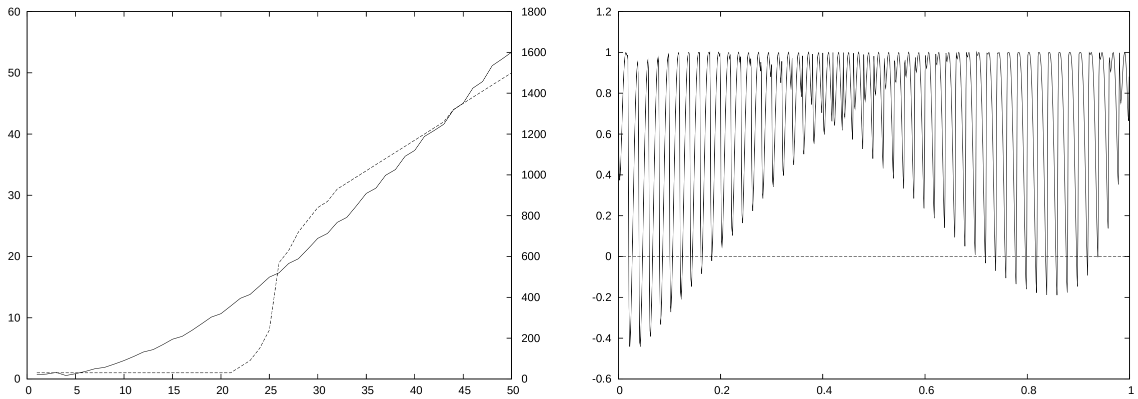

Figure III.4: Unfolding of the invariant curve corresponding to the point $z_0 = 0.8$ for the quadratic polynomial. Left plot: we show, versus the integer s, the value of L for which the curve $\gamma^{(L,\theta^{(s)})}$ is "almost" a circle (solid line, right vertical axis) and the minimum value of L for which the projection defines a circle map (dashed line, left vertical axis). Right plot: we plot function (III.23) for the averaged curve of Figure III.3 (left plot) versus the arc parameter α on $\mathbb{T}$ described in Remark III.3.1.

III.3.1 Siegel domain of a quadratic polynomial

Let $F : U \to \mathbb{C}$ be an analytic map, where $U \subset \mathbb{C}$ is an open set, such that $F(0) = 0$ and $F'(0) = \mathrm{e}^{2\pi\mathrm{i}\theta}$. It is well known that if θ is a Brjuno number, then there exists a conformal isomorphism that conjugates F to a rotation around the origin of angle $2\pi\theta$ (see [Yoc95]). The conjugation determines a maximal set (called *Siegel disk*) which is foliated by invariant curves of rotation number θ.

In particular, we consider the case of the quadratic polynomial $F(z) = \lambda(z - \frac{1}{2}z^2)$,with $\lambda = \mathrm{e}^{2\pi\mathrm{i}\theta}$, for several rotation numbers θ. Concretely, we use $\theta = \theta^{(s)} \in (0, 1)$ which is a zero of $\theta^2 + s\theta - 1 = 0$, with $s \in \mathbb{N}$. It is clear that $\theta^{(s)}$ is a Diophantine number for any

s but with a larger constant C (recall Definition III.2.3) when s increases. Note that $\theta^{(s)} = 1/s - 6/s^3 + \mathcal{O}(1/s^5)$ shows that, for large s, $\theta^{(s)}$ is "close" to a rational number.

Even in this simple example, the direct projection on the angular variable does not always give a diffeomorphism on $\mathbb{T}$. For example, in the left plot of Figure III.3 we show the curve that corresponds to $\theta = \theta^{(50)}$ for the initial condition $z_0 = 0.8$. The right plot of Figure III.3 corresponds to the averaged iterates $z_n^{(200,\theta^{(50)})}$ according to (III.3). As expected, the new curve is closer to a circle centered at the origin.

Our first goal is to emphasize how difficult it becomes to unfold the invariant curve of the point $z_0 = 0.8$ depending on the chosen rotation number $\theta^{(s)}$ (as s increases). To this end, we introduce a criterion to decide when the curve is "close enough" to be a circle. Given a fixed value of $\theta^{(s)}$, we choose $L = 1$ and compute $\{z_n^{(L,\theta^{(s)})}\}_{n=1,\ldots,50000}$ iterates, the mean value of their modulus, and the corresponding standard deviation. Then, if the relative standard deviation is less than 0.5%, we consider that the curve $\gamma^{(L,\theta^{(s)})}$ is close enough to be a circle or we increase L otherwise. The continuous line in the left plot of Figure III.4 (using the vertical axis on the right) shows the obtained value of L versus the integer s that labels the rotation number $\theta^{(s)}$. As expected, we require larger values of L when the rotation number is closer to a rational number.

Let us observe that in practice we do not require to take such large values of L to unfold the curve. Rather than obtaining a circle, we are interested in unfolding the curve in a way that can be projected smoothly into a circle. To this end, we propose a simple criterion to decide if the curve is already unfolded or not. This can be done by computing the changes of sign of the function

$$z \in \gamma(\mathbb{T}) \longmapsto \det(v_t(z), v_r(z)) \in [-1, 1], \tag{III.23}$$

where $v_t(z)$ is the oriented (in the sense of the dynamics) unitary tangent vector of γ at the point z and $v_r(z)$ the corresponding unitary radial vector with respect to the origin. It is clear that if $\det(v_t(z), v_r(z))$ changes sign at some point, then the curve γ is still folded. Moreover, if $\gamma(\mathbb{T})$ is exactly a circle, we have that $\det(v_t(z), v_r(z))$ is constant for all $z \in \gamma(\mathbb{T})$, taking the value -1 or 1, depending if the iterates rotate clockwise or counterclockwise.

As an example, we apply this criterion to the invariant curve shown in the left plot of Figure III.3. The function (III.23) is shown in the right plot of Figure III.4. The horizontal axis in this plot corresponds to the sampling of points on the curve distributed according to Remark III.3.1. Let us observe that this function oscillates due to the folds of the invariant curve, and there are changes of sign since the projection is not well posed. Now, we unfold this invariant curve for different values of $\theta^{(s)}$ looking for the minimum value of L such that $\min(\det(v_t(z), v_r(z))) > 0$. The discontinuous line in the left plot of Figure III.4 (using the vertical axis on the left) shows this value of L versus s, and we observe that it is much smaller than the value of L for which the curve is almost a circle.

Remark III.3.1. *Function* (III.23) *is evaluated by computing the tangent vector $v_t(z)$ using finite differences. To this end, we require a good distribution of points along the curve γ that are obtained using the fact that we know the rotation number (at least approximately). In particular,*

coefficient	value	estimated error	real error
$\hat{\gamma}_{-1}$	$1.4 - 2\mathrm{i}$	$6 \cdot 10^{-40}$	$1 \cdot 10^{-40}$
$\hat{\gamma}_0$	$4.1 + 1.34\mathrm{i}$	$6 \cdot 10^{-41}$	$9 \cdot 10^{-42}$
$\hat{\gamma}_1$	$-2 + 2.412\mathrm{i}$	$5 \cdot 10^{-41}$	$8 \cdot 10^{-42}$
$\hat{\gamma}_2$	$-2.5 - 1.752\mathrm{i}$	$4 \cdot 10^{-40}$	$8 \cdot 10^{-41}$

Table III.1: Fourier coefficients defining expression (III.24) and the numerical error obtained in their approximation using the method of Section III.2.7. The estimated error is obtained by means of formula (III.22).

we fix a number of points M to discretize the curve parameterized by an "arc parameter" $\alpha \in [0, 1)$ on $\mathbb{T}$ defined by the quasi-periodic dynamics. We start with a point $z_0 \in \gamma(\mathbb{T})$ that we identify with the reference parameter $\alpha = 0$. Then, we compute the next iterate of the map and we update the parameter $\alpha \leftarrow \alpha + \theta_0 (\mathrm{mod}\, 1)$. Defining i as the integer part of αM and if $d_{\mathbb{T}}(\alpha, i/M) < 10^{-4}$ we store the iterate in the position ith of an array. We iterate this process till we store M points on $\gamma(\mathbb{T})$. Observe that these computed points are ordered following the dynamics of the curve so we can compute the tangent vector just by finite differences. We will consider that $\min(\det(v_t(z), v_r(z))) > 0$ if the minimum value at the M selected points is positive.

III.3.2 Study of a quasi-periodic signal

We consider a quasi-periodic signal $z_n = \gamma(n\theta)$, with $\theta = (\sqrt{5} - 1)/2 \in \mathcal{D}$, as introduced in Definition III.2.1. Interest is focused in the case where γ is not an embedding, and hence the corresponding orbit is not related to a planar map. In particular, we consider

$$\gamma(x) = \hat{\gamma}_{-1} \mathrm{e}^{-2\pi \mathrm{i} x} + \hat{\gamma}_0 + \hat{\gamma}_1 \mathrm{e}^{2\pi \mathrm{i} x} + \hat{\gamma}_2 \mathrm{e}^{4\pi \mathrm{i} x}, \tag{III.24}$$

where the Fourier coefficients, given in Table III.1, have been selected in such a way that the curve $\gamma(\mathbb{T})$ intersects itself.

First, we show the initial curve $\gamma(\mathbb{T})$ (left plot of Figure III.5) and averaged curves $\gamma^{(L,\theta)}(\mathbb{T})$ (right plot of Figure III.5) corresponding to $L = 2, 3, 4$. Observe that we are using the exact value of the rotation number to compute the averaged iterates given by (III.3). As expected, the new curves are unfolded and become close to a circle.

Since in this problem we know the rotation number, we can perform a simple test of the method presented in Section III.2.7 computing the non-zero Fourier coefficients of the initial curve. To this end we use 64-digit arithmetics (*quadruple-double* data type from [HLB05]) taking an averaging order $p = 9$ and $N = 2^{23}$ iterates of the map. Both the estimated extrapolation error using (III.22) and the real one are shown in Table III.1 observing a very good agreement.

Our goal now is to study the dependence with respect to L of the norms in Lemma III.2.6 and Proposition III.2.8. In Figure III.6 (left) we plot the function

$$L \longmapsto \log_{10} \left(\frac{||\gamma^{(L,\theta)} - \gamma_1[\hat{\gamma}_1]||}{|\hat{\gamma}_1|} \right), \tag{III.25}$$

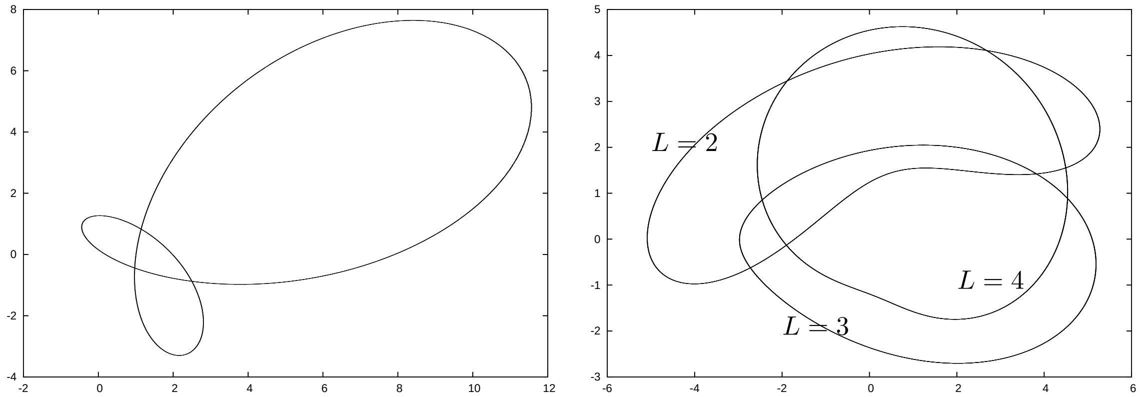

Figure III.5: Left: Curve $\gamma(\mathbb{T})$ in the complex plane corresponding to the parameterization in (III.24). Right: Unfolded curves $\gamma^{(L,\theta)}(\mathbb{T})$, for $L = 2, 3$ and 4, using the known value of the rotation number $\theta = (\sqrt{5} - 1)/2$.

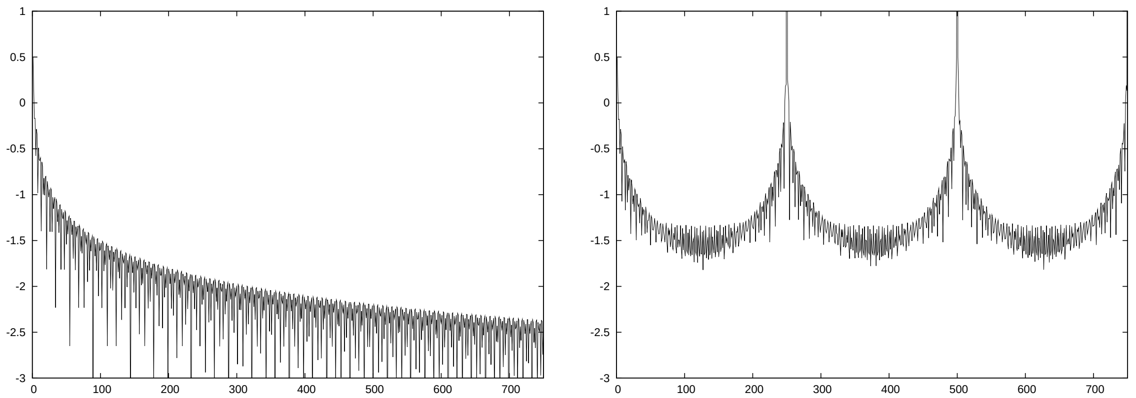

Figure III.6: Left: we plot function (III.25) versus L. Right: we plot function (III.26) versus L using the approximation $\theta_0 = \theta + 1/250$.

that can be evaluated from expression (III.5) using the exact value of the rotation number. We observe that the computed points can be bounded from above in a sharp way by $3.19/L$.

On the other hand, assuming that we only have an approximation θ_0 of the rotation number θ, we want to study the estimate (III.8) of Proposition III.2.8. Concretely, we compute the function

$$L \longmapsto \log_{10}\left(\frac{||\gamma^{(L,\theta_0)} - \gamma_1[\hat{\gamma}_1^{(L,\theta_0)}]||}{|\hat{\gamma}_1^{(L,\theta_0)}|}\right) \qquad \text{(III.26)}$$

using the approximation $\theta_0 = \theta + 1/250$. In the right plot of Figure III.6 we observe that this function is close to be periodic, of period approximately 250, and it reaches a minimum at $L \approx 125 \pmod{250}$ so the bound given in Proposition III.2.8 turns out to be quite good in this case.

It is very interesting to study the effect of a random error in the evaluation of iterates, trying

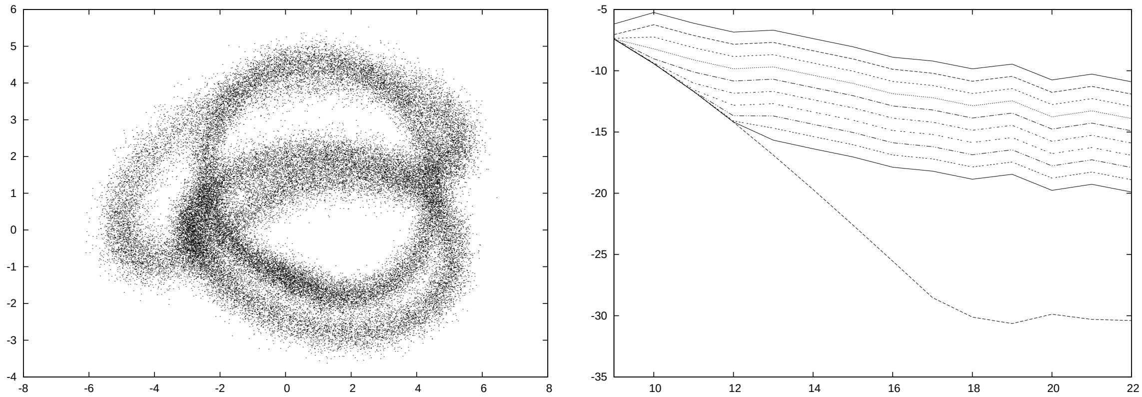

Figure III.7: Effect of a random noise in the quasi-periodic signal (III.24). Left: Unfolded clouds of points (in the complex plane) corresponding to the curves in the right plot of Figure III.5 using $\varepsilon = 0.5$ (see text for details). Right: For different noises, taken as $\varepsilon = 10^{-\delta}$ for $\delta = 1, \ldots, 10, \infty$, we plot $\log_{10}$ of the real error versus q in the computation of the rotation number (using $p = 9$ and 2^q iterates). The data is "unfolded" using $L = 10$ and $\theta_0 = \theta + 1/250$.

to simulate that the source of our quasi-periodic signal is experimental data. Concretely, we consider the iterates $z_n = \gamma(n\theta) + \varepsilon x_n$, where the real and the imaginary parts of the noise x_n are normally distributed with zero mean and unit variance. Of course, the new iterates z_n do not belong to a curve but they are distributed in a cloud around the curve in Figure III.5 (left plot). If we compute the iterates $z_n^{(L,\theta_0)}$, using the approximation $\theta_0 = \theta + 1/250$, then it turns out that we can "unfold" the cloud of points in a similar way. For example, in the left plot of Figure III.7 we show unfolded clouds for an error of size $\varepsilon = 0.5$, using $L = 2, 3, 4$, i.e., the same values that we used in Figure III.5 (right plot).

Now we focus on the effect of this external noise when computing the rotation number θ of the "circle map" thus obtained. The size of the considered noise ranges as $\varepsilon = 10^{-\delta}$ for $\delta = 1, \ldots, 10, \infty$. Although we observe in Figures III.5 (right plot) and III.7 (left plot) that the "projection" is well defined for $L = 4$, in the following computations we take $L = 10$ since for this value the corresponding curve is almost a circle. To the constructed "circle map" we apply Algorithm II.2.5 to refine the numerical computation of the rotation number. As implementation parameters we take an averaging order $p = 9$ and $N = 2^q$ iterates of the map, with $q = 9, \ldots, 22$. The random numbers x_n are generated using the routine `gasdev` from [PTVF02] for generating normal (Gaussian) deviates. Computations are performed using 32-digit arithmetics (*double-double* data type from [HLB05]).

In the right plot of Figure III.7 we show, in $\log_{10}$ scale, the error in the computation of the rotation number with respect to q for different values of ε. Let us observe that for $\varepsilon = 0$ (lowest curve) the extrapolation error is saturated around 10^{-31} for $q \geq 18$. We notice that this error is of the order of the selected arithmetics. The other curves in the right Figure III.7 correspond to increasing values of ε (from bottom to top). Let us remark that in all cases the random error is

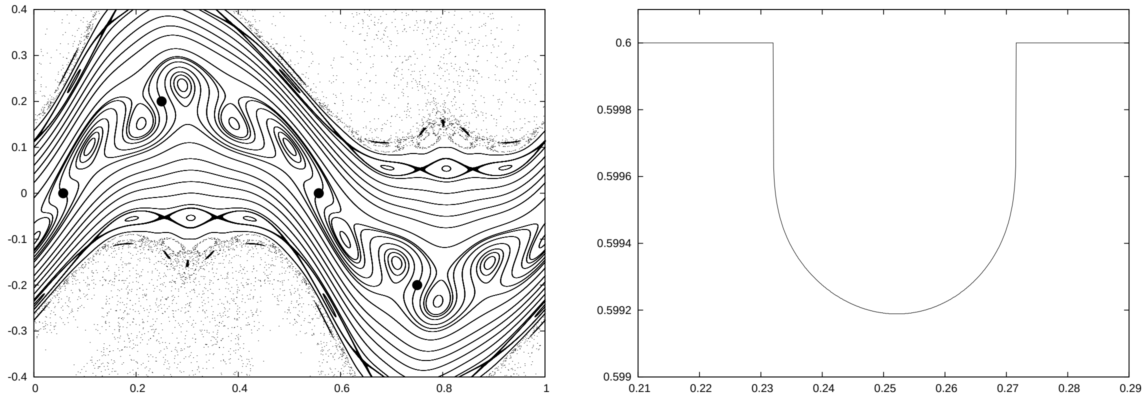

Figure III.8: Left: some meandering curves, in the xy-plane, of the map (III.27) around the shearless invariant curve corresponding to $a = 0.615$ and $b = 0.4$; four dots represent the corresponding indicator points for this curve. Right: Rotation number versus x along the straight line connecting the points $(x, y) = (0.21, 0.15)$ and $(x, y) = (0.29, 0.235)$ in the left plot.

averaged in a very efficient way, and it turns out that the rotation number is approximated with an error of order $\varepsilon \cdot 10^{-10}$.

III.3.3 Study of invariant curves in non-twist maps

Finally, we apply the developed methodology to the study of quasi-periodic invariant curves of non-twist maps. It is known that the Aubry-Mather variational theory for twist maps does not generalize to the non-twist case, but there is an analogue of KAM theory (see for example the works of [DdlL00, Sim98b]). However, the loss of the twist condition introduces different properties than in the twist case, for example the fact that the Birkhoff Graph Theorem does not generalize. A classical mechanism that creates folded invariant curves is called *reconnection*. Reconnection is a global bifurcation of the invariant manifolds of two or more distinct hyperbolic periodic orbits having the same winding number (we refer to [dCNGM96, Sim98b, WAFM05] and references therein for discussion of this bifurcation).

Let us start by considering the family of area preserving non-twist maps given by

$$F_{a,b} : (x, y) \longmapsto (\bar{x}, \bar{y}) = \big(x + a(1 - \bar{y}^2), y - b\sin(2\pi x)\big)\,, \qquad \text{(III.27)}$$

where $(x, y) \in \mathbb{T} \times \mathbb{R}$ are phase space coordinates and a, b parameters. This family is usually called *standard non-twist map* and it is studied as a paradigmatic example of a non-twist family. Although this family is non-generic (it is degenerate in the sense that it contains just one harmonic), it describes the essential features of non-twist systems with a local quadratic extremum in the rotation number.

It is clear that the standard non-twist map violates the twist condition along the curve $y = b\sin(2\pi x)$ which is called *non-monotone curve*. Only orbits with points falling on this curve and orbits with points on both sides of it are affected by the non-twist property. Among these orbits, of special interest is the one that corresponds to an invariant curve having a local extremum in the rotation number, called *shearless invariant curve* γ_S. For the standard non-twist map this curve is characterized by the fact that, when it exists, it must contain the points

$$(x_{\pm}^{(0)}, y_{\pm}^{(0)}) = \left(\pm\frac{1}{4}, \pm\frac{b}{2}\right) \qquad (x_{\pm}^{(1)}, y_{\pm}^{(1)}) = \left(\frac{a}{2} \pm \frac{1}{4}, 0\right),$$

that are called *indicator points* (see [SA98]). These points are used extensively in the literature to study the breakdown of shearless invariant curves (we refer for example to [AdlLP05, WAFM05]).

In the left plot of Figure III.8 we show some invariant curves close to a reconnection scenario for $a = 0.615$ and $b = 0.4$. We observe some *meandering curves*, i.e., curves that are folded around periodic orbits in such a way that they are not graphs over x. In addition, we plot the four indicator points in order to identify the shearless invariant curve. Actually, the invariant curve that we used as an illustration in Section III.2.1 (see Figure III.1) is precisely this shearless curve in the complex variable $z = \mathrm{e}^y \mathrm{e}^{\mathrm{i}x}$.

First, we focus on this shearless curve computing its rotation number by applying the extrapolation method of Algorithm II.2.5 to the circle correspondence obtained by direct projection (see the discussion in Sections III.1 and III.2.1) on the angular variable x. Since the folds of this example are relatively small and the rotation number is quite big —it is close to 0.6, i.e., the winding number of the nearby periodic orbits— this direct projection allows us to compute numerically the lift of this circle correspondence without unfolding the curve. In Table III.2 we give the estimated extrapolation error, by means of formula (II.13), in the computation of the rotation number of γ_S, for different values of the extrapolation order p and number of iterates 2^q. Computations are performed using 32-digit arithmetics (*double-double* data type from [HLB05]). Let us observe that the extrapolation method allows us to obtain a very good approximation of the rotation number with a relative small number of iterates, in contrast with $p = 0$ which corresponds to the definition of the rotation number —let us mention that some recent works like [WAFM05] use the definition to approximate the rotation number. According to our estimates, the best computed approximation of the rotation number turns out to be

$$\theta \simeq \Theta_{21,7} = 0.59918902772269558576430971159247.$$

In addition, we compute the rotation number of meandering curves in the left plot of Figure III.8 using an averaging order $p = 7$ and 2^{21} iterates of the map. In Figure III.8 (right plot) we show the rotation number profile in this reconnection scenario. Concretely, we compute the rotation number for the orbits corresponding to 1000 points along an straight line connecting $(x, y) = (0.21, 0.15)$ and $(x, y) = (0.29, 0.235)$, that are close to the elliptic periodic orbits of winding number $3/5$. As far as the estimated extrapolation error is concerned, 93% of the points have an error less than 10^{-26} and 98% of the points have an error less than 10^{-20}. The minimum value in the profile corresponds to the rotation number of γ_S, and we observe the loss of uniqueness of invariant curves when the twist condition fails.

p	$q=10$	$q=11$	$q=12$	$q=13$	$q=14$	$q=15$	$q=16$	$q=17$	$q=18$	$q=19$	$q=20$	$q=21$
0	1.4e-03	7.3e-04	1.9e-04	1.9e-04	8.1e-05	3.6e-05	1.1e-05	1.0e-05	5.0e-06	2.5e-06	1.0e-06	3.4e-07
1	7.9e-06	5.5e-06	2.6e-07	8.1e-07	1.4e-07	6.9e-08	3.7e-09	1.9e-09	1.2e-09	2.5e-11	8.8e-12	6.4e-12
2	3.8e-07	8.6e-09	3.0e-08	4.5e-08	4.6e-09	2.1e-10	3.7e-11	1.4e-12	8.5e-13	1.9e-14	4.1e-16	1.3e-16
3	1.6e-05	9.5e-11	1.0e-09	8.5e-10	7.0e-12	1.2e-11	5.0e-13	6.5e-14	2.3e-16	1.2e-16	1.6e-17	7.1e-22
4	1.6e-05	1.2e-07	1.9e-09	5.2e-11	2.2e-12	8.9e-13	2.0e-14	2.8e-16	1.0e-17	1.9e-19	1.2e-20	8.8e-23
5	4.9e-06	7.0e-07	2.1e-08	5.0e-12	6.9e-14	2.2e-14	3.3e-17	1.7e-17	2.0e-19	6.0e-21	4.0e-24	8.2e-25
6	2.6e-06	1.0e-06	2.8e-08	4.5e-11	4.5e-13	3.5e-15	3.6e-18	1.9e-18	9.3e-21	4.5e-23	3.3e-25	2.0e-27
7	4.7e-06	8.8e-07	8.5e-09	4.5e-10	2.7e-12	7.7e-16	1.6e-18	6.9e-20	2.1e-23	2.5e-24	8.6e-27	5.9e-29

Table III.2: Estimated extrapolation error, using formula (II.13), in the approximation of the rotation number of the shearless curve of the map (III.27) that corresponds to $a = 0.615$ and $b = 0.4$.

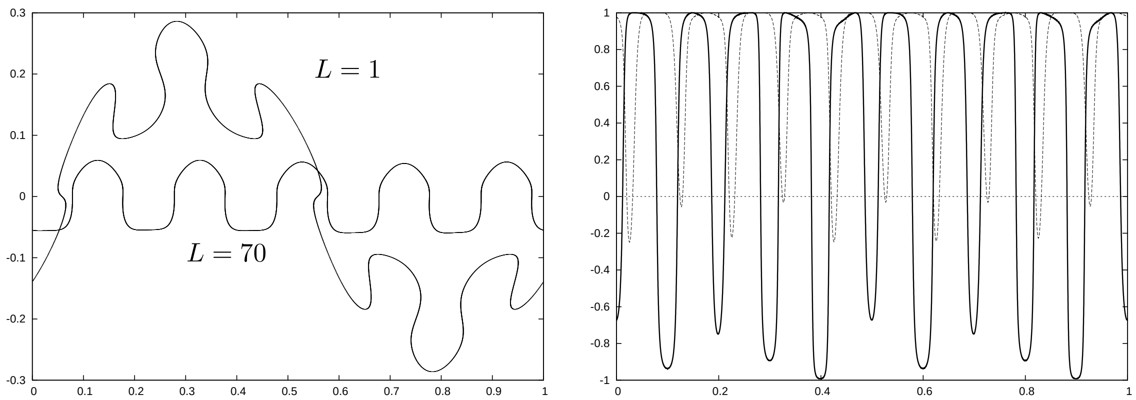

Figure III.9: Left: we plot in the xy-plane the shearless curve $\gamma_S = \gamma^{(1,\theta_0)}$ in Figure III.8 and the averaged curve $\gamma^{(70,\theta_0)}$, where $\theta_0 = \Theta_{21,7}$. Right: we plot function (III.23) on the previous curves versus the arc parameter α on $\mathbb{T}$ described in Remark III.3.1.

Remark III.3.2. *Of course, the points in the left plot of Figure III.8 to which we assign a rational rotation number (the profile is locally constant) cannot belong to an invariant curve. These points correspond to "secondary invariant curves" or "islands", which are invariant curves of a suitable power of the map, that appear close to the elliptic periodic orbits. Thus, for a point on these islands, what we obtain is the "winding number" of the periodic orbit in the middle of the island. We refer to discussions in [SV06] and Chapter II.*

Now, let us illustrate the methodology of Section III.2 in order to unfold the shearless invariant curve γ_S. To this end, we complexify phase space by means of the change of variables $z = e^y e^{2\pi i x}$ and compute the new quasi-periodic signal $\{z_n^{(L,\theta_0)}\}$ using the approximation $\theta_0 = \Theta_{21,7}$. In Figure III.9 (left plot) we show the original curve γ_S, corresponding to $L = 1$, together with the curve $\gamma^{(70,\theta_0)}$ that is less folded but its projection is not well defined yet. In Figure III.9 (right plot) we show function (III.23) for γ_S and $\gamma^{(70,\theta_0)}$. Actually, the projection into a circle is well posed for $L = 75$ and the curve is very close to be a circle —the minimum of function (III.23) is $\simeq 0.999897$— for $L = 240$.

Furthermore, we try to unfold γ_S using higher order averages to accelerate the convergence to a circle. To this end, we still fix $\theta_0 = \Theta_{21,7}$ and apply the higher order method explained

q	$p=1$	$p=2$	$p=3$	$p=4$	$p=5$
6	2.8e-02 ($-$)	5.6e-02 ($-$)	7.1e-02 ($-$)	7.5e-02 ($-$)	7.4e-02 ($-$)
7	1.1e-01 ($+$)	1.5e-01 ($=$)	1.2e-01 ($=$)	7.7e-02 ($-$)	4.2e-02 ($-$)
8	4.0e-02 ($+$)	5.6e-02 ($+$)	9.6e-02 ($=$)	1.4e-01 ($=$)	1.4e-01 ($=$)
9	1.8e-03 ($+$)	4.4e-03 ($+$)	3.4e-02 ($+$)	6.7e-02 ($+$)	6.4e-02 ($=$)
10	2.8e-03 ($+$)	3.6e-05 ($+$)	7.8e-03 ($+$)	1.2e-02 ($+$)	1.1e-02 ($+$)
11	2.0e-03 ($+$)	3.5e-05 ($+$)	1.0e-06 ($+$)	2.6e-04 ($+$)	1.9e-03 ($+$)
12	3.0e-03 ($+$)	2.1e-04 ($+$)	7.6e-06 ($+$)	2.4e-06 ($+$)	1.3e-04 ($+$)
13	9.6e-04 ($+$)	1.1e-04 ($+$)	3.0e-06 ($+$)	9.8e-08 ($+$)	1.5e-08 ($+$)
14	8.0e-04 ($+$)	2.0e-05 ($+$)	3.8e-07 ($+$)	7.7e-08 ($+$)	4.5e-09 ($+$)
15	3.0e-04 ($+$)	3.5e-06 ($+$)	1.7e-07 ($+$)	1.9e-08 ($+$)	8.0e-10 ($+$)
16	2.0e-05 ($+$)	8.4e-07 ($+$)	1.9e-08 ($+$)	9.3e-10 ($+$)	1.8e-11 ($+$)

Table III.3: Estimated distance to be a circle of the higher order averaged curve $\gamma^{(2^q,\theta_0,p)}$, where $\theta_0 = \Theta_{21,7}$, for the shearless curve γ_S in Figure III.8. The meaning of the symbols is the following: ($-$) if the curve is still folded, ($=$) if the curve is unfolded but not close enough to a circle and ($+$) if the curve is close to be a circle.

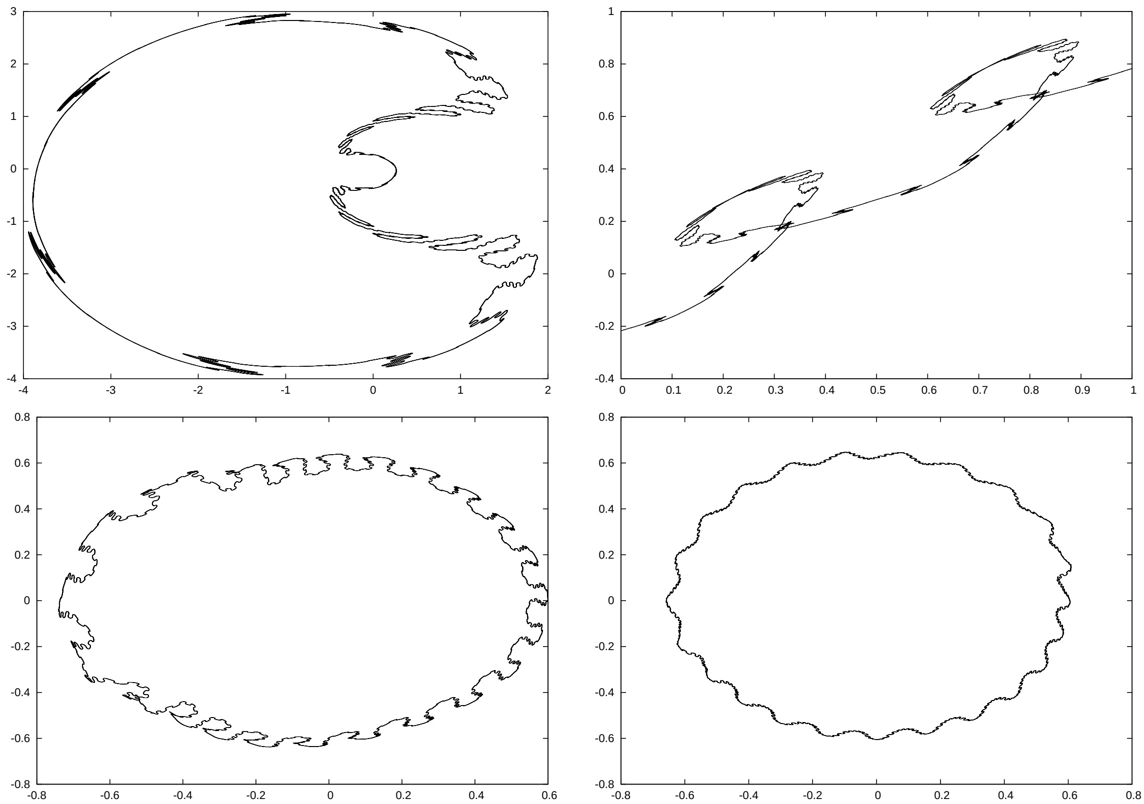

Figure III.10: Top-left: Higher order meandering curve for the standard non-twist map in the complexified phase space using $z = \mathrm{e}^{\sqrt{2\pi|a|}y}\mathrm{e}^{2\pi\mathrm{i}x}$. Top-right: Circle correspondence obtained from the direct projection —defined in (III.1)— of the iterates of the curve in top-left plot. Bottom: We show the curves $\gamma^{(L,\theta_0)}$, where $\theta_0 = 0.0429853252$, for $L = 50$ (left) and $L = 150$ (right) in the complex plane.

in Section III.2.6 to unfold the curve, using different values of the extrapolation order p and

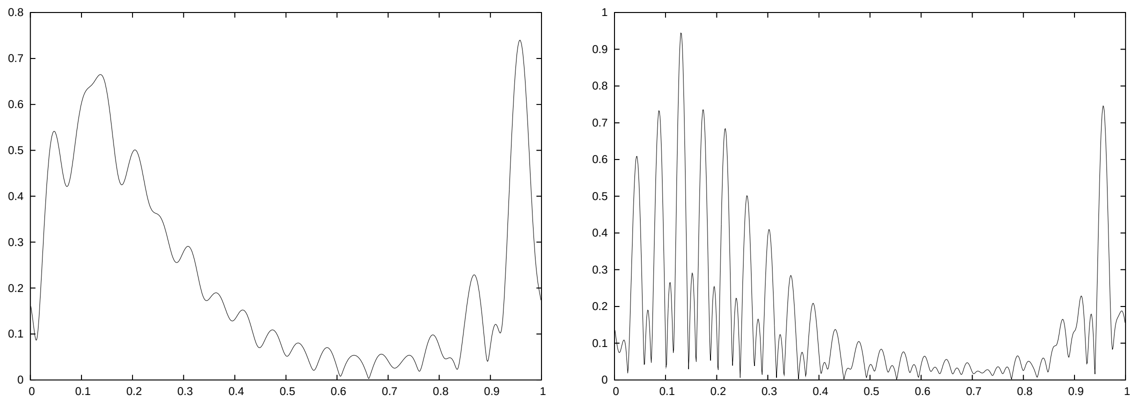

Figure III.11: We plot the function $\theta_0 \mapsto |z_0^{(L,\theta_0)}|$ using the values $L = 30$ (left) and 60 (right) for the meandering curve in Figure III.10. The real value of the rotation number is around 0.04332244074906551 (see the text for details).

number of iterates $L = 2^q$. In Table III.3 we present the estimated error when we compare the new curve with a circle. Since we have constructed a sequence of curves tending to a circle —up to small error— we estimate the distance of these curves to be a circle just looking at the number of digits that coincide for a point on these curves when we increase the number of iterates from 2^{q-1} to 2^q, this is, we use the formula $|z_0^{(2^q,\theta_0,p)} - z_0^{(2^{q-1},\theta_0,p)}|$, where $\theta_0 = \Theta_{21,7}$. The adjacent symbol in this table indicates if the projection of the invariant curve induces a circle map, according to the criterion of the sign of function (III.23) —as explained in Remark III.3.1. We put the symbol $(-)$ if the minimum of function (III.23) is negative, we put $(=)$ if this minimum is positive but less than 0.99 and we put $(+)$ if this minimum is between 0.99 and 1. As expected, for $p = 1$ and $q = 6$ the curve is still folded since $z_n^{(2^q,\theta_0)} = z_n^{(2^q,\theta_0,1)}$ and $2^6 = 64 < 75$. Let us observe that when we increase the order of averaging we require more iterates in order to appreciate an improvement in the extrapolation. Nevertheless, after this transition the method turns out to be much more efficient (for example, see the row for $q = 16$).

A further step is to consider the case of the so-called *higher order meanderings* that appear due to reconnections involving periodic orbits in a neighborhood of a meandering curve. The considered example is selected from [Sim98b] and we refer there for a constructive explanation. Concretely, we consider the values $a = -0.071963192$ and $b = -0.44614508325727$ and the curve that corresponds to the initial condition $(x_0, y_0) = (0, -2.00377736103447)$. In Figure III.10 (left plot) we show this higher order meandering curve in the complex plane by means of the change of variables $z = \mathrm{e}^{\sqrt{2\pi|a|}y}\mathrm{e}^{2\pi \mathrm{i} x}$.

Let us observe that the invariant curve of this example is folded in a very wild way. Actually, in the top-right plot of Figure III.10 we show the "lift" to $\mathbb{R}$ of the circle correspondence that we obtain by means of the direct projection on the angular variable x. This plot has been obtained projecting iterates of an orbit of the curve, and we observe that it is not easy to compute numerically this lift just looking at isolated iterates of the circle correspondence —without any "a-

$L = 30$		$L = 60$		$L = 90$		$L = 120$		$L = 150$	
0.9577778096	(0.74)	0.1303407362	(0.95)	0.1302228135	(0.86)	0.1302697604	(0.82)	0.1301364744	(0.84)
0.1371713373	(0.67)	0.9562939918	(0.75)	0.9565703718	(0.77)	0.9565427675	(0.72)	0.9569018161	(0.70)
0.0465959056	(0.54)	0.1734756047	(0.74)	0.0870283875	(0.68)	0.0870160838	(0.62)	0.0865975344	(0.65)
0.2046548291	(0.50)	0.0872525690	(0.73)	0.1733546296	(0.64)	**0.0437697415**	(0.59)	**0.0429853252**	(0.62)
0.3084897181	(0.29)	0.2166379341	(0.68)	**0.0437231352**	(0.61)	0.1737224972	(0.58)	0.1736243983	(0.61)
0.8683260095	(0.23)	**0.0435679422**	(0.61)	0.2169984090	(0.58)	0.2172096615	(0.54)	0.2171559900	(0.57)
0.3631744037	(0.19)	0.2593740449	(0.50)	0.2601564576	(0.42)	0.2604187823	(0.38)	0.2606176626	(0.40)
0.4154453760	(0.15)	0.3024868348	(0.41)	0.3037390964	(0.34)	0.3041163591	(0.32)	0.3041431703	(0.34)

Table III.4: Relevant maxima of the function $\theta_0 \in [0,1] \mapsto |z_0^{(L,\theta_0)}|$ for different values of L corresponding to the invariant curve in Figure III.10 (top-left plot). In parentheses we show the value of the function at the local maxima. We write in bold the value of the maximum that approximates the rotation number of the curve.

i	k_{i1}	k_{i2}	k_{i3}	k_{i4}	k_{i5}	k_{i6}	k_{i7}	k_{i8}	κ_i
1	1	-8	9	-7	-6	8	2	10	51
2	-3	1	-4	-2	-5	-1	-6	-7	29
3	-5	-6	1	-11	7	6	13	-4	53
4	-10	-12	2	1	-9	-11	3	-8	56
5	-4	9	10	5	1	-9	-8	6	52
6	3	-1	4	2	5	1	6	7	29
7	-11	-4	-7	8	-3	4	1	5	43
8	7	-10	-6	-3	4	10	-9	1	50

Table III.5: Indicators k_{ij} and κ_i corresponding to the step "rotation number selection" described in Section III.2.4. We use $L = 60$.

priori" information on the rotation number. Therefore, we apply Laskar's method of frequency analysis (implemented as described in Section III.2.4) to obtain a sufficiently good approximation of the rotation number in order to unfold the curve. *Maxima chasing* is performed looking for local maxima of $\theta_0 \mapsto |z_0^{(L,\theta_0)}|$ in the interval $[0,1]$. These maxima are obtained using a partition of 500 points of this interval and asking for a tolerance $\varepsilon_{GSS} = 10^{-10}$ in the golden section search. In Table III.4 we show the relevant maxima, using $\nu = 5$ in equation (III.13), for several values of L. Moreover, in Figure III.11 we plot the function $\theta_0 \mapsto |z_0^{(L,\theta_0)}|$, for $L = 30$ (left plot) and $L = 60$ (right plot). Of course, when we increase L we observe a larger number of maxima but the width of the peaks is reduced so their approximation is improved.

As explained in Section III.2.4, we assume that one of these maxima approximates the rotation number of the curve in such a way that the rest are multiples of it (modulo 1). To select which maximum approximates the rotation number, we compute the indicators k_{ij}, κ_i, d_{ij} and δ_i given in equations (III.14) and (III.15). For example, the values corresponding to $L = 60$ are given in Tables III.5 and III.6. Let us point out that the maximum corresponding to the ith row in these tables can be read in the ith row of Table III.4. If we select the 2nd or the 6th maxima as the approximation θ_0 of the rotation number, then the remaining peaks are approximated —with the error shown in Table III.6— as multiples $k\theta_0$ (modulo 1) using smaller values for k than if we make any other choice (see Table III.5). These two peaks correspond to approximations of $1 - \theta$ and θ, and we select the 6th one as an approximation of the rotation number since is follows the positive orientation of the dynamics. We write this peak in bold in Table III.4.

i	d_{i1}	d_{i2}	d_{i3}	d_{i4}	d_{i5}	d_{i6}	d_{i7}	d_{i8}	δ_i
1	0.00e+00	9.80e-04	4.00e-04	3.60e-04	1.30e-03	8.40e-04	1.30e-03	9.20e-04	6.10e-03
2	7.70e-04	0.00e+00	1.30e-03	1.50e-04	1.80e-03	1.30e-04	2.80e-03	3.40e-03	1.00e-02
3	2.20e-03	2.80e-03	0.00e+00	4.50e-03	2.30e-03	2.70e-03	4.10e-03	3.60e-03	2.20e-02
4	2.80e-03	3.30e-03	1.00e-03	0.00e+00	1.90e-03	3.30e-03	2.30e-03	5.00e-04	1.50e-02
5	3.10e-03	6.50e-03	7.10e-03	4.00e-03	0.00e+00	6.60e-03	7.50e-03	2.60e-03	3.70e-02
6	3.60e-04	1.30e-04	7.90e-04	1.10e-04	1.20e-03	0.00e+00	2.00e-03	2.40e-03	7.10e-03
7	1.60e-02	6.20e-03	1.00e-02	1.20e-02	5.20e-03	6.00e-03	0.00e+00	5.60e-03	6.20e-02
8	1.20e-02	1.80e-02	1.10e-02	5.20e-03	6.60e-03	1.80e-02	1.80e-02	0.00e+00	9.20e-02

Table III.6: Indicators d_{ij} and δ_i corresponding to the step "rotation number selection" described in Section III.2.4. We use $L = 60$.

Let us refine the approximation of the rotation number from the 6th peak. In Figure III.10 we plot $\gamma^{(L,\theta_0)}$ using the approximation $\theta_0 = 0.0429853252$ for $L = 50$ (bottom-left plot) and $L = 150$ (bottom-right plot). We observe that for $L = 50$ the curve is still very folded but if we take $L = 150$, even thought there are still some harmonics that fold the curve, we are close to a circle centered at the origin and the projection gives us a circle correspondence that we can "lift" to $\mathbb{R}$ easily, since the size of the folds is small when compared with θ_0. Finally, we compute the rotation number of this invariant curve from the circle correspondence obtained by means of this unfolding procedure. We apply Algorithm II.2.5, using $p = 7$ and $q = 21$, to the iterates $x_n = \arg(z_n^{(150,\theta_0)})/2\pi$ thus obtaining the approximation $\Theta_{7,21} = 0.04332244074906551$ with an estimated error $1.7 \cdot 10^{-13}$.

We observe that the computation of the rotation number does not work as well as in previous examples, but an error of order 10^{-13} is very satisfying in this context since a huge number of Fourier coefficients is required to approximate the curve with this error and the rotation number is close to resonance (see also the discussion regarding Hénon's map example). To show this, let us compute the Fourier coefficients corresponding to $|k| \leq 750$ by means of Algorithm III.2.16 using $p = 7$, $q = 21$ and $\theta \simeq \Theta_{21,7}$. Computations are performed using 32-digit arithmetics (*double-double* data type from [HLB05]). The modulus of the obtained values $\Phi_{21,7,k}$ are shown in the left plot of Figure III.12. The extrapolation error, estimated using (III.22), typically ranges between 10^{-10} and 10^{-8}. As expected, the decay of these coefficients is very mild and we point out that coefficients for $|k| \simeq 750$ are still of order 10^{-5}. It is interesting to compare this behavior to that corresponding to the shearless curve in Figure III.8 since the extrapolation methods can be applied successfully to this case. For this curve the Fourier coefficients —computed using the same implementation parameters— decay much faster (see the right plot in Figure III.12). In this case, the estimated extrapolation error typically ranges between 10^{-16} and 10^{-24}, and Fourier coefficients for $|k| \geq 400$ are so small that we cannot compute any significant digit (we detect this fact because $|\Phi_{21,7,k}|$ is of the same order as the extrapolation error). A final remark is that in this plot we observe an increase of the size of Fourier coefficients around $|k| \simeq 250$ and after that they decay again at the same rate. The rotation number of this curve has the convergent $145/242$, so the small divisor for the corresponding Fourier coefficient turns out to be very small. Precisely we observe that $|\Phi_{21,7,243}| \simeq 214|\Phi_{21,7,242}|$.

In the next example we want to illustrate the computation of derivatives of the rotation

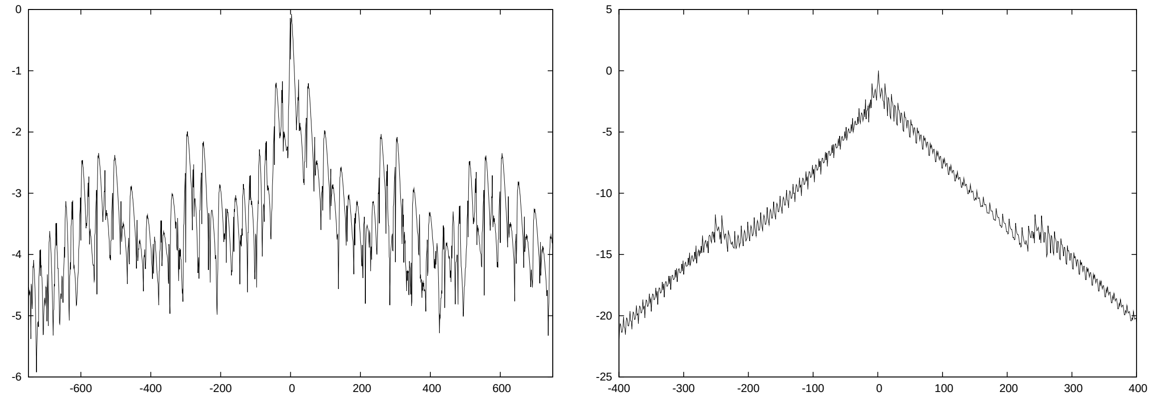

Figure III.12: We plot $\log_{10}|\Phi_{21,7,k}|$ versus k corresponding to the approximated Fourier coefficients of two studied invariant curves. Left: Higher order meandering curve in Figure III.10 (top-left plot). Right: Shearless invariant curve in Figure III.8 (left plot).

number by applying Algorithm II.3.2 to a circle correspondence that we can "lift" numerically to $\mathbb{R}$. Moreover, we also want to stress how extrapolation methods are less accurate when the rotation number is a Diophantine number "close to a rational" —thus having a large constant C in (III.7). Therefore, we select a problem very close to a resonance. Comparing with previous examples, we point out that the shearless curve in Figure III.8 is not too resonant —actually the continued fraction of the rotation number in that case is $[0, 1, 1, 2, 48, 1, \ldots]$. Meanwhile the case of higher order meandering in Figure III.10 is more resonant but the curve extremely complicated. Let us consider the well-known Hénon family, which is a paradigmatic example since it appears generically in the study of a saddle-node bifurcation. This family can be written as

$$H_\alpha : \begin{pmatrix} u \\ v \end{pmatrix} \longmapsto \begin{pmatrix} \cos(2\pi\alpha) & -\sin(2\pi\alpha) \\ \sin(2\pi\alpha) & \cos(2\pi\alpha) \end{pmatrix} \begin{pmatrix} u \\ v - u^2 \end{pmatrix}.$$

It is not difficult to check that we can perform a close to the identity change of variables to guarantee the twist condition close to the origin, except for the values $\alpha = 1/3, 2/3$. Then, for values of α close to $1/3$ and $2/3$, reconnection takes places and meandering phenomena arises, i.e., there are folded invariant curves. Next we want to illustrate the computation of the derivatives of the rotation number for α close to $1/3$.

In the left plot of Figure III.13 we show meandering curves for $\alpha = 0.299544$. We use the direct projection and we apply Algorithm II.3.2 to the "lift" of the circle correspondence thus obtained, in order to compute the derivative of the rotation number with respect to the initial condition u_0, for 6000 points of the form $(u_0, 0.15)$ along the dotted line in the figure. Computations are performed taking $p = 8$, $q = 23$, and using 32-digit arithmetics (*double-double* data type from [HLB05]). The corresponding profile is shown in the right plot of Figure III.13, and we observe that the sign of the derivative changes when we pass from one twist zone to

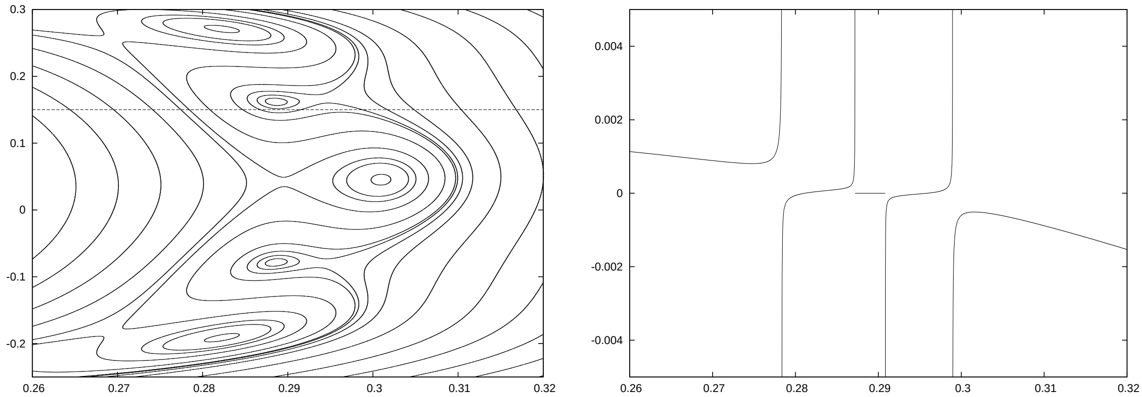

Figure III.13: Left: Phase space in the uv-plane of Hénon's map for $\alpha = 0.299544$ showing meandering curves close to periodic orbits of period $3/10$. Right: we plot the derivative of the rotation number $u \mapsto D_u\rho$ along the straight line in the left plot.

another. The isolated points where the derivative vanishes correspond to the *shearless invariant curve* while the points where it is locally constant corresponds to the chain of islands (secondary tori) around the elliptic periodic points (see Remark III.3.2). We want to stress that since we are very close to a resonance, the convergence of the computations in this example is not as good as in the far-from-resonance case. For example, for $u_0 = 0.28$ we obtain the approximation $\Theta_{23,8} = 0.299999020519$ of the rotation number with an estimated error, using (II.13), of order 10^{-12} and the derivative $\Theta^1_{23,8,7} = -6.027735852 \cdot 10^{-5}$ with an estimated error, using (II.19), of order 10^{-10}. Actually, the continued fraction expansion of $\Theta_{23,8}$ is given by $[0, 3, 2, 1, 10203, 2, \ldots]$ which is close to $3/10$ (winding number of the periodic orbit). It is interesting to compare the error of order 10^{-12} in the computation of the rotation number with the error obtained when dealing with good Diophantine numbers, that are typically of order 10^{-30} for the used implementation parameters. We refer to several comments given in [LV08] and in Section II.4.1.

We remark that all the previous examples considered in this section contain only one harmonic or are written as perturbations of maps that have strong twist behavior. For this reason we have not shown all the possibilities of our methodology. As is pointed out in [Sim98b], folded invariant curves appear in a natural way, when we introduce more harmonics in the studied family of maps. Since these curves can be constructed following arbitrarily complicated paths in phase space, they got the name of *labyrinthic curves*. We shall consider the following family of maps $F_\varepsilon = F_{s_1,s_2,c_1,y_2,y_3,\gamma,\varepsilon}$ defined on $\mathbb{R}/(2\pi\mathbb{Z}) \times \mathbb{R}$:

$$F_\varepsilon : \begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} \bar{x} \\ \bar{y} \end{pmatrix} = \begin{pmatrix} x + \varepsilon P(\bar{y}) \\ y + \varepsilon\gamma T(x) \end{pmatrix}. \tag{III.28}$$

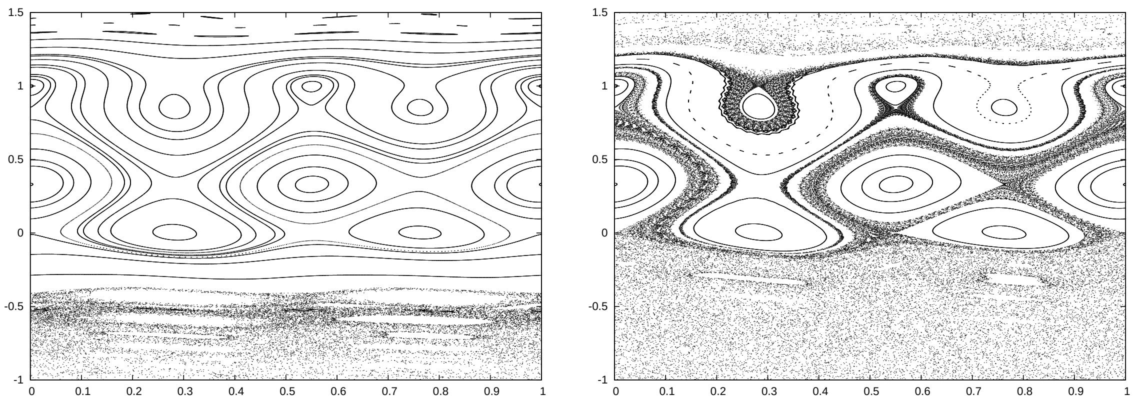

Figure III.14: Phase space in the xy-plane of family (III.28). Left: $\varepsilon = 5$. Right: $\varepsilon = 10$.

where

$$P(y) = y(y - y_2)(y - y_3)(y - 1)$$
$$T(x) = s_1 \sin(x) + s_2 \sin(2x) + c_1(\cos(x) - \cos(2x)).$$

From now on, following [Sim98b], we fix

$$y_2 = 0.33040195,$$
$$y_3 = 0.84999789,$$
$$s_1 = 0.1608819674465999,$$
$$s_2 = 0.9444712344787136,$$
$$c_1 = 0.2865154093461046,$$
$$\gamma = -0.0049.$$

In the left plot of Figure III.14 we show some iterates for $\varepsilon = 5$ corresponding to several initial conditions. We observe that, if we consider the main elliptic islands as holes on the cylinder, then we find invariant curves of different homotopy classes. Since these curves are folded in a very complicated way, it is difficult to face the systematic computation of the "lift" of the direct projection of the iterates. Then, the unfolding method turns out to be very useful to compute these invariant curves.

For example, we consider the curve associated to the initial condition $(x_0, y_0) = (0, 2.2)$, for $\varepsilon = 5$. We first obtain an approximation of the rotation number by means of frequency analysis as is explained in Section III.2.4. In particular, using the complex variable $z = \mathrm{e}^y \mathrm{e}^{\mathrm{i}x}$ and $L = 100$ we obtain the approximation of the rotation number $\theta_0 = 7.46161 \cdot 10^{-3}$ (details are omitted since they are very similar of those corresponding to the previously discussed examples). Using this approximation θ_0 we can unfold the invariant curve computing the averages (III.3) for $L = 150$ (for this values the curve is almost a circle). Then, we apply Algorithm II.2.5 using

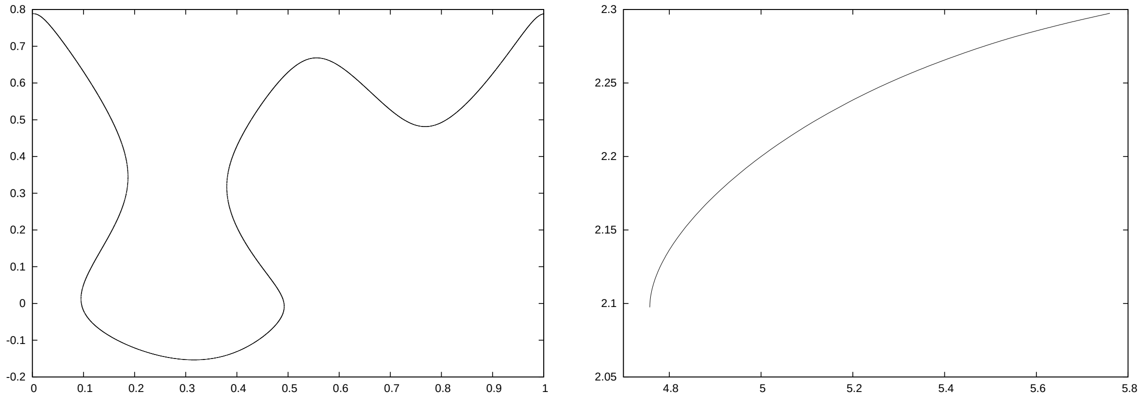

Figure III.15: Left: Labyrinthic invariant curve in the xy-plane of (III.28) corresponding to the initial condition $(x_0, y_0) = (0, 2.2)$ for $\varepsilon = 5$. Right: We plot $y(\varepsilon)$ versus ε, in the continuation of the invariant curve on the left.

$p = 9$ and $q = 23$. Computations are performed using 32-digit arithmetics (*double-double* data type from [HLB05]). We obtain the approximation

$$\theta \simeq \Theta_{23,9} = 0.0080998168999841202002324221453501 \qquad \text{(III.29)}$$

of the rotation number with an estimated extrapolation error of order 10^{-34}, that corresponds with the arithmetic precision of our computations. Let us observe that the number (III.29) is very close to zero, but its continued fraction is $[0, 123, 2, 5, 1, 2, \ldots]$, so we conclude that it is not so resonant as the example of Hénon's map that we considered before —for this reason the obtained computations are more accurate.

Furthermore, we use the methods of the chapter to follow the evolution, when ε changes, of the invariant curve of F_ε having a prefixed rotation number θ. Concretely, let $(x_0, y(\varepsilon_0))$ be a point on an invariant curve of rotation number θ of the map F_{ε_0}. Then, given ε close to ε_0, we want to compute an initial condition $(x_0, y(\varepsilon))$ that corresponds to the invariant curve —if it exists— of F_ε that has the same rotation number. Indeed, if we denote by $\rho(\varepsilon, y_0)$ the function that gives the rotation number of the invariant curve of F_ε for the point (x_0, y_0) —if it exists— then we look for solutions with respect to $y_0 = y(\varepsilon)$ of the equation $\rho(\varepsilon, y_0) = \theta$. The computation of the point $y(\varepsilon)$ is performed by means of the secant method as is done in [SV06] for the standard map. Another possibility is to use a Newton scheme like in Section II.5.2 for Hénon's map, computing derivatives of the rotation number as described in Section III.2.5, but the secant method is enough for our purposes.

We continue the curve of rotation number given by (III.29) starting from the values $\varepsilon_0 = 5$ and $(x_0, y_0) = (0, 2.2)$. Computations are performed using 32-digit arithmetics (*double-double* data type from [HLB05]). In order to unfold invariant curves we use as an approximation the prefixed value of the rotation number (curves nearby have a similar rotation number) and we take $L = 150$. For computing the rotation number, we apply Algorithm II.2.5 using $p = 6$ and 2^{16} iterates of the map at most. We estimate the error of the rotation number by using (II.13) and we validate this computation if it is smaller than 10^{-16}. For the secant method we require an error smaller than 10^{-14}.

The continuation of the invariant curve is performed successfully for values of ε in the interval $(4.75745894, 5.75985518)$ —in Figure III.15 we plot $y(\varepsilon)$ in this interval. For $\varepsilon = 4.75745894$ the invariant curve turns shearless and then it disappears after the collapse with the other invariant curve of the same rotation number on the other side of the meandering, so we have a turning point in the continuation. For $\varepsilon = 5.75985518$ we stop because we are very close to the breakdown of the curve. For $\varepsilon > 5.75985518$ we have observed that the chaotic zone that appears at the breakdown of this invariant curve is very narrow and there are still many invariant curves nearby. In the right plot of Figure III.14 we include the phase space that corresponds to $\varepsilon = 10$ in order to show the chaotic zone that is created when most of the invariant curves around the continued one are destroyed.

Chapter IV

A KAM theorem without action-angle variables for elliptic lower dimensional tori

Abstract of the chapter. *We study elliptic lower dimensional invariant tori of Hamiltonian systems via parameterizations. The method is based in solving iteratively the functional equations that stand for invariance and reducibility. In contrast with classical methods, we do not assume that the system is close to integrable nor that is written in action-angle variables. We only assume that we have an approximation of an invariant torus of fixed vector of basic frequencies and a basis along the torus that approximately reduces the normal variational equations to constant coefficients. We want to hightlight that this approach presents many advantages compared with methods which are built in terms of canonical transformations, e.g., it produces simpler and more constructive proofs that lead to more efficient numerical algorithms for the computation of these objects. Such numerical algorithms are suitable to be adapted in order to perform computer assisted proofs.*

IV.1 Introduction

Persistence of quasi-periodic solutions has been for long time a subject of remarkable importance in dynamical systems. Roughly speaking, KAM theory —named after A.N. Kolmogorov [Kol54], V.I. Arnold [Arn63] and J.K. Moser [Mos62]— deals with the effect of small perturbations on dynamical systems (typically Hamiltonian) which admit invariant tori carrying quasi-periodic motion. Nowadays, KAM theory is a vast area of research that involves a large collection of methods and applications to a wide set of contexts: Hamiltonian systems, reversible systems, volume-preserving systems, symplectic maps, PDEs and lattices, just to mention a few. We refer to [Bos86, BHS96, dlL01, Pös01, Sev07, SM71] for different surveys or tutorials that collect many aspect of the theory and cover a large amount of bibliography.

In this work we are concerned with lower dimensional (isotropic) tori of Hamiltonian systems. Thus, let us consider a real analytic Hamiltonian system with n degrees of freedom having an invariant torus of dimension $r < n$, carrying quasi-periodic dynamics with vector of basic frequencies $\omega \in \mathbb{R}^r$. The variational equations around such a torus correspond to a $2n$-dimensional linear quasi-periodic system with vector of frequencies ω. For this linear system we have $2r$ trivial directions (i.e., zero eigenvalues of the reduced matrix of the system restricted to these directions) associated to the tangent directions of the torus and the symplectic conjugate ones (these trivial directions are usually referred as the *central directions* of the torus). If the remaining $2(n-r)$ directions (*normal directions* of the torus) are hyperbolic, we say that the torus is *hyperbolic* or *whiskered*. Hyperbolic tori are very robust under perturbations [dlLW04, FdlLS09, Gra74, HL00, JdlLZ99, LY05, Val00, Zeh76, ZLL08]. For example, if we consider a perturbation of the system depending analytically on external parameters, it can be established, under suitable conditions, the existence of an analytic (with respect to these parameters) family of hyperbolic tori having the same basic frequencies. In the above setting, if the torus possesses some elliptic (oscillatory) normal directions we say that it is *elliptic* or *partially elliptic*. In this case the situation is completely different, since we have to take into account combinations between basic and normal frequencies in the small divisors that appear in the construction of these tori (the corresponding non-resonance conditions are usually referred as Melnikov conditions [Mel65]). As a consequence, families of elliptic or partially elliptic invariant tori with fixed basic frequencies ω cannot be continuous in general, but they turn out to be Cantorian with respect to parameters. First rigorous proofs of existence of elliptic tori were given in [Mos67] for $r = n-1$, and in [Eli88, Kuk88] for $r < n$. We refer also to [BHS96, HLY06, JdlLZ99, JV97a, JV97b, Pös89, Sev96, XY01] as interesting contributions covering different points of view.

The main source of difficulty in presence of elliptic normal directions is the so-called *lack of parameters* problem [BHS96, Eli88, JV97a, JV97b, Mos67, Sev99]. Basically, since we have only as many internal parameters ("actions") as the number of basic frequencies of the torus, we cannot control simultaneously the normal ones, so we cannot prevent them from "falling into resonance". This is equivalent to say that, for a given Hamiltonian system, we cannot construct a torus with fixed basic and normal frequencies, because there are not enough internal parameters. The previous fact leads to the exclusion of a small set of these internal parameters in order to avoid resonances involving normal frequencies. To perform the previous exclusion, it is necessary to assume that the normal frequencies "move" as a function of the internal parameters. Another possibility to overcome this problem is to apply the so-called Broer-Huitema-Takens theory (see [BHTB90]). This consists in adding as many (external) parameters as needed to control simultaneously the values of both basic and normal frequencies (this process is referred as *unfolding*). Then, under suitable (generic) hypotheses on the parameter dependence, we can prove that —under small perturbations— there exist invariant tori over a "Cantor set" of large measure in parameter space. Finally, if we are interested in recovering invariant tori for the original system, it is possible to obtain measure estimates for the corresponding set of projected parameters by means of results on *Diophantine approximation on submanifolds* (see [BHS96, Sev99]).

Another difficulty linked to persistence of lower dimensional invariant tori refers to reducibility of the normal variational equations (at least in the elliptic directions) which is usually asked in order to simplify the study of the linearized equations involved. In order to achieve this reducibility, it is typical to consider second order Melnikov conditions [Mel65] to control the small divisors of the cohomological equations appearing in the construction of the reduced matrix. Other approaches for studying persistence of invariant tori in the elliptic context, without second order Melnikov conditions, are discussed in Remark IV.3.10.

Classical methods for studying persistence of lower dimensional tori are based on canonical transformations performed on the Hamiltonian function. These methods typically deal with a perturbative setting in such a way that the problem is written as a perturbation of an "integrable" Hamiltonian (in the sense that it has a continuous family of reducible invariant tori), and take advantage of the existence of action-angle-like coordinates for the unperturbed integrable Hamiltonian system. These coordinates play an important role not only in solving the cohomological equations involved in the iterative KAM process, but also allowing us to control the isotropic[1] character of the tori thus simplifying a lot of details. However, classical approaches present some shortcomings, mainly due to the fact that they only allow us to face perturbative problems. For example

- In many practical applications (design of space missions [GJSM01a, GJSM01b], study of models in Celestial Mechanics [SM71], Molecular Dynamics [PCU08, TLBB01] or Plasma-Beam Physics [Mic95], just to mention a few) we have to consider non-perturbative systems. For such systems we can obtain approximate invariant tori by means of numerical computations or asymptotic expansions, but in general we cannot apply the classical results to prove the existence of these objects. Furthermore, in some cases it is possible to identify an integrable approximation of a given system but the remaining part cannot be considered as an arbitrarily small perturbation.

- Even if we are studying a concrete perturbative problem, sometimes it is very complicated to establish action-angle variables for the unperturbed Hamiltonian. In some cases, action-angle variables are not explicit, become singular or introduce problems of regularity (for example, when we approach to a separatrix). Although in many contexts this shortcoming has been solved by means of several techniques (see for example [DG96b, Her83, Neĭ84] for a construction in the case of an integrable Hamiltonian or [JV98, Sev96] for a construction around a particular object), it introduces more technical difficulties in the problem.

- From the computational viewpoint, methods based on transformations are sometimes inefficient and quite expensive. This is a serious difficulty in order to implement numerical methods or computer assisted proofs based on them.

[1] If we pull-back an isotropic torus by means of a symplectomorphism, the isotropic character is preserved.

An alternative to the classical approach is the use of so-called parameterization methods, which consist in performing an iterative scheme to solve the invariance equation of the torus. Instead of performing canonical transformations, this scheme is carried out by adding a small function to the previous approximation of the torus, and thus, it is suitable for studying existence of invariant tori of Hamiltonian systems without using neither action-angle variables nor a perturbative setting. The previous correction is obtained iteratively by solving (approximately) the corresponding linearized equations (Newton method). We point out that the geometry of the problem plays an important role in the study of these equations. Such geometric approach —also refered as KAM theory without action-angle variables— was introduced in [dlLGJV05] for Lagrangian tori and extended in [FdlLS09] to hyperbolic lower dimensional tori, following long-time developed ideas (relevant work can be found in [JdlLZ99, Rüs76, SZ89, Zeh76] but we refer to [dlL01, dlLGJV05] for a detailed list). Roughly speaking, the insight of these methods is summarized in the following quote from [JdlLZ99]: *"...near approximate solutions of certain equations satisfying certain non-degeneracy assumptions, we can find true solutions defined on a large set."*

The aim of this chapter is to adapt KAM methods without action-angle variables to the normally elliptic context in a non-perturbative setting. Concretely, we assume that we have a 1-parameter family of Hamiltonian systems for which we know a 1-parameter family of approximately invariant r-dimensional elliptic tori —all of them with the same vector of basic frequencies— and also approximations of the vectors of normal frequencies and the corresponding normal directions associated to these frequencies (i.e., a basis of the normal directions along each torus that approximately reduces the normal variational equations to constant coefficients). Then we show that, under suitable hypotheses of non-resonance and non-degeneracy, for a Cantorian subset of parameters —of large relative Lebesgue measure— there exist true elliptic tori close to the approximate ones, having the same vector of basic frequencies and slightly modified vector of normal frequencies. The scheme to deal with reducibility of the normal directions of these tori is the main contribution of this chapter, and it consists in performing suitable (small) corrections in the normal directions at each step of the iterative procedure.

We want to highlight that parameterization methods, as presented above, are computationally oriented in the sense that they can be implemented numerically, thus obtaining very efficient algorithms for the computation of invariant tori. For example, if we approximate a torus by using N Fourier modes, such algorithms allow us to compute the object with a cost of order $\mathcal{O}(N \log N)$ in time and $\mathcal{O}(N)$ in memory (see Remark IV.3.11). This is another advantadge of our approach in contrast with other existing methods to deal with elliptic lower dimensional invariant tori. We refer to [dlLHS] for the implementation of the ideas in [dlLGJV05, FdlLS09] for Lagrangian and whiskered tori (see also [CdlL09] for the case of lattices and twist maps).

The setting discussed in this chapter has been selected in order to simplify some (well known) technical aspects of the result —both in the assumptions and in the proof— but we want to point out that all the basic geometric ideas linked to parameterization methods, without using action-angle variables, for reducible lower-dimensional tori are present in our approch. In

Section 3 we discuss several extensions and generalizations that can be tackled with the method presented in this work.

Finally, we observe that in presence of hyperbolic directions one can approach the problem by combining techniques in [FdlLS09] (for studying hyperbolic directions) together with those presented here (for studying elliptic directions). Indeed, the methodology presented in this work can be adapted to deal with invariant tori with reducible hyperbolic directions, but this assumption is fairly restrictive (see [HdlL06, HdlL07]) in the hyperbolic context (reducibility is not required in [FdlLS09]).

The chapter is organized as follows. In Section IV.2 we provide some notations, definitions and background of the problem. In Section IV.3 we state the main result of this chapter and we discuss several extensions and generalizations of the presented method presented. A motivating sketch of the construction performed in the proof of this result is given in Section IV.4, together with a detailed description of some geometric properties of elliptic lower dimensional invariant tori of Hamiltonian systems. Next, in Section IV.5 we perform one step of the iterative method to correct both an approximation of an elliptic invariant torus and a basis along this torus that approximately reduces the normal variational equations to constant coefficients. The new errors in invariance and reducibility are quadratic in terms of the previous ones. The main result is proved in Section IV.6.

IV.2 General background

In this section we introduce some notation and, in order to help the reader, we recall the basic terminology and concepts related to the problem. Thus, after setting the notation used along the chapter in Section IV.2.1, we provide the basic definitions regarding lower dimensional invariant tori of Hamiltonian systems (Section IV.2.2) and their normal behavior (Section IV.2.3).

IV.2.1 Basic notations

Given a real or complex function f of several variables, we denote Df the Jacobian matrix, $\mathrm{grad}\, f = Df^\top$ the gradient vector and $\mathrm{hess}\, f = D^2 f$ the Hessian matrix, respectively.

For any complex number $z \in \mathbb{C}$ we denote $z^* \in \mathbb{C}$ its complex conjugate number and $\mathrm{Re}(z)$, $\mathrm{Im}(z)$ the real and imaginary parts of z, respectively. We extend these notations to complex vectors and matrices.

Given a complex vector $v \in \mathbb{C}^l$ we denote by $\mathrm{diag}\,(v) \in M_{l\times l}(\mathbb{C})$ the diagonal matrix having the components of v in the diagonal. Moreover, given $Z \in M_{l\times l}(\mathbb{C})$, we denote by $\mathrm{diag}\,(Z) \in M_{l\times l}(\mathbb{C})$ the diagonal matrix having the same diagonal entries as Z.

For any $k \in \mathbb{Z}^r$, we denote $|k|_1 = |k_1| + \ldots + |k_r|$. Given a vector $x \in \mathbb{C}^l$, we set the supremum norm $|x| = \sup_{j=1,\ldots,l} |x_j|$ and we extend the notation to the induced norm for complex matrices. Furthermore, given an analytic function f, with bounded derivatives in a complex domain $\mathcal{U} \subset \mathbb{C}^l$, and $m \in \mathbb{N}$ we introduce the $\mathcal{C}^m$-norm for f as

$$|f|_{\mathcal{C}^m,\mathcal{U}} = \sup_{\substack{k\in(\mathbb{N}\cup\{0\})^l \\ 0\le|k|_1\le m}} \sup_{z\in\mathcal{U}} |D^k f(z)|.$$

We denote by $\mathbb{T}^r = \mathbb{R}^r/(2\pi\mathbb{Z})^r$ the real r-dimensional torus, with $r \geq 1$. We use the $|\cdot|$-norm introduced above to define the complex strip around $\mathbb{T}^r$ of width $\rho > 0$ as

$$\Delta(\rho) = \{\theta \in \mathbb{C}^r/(2\pi\mathbb{Z})^r : |\mathrm{Im}(\theta)| \le \rho\}.$$

Accordingly we will consider the Banach space of analytic functions $f : \Delta(\rho) \to \mathbb{C}$ equipped with the norm

$$\|f\|_\rho = \sup_{\theta\in\Delta(\rho)} |f(\theta)|,$$

and similarly, if f takes values in $\mathbb{C}^l$, we set $\|f\|_\rho = |(\|f_1\|_\rho, \ldots, \|f_l\|_\rho)|$. If f is a matrix valued function, we extend $\|f\|_\rho$ by computing the $|\cdot|$-norm of the constant matrix defined by the $\|\cdot\|_\rho$-norms of the entries of f. We observe that if the matrix product is defined then this space is a Banach algebra and we have $\|f_1 f_2\|_\rho \le \|f_1\|_\rho \|f_2\|_\rho$. In addition, we can use Cauchy estimates

$$\left\|\frac{\partial f}{\partial\theta_j}\right\|_{\rho-\delta} \le \frac{\|f\|_\rho}{\delta}, \qquad j = 1, \ldots, r.$$

For any function f analytic on $\mathbb{T}^r$ and taking values in $\mathbb{C}$, $\mathbb{C}^l$ or in a space of complex matrices, we denote its Fourier series as

$$f(\theta) = \sum_{k\in\mathbb{Z}^n} \hat{f}_k \mathrm{e}^{\mathrm{i}\langle k,\theta\rangle}, \qquad \hat{f}_k = \frac{1}{(2\pi)^r}\int_{\mathbb{T}^r} f(\theta)\mathrm{e}^{-\mathrm{i}\langle k,\theta\rangle} d\theta$$

and its average as $[f]_{\mathbb{T}^r} = \hat{f}_0$. We also set $\widetilde{f}(\theta) = f(\theta) - [f]_{\mathbb{T}^r}$. Moreover, we have the following bounds

$$|\,[f]_{\mathbb{T}^r}\,| \le \|f\|_\rho, \qquad \|\widetilde{f}\|_\rho \le 2\|f\|_\rho, \qquad |\hat{f}_k| \le \|f\|_\rho \mathrm{e}^{-\rho|k|_1}.$$

Now, we introduce some notation regarding Lipschitz regularity. Assume that $f(\mu)$ is a function defined for $\mu \in I \subset \mathbb{R}$ —the subset I may not be an interval— taking values in $\mathbb{C}$, $\mathbb{C}^l$ or $M_{l_1\times l_2}(\mathbb{C})$. We say that f is Lipschitz with respect to μ on the set I if

$$\mathrm{Lip}_I(f) = \sup_{\substack{\mu_1,\mu_2\in I \\ \mu_1\ne\mu_2}} \frac{|f(\mu_2) - f(\mu_1)|}{|\mu_2 - \mu_1|} < \infty.$$

The value $\mathrm{Lip}_I(f)$ is called the Lipschitz constant of f on I. For these functions we define $\|f\|_I = \sup_{\mu\in I} |f(\mu)|$. Similarly, if we have a family $\mu \in I \subset \mathbb{R} \mapsto f_\mu$, where f_μ is a function on $\mathbb{T}^r$ taking values in $\mathbb{C}$, $\mathbb{C}^l$ or $M_{l_1\times l_2}(\mathbb{C})$, we extend the previous notations as

$$\mathrm{Lip}_{I,\rho}(f) = \sup_{\substack{\mu_1,\mu_2\in I \\ \mu_1\ne\mu_2}} \frac{\|f_{\mu_2} - f_{\mu_1}\|_\rho}{|\mu_2 - \mu_1|} < \infty, \qquad \|f\|_{I,\rho} = \sup_{\mu\in I} \|f_\mu\|_\rho.$$

Analogously, we say that f is Lipschitz from below with respect to μ on the set I if

$$\mathrm{lip}_I(f) = \inf_{\substack{\mu_1,\mu_2\in I\\ \mu_1\neq\mu_2}} \frac{|f(\mu_2)-f(\mu_1)|}{|\mu_2-\mu_1|} < \infty.$$

Finally, given a family $\mu \in I \subset \mathbb{R} \mapsto f_\mu$, where f_μ is an analytic function with bounded derivatives in a complex domain $\mathcal{U} \subset \mathbb{C}^l$, we introduce for $m \in \mathbb{N}$

$$\mathrm{Lip}_{I,\mathcal{C}^m,\mathcal{U}}(f) = \sup_{\substack{\mu_1,\mu_2\in I\\ \mu_1\neq\mu_2}} \frac{|f_{\mu_1}-f_{\mu_2}|_{\mathcal{C}^m,\mathcal{U}}}{|\mu_1-\mu_2|}, \qquad \|f\|_{I,\mathcal{C}^m,\mathcal{U}} = \sup_{\mu\in I} |f_\mu|_{\mathcal{C}^m,\mathcal{U}}.$$

In this work we are concerned with Hamiltonian systems in $\mathbb{R}^{2n}$ with respect to the standard symplectic form Ω^0, given by $\Omega^0(\xi,\eta) = \xi^\top J_n \eta$ where

$$J_n = \begin{pmatrix} 0 & \mathrm{Id}_n \\ -\mathrm{Id}_n & 0 \end{pmatrix}$$

is the canonical skew-symmetric matrix. We extend the notation above and we write J_j for any $1 \leq j \leq n$, and Id_j for any $1 \leq j \leq 2n$. For the sake of simplicity, we denote $J = J_n$ and $\mathrm{Id} = \mathrm{Id}_{2n}$.

Finally, given functions $A : \mathbb{T}^r \to M_{2n\times l_a}(\mathbb{C})$ and $B : \mathbb{T}^r \to M_{2n\times l_b}(\mathbb{C})$, we set the notations $G_{A,B}(\theta) = A(\theta)^\top B(\theta)$, $\Omega_{A,B}(\theta) = A(\theta)^\top J B(\theta)$, $G_A(\theta) = G_{A,A}(\theta)$ and $\Omega_A(\theta) = \Omega_{A,A}(\theta)$.

IV.2.2 Invariant and approximately invariant tori

Given a Hamiltonian function $h : U \subset \mathbb{R}^{2n} \to \mathbb{R}$, we study the existence of lower dimensional quasi-periodic invariant tori for the Hamiltonian vector field $X_h(x) = J\,\mathrm{grad}\, h(x)$.

Definition IV.2.1. *For any integer* $1 \leq r \leq n$, $\mathscr{T} \subset U$ *is an* r-dimensional quasi-periodic invariant torus with basic frequencies $\omega \in \mathbb{R}^r$ for X_h, *if* $\mathscr{T}$ *is invariant under the flow of* X_h *and there exists a* parameterization *given by an embedding* $\tau : \mathbb{T}^r \to U$ *such that* $\mathscr{T} = \tau(\mathbb{T}^r)$, *making the following diagram commute*

$$\begin{array}{ccc} \mathbb{T}^r & \xrightarrow{T_{t,\omega}} & \mathbb{T}^r \\ {\scriptstyle\tau}\downarrow & & \downarrow{\scriptstyle\tau} \\ \mathscr{T} & \xrightarrow{\phi_t|_{\mathscr{T}}} & \mathscr{T} \end{array} \tag{IV.1}$$

where $T_{t,\omega}(x) = x + \omega t$ *is the (parallel) flow of the constant vector field*

$$L_\omega = \omega_1 \frac{\partial}{\partial\theta_1} + \ldots + \omega_r \frac{\partial}{\partial\theta_r}$$

and ϕ_t is the flow of the Hamiltonian vector field X_h. In addition, if

$$\langle k, \omega \rangle \neq 0, \qquad \forall k \in \mathbb{Z}^r \backslash \{0\}, \tag{IV.2}$$

then we say that ω is non-resonant.

Let us observe that, if $\omega \in \mathbb{R}^r$ is non-resonant, then the quasi-periodic function $z(t) = \tau(\omega t + \theta_0)$ is an integral curve of X_h for any $\theta_0 \in \mathbb{T}^n$ that fills densely $\mathscr{T}$. Equivalently, we have that the embedding τ satisfies

$$L_\omega \tau(\theta) = X_h(\tau(\theta)). \tag{IV.3}$$

By means of τ we can pull-back to $\mathbb{T}^r$ both the restrictions to $\mathscr{T}$ of the standard metric and the symplectic structure, obtaining the following matrix representations

$$G_{D\tau}(\theta) := D\tau(\theta)^\top D\tau(\theta), \qquad \Omega_{D\tau}(\theta) := D\tau(\theta)^\top J D\tau(\theta), \qquad \theta \in \mathbb{T}^r.$$

Remark IV.2.2. *We note that as τ is an embedding we have $\operatorname{rank}(D\tau(\theta)) = r$ for every $\theta \in \mathbb{T}^r$, so it turns out that $\det G_{D\tau}(\theta) \neq 0$ for every $\theta \in \mathbb{T}^r$. Moreover, we see that the average $[\Omega_{D\tau}]_{\mathbb{T}^r}$ is zero since if we write $\tau(\theta) = (x(\theta), y(\theta))$ then we have $\Omega_{D\tau}(\theta) = D\alpha(\theta) - D\alpha(\theta)^\top$, where $\alpha(\theta) = Dx(\theta)^\top y(\theta)$ and, by definition, $[D\alpha]_{\mathbb{T}^r} = 0$.*

Lemma IV.2.3. *Let $h : U \subset \mathbb{R}^{2n} \to \mathbb{R}$ be a Hamiltonian function and $\mathscr{T}$ an r-dimensional invariant torus for X_h of non-resonant frequencies ω. Then, the submanifold $\mathscr{T}$ is isotropic, i.e., $\Omega_{D\tau}(\theta) = 0$ for every $\theta \in \mathbb{T}^r$. In particular, if $r = n$ then $\mathscr{T}$ is Lagrangian.*

Remark IV.2.4. *Along the text there appear many functions depending on $\theta \in \mathbb{T}^r$. In order to simplify the notation sometimes we omit the dependence on θ —eventually we even omit the fact that some functions are evaluated at $\tau(\theta)$ if there is no source of confusion.*

Proof of Lemma IV.2.3. The isotropic character of $\mathscr{T}$ is obtained as it was done in [dlLGJV05]. First, we compute

$$\begin{aligned} L_\omega(\Omega_{D\tau}) &= L_\omega(D\tau^\top J D\tau) = [D(L_\omega \tau)]^\top J D\tau + D\tau^\top J D(L_\omega \tau) \\ &= [J \operatorname{hess} h(\tau) D\tau]^\top J D\tau + D\tau^\top J J \operatorname{hess} h(\tau) D\tau = 0, \end{aligned}$$

where we used that $D \circ L_\omega = L_\omega \circ D$, the hypothesis $L_\omega \tau(\theta) = X_h(\tau(\theta))$ and the properties $J^\top = -J$ and $J^2 = -\mathrm{Id}$. Then, since ω is non-resonant, the fact that the derivative L_ω vanishes implies that $\Omega_{D\tau} = [\Omega_{D\tau}]_{\mathbb{T}^r}$. Finally, from Remark IV.2.2 we conclude that $\Omega_{D\tau} = 0$. □

Finally, we set the idea of parameterization of an approximately invariant torus. Essentially, we measure how far to commute is diagram (IV.1).

Definition IV.2.5. *Given a Hamiltonian* $h : U \subset \mathbb{R}^{2n} \to \mathbb{R}$ *and an integer* $1 \leq r \leq n$, *we say that* $\mathscr{T} \subset U$ *is an* r-dimensional approximately quasi-periodic invariant torus with non-resonant basic frequencies $\omega \in \mathbb{R}^r$ for X_h *provided that there exists an embedding* $\tau : \mathbb{T}^r \to U$, *such that* $\mathscr{T} = \tau(\mathbb{T}^r)$, *satisfying*

$$L_\omega \tau(\theta) = J \operatorname{grad} h(\tau(\theta)) + e(\theta),$$

where $e : \mathbb{T}^r \to \mathbb{R}^{2n}$ *is "small" in a suitable norm.*

Among the conditions needed to find a true invariant torus around an approximately invariant one, we are concerned with Diophantine conditions on the vector of basic frequencies.

Definition IV.2.6. *We say that* $\omega \in \mathbb{R}^r$ *satisfies Diophantine conditions of* (γ, ν)*-type, for* $\gamma > 0$ *and* $\nu > r - 1$, *if*

$$|\langle k, \omega \rangle| \geq \frac{\gamma}{|k|_1^\nu}, \qquad k \in \mathbb{Z}^r \backslash \{0\}. \tag{IV.4}$$

It is well-known that if we consider a fixed ν then, for almost every $\omega \in \mathbb{R}^r$, there is $\gamma > 0$ for which (IV.4) is fulfilled (see [LM88]).

IV.2.3 Linear normal behavior of invariant tori

In order to study the behavior of the solutions in a neighborhood of an r-dimensional quasi-periodic invariant torus of basic frequencies ω —parameterized by τ— it is usual to consider the variational equations around the torus, given by

$$L_\omega \xi(\theta) = J \operatorname{hess} h(\tau(\theta)) \xi(\theta). \tag{IV.5}$$

If $r = 1$ the system (IV.5) is $2\pi/\omega$-periodic. Then, following Floquet's theorem, there exists a linear periodic change of variables that reduces the system to constants coefficients. If $r > 1$, then we consider reducibility to constant coefficients (in the sense of Lyapunov-Perron) as follows.

Definition IV.2.7. *Let* $h : U \subset \mathbb{R}^{2n} \to \mathbb{R}$ *be a Hamiltonian function and* $\mathscr{T}$ *be an* r*-dimensional invariant torus with basic frequencies* ω *(see Definition IV.2.1). Then,* $\mathscr{T}$ *is* reducible *if there exists a linear change of coordinates* $\xi = M(\theta)\eta$, *defined for* $\theta \in \mathbb{T}^r$, *such that the variational equations* (IV.5) *turn out to be* $L_\omega \eta(\theta) = B\eta(\theta)$, *where* $B \in M_{2n \times 2n}(\mathbb{C})$.

It is immediate to check that this property is equivalent to the fact that the transformation given by M satisfies the differential equation

$$L_\omega M(\theta) = J \operatorname{hess} h(\tau(\theta)) M(\theta) - M(\theta) B. \tag{IV.6}$$

In the Lagrangian case $r = n$, under regularity assumptions, such transformation exists provided ω satisfies (IV.4) due to the geometric constrains of the problem (as it is discussed in [dlLGJV05]). Indeed, we can take derivatives at both sides of the invariance equation (IV.3), thus obtaining

$$L_\omega D\tau(\theta) = J\text{hess}\, h(\tau(\theta))D\tau(\theta).$$

Then, we can choose a suitable $n \times n$ matrix $C(\theta)$ —given by the solution of certain cohomological equation— in such a way that the columns of $D\tau(\theta)$ and $JD\tau(\theta)G_{D\tau}^{-1}(\theta) + D\tau(\theta)C(\theta)$ give us the matrix $M(\theta)$. Then, the reduced matrix turns out to be of the form

$$B = \begin{pmatrix} 0 & B^C \\ 0 & 0 \end{pmatrix},$$

where $B^C \in M_{n\times n}(\mathbb{R})$ is symmetric. The $2n$ zero eigenvalues correspond to the tangent directions to the torus together with their symplectic conjugate ones, meanwhile the matrix B^C controls the variation of the frequencies of the torus —the twist condition reads $\det B^C \neq 0$— when moving the "actions" of the system.

In the lower dimensional case $1 < r < n$ we cannot guarantee, in general, reducibility to constant coefficients (we refer to [JM82, JS81, Pui]). Nevertheless, if we consider a family of quasi-periodic linear perturbations of a linear system with constant coefficients then, under some generic hypothesis of non-resonance and non-degeneracy, we can state the reducibility of a large subfamily. On the one hand, if we restrict $M(\theta)$ to the space of close-to-the-identity matrices, then we can prove that the reducible subfamily is Cantorian and has large Lebesgue measure (we refer to [JS92, JS96]). On the other hand, considering a more general class of matrices (see ideas introduced in [Eli92, Eli01, Kri99]) this result can be extended to a full measure subfamily (this was conjectured in [Eli01] and proved in [HY08]).

If the system (IV.5) is reducible, it turns out that the geometry of the problem allows us to choose the matrix B with the following block structure

$$B = \left(\begin{array}{cc|c} 0 & B^C & 0 \\ 0 & 0 & 0 \\ \hline 0 & 0 & B^N \end{array}\right),$$

where $B^C \in M_{r\times r}(\mathbb{R})$ is symmetric (it plays the same "twist" role as in the Lagrangian case), and $B^N \in M_{2(n-r)\times 2(n-r)}(\mathbb{C})$ can be written as $B^N = J_{n-r}S$, where S is also symmetric. In this context, B^N gives the normal linear behavior of the torus. The real parts of the eigenvalues of B^N correspond to Lyapunov exponents and their imaginary parts to normal frequencies. As discussed in the introduction, in this work we are interested in the normally elliptic case, in which all the eigenvalues of B^N have vanishing real part, i.e.,

$$\text{spec}\,(B^N) = \{\mathrm{i}\lambda_1, \ldots, \mathrm{i}\lambda_{n-r}, -\mathrm{i}\lambda_1, \ldots, -\mathrm{i}\lambda_{n-r}\},$$

where $\lambda_j \in \mathbb{R}\backslash\{0\}$ are the so-called *normal frequencies*. Thoughtout the chapter we assume that they have different modulus.

In order to simplify the resolution of the obtained cohomological equations, it is convenient to put the matrix B^N in diagonal form. In the classical KAM approach —using symplectic transformations and action-angle variables adapted to the torus— this is possible with a complex canonical change of coordinates, that transforms the initial real Hamiltonian into a complex one, having some symmetries. As these symmetries are preserved by the canonical transformations performed along these classical proofs, the final Hamiltonian can be realified and thus the obtained tori are real. In this work we perform this complexification by selecting a complex matrix function $N : \mathbb{T}^r \to M_{2n\times(n-r)}(\mathbb{C})$ associated to the eigenfunctions of eigenvalues $\mathrm{i}\lambda_1, \ldots, \mathrm{i}\lambda_{n-r}$. It is clear that the real and imaginary parts of these vectors span the associated real normal subspace at any point of the torus. Indeed, from Equation (IV.6), the matrix function N satisfies

$$L_\omega N(\theta) = J \operatorname{hess} h(\tau(\theta))N(\theta) - N(\theta)\Lambda,$$

where $\Lambda = \operatorname{diag}(\mathrm{i}\lambda) = \operatorname{diag}(\mathrm{i}\lambda_1, \ldots, \mathrm{i}\lambda_{n-r})$. Then, together with these vectors, we resort to the use of the complex conjugate ones, that clearly satisfy

$$L_\omega N^*(\theta) = J \operatorname{hess} h(\tau(\theta))N^*(\theta) + N^*(\theta)\Lambda,$$

to span a basis of the complexified normal space along the torus (this is guaranteed by the conditions $\det G_{N,N^*} \neq 0$ on $\mathbb{T}^r$).

As we have pointed out in the introduction of the chapter, in order to face the resolution of the cohomological equations standing for invariance and reducibility of elliptic tori, we assume additional non-resonance conditions apart from (IV.2).

Definition IV.2.8. *We say that the normal frequencies $\lambda \in \mathbb{R}^{n-r}$ are non-resonant with respect to $\omega \in \mathbb{R}$, if $\lambda_j \in \mathbb{R}\backslash\{0\}$ have all different modulus and we have*

$$\langle k, \omega\rangle + \lambda_i \neq 0, \qquad \forall k \in \mathbb{Z}^r\backslash\{0\}, \qquad i = 1, \ldots, n-r, \tag{IV.7}$$

and

$$\langle k, \omega\rangle + \lambda_i \pm \lambda_j \neq 0, \qquad \forall k \in \mathbb{Z}^r\backslash\{0\}, \qquad i, j = 1, \ldots, n-r. \tag{IV.8}$$

Conditions (IV.7) *and* (IV.8) *are refered as* first and second order Melnikov conditions, *respectively.*

In the spirit of Definition IV.2.5, we introduce the idea of approximate reducibility as follows.

Definition IV.2.9. *Let us consider a Hamiltonian $h : U \subset \mathbb{R}^{2n} \to \mathbb{R}$ and $\mathscr{T} \subset U$ an r-dimensional* approximately quasi-periodic invariant torus *for X_h with non-resonant basic frequencies $\omega \in \mathbb{R}^r$ (see Definition IV.2.5). Then, we say that $\mathscr{T}$ is* approximately elliptic *if there exists a map $N : \mathbb{T}^r \to M_{2n\times(n-r)}(\mathbb{C})$ and normal frequencies $\lambda \in \mathbb{R}^{n-r}$, which are non-resonant with respect to ω, satisfying*

$$L_\omega N(\theta) = J \operatorname{hess} h(\tau(\theta))N(\theta) - N(\theta)\Lambda + R(\theta), \tag{IV.9}$$

where $\Lambda = \mathrm{diag}\,(\mathrm{i}\lambda)$, $\det G_{N,N^*} \neq 0$ *on* $\mathbb{T}^r$ *and* $R : \mathbb{T}^r \to M_{2n\times(n-r)}(\mathbb{C})$ *is "small" in a suitable norm.*

In order to avoid the effect of the small divisors associated with the additional resonances, we assume standard Diophantine conditions on (IV.7) and (IV.8).

Definition IV.2.10. *Let us consider non-resonant basic and normal frequencies* $(\omega, \lambda) \in \mathbb{R}^r \times \mathbb{R}^{n-r}$ *and constants* $\gamma > 0$ *and* $\nu > r-1$. *We say that* λ *satisfies Diophantine conditions of* (γ,ν)*-type with respect to* ω *if*

$$|\langle k, \omega\rangle + \lambda_i| \geq \frac{\gamma}{|k|_1^\nu}, \qquad |\langle k, \omega\rangle + \lambda_i \pm \lambda_j| \geq \frac{\gamma}{|k|_1^\nu}, \tag{IV.10}$$

$\forall k \in \mathbb{Z}^r \backslash \{0\}$ *and* $i, j = 1, \ldots, n-r$, *with* $i \neq j$.

IV.3 Statement of the main result

In this section we state the main result of the chapter. Concretely, if we have a 1-parameter family of Hamiltonian systems for which we know a family of approximately (with small error) elliptic lower dimensional invariant tori, with the same basic frequencies and satisfying certain non-degeneracy conditions, we use the parameter to control the normal frequencies in order to prove that there exists a large set of parameters for which we have a true elliptic invariant torus close to the approximate one. We emphasize that we do not assume that the system is given in action-angle-like coordinates nor that the Hamiltonians are close to integrable.

Theorem IV.3.1. *Let us consider a family of Hamiltonians* $\mu \in I \subset \mathbb{R} \mapsto h_\mu$ *with* $h_\mu : U \subset \mathbb{R}^{2n} \to \mathbb{R}$, *where* I *is a finite interval and* U *is an open set. Let* $\omega \in \mathbb{R}^r$ *be a vector of basic frequencies satisfying Diophantine conditions* (IV.4) *of* $(\hat{\gamma}, \nu)$*-type, with* $\hat{\gamma} > 0$ *and* $\nu > r-1$. *Assume that the following hypotheses hold:*

- H_1 *The functions* h_μ *are real analytic and can be holomorphically extended to some complex neighborhood* $\mathcal{U}$ *of* U. *Moreover, we assume that* $|h|_{I,\mathcal{C}^3,\mathcal{U}} \leq \sigma_0$.
- H_2 *There exists a family of approximate invariant and elliptic tori of* h_μ *in the sense of Definitions IV.2.5 and IV.2.9, i.e., we have families of embeddings* $\mu \in I \subset \mathbb{R} \mapsto \tau_\mu$, *matrix functions* $\mu \in I \subset \mathbb{R} \mapsto N_\mu$ *and approximated normal eigenvalues* $\Lambda_\mu = \mathrm{diag}\,(\mathrm{i}\lambda_\mu)$, *with* $\lambda_\mu \in \mathbb{R}^{n-r}$, *satisfying*

  $$L_\omega \tau_\mu(\theta) = J\mathrm{grad}\, h_\mu(\tau_\mu(\theta)) + e_\mu(\theta),$$
  $$L_\omega N_\mu(\theta) = J\mathrm{hess}\, h_\mu(\tau_\mu(\theta))N_\mu(\theta) - N_\mu(\theta)\Lambda_\mu + R_\mu(\theta),$$

  for certain error functions e_μ *and* R_μ, *where* τ_μ *and* N_μ *are analytic and can be holomorphically extended to* $\Delta(\rho)$ *for certain* $0 < \rho < 1$, *satisfying* $\tau_\mu(\Delta(\rho)) \subset \mathcal{U}$. *Assume also that we have constants* σ_1, σ_2 *such that*

  $$\|D\tau\|_{I,\rho}, \|N\|_{I,\rho}, \|G_{D\tau}^{-1}\|_{I,\rho}, \|G_{N,N^*}^{-1}\|_{I,\rho} < \sigma_1, \qquad \mathrm{dist}(\tau_\mu(\Delta(\rho)), \partial\mathcal{U}) > \sigma_2 > 0,$$

for every $\mu \in I$*, where* $\partial\mathcal{U}$ *stands for the boundary of* $\mathcal{U}$*.*

H_3 *We have* $\operatorname{diag}\left[\Omega_{N_\mu,N_\mu^*}\right]_{\mathbb{T}^r} = \mathrm{i}\mathrm{Id}_{n-r}$ *for every* $\mu \in I$*.*

H_4 *The family of matrix functions*

$$A_{1,\mu}(\theta) := G_{D\tau_\mu}^{-1}(\theta) D\tau_\mu(\theta)^\top (T_{1,\mu}(\theta) + T_{2,\mu}(\theta) + T_{2,\mu}(\theta)^\top) D\tau_\mu(\theta) G_{D\tau_\mu}^{-1}(\theta),$$

where

$$T_{1,\mu}(\theta) := J^\top \operatorname{hess} h_\mu(\tau_\mu(\theta)) J - \operatorname{hess} h_\mu(\tau_\mu(\theta)),$$
$$T_{2,\mu}(\theta) := T_{\mu,1}(\theta) J \left[D\tau_\mu(\theta) G_{D\tau_\mu}(\theta)^{-1} D\tau_\mu(\theta)^\top - \mathrm{Id}\right] \mathbf{Re}(\mathrm{i} N_\mu(\theta) N_\mu^*(\theta)^\top),$$

satisfies the non-degeneracy (twist) condition $\| \, [A_1]_{\mathbb{T}^r}^{-1} \, \|_I < \sigma_1$*.*

H_5 *There exist constants* σ_3, σ_4 *such that for every* $\mu \in I$ *the approximated normal frequencies* $\lambda_\mu = (\lambda_{1,\mu}, \ldots, \lambda_{n-r,\mu})$ *satisfy*

$$0 < \frac{\sigma_3}{2} < |\lambda_{i,\mu}| < \frac{\sigma_4}{2}, \qquad 0 < \sigma_3 < |\lambda_{i,\mu} \pm \lambda_{j,\mu}|,$$

for $i, j = 1, \ldots, n-r$*, with* $i \neq j$*.*

H_6 *The objects* h_μ*,* τ_μ*,* N_μ *and* λ_μ *are at least* $\mathcal{C}^1$ *with respect to* μ*, and we have*

$$\left\|\frac{dh}{d\mu}\right\|_{I,\mathcal{C}^3,\mathcal{U}}, \left\|\frac{dD\tau}{d\mu}\right\|_{I,\rho}, \left\|\frac{dN}{d\mu}\right\|_{I,\rho}, \left\|\frac{d\lambda_i}{d\mu}\right\|_{I} < \sigma_5,$$

for $i = 1, \ldots, n-r$*. Moreover, we have the next separation condition*

$$0 < \frac{\sigma_6}{2} < \left|\frac{d}{d\mu}\lambda_{i,\mu}\right|, \qquad 0 < \sigma_6 < \left|\frac{d}{d\mu}\lambda_{i,\mu} \pm \frac{d}{d\mu}\lambda_{j,\mu}\right|,$$

for $i, j = 1, \ldots, n-r$*, with* $i \neq j$*.*

Under these assumptions, if

$$\varepsilon = \frac{16}{\rho}\left(\|e\|_{I,\rho} + \left\|\frac{de}{d\mu}\right\|_{I,\rho}\right) + \|R\|_{I,\rho} + \left\|\frac{dR}{d\mu}\right\|_{I,\rho}$$

is sufficiently small —condition that depends on $|\omega|$*,* ν*,* r*,* n*,* σ_0*,* σ_1*,* σ_2*,* σ_3*,* σ_4*,* σ_5 *and* σ_6 *— then there exist a Cantorian subset* $I_{(\infty)} \subset I$ *such that* $\forall \mu \in I_{(\infty)}$ *the Hamiltonian* h_μ *has an* r*-dimensional elliptic invariant torus* $\mathscr{T}_{\mu,(\infty)}$ *with basic frequencies* ω *and normal frequencies* $\lambda_{\mu,(\infty)}$ *that satisfy Diophantine conditions of the form*

$$|\langle k, \omega\rangle + \lambda_{i,(\infty)}| \geq \frac{C\varepsilon^{1/10}}{|k|_1^\nu}, \qquad |\langle k, \omega\rangle + \lambda_{i,(\infty)} \pm \lambda_{j,(\infty)}| \geq \frac{C\varepsilon^{1/10}}{|k|_1^\nu}, \qquad \forall k \in \mathbb{Z}^r \backslash \{0\},$$

for $i, j = 1, \ldots, n-r$, with $i \neq j$. Moreover, $I_{(\infty)}$ has big relative Lebesgue measure

$$\mathrm{meas}_{\mathbb{R}}(I \backslash I_{(\infty)}) \leq \alpha \varepsilon^{\beta} \tag{IV.11}$$

where $\beta = \frac{3\nu - 2r + 2}{10\nu}$ and C, α depend on ν, r, n, σ_0, σ_1, σ_3, σ_4, σ_5 and σ_6. Finally, we have

$$\|\tau_{(\infty)} - \tau\|_{I_{(\infty)},\rho/2} = \mathcal{O}(\varepsilon^{4/5}), \qquad \|N_{(\infty)} - N\|_{I_{(\infty)},\rho/2} = \mathcal{O}(\varepsilon^{3/5}), \tag{IV.12}$$

and for $i = 1, \ldots, n-r$,

$$\|\lambda_{i,(\infty)} - \lambda_i\|_{I_{(\infty)}} = \mathcal{O}(\varepsilon^{4/5}). \tag{IV.13}$$

Remark IV.3.2. *We will see that, if $\|e\|_{I,\rho}$ and $\|R\|_{I,\rho}$ are small enough, hypothesis H_2 and H_5, together with suitable Diophantine conditions on ω and λ, imply that the matrix Ω_{N,N^*} is pure imaginary, approximately constant and close to diagonal (see Propositions IV.4.1 and IV.5.3 for details). In order to follow our approach for constructing an approximately symplectic basis along the torus, we assume that the average of this matrix is non-singular. According to this, it is clear that we can assume (after a suitable ordering and scaling of the columns) that* $\mathrm{diag}\ [\Omega_{N,N^*}]_{\mathbb{T}^r} = \mathrm{i}\mathrm{Id}_{n-r}$, *as it is done in hypothesis* H_3 *of Theorem IV.3.1.*

Remark IV.3.3. *As it is customary in parameterization methods —we encourage the reader to compare this result with those in [dlLGJV05, FdlLS09, JdlLZ99]— the conditions of Theorem IV.3.1 can be verified using information provided by the initial approximations. This fact is useful in the validation of numerical computations that consist in looking for trigonometric functions that satisfy invariance and reducibility equations approximately [HdlL07, JO]. Concretely, let us assume that for a given parameter $\mu_0 \in I$ we have computed approximations τ_{μ_0}, N_{μ_0} and Λ_{μ_0} satisfying the explicit conditions of Theorem IV.3.1 for certain $\omega \in \mathbb{R}^r$. Then, for many values of μ close to μ_0, there exist an elliptic quasi-periodic invariant torus nearby, whose normal frequencies are just slightly changed. Furthermore, the proof of Theorem IV.3.1 can be implemented in a computer, thus leading to efficient numerical algorithms for the computation of these objects (as it is done in [dlLHS] for the Lagrangian and hyperbolic cases).*

Remark IV.3.4. *Hypothesis* H_4 *is called* twist condition *because when applying this result in a perturbative setting it stands for the Kolmogorov non-degeneracy condition (see the computations performed for Hamiltonian* (IV.15) *below). Observe that in the Lagrangian case hypothesis $A_{1,\mu}$ reads as $A_{1,\mu} = G_{D\tau_\mu}^{-1} D\tau_\mu^\top T_{1,\mu} D\tau G_{D\tau_\mu}^{-1}$ for the same matrix $T_{1,\mu}$, thus recovering the condition in [dlLGJV05].*

Remark IV.3.5. *Let us assume that for $\mu = 0$ we have a true elliptic quasi-periodic invariant torus satisfying Diophantine and non-degeneracy conditions. In this case, it is expected that the measure of true invariant tori nearby is larger that the one predicted by Theorem IV.3.1. Actually, it is known that the complementary set $[-\mu_0, \mu_0] \backslash I_{(\infty)}$ has measure exponentially small when $\mu_0 \to 0$ (see for example [JV97a]). To obtain such estimates we would need to modify slightly some details of the proof performed here —but not the scheme— asking for Diophantine conditions as those used in [JS96, JV97b].*

Remark IV.3.6. *Theorem IV.3.1 can be extended to the case of exact symplectic maps. Actually, the parameterization approach in the context of maps is the main setting in [dlLGJV05, FdlLS09, GdlLH, JdlLZ99]. To this end, we should "translate" the computations performed here to the context of maps, following the "dictionary" of the previous references. Attemption should be taken in order to adapt the geometric conditions that we highlight in Remarks IV.4.5 and IV.4.7, which are not true for maps, but satisfied up to quadratic terms (this is enough for the convergence of the scheme).*

Remark IV.3.7. *It would be also interesting to extend the result in order to deal with symplectic vector fields or symplectic maps. Let us recall that a vector field* X *on a symplectic manifold with* 2*-form* Ω *is said to be symplectic if* $L_X\Omega = 0$*, i.e., if the* 2*-form is preserved along the flow of* X *(symplectic vector fields that are not Hamiltonian can be found for example in the context of magnetic fields). In this situation, the method of "translated torus" should be adapted as it is done in [FdlLS09] for the hyperbolic case. To this end, it must be taken into account that the cohomology of the torus must be compatible with the cohomology class of the contraction* $\Omega(\cdot, X)$.

Remark IV.3.8. *The scheme of the proof of Theorem IV.3.1 can be also used for proving the existence of reducible tori having some hyperbolic directions, under the assumption of first and second order Melnikov. In this case, we need to adapt the geometrical ideas in order to deal simultaneously with elliptic and hyperbolic directions. In this case, we have to be careful because hyperbolic tori are known to exist beyond the breakdown of reducibility (see [HdlL07]). Nevertheless, as we already pointed out in Section IV.1 we can approach the problem of partially elliptic tori by combining techniques in [FdlLS09] (for studying hyperbolic directions) together with those presented here (for studying elliptic directions).*

Remark IV.3.9. *As it was explained in the introduction, the scheme can be also adapted to deal with the classical Broer-Huitema-Takens approach (see [BHTB90]). On the one hand, this allows obtaning* $\mathcal{C}^\infty$*-Whitney regularity for the constructed tori, and on the other hand this permits to deal with degenerate cases where Kolmogorov condition does not hold, but we have other higher-order non-degeneracy conditions such as the so-called Rüssmann's non-degeneracy condition (see [Sev95]).*

Remark IV.3.10. *After the work in [Bou97, Bou98, Eli01, HY08, XY01] it is known that second order Melnikov conditions are not necessary for proving existence of lower dimensional tori. For example, Bourgain approached the problem without asking for reducibility, thus avoiding to ask for these non-resonance conditions. However, cumbersome multiescale analysis is required to approximate the solution of truncated cohomological equations, thus leading to a process which is not suitable for numerical implementations: at each step, one has to invert a large matrix which has a huge computational cost. Nevertheless, asking for reducibility we end up inverting a diagonal matrix in Fourier space (see Remark IV.3.11). Another approach to avoid second Melnikov conditions was proposed by Eliasson and consists in performing a far-from-identity transformation when we have to deal with such resonant frequencies. Concretely, if*

$\lambda \in \mathbb{R}^{n-r}$ does not satisfies second Melnikov conditions, then we can introduce new normal frequencies

$$\widetilde{\lambda}_j = \lambda_j - \mathrm{i}\langle m_j, \frac{\omega}{2}\rangle, \tag{IV.14}$$

and we can choose carefully the vectors $m_j \in \mathbb{Z}^r$ in such a way that second Melnikov conditions are satisfied (it is also recessary to work in the double covering $2\mathbb{T}^r = \mathbb{R}^r/(4\pi\mathbb{Z})^r$ of the torus). In order to enhance the main contribution of this work (which are the geometric ideas linked with parameterization methods), we ask for second Melnikov conditions thus simplifying these technical details —we pay the price of excluding a very small set of invariant tori. Nevertheless, when implementing numerically this method, the use of transformation (IV.14) *can be very useful (this was used in [HdlL07] to continue elliptic tori beyond their bifurcation to hyperbolic tori).*

Remark IV.3.11. *All the computations performed in the proof can be implemented very efficiently in a computer. For example, the solution of cohomological equations with constant coefficients and the computation of derivatives like $D\tau$ or $L_\omega\tau$, correspond to diagonal operators in Fourier space. Other algebraic manipulations can be performed efficiently in real space and there are very fast and robust FFT algorithms that allows moving from real (or complex) to Fourier space (and "vice versa"). Accordingly, if we approximate a torus by using N Fourier modes, we can implement an algorithm to compute the object with a cost of order $\mathcal{O}(N \log N)$ in time and $\mathcal{O}(N)$ in memory. We refer to the works [CdlL09, dlLHS] to analogous algorithms in several contexts. Therefore, this approach present significant advantages in contrast with methods which require to deal with large matrices, since they represent a cost of $\mathcal{O}(N^2)$ in memory and $\mathcal{O}(N^3)$ in time (we refer for example to [CJ00, JO]).*

Although one of the main features of both the formulation and the proof of Theorem IV.3.1 is that we do not require to write the problem in action-angle coordinates, we think that it can be illustrative to express this result for a close-to-integrable system, in order to clarify the meaning of hypotheses H_3 and H_4 in this context. Indeed, let us consider the following family of Hamiltonian systems written in action-angle-like coordinates $(\varphi, y, z) \in \mathbb{T}^r \times \mathbb{R}^r \times \mathbb{R}^{2(n-r)}$

$$h_\mu(\varphi, y, z) = h_0(y, z) + \mu f(\varphi, y, z) \tag{IV.15}$$

such that for $y = 0$, we have that $z = 0$ is an elliptic non-degenerate equilibrium for the system $h_0(y, z)$. This means that $\tau_0(\theta) = (\theta, 0, 0)$ gives a parameterization of an invariant torus of h_0 with basic frequencies $\omega = \mathrm{grad}_y h_0(0, 0) \in \mathbb{R}^r$. By performing a suitable change of variables in order to eliminate crossed quadratic terms in (y, z), we can assume that

$$h_0(y, z) = \langle \omega, y\rangle + \frac{1}{2}\langle y, Ay\rangle + \frac{1}{2}\langle z, Bz\rangle + \mathcal{O}_3(y, z)$$

close to $(y, z) = (0, 0)$, where A and B are symmetric matrices, such that

$$\mathrm{spec}\,(J_{n-r}B) = \{\mathrm{i}\lambda_1, \ldots, \mathrm{i}\lambda_{n-r}, -\mathrm{i}\lambda_1, \ldots, -\mathrm{i}\lambda_{n-r}\},$$

are the normal eigenvalues of the torus given by $\mathbb{T}^r \times \{0\} \times \{0\}$. The associated normal directions are given by the real and imaginary parts of the matrix of eigenvectors satisfying $J_{n-r}B\hat{N} = \hat{N}\Lambda$, where $\Lambda = \mathrm{diag}\,(\mathrm{i}\lambda) = \mathrm{diag}\,(\mathrm{i}\lambda_1, \ldots, \mathrm{i}\lambda_{n-r})$. Using symplectic properties, we can select this complex matrix in such a way that it satisfies $\hat{N}^\top J_{n-r}\hat{N}^* = \mathrm{iId}_{n-r}$.

Then, to apply Theorem IV.3.1 to the family of Hamiltonians h_μ given by (IV.15), for small $|\mu|$, we can consider the family of approximately elliptic invariant tori $\tau_\mu(\theta) = \tau_0(\theta) + \mathcal{O}(\mu)$ with normal frequencies $\lambda_\mu = \lambda + \mathcal{O}(\mu)$ and normal vectors $N_\mu(\theta) = (0, 0, \hat{N})^\top + \mathcal{O}(\mu)$, where the terms $\mathcal{O}(\mu)$ stand for the first orden corrections in μ —they can be computed by means of Lindstedt series or normal forms with respect to μ— that we need to check if the normal frequencies "move" as a function of μ, thus allowing to control the measure of the Cantor set $I_{(\infty)}$. This family satifies

$$\begin{aligned} L_\omega \tau_\mu(\theta) &= J\mathrm{grad}\, h_\mu(\tau_\mu(\theta)) + \mathcal{O}_2(\mu), \\ L_\omega N_\mu(\theta) &= J\mathrm{hess}\, h_\mu(\tau_\mu(\theta))N_\mu(\theta) - N_\mu(\theta)\Lambda_\mu + \mathcal{O}_2(\mu), \end{aligned}$$

and, for $\mu = 0$, we have

$$D\tau_0(\theta) = \begin{pmatrix} \mathrm{Id}_r \\ 0 \\ 0 \end{pmatrix}, \qquad N_0(\theta) = \begin{pmatrix} 0 \\ 0 \\ \hat{N} \end{pmatrix}, \qquad G^{-1}_{D\tau_0}(\theta) = \mathrm{Id}_r, \qquad \Omega_{N_0, N_0^*}(\theta) = \mathrm{iId}_{n-r}.$$

Moreover, it is not difficult to check that the matrix $A_{1,\mu}(\theta)$ in H_4 at $\mu = 0$ reads as $A_{1,0}(\theta) = A$, which implies that H_4 is equivalent to the standard (Kolmogorov) non-degeneracy condition for the unperturbed system.

IV.4 Overview and heuristics of the method

In this section we outline the main ideas of the presented approach emphasizing the geometric interpretation of our construction and highlighting the additional difficulties with respect to the Lagrangian and normally hyperbolic cases. First, in Section IV.4.1, we sketch briefly the proof of Theorem IV.3.1. Our aim is to highlight that, even though some parts of the proof involve quite cumbersome computations, the construction of the iterative procedure is fairly natural. Then, in Section IV.4.2 we focus on the geometric properties of the invariant and elliptic case that allow us to obtain approximate solutions for the equations derived in Section IV.4.1 associated to approximately invariant and elliptic tori.

IV.4.1 Sketch of the proof

Let $h : U \subset \mathbb{R}^{2n} \to \mathbb{R}$ be a Hamiltonian function and let us suppose that $\mathscr{T}$ is an approximately invariant and elliptic torus of basic frequencies $\omega \in \mathbb{R}^r$ and normal ones $\lambda \in \mathbb{R}^{n-r}$, satisfying non-resonance conditions (IV.4) and (IV.10). The translation of Definitions IV.2.5 and IV.2.9 into a functional setting is

$$\mathcal{F}(\tau) = e, \qquad \mathcal{G}(\tau, N, \Lambda) = R,$$

with $\Lambda = \mathrm{diag}\,(\mathrm{i}\lambda)$, where we have introduced the following operators

$$\begin{aligned} \mathcal{F}(\tau) &:= L_\omega \tau - J\mathrm{grad}\, h(\tau), \\ \mathcal{G}(\tau, N, \Lambda) &:= L_\omega N - J\mathrm{hess}\, h(\tau) N + N\Lambda. \end{aligned}$$

Then, we look for an embedding $\bar{\tau} : \mathbb{T}^r \to U$ and a set of normal vectors $\bar{N} : \mathbb{T}^r \to M_{2n\times(n-r)}(\mathbb{C})$, of normal frequencies $\bar{\lambda}$, satisfying

$$\mathcal{F}(\bar{\tau}) = 0, \qquad \mathcal{G}(\bar{\tau}, \bar{N}, \bar{\Lambda}) = 0,$$

with $\bar{\Lambda} = \mathrm{diag}\,(\mathrm{i}\bar{\lambda})$. Since these equations have triangular structure, we approach first the correction of the parameterization of the torus, i.e., we look for $\bar{\tau} = \tau + \Delta_\tau$ satisfying the above expressions. We write the first equation as

$$\mathcal{F}(\tau + \Delta_\tau) = e + L_\omega \Delta_\tau - J\mathrm{hess}\, h(\tau)\Delta_\tau + O_2(\Delta_\tau) = 0.$$

If we neglect terms $O_2(\Delta_\tau)$ we obtain the following linearized equation (Newton method)

$$L_\omega \Delta_\tau - J\mathrm{hess}\, h(\tau)\Delta_\tau = -e, \tag{IV.16}$$

that allows us to correct the invariance of the torus up to terms of second order in e. In a similar way, we look for $\bar{N} = N + \Delta_N$ and $\bar{\Lambda} = \Lambda + \Delta_\Lambda$ such that

$$\mathcal{G}(\bar{\tau}, \bar{N}, \bar{\Lambda}) = \hat{R} + L_\omega \Delta_N - J\mathrm{hess}\, h(\tau)\Delta_N + N\Delta_\Lambda + \Delta_N \Lambda + O_2(\Delta_N, \Delta_\Lambda) = 0,$$

where

$$\hat{R} = R + J\mathrm{hess}\, h(\tau) N - J\mathrm{hess}\, h(\bar{\tau}) N \tag{IV.17}$$

includes both the error in reducibility and the one introduced when correcting the torus (which is expected to be of order of the size of e). Hence, in order to apply one step of the Newton method to correct reducibility, we have to solve the following linearized equation for Δ_N and Δ_Λ

$$L_\omega \Delta_N - J\mathrm{hess}\, h(\tau)\Delta_N + N\Delta_\Lambda + \Delta_N \Lambda = -\hat{R}. \tag{IV.18}$$

For convenience, once we fix τ, N and Λ, we define the following differential operators (acting on vectors or matrices of $2n$ rows)

$$\mathcal{R}(\xi) := L_\omega \xi - J\mathrm{hess}\, h(\tau)\xi, \tag{IV.19}$$

$$\mathcal{S}(\xi, \eta) := \mathcal{R}(\xi) + N\eta + \xi\Lambda, \tag{IV.20}$$

so Equations (IV.16) and (IV.18) are equivalent to invert $\mathcal{R}$ and $\mathcal{S}$

$$\mathcal{R}(\Delta_\tau) = -e, \qquad \mathcal{S}(\Delta_N, \Delta_\Lambda) = -\hat{R}. \tag{IV.21}$$

As it was done in [dlLGJV05], the main idea is to use the geometric properties of the problem to prove that the linearized Equation (IV.16) can be transformed, using a suitable basis along the approximate torus, into a simpler linear equation —with constant coefficients— that can be approximately solved by means of Fourier series. Indeed, an approximate solution with an error of quadratic size in e and R is enough for the convergence of the scheme —the Newton method still converges quadratically if we have a good enough approximation of the Jacobian matrix. Under suitable conditions of non-resonance and non-degeneracy, iteration of this process leads to a quadratic scheme that allows us to overcome the effect of the small divisors of the problem. The main contribution of this work is to adapt this construction (that we describe next in a more precise way) to deal with Equations (IV.16) and (IV.18) simultaneously.

Let us discuss the construction of the basis mentioned above. In the Lagrangian case we only have to deal with Equation (IV.16) and the columns of the matrices $D\tau$ and $JD\tau G_{D\tau}^{-1}$ give us an approximately symplectic basis of $\mathbb{R}^{2n}$ at any point of the torus. Moreover, it turns out that $\mathcal{R}(D\tau) = 0 + \mathcal{O}(e)$ and $\mathcal{R}(JD\tau G_{D\tau}^{-1}) = D\tau A_1 + \mathcal{O}(e)$, where $A_1 : \mathbb{T}^n \to M_{n\times n}(\mathbb{R})$ is a symmetric matrix. Using this basis we can write the linearized equation (IV.16) in "triangular form" with respect to the projections of Δ_τ over $D\tau$ and $JD\tau G_{D\tau}^{-1}$, in such a way that the problem is reduced to solve two cohomological equations with constant coefficients. However, in the lower dimensional case the previous construction is not enough since we also have to take into account the normal directions of the torus. As mentioned in the introduction, this scheme has been recently adapted in [FdlLS09] for the normally hyperbolic case, without requiring reducibility of the normal variational equations. The main ingredient is that there exists a splitting between the center and the hyperbolic directions of the torus and we can reduce the study of Equation (IV.16) to the projections according to this splitting. The dynamics on the hyperbolic directions is characterized by asymptotic (geometric) growth conditions[2] —both in the future and in the past— and the linearized equation (IV.16) restricted to the center subspace follows as in the Lagragian case (now the ambient space is $\mathbb{R}^{2r}$).

In the normally elliptic context, we ask for reducibility in order to express equation (IV.16) in a simple form. Hence, we solve simultaneously equation (IV.18), thus obtaining a basis that reduces the normal variational equations of the torus to constants coefficients up to a quadratic error. In this case, the approximately (with an error of the order of the size of e and R) symplectic basis is obtained by completing the columns of $D\tau$, N and $\mathrm{i}N^*$ with the columns of a suitably constructed matrix $V : \mathbb{T}^r \to M_{2n\times r}(\mathbb{R})$. Basically, we take advantage of the fact that V satisfies $\mathcal{R}(V) = D\tau A_1$ modulo terms of order e and R, where $A_1 : \mathbb{T}^r \to M_{r\times r}(\mathbb{R})$ will be specified later on. Hence, we find approximately solutions for equations (IV.21) in terms of the

[2]Concretely, the solution for the equations projected into the hyperbolic directions are obtained by means of absolutely convergent power series. See details in [FdlLS09].

constructed basis as follows

$$\Delta_\tau = D\tau\Delta_1 + V\Delta_2 + N\Delta_3 + \mathrm{i}N^*\Delta_4$$
$$\Delta_N = D\tau P_1 + VP_2 + NP_3 + \mathrm{i}N^*P_4$$

where $\{\Delta_i\}, \{P_i\}$, with $i = 1, \ldots, 4$, are the solutions of cohomological equations (IV.39)-(IV.42), and (IV.44)- (IV.47), respectively. The correction Δ_Λ in the normal eigenvalues is determined from the compatibitity condition of these last equations.

Let us observe that in order to correct the reducibility of the torus we have to change slightly the normal directions and the normal frequencies. Since the normal frequencies λ are modified at each step of the process, we do not know in advance if they will satisfy the required Diophantine conditions for all steps —unless we have enough parameters to control the value of all of them simultaneously. To deal with this problem we require some control on the change of these frequencies, in such a way that we can remove parameters that give rise to resonant frequencies. Since at every step of the inductive process we are removing a dense set of parameters, this does not allow us to keep any kind of smooth dependence with respect to parameters (because now they move on a set of empty interior).

There are several methods in the literature to deal with this problem. The first approach was due to Arnold (see [Arn63]) and it consists in working, at every step of the inductive procedure, with a finite number of terms in the Fourier expansions ("ultraviolet cut-off"). Then, since we only need to deal with a finite number of resonances at every step, we can work on open sets of parameters and keep the smooth dependence on these steps. Another possibility is to consider Lipschitz parametric dependence and to check that this dependence is preserved along the iterative procedure (this is the method used in [JS92, JS96, JV97a, JV97b]). Lipschitz regularity suffices to control the measure of the resonant sets. In this work we follow the Lipschitz approach because it does not forces to modify, by the effect of the "ultraviolet cut-off", the geometric construction we have developed in the Diophantine case.

IV.4.2 Characterization of the invariant and reducible case

Our goal now is to formally "invert" the linear operators $\mathcal{R}$ given by (IV.19) and $\mathcal{S}$ given by (IV.20) —see Propositions IV.4.4 and IV.4.6, respectively— when the corresponding torus $\mathscr{T}$ is invariant and normally elliptic. In order to do this, first we characterize at a formal level some geometric properties of lower dimensional elliptic invariant tori. Later on, the same construction provided in this section will be used to study approximately invariant tori in order to solve equations in (IV.21) with a small error (controlled by the errors of invariance and reducibility).

All along this section we consider an r-dimensional normally elliptic quasi-periodic invariant torus $\mathscr{T}$ for a Hamiltonian h, of basic frequencies $\omega \in \mathbb{R}^r$ and normal frequencies $\lambda \in \mathbb{R}^{n-r}$ satisfying non-resonance conditions (IV.2), (IV.7) and (IV.8), i.e., we have

$$L_\omega \tau(\theta) = J\mathrm{grad}\, h(\tau(\theta)), \tag{IV.22}$$
$$L_\omega N(\theta) = J\mathrm{hess}\, h(\tau(\theta))N(\theta) - N(\theta)\Lambda, \tag{IV.23}$$

with $\Lambda = \mathrm{diag}\,(\mathrm{i}\lambda)$. We assume also that the matrices $G_{D\tau}(\theta)$ and $G_{N,N^*}(\theta)$ are invertible for every $\theta \in \mathbb{T}^r$. Then, we claim (see the proof if Proposition IV.4.1) that under these conditions Ω_{N,N^*} is constant, pure imaginary and diagonal. If we assume that this matrix is non-singular, then we can suppose that (see Remark IV.3.2)

$$\Omega_{N,N^*}(\theta) = \mathrm{i}\mathrm{Id}_{n-r}. \tag{IV.24}$$

Proposition IV.4.1. *Given $\mathscr{T}$ an invariant and elliptic torus as above, we define the matrix functions*

$$N_1(\theta) = N(\theta), \qquad N_2(\theta) = \mathrm{i}N^*(\theta),$$

and the real matrix

$$V(\theta) = JD\tau(\theta)G_{D\tau}^{-1}(\theta) + N_1(\theta)B_1(\theta) + N_2(\theta)B_2(\theta) + D\tau(\theta)B_3(\theta), \tag{IV.25}$$

where

$$B_1(\theta) = G_{N_2,D\tau}(\theta)G_{D\tau}^{-1}(\theta), \tag{IV.26}$$
$$B_2(\theta) = -G_{N_1,D\tau}(\theta)G_{D\tau}^{-1}(\theta), \tag{IV.27}$$
$$B_3(\theta) = \mathrm{Re}(G_{B_2,B_1}(\theta)). \tag{IV.28}$$

Then, the columns of the matrices $D\tau(\theta)$, $V(\theta)$, $N_1(\theta)$ and $N_2(\theta)$ form a symplectic basis for any $\theta \in \mathbb{T}^r$, in the sense that the matrices $\Omega_{D\tau}(\theta), \Omega_V(\theta), \Omega_{N_i}(\theta), \Omega_{D\tau,N_i}(\theta)$ and $\Omega_{N_i,V}(\theta)$ vanish, for $i = 1, 2$, and

$$\Omega_{N_2,N_1}(\theta) = \mathrm{Id}_{n-r}, \qquad \Omega_{V,D\tau}(\theta) = \mathrm{Id}_r.$$

Proof. To obtain the geometric properties associated to the matrices $D\tau$, N_1 and N_2 we proceed as in the proof of Lemma IV.2.3, where we proved that $\Omega_{D\tau} = 0$. Let us start studying the matrix Ω_{N_2,N_1} by computing

$$\begin{aligned} L_\omega \Omega_{N_2,N_1} &= L_\omega(N_2^\top J N_1) = (L_\omega N_2)^\top J N_1 + N_2^\top J L_\omega N_1 \\ &= (J\mathrm{hess}\, h(\tau)N_2 + N_2\Lambda)^\top J N_1 + N_2^\top J(J\mathrm{hess}\, h(\tau)N_1 - N_1\Lambda) \\ &= \Lambda N_2^\top J N_1 - N_2^\top J N_1 \Lambda = \Lambda\Omega_{N_2,N_1} - \Omega_{N_2,N_1}\Lambda. \end{aligned}$$

Then, if we expand Ω_{N_2,N_1} in Fourier series we obtain

$$\Big(\langle k, \omega\rangle - \lambda_i + \lambda_j\Big)(\widehat{\Omega}_{N_2,N_1})_k^{(i,j)} = 0,$$

where $(\Omega_{N_2,N_1})^{(i,j)}$ denotes the (i,j)-th entry of Ω_{N_2,N_1}. Recalling the non-resonance hypothesis $\langle k, \omega\rangle - \lambda_i + \lambda_j \neq 0$ (if $i \neq j$ or $k \neq 0$) we obtain that $(\widehat{\Omega}_{N_2,N_1})_k^{(i,j)} = 0$, for all

$k \in \mathbb{Z}^r \backslash \{0\}$, and $(\widehat{\Omega}_{N_2,N_1})_0^{(i,j)} = 0$ if $i \neq j$, so this matrix is constant and diagonal. Moreover, $\Omega_{N_2,N_1}^\top = \Omega_{N_2,N_1}^*$ so its entries are real. Finally, using hypothesis (IV.24) we write $\Omega_{N_2,N_1} = \mathrm{i}\Omega_{N^*,N} = -\mathrm{i}\Omega_{N,N^*}^\top = \mathrm{Id}_{n-r}$.

To prove that Ω_{N_1} vanishes we compute

$$L_\omega \Omega_{N_1} = -\Lambda \Omega_{N_1} - \Omega_{N_1}\Lambda,$$

in a similar way as above. Now, the Fourier coefficients of Ω_{N_1} satisfy the expression

$$\Big(\langle k, \omega\rangle + \lambda_i + \lambda_j\Big)(\widehat{\Omega}_{N_1})_k^{(i,j)} = 0,$$

so it turns out that all of them vanish (using the non-resonance conditions). Moreover, taking derivatives at Equation (IV.22) we obtain

$$L_\omega D\tau = J \operatorname{hess} h D\tau,$$

that together with Equation (IV.23), leads to

$$L_\omega \Omega_{D\tau,N_1} = -\Omega_{D\tau,N_1}\Lambda,$$

which implies that $\Omega_{D\tau,N_1} = 0$, since the Fourier coefficients of the component functions satisfy the equation

$$\Big(\langle k, \omega\rangle + \lambda_i\Big)(\widehat{\Omega}_{D\tau,N_1})_k^{(i,j)} = 0.$$

Finally, it is easy to see that $\Omega_{N_2} = -\Omega_{N_1}^*$ and $\Omega_{D\tau,N_2} = \mathrm{i}\Omega_{D\tau,N_1}^*$, so these matrices also vanish.

Next, we see that the columns of the (real) matrices $D\tau$, $JD\tau G_{D\tau}^{-1}$, $\mathrm{Re}(N)$ and $\mathrm{Im}(N)$ form a $\mathbb{R}$-basis of $\mathbb{R}^{2n}$. To this end, it suffices to check that the columns of $D\tau$, $JD\tau G_{D\tau}^{-1}$, N_1 and N_2 are $\mathbb{C}$-independent on $\mathbb{C}^{2n}$. Thus, let us consider a linear combination

$$D\tau a + JD\tau G_{D\tau}^{-1} b + N_1 c + N_2 d = 0,$$

for vector functions $a, b : \mathbb{T}^r \to \mathbb{C}^r$ and $c, d : \mathbb{T}^r \to \mathbb{C}^{n-r}$. Multiplying by $D\tau^\top$, $D\tau^\top J$, $N_2^\top J$ and $N_1^\top J$ and using the geometric properties proved above, we obtain the following system of equations

$$\underbrace{\begin{pmatrix} G_{D\tau} & 0 & G_{D\tau,N_1} & G_{D\tau,N_2} \\ 0 & -\mathrm{Id}_r & 0 & 0 \\ 0 & -G_{N_2,D\tau}G_{D\tau}^{-1} & \mathrm{Id}_{n-r} & 0 \\ 0 & -G_{N_1,D\tau}G_{D\tau}^{-1} & 0 & -\mathrm{Id}_{n-r} \end{pmatrix}}_{M_1} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \tag{IV.29}$$

where $\det M_1 = \det G_{D\tau} \neq 0$, so we conclude that $a = b = 0$ and $c = d = 0$.

To check that the matrix V is real, we use the expressions $N_1^* = -\mathrm{i}N_2$ and $B_1^* = \mathrm{i}B_2$ that are obtained in a straightforward way. Then, we compute $N_1^* B_1^* = -\mathrm{i}^2 N_2 B_2 = N_2 B_2$, thus concluding that $V^* = V$.

Finally, the following computations are straightforward

$$\begin{aligned}
\Omega_{D\tau,V} &= -\operatorname{Id}_r + \Omega_{D\tau,N_1} B_1 + \Omega_{D\tau,N_2} B_2 + \Omega_{D\tau} B_3 = -\operatorname{Id}_r, \\
\Omega_{N_1,V} &= -G_{N_1,D\tau} G_{D\tau}^{-1} + \Omega_{N_1} B_1 + \Omega_{N_1,N_2} B_2 + \Omega_{N_1.D\tau} B_3 \\
&= -G_{N_1,D\tau} G_{D\tau}^{-1} - B_2 = 0, \\
\Omega_{N_2,V} &= -G_{N_2,D\tau} G_{D\tau}^{-1} + B_1 = 0, \\
\Omega_V &= (G_{D\tau}^{-1} D\tau^\top J^\top + B_1^\top N_1^\top + B_2^\top N_2^\top + B_3^\top D\tau^\top) JV \\
&= G_{D\tau}^{-1} G_{D\tau,N_1} B_1 + G_{D\tau}^{-1} G_{N_2,D\tau} B_2 + B_3 - B_3^\top \\
&= -G_{B_2,B_1} + G_{B_1,B_2} + B_3 - B_3^\top \\
&= \mathrm{i}\operatorname{Im}(G_{B_1,B_2} - G_{B_2,B_1}) = 0.
\end{aligned}$$

In the last computation we used that V is real. □

Remark IV.4.2. *Notice that the matrix B_3 can be defined modulo the addition of a symmetric real matrix. This freedom can be used to ask for reducibility also in the "central direction" of the torus. Hence, instead of the matrix A_1 that appears in Lemma IV.4.3 we would obtain its average $[A_1]_{\mathbb{T}^r}$. Since this does not give us any significant advantage, we do not resort to this fact.*

In the invariant and reducible case, we characterize the action of $\mathcal{R}$ on $D\tau(\theta)$, $N_1(\theta)$ and $N_2(\theta)$ in a very simple way

$$\mathcal{R}(D\tau(\theta)) = 0, \qquad \mathcal{R}(N_1(\theta)) = -N_1(\theta)\Lambda, \qquad \mathcal{R}(N_2(\theta)) = N_2(\theta)\Lambda. \tag{IV.30}$$

The first expression follows immediately from equation (IV.22) —invariance— and the other ones from equation (IV.23) —reducibility. Moreover, we have the following result for $V(\theta)$.

Lemma IV.4.3. *Under the setting of Proposition IV.4.1, we have that*

$$\mathcal{R}(V(\theta)) = D\tau(\theta) A_1(\theta),$$

where $A_1 : \mathbb{T}^r \to M_{r\times r}(\mathbb{R})$ is given by the real symmetric matrix

$$A_1(\theta) = G_{D\tau}^{-1}(\theta) D\tau(\theta)^\top (T_1(\theta) + T_2(\theta) + T_2(\theta)^\top) D\tau(\theta) G_{D\tau}^{-1}(\theta), \tag{IV.31}$$

where

$$\begin{aligned}
T_1(\theta) &= J^\top \operatorname{hess} h(\tau(\theta)) J - \operatorname{hess} h(\tau(\theta)), \\
T_2(\theta) &= T_1 J \left[D\tau(\theta) G_{D\tau}(\theta)^{-1} D\tau(\theta)^\top - \operatorname{Id} \right] \operatorname{Re}(N_1(\theta) N_2(\theta)^\top).
\end{aligned} \tag{IV.32}$$

Proof. We only have to write the expression for $\mathcal{R}(V)$ in terms of the previously constructed symplectic basis

$$\mathcal{R}(V) = D\tau A_1 + VA_2 + N_1 A_3 + N_2 A_4, \tag{IV.33}$$

and then show that A_1 is given by (IV.31) and $A_2 = A_3 = A_4 = 0$. First, we use (IV.25) and (IV.30) to express $\mathcal{R}(V)$ as

$$\mathcal{R}(V) = \mathcal{R}(JD\tau G_{D\tau}^{-1}) + N_1(L_\omega B_1 - \Lambda B_1) + N_2(L_\omega B_2 + \Lambda B_2) + D\tau L_\omega B_3.$$

Then, multiplying at both sides of equation (IV.33) by $V^\top J$, $D\tau^\top J$, $N_2^\top J$, $N_1^\top J$ and using the symplectic properties of the basis we obtain the following expressions:

$$A_1 = L_\omega B_3 + V^\top J\mathcal{R}(JD\tau G_{D\tau}^{-1}), \tag{IV.34}$$

$$A_2 = -D\tau^\top J\mathcal{R}(JD\tau G_{D\tau}^{-1}), \tag{IV.35}$$

$$A_3 = L_\omega B_1 - \Lambda B_1 + N_2^\top J\mathcal{R}(JD\tau G_{D\tau}^{-1}), \tag{IV.36}$$

$$A_4 = L_\omega B_2 + \Lambda B_2 - N_1^\top J\mathcal{R}(JD\tau G_{D\tau}^{-1}). \tag{IV.37}$$

First, introducing that $B_1 = G_{N_2,D\tau} G_{D\tau}^{-1}$ into equation (IV.36), we obtain

$$\begin{aligned} A_3 &= L_\omega(N_2^\top D\tau G_{D\tau}^{-1}) - \Lambda N_2^\top D\tau G_{D\tau}^{-1} + N_2^\top J\mathcal{R}(JD\tau G_{D\tau}^{-1}) \\ &= L_\omega N_2^\top D\tau G_{D\tau}^{-1} + N_2^\top L_\omega(D\tau G_{D\tau}^{-1}) - \Lambda N_2^\top D\tau G_{D\tau}^{-1} \\ &\quad + N_2^\top J L_\omega(JD\tau G_{D\tau}^{-1}) + N_2^\top \operatorname{hess} h JD\tau G_{D\tau}^{-1} \\ &= (\underbrace{L_\omega N_2 - J\operatorname{hess} h N_2}_{\mathcal{R}(N_2)} - N_2\Lambda)^\top D\tau G_{D\tau}^{-1} = 0, \end{aligned}$$

where we used the property (IV.30) for N_2. Recalling that $N_2 = \mathrm{i}N_1^*$ we observe that $A_3^* = \mathrm{i}A_4$ so we also have $A_4 = 0$. In order to see that $A_2 = 0$ we expand the expression for $\mathcal{R}(JD\tau G_{D\tau}^{-1})$, obtaining

$$\begin{aligned} \mathcal{R}(JD\tau G_{D\tau}^{-1}) &= \mathcal{R}(JD\tau)G_{D\tau}^{-1} + JD\tau L_\omega(G_{D\tau}^{-1}) \\ &= -\operatorname{hess} h D\tau G_{D\tau}^{-1} - JD\tau G_{D\tau}^{-1}(D\tau^\top J\operatorname{hess} h - D\tau^\top \operatorname{hess} h J)D\tau G_{D\tau}^{-1} \\ &\quad - J\operatorname{hess} h JD\tau G_{D\tau}^{-1} = (\mathrm{Id}_r + JD\tau G_{D\tau}^{-1} D\tau^\top J)T_1 D\tau G_{D\tau}^{-1}, \end{aligned}$$

where we used expression (IV.32) for T_1. Then, on the one hand we have $D\tau^\top J\mathcal{R}(JD\tau G_{D\tau}^{-1}) = 0$ and on the other hand we have

$$\begin{aligned} V^\top J\mathcal{R}(JD\tau G_{D\tau}^{-1}) &= (B_3^\top D\tau^\top + B_2^\top N_2^\top + B_1^\top N_1^\top + G_{D\tau}^{-1} D\tau^\top J^\top)J\mathcal{R}(JD\tau G_{D\tau}^{-1}) \\ &= -B_2^\top(L_\omega B_1 - \Lambda B_1) + B_1^\top(L_\omega B_2 + \Lambda B_2) + G_{D\tau}^{-1} D\tau^\top T_1 D\tau G_{D\tau}^{-1}, \end{aligned}$$

where we have used equations (IV.36) and (IV.37) taking into account that $A_3 = A_4 = 0$. Finally, we introduce this last expression into (IV.34) and recall that $B_3 = \mathrm{Re}(G_{B_2,B_1})$ in order to obtain

$$\begin{aligned} A_1 &= \mathrm{Re}(L_\omega(B_2^\top B_1)) - B_2^\top(L_\omega B_1 - \Lambda B_1) + B_1^\top(L_\omega B_2 + \Lambda B_2) + G_{D\tau}^{-1} D\tau^\top T_1 D\tau G_{D\tau}^{-1} \\ &= \mathrm{Re}(L_\omega B_2^\top B_1 - B_2^\top L_\omega B_1 + 2B_2^\top \Lambda B_1) + G_{D\tau}^{-1} D\tau^\top T_1 D\tau G_{D\tau}^{-1}, \end{aligned}$$

where we used that $(B_1^\top(L_\omega B_2 + \Lambda B_2))^* = -B_2^\top(L_\omega B_1 - \Lambda B_1)$. Now we replace B_1 and B_2 by equations (IV.26) and (IV.27) respectively, and we expand the expression for $L_\omega B_1$ and $L_\omega B_2$ as follows (we also use that $B_1 = -\mathrm{i}B_2^*$)

$$\begin{aligned} L_\omega B_2 &= -L_\omega N_1^\top D\tau G_{D\tau}^{-1} - N_1^\top L_\omega(D\tau G_{D\tau}^{-1}) \\ &= \Lambda N_1^\top D\tau G_\tau^{-1} - N_1^\top \mathrm{hess}\, h J^\top D\tau G_{D\tau}^{-1} - N_1^\top L_\omega(D\tau G_{D\tau}^{-1}) \\ &= -\Lambda B_2 + N_1^\top J T_1 D\tau G_{D\tau}^{-1} - N_1^\top D\tau L_\omega(G_{D\tau}^{-1}) \\ &= -\Lambda B_2 + N_1^\top J T_1 D\tau G_{D\tau}^{-1} - N_1^\top D\tau G_{D\tau}^{-1} D\tau^\top J T_1 D\tau G_{D\tau}^{-1}. \\ L_\omega B_1 &= \Lambda B_1 - N_2^\top J T_1 D\tau G_{D\tau}^{-1} + N_2^\top D\tau G_{D\tau}^{-1} D\tau^\top J T_1 D\tau G_{D\tau}^{-1}. \end{aligned}$$

From this expressions we observe that term $2B_2^\top \Lambda B_1$ in A_1 is cancelled. Finally, since $(N_1 N_2^\top)^* = -N_2 N_1^\top = (-N_1 N_2^\top)^\top$, it turns out that $\mathrm{Re}(N_1 N_2^\top) = \mathrm{Re}((-N_1 N_2^\top)^\top)$ so we obtain the expression (IV.31) for A_1. □

Now we have all the ingredients for inverting formally the operator $\mathcal{R}$.

Proposition IV.4.4. *Under the setting of Proposition IV.4.1, we assume that the matrix A_1 given in* (IV.31) *satisfies the twist condition* $\det [A_1]_{\mathbb{T}^r} \neq 0$. *Then, given a function $e : \mathbb{T}^r \to \mathbb{R}^{2n}$ satisfying $\left[D\tau^\top J e\right]_{\mathbb{T}^r} = 0$, we obtain a formal solution for the equation*

$$\mathcal{R}(\Delta_\tau(\theta)) = L_\omega \Delta_\tau - J \mathrm{hess}\, h(\tau)\Delta_\tau = -e(\theta),$$

which is unique up to a terms in $\ker(\mathcal{R}) = \{D\tau A \,:\, A \in M_{r\times r}(\mathbb{R})\}$.

Proof. We express the unknown $\Delta_\tau(\theta)$ in terms of the constructed symplectic basis

$$\Delta_\tau = D\tau\Delta_1 + V\Delta_2 + N_1\Delta_3 + N_2\Delta_4, \tag{IV.38}$$

expand $\mathcal{R}(\Delta_\tau)$ and project to compute the functions $\{\Delta_i\}_{i=1,\ldots,4}$. Concretely, we have

$$\begin{aligned} \mathcal{R}(\Delta_\tau) &= \mathcal{R}(D\tau)\Delta_1 + \mathcal{R}(V)\Delta_2 + \mathcal{R}(N_1)\Delta_3 + \mathcal{R}(N_2)\Delta_4 \\ &\quad + D\tau L_\omega\Delta_1 + V L_\omega\Delta_2 + N_1 L_\omega\Delta_3 + N_2 L_\omega \Delta_4 \\ &= D\tau(L_\omega\Delta_1 + A_1\Delta_2) + V L_\omega\Delta_2 + N_1(L_\omega\Delta_3 - \Lambda\Delta_3) + N_2(L_\omega\Delta_4 + \Lambda\Delta_4). \end{aligned}$$

Multiplying at both sides of this expression by $V^\top J$, $D\tau^\top J$, $N_2^\top J$ and $N_1^\top J$, we obtain the following four cohomological equations:

$$L_\omega \Delta_1 + A_1 \Delta_2 = -V^\top J e, \tag{IV.39}$$

$$L_\omega \Delta_2 = D\tau^\top J e, \tag{IV.40}$$

$$L_\omega \Delta_3 - \Lambda \Delta_3 = -N_2^\top J e, \tag{IV.41}$$

$$L_\omega \Delta_4 + \Lambda \Delta_4 = N_1^\top J e. \tag{IV.42}$$

As $\left[D\tau^\top J e\right]_{\mathbb{T}^r} = 0$, the solution of equation (IV.40) is unique, up to an arbitrary average $[\Delta_2]_{\mathbb{T}^r}$, provided that the non-resonance condition (IV.2) holds. Then, using the non-degeneracy condition $\det [A_1]_{\mathbb{T}^r} \neq 0$, we choose

$$[\Delta_2]_{\mathbb{T}^r} = [A_1]_{\mathbb{T}^r}^{-1} \left(\left[-V^\top J e\right]_{\mathbb{T}^r} - \left[\widetilde{A}_1 \widetilde{\Delta}_2\right]_{\mathbb{T}^r} \right) \tag{IV.43}$$

in such a way that $\left[A_1 \Delta_2 + V^\top J e\right]_{\mathbb{T}^r} = 0$ so we have a unique solution for Δ_1 up to the freedom of fixing $[\Delta_1]_{\mathbb{T}^r}$. Actually, it is easy to check (IV.41) and (IV.42) have unique solution for Δ_3 and Δ_4 provided that the non-resonance condition (IV.10) is fulfilled. Moreover, since e is a real function, we conclude that $\Delta_3^* = \mathrm{i}\Delta_4$ and this allows us to guarantee that the expression (IV.38) is also real. □

Remark IV.4.5. *We will see that the compatibility condition* $\left[D\tau^\top J e\right]_{\mathbb{T}^r} = 0$ *is automatically fulfilled if* e *is the error in the invariance of the torus —see computations in* (IV.98).

Proposition IV.4.6. *Under the setting of Proposition IV.4.1, given a function* $\hat{R} : \mathbb{T}^r \mapsto M_{2n\times(n-r)}(\mathbb{C})$*, we obtain a solution for the equation*

$$\mathcal{S}(\Delta_N, \Delta_\Lambda) = \mathcal{R}(\Delta_N) + N\Delta_\Lambda + \Delta_N \Lambda = -\hat{R},$$

which is unique for Δ_Λ *and for* Δ_N *up to terms in* $\ker(\mathcal{S}) = \{ND \,:\, D = \mathrm{diag}\,(d), d \in \mathbb{C}^{n-r}\}$.

Proof. As before, we write the solution Δ_N of this equation in terms of the symplectic basis as

$$\Delta_N = D\tau P_1 + V P_2 + N_1 P_3 + N_2 P_4.$$

Then, we compute the action of $\mathcal{S}$ on the pair $(\Delta_N, \Delta_\Lambda)$, thus obtaining

$$\begin{aligned}
\mathcal{S}(\Delta_N, \Delta_\Lambda) &= \mathcal{R}(\Delta_N) + N_1 \Delta_\Lambda + \Delta_N \Lambda \\
&= \mathcal{R}(D\tau)P_1 + \mathcal{R}(V)P_2 + \mathcal{R}(N_1)P_3 + \mathcal{R}(N_2)P_4 \\
&\quad + D\tau L_\omega P_1 + V L_\omega P_2 + N_1 L_\omega P_3 + N_2 L_\omega P_4 + N_1 \Delta_\Lambda \\
&\quad + D\tau P_1 \Lambda + V P_2 \Lambda + N_1 P_3 \Lambda + N_2 P_4 \Lambda \\
&= D\tau(L_\omega P_1 + P_1 \Lambda + A_1 P_2) + V(L_\omega P_2 + P_2 \Lambda)
\end{aligned}$$

$$+ N_1(L_\omega P_3 + P_3\Lambda - \Lambda P_3 + \Delta_\Lambda) + N_2(L_\omega P_4 + P_4\Lambda + \Lambda P_4) = \hat{R}.$$

If we multiply this expression by $V^\top J$, $D\tau^\top J$, $N_2^\top J$ and $N_1^\top J$, we end up with the following four cohomological equations:

$$L_\omega P_1 + P_1\Lambda + A_1 P_2 = -V^\top J\hat{R}, \tag{IV.44}$$

$$L_\omega P_2 + P_2\Lambda = D\tau^\top J\hat{R}, \tag{IV.45}$$

$$L_\omega P_3 + P_3\Lambda - \Lambda P_3 = -N_2^\top J\hat{R} - \Delta_\Lambda, \tag{IV.46}$$

$$L_\omega P_4 + P_4\Lambda + \Lambda P_4 = N_1^\top J\hat{R}. \tag{IV.47}$$

Let us observe that, under the assumed non-resonance conditions (IV.7) and (IV.8), the only unavoidable resonances are those in the diagonal of the average of equation (IV.46), so we require that the diagonal of the average of the right-hand side of this equation vanishes. This is attained by fixing the correction of the normal eigenvalues $\Delta_\Lambda = -\mathrm{diag}\,[N_2^\top J\hat{R}]_{\mathbb{T}^r}$. Therefore, we obtain a unique solution P_1, P_2, P_3, P_4 and Δ_Λ —modulo terms in $\mathrm{diag}\,[P_3]_{\mathbb{T}^r}$— of this system of equations. □

Remark IV.4.7. *We will see that if $\hat{R}$ corresponds to the error in reducibility as defined in equation* (IV.17) *then the geometry imposes that the correction Δ_Λ is a pure imaginary diagonal matrix, thus preserving the elliptic normal behavior —see computations in* (IV.101).

IV.5 One step of the Newton method

In this section we perform one step of the Newton method to correct an approximately invariant and elliptic torus. To this end, we follow the scheme presented in Section IV.4.2 for the case of a true elliptic invariant torus. The main difficulty is that we have to handle with "noise" introduced by the approximately invariant an reducible character.

Proposition IV.5.1. *Let us consider a Hamiltonian $h : U \subset \mathbb{R}^{2n} \to \mathbb{R}$, where U is an open set, and a vector of basic frequencies $\omega \in \mathbb{R}^r$. Let us assume that the following hypotheses hold:*

- H_1 *The Hamiltonian h is real analytic and can be holomorphically extended to some complex neighborhood $\mathcal{U}$ of U. Moreover, we assume that $|h|_{\mathcal{C}^3,\mathcal{U}} \le \sigma_0$.*

- H_2 *There exists an approximate invariant and elliptic torus in the sense of Definitions IV.2.5 and IV.2.9, i.e., we have an embedding τ, a matrix function N and approximated normal eigenvalues $\Lambda = \mathrm{diag}\,(\mathrm{i}\lambda)$, with $\lambda \in \mathbb{R}^{n-r}$, satisfying*

$$L_\omega \tau(\theta) = J\mathrm{grad}\, h(\tau(\theta)) + e(\theta), \tag{IV.48}$$

$$L_\omega N(\theta) = J\mathrm{hess}\, h(\tau(\theta))N(\theta) - N(\theta)\Lambda + R(\theta), \tag{IV.49}$$

for certain error functions e and R, where the functions τ and N are analytic and can be holomorphically extended to $\Delta(\rho)$ for certain $0 < \rho < 1$, satisfying $\tau(\Delta(\rho)) \subset \mathcal{U}$. Assume also that we have constants σ_1, σ_2 such that

$$\|D\tau\|_\rho, \|N\|_\rho, \|G_{D\tau}^{-1}\|_\rho, \|G_{N,N^*}^{-1}\|_\rho < \sigma_1, \qquad \mathrm{dist}(\tau(\Delta(\rho)), \partial\mathcal{U}) > \sigma_2 > 0.$$

H_3 *We have* $\mathrm{diag}\,[\Omega_{N,N^*}]_{\mathbb{T}^r} = \mathrm{iId}_{n-r}$.

H_4 *The real symmetric matrix A_1 given by*

$$A_1(\theta) := G_{D\tau}^{-1}(\theta) D\tau(\theta)^\top (T_1(\theta) + T_2(\theta) + T_2(\theta)^\top) D\tau(\theta) G_{D\tau}^{-1}(\theta), \tag{IV.50}$$

where

$$T_1(\theta) := J^\top \mathrm{hess}\, h(\tau(\theta)) J - \mathrm{hess}\, h(\tau(\theta)), \tag{IV.51}$$

$$T_2(\theta) := T_1(\theta) J \left[D\tau(\theta) G_{D\tau}(\theta)^{-1} D\tau(\theta)^\top - \mathrm{Id}\right] \mathrm{Re}(\mathrm{i}N(\theta)N^*(\theta)^\top), \tag{IV.52}$$

satisfies the non-degeneracy (twist) condition $|\,[A_1]_{\mathbb{T}^r}^{-1}\,| < \sigma_1$.

H_5 *There exist constants σ_3, σ_4 such that the approximated normal frequencies $\lambda = (\lambda_1, \ldots, \lambda_{n-r})$ satisfy*

$$0 < \frac{\sigma_3}{2} < |\lambda_i| < \frac{\sigma_4}{2}, \qquad 0 < \sigma_3 < |\lambda_i \pm \lambda_j|,$$

for $i, j = 1, \ldots, n-r$, with $i \neq j$.

H_6 *The basic frequencies $\omega \in \mathbb{R}^r$ and the normal frequencies $\lambda \in \mathbb{R}^{n-r}$ satisfy Diophantine conditions* (IV.4) *and* (IV.10) *of (γ, ν)-type, for certain $0 < \gamma < 1$ and $\nu > r - 1$.*

Then, there exist a constant $\bar{\alpha} > 1$ depending on ν, r, n, $|\omega|$, σ_0, σ_1, σ_2, σ_3 and σ_4 such that if the following bounds are satisfied

$$\frac{\bar{\alpha}}{\gamma^4 \delta^{4\nu-1}} \left(\frac{\|e\|_\rho}{\delta} + \|R\|_\rho \right) < \min\{1, \sigma_1 - \sigma^*\}, \tag{IV.53}$$

$$\mathrm{dist}(\tau(\Delta(\rho)), \partial\mathcal{U}) - \frac{\bar{\alpha}}{\gamma^2 \delta^{2\nu}} \|e\|_\rho > \sigma_2, \tag{IV.54}$$

$$\min_{i \neq j} |\lambda_i \pm \lambda_j| - \frac{\bar{\alpha}}{\gamma^2 \delta^{2\nu-1}} \left(\frac{\|e\|_\rho}{\delta} + \|R\|_\rho \right) > \sigma_3, \tag{IV.55}$$

$$\min_i |\lambda_i| - \frac{\bar{\alpha}}{\gamma^2 \delta^{2\nu-1}} \left(\frac{\|e\|_\rho}{\delta} + \|R\|_\rho \right) > \frac{\sigma_3}{2}, \tag{IV.56}$$

$$\max_i |\lambda_i| + \frac{\bar{\alpha}}{\gamma^2 \delta^{2\nu-1}} \left(\frac{\|e\|_\rho}{\delta} + \|R\|_\rho \right) < \frac{\sigma_4}{2}, \tag{IV.57}$$

where

$$\sigma^* = \max\left\{\|D\tau\|_\rho, \|N\|_\rho, \|G_{D\tau}^{-1}\|_\rho, \|G_{N,N^*}^{-1}\|_\rho, |\,[A_1]_{\mathbb{T}^r}^{-1}\,|\right\},$$

for some $0 < \delta < \rho/4$*, then we have an approximate invariant and elliptic torus* $\bar{\mathscr{T}}$ *for* X_h *of the same basic frequencies* ω*, i.e., we have an embedding* $\bar\tau = \tau + \Delta_\tau$*, with* $\bar\tau(\mathbb{T}^r) = \bar{\mathscr{T}}$*, a matrix function* $\bar N = N + \Delta_N$*, which are analytic in* $\Delta(\rho - 2\delta)$ *and* $\Delta(\rho - 4\delta)$*, respectively, and approximated normal eigenvalues* $\bar\Lambda = \operatorname{diag}(\mathrm{i}\bar\lambda) = \Lambda + \Delta_\Lambda$*, with* $\bar\lambda \in \mathbb{R}^{n-r}$*, such that*

$$\begin{aligned} L_\omega \bar\tau(\theta) &= J \operatorname{grad} h(\bar\tau(\theta)) + \bar e(\theta), \\ L_\omega \bar N(\theta) &= J \operatorname{hess} h(\bar\tau(\theta)) \bar N(\theta) - \bar N(\theta)\bar\Lambda + \bar R(\theta). \end{aligned}$$

In addition, the following estimates hold

$$\|\Delta_\tau\|_{\rho-2\delta} \leq \frac{\bar\alpha}{\gamma^2\delta^{2\nu}} \|e\|_\rho, \tag{IV.58}$$

$$\|\bar e\|_{\rho-3\delta} \leq \frac{\bar\alpha}{\gamma^4\delta^{4\nu-1}} \left(\frac{\|e\|_\rho}{\delta} + \|R\|_\rho\right)\|e\|_\rho, \tag{IV.59}$$

$$|\Delta_\Lambda| \leq \frac{\bar\alpha}{\gamma^2\delta^{2\nu-1}} \left(\frac{\|e\|_\rho}{\delta} + \|R\|_\rho\right), \tag{IV.60}$$

$$\|\Delta_N\|_{\rho-4\delta} \leq \frac{\bar\alpha}{\gamma^4\delta^{4\nu-1}} \left(\frac{\|e\|_\rho}{\delta} + \|R\|_\rho\right), \tag{IV.61}$$

$$\|\bar R\|_{\rho-4\delta} \leq \frac{\bar\alpha}{\gamma^8\delta^{8\nu-2}} \left(\frac{\|e\|_\rho}{\delta} + \|R\|_\rho\right)^2, \tag{IV.62}$$

$$\|G_{D\bar\tau}^{-1} - G_{D\tau}^{-1}\|_{\rho-3\delta} \leq \frac{\bar\alpha}{\gamma^2\delta^{2\nu+1}} \|e\|_\rho, \tag{IV.63}$$

$$\|G_{\bar N,\bar N^*}^{-1} - G_{N,N^*}^{-1}\|_{\rho-4\delta} \leq \frac{\bar\alpha}{\gamma^4\delta^{4\nu-1}} \left(\frac{\|e\|_\rho}{\delta} + \|R\|_\rho\right), \tag{IV.64}$$

$$\|\,\left[\bar A_1\right]_{\mathbb{T}^r}^{-1} - [A_1]_{\mathbb{T}^r}^{-1}\,\|_{\rho-3\delta} \leq \frac{\bar\alpha}{\gamma^4\delta^{4\nu-1}} \left(\frac{\|e\|_\rho}{\delta} + \|R\|_\rho\right). \tag{IV.65}$$

Furthermore, the new objects satisfy the following conditions

$$\operatorname{dist}(\bar\tau(\Delta(\rho-2\delta)), \partial\mathcal{U}) > \sigma_2, \qquad \frac{\sigma_3}{2} < |\bar\lambda_j| < \frac{\sigma_4}{2}, \qquad \sigma_3 < |\bar\lambda_i \pm \bar\lambda_j|, \tag{IV.66}$$

for $i, j = 1, \ldots, n-r$*, with* $i \neq j$*, and*

$$\max\left\{\|D\bar\tau\|_{\rho-3\delta}, \|\bar N\|_{\rho-4\delta}, \|G_{D\bar\tau}^{-1}\|_{\rho-3\delta}, \|G_{\bar N,\bar N^*}^{-1}\|_{\rho-4\delta}, |\,\left[\bar A_1\right]_{\mathbb{T}^r}^{-1}\,|\right\} < \sigma_1, \tag{IV.67}$$

where $\bar A_1$ *corresponds to formulas* (IV.50), (IV.51) *and* (IV.52) *for* $\bar\tau$ *and* $\bar N$ *. Moreover, the columns of* $\bar N$ *are normalized in such a way that* $\operatorname{diag}\left[\Omega_{\bar N,\bar N^*}\right]_{\mathbb{T}^r} = \mathrm{i}\mathrm{Id}_{n-r}$.

To prove this result, we first construct an approximately symplectic basis along the torus following the ideas of Section IV.4.2. This is done in Proposition IV.5.3. The geometric properties of this basis will allow us to approximately invert the operators $\mathcal{R}$ and $\mathcal{S}$ —given by (IV.19) and (IV.20), respectively— as it is required to obtain the iterative result of Proposition IV.5.1. Basically, it turns out that the solutions of the cohomological equations derived in Section IV.4.2 are enough to get the desired result. Before that, we state the following standard result that allows us to control the small divisors.

Lemma IV.5.2 (Rüssmann estimates)**.** *Let $g : \mathbb{T}^r \to \mathbb{C}$ be an analytic function on $\Delta(\rho)$ and bounded in the closure. Given $\omega \in \mathbb{R}^r \backslash \{0\}$ and $d \in \mathbb{R}\backslash\{0\}$ we consider the sets of complex numbers $\{d_k^0\}_{k\in\mathbb{Z}^r\backslash\{0\}}$, $\{d_k^1\}_{k\in\mathbb{Z}^r}$ given by $d_k^0 = \langle k, \omega\rangle$, $d_k^1 = \langle k,\omega\rangle + d$, satisfying*

$$|d_k^0|, |d_k^1| \geq \gamma/|k|_1^\nu, \qquad k \in \mathbb{Z}^r\backslash\{0\}$$

for certain $\gamma > 0$ and $\nu > r-1$. Then, the functions f^0 and f^1 whose Fourier coefficients are given by

$$\hat{f}_k^0 = \hat{g}_k/d_k^0, \qquad k \in \mathbb{Z}^r\backslash\{0\}, \qquad \hat{f}_0^0 = 0,$$
$$\hat{f}_k^1 = \hat{g}_k/d_k^1, \qquad k \in \mathbb{Z}^r,$$

satisfy

$$\|f^0\|_{\rho-\delta} \leq \frac{\alpha_0}{\gamma\delta^\nu}\|g\|_\rho,$$
$$\|f^1\|_{\rho-\delta} \leq \left(\frac{1}{|d|} + \frac{\alpha_0}{\gamma\delta^\nu}\right)\|g\|_\rho,$$

for any $\delta \in (0, \min\{1,\rho\})$, where $\alpha_0 \geq 1$ is a constant depending on r and ν.

Proof. We can control the functions $\widetilde{f^i}(\theta) = f^i(\theta) - [f^i]_{\mathbb{T}^r}$ as

$$\|\widetilde{f^i}\|_{\rho-\delta} \leq \sum_{k\in\mathbb{Z}^r\backslash\{0\}} \frac{|\hat{g}_k|}{|d_k^i|}\mathrm{e}^{|k|_1(\rho-\delta)} \leq \left(\sum_{k\in\mathbb{Z}^r\backslash\{0\}} |\hat{g}_k|^2 \mathrm{e}^{2|k|_1\rho}\right)^{1/2}\left(\sum_{k\in\mathbb{Z}^r\backslash\{0\}} \frac{1}{|d_k^i|^2}\mathrm{e}^{-2|k|_1\delta}\right)^{1/2},$$

for $i = 0, 1$, where we used Cauchy-Schwarz inequality. On the one hand, it is not difficult to see —using Bessel's inequality, see details in [Rüs75]— that the first term can be bounded by

$$\sum_{k\in\mathbb{Z}^r\backslash\{0\}} |\hat{g}_k|^2\mathrm{e}^{2|k|_1\rho} \leq 2^r\|\widetilde{g}\|_\rho^2,$$

and on the other hand, the second term is controlled by estimating the sum

$$\sum_{k\in\mathbb{Z}^r\backslash\{0\}} \frac{1}{|d_k^i|^2}\mathrm{e}^{-2|k|_1\delta} = \sum_{l=1}^{\infty}\left(\sum_{\substack{k\in\mathbb{Z}^r\backslash\{0\}\\ |k|_1\leq l}} \frac{1}{|d_k^i|^2}\right)\mathrm{e}^{-2l\delta}. \qquad \text{(IV.68)}$$

Now, we study in detail the case of d_k^1 (the case of d_k^0 is analogous). First, we observe that the divisors $d_k^1 = \langle k, \omega \rangle + d$ satisfy $d_{k_1} \neq d_{k_2}$ if $k_1 \neq k_2$. Then, given $l \in \mathbb{N}$, we define

$$\mathcal{D}_l = \{k \in \mathbb{Z}^r \backslash \{0\} \,:\, |k|_1 \leq l \text{ and } d_k^1 > 0\}$$

and we sort the divisors according to $0 < d_{k_1} < \ldots < d_{k_{\#\mathcal{D}_l}}$ with $k_j \in \mathcal{D}_l$, for $j = 1, \ldots, \#\mathcal{D}_l$. Then, we observe that (since $|k_j - k_{j-1}| \leq 2l$)

$$d_{k_j}^1 - d_{k_{j-1}}^1 = |\langle k_j - k_{j-1}, \omega \rangle| \geq d_{2l,\min}^0, \tag{IV.69}$$

where the have introduced the notation

$$d_{l,\min}^i := \min_{\substack{k \in \mathbb{Z}^r \backslash \{0\} \\ |k|_1 \leq l}} |d_k^i|.$$

From expression (IV.69) we obtain recursively

$$d_{k_j}^1 = d_{k_{j-1}}^1 + d_{k_j}^1 - d_{k_{j-1}}^1 \geq d_{k_{j-1}}^1 + d_{2l,\min}^0 \geq d_{l,\min}^1 + (j-1) d_{2l,\min}^0.$$

Then, using that $d_{2l,\min}^0 \geq \gamma/(2l)^\nu$ and $d_{l,\min}^1 \geq \gamma/l^\nu$, we have

$$\sum_{j=1}^{\#\mathcal{D}_l} \frac{1}{(d_{k_j}^1)^2} \leq \sum_{j=1}^{\#\mathcal{D}_l} \frac{1}{(d_{l,\min}^1 + (j-1) d_{2l,\min}^0)^2} \leq \sum_{j=1}^{\infty} \frac{l^{2\nu}}{\gamma^2 (1 + (j-1) 2^{-\nu})^2} \leq \frac{\alpha(\nu)}{\gamma^2} l^{2\nu},$$

and using a similar argument for $d_{k_j}^1 < 0$, we obtain

$$\sum_{\substack{k \in \mathbb{Z}^r \backslash \{0\} \\ |k|_1 \leq l}} \frac{1}{|d_k^1|^2} \leq \frac{2\alpha(\nu)}{\gamma^2} l^{2\nu},$$

so we can control the sum (IV.68) as follows

$$\begin{aligned}
\sum_{k \in \mathbb{Z}^r \backslash \{0\}} \frac{1}{|d_k^i|^2} \mathrm{e}^{-2|k|_1 \delta} &= \sum_{l=1}^{\infty} \Bigg(\sum_{\substack{k \in \mathbb{Z}^r \backslash \{0\} \\ |k|_1 \leq l}} \frac{1}{|d_k^i|^2} \Bigg) (\mathrm{e}^{-2l\delta} - \mathrm{e}^{-2(l+1)\delta}) \\
&\leq \sum_{l=1}^{\infty} \frac{2\delta \alpha(\nu)}{\gamma^2} \int_l^{l+1} x^{2\nu} \mathrm{e}^{-2\delta x} dx \leq \frac{\alpha(\nu)}{\gamma^2 (2\delta)^{2\nu}} \Gamma(2\nu + 1).
\end{aligned}$$

Combining the obtained expressions —and using that $|\,[f^1]_{\mathbb{T}^r}\,| = |\hat{g}_0|/|d|$— we end up with the stated estimates. ◻

Proposition IV.5.3. *Under the same notations and assumptions of Proposition IV.5.1, we define the matrix functions*

$$N_1(\theta) = N(\theta), \qquad N_2(\theta) = \mathrm{i}N^*(\theta),$$

and the real analytic matrix

$$V(\theta) = JD\tau(\theta)G_{D\tau}^{-1}(\theta) + N_1(\theta)B_1(\theta) + N_2(\theta)B_2(\theta) + D\tau(\theta)B_3(\theta), \tag{IV.70}$$

where B_1, B_2 and B_3 are given by (IV.26), (IV.27) *and* (IV.28), *respectively. Then, for any $0 < \delta < \rho/2$ the following estimates hold:*

$$\|\Omega_{D\tau}\|_{\rho-2\delta} \leq \frac{\hat{\alpha}}{\gamma\delta^{\nu+1}}\|e\|_\rho, \tag{IV.71}$$

$$\|\Omega_{N_i}\|_{\rho-\delta} \leq \frac{\hat{\alpha}}{\gamma\delta^{\nu}}\|R\|_\rho, \tag{IV.72}$$

$$\|\Omega_{D\tau,N_i}\|_{\rho-2\delta} \leq \frac{\hat{\alpha}}{\gamma\delta^{\nu}}\left(\frac{\|e\|_\rho}{\delta} + \|R\|_\rho\right), \tag{IV.73}$$

$$\|\Omega_{N_2,N_1} - \mathrm{Id}_{n-r}\|_{\rho-\delta} \leq \frac{\hat{\alpha}}{\gamma\delta^{\nu}}\|R\|_\rho, \tag{IV.74}$$

$$\|\Omega_{V,D\tau} - \mathrm{Id}_r\|_{\rho-2\delta} \leq \frac{\hat{\alpha}}{\gamma\delta^{\nu}}\left(\frac{\|e\|_\rho}{\delta} + \|R\|_\rho\right), \tag{IV.75}$$

$$\|\Omega_{V,N_i}\|_{\rho-2\delta} \leq \frac{\hat{\alpha}}{\gamma\delta^{\nu}}\left(\frac{\|e\|_\rho}{\delta} + \|R\|_\rho\right), \tag{IV.76}$$

$$\|\Omega_{V}\|_{\rho-2\delta} \leq \frac{\hat{\alpha}}{\gamma\delta^{\nu}}\left(\frac{\|e\|_\rho}{\delta} + \|R\|_\rho\right), \tag{IV.77}$$

for $i = 1, 2$, where $\hat{\alpha} > 1$ is a constant depending on ν, r, n, $|\omega|$, σ_0, σ_1, σ_3 and σ_4. Furthermore, if the errors $\|e\|_\rho$ and $\|R\|_\rho$ satisfy

$$\frac{\hat{\alpha}}{\gamma\delta^{\nu}}\left(\frac{\|e\|_\rho}{\delta} + \|R\|_\rho\right) \leq \frac{1}{2}, \tag{IV.78}$$

then the columns of $D\tau(\theta)$, $V(\theta)$, $N_1(\theta)$, $N_2(\theta)$ form an approximately symplectic basis for every $\theta \in \mathbb{T}^r$. In addition, it turns out that the action of the operator $\mathcal{R}$ given in (IV.19) *on V is expressed in terms of this basis as*

$$\mathcal{R}(V(\theta)) = D\tau(\theta)(A_1(\theta) + A_1^+(\theta)) + V(\theta)A_2^+(\theta) + N_1(\theta)A_3^+(\theta) + N_2(\theta)A_4^+(\theta), \tag{IV.79}$$

where A_1 is the matrix (IV.50) *and A_1^+, A_2^+, A_3^+ and A_4^+ satisfy the estimate*

$$\|A_i^+\|_{\rho-2\delta} \leq \frac{\hat{\alpha}}{\gamma\delta^{\nu+1}}\left(\frac{\|e\|_\rho}{\delta} + \|R\|_\rho\right), \tag{IV.80}$$

for $i = 1, \ldots, 4$.

Proof. For the sake of simplicity, we redefine (enlarge) the constant $\hat{\alpha}$ along the proof to meet the different conditions given in the statement. For example, we observe that there exist a constant $\hat{\alpha} > 0$, depending on r, n, $|\omega|$, σ_0 and σ_1, such that

$$\|B_i\|_\rho, \|T_1\|_\rho, \|T_2\|_\rho, \|A_1\|_\rho, \|V\|_\rho \leq \hat{\alpha}, \qquad \|L_\omega B_i\|_{\rho-\delta} \leq \frac{\hat{\alpha}}{\delta}, \tag{IV.81}$$

for $i = 1, 2, 3$ —we recall that T_1 and T_2 are given in (IV.51) and (IV.52), respectively. Now we take derivatives at both sides of the approximated invariance equation in (IV.48) and we read the reducibility equations in (IV.49) for N_1 and N_2

$$L_\omega D\tau = J\mathrm{hess}\, h(\tau) D\tau + De, \tag{IV.82}$$
$$L_\omega N_1 = J\mathrm{hess}\, h(\tau) N_1 - N_1\Lambda + R, \tag{IV.83}$$
$$L_\omega N_2 = J\mathrm{hess}\, h(\tau) N_2 + N_2\Lambda + \mathrm{i}R^*. \tag{IV.84}$$

Using the previous expressions, we compute the derivate L_ω of the matrices $\Omega_{D\tau}$, Ω_{N_1}, $\Omega_{D\tau,N_1}$ and Ω_{N_2,N_1} thus obtaining

$$L_\omega(\Omega_{D\tau}) = \Omega_{De,D\tau} + \Omega_{D\tau,De}, \tag{IV.85}$$
$$L_\omega(\Omega_{N_1}) = -\Lambda\Omega_{N_1} - \Omega_{N_1}\Lambda + \Omega_{R,N_1} + \Omega_{N_1 R}, \tag{IV.86}$$
$$L_\omega(\Omega_{D\tau,N_1}) = -\Omega_{D\tau,N_1}\Lambda + \Omega_{De,N_1} + \Omega_{D\tau,R}, \tag{IV.87}$$
$$L_\omega(\Omega_{N_2,N_1}) = \Lambda\Omega_{N_2,N_1} - \Omega_{N_2,N_1}\Lambda + \mathrm{i}\Omega_{R^*,N_1} + \Omega_{N_2,R}. \tag{IV.88}$$

First, we get estimate (IV.71) for $\Omega_{D\tau}$ by applying Lemma IV.5.2 to the (i,j)-component of $\Omega_{D\tau}$ obtained from (IV.85), i.e., taking $d_k^0 = \langle \omega, k \rangle$ and $g = -\mathrm{i}(\Omega_{De,D\tau} + \Omega_{D\tau,De})^{(i,j)}$ that (using Cauchy estimates) is analytic in $\Delta(\rho-\delta)$. Moreover, since $[\Omega_{D\tau}]_{\mathbb{T}^r} = 0$ (see Remark IV.2.2), we obtain

$$\|\Omega_{D\tau}\|_{\rho-2\delta} \leq \frac{\alpha_0}{\gamma\delta^\nu}\|g\|_{\rho-\delta} \leq \frac{\hat{\alpha}}{\gamma\delta^{\nu+1}}\|e\|_\rho.$$

Then, we proceed in a similar way to get (IV.72) for N_1, by applying Lemma IV.5.2 to the (i,j)-component of Ω_{N_1} obtained from (IV.86), i.e., taking $d_k^1 = \langle \omega, k \rangle + \lambda_i + \lambda_j$ and $g = -\mathrm{i}(\Omega_{R,N_1} + \Omega_{N_1 R})^{(i,j)}$, analytic in $\Delta(\rho)$. To bound the average, we use hypothesis H_4 of Proposition IV.5.1.

$$\|\Omega_{N_1}\|_{\rho-\delta} \leq \left(\frac{1}{\min_{i,j}|\lambda_i + \lambda_j|} + \frac{\alpha_0}{\gamma\delta^\nu}\right)\|g\|_\rho \leq \left(\frac{1}{\sigma_3} + \frac{\alpha_0}{\gamma\delta^\nu}\right) \leq \frac{\hat{\alpha}}{\gamma\delta^\nu}\|R\|_\rho,$$

Analogous computations from Equations (IV.87) and (IV.88) allow us to obtain estimate (IV.73) for N_1 and (IV.74). Of course, to obtain (IV.74) we resort to the hypothesis $\mathrm{diag}\,[\Omega_{N,N^*}]_{\mathbb{T}^r} = \mathrm{i}\mathrm{Id}_{n-r}$ in H_3 of Proposition IV.5.1. The corresponding estimates (IV.72) and (IV.73) for N_2 are straightforward using that $\Omega_{N_2} = -\Omega^*_{N_1}$ and $\Omega_{D\tau,N_2} = \mathrm{i}\Omega^*_{D\tau,N_1}$.

Next we show that the columns of $D\tau$, $JD\tau G_{D\tau}^{-1}$, $\mathrm{Re}(N_1)$ and $\mathrm{Im}(N_1)$ form a $\mathbb{R}$-basis of $\mathbb{R}^{2n}$. As in the proof of Proposition IV.4.1, we consider a linear combination

$$D\tau a + JD\tau G_{D\tau}^{-1} b + N_1 c + N_2 d = 0,$$

for functions $a, b : \mathbb{T}^r \to \mathbb{C}^r$ and $c, d : \mathbb{T}^r \to \mathbb{C}^{n-r}$. We project this equation multiplying by $D\tau^\top$, $D\tau^\top J$, $N_2^\top J$ and $N_1^\top J$, thus obtaining

$$\left(M_1 + \underbrace{\begin{pmatrix} 0 & \Omega_{D\tau} G_{D\tau}^{-1} & 0 & 0 \\ \Omega_{D\tau} & 0 & \Omega_{D\tau,N_1} & \Omega_{D\tau,N_2} \\ \Omega_{N_2,D\tau} & 0 & \Omega_{N_2,N_1} - \mathrm{Id}_{n-r} & \Omega_{N_2} \\ \Omega_{N_1,D\tau} & 0 & \Omega_{N_1} & \Omega_{N_1,N_2} + \mathrm{Id}_{n-r} \end{pmatrix}}_{M_2} \right) \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

where M_1 is the same matrix that appears in equation (IV.29). Now, we have to invert the matrix $M_1 + M_2 = M_1(\mathrm{Id} + M_1^{-1} M_2)$, where

$$M_1^{-1} = \begin{pmatrix} G_{D\tau}^{-1} & M_{1,2} & -G_{D\tau}^{-1} G_{D\tau,N_1} & G_{D\tau}^{-1} G_{D\tau,N_2} \\ 0 & -\mathrm{Id}_r & 0 & 0 \\ 0 & -G_{N_2,D\tau} G_{D\tau}^{-1} & \mathrm{Id}_{n-r} & 0 \\ 0 & G_{N_1,D\tau} G_{D\tau}^{-1} & 0 & -\mathrm{Id}_{n-r} \end{pmatrix},$$

with $M_{1,2} = G_{D\tau}^{-1}(G_{D\tau,N_1} G_{N_2,D\tau} - G_{D\tau,N_2} G_{N_1,D\tau}) G_{D\tau}^{-1}$, so it is clear that $\|M_1^{-1}\|_\rho \le \hat{\alpha}$. By means of Neumann series we obtain

$$\|(\mathrm{Id} + M_1^{-1} M_2)^{-1}\|_{\rho-2\delta} \le \frac{1}{1 - \|M_1^{-1} M_2\|_{\rho-2\delta}},$$

that it is well posed since (using bounds (IV.71)-(IV.74))

$$\|M_1^{-1} M_2\|_{\rho-2\delta} \le \frac{\hat{\alpha}}{\gamma \delta^\nu} \left(\frac{\|e\|_\rho}{\delta} + \|R\|_\rho \right) \le \frac{1}{2},$$

and applying hypothesis (IV.78). Then, it must be $a = b = 0$ and $c = d = 0$ along $\mathbb{T}^r$.

Now, we replace the columns of $JD\tau G_{D\tau}^{-1}$ by those of (IV.70) and we characterize the fact that the new basis is approximately symplectic. It is straightforward to compute

$$\begin{aligned} \Omega_{D\tau,V} &= -\mathrm{Id}_r + \Omega_{D\tau,N_1} B_1 + \Omega_{D\tau,N_2} B_2 + \Omega_{D\tau} B_3, \\ \Omega_{N_1,V} &= -G_{N_1,D\tau} G_{D\tau}^{-1} + \Omega_{N_1} B_1 + \Omega_{N_1,N_2} B_2 + \Omega_{N_1,D\tau} B_3 \\ &= \Omega_{N_1} B_1 + (\Omega_{N_1,N_2} + \mathrm{Id}_{n-r}) B_2 + \Omega_{N_1,D\tau} B_3, \\ \Omega_V &= B_3^\top (\Omega_{D\tau,V} + \mathrm{Id}_r) + B_1^\top \Omega_{N_1,V} + B_2^\top \Omega_{N_2,V} + G_{D\tau}^{-1} \Omega_{D\tau} G_{D\tau}^{-1}, \end{aligned}$$

and $\Omega_{N_2,V} = \mathrm{i}\Omega^*_{N_1,V}$. Then, estimates (IV.75)-(IV.77) follow from (IV.71)-(IV.74) and (IV.81).

Let us characterize the action of the linear operator $\mathcal{R}$ on the elements of this basis. By hypothesis, we immediately have that

$$\mathcal{R}(D\tau) = De, \qquad \mathcal{R}(N_1) = -N_1\Lambda + R, \qquad \mathcal{R}(N_2) = N_2\Lambda + \mathrm{i}R^*, \tag{IV.89}$$

and we have to see that if we write

$$\mathcal{R}(V) = D\tau(A_1 + A_1^+) + VA_2^+ + N_1A_3^+ + N_2A_4^+,$$

where A_1 is the matrix (IV.31), then the functions A_1^+, A_2^+, A_3^+ and A_4^+ are small —i.e., they satisfy (IV.80). To this end, expanding $\mathcal{R}(V)$ in the previous expression as

$$\begin{aligned}\mathcal{R}(V) = {} & \mathcal{R}(JD\tau G_{D\tau}^{-1}) + N_1(L_\omega B_1 - \Lambda B_1) + N_2(L_\omega B_2 + \Lambda B_2) + D\tau L_\omega B_3 \\ & + RB_1 + \mathrm{i}R^*B_2 + DeB_3,\end{aligned}$$

and multiplying at both sides of this equation by $V^\top J$, $D\tau^\top J$, $N_2^\top J$ and $N_1^\top J$, we obtain the linear system

$$\big(\mathrm{Id} + M_3\big)\begin{pmatrix} A_1^+ \\ A_2^+ \\ A_3^+ \\ A_4^+ \end{pmatrix} = \begin{pmatrix} C_1 \\ C_2 \\ C_3 \\ C_4 \end{pmatrix}, \tag{IV.90}$$

where

$$M_3 = \begin{pmatrix} \Omega_{V,D\tau} - \mathrm{Id}_r & \Omega_V & \Omega_{V,N_1} & \Omega_{V,N_2} \\ -\Omega_{D\tau} & \Omega_{V,D\tau} - \mathrm{Id}_r & -\Omega_{D\tau,N_1} & -\Omega_{D\tau,N_2} \\ \Omega_{N_2,D\tau} & \Omega_{N_2,V} & \Omega_{N_2,N_1} - \mathrm{Id}_{n-r} & \Omega_{N_2} \\ -\Omega_{N_1,D\tau} & -\Omega_{N_1,V} & -\Omega_{N_1} & \Omega_{N_2,N_1} - \mathrm{Id}_{n-r} \end{pmatrix} \tag{IV.91}$$

and the functions C_1, C_2, C_3 and C_4 have the following form

$$\begin{aligned}
C_1 = {} & \overbrace{V^\top J\mathcal{R}(JD\tau G_{D\tau}^{-1}) + L_\omega B_3 - A_1}^{C_1^+} + \Omega_{V,N_1}(L_\omega B_1 - \Lambda B_1) + \Omega_{V,N_2}(L_\omega B_2 + \Lambda B_2) \\
& + V^\top J(RB_1 + \mathrm{i}R^*B_2 + DeB_3) + (\Omega_{V,D\tau} - \mathrm{Id}_r)(L_\omega B_3 - A_1), \\
C_2 = {} & \overbrace{-D\tau^\top J\mathcal{R}(JD\tau G_{D\tau}^{-1})}^{C_2^+} - \Omega_{D\tau,N_1}(L_\omega B_1 - \Lambda B_1) - \Omega_{D\tau,N_2}(L_\omega B_2 + \Lambda B_2) \\
& - D\tau^\top J(RB_1 + \mathrm{i}R^*B_2 + DeB_3) + \Omega_{D\tau}(A_1 - L_\omega B_3), \\
C_3 = {} & \overbrace{N_2^\top J\mathcal{R}(JD\tau G_{D\tau}^{-1}) + L_\omega B_1 - \Lambda B_1}^{C_3^+} + \Omega_{N_2}(L_\omega B_2 + \Lambda B_2) + \Omega_{N_2,D\tau}(L_\omega B_3 - A_1) \\
& + (\Omega_{N_2,N_1} - \mathrm{Id}_{n-r})(L_\omega B_1 - \Lambda B_1) + N_2^\top J(RB_1 + \mathrm{i}R^*B_2 + DeB_3),
\end{aligned}$$

$$C_4 = \overbrace{-N_1^\top J\mathcal{R}(JD\tau G_{D\tau}^{-1}) + L_\omega B_2 + \Lambda B_2}^{C_4^+} - \Omega_{N_1}(L_\omega B_1 - \Lambda B_1) + \Omega_{N_1,D\tau}(A_1 - L_\omega B_3)$$
$$- (\Omega_{N_1,N_2} + \mathrm{Id}_{n-r})(L_\omega B_2 + \Lambda B_2) - N_1^\top J(RB_1 + \mathrm{i}R^* B_2 + DeB_3),$$

and we observe that $C_4 = -\mathrm{i}C_3^*$ and $C_4^+ = -\mathrm{i}(C_3^+)^*$. Apart from C_1^+, C_2^+, C_3^+ and C_4^+, the size of the other terms that appear in the above expressions are easily controlled in terms of $\|e\|_\rho$ and $\|R\|_\rho$ —using approximately symplectic properties in (IV.71)-(IV.77). We see next that C_j^+, for $j = 1, \ldots, 4$, are also controlled in a similar way, since they are given by equation which are close to (IV.34)-(IV.37) for the invariant and reducible case. For example, using equation (IV.26) for B_1 in the expressions of C_3^+ we obtain

$$C_3^+ = (L_\omega N_2 - J\mathrm{hess}\, hN_2 - N_2\Lambda)^\top D\tau G_{D\tau}^{-1} = \mathrm{i}G_{R^*,D\tau}G_{D\tau}^{-1}, \qquad \text{(IV.92)}$$

where we used equation (IV.84). To control C_1^+ and C_2^+ we have to compute the action of $\mathcal{R}$ on the matrix $JD\tau G_{D\tau}^{-1}$

$$\mathcal{R}(JD\tau G_{D\tau}^{-1}) = \mathcal{R}(JD\tau)G_{D\tau}^{-1} + JD\tau L_\omega(G_{D\tau}^{-1}) = (\mathrm{Id} + JD\tau G_{D\tau}^{-1}D\tau^\top J)T_1 D\tau G_{D\tau}^{-1}$$
$$+ JDeG_{D\tau}^{-1} - JD\tau G_{D\tau}^{-1}[G_{De,D\tau} + G_{D\tau,De}]G_{D\tau}^{-1}.$$

where T_1 is given by (IV.51). Then, if we multiply this expression by $D\tau^\top J$ we get

$$C_2^+ = -G_{De,D\tau}G_{D\tau}^{-1} \qquad \text{(IV.93)}$$

and if we multiply by $V^\top J$ and use the definitions of C_2^+, C_3^+ and C_4^+, we obtain

$$\begin{aligned} C_1^+ &= (B_3^\top D\tau^\top + B_2^\top N_2^\top + B_1^\top N_1^\top + G_{D\tau}^{-1}D\tau^\top J^\top)J\mathcal{R}(JD\tau G_{D\tau}^{-1}) + L_\omega B_3 - A_1 \\ &= -B_3^\top C_2^+ + B_2^\top(C_3^+ - L_\omega B_1 + \Lambda B_1) + B_1^\top(-C_4^+ + L_\omega B_2 + \Lambda B_2) \\ &\quad + G_{D\tau}^{-1}D\tau^\top \mathcal{R}(JD\tau G_{D\tau}^{-1}) + L_\omega B_3 - A_1 \\ &= -B_3^\top C_2^+ + B_2^\top C_3^+ - B_1^\top C_4^+ + G_{D\tau}^{-1}\Omega_{D\tau}G_{D\tau}^{-1}D\tau^\top JT_1 D\tau G_{D\tau}^{-1} \\ &\quad + G_{D\tau}^{-1}\Omega_{D\tau,De}G_{D\tau}^{-1} - G_{D\tau}^{-1}\Omega_{D\tau}G_{D\tau}^{-1}(G_{De,D\tau} + G_{D\tau,De})G_{D\tau}^{-1} + C_1^{++}, \end{aligned} \qquad \text{(IV.94)}$$

where C_1^{++} is given as

$$\begin{aligned} C_1^{++} &= G_{D\tau}^{-1}D\tau^\top T_1 D\tau G_{D\tau}^{-1} + B_1^\top(L_\omega B_2 + \Lambda B_2) - B_2^\top(L_\omega B_1 - \Lambda B_1) + L_\omega B_3 - A_1 \\ &= G_{D\tau}^{-1}D\tau^\top T_1 D\tau G_{D\tau}^{-1} + \mathrm{Re}(L_\omega B_2^\top B_1 - B_2^\top L_\omega B_1 + 2B_2^\top \Lambda B_1) - A_1, \end{aligned}$$

where we used that $B_3 = \mathrm{Re}(G_{B_2,B_1})$ and $(B_1^\top(L_\omega B_2 + \Lambda B_2))^* = -B_2^\top(L_\omega B_1 - \Lambda B_1)$. By introducing the expression (IV.50) for A_1, expanding $L_\omega B_1$ and $L_\omega B_2$ as in Lemma IV.4.3

$$\begin{aligned} L_\omega B_2 &= -\Lambda B_2 + N_1^\top JT_1 D\tau G_{D\tau}^{-1} - N_1^\top D\tau G_{D\tau}^{-1}D\tau^\top JT_1 D\tau G_{D\tau}^{-1} \\ &\quad - G_{R,D\tau}G_{D\tau}^{-1} - G_{N_1,De}G_{D\tau}^{-1} + G_{N_1,D\tau}G_{D\tau}^{-1}(G_{De,D\tau} + G_{D\tau,De})G_{D\tau}^{-1}. \\ L_\omega B_1 &= \Lambda B_1 - N_2^\top JT_1 D\tau G_{D\tau}^{-1} + N_2^\top D\tau G_{D\tau}^{-1}D\tau^\top JT_1 D\tau G_{D\tau}^{-1} \\ &\quad + \mathrm{i}G_{R^*,D\tau}G_{D\tau}^{-1} + G_{N_2,De}G_{D\tau}^{-1} - G_{N_2,D\tau}G_{D\tau}^{-1}(G_{De,D\tau} + G_{D\tau,De})G_{D\tau}^{-1}, \end{aligned}$$

and using that $\mathrm{Re}(N_1 N_2^\top) = -\mathrm{Re}((N_1 N_2^\top)^\top)$ and $\mathrm{Re}(G_{D\tau,R} G_{N_2,D\tau})^\top = -\mathrm{Re}(\mathrm{i} G_{D\tau,N_1} G_{R^*,D\tau})$, we obtain (after some cancellations)

$$C_1^{++} = \mathrm{Re}(T_3 + T_3^\top), \tag{IV.95}$$

where

$$T_3 := -G_{D\tau}^{-1}\Big(G_{D\tau,R} + G_{De,N_1} - (G_{De,D\tau} + G_{D\tau,De}) G_{D\tau}^{-1} G_{D\tau,N_1}\Big) G_{N_2,D\tau} G_{D\tau}^{-1}. \tag{IV.96}$$

Now, we control the expressions (IV.93), (IV.92), (IV.96) and (IV.95) as

$$\|C_2^+\|_{\rho-\delta} \leq \frac{\hat{\alpha}}{\delta}\|e\|_\rho, \quad \|C_3^+\|_{\rho-\delta}, \|C_4^+\|_{\rho-\delta} \leq \hat{\alpha}\|R\|_\rho, \quad \|T_3\|_{\rho-\delta}, \|C_1^{++}\|_{\rho-\delta} \leq \hat{\alpha}\left(\frac{\|e\|_\rho}{\delta} + \|R\|_\rho\right)$$

and we use these bounds to control the expression (IV.94) as follows

$$\|C_1^+\|_{\rho-2\delta} \leq \hat{\alpha}\left(\frac{\|e\|_\rho}{\gamma\delta^{\nu+1}} + \frac{\|e\|_\rho^2}{\gamma\delta^{\nu+2}} + \|R\|_\rho\right),$$

and we use hypothesis (IV.78) to get rid of the quadratic terms, thus obtaining

$$\|C_1^+\|_{\rho-2\delta} \leq \frac{\hat{\alpha}}{\gamma\delta^\nu}\left(\frac{\|e\|_\rho}{\delta} + \|R\|_\rho\right).$$

Therefore, we have

$$\|C_i\|_{\rho-2\delta} \leq \frac{\hat{\alpha}}{\gamma\delta^{\nu+1}}\left(\frac{\|e\|_\rho}{\delta} + \|R\|_\rho\right),$$

for $i = 1, \ldots, 4$. Finally, we obtain estimates for the inverse of the matrix $\mathrm{Id} + M_3$ that appears in system (IV.90), given by

$$\|(\mathrm{Id} + M_3)^{-1}\|_{\rho-2\delta} \leq \frac{1}{1 - \|M_3\|_{\rho-2\delta}}, \tag{IV.97}$$

that, by using hypothesis (IV.78) again, is well-posed since

$$\|M_3\|_{\rho-2\delta} \leq \frac{\hat{\alpha}}{\gamma\delta^\nu}\left(\frac{\|e\|_\rho}{\delta} + \|R\|_\rho\right) \leq \frac{1}{2}.$$

Therefore, we obtain (IV.80) for the functions $\{A_i^+\}_{i=1,\ldots,4}$. □

Proof of Proposition IV.5.1. We organize the proof of this iterative procedure in three parts. In part **I**), we correct the invariance of the torus by approximately solving the linearized equation $\mathcal{R}(\Delta_\tau) = -e$, given by (IV.19), as it was explained in Proposition IV.4.4. Analogously, in part **II**) we correct the reducibility of the torus by approximately solving the linearized equation $\mathcal{S}(\Delta_N, \Delta_\Lambda) = -\hat{R}$, given by (IV.17) and (IV.20), as it was explained in Proposition IV.4.6. Finally, in part **III**) we compute some additional estimates regarding the non-degeneracy condition for the new torus.

Firstly, let us observe that condition (IV.53) implies condition (IV.78) in Proposition IV.5.3 by taking a constant $\bar{\alpha}$ larger than $\hat{\alpha}$. Then, we use Proposition IV.5.3 construct an approximately symplectic basis at every point of the torus. As before, we redefine (enlarge) the constant $\bar{\alpha}$ along the proof to meet the different conditions given in the statement.

I) *Correction of the torus:* The idea is that the solution of the equation $\mathcal{R}(\Delta_\tau) = -e$ obtained in the invariant and reducible case —as discussed in Proposition IV.4.4— provides an approximate solution in the approximately invariant case. To this end, we consider the function

$$\Delta_\tau = D\tau\Delta_1 + V\Delta_2 + N_1\Delta_3 + N_2\Delta_4,$$

where Δ_i, for $i = 1, \ldots, 4$, are solutions of the cohomological equations (IV.39)-(IV.42), taking $[\Delta_1]_{\mathbb{T}^r} = 0$ and $[\Delta_2]_{\mathbb{T}^r}$ given by (IV.43). Then we claim that the new embedding $\bar{\tau} = \tau + \Delta_\tau$ parameterizes an approximate reducible and invariant torus $\bar{\mathscr{T}}$ with an error which is quadratic in $\|e\|_\rho$ and $\|R\|_\rho$. Of course, first we have to check the compatibility condition $\left[D\tau^\top Je\right]_{\mathbb{T}^r} = 0$, that follows from the next computation

$$D\tau^\top Je = D\tau^\top J(L_\omega\tau - J\mathrm{grad}\, h(\tau)) = \Omega_{D\tau}\omega + \mathrm{grad}\,_\theta(h(\tau)) \tag{IV.98}$$

by observing that both terms at the right hand side have zero average (see Remark IV.2.2). It is important to observe that $\Delta_3^* = \mathrm{i}\Delta_4$ so the correction Δ_τ is real analytic.

As far as the estimates are concerned, we have (using Lemma IV.5.2 to control the solution of the cohomological equations)

$$\|\Delta_1\|_{\rho-2\delta} \leq \frac{\bar{\alpha}}{\gamma^2\delta^{2\nu}}\|e\|_\rho, \qquad \|\Delta_i\|_{\rho-\delta} \leq \frac{\bar{\alpha}}{\gamma\delta^{\nu}}\|e\|_\rho,$$

for $i = 2, 3, 4$, so we can control the correction Δ_τ in the parameterization as follows

$$\|\Delta_\tau\|_{\rho-2\delta} \leq \frac{\bar{\alpha}}{\gamma^2\delta^{2\nu}}\|e\|_\rho,$$

thus obtaining estimate (IV.58). Moreover, we observe that the derivative of the new parameterization can be controlled easily as follows

$$\|D\bar{\tau}\|_{\rho-3\delta} \leq \|D\tau\|_\rho + \|D\Delta_\tau\|_{\rho-3\delta} \leq \|D\tau\|_\rho + \frac{\bar{\alpha}}{\gamma^2\delta^{2\nu+1}}\|e\|_\rho < \sigma_1,$$

where we used hypothesis (IV.53), and also the distance of $\bar{\tau}(\Delta(\rho-2\delta))$ to the boundary of $\mathcal{U}$

$$\begin{aligned} \operatorname{dist}(\bar{\tau}(\Delta(\rho-2\delta)),\partial\mathcal{U}) &\geq \operatorname{dist}(\tau(\Delta(\rho)),\partial\mathcal{U}) - \|\Delta_\tau\|_{\rho-2\delta} \\ &\geq \operatorname{dist}(\tau(\Delta(\rho)),\partial\mathcal{U}) - \frac{\bar{\alpha}}{\gamma^2\delta^{2\nu}}\|e\|_\rho > \sigma_2, \end{aligned}$$

where we used hypothesis (IV.54). Notice that we have achieved part of (IV.66) and (IV.67).

Next we control the new error in the invariance. To this end, we first introduce Δ_τ into $\mathcal{R}(\Delta_\tau)+e$ and we use the properties (IV.79) and (IV.89) of the operator $\mathcal{R}$ and also the cohomological equations (IV.39)-(IV.42), thus obtaining

$$\begin{aligned} \mathcal{R}(\Delta_\tau)+e &= \mathcal{R}(D\tau)\Delta_1 + \mathcal{R}(V)\Delta_2 + \mathcal{R}(N_1)\Delta_3 + \mathcal{R}(N_2)\Delta_4 \\ &\quad + D\tau L_\omega\Delta_1 + VL_\omega\Delta_2 + N_1L_\omega\Delta_3 + N_2L_\omega\Delta_4 + e \\ &= De\Delta_1 + (D\tau A_1^+ + VA_2^+ + N_1A_3^+ + N_2A_4^+)\Delta_2 + R\Delta_3 + \mathrm{i}R^*\Delta_4 \\ &\quad \underbrace{-D\tau V^\top Je + VD\tau^\top Je - N_1N_2^\top Je + N_2N_1^\top Je + e}_{e^+} \end{aligned} \tag{IV.99}$$

We note that the terms not included in e^+ are clearly quadratic in e and R, since the functions $\{A_i^+\}_{i=1,\ldots,4}$ and $\{\Delta_i\}_{i=1,\ldots,4}$ are controlled by $\|e\|_\rho$ and $\|R\|_\rho$. Then, it suffices to study the remaining part e^+. To this end, we write e^+ in terms of the constructed basis

$$e^+ = D\tau e_1^+ + Ve_2^+ + N_1e_3^+ + N_2e_4^+,$$

and obtain $\{e_i^+\}_{i=1,\ldots,4}$ by multiplying at both sides by $V^\top J$, $D\tau^\top J$, $N_2^\top J$ and $N_1^\top J$. This leads to study the linear system

$$\big(\mathrm{Id}+M_3\big)\begin{pmatrix} e_1^+ \\ e_2^+ \\ e_3^+ \\ e_4^+ \end{pmatrix} = \begin{pmatrix} D_1 \\ D_2 \\ D_3 \\ D_4 \end{pmatrix}, \tag{IV.100}$$

where M_3 is given in (IV.91) and the matrices in the right-hand side are the following

$$\begin{aligned} D_1 &= -(\Omega_{V,D\tau}-\mathrm{Id}_r)V^\top Je + \Omega_V D\tau^\top Je - \Omega_{V,N_1}N_2^\top Je + \Omega_{V,N_2}N_1^\top Je, \\ D_2 &= \Omega_{D\tau}V^\top Je - (\Omega_{D\tau,V}+\mathrm{Id}_r)D\tau^\top Je + \Omega_{D\tau,N_1}N_2^\top Je - \Omega_{D\tau,N_2}N_1^\top Je, \\ D_3 &= -\Omega_{N_2,D\tau}V^\top Je + \Omega_{N_2,V}D\tau^\top Je - (\Omega_{N_2,N_1}-\mathrm{Id}_{n-r})N_2^\top Je + \Omega_{N_2}N_1^\top Je, \\ D_4 &= -\mathrm{i}D_3^*. \end{aligned}$$

Now we control these functions using estimates (IV.71)-(IV.77) in Proposition IV.5.3

$$\|D_i\|_{\rho-2\delta} \leq \frac{\bar{\alpha}}{\gamma\delta^\nu}\left(\frac{\|e\|_\rho}{\delta} + \|R\|_\rho\right)\|e\|_\rho,$$

for $i = 1, \ldots, 4$. We have shown in the proof of Proposition IV.5.3 that the matrix $\mathrm{Id} + M_3$ is invertible and that $\|(\mathrm{Id} + M_3)^{-1}\|_{\rho-2\delta} \leq 2$ (see (IV.97)) so we conclude that

$$\|e^+\|_{\rho-2\delta} \leq \frac{\bar{\alpha}}{\gamma\delta^\nu}\left(\frac{\|e\|_\rho}{\delta} + \|R\|_\rho\right)\|e\|_\rho.$$

Going back to equation (IV.99) we get

$$\|\mathcal{R}(\Delta_\tau) + e\|_{\rho-3\delta} \leq \frac{\bar{\alpha}}{\gamma^2\delta^{2\nu+1}}\left(\frac{\|e\|_\rho}{\delta} + \|R\|_\rho\right)\|e\|_\rho,$$

and therefore, we conclude that $\mathcal{R}(\Delta_\tau) = -e$ is solved modulo quadratic terms in the errors. Then, we observe that

$$\begin{aligned}\bar{e} &= L_\omega\bar{\tau} - J\mathrm{grad}\, h(\bar{\tau}) \\ &= \mathcal{R}(\Delta_\tau) + e + J(\mathrm{grad}\, h(\tau) + \mathrm{hess}\, h(\tau)\Delta_\tau - \mathrm{grad}\, h(\tau + \Delta_\tau))\end{aligned}$$

and control the last terms by estimating the residue of the Taylor expansion of h up to second order, thus obtaining

$$\|\mathrm{grad}\, h(\tau) + \mathrm{hess}\, h(\tau)\Delta_\tau - \mathrm{grad}\, h(\tau + \Delta_\tau)\|_{\rho-2\delta} \leq \frac{\bar{\alpha}}{\gamma^4\delta^{4\nu}}\|e\|_\rho^2.$$

Hence, we end up with

$$\|\bar{e}\|_{\rho-3\delta} \leq \frac{\bar{\alpha}}{\gamma^4\delta^{4\nu-1}}\left(\frac{\|e\|_\rho}{\delta} + \|R\|_\rho\right)\|e\|_\rho,$$

where we used that $\nu > r - 1 \geq 1$, finally obtaining estimate (IV.59).

II) *Correction of the reducibility:* To square the error in reducibility of the new torus $\bar{\mathscr{T}}$ we have to deal with the equation $\mathcal{S}(\Delta_N, \Delta_\Lambda) = -\hat{R}$, given by (IV.17) and (IV.20). As before, we solve approximately this equation by taking (the reason of writing $\hat{\Delta}_N$ rather than Δ_N will be clear later on)

$$\hat{\Delta}_N = D\tau P_1 + VP_2 + N_1P_3 + N_2P_4,$$

$\{P_i\}_{i=1,\ldots,4}$ and Δ_Λ being the solution of the cohomological equations (IV.44)-(IV.47) for

$$\hat{R} = R + J\mathrm{hess}\, h(\tau)N - J\mathrm{hess}\, h(\bar{\tau})N,$$

and fixing $\mathrm{diag}\ [P_3]_{\mathbb{T}^r} = 0$. The formal solution of these equation has been discussed in Proposition IV.4.6 so we know that we must take $\Delta_\Lambda = -\mathrm{diag}\,[N_2^\top J\hat{R}]_{\mathbb{T}^r}$.

Firstly, we claim that the geometry of the problem imposes that the selected Δ_Λ is pure imaginary, so our procedure automatically preserves the approximately elliptic character of the torus. To see that, we observe that transposing equation (IV.88) leads to

$$L_\omega\Omega_{N_1,N_2} = -\Lambda\Omega_{N_1,N_2} + \Omega_{N_1,N_2}\Lambda + \Omega_{R,N_2} + \mathrm{i}\Omega_{N_1,R^*}.$$

Since the left-hand side of this expression has vanishing average and $\mathrm{diag}\ [\Omega_{N_1,N_2}]_{\mathbb{T}^r} = -\mathrm{Id}_{n-r}$, it turns out that

$$\mathrm{diag}\ [\Omega_{R,N_2} + \mathrm{i}\Omega_{N_1,R^*}]_{\mathbb{T}^r} = 0,$$

and so $\mathrm{diag}\ [\mathrm{i}\Omega_{N_1,R^*}]_{\mathbb{T}^r} = \mathrm{diag}\ [\Omega_{N_2,R}]_{\mathbb{T}^r}^\top$. Then, it is straightforward to compute

$$\begin{aligned}\Delta_\Lambda^* = &-\mathrm{diag}\ \left[N_2^\top J\hat{R}\right]_{\mathbb{T}^r}^* = -\mathrm{diag}\ \left[-\mathrm{i}\Omega_{N_1,R^*} + N_1^\top(\mathrm{hess}\, h(\tau) - \mathrm{hess}\, h(\bar\tau))N_2\right]_{\mathbb{T}^r} \\ &= \mathrm{diag}\ [\Omega_{N_2,R}]_{\mathbb{T}^r}^\top - \mathrm{diag}\ \left[N_2^\top(\mathrm{hess}\, h(\tau) - \mathrm{hess}\, h(\bar\tau))N_1\right]_{\mathbb{T}^r}^\top \\ &= \mathrm{diag}\ \left[\Omega_{N_2,R} - N_2^\top(\mathrm{hess}\, h(\tau) - \mathrm{hess}\, h(\bar\tau))N_1\right]_{\mathbb{T}^r}^\top = -\Delta_\Lambda^\top = -\Delta_\Lambda,\end{aligned} \tag{IV.101}$$

so Δ_Λ is pure imaginary.

Now obtaining estimates for the solution of the cohomological equations is straightforward after controlling

$$\|\hat{R}\|_{\rho-2\delta} \leq \frac{\bar\alpha}{\gamma^2\delta^{2\nu-1}}\left(\frac{\|e\|_\rho}{\delta} + \|R\|_\rho\right), \qquad |\Delta_\Lambda| \leq \frac{\bar\alpha}{\gamma^2\delta^{2\nu-1}}\left(\frac{\|e\|_\rho}{\delta} + \|R\|_\rho\right),$$

and applying Lemma IV.5.2

$$\|P_1\|_{\rho-4\delta} \leq \frac{\bar\alpha}{\gamma^4\delta^{4\nu-1}}\left(\frac{\|e\|_\rho}{\delta} + \|R\|_\rho\right), \qquad \|P_i\|_{\rho-3\delta} \leq \frac{\bar\alpha}{\gamma^3\delta^{3\nu-1}}\left(\frac{\|e\|_\rho}{\delta} + \|R\|_\rho\right)$$

for $i = 2, 3, 4$. With these estimates we check condition (IV.66) for the new approximate normal frequencies $\bar\lambda$. For example,

$$|\bar\lambda_i \pm \bar\lambda_j| \geq |\lambda_i \pm \lambda_j| - 2|\Delta_\Lambda| \geq \min_{i\neq j} |\lambda_i \pm \lambda_j| - 2|\Delta_\Lambda| > \sigma_3,$$

where we used (IV.55). Similar computations allow us to see that $\frac{\sigma_3}{2} < |\bar\lambda_j| < \frac{\sigma_4}{2}$, using (IV.56) and (IV.57), respectively.

We also have

$$\|\hat\Delta_N\|_{\rho-4\delta} \leq \frac{\bar\alpha}{\gamma^4\delta^{4\nu-1}}\left(\frac{\|e\|_\rho}{\delta} + \|R\|_\rho\right)$$

and we observe that, if we introduce $\hat N = N + \hat\Delta_N$, using (IV.53) we obtain that

$$\|\hat N\|_{\rho-4\delta} \leq \|N\|_\rho + \|\hat\Delta_N\|_{\rho-4\delta} \leq \|N\|_\rho + \frac{\bar\alpha}{\gamma^4\delta^{4\nu-1}}\left(\frac{\|e\|_\rho}{\delta} + \|R\|_\rho\right) < \sigma_1, \tag{IV.102}$$

and that the matrix $\mathrm{diag}\,[\Omega_{\hat N,\hat N^*}]_{\mathbb{T}^r}$ is constant, diagonal and pure imaginary, but it is not iId_{n-r} as we want. Nevertheles, from the following expression

$$\Omega_{\hat N,\hat N^*} - \Omega_{N,N^*} = \Omega_{N,\hat\Delta_N^*} + \Omega_{\hat\Delta_N,\hat N^*}$$

and using hypothesis (IV.53) we obtain

$$\|\Omega_{\hat{N},\hat{N}^*} - \Omega_{N,N^*}\|_{\rho-4\delta} \leq \frac{\bar{\alpha}}{\gamma^4\delta^{4\nu-1}}\left(\frac{\|e\|_\rho}{\delta} + \|R\|_\rho\right),$$

where we recall that $\mathrm{diag}\,[\Omega_{N,N^*}]_{\mathbb{T}^r} = \mathrm{i}\mathrm{Id}_{n-r}$. Hence, we have that the elements of $\mathrm{diag}\,[\Omega_{\hat{N},\hat{N}^*}]_{\mathbb{T}^r}$ are of the form $\mathrm{i}(1+d_i)$ with

$$|d_i| \leq \frac{\bar{\alpha}}{\gamma^4\delta^{4\nu-1}}\left(\frac{\|e\|_\rho}{\delta} + \|R\|_\rho\right).$$

Hence, using again hypothesis (IV.53), we have that $|d_i| \leq 1/2$ for $i = 1, \ldots, n-r$, so we can normalize $\hat{N}$ in order to preserve hypothesis H_3. To this end, we define the real matrix

$$B = \mathrm{diag}\,(b_1, \ldots, b_{n-r}), \quad \text{with} \qquad b_i = \sqrt{\frac{1}{1+d_i}},$$

and it turns out that the matrix $\bar{N} = \hat{N}B$ satisfies $\mathrm{diag}\,\left[\Omega_{\bar{N},\bar{N}^*}\right]_{\mathbb{T}^r} = \mathrm{i}\mathrm{Id}_{n-r}$. Let us observe that the performed correction is small, since if we take $\bar{N} = N + \Delta_N$ we have that

$$\Delta_N = N(B - \mathrm{Id}_{n-r}) + \hat{\Delta}_N B,$$

and so

$$\|\Delta_N\|_{\rho-4\delta} \leq \frac{\bar{\alpha}}{\gamma^4\delta^{4\nu-1}}\left(\frac{\|e\|_\rho}{\delta} + \|R\|_\rho\right),$$

that corresponds to estimate (IV.61). We see that $\|\bar{N}\|_{\rho-4\delta} < \sigma_1$ by similar computations as in (IV.102), thus obtaining the corresponding condition in (IV.67).

The rest of this part is devoted to check that, using $\bar{N}$ and $\bar{\Lambda}$, the new approximately invariant torus $\bar{\mathscr{T}}$ is approximately elliptic up to a quadratic error. To this end, we compute

$$\begin{aligned} \bar{R} &= L_\omega \bar{N} - J\mathrm{hess}\, h(\bar{\tau})\bar{N} + \bar{N}\bar{\Lambda} \\ &= \mathcal{S}(\Delta_N, \Delta_\Lambda) + \hat{R} + J(\mathrm{hess}\, h(\tau) - \mathrm{hess}\, h(\bar{\tau}))\Delta_N + \Delta_N\Delta_\Lambda, \end{aligned} \tag{IV.103}$$

where the action of $\mathcal{S}$ on Δ_N is written in terms of the action on $\hat{\Delta}_N$ as follows

$$\begin{aligned} &\mathcal{S}(\Delta_N, \Delta_\Lambda) + \hat{R} \\ &\quad = \mathcal{S}(N(B - \mathrm{Id}_{n-r}) + \hat{\Delta}_N B, \Delta_\Lambda) + \hat{R} \\ &\quad = \mathcal{R}(N(B - \mathrm{Id}_{n-r})) + \mathcal{R}(\hat{\Delta}_N B) + N\Delta_\Lambda + N(B - \mathrm{Id}_{n-r})\Lambda + \hat{\Delta}_N B\Lambda + \hat{R} \\ &\quad = R(B - \mathrm{Id}_{n-r}) + \mathcal{R}(\hat{\Delta}_N)B + N\Delta_\Lambda + \hat{\Delta}_N B\Delta + \hat{R} \\ &\quad = (\mathcal{S}(\hat{\Delta}_N, \Delta_\Lambda) + \hat{R})B + (R - \hat{R} - N\Delta_\Lambda + \hat{\Delta}_N\Lambda)(B - \mathrm{Id}_{n-r}), \end{aligned}$$

where we used that $\mathcal{R}(N) = -N\Lambda + R$ and $B\Lambda = \Lambda B$.

Then, we introduce $\hat{\Delta}_N$ and Δ_Λ in $\mathcal{S}(\hat{\Delta}_N, \Delta_\Lambda) + \hat{R}$ and we use the properties (IV.79) and (IV.89) of the operator $\mathcal{R}$ and also the cohomological equations (IV.44)-(IV.47), thus obtaining

$$\begin{aligned} \mathcal{S}(\hat{\Delta}_N, \Delta_\Lambda) + \hat{R} &= \mathcal{R}(\hat{\Delta}_N) + N_1\Delta_\Lambda + \hat{\Delta}_N\Lambda + \hat{R} \\ &= \mathcal{R}(D\tau)P_1 + \mathcal{R}(V)P_2 + \mathcal{R}(N_1)P_3 + \mathcal{R}(N_2)P_4 \\ &\quad + D\tau L_\omega P_1 + VL_\omega P_2 + N_1 L_\omega P_3 + N_2 L_\omega P_4 + N_1\Delta_\Lambda \\ &\quad + D\tau P_1\Lambda + VP_2\Lambda + N_1P_3\Lambda + N_2P_4\Lambda + \hat{R} \\ &= DeP_1 + (D\tau A_1^+ + VA_2^+ + N_1A_3^+ + N_2A_4^+)P_2 + RP_3 + \mathrm{i}R^*P_4 \\ &\quad \underbrace{-D\tau V^\top J\hat{R} + VD\tau^\top J\hat{R} - N_1N_2^\top J\hat{R} + N_2N_1^\top J\hat{R} + \hat{R}}_{R^+}. \end{aligned} \tag{IV.104}$$

As we made in equation (IV.99), the terms not included in R^+ are clearly quadratic in e and R. Then, we express R^+ in terms of the basis

$$R^+ = D\tau R_1^+ + VR_2^+ + N_1R_3^+ + N_2R_4^+,$$

and for R_j^+ we get a system like (IV.100) for e_j^+, simply by replacing e with $\hat{R}$ in the definition of D_j. Hence,

$$\|R^+\|_{\rho-2\delta} \leq \frac{\bar{\alpha}}{\gamma^3\delta^{3\nu-1}}\left(\frac{\|e\|_\rho}{\delta} + \|R\|_\rho\right)^2.$$

Therefore, we can compute a bound for the error in the solution of the linear equation that corrects reducibility

$$\|\mathcal{S}(\hat{\Delta}_N, \Delta_\Lambda) + \hat{R}\|_{\rho-4\delta} \leq \frac{\bar{\alpha}}{\gamma^4\delta^{4\nu}}\left(\frac{\|e\|_\rho}{\delta} + \|R\|_\rho\right)^2,$$

so we obtain —again, we use hypothesis (IV.53) to control the quadratic terms—

$$\|\mathcal{S}(\Delta_N, \Delta_\Lambda) + \hat{R}\|_{\rho-4\delta} \leq \frac{\bar{\alpha}}{\gamma^8\delta^{8\nu-2}}\left(\frac{\|e\|_\rho}{\delta} + \|R\|_\rho\right)^2.$$

Therefore, recalling (IV.103), we easily show that the new error (IV.62) in reducibility is quadratic

$$\|\bar{R}\|_{\rho-4\delta} \leq \frac{\bar{\alpha}}{\gamma^8\delta^{8\nu-2}}\left(\frac{\|e\|_\rho}{\delta} + \|R\|_\rho\right)^2.$$

III) *Additional estimates:* Finally, we have to check estimates that allow us to control the non-degeneracy of the basis and the twist condition. Using that

$$G_{D\bar{\tau}} - G_{D\tau} = G_{D\tau, D\Delta_\tau} + G_{D\Delta_\tau, D\bar{\tau}}$$

and recalling (IV.53) and (IV.58), we get

$$\|G_{D\bar{\tau}} - G_{D\tau}\|_{\rho-3\delta} \leq \frac{\bar{\alpha}}{\gamma^2\delta^{2\nu+1}}\|e\|_\rho.$$

Now, we observe that $G_{D\bar{\tau}}^{-1} = (\mathrm{Id}_r + G_{D\tau}^{-1}(G_{D\bar{\tau}} - G_{D\tau}))^{-1}G_{D\tau}^{-1}$ so we can compute the following —again, we make use of (IV.53)—

$$\begin{aligned}\|G_{D\bar{\tau}}^{-1} - G_{D\tau}^{-1}\|_{\rho-3\delta} &\leq \|G_{D\tau}^{-1}\|_\rho \|(\mathrm{Id}_r + G_{D\tau}^{-1}(G_{D\bar{\tau}} - G_{D\tau}))^{-1} - \mathrm{Id}_r\|_{\rho-3\delta} \\ &\leq \frac{\|G_{D\tau}^{-1}\|_\rho^2 \|G_{D\bar{\tau}} - G_{D\tau}\|_{\rho-3\delta}}{1 - \|G_{D\tau}^{-1}\|_\rho \|G_{D\bar{\tau}} - G_{D\tau}\|_{\rho-3\delta}} \leq \frac{\bar{\alpha}}{\gamma^2\delta^{2\nu+1}}\|e\|_\rho,\end{aligned}$$

thus obtaining (IV.63) and the term in (IV.67) that corresponds to $G_{D\bar{\tau}}^{-1}$. Similar computations allow us to control the non-degeneracy of the set of normal vectors, thus geting (IV.64) and (IV.67) for $G_{\bar{N},\bar{N}^*}^{-1}$. Now, we are able to estimate the new twist condition for

$$\bar{A}_1(\theta) := G_{D\bar{\tau}}^{-1}(\theta) D\bar{\tau}(\theta)^\top (\bar{T}_1(\theta) + \bar{T}_2(\theta) + \bar{T}_2(\theta)^\top) D\bar{\tau}(\theta) G_{D\bar{\tau}}^{-1}(\theta),$$

where

$$\begin{aligned}\bar{T}_1(\theta) &:= J^\top \operatorname{hess} h(\bar{\tau}(\theta)) J - \operatorname{hess} h(\bar{\tau}(\theta)),\\ \bar{T}_2(\theta) &:= \bar{T}_1 J \left[D\bar{\tau}(\theta) G_{D\bar{\tau}}(\theta)^{-1} D\bar{\tau}(\theta)^\top - \mathrm{Id}\right] \mathrm{Re}(\mathrm{i}\bar{N}(\theta)\bar{N}^*(\theta)^\top).\end{aligned}$$

As before, we first bound

$$|\left[\bar{A}_1\right]_{\mathbb{T}^r} - \left[A_1\right]_{\mathbb{T}^r}| \leq \frac{\bar{\alpha}}{\gamma^4\delta^{4\nu-1}}\left(\frac{\|e\|_\rho}{\delta} + \|R\|_\rho\right).$$

Now we estimate the inverse of $\left[\bar{A}_1\right]_{\mathbb{T}^r}$ by using the fact $\bar{A}_1 = A_1 + \bar{A}_1 - A_1$ and we can repeat the same argument used before, using hypothesis (IV.53), thus obtaining bounds (IV.65) and (IV.67) for $\left[\bar{A}_1\right]_{\mathbb{T}^r}^{-1}$. □

IV.6 Proof of the main result

In this section we prove Theorem IV.3.1 by applying inductively Proposition IV.5.1. First, in Section IV.6.1 we study the convergence of the obtained iterative scheme, without worrying about the exclusion of parameters that lead to resonances. As usual, the quadratic convergence overcomes the effect of small divisors. Then, in Section IV.6.2 we prove that Lipschitz regularity is preserved along the iterative procedure. Finally, in Section IV.6.3, we estimate the measure of the set of excluded parameters.

IV.6.1 Convergence of the Newton scheme

First, we observe that all the bounds are uniform with respect to μ so from now on we omit the dependence on the parameter. This is, for any fixed $\mu \in I$ we denote the objects that characterize approximately elliptic invariant tori as

$$\tau_{(0)} := \tau_\mu, \qquad N_{(0)} := N_\mu, \qquad \Lambda_{(0)} := \Lambda_\mu,$$

and we introduce also

$$e_{(0)} := e_\mu, \qquad R_{(0)} := R_\mu, \qquad A_{1,(0)} := A_{1,\mu}, \qquad \lambda_{(0)} = (\lambda_1^{(0)}, \ldots, \lambda_{n-r}^{(0)}) := \lambda_\mu.$$

where we recall that $\Lambda_{(0)} = \operatorname{diag}(\mathrm{i}\lambda_{(0)})$. Moreover, given certain $\gamma_{(0)} > 0$ —to be precised later— such that $\gamma_{(0)} \leq \frac{1}{2}\min\{1, \hat{\gamma}\}$, we define the following quantities (let us recall from H_2 that $0 < \rho < 1$)

$$\rho_{(0)} = \rho, \quad \delta_{(0)} = \frac{\rho_{(0)}}{16}, \quad \rho_{(s)} = \rho_{(s-1)} - 4\delta_{(s-1)}, \quad \delta_{(s)} = \frac{\delta_{(0)}}{2^s}, \quad \gamma_{(s)} = (1 + 2^{-s})\gamma_{(0)},$$

for any $s \geq 1$, and consider the normalized error

$$\varepsilon_{(0)} := \frac{\|e_{(0)}\|_{\rho_{(0)}}}{\delta_{(0)}} + \|R_{(0)}\|_{\rho_{(0)}}.$$

Then, we are goint to show that, considering the constant $\bar{\alpha}$ provided by Proposition IV.5.1, which depends on the quantities ν, r, n, $|\omega|$, σ_0, σ_1, σ_2, σ_3 and σ_4 in the statement of Theorem IV.3.1, if the normalized error $\varepsilon_{(0)}$ is sufficiently small so that

$$\frac{2^{8\nu-1}\bar{\alpha}\varepsilon_{(0)}}{\gamma_{(0)}^8 \delta_{(0)}^{8\nu-2}} < \frac{1}{2}\min\Big\{1, \sigma_1 - \sigma^*, \operatorname{dist}(\tau_{(0)}(\Delta(\rho)), \partial\mathcal{U}) - \sigma_2, \sigma^{**} - \sigma_3, \sigma_4 - 2\max_j |\lambda_j^{(0)}|\Big\}, \tag{IV.105}$$

where

$$\sigma^* = \max\Big\{\|D\tau_{(0)}\|_{\rho_{(0)}}, \|N_{(0)}\|_{\rho_{(0)}}, \|G_{D\tau_{(0)}}^{-1}\|_{\rho_{(0)}}, \|G_{N_{(0)},N_{(0)}^*}^{-1}\|_{\rho_{(0)}}, |\left[A_{1,(0)}\right]_{\mathbb{T}^r}^{-1}|\Big\}, \tag{IV.106}$$

$$\sigma^{**} = \min\Big\{\min_{i\neq j}|\lambda_i^{(0)} \pm \lambda_j^{(0)}|, 2\min_j|\lambda_j^{(0)}|\Big\}. \tag{IV.107}$$

then we can apply recursively Proposition IV.5.1 to the initial approximation, thus obtaining a sequence

$$\tau_{(s)} = \bar{\tau}_{(s-1)} = \tau_{(s-1)} + \Delta_{\tau_{(s-1)}},$$
$$N_{(s)} = \bar{N}_{(s-1)} = N_{(s-1)} + \Delta_{N_{(s-1)}},$$
$$\Lambda_{(s)} = \bar{\Lambda}_{(s-1)} = \Lambda_{(s-1)} + \Delta_{\Lambda_{(s-1)}},$$

$$\begin{aligned} e_{(s)} &= \bar{e}_{(s-1)}, \\ R_{(s)} &= \bar{R}_{(s-1)}, \\ A_{1,(s)} &= \bar{A}_{1,(s-1)}, \end{aligned}$$

all these objects being analytic in $\Delta(\rho_{(s)})$. This sequence is well defined up to the sth term provided that the parameter μ belongs to the set $I_{(s)}$ defined iterative by $I_{(-1)} = I$ and

$$I_{(s)} = \{\mu \in I_{(s-1)} : \lambda_{(s-1)} \text{ satisfies Diophantine conditions (IV.10)} \\ \text{of } (\gamma_{(s)}, \nu)\text{-type with respect to } \omega\}.$$

Let us observe that, since the basic frequencies ω are fixed along the procedure and $2\gamma_{(0)} \leq \hat{\gamma}$, they automatically satisfy Diophantine conditions (IV.4) of $(\gamma_{(s)}, \nu)$-type, for every $s \geq 1$.

Now, let us assume that we have been able to apply s times Proposition IV.5.1. To this end, we define $\sigma^*_{(s)}$ and $\sigma^{**}_{(s)}$ as in (IV.106) and (IV.107), just by replacing the (0)-objects with (s)-ones. First, we observe that by hypothesis

$$\sigma^*_{(s)} < \sigma_1, \quad \operatorname{dist}(\tau_{(s)}(\Delta(\rho_{(s)})), \partial\mathcal{U}) > \sigma_2, \quad \sigma^{**}_{(s)} > \sigma_3, \quad \max_j |\lambda_j^{(s)}| < \frac{\sigma_4}{2},$$

so the construction of the constant $\bar{\alpha}$ is uniform for all iterative steps —it depends on the constants σ_1, σ_2, σ_3 and σ_4 that remain unchanged along the procedure— and so, conditions (IV.53)-(IV.57) are fullfilled provided that the normalized error $\varepsilon_{(s)}$ satisfies —let us recall that $\gamma_{(s)} < 1$ and $\delta_{(s)} < 1$—

$$\frac{\bar{\alpha}\varepsilon_{(s)}}{\gamma^4_{(s)}\delta^{4\nu-1}_{(s)}} < \frac{1}{2}\min\Big\{1, \sigma_1 - \sigma^*_{(s)}, \operatorname{dist}(\tau_{(s)}(\Delta(\rho_{(s)})), \partial\mathcal{U}) - \sigma_2, \sigma^{**}_{(s)} - \sigma_3, \sigma_4 - 2\max_j |\lambda_j^{(s)}|\Big\}. \tag{IV.108}$$

In order to verify this inequality, we stat by computing the normalized error at the sth step

$$\varepsilon_{(s)} = \frac{\|e_{(s)}\|_{\rho_{(s)}}}{\delta_{(s)}} + \|R_{(s)}\|_{\rho_{(s)}} \leq \frac{2\bar{\alpha}}{\gamma^8_{(s-1)}\delta^{8\nu-2}_{(s-1)}}\varepsilon^2_{(s-1)} \leq \frac{2^{(s-1)(8\nu-2)+1}\bar{\alpha}}{\gamma^8_{(0)}\delta^{8\nu-2}_{(0)}}\varepsilon^2_{(s-1)},$$

where we used that $\gamma_{(s-1)} \geq \gamma_{(0)}$. Then, by iterating this sequence backwards, we obtain that

$$\varepsilon_{(s)} \leq \frac{\gamma^8_{(0)}\delta^{8\nu-2}_{(0)}}{2\bar{\alpha}} 2^{-(s+1)(8\nu-2)} \left(\frac{\bar{\alpha}2^{8\nu-1}\varepsilon_{(0)}}{\gamma^8_{(0)}\delta^{8\nu-2}_{(0)}}\right)^{2^s}.$$

Using this expression of the error, we verify condition (IV.108) in order to perform the step $s+1$. For example, the first term in this condition is straightforward

$$\frac{\bar{\alpha}\,\varepsilon_{(s)}}{\gamma^4_{(s)}\delta^{4\nu-1}_{(s)}} \leq \frac{1}{2}\gamma^4_{(0)}\delta^{4\nu-1}_{(0)}2^{-(4\nu-1)s-8\nu+2}\left(\frac{\bar{\alpha}2^{8\nu-1}\varepsilon_{(0)}}{\gamma^8_{(0)}\delta^{8\nu-2}_{(0)}}\right)^{2^s} < \frac{1}{2},$$

recalling that $\nu > r-1 \geq 1$ and (IV.105). In order to verify the remaining conditions in (IV.108), we have to control also the objects $\|D\tau_{(s)}\|_{\rho_{(s)}}$, $\|N_{(s)}\|_{\rho_{(s)}}$, $|\Lambda_{(s)}|$, etc. For example, we discuss in detail the following inequality

$$\|D\tau_{(s)}\|_{\rho_{(s)}} + \frac{\bar{\alpha}\,\varepsilon_{(s)}}{\gamma_{(s)}^4 \delta_{(s)}^{4\nu-1}} < \sigma_1. \tag{IV.109}$$

By using $D\tau_{(s)} = D\tau_{(s-1)} + D\Delta_{\tau_{(s-1)}}$ recursively as follows

$$\begin{aligned}
\|D\tau_{(s)}\|_{\rho_{(s)}} + \frac{\bar{\alpha}\,\varepsilon_{(s)}}{\gamma_{(s)}^4 \delta_{(s)}^{4\nu-1}} &\leq \|D\tau_{(s-1)}\|_{\rho_{(s-1)}} + \|D\Delta_{\tau_{(s-1)}}\|_{\rho_{(s)}} + \frac{\bar{\alpha}\,\varepsilon_{(s)}}{\gamma_{(s)}^4 \delta_{(s)}^{4\nu-1}} \\
&\leq \|D\tau_{(0)}\|_{\rho_{(0)}} + \sum_{j=0}^{s-1} \|D\Delta_{\tau_{(j)}}\|_{\rho_{(j+1)}} + \frac{\bar{\alpha}\,\varepsilon_{(s)}}{\gamma_{(s)}^4 \delta_{(s)}^{4\nu-1}} \\
&\leq \|D\tau_{(0)}\|_{\rho_{(0)}} + \sum_{j=0}^{s-1} \frac{\bar{\alpha}\varepsilon_{(j)}}{\gamma_{(j)}^2 \delta_{(j)}^{2\nu}} + \frac{\bar{\alpha}\,\varepsilon_{(s)}}{\gamma_{(s)}^4 \delta_{(s)}^{4\nu-1}} \\
&\leq \|D\tau_{(0)}\|_{\rho_{(0)}} + \sum_{j=0}^{s} \frac{\bar{\alpha}\,\varepsilon_{(j)}}{\gamma_{(j)}^4 \delta_{(j)}^{4\nu-1}}.
\end{aligned}$$

Notice that in the above computations we used estimate (IV.58) in Proposition IV.5.1 and the fact that $\gamma_{(s)}, \delta_{(s)} < 1$. Then, we introduce the expression for the errors $\varepsilon_{(j)}$ previously computed and use that $j+1 \leq 2^j$ in order to get

$$\begin{aligned}
\|D\tau_{(s)}\|_{\rho_{(s)}} + \frac{\bar{\alpha}\,\varepsilon_{(s)}}{\gamma_{(s)}^4 \delta_{(s)}^{4\nu-1}} &\leq \|D\tau_{(0)}\|_{\rho_{(0)}} + \gamma_{(0)}^4 \delta_{(0)}^{4\nu-1} 2^{-8\nu+1} \sum_{j=0}^{s} 2^{-(4\nu-1)j} \left(\frac{\bar{\alpha} 2^{8\nu-1} \varepsilon_{(0)}}{\gamma_{(0)}^8 \delta_{(0)}^{8\nu-2}} \right)^{2^j} \\
&\leq \|D\tau_{(0)}\|_{\rho_{(0)}} + \gamma_{(0)}^4 \delta_{(0)}^{4\nu-1} 2^{-8\nu+1} \sum_{j=0}^{\infty} \left(\frac{\bar{\alpha} 2^{8\nu-1} \varepsilon_{(0)}}{\gamma_{(0)}^8 \delta_{(0)}^{8\nu-2}} \right)^{j+1} \\
&< \|D\tau_{(0)}\|_{\rho_{(0)}} + \frac{2\bar{\alpha}\varepsilon_{(0)}}{\gamma_{(0)}^4 \delta_{(0)}^{4\nu-1}} < \sigma_1,
\end{aligned}$$

where in the last inequality we have used hypothesis (IV.105) in order to bound the expression by the sum of a geometric progression of ratio $1/2$. Analogous computations can be performed for the objects $\|N_{(s)}\|_{\rho_{(s)}}$, $\|G_{D\tau_{(s)}}^{-1}\|_{\rho_{(s)}}$, $\|G_{N_{(s)},N_{(s)}^*}^{-1}\|_{\rho_{(s)}}$ and $|\left[A_{(s),1}\right]_{\mathbb{T}^r}^{-1}|$, in order to verify that —we use estimates (IV.61), (IV.63), (IV.64) and (IV.65), respectively—

$$\|N_{(s)}\|_{\rho_{(s)}} + \frac{\bar{\alpha}\,\varepsilon_{(s)}}{\gamma_{(s)}^4 \delta_{(s)}^{4\nu-1}} < \sigma_1, \qquad \|G_{D\tau_{(s)}}^{-1}\|_{\rho_{(s)}} + \frac{\bar{\alpha}\,\varepsilon_{(s)}}{\gamma_{(s)}^4 \delta_{(s)}^{4\nu-1}} < \sigma_1,$$

$$\|G^{-1}_{N_{(s)},N^*_{(s)}}\|_{\rho_{(s)}} + \frac{\bar{\alpha}\,\varepsilon_{(s)}}{\gamma^4_{(s)}\delta^{4\nu-1}_{(s)}} < \sigma_1, \qquad |\left[A_{(s)}\right]^{-1}_{\mathbb{T}^r}| + \frac{\bar{\alpha}\,\varepsilon_{(s)}}{\gamma^4_{(s)}\delta^{4\nu-1}_{(s)}} < \sigma_1,$$

thus obtaining the second condition in (IV.108). Next, to verify the inequality which corresponds to the third term in (IV.108) we observe that

$$\text{dist}\,(\tau_{(s)}(\Delta(\rho_{(s)})), \partial\mathcal{U}) \geq \text{dist}\,(\tau_{(s-1)}(\Delta(\rho_{(s-1)})), \partial\mathcal{U}) - \|\Delta_{\tau_{(s-1)}}\|_{\rho_{(s)}},$$

and we use again (IV.58) and (IV.105), thus obtaining —computations are anologous as those performed for $D\tau_{(s)}$ above—

$$\text{dist}\,(\tau_{(s)}(\Delta(\rho_{(s)})), \partial\mathcal{U}) - \frac{\bar{\alpha}\,\varepsilon_{(s)}}{\gamma^4_{(s)}\delta^{4\nu-1}_{(s)}} \geq \text{dist}\,(\tau_{(0)}(\Delta(\rho_{(0)})), \partial\mathcal{U}) - \frac{2\bar{\alpha}\varepsilon_{(0)}}{\gamma^4_{(0)}\delta^{4\nu-1}_{(0)}} > \sigma_2.$$

Checking fourth and fifth conditions in (IV.108) —which involves estimates (IV.60) for the normal frequencies— is left to the reader, since it follows in the same way.

We now observe that hypotheses H_1, H_2, H_3, H_4 and H_5 are automatically (actually by inductive hypothesis) satisfied for the s-objects and Diophantine conditions in H_6 are guaranteed after defining the sets $I_{(s)}$ of "good parameters". Then, we can apply Proposition IV.5.1 in the step $s+1$, thus obtaining new objects satisfying

$$\sigma^*_{(s+1)} < \sigma_1, \quad \text{dist}\,(\tau_{(s+1)}(\Delta(\rho_{(s+1)})), \partial\mathcal{U}) > \sigma_2, \quad \sigma^{**}_{(s+1)} > \sigma_3, \quad \max_j |\lambda^{(s+1)}_j| < \frac{\sigma_4}{2}.$$

Therefore, we can apply inductively this scheme and, since the sequence of normalized errors satisfies $\varepsilon_{(s)} \overset{s\to\infty}{\longrightarrow} 0$ due to hypothesis (IV.105), we converge to a true quasi-periodic invariant torus for every μ in the set

$$I_{(\infty)} = \bigcap_{s\geq 0} I_{(s)}. \tag{IV.110}$$

Notice also that

$$\rho_{(\infty)} = \lim_{s\to\infty} \rho_{(s)} = \rho_{(0)} - 4\sum_{s=0}^{\infty} \delta_{(s)} = \rho_{(0)} - 8\delta_{(0)} = \frac{\rho_{(0)}}{2}$$

and that the limit objects are close to the initial (approximate) ones —computations are analogous as those performed to check (IV.109)—

$$\|\tau_{(\infty)} - \tau_{(0)}\|_{I_{(\infty)},\rho_{(0)}/2} \leq \sum_{j=0}^{\infty} \|\Delta_{\tau_{(j)}}\|_{\rho_{(j+1)}} \leq \frac{2\bar{\alpha}\varepsilon_{(0)}}{\gamma^2_{(0)}\delta^{2\nu}_{(0)}}, \tag{IV.111}$$

$$\|N_{(\infty)} - N_{(0)}\|_{I_{(\infty)},\rho_{(0)}/2} \leq \sum_{j=0}^{\infty} \|\Delta_{N_{(j)}}\|_{\rho_{(j+1)}} \leq \frac{2\bar{\alpha}\varepsilon_{(0)}}{\gamma^4_{(0)}\delta^{4\nu-1}_{(0)}}, \tag{IV.112}$$

$$|\lambda_{i,(\infty)} - \lambda_{i,(0)}|_{I_{(\infty)}} \leq \sum_{j=0}^{\infty} |\Delta_{\Lambda_{(j)}}| \leq \frac{2\bar{\alpha}\varepsilon_{(0)}}{\gamma_{(0)}^2 \delta_{(0)}^{2\nu-1}}, \tag{IV.113}$$

for $i = 1, \ldots, n-r$.

Remark IV.6.1. *According to these computations —see condition* (IV.105)— *we observe that the Diophantine constant* $\gamma_{(0)}$ *of the constructed invariant tori depends on the error as* $\gamma_{(0)} \geq \mathcal{O}(\varepsilon_{(0)}^{1/8})$*. However, in Section IV.6.2 —when controlling the Lipschitz regularity of the constructed objects— we will require the stronger condition* $\gamma_{(0)} \geq \mathcal{O}(\varepsilon_{(0)}^{1/10})$*. Therefore, from expressions* (IV.111) *and* (IV.112) *we obtain the bounds in* (IV.12) *stated in the theorem. Analogously, the bound in* (IV.13) *follows from* (IV.113).

IV.6.2 Lipschitz regularity

As we pointed out before, to control the measure of the set of removed parameters we cannot use any kind of smooth dependence with respect to μ, because the sets $I_{(s)}$ have empty interior. Then, following closelly [JS96], to control this measure we will use a Lipschitz condition from below with respect to μ on the eigenvalues of the matrix $\Lambda_{(s)}$, for $s \geq 0$. In order to guarantee this condition we prove that $\Lambda_{(s)}$ is Lipschitz and then, using that $\Lambda_{(s)}$ is close to $\Lambda_{(0)}$, we can ensure a-posteriori that $\Lambda_{(s)}$ is Lipschitz from below (see Lemma IV.6.4). For the sake of completeness, we provide next some basic results related to Lipschitz dependence.

Lemma IV.6.2. *Given Lipschitz functions* $f, g : I \subset \mathbb{R} \to \mathbb{C}$*, we have*

(i) $\mathrm{Lip}_I(f+g) \leq \mathrm{Lip}_I(f) + \mathrm{Lip}_I(g)$.

(ii) $\mathrm{Lip}_I(fg) \leq \mathrm{Lip}_I(f)\|g\|_I + \|f\|_I \,\mathrm{Lip}_I(g)$.

(iii) $\mathrm{Lip}_I(1/f) \leq \|1/f\|_I^2 \,\mathrm{Lip}_I(f)$*, if* f *does not vanish in* I.

Moreover, an equivalent result holds if f *and* g *take values in spaces of complex matrices (*f *must be invertible in the third item) and also for families* $\mu \mapsto f_\mu$ *of functions on* $\mathbb{T}^r$*, using* $\mathrm{Lip}_{I,\rho}(f)$ *and* $\|f\|_{I,\rho}$.

(iv) Given a family $\mu \in I \subset \mathbb{R} \mapsto f_\mu$*, where* $f_\mu : \mathcal{U} \subset \mathbb{C}^l \to \mathbb{C}$ *is an analytic function with bounded derivatives in* $\mathcal{U}$*, and given a family* $\mu : I \subset \mathbb{R} \mapsto g_\mu$*, where* $g_\mu : \mathbb{T}^r \to \mathcal{U}$ *is analytic in* $\Delta(\rho)$*, we have that*

$$\mathrm{Lip}_{I,\rho}(f \circ g) = \mathrm{Lip}_{I,\mathcal{U}}(f) + \|f\|_{I,\mathcal{U}} \mathrm{Lip}_{I,\rho}(g).$$

Proof. Items (i), (ii) and (iii) are straightforward. Item (iv) follows from applying formally the chain rule. □

Lemma IV.6.3. *Let $\mu \in I \subset \mathbb{R} \mapsto g_\mu$ be a family of functions $g_\mu : \mathbb{T}^r \to \mathbb{C}$ that are analytic in $\Delta(\rho)$ and satisfying $\mathrm{Lip}_{I,\rho}(g) < \infty$. If we expand g in Fourier series*

$$g_\mu(\theta) = \sum_{k \in \mathbb{Z}^r} \hat{g}_k(\mu) \mathrm{e}^{\mathrm{i}\langle k, \theta \rangle},$$

then, we have

(i) $\mathrm{Lip}_I(\hat{g}_k) \leq \mathrm{Lip}_{I,\rho}(g) \mathrm{e}^{-|k|_1 \rho}$.

(ii) $\mathrm{Lip}_{I,\rho-\delta}\left(\dfrac{\partial g}{\partial \theta_j}\right) \leq \dfrac{1}{\delta} \mathrm{Lip}_{I,\rho}(g)$, *for* $j = 1, \ldots, r$.

(iii) Given $\omega \in \mathbb{R}^r \backslash \{0\}$ and a Lipschitz function $d : I \subset \mathbb{R} \to \mathbb{C}$, we consider sequences $\{d_k^0\}_{k \in \mathbb{Z}^r \backslash \{0\}}$, $\{d_k^1\}_{k \in \mathbb{Z}^r}$ of complex functions of μ given by $d_k^0 = \langle k, \omega \rangle$, $d_k^1 = \langle k, \omega \rangle + d(\mu)$, satisfying $|d_k^0|, |d_k^1| \geq \gamma / |k|_1^\nu$, with $|k|_1 \neq 0$, for certain $1 > \gamma > 0$ and $\nu > r - 1$. Then, the functions f^0 and f^1 whose Fourier coefficients are given by

$$\begin{aligned} \hat{f}_k^0 &= \hat{g}_k / d_k^0, \qquad k \in \mathbb{Z}^r \backslash \{0\}, \qquad \hat{f}_0^0 = 0, \\ \hat{f}_k^1 &= \hat{g}_k / d_k^1, \qquad k \in \mathbb{Z}^r, \end{aligned}$$

satisfy

$$\begin{aligned} \mathrm{Lip}_{I,\rho-\delta}(f^0) &\leq \frac{\alpha_0}{\gamma \delta^\nu} \mathrm{Lip}_{I,\rho-\delta}(g), \\ \mathrm{Lip}_{I,\rho-\delta}(f^1) &\leq \beta_0 \bigg(\frac{\mathrm{Lip}_{I,\rho}(g)}{\gamma \delta^\nu} + \frac{\mathrm{Lip}_I(d)}{\gamma^2 \delta^{2\nu}} \bigg) + \mathrm{Lip}_{I,\rho}(g) \left\| \frac{1}{d} \right\|_I + \left\| \frac{1}{d} \right\|_I^2 \mathrm{Lip}_I(d) \|g\|_{I,\rho}, \end{aligned}$$

for any $\delta \in (0, \min\{1, \rho\})$, where $\alpha_0 \geq 1$ is the constant that appears in Lemma IV.5.2, and $\beta_0 \geq \alpha_0$ is a constant depending on r, ν and α_0.

Proof. Items (i) and (ii) are straightforward (see [JS96]). Item (iii) follows from the same arguments used in Lemma IV.5.2 and applying the properties in Lemma IV.6.2. □

Lemma IV.6.4. *Let $f : I \to \mathbb{C}$ be a function satisfying*

$$\mathrm{lip}_I(f) |\mu_2 - \mu_1| \leq |f(\mu_2) - f(\mu_1)| \leq \mathrm{Lip}_I(f) |\mu_2 - \mu_1|,$$

and let $g : I \to \mathbb{C}$ be a Lipschitz function with constant $\mathrm{Lip}_I(d) < \mathrm{lip}_I(f)$. Then, the function $h = f + g$ is Lipschitz with constant $\mathrm{lip}_I(h) = \mathrm{lip}_I(f) - \mathrm{Lip}_I(g)$ and Lipschitz from below with constant $\mathrm{Lip}_I(h) = \mathrm{Lip}_I(f) + \mathrm{Lip}_I(g)$.

Proof. It is straightforward. □

Now, we use these elementary results to control recursively the Lipschitz dependence of the objects that we constructed in Proposition IV.5.1. To this end, we obtain an upgraded version of Proposition IV.5.1.

Lemma IV.6.5 (Addenda to Proposition IV.5.1)**.** *Let us consider a Lipschitz family of Hamiltonian systems $\mu \in I \subset \mathbb{R} \mapsto h_\mu$, where I is not necessarily an interval, with $h_\mu : U \subset \mathbb{R}^{2n} \to \mathbb{R}$, and a vector of basic frequencies $\omega \in \mathbb{R}^r$. Assume that there exist families $\mu \in I \mapsto \tau_\mu, N_\mu, \Lambda_\mu, e_\mu, R_\mu$ satisfying all the hypotheses of Theorem IV.5.1 and also that*

$$\mathrm{Lip}_{I,\mathcal{C}^3,\mathcal{U}}(h), \mathrm{Lip}_{I,\rho}(\tau), \mathrm{Lip}_{I,\rho}(D\tau), \mathrm{Lip}_{I,\rho}(N), \mathrm{Lip}_I(\Lambda) < \sigma_5.$$

Then, if

$$\varepsilon := \frac{\mathrm{Lip}_{I,\rho}(e) + \|e\|_{I,\rho}}{\delta} + \mathrm{Lip}_{I,\rho}(R) + \|R\|_{I,\rho}$$

is sufficiently small —quantitative estimates are obtained analogously as those in Proposition IV.5.1— we have that the families $\mu \mapsto D\bar{\tau}_\mu, \bar{N}_\mu, \bar{\Lambda}_\mu, \bar{e}_\mu, \bar{R}_\mu$ obtained in the output Proposition IV.5.1 satisfy

$$\mathrm{Lip}_{I,\rho-2\delta}(\bar{\tau}), \mathrm{Lip}_{I,\rho-3\delta}(D\bar{\tau}), \mathrm{Lip}_{I,\rho-4\delta}(\bar{N}), \mathrm{Lip}_I(\bar{\Lambda}) < \sigma_5, \tag{IV.114}$$

and

$$\mathrm{Lip}_{I,\rho-3\delta}(\bar{e}) \leq \frac{\bar{\beta}\varepsilon^2}{\gamma^4\delta^{4\nu-2}}, \qquad \mathrm{Lip}_{I,\rho-4\delta}(\bar{R}) \leq \frac{\bar{\beta}\varepsilon^2}{\gamma^{10}\delta^{10\nu-2}},$$

for some constant $\bar{\beta} > 1$ that depends on r, n, ν, $|h_\mu|_{\mathcal{C}^3,\mathcal{U}}$, $|\omega|$, σ_1, σ_2, σ_3, σ_4 and σ_5.

Proof. Basically, it consists in using the properties in Lemmata IV.6.2 and IV.6.3 to control the Lipschitz constant of the different functions that appear along the proofs of Propositions IV.5.3 (construction of the approximately symplectic basis) and Proposition IV.5.1 (iterative procedure).

First, let us study if objects in Proposition IV.5.3. To this end, we find a constant $\hat{\beta}$ (which is enlarged along the proof in order to include dependence on r, n, ν, $|h_\mu|_{\mathcal{C}^3,\mathcal{U}}$, $|\omega|$, σ_1, σ_2, σ_3, σ_4 and σ_5) such that

$$\mathrm{Lip}_{I,\rho}(G_{D\tau}^{-1}) \leq \|G_{D\tau}^{-1}\|_{I,\rho}^2 \mathrm{Lip}_{I,\rho}(G_{D\tau}) \leq 2\|G_{D\tau}^{-1}\|_{I,\rho}^2 \|D\tau\|_{I,\rho} \mathrm{Lip}_{I,\rho}(D\tau) \leq \hat{\beta},$$

(similar computations allows us to control $\mathrm{Lip}_{I,\rho}(G_{N,N^*}^{-1}) \leq \hat{\beta}$ and $\mathrm{Lip}_{I,\rho}([A_1]_{\mathbb{T}^r}^{-1}) \leq \hat{\beta}$) and also such that

$$\mathrm{Lip}_{I,\rho}(T_1), \mathrm{Lip}_{I,\rho}(T_2), \mathrm{Lip}_{I,\rho}(A_1), \mathrm{Lip}_{I,\rho}(V), \mathrm{Lip}_{I,\rho}(B_i) \leq \hat{\beta},$$

$$\mathrm{Lip}_{I,\rho-2\delta}(L_\omega B_i) \leq \frac{\hat{\beta}}{\delta}$$

for $i = 1, 2, 3$.

Then, we estimate Lipschitz constants for the matrixes $\Omega_{D\tau}$, Ω_{N_1}, ..., Ω_V that characterize the approximately symplectic character of the basis in Propositions IV.5.3. For example, we get $\mathrm{Lip}_{I,\rho-2\delta}(\Omega_{D\tau})$ by applying item (iii) of Lemma IV.6.3 to the (i, j)-component of $\Omega_{D\tau}$ obtained from equation (IV.85), i.e., taking $d_k^0 = \langle \omega, k\rangle$ and $g = -\mathrm{i}(\Omega_{De,D\tau} + \Omega_{D\tau,De})^{(i,j)}$, thus obtaining

$$\mathrm{Lip}_{I,\rho-\delta}(\Omega_{D\tau}) \leq \frac{\hat{\beta}}{\gamma\delta^{\nu+1}}(\mathrm{Lip}_{I,\rho}(e) + \|e\|_{I,\rho})$$

and it is straightforward to check that

$$\mathrm{Lip}_{I,\rho-2\delta}(\Omega_{N_i}) \leq \frac{\hat{\beta}}{\gamma^2\delta^{2\nu}}(\mathrm{Lip}_{I,\rho}(R) + \|R\|_{I,\rho}),$$
$$\mathrm{Lip}_{I,\rho-2\delta}(\Omega_{N_2,N_1} - \mathrm{Id}_{n-r}) \leq \frac{\hat{\beta}}{\gamma^2\delta^{2\nu}}(\mathrm{Lip}_{I,\rho}(R) + \|R\|_{I,\rho}),$$

and also

$$\mathrm{Lip}_{I,\rho-2\delta}(\Omega_{D\tau,N_i}) \leq \frac{\hat{\beta}\varepsilon}{\gamma^2\delta^{2\nu}}, \qquad \mathrm{Lip}_{I,\rho-2\delta}(\Omega_{V,D\tau} - \mathrm{Id}_r) \leq \frac{\hat{\beta}\varepsilon}{\gamma^2\delta^{2\nu}},$$
$$\mathrm{Lip}_{I,\rho-2\delta}(\Omega_{V,N_i}) \leq \frac{\hat{\beta}\varepsilon}{\gamma^2\delta^{2\nu}}, \qquad \mathrm{Lip}_{I,\rho-2\delta}(\Omega_V) \leq \frac{\hat{\beta}\varepsilon}{\gamma^2\delta^{2\nu}},$$

for $i = 1, 2$. Furthermore, by performing similar computations to estimate the Lipschitz constants of M_3 in (IV.91), C_1^+ in (IV.94), C_2^+ in (IV.93) and C_3^+ in (IV.92) we obtain that the functions A_i^+, for $i = 1, \ldots, 4$, in the statement of Proposition IV.5.3 are controlled by

$$\mathrm{Lip}_{I,\rho-3\delta}(A_i^+) \leq \frac{\hat{\beta}\varepsilon}{\gamma^2\delta^{2\nu+1}},$$

provided ε is small enough.

Now we can estimate the Lipschitz constant of Δ_i, $i = 1, \ldots, 4$, defined as the solutions of cohomological equations (IV.39)-(IV.42). In analogy with the notation in Proposition IV.5.1, we introduce a constant $\bar{\beta}$ that depends on r, n, ν, $|h_\mu|_{\mathcal{C}^3,\mathcal{U}}$, $|\omega|$, σ_1, σ_2, σ_3, σ_4 and σ_5. It is straightforward to obtain

$$\mathrm{Lip}_{I,\rho-2\delta}(\Delta_1) \leq \frac{\hat{\beta}}{\gamma^2\delta^{2\nu}}(\mathrm{Lip}_{I,\rho}(e) + \|e\|_{I,\rho}),$$
$$\mathrm{Lip}_{I,\rho-\delta}(\Delta_2) \leq \frac{\hat{\beta}}{\gamma\delta^{\nu}}(\mathrm{Lip}_{I,\rho}(e) + \|e\|_{I,\rho}),$$
$$\mathrm{Lip}_{I,\rho-\delta}(\Delta_i) \leq \frac{\hat{\beta}}{\gamma^2\delta^{2\nu}}(\mathrm{Lip}_{I,\rho}(e) + \|e\|_{I,\rho}),$$

for $i = 3, 4$. Then, the Lipschitz character of the new embedding $\bar{\tau} = \tau + \Delta_\tau$ can be controlled as

$$\mathrm{Lip}_{I,\rho-2\delta}(\bar{\tau}) \leq \mathrm{Lip}_{I,\rho}(\tau) + \frac{\hat{\beta}\varepsilon}{\gamma^2\delta^{2\nu-1}},$$

and we observe that for ε small enough, we obtain (IV.114) (and also for the derivative $D\bar{\tau}$). Similarly we control the correction in the normal eigenvalues as

$$\mathrm{Lip}_I(\Delta_\Lambda) \leq \frac{\hat{\beta}\varepsilon}{\gamma^2\delta^{2\nu-1}}.$$

In order to avoid an unnecessary extension in the length of the document, we leave to the reader the intermediate computations required to control the Lipschitz constant of the new error in reducibility $\bar{R}$. These computations are analogous as those performed above. □

Taking into account Lemma IV.6.5, we repeat the computations in Section IV.6.1 in order to guarantee that the scheme converges. To characterize the Lipschitz dependence of the initial family of approximate tori we use hypothesis H_6 of Theorem IV.3.1 and so, we can control the Lipschitz constants of the functions $h(\mu)$, $D\tau_{(0)}(\mu)$, $N_{(0)}(\mu)$ and $\Lambda_{(0)}(\mu)$ by σ_6. Then, it is straightforward to see that the new quadratic procedure converges by asking for $\gamma_{(0)} \geq \hat{C}\varepsilon_0^{1/10}$, where $\hat{C}$ depends on ν, r, n, $|h_\mu|_{\mathcal{C}^3,\mathcal{U}}$, $|\omega|$, σ_1, σ_2, σ_3, σ_4 and σ_5.

Finally, we show that the functions

$$\mu \mapsto \lambda_{i,(s)}(\mu), \qquad \mu \mapsto \lambda_{i,(s)}(\mu) \pm \lambda_{j,(s)}(\mu),$$

for $i \neq j = 1, \ldots, n-r$, are Lipschitz from below with constants that do not depend on the step s —notice that Lipschitz (from above) constants are controlled for every s as (IV.114). For example, we have that

$$\Lambda_{(s)} - \Lambda_{(0)} = \sum_{j=0}^{s-1} \Delta_{\Lambda_{(j)}},$$

so we obtain

$$\mathrm{Lip}_{I_{(s)}}(\Lambda_{(s)} - \Lambda_{(0)}) \leq \sum_{j=0}^{s-1} \mathrm{Lip}_{I_{(j)}}(\Delta_{\Lambda_{(j)}}) \leq \frac{2\bar{\beta}\varepsilon_{(0)}}{\gamma_{(0)}^2\delta_{(0)}^{2\nu-1}}.$$

Hence, we can apply Lemma IV.6.4 taking the functions $f(\mu) = \lambda_{i,(0)}(\mu)$ and $g(\mu) = \lambda_{i,(s)}(\mu) - \lambda_{i,(0)}(\mu)$, for $i = 1, \ldots, n-r$, and provided

$$\frac{2\bar{\beta}\varepsilon_{(0)}}{\gamma_{(0)}^4\delta_{(0)}^{4\nu-1}} < \frac{\sigma_6}{2},$$

(we have used H_6 in Theorem IV.3.1 to control the Lipschitz constant from below of f) we conclude that $\mu \mapsto \lambda_{i,(s)}(\mu)$ is Lipschitz from below. Notice that, in order to achieve this condition have to ask for $\gamma_{(0)} \geq C\varepsilon_0^{1/10}$, where $C \geq \hat{C}$ depends on ν, r, n, $|h_\mu|_{\mathcal{C}^3,\mathcal{U}}$, $|\omega|$, σ_1, σ_2, σ_3, σ_4, σ_5 and σ_6. The same arguments apply for $\mu \mapsto \lambda_{i,(s)}(\mu) \pm \lambda_{j,(s)}(\mu)$.

IV.6.3 Measure of the set of excluded parameters

It remains to bound the measure of the set $I_{(\infty)}$, given by (IV.110), for which all steps performed along the iterative procedure of Section IV.6.1 are well-posed. Let us recall that $I_{(\infty)}$ is constructed by taking out, in recursive form, the set of parameters μ for which (IV.4) and (IV.10) do not hold. Instead of controlling $I_{(\infty)}$, it is typical to bound the measure of the complementary set

$$I \backslash I_{(\infty)} = \bigcup_{s \geq 0} I_{(s-1)} \backslash I_{(s)}.$$

In order to simplify the notation, we assume that we have divisors of the form $|\langle \omega, k \rangle + d_{(s)}(\mu)|$ in the definition of $I_{(s)}$, where d is Lipschitz from below, and then we multiply the measure of the obtained set by $(n-r)^2$ to take into account all possible combinations of normal frequencies (i.e., applying the arguments to $d = \lambda_j$ and $d = \lambda_j \pm \lambda_i$, for $i, j = 1, \ldots, n-r$, $i \neq j$). Given $\gamma_{(0)} > 0$ and $\nu > r-1$, let us introduce the *k-th resonant set at the s-th step* as

$$\mathrm{Res}_k^{(s)} = \left\{ \mu \in I_{(s-1)} \;:\; |\langle \omega, k \rangle + d_{(s)}(\mu)| < \frac{(1+2^{-s})\gamma_{(0)}}{|k|_1^{\nu}} \right\},$$

and we claim that for every $k \in \mathbb{Z}^r \backslash \{0\}$ fixed, there exists $s^*(|k|_1) \in \mathbb{N}$ such that if $s \geq s^*(|k|_1)$ then $\mathrm{Res}_k^{(s)} = \varnothing$. Hence, the measure that we are looking for can me computed as follows

$$\mathrm{meas}_{\mathbb{R}}(I \backslash I_{(\infty)}) = \sum_{k \in \mathbb{Z}^r \backslash \{0\}} \sum_{s=1}^{\infty} \mathrm{meas}_{\mathbb{R}}(\mathrm{Res}_k^{(s)}) = \sum_{k \in \mathbb{Z}^r \backslash \{0\}} \sum_{s=1}^{s^*(|k|_1)} \mathrm{meas}_{\mathbb{R}}(\mathrm{Res}_k^{(s)}). \tag{IV.115}$$

To prove the claim, we simply observe that the correction of the normal frequencies —in general, of $d_{(s)}(\mu)$— is smaller at each step of the iterative procedure. Actually, we have that (recall the computations performed in Section IV.6.1)

$$\|d_{(s)} - d_{(s-1)}\|_{I_{(s)}} \leq \gamma_{(0)}^6 \delta_{(0)}^{6\nu-2} 2^{-(8\nu-2)s-6\nu+1} \widetilde{\varepsilon}^{2^s}, \qquad \widetilde{\varepsilon} = \frac{\bar{\alpha} 2^{8\nu-1} \varepsilon_{(0)}}{\gamma_{(0)}^8 \delta_{(0)}^{8\nu-2}}. \tag{IV.116}$$

Then, it turns out that if $\mu \notin \mathrm{Res}_k^{(s-1)}$ it must be $\mu \notin \mathrm{Res}_k^{(s)}$ provided $\widetilde{\varepsilon}$ is sufficiently small. Indeed, we are asking the quantity

$$|\langle \omega, k \rangle + d_{(s)}(\mu)| \geq (1 + 2^{-s+1}) \gamma_{(0)} |k|_1^{-\nu} - \|d_{(s)} - d_{(s-1)}\|_{I_{(s)}}$$

to be larger that $(1+2^{-s})\gamma_{(0)}|k|_1^{-\nu}$, and using (IV.116), this is equivalent to ask for

$$\gamma_{(0)}^6 \delta_{(0)}^{6\nu-2} 2^{-(8\nu-2)s-6\nu+1} \widetilde{\varepsilon}^{2^s} \leq 2^{-s} \gamma_{(0)} |k|_1^{-\nu}.$$

This inequality is satisfied for every $s \in \mathbb{N}$ such that $\widetilde{\varepsilon}^{2^s} \leq |k|_1^{-\nu}$, or equivalently, for every $s \geq s^*(|k|_1) \in \mathbb{N}$, where

$$s^*(|k|_1) = \left\lfloor \log_2 \left(-\nu \frac{\log |k|_1}{\log \widetilde{\varepsilon}} \right) \right\rfloor + 1$$

Furthermore, let us remark that, for every fixed $\widetilde{\varepsilon}$, the expression for $s^*(|k|_1)$ only makes sense if $|k|_1 > \widetilde{\varepsilon}^{-1/\nu}$ so from now on we omit the corresponding terms in the sum (IV.115).

The last ingredient is to estimate the measure of a given resonant set. To this end, we use the Lipschitz property and geometric properties of Diophantine number (see [dlL01]), thus obtaining

$$\mathrm{meas}_{\mathbb{R}}(\mathrm{Res}_k^{(s)}) \leq \frac{4\sqrt{2}\gamma_{(0)}}{\mathrm{lip}_{I_{(\infty)}}(d)}|k|_1^{-(\nu+1)} < \frac{2\sqrt{2}\gamma_{(0)}}{\sigma_6}|k|_1^{-(\nu+1)}.$$

Then, we control (IV.115) as follows

$$\begin{aligned}
\mathrm{meas}_{\mathbb{R}}(I\backslash I_{(\infty)}) &< \sum_{\substack{k\in\mathbb{Z}^r\backslash\{0\}\\ |k|_1\leq\widetilde{\varepsilon}^{-1/\nu}}} \sum_{s=1}^{s^*(|k|_1)} \frac{2\sqrt{2}\gamma_{(0)}}{\sigma_6}|k|_1^{-(\nu+1)} \leq \sum_{l=\varepsilon^{-1/\nu}}^{\infty}\sum_{s=1}^{s^*(l)} \frac{2\sqrt{2}\gamma_{(0)}C(r)}{\sigma_6} l^{l-\nu-2} \\
&\leq \sum_{l=\varepsilon^{-1/\nu}}^{\infty} \left(\log_2\left(-\nu\frac{\log l}{\log\widetilde{\varepsilon}}\right)+1\right)\frac{2\sqrt{2}\gamma_{(0)}C(r)}{\sigma_6} l^{l-\nu-2} \\
&\leq -\frac{4\sqrt{2}\gamma_{(0)}C(r)}{\sigma_6\log\widetilde{\varepsilon}}\int_{\varepsilon^{-1/\nu}}^{\infty}\frac{\log x}{x^{\nu-r+2}}dx \leq \frac{4\sqrt{2}\gamma_{(0)}C(r)}{\sigma_6(\nu-r+1)}\left(\frac{\bar{\alpha}2^{8\nu-1}\varepsilon_{(0)}}{\gamma_{(0)}^8\delta_{(0)}^{8\nu-2}}\right)^{\frac{\nu-r+1}{\nu}} \\
&\leq \frac{4\sqrt{2}C(r)}{\sigma_6(\nu-r+1)}\left(\frac{\bar{\alpha}2^{8\nu-1}}{\delta_{(0)}^{8\nu-2}}\right)^{\frac{\nu-r+1}{\nu}}\gamma_{(0)}^{\frac{-7\nu+8r-8}{\nu}}\varepsilon_0^{\frac{\nu-r+1}{\nu}} \leq \alpha\varepsilon_0^{\beta},
\end{aligned}$$

where α —which depends on ν, r, n, σ_0, σ_1, σ_2, σ_3, σ_4, σ_5 and σ_6— and $\beta = \frac{3\nu-2r+2}{10\nu}$ are the constants that appear in equation (IV.11) of Theorem IV.3.1.

Bibliography

[AdlLP05] A. Apte, R. de la Llave, and N.P. Petrov. Regularity of critical invariant circles of the standard nontwist map. *Nonlinearity*, 18(3):1173–1187, 2005.

[AM78] R. Abraham and J.E. Marsden. *Foundations of mechanics*. Benjamin/Cummings Publishing Co. Inc. Advanced Book Program, Reading, Mass., 1978.

[AP90] D.K. Arrowsmith and C.M. Place. *An introduction to dynamical systems*. Cambridge University Press, 1990.

[Arn61] V.I. Arnold. Small denominators. I. Mapping the circle onto itself. *Izv. Akad. Nauk SSSR Ser. Mat.*, 25:21–86, 1961.

[Arn63] V.I. Arnol'd. Proof of a theorem of A. N. Kolmogorov on the invariance of quasi-periodic motions under small perturbations. *Russian Math. Surveys*, 18(5):9–36, 1963.

[Arn88] V.I. Arnold. *Geometrical methods in the theory of ordinary differential equations*. Springer-Verlag, New York, second edition, 1988.

[Arn89] V.I. Arnol'd. *Mathematical methods of classical mechanics*, volume 60 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1989.

[BGGS84] G. Benettin, L. Galgani, A. Giorgilli, and J.-M. Strelcyn. A proof of Kolmogorov's theorem on invariant tori using canonical transformations defined by the Lie method. *Nuovo Cimento B (11)*, 79(2):201–223, 1984.

[BHS96] H.W. Broer, G.B. Huitema, and M.B. Sevryuk. *Quasi-periodic motions in families of dynamical systems*. Lecture Notes in Mathematics. Springer-Verlag, Berlin, 1996. Order amidst chaos.

[BHTB90] H. W. Broer, G. B. Huitema, F. Takens, and B. L. J. Braaksma. Unfoldings and bifurcations of quasi-periodic tori. *Mem. Amer. Math. Soc.*, 83(421):viii+175, 1990.

[Bos86] J.B. Bost. Tores invariants des systèmes dynamiques hamiltoniens (d'après Kolmogorov, Arnold, Moser, Rüssmann, Zehnder, Herman, Pöschel, . . .). *Astérisque*, (133-134):113–157, 1986. Seminar Bourbaki, Vol. 1984/85.

[Bou97] J. Bourgain. On Melnikov's persistency problem. *Math. Res. Lett.*, 4(4):445–458, 1997.

[Bou98] J. Bourgain. Quasi-periodic solutions of hamiltonian perturbations of 2d linear schrdinger equations. *Ann. of Math.*, 2(148):363–439, 1998.

[Bru92] H. Bruin. Numerical determination of the continued fraction expansion of the rotation number. *Phys. D*, 59(1-3):158–168, 1992.

[Bry91] R.L. Bryant. An introduction to lie groups and symplectic geometry, 1991. A series of nine lectures on Lie groups and symplectic geometry delivered at the Regional Geometry Institute in Park City, Utah.

[BS00] H. Broer and C. Simó. Resonance tongues in Hill's equations: a geometric approach. *J. Differential Equations*, 166(2):290–327, 2000.

[CC07] A. Celletti and L. Chierchia. KAM stability and celestial mechanics. *Mem. Amer. Math. Soc.*, 187(878):viii+134, 2007.

[CdlL09] R. Calleja and R. de la Llave. Fast numerical computation of quasi-periodic equilibrium states in 1-d statistical mechanics. *Nonlinearity*, 22(1):1311–1336, 2009.

[CJ00] E. Castellà and À. Jorba. On the vertical families of two-dimensional tori near the triangular points of the bicircular problem. *Celestial Mech. Dynam. Astronom.*, 76(1):35–54, 2000.

[dCNGM96] D. del Castillo-Negrete, J.M. Greene, and P.J. Morrison. Area preserving nontwist maps: periodic orbits and transition to chaos. *Phys. D*, 91(1-2):1–23, 1996.

[DdlL00] A. Delshams and R. de la Llave. KAM theory and a partial justification of Greene's criterion for nontwist maps. *SIAM J. Math. Anal.*, 31(6):1235–1269 (electronic), 2000.

[DG96a] A. Delshams and P. Gutiérrez. Effective stability and KAM theory. *J. Differential Equations*, 128(2):415–490, 1996.

[DG96b] A. Delshams and P. Gutiérrez. Estimates on invariant tori near an elliptic equilibrium point of a Hamiltonian system. *J. Differential Equations*, 131(2):277–303, 1996.

[dlL01] R. de la Llave. A tutorial on KAM theory. In *Smooth ergodic theory and its applications (Seattle, WA, 1999)*, volume 69 of *Proc. Sympos. Pure Math.*, pages 175–292. Amer. Math. Soc., 2001.

[dlLGJV05] R. de la Llave, A. González, À. Jorba, and J. Villanueva. KAM theory without action-angle variables. *Nonlinearity*, 18(2):855–895, 2005.

[dlLHS] R. de la Llave, G. Huguet, and Y. Sire. Fast numerical algorithms for the computation of invariant tori in Hamiltonian Systems. Preprint available electronically at http://www.ma.utexas.edu/mp_arc-bin/mpa?yn=09-2.

[dlLO06] Rafael de la Llave and Arturo Olvera. The obstruction criterion for non-existence of invariant circles and renormalization. *Nonlinearity*, 19(8):1907–1937, 2006.

[dlLP02] R. de la Llave and N.P. Petrov. Regularity of conjugacies between critical circle maps: an experimental study. *Experiment. Math.*, 11(2):219–241, 2002.

[dlLW04] R. de la Llave and C. E. Wayne. Whiskered and low dimensional tori in nearly integrable Hamiltonian systems. *Math. Phys. Electron. J.*, 10:Paper 5, 45 pp. (electronic), 2004.

[DMS00] H.R. Dullin, J.D. Meiss, and D. Sterling. Generic twistless bifurcations. *Nonlinearity*, 13(1):203–224, 2000.

[dMvS93] W. de Melo and S. van Strien. *One-dimensional dynamics*, volume 25 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]*. Springer-Verlag, Berlin, 1993.

[Eli88] L.H. Eliasson. Perturbations of stable invariant tori for Hamiltonian systems. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 15(1):115–147 (1989), 1988.

[Eli92] L.H. Eliasson. Floquet solutions for the 1-dimensional quasi-periodic Schrödinger equation. *Comm. Math. Phys.*, 146(3):447–482, 1992.

[Eli01] L.H. Eliasson. Almost reducibility of linear quasi-periodic systems. In *Smooth ergodic theory and its applications (Seattle, WA, 1999)*, volume 69 of *Proc. Sympos. Pure Math.*, pages 679–705. Amer. Math. Soc., Providence, RI, 2001.

[FdlLS09] E. Fontich, R. de la Llave, and Y. Sire. Construction of invariant whiskered tori by a parametrization method. Part I: Maps and flows in finite dimensions. *J. Differential Equations*, 246:3136–3213, 2009.

[Féj04] Jacques Féjoz. Démonstration du ‘théorème d’Arnold’ sur la stabilité du système planétaire (d’après Herman). *Ergodic Theory Dynam. Systems*, 24(5):1521–1582, 2004.

[Gal83] G. Gallavotti. Perturbation theory for classical Hamiltonian systems. In *Scaling and self-similarity in physics (Bures-sur-Yvette, 1981/1982)*, volume 7 of *Progr. Phys.*, pages 359–426. Birkhäuser Boston, Boston, MA, 1983.

[GdlL08] A. González and R. de la Llave. Analytic smoothing of geometric maps with applications to KAM theory. *J. Differential Equations*, 245(5):1243–1298, 2008.

[GdlLH] A. González, R. de la Llave, and A. Haro. Non-twist KAM theory. In preparation.

[GJMS93] G. Gómez, À. Jorba, J. Masdemont, and C. Simó. *Study of Poincaré maps for orbits near Lagrangian points*. ESA-ESOC contract 8711/91/D/IM/(SC), Darmstadt, Germany, 1993.

[GJSM01a] G. Gómez, À. Jorba, C. Simó, and J. Masdemont. *Dynamics and mission design near libration points. Vol. III*, volume 4 of *World Scientific Monograph Series in Mathematics*. World Scientific Publishing Co. Inc., River Edge, NJ, 2001. Advanced methods for triangular points.

[GJSM01b] G. Gómez, À. Jorba, C. Simó, and J. Masdemont. *Dynamics and mission design near libration points. Vol. IV*, volume 5 of *World Scientific Monograph Series in Mathematics*. World Scientific Publishing Co. Inc., River Edge, NJ, 2001. Advanced methods for triangular points.

[GMS] G. Gómez, J.M. Mondelo, and C. Simó. *Refined Fourier analysis: procedures, error estimates and applications.* Preprint 2001, available electronically at `http://www.maia.ub.es/dsg/2001/index.html`.

[Gol01] C. Golé. *Symplectic Twist Maps: Global Variational Techniques*. World Scientific Publishing, 2001.

[Gra74] S.M. Graff. On the conservation of hyperbolic invariant tori for Hamiltonian systems. *J. Differential Equations*, 15:1–69, 1974.

[Gre79] J.M. Greene. A method for determining a stochastic transition. *J. Math. Phys.*, 20:1183–1201, 1979.

[HdlL06] À. Haro and R. de la Llave. Manifolds on the verge of a hyperbolicity breakdown. *Chaos*, 16(1):013120, 8, 2006.

[HdlL07] A. Haro and R. de la Llave. A parameterization method for the computation of invariant tori and their whiskers in quasi-periodic maps: explorations and mechanisms for the breakdown of hyperbolicity. *SIAM J. Appl. Dyn. Syst.*, 6(1):142–207 (electronic), 2007.

[Her77] M.R. Herman. Mesure de Lebesgue et nombre de rotation. In *Geometry and topology (Proc. III Latin Amer. School of Math., Inst. Mat. Pura Aplicada CNPq, Rio de Janeiro, 1976)*, pages 271–293. Lecture Notes in Math., Vol 597. Springer, Berlin, 1977.

[Her79] M.R. Herman. Sur la conjugaison différentiable des difféomorphismes du cercle à des rotations. *Inst. Hautes Études Sci. Publ. Math.*, (49):5–233, 1979.

[Her83] M.R. Herman. *Sur les courbes invariantes par les difféomorphismes de l'anneau. Vol. 1*, volume 103 of *Astérisque*. Société Mathématique de France, Paris, 1983. With an appendix by Albert Fathi, With an English summary.

[HI04] John Hubbard and Yulij Ilyashenko. A proof of Kolmogorov's theorem. *Discrete Contin. Dyn. Syst.*, 10(1-2):367–385, 2004.

[HL00] D. Huang and Z. Liu. On the persistence of lower-dimensional invariant hyperbolic tori for smooth Hamiltonian systems. *Nonlinearity*, 13(1):189–202, 2000.

[HLB05] Y. Hida, X. Li, and D. H. Bailey. *QD (quad-double/double-double computation package)*, 2005. Available electronically at `http://crd.lbl.gov/~dhbailey/mpdist/`.

[HLY06] Y. Han, Y. Li, and Y. Yi. Degenerate lower-dimensional tori in Hamiltonian systems. *J. Differential Equations*, 227(2):670–691, 2006.

[HY08] Hai-Long Her and Jiangong You. Full measure reducibility for generic one-parameter family of quasi-periodic linear systems. *J. Dynam. Differential Equations*, 20(4):831–866, 2008.

[HZ94] H. Hofer and E. Zehnder. *Symplectic invariants and Hamiltonian dynamics*. Birkhäuser Verlag, Basel, 1994.

[JdlLZ99] À. Jorba, R. de la Llave, and M. Zou. Lindstedt series for lower-dimensional tori. In *Hamiltonian systems with three or more degrees of freedom (S'Agaró, 1995)*, volume 533 of *NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci.*, pages 151–167. 1999.

[JM82] R. Johnson and J. Moser. The rotation number for almost periodic potentials. *Comm. Math. Phys.*, 84(3):403–438, 1982.

[JO] À. Jorba and E. Olmedo. On the computation of reducible invariant tori in a parallel computer. Preprint available electronically at `http://www.maia.ub.es/dsg/2008/index.shtml`.

[JRRV97] À. Jorba, R. Ramírez-Ros, and J. Villanueva. Effective reducibility of quasi-periodic linear equations close to constant coefficients. *SIAM J. Math. Anal.*, 28(1):178–188, 1997.

[JS81] R. A. Johnson and G. R. Sell. Smoothness of spectral subbundles and reducibility of quasiperiodic linear differential systems. *J. Differential Equations*, 41(2):262–288, 1981.

[JS92] À. Jorba and C. Simó. On the reducibility of linear differential equations with quasiperiodic coefficients. *J. Differential Equations*, 98(1):111–124, 1992.

[JS96] À. Jorba and C. Simó. On quasi-periodic perturbations of elliptic equilibrium points. *SIAM J. Math. Anal.*, 27(6):1704–1737, 1996.

[JS98] J.V. José and E.J. Saletan. *Classical dynamics*. Cambridge University Press, Cambridge, 1998. A contemporary approach.

[JV97a] À. Jorba and J. Villanueva. On the normal behaviour of partially elliptic lower-dimensional tori of Hamiltonian systems. *Nonlinearity*, 10(4):783–822, 1997.

[JV97b] À. Jorba and J. Villanueva. On the persistence of lower-dimensional invariant tori under quasi-periodic perturbations. *J. Nonlinear Sci.*, 7(5):427–473, 1997.

[JV98] À. Jorba and J. Villanueva. Numerical computation of normal forms around some periodic orbits of the restricted three-body problem. *Phys. D*, 114(3-4):197–229, 1998.

[KH95] A. Katok and B. Hasselblatt. *Introduction to the modern theory of dynamical systems*, volume 54 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, 1995.

[KO89] Y. Katznelson and D. Ornstein. The differentiability of the conjugation of certain diffeomorphisms of the circle. *Ergodic Theory Dynam. Systems*, 9(4):643–680, 1989.

[Kol54] A. N. Kolmogorov. On conservation of conditionally periodic motions for a small change in Hamilton's function. *Dokl. Akad. Nauk SSSR (N.S.)*, 98:527–530, 1954.

[KP02] S.G. Krantz and H.R. Parks. *A primer of real analytic functions*. Birkhäuser Advanced Texts: Basler Lehrbücher. [Birkhäuser Advanced Texts: Basel Textbooks]. Birkhäuser Boston Inc., Boston, MA, second edition, 2002.

[Kri99] R. Krikorian. Réductibilité presque partout des flots fibrés quasi-périodiques à valeurs dans des groupes compacts. *Ann. Sci. École Norm. Sup. (4)*, 32(2):187–240, 1999.

[Kuk88] S.B. Kuksin. Perturbation of conditionally periodic solutions of infinite-dimensional Hamiltonian systems. *Izv. Akad. Nauk SSSR Ser. Mat.*, 52(1):41–63, 240, 1988.

[Laz93] V.F. Lazutkin. *KAM theory and semiclassical approximations to eigenfunctions*. Springer-Verlag, Berlin, 1993.

[LFC92] J. Laskar, C. Froeschlé, and A. Celletti. The measure of chaos by the numerical analysis of the fundamental frequencies. Application to the standard mapping. *Phys. D*, 56(2-3):253–269, 1992.

[LM88] P. Lochak and C. Meunier. *Multiphase Averaging for Classical Systems*, volume 72 of *Appl. Math. Sci.* Springer, New York, 1988.

[LV08] A. Luque and J. Villanueva. Computation of derivatives of the rotation number for parametric families of circle diffeomorphisms. *Phys. D*, 237(20):2599–2615, 2008.

[LV09] A. Luque and J Villanueva. Numerical computation of rotation numbers for quasi-periodic planar curves. *Phys. D*, 238(20):2025–2044, 2009.

[LY05] Y. Li and Y. Yi. Persistence of hyperbolic tori in Hamiltonian systems. *J. Differential Equations*, 208(2):344–387, 2005.

[Mel65] V. K. Melnikov. On certain cases of conservation of almost periodic motions with a small change of the Hamiltonian function. *Dokl. Akad. Nauk SSSR*, 165:1245–1248, 1965.

[Mic95] L. Michelotti. *Intermediate classical dynamics with applications to beam physics*. Wiley Series in Beam Physics and Accelerator Technology. John Wiley & Sons Inc., New York, 1995. A Wiley-Interscience Publication.

[Mos62] J. Moser. On invariant curves of area-preserving mappings of an annulus. *Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. II*, 1962:1–20, 1962.

[Mos67] J. Moser. Convergent series expansions for quasi-periodic motions. *Math. Ann.*, 169:136–176, 1967.

[Neĭ84] A. I. Neĭshtadt. The separation of motions in systems with rapidly rotating phase. *Prikl. Mat. Mekh.*, 48(2):197–204, 1984.

[Nek94] N. N. Nekhoroshev. The Poincaré-Lyapunov-Liouville-Arnol′d theorem. *Funktsional. Anal. i Prilozhen.*, 28(2):67–69, 1994.

[OP08] Arturo Olvera and Nikola P. Petrov. Regularity properties of critical invariant circles of twist maps, and their universality. *SIAM J. Appl. Dyn. Syst.*, 7(3):962–987, 2008.

[OS87] Arturo Olvera and Carles Simó. An obstruction method for the destruction of invariant curves. *Phys. D*, 26(1-3):181–192, 1987.

[Par] *PARI/GP Development Headquarter*. `http://pari.math.u-bordeaux.fr/`.

[Pav95] R. Pavani. A numerical approximation of the rotation number. *Appl. Math. Comput.*, 73(2-3):191–201, 1995.

[PCU08] R. Paskauskas, C. Chandre, and T. Uzer. Dynamical bottlenecks to intramolecular energy flow. *Phys. Rev. Let.*, 100(8):083001, 2008.

[Pös82] J. Pöschel. Integrability of Hamiltonian systems on Cantor sets. *Comm. Pure Appl. Math.*, 35(5):653–696, 1982.

[Pös89] J. Pöschel. On elliptic lower-dimensional tori in Hamiltonian systems. *Math. Z.*, 202(4):559–608, 1989.

[Pös01] J. Pöschel. A lecture on the classical KAM theorem. In *Smooth ergodic theory and its applications (Seattle, WA, 1999)*, volume 69 of *Proc. Sympos. Pure Math.*, pages 707–732. Amer. Math. Soc., 2001.

[PTVF02] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery. *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, second edition edition, 2002.

[Pui] J. Puig. Reducibility of linear differential equations with quasi-periodic coefficients: a survey. Preprint available electronically at `http://www.ma1.upc.edu/~jpuig/preprints/qpred.ps`.

[RGJ05] P. Robutel, F. Gabern, and A. Jorba. The observed Trojans and the global dynamics around the Lagrangian points of the Sun-Jupiter system. *Celestial Mech. Dynam. Astronom.*, 92(1-3):53–69, 2005.

[Ris99] E. Risler. Linéarisation des perturbations holomorphes des rotations et applications. *Mém. Soc. Math. Fr. (N.S.)*, (77):viii+102, 1999.

[Rüs75] Helmut Rüssmann. On optimal estimates for the solutions of linear partial differential equations of first order with constant coefficients on the torus. In *Dynamical systems, theory and applications (Rencontres, Battelle Res. Inst., Seattle, Wash., 1974)*, pages 598–624.Lecture Notes in Phys., Vol. 38. Springer, Berlin, 1975.

[Rüs76] Helmut Rüssmann. On a new proof of Moser's twist mapping theorem. In *Proceedings of the Fifth Conference on Mathematical Methods in Celestial Mechanics (Oberwolfach, 1975), Part I*, number 1, pages 19–31, 1976.

[SA98] S. Shinohara and Y. Aizawa. Indicators of reconnection processes and transition to global chaos in nontwist maps. *Progr. Theoret. Phys.*, 100(2):219–233, 1998.

[Sev95] M. B. Sevryuk. KAM-stable Hamiltonians. *J. Dynam. Control Systems*, 1(3):351–366, 1995.

[Sev96] M. B. Sevryuk. Invariant tori of Hamiltonian systems that are nondegenerate in the sense of Rüssmann. *Dokl. Akad. Nauk*, 346(5):590–593, 1996.

[Sev99] M. B. Sevryuk. The lack-of-parameters problem in the KAM theory revisited. In *Hamiltonian systems with three or more degrees of freedom (S'Agaró, 1995)*, volume 533 of *NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci.*, pages 568–572. Kluwer Acad. Publ., 1999.

[Sev07] Mikhail B. Sevryuk. Invariant tori in quasi-periodic non-autonomous dynamical systems via Herman's method. *Discrete Contin. Dyn. Syst.*, 18(2-3):569–595, 2007.

[Sim98a] C. Simó. Effective computations in celestial mechanics and astrodynamics. In *Modern methods of analytical mechanics and their applications (Udine, 1997)*, pages 55–102. Springer, Vienna, 1998.

[Sim98b] C. Simó. Invariant curves of analytic perturbed nontwist area preserving maps. *Regul. Chaotic Dyn.*, 3(3):180–195, 1998. J. Moser at 70.

[SK89] Ya. G. Sinaĭ and K. M. Khanin. Smoothness of conjugacies of diffeomorphisms of the circle with rotations. *Uspekhi Mat. Nauk*, 44(1(265)):57–82, 247, 1989.

[SM71] C. L. Siegel and J. K. Moser. *Lectures on Celestial Mechanics*. Springer-Verlag, New York, 1971. Translation by C. I. Kalme, Die Grundlehren der mathematischen Wissenschaften, Band 187.

[SV06] T.M. Seara and J. Villanueva. On the numerical computation of Diophantine rotation numbers of analytic circle maps. *Phys. D*, 217(2):107–120, 2006.

[SV09] T.M. Seara and J. Villanueva. Numerical computation of the asymptotic size of the rotation domain for the Arnold family. *Phys. D*, 238(2):197–208, 2009.

[SZ89] D. Salamon and E. Zehnder. KAM theory in configuration space. *Comment. Math. Helv.*, 64(1):84–132, 1989.

[TLBB01] A.M. Tarquis, J.C. Losada, R.M. Benito, and F. Borondo. Multifractal analysis of tori destruction in a molecular Hamiltonian system. *Phys. Rev. E*, 65(1):016213, 2001.

[Tre89] D. V. Treshchëv. A mechanism for the destruction of resonance tori in Hamiltonian systems. *Mat. Sb.*, 180(10):1325–1346, 1439, 1989.

[Val00] Enrico Valdinoci. Families of whiskered tori for a-priori stable/unstable Hamiltonian systems and construction of unstable orbits. *Math. Phys. Electron. J.*, 6:Paper 2, 31 pp. (electronic), 2000.

[Van02] J. Vano. *A Nash-Moser implicit function theorem with Whitney regularity and applications*. PhD thesis, University of Texas at Austin, available electronically at `http://hdl.handle.net/2152/655`, 2002.

[Vil08] Jordi Villanueva. Kolmogorov theorem revisited. *J. Differential Equations*, 244(9):2251–2276, 2008.

[vV88] M. van Veldhuizen. On the numerical approximation of the rotation number. *J. Comput. Appl. Math.*, 21(2):203–212, 1988.

[WAFM05] A. Wurm, A. Apte, K. Fuchss, and P. J. Morrison. Meanders and reconnection-collision sequences in the standard nontwist map. *Chaos*, 15(2):023108, 13, 2005.

[Wei79] A. Weinstein. *Lectures on symplectic manifolds*, volume 29 of *CBMS Regional Conference Series in Mathematics*. American Mathematical Society, Providence, R.I., 1979.

[XY01] J. Xu and J. You. Persistence of lower-dimensional tori under the first Melnikov's non-resonance condition. *J. Math. Pures Appl. (9)*, 80(10):1045–1067, 2001.

[Yoc84a] J.C. Yoccoz. Conjugaison différentiable des difféomorphismes du cercle dont le nombre de rotation vérifie une condition diophantienne. *Ann. Sci. École Norm. Sup. (4)*, 17(3):333–359, 1984.

[Yoc84b] J.C. Yoccoz. Il n'y a pas de contre-exemple de Denjoy analytique. *C. R. Acad. Sci. Paris Sér. I Math.*, 298(7):141–144, 1984.

[Yoc95] J.C. Yoccoz. Théorème de Siegel, nombres de Bruno et polynômes quadratiques. *Astérisque*, (231):3–88, 1995.

[Zeh75] E. Zehnder. Generalized implicit function theorems with applications to some small divisor problems. I. *Comm. Pure Appl. Math.*, 28:91–140, 1975.

[Zeh76] E. Zehnder. Generalized implicit function theorems with applications to some small divisor problems. II. *Comm. Pure Appl. Math.*, 29(1):49–111, 1976.

[ZLL08] W. Zhu, B. Liu, and Z. Liu. The hyperbolic invariant tori of symplectic mappings. *Nonlinear Anal.*, 68(1):109–126, 2008.