UNIVERSITÀ DEGLI STUDI ROMA TRE
FACOLTÀ DI SCIENZE M.F.N.

Graduation Thesis in Mathematics
of

Fabio Giuseppe Pusateri

ANALYTIC KAM TORI for the PLANETARY $(n+1)$-BODY PROBLEM

Adviser
Prof. Luigi Chierchia

The Candidate | The Adviser

ACADEMIC YEAR 2005 - 2006
October 2006

Key words: Kolmogorov-Arnold-Moser theory, small denominators, Rüßmann's theory for analytic nearly integrable Hamiltonian systems, properly degenerate Hamiltonian systems, "Arnold's Theorem" on the stability of planetary motions.

Abstract

In [Féj04] Jacques Féjoz completed and gave the details of Michel Herman's proof of Arnold's 1963 theorem on the stability of planetary motions. This result provided the existence of C^∞ maximal invariant tori for the planetary $(n+1)$-body problem, with $n \geq 2$, in a neighborhood of Keplerian circular and coplanar movements, under the hypothesis that the masses of the planets are sufficiently small with respect to the mass of the "Sun". In this thesis we prove an analogous result in analytic class, i.e., we prove, under the same hypotheses listed above, the existence of real-analytic maximal invariant tori for the planetary $(n+1)$-body problem. The proof is based on the cited article by J. Féjoz and on a 2001 paper by H. Rüßmann. First we prove a general quantitative theorem about existence of maximal KAM tori for nearly-integrable Hamiltonian systems near elliptic lower dimensional tori. Then, using [Féj04], we obtain a set of initial data, in the phase space of the Hamiltonian model for a planetary system, with strictly positive Lebesgue measure, leading to quasi-periodic motions with $3n-1$ frequencies.

In appendix A and B we give a complete and detailed proof of Kolmogorov's original 1954 KAM theorem. In appendix C we briefly review Rüßmann's theory, contained in [Rüßm01], about lower dimensional elliptic invariant tori for nearly-integrable Hamiltonian systems.

Contents

0.1 Notations

1. Let $x \in \mathbb{R}^n$ or $\mathbb{C}^n$, $|x|$ always means $|x|_2 = \sqrt{|x_1|^2 + \cdots + |x_n|^2}$ if not otherwise specified.

2. Let $\Omega \subseteq \mathbb{C}^n$ and $\mathbb{T}^d$ the usual d-dimensional torus $\mathbb{R}^d/2\pi\mathbb{Z}^d$, we define for r and $\sigma > 0$ the following sets:

$$\Omega_r := \Omega + r := \bigcup_{x\in\Omega} \{y \in \mathbb{C}^n \ : \ |y - x| < r\} \,,$$

$$\mathbb{T}^d_\sigma := \{x \in \mathbb{C}^d \ : \ |\ \mathrm{Im}\ x_j| < \sigma,\ \mathrm{Re}\ x_j \in \mathbb{T}^d \ \ \forall\ j = 1 \ldots d\}\,;$$

   we will also use sometimes the notation $\mathbb{T}^d(\sigma)$ instead of $\mathbb{T}^d_\sigma$; if $\Omega \subseteq \mathbb{R}^n$ we denote

$$\Omega +_{\mathbb{R}} r := (\Omega + r) \cap \mathbb{R}^n \,.$$

3. Let $A \subseteq \mathbb{C}^d$ a domain and $f : A \longrightarrow \mathbb{C}^m$ a real-analytic function, we put

$$|f|_A := \sup_{x\in A} |f(x)|$$

   and if f can be holomorphically extended on a $A + r$ we denote

$$|f|_r := |f|_{A+r} = \sup_{x\in A+r} |f(x)|\,.$$

4. Let $B \subseteq \mathbb{K}^n$, with $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, a domain and $g : B \longrightarrow \mathbb{K}^d \times \mathbb{K}^{q\times q}$ a ν-times continuously differentiable function. We denote

$$(a_1, \ldots, a_\nu) \in (\mathbb{K}^n)^\nu \longmapsto D^\nu g(b)(a_1, \ldots, a_\nu) \in \mathbb{K}^d \times \mathbb{K}^{q\times q}$$

   as the ν-th derivative of g in $b \in B$ (with the convention $D^\nu g(b)(a) := g(b)$ for $\nu = 0$); moreover we write $|D^\nu g(b)(a^\nu)| = |D^\nu g(b)(a, \ldots, a)|$ and define

$$|D^\nu g(b)| := \max_{a\in\mathbb{K}^n\,,\,|a|_2=1} |D^\nu g(b)(a^\nu)|$$

   where $|\cdot|$ is some norm in $\mathbb{K}^d \times \mathbb{K}^{q\times q}$. For any $A \subseteq B$ we write

$$|D^\nu g(b)|_A := \sup_{b\in A} |D^\nu g(b)| \leq \infty \ .$$

   The Banach space of all μ-times continuously differentiable function $g : B \longrightarrow \mathbb{R}^d \times \mathbb{R}^{q\times q}$, (where B is an open subset of $\mathbb{R}^n$), with bounded derivatives up to order μ is defined by $C^\mu(B, \mathbb{R}^d \times \mathbb{R}^{q\times q})$ provided with norm

$$|g|^\mu_B := \max_{0\leq\nu\leq\mu} |D^\nu g|_B < \infty\,.$$

5. J_{2n}, for $n \in \mathbb{N}_+$, denotes the standard symplectic matrix

$$J_{2n} := \begin{pmatrix} 0 & -\mathbb{I}_n \\ \mathbb{I}_n & 0 \end{pmatrix} .$$

6. For any $p, j \in \mathbb{N}_+$ we define

$$\mathbb{Z}^p_j := \{n \in \mathbb{Z}^p \mid |n|_1 = j\} .$$

Other definitions or notations that are used only in some parts of the thesis will be introduced when necessary.

Chapter 1

Summary

" ...l'Amor che move il sole e l'altre stelle. "
Dante Alighieri. Divina Commedia, Paradiso, Canto XXXIII.

The main body of this thesis is divided into four chapters and three appendices.

In chapter 2 we review H. Rüßmann's theory about existence of analytic maximal KAM tori for nearly integrable Hamiltonian systems. All results provided are obtained from the paper [Rüßm01] where, actually, Rüßmann proves a much more general result about existence of lower dimensional invariant tori.

At the beginning of section 2.1 we introduce two definitions of weak non-degeneracy conditions; the term "weak" is used here to underline the difference with respect to the classical non-degeneracy conditions in the history of KAM theory (see for example A.N. Kolmogorov's 1954 theorem in appendix A). The first definition presented is the so called "Arnold-Pyartli condition", used by M. Herman to prove a very general and elegant theorem about existence of invariant tori for C^∞ nearly-integrable Hamiltonian systems. Immediately after the definitions of non-degeneracy, weak non-degeneracy and extreme non-degeneracy in the sense of Rüßmann are given. Some simple examples and remarks follow together with a brief discussion of the relations between the two different non-degeneracy conditions.

In section 2.2 we prepare for the statement of Rüßmann's theorem on maximal invariant tori; we introduce the concept of approximation function, used to control in a more general way (than the classical diophantine inequalities) the small denominators appearing in the problem and give the fundamental definition of index and amount of non-degeneracy of a real-analytic function. Soon after, we present a somewhat quantitative statement of Rüßmann's theorem obtained putting together different results contained in the article cited above.

In section 2.4 we provide an explicit estimate for the size of the perturbation in Rüßmann's theorem. We remark that this estimate has been deduced considering the case of maximal tori all along Rüßmann's paper and simplifying some numerical values with the imposition of more strict upper bounds.

As ending to this chapter, in section 2.5 we discuss with some more details Rüßmann's theorem about existence of non-resonant frequencies under the hypothesis of non-degeneration of the frequency application (i.e., the gradient of the integrable part of the considered Hamiltonian function). We have chosen to investigate this aspect of Rüßmann's work (see [Rüßm01, Part IV] for complete details) not only for its technical beauty, but to serve an important purpose. We aim to prove that it is not necessary to use the literal definition of index of non-degeneracy μ given by Rüßmann (see section 2.2 and definition 2.2.2); indeed, it will be shown that this quantity may be replaced by any integer $\bar{\mu} \geq \mu$, as well as the related amount of non-degeneracy may be substituted by the amount corresponding to $\bar{\mu}$ (see remarks in subsection 4.3.3). Actually, there is also another reason for the exposition of this matter: by the construction described the reader will understand how it is possible to control the non-resonant frequencies with the classical diophantine inequalities instead of using an approximation function (see subsection 4.3.5 for details). This aspect, apparently trivial, is fundamental to apply Rüßmann's theorem to properly degenerate Hamiltonian systems, i.e., nearly-integrable systems whose integrable part does not depend upon all the action variables, as for example the Hamiltonian model for the many-body problem.

In chapter 3 we prove the first part of a general theorem about analytic and properly degenerate Hamiltonian systems. We infer existence of a positive measure set of initial data, in the phase space of the considered Hamiltonian function, leading to quasi-periodic motions laying on analytic Lagrangian (maximal) KAM tori. This theorem is an analogous, performed in analytic class, of M. Herman's KAM theorem contained in [Her98] (a proof can be also found in [Féj04]). Herman's theorem is extremely elegant and very general since it establishes also the existence of lower dimensional tori for properly degenerate Hamiltonian systems (while we will only prove the existence of maximal KAM tori in the general theorem as well as in the application to planetary systems). A non-properly degenerate version of Herman's theorem is Rüßmann's theorem discussed in chapter 2 (case of maximal tori) and appendix C (general case of lower dimensional tori). Indeed, this last theorem is the one we will apply to prove our KAM theorem.

In section 3.1 we start introducing the general Hamiltonian setting. We consider a real-analytic nearly-integrable Hamiltonian in the form $H = h + \epsilon f$, where ϵ is a "small" positive parameter; in particular we require that the average of the perturbation with respect to the angles, we shall call f_0, possesses an elliptic equilibrium point, say, in the origin. Then we give general formulations of the results proved in chapter 3 and 4; theorem 3.1.1 states the main result and is followed by a brief description of the strategy of the proof made through theorems 3.1.2, 3.1.3 and 3.1.4. The first of this three theorems is proved in the whole remaining of chapter 3. The scheme adopted is a classical one, as well as the majority of the results used and applied, and reflects for most aspects Herman's scheme in his general KAM theorem (see [Féj04, theorem 60]). The proof runs as follows. Section 3.2 contains the exposition of a well known theory of dynamical systems that is averaging theory. Using the hypothesis of non-degeneracy in the sense of Rüßmann of $o := \nabla\omega$, we localize the initial action variables in an open ball where we can apply a corollary of a general result taken from [Val03] (theorem 3.2.1 in this thesis). Averaging theory permits the removal of the dependence upon the angle variables up to any chosen and fixed order greater than 1. Unfortunately, it causes at the same time a little shift of order $O(\epsilon)$ of the elliptic equilibrium initially possessed by f_0. Thus, in section 3.3 we find a symplectic transformation that restores the original equilibrium. Subsequently, we use a well known result by K.Weierstraß to diagonalize the quadratic part of the function obtained composing f_0 with the symplectic transformations made so far.

Next, in section 3.4, a general formulation of a classical theorem, usually called "Birkhoff's normal form theorem", is provided. Following the ideas in [Zeh94, pages 43-44], we give a detailed and quantitative proof about existence of a symplectic map that puts an Hamiltonian into Birkhoff's normal form up to a certain arbitrarily fixed order. The most important aspect to underline is that we give an explicit evaluation of the domains of the transformation.

In the first part of section 3.5 we perform the passage to symplectic polar coordinates that represent angle-action variables for the integrable part of the Hamiltonian considered. Subsequently we make a rescaling of the new angle-variables with a conformally symplectic map which casts the Hamiltonian in a simple form very similar to that considered in Rüßmann's theorem 2.3.1. More precisely, the Hamiltonian system obtained so far is a *degenerate* case of the system considered in Rüßmann's theorem, i.e., the Hamiltonian function obtained has the same form of Rüßmann's Hamiltonian, with the difference that in the first the dependence upon the small parameter ϵ appears also in its integrable part.

Chapter 4 is dedicated to the application of Rüßmann's theorem and constitutes the most original part of this thesis.

In section 4.1 we analyze the frequency application of the integrable part of the Hamiltonian $\hat{H}$ obtained through the different steps performed in chapter 3. This "new" frequency application is related to the initial one by the following two aspects: their difference is $O(\epsilon)$ and the last components of the new frequency application (that is the components given by the first Birkhoff's invariant) are multiplied by a factor proportional to ϵ. Thus, assuming sufficient condition on ϵ, we prove that the frequency application of $\hat{H}$ is non-degenerate in the sense of Rüßmann. Moreover, imposing further upper bounds, we are able to control the index and amount of non-degeneracy of this frequency application and relate them to the index and amount of the initial one. This key passage, necessary to apply Rüßmann's theorem and at the same time to give an explicit estimate for the size of the initial perturbation, is performed in section 4.2.

The following section 4.3 describes with full details the application of Rüßmann's theorem for maximal invariant tori. We consider and analyze every single quantity and constant involved in Rüßmann's estimate for the size of the perturbation; in particular we focus our attention on how they change order in ϵ when our degenerate case is taken into consideration. This allows us to determine, in subsection 4.3.11, a sufficient lower bound for the order to which we may remove the dependence upon the initial angles (with averaging theorem) and the order up to which we may apply Birkhoff's normal form theorem. Particular attention may be paid to subsection 4.3.5 where we explain the different choice, with respect to Rüßmann, for the control of the small denominators and show how our choice is effectively possible in Rüßmann's theorem. In subsection 4.3.11 we determine sufficient conditions on ϵ to apply Rüßmann's theorem and finally, in section 4.4, we gather all other conditions imposed on ϵ in chapter 3 and in sections 4.1 and 4.2.

The main part of this thesis ends with chapter 5 where we review the results contained in [Féj04, pages 45-62]. In the cited pages J. Féjoz completes M. Herman's work on the spatial planetary $(n+1)$-body problem proving the non-planarity of the planetary frequency application. Then, applying Herman's general KAM theorem about C^∞ invariant tori he finally obtains a proof of Arnold's 1963 theorem on planetary motions in the C^∞ case (theorem 63 in [Féj04] [1]). In particular J. Féjoz proves that the spatial planetary frequency map α, formed by $3n-1$ frequencies, does not satisfy any linear relation on an open and dense subset with full Lebesgue measure of the secular space (i.e., the "collisionless" space of the ellipses described by the planets). This is equivalent to prove the non-degeneracy of α in the sense of Rüßmann on the same open set and is the fundamental result we use to apply our KAM theorem for analytic invariant tori to the spatial planetary problem.

[1] Actually, in [Arn63b] V.I. Arnold announced a somewhat stronger result: "If the masses, eccentricities and inclinations of the planets are sufficiently small, then for the majority of initial conditions the true motion is conditionally periodic and differs little from Lagrangian motion with suitable initial conditions throughout an infinite interval of time $\infty < t < \infty$". However Arnold proved this statement only in the case of the planar three-body problem and gave indications on how to generalize this result but, apparently, nobody has ever succeeded in implementing Arnold's indications.

We can describe the structure of chapter 5 as follows. First we introduce the general setting of the Hamiltonian model for the spatial planetary problem; starting from Newton's equation we pass to heliocentric coordinates and regard the problem considered as a perturbation of n decoupled two-body problems. Further on in section 5.1, we perform the passage to Poincaré coordinates and describe their relations with the elliptic orbital elements.

In section 5.2 we define the secular Hamiltonian as the average, made with respect to the angle-variables, of the perturbative function expressed in Poincaré variables. Then we summarize how is it possible to put the secular Hamiltonian into the form considered in the general KAM theorem stated in section 3.1. In particular, results contained in [Poi07], [Rob95] and [Las91] (to which we always refer for all the details) are used to express, with the help of Laplace's coefficients, the quadratic part of the secular Hamiltonian in a very remarkable and simple form.

Section 5.3 starts with some preliminary lemmata and a proposition concerning the eigenvalues of a symmetric matrix which is a "perturbation" of a diagonal matrix. In subsection 5.3.1 we use the results just shown to check Arnold-Pyartli condition for the planetary $(n + 1)$-body problem in the plane. First of all the frequency map, composed by $2n$ frequencies, is complexified on a certain connected domain, using the analycity of the Hamiltonian function expressed in Poincaré variables and the holomorphic extension of Laplace's coefficients. Then, the development of the quadratic part of the secular Hamiltonian for small ratios of the semi major axes yields the non-planarity of the frequency map on an open and dense subset of the secular space with full Lebesgue measure.

In subsection 5.3.2 the non-planarity of the frequency application for the spatial problem is discussed. It is immediately observed that the non-degeneracy condition is not verified in this case. In fact, the spatial planetary frequency application satisfies two (and only two) linear relations: the first is due to the presence of a null frequency whereas the second is a strange resonance remarked for the first time in his generality by M. Herman (indeed it is usually called "Herman's resonance").

To suppress this two linear relations Féjoz follows an intermediate strategy between Herman and Arnold. If we denote H_{plt} the Hamiltonian function of the spatial planetary problem expressed in Poincaré variables, Féjoz' idea is to consider a modified Hamiltonian $H_\delta := H_{\text{plt}} - \delta C_z^2$ where δ is a real parameter and C_z denotes the third component of the total angular momentum. The frequency application of H_δ extends for all $\delta \neq 0$ the frequency application of H_{plt} and does not verify "Herman's resonance". Moreover, the restriction to the symplectic submanifold of vertical total angular momentum suppresses also the linear relation given by the null eigenvalue so that the non-planarity of the frequency application is satisfied for H_δ. Thus, theorem 3.1.1 gives an open set of initial data with full measure, leading to quasi periodic motions for the flow of H_δ laying on invariant analytic Lagrangian tori. Finally, since H_{plt} and H_δ commute, they possess the same Lagrangian tori so that the result is established also for H_{plt}.

The second part of this thesis is constituted by three appendices: in appendix A and B we review the classical KAM theory while in appendix C a summary of the more recent Rüßmann's theory on the construction of analytic lower dimensional invariant tori is presented.

In appendix A we discuss the result gave by Kolmogorov in his 1954 theorem on the persistence of quasi-periodic motions for analytic nearly-integrable Hamiltonian systems. In [Kol54] the Russian mathematician stated a general theorem and outlined a brief sketch of the proof; he showed how periodic motions for integrable systems still persist under the addition of a perturbation, provided that this last is small enough and assuming a non-degeneracy hypothesis on the quadratic part of the integrable part of the Hamiltonian function. Using his brilliant ideas we prove a quantitative version of his results (theorem A.1.2), providing also an explicit estimate for the size of the perturbation.

In subsections A.1.1 and A.1.2 some useful preliminary estimates are given. In section A.2 we describe Kolmogorov's main idea and perform the first step of the iteration process needed to put the Hamiltonian function considered into "Kolmogorov's normal form". Then, in section A.3 we discuss the iterative scheme and finally prove the convergence.

Appendix B contains a classical result concerning the measure of maximal invariant tori carrying quasi-periodic motions (usually called "Kolmogorov's tori") for nearly-integrable analytic Hamiltonian systems. A brief description of the general setting is made in the first two sections. In section B.3 we use Whitney's extension theorem (formulated and proved in [Whi34]) to extend the symplectic map gave by Kolmogorov's theorem (i.e., the map that puts the given Hamiltonian function into Kolmogorov's normal form) to a C^1 map. From the regularity of this extension, with the help of some classical diophantine estimates, we obtain the main result in theorem B.4.1: the measure of the complement of the union of maximal invariant KAM tori in the phase space of a nearly-integrable analytic Hamiltonian system is proportional to the size of the perturbation.

The thesis ends with appendix C where we describe Rüßmann's theorem for the existence of lower dimensional tori in nearly-integrable Hamiltonian systems. This remarkable result, contained in [Rüßm01], is obtained with the hypothesis of weak and extreme non-degeneration in the sense of Rüßmann of the frequency application. We remark once again how Rüßmann's theory for nearly-integrable Hamiltonian system and his concept of non-degeneracy are fundamental in our work (it is obvious!) but have also influenced Herman's work on the matter and vice versa [2].

The contents of appendix C can be briefly described as follows. Section C.1 contains a bit of preliminaries: we recall the definition of approximation function; we expose with some details the complete theory of index and amount of non-degeneracy for a weakly non-degenerate and extreme real-analytic function; we define the amount of degeneracy for such functions and give a useful estimate for the exponent of a matrix function. Then we state a quantitative version of Rüßmann's main theorem and in subsection C.2.1 we give an explicit estimate for the size of the perturbation. The two following subsections are dedicated to a brief description of Rüßmann's strategy and iterative scheme. Subsection C.2.4 contains the survey of Rüßmann's conditions needed to carry out the n-step of the iterative scheme for $n = 0$. It is motivated by our statement of Rüßmann's theorem which is slightly different from the original. In fact, in our formulation we add a term to the Hamiltonian function considered by Rüßmann; then, we verify that the initial conditions indicated by Rüßmann hold for this new Hamiltonian so that we are allowed to enter Rüßmann's scheme without any further modifications.

[2] In [Rüßm01] we read: "I thank M. Herman for his permanent interest in the realization of this work and I thank him also that he has emphasized, at every turn, that the discovery of invariant tori under the condition of weak non-degeneracy is my work, in particular in the case $p = q = 0$." (here $p = q = 0$ means the case of maximal tori).

Chapter 2

Rüßmann's theory on maximal KAM tori

2.1 Non-degeneracy conditions

In this section we give definitions of some different kinds of non-degeneracy conditions under which invariant tori for an unperturbed Hamiltonian system persist with the addition of a small perturbation. In particular we are interested in the weak non-degeneracy conditions defined by A. S. Pyartli (and widely used and developed by M.Herman and J.Féjoz in [Féj04]) and by H. Rüßmann.

Definition 2.1.1. *Let $\omega : I \subseteq \mathbb{R} \longrightarrow \mathbb{R}^m$ a parametrized curve of class $C^{m-1}(I)$, where I is a compact interval, ω is said to be* **non-planar** *at $t_0 \in I$ if*

$$\det\left[\omega(t_0), \omega'(t_0), \ldots, \omega^{(m-1)}(t_0)\right] \neq 0 \ ; \tag{2.1.1}$$

ω is said to be **non-planar** *(homogeneously non-planar for Herman) if it is non-planar at all points $t \in I$ (or equivalently if it is not contained in any vectorial hyperplane). ω is said to be* **planar** *at $t_0 \in I$ (in I) if it is not non-planar at $t_0 \in I$ (in I). Instead ω is called* **essentially non-planar** *(non-planar for Herman) if for every open subset $J \subset I$ the image $\omega(J)$ is not contained in an affine hyperplane.*

Let $B \subset \mathbb{R}^n$ be a closed ball and $\omega : B \longrightarrow \mathbb{R}^m$ a parametrized curve; ω is called non-planar at the point $t \in B$ if there exist an immersion $c : (I, s) \to (B, t)$, where I is an interval of $\mathbb{R}$, such that the curve $\omega \circ c$ is non-planar at s; it is non-planar if it is non-planar at all point $t \in B$ while it is essentially non-planar if its local image is nowhere contained in an affine hyperplane.

The so called "non-degeneracy condition of Arnold-Pyartli" consists in the non-planarity of the frequency vector of the unperturbed Hamiltonian system. The most important issues related to this non-degeneracy condition and the properties of such non-degenerate functions, are extensively discussed in [Pya69].

Definition 2.1.2 (Rüßmann non-degeneracy condition). *Let B a domain (a non-void open connected set) of $\mathbb{R}^n$:*

- *a real-analytic function $\omega = (\omega_1, \dots, \omega_m) : B \longrightarrow \mathbb{R}^m$ is called* **non-degenerate** *if for any $(c_1, \dots, c_m) \in \mathbb{R}^m \setminus \{0\}$*

  $$c_1\omega_1 + \cdots + c_m\omega_m \neq 0$$

  (if and only if the range $f(B)$ of f does not lie in any $(m-1)$-dimensional linear subspace of $\mathbb{R}^m$); we call ω **degenerate** *if it is not non-degenerate;*

- *a real-analytic function $(\omega, \Omega) : B \longrightarrow \mathbb{R}^m \times \mathbb{R}^p$ is* **weakly non-degenerate** *if ω is non-degenerate;*

- *defining the following subset of $\mathbb{Z}^p$*

  $$\begin{array}{rcl} \mathbb{Z}^p_\Omega & = & \{l \in \mathbb{Z}^p \,|\, \langle l, \Omega\rangle \neq 0\,, |l|_1 \leq 2\}\ , \\ \mathbb{Z}^p_{\Omega\omega} & = & \{l \in \mathbb{Z}^p \,|\, \langle k, \omega\rangle + \langle l, \Omega\rangle \neq 0,\ \ \forall\ k \in \mathbb{Z}^m, 0 < |l|_1 \leq 2\} \end{array}$$

  we call a weakly non-degenerate function (ω, Ω) **extreme** *if $\mathbb{Z}^p_\Omega = \mathbb{Z}^p_{\Omega\omega}$.*

We remark that in this chapter we will use exclusively the hypothesis of non-degeneration since we discuss Rüßmann's theory only in the case of maximal tori. On the other hand, in appendix C, the condition of weakly and extreme non-degeneracy will be considered since Rüßmann's complete results in the case of lower dimensional tori are exposed.

2.1.1 Remarks and examples

We now make some simple but important remarks:

1. We start remarking the equivalence of the non-planarity of a curve $\omega : I \longrightarrow \mathbb{R}^m$ and the fact that ω is not contained in any vectorial hyperplane (or any linear subspace of positive codimension in $\mathbb{R}^m$).
   In fact if ω is contained in a linear subspace $V \subsetneq \mathbb{R}^m$ there exists $a \neq 0$ in $\mathbb{R}^m$ such that $a \cdot \omega(t) = 0$ for every $t \in I$; by derivation

   $$a \cdot \omega(t) = 0 = a \cdot \omega'(t) = \cdots = a \cdot \omega^{(m-1)}(t)$$

   and therefore

   $$a \cdot \omega(t) + a \cdot \omega'(t) + \cdots + a \cdot \omega^{(m-1)}(t) = 0$$

for every $t \in I$; then, since $\omega, \omega', \dots, \omega^{(m-1)}$ verify a linear combination we have

$$\det\left[\omega(t), \omega'(t), \cdots, \omega^{(m-1)}(t)\right] = 0 \ . \tag{2.1.2}$$

for every $t \in I$ that is the definition of planarity.
On the other hand if ω is planar on I, since it verifies equation (2.1.2) there exists $b = (b_0, \dots, b_{m-1}) \in \mathbb{R}^m$, $b \neq 0$, such that

$$b_0\omega(t) + b_1\omega'(t) + \cdots + b_{m-1}\omega^{(m-1)}(t) = 0 \ .$$

for all $t \in I$. Let $j_0 := \max\{j = 0, \dots, m-1 \,|\, b_j \neq 0\}$, then, by the last equation, ω verifies on I the ordinary differential equation of order j_0

$$\omega^{(j_0)}(t) = -\frac{1}{b_{j_0}}\left(b_0\omega(t) + b_1\omega'(t) + \cdots + b_{j_0-1}\omega^{(j_0-1)}(t)\right) \ . \tag{2.1.3}$$

Choose now $v \in \mathbb{R}^m$ in the orthogonal space generated by the vectors $\{\omega^{(j)}(t_0)\}_{j=0,\dots,j_0-1}$ for some t_0 in $I \setminus \partial I$; then $v \cdot \omega(t)$ verifies equation (2.1.3) with the initial conditions

$$v \cdot \omega(t_0) = 0 = v \cdot \omega'(t_0) = \cdots = v \cdot \omega^{(j_0-1)}(t_0)$$

and therefore $v \cdot \omega(t) = 0$ in I.

2. It can be easily observed that the essential non-planarity implies the non-planarity since ω is essentially non-planar from definition 2.1.1 if for every open $J \subset I$

$$v \cdot \omega(t) + c \neq 0 \,, \quad \forall\ v \in \mathbb{R}^m, \ \forall\ c \in \mathbb{R}, \ \forall\ t \in J \tag{2.1.4}$$

therefore by taking $c = 0$, $J = I \setminus \partial I$ and using the continuity of ω we have

$$v \cdot \omega(t) \neq 0 \,, \quad \forall\ v \in \mathbb{R}^m, \ \forall\ t \in I \tag{2.1.5}$$

that is, by the previous remark, the non-planarity of ω

3. By remark 1 and definition 2.1.2 we obtain for a curve $\omega : I \subset \mathbb{R} \longrightarrow \mathbb{R}^m$ the equivalence

$$\omega \ \text{ is non-planar in } t \in I \Longleftrightarrow \omega \ \text{ is non-degenerate in } I. \tag{2.1.6}$$

This equivalence underlines very well the different aspect of the two definitions which, nevertheless, describe the same geometrical property of a function. In particular, the non-planarity describes a local behavior of the considered function while Rüßmann's non-degeneracy condition has a global nature. This is also clear from the two different settings in which this definitions are given.

The same equivalence in (2.1.6) holds if ω is defined on a closed ball $B \subset \mathbb{R}^n$.

4. As far as an analytic function ω is considered, the notion of essential non-planarity can be expressed by

$$v \cdot \omega(t) + c \neq 0\,, \quad \forall\ v \in \mathbb{R}^n,\ \ \forall\ c \in \mathbb{R},\ \ \forall\ t \in I\ . \tag{2.1.7}$$

To be as much clearer as possible we provide some elementary examples of functions belonging to the different categories of non-degeneration:

1. A non-planar (or non degenerate) but not essentially non-planar curve is $\omega(t) = (1, t) \in \mathbb{R}^2$; in fact $\det[\omega(t), \omega'(t)] = 1$ for every $t \in I$ while $\omega(t) \cdot (1, 0) - 1 = 0$. An example of a non-planar curve that is not essentially non-planar but whose (global) image $\omega(I)$ is not contained in any affine plane is
$$\omega : t \in [-1, 1] \to (t, 1)\chi_{[-1,0]} + (t, 1 + t^2)\chi_{(0,1]}.$$

2. A degenerate function that is weakly non-degenerate but not extreme is $(\omega, \Omega) : (0, 1) \longrightarrow \mathbb{R}^2 \times \mathbb{R}$ defined by $\omega(t) = (1, t)$ and $\Omega(t) = \frac{1}{2}t$; in fact $\mathbb{Z}^1_\Omega = \mathbb{Z}^1_1 \cup \mathbb{Z}^1_2$ while $\mathbb{Z}^1_{\omega\Omega} = \mathbb{Z}^1_1$ (since $(0, \pm 1) \cdot \omega(t) \mp 2\Omega(t) \equiv 0$).

3. A degenerate function that is weakly non-degenerate and extreme is given by $(\omega, \Omega) = \big((1, t), \frac{1}{3}t\big)$ since $\mathbb{Z}^1_{\omega\Omega} = \mathbb{Z}^1_\Omega = \mathbb{Z}^1_1 \cup \mathbb{Z}^1_2$;
an example of a degenerate function, weakly non-degenerate and extreme with $\mathbb{Z}^2_{\omega\Omega} = \mathbb{Z}^2_\Omega \subsetneq \mathbb{Z}^2_1 \cup \mathbb{Z}^2_2$ is $(\omega, \Omega) = \big((1, t), (\frac{1}{3}t, \frac{1}{3}t)\big)$.

Next we state some simple results that will help to better comprehend the strong relation between non-planarity and non-degeneration. As already observed, the main difference consists in the regularity of the functions considered in this two definitions. In fact, both may be regarded as geometrical properties with the only difference that local properties of analytic functions may become global properties (for instance the property of being not contained in any vectorial hyperplane) while this is not true for finitely many times differentiable functions.

Lemma 2.1.1. *Let I a non-void interval of $\mathbb{R}$ and $\nu : I \longrightarrow \mathbb{R}^m$ a parametrized curve of class $C^{p-1}(I)$. If the image of ν is contained in a vectorial hyperplane of $\mathbb{R}^m$ then function*

$$D(t) = \det\left[\nu(t), \nu'(t), \ldots, \nu^{(m-1)}(t)\right]$$

is constantly vanishing on I. On the other hand, if $D(t) = 0$ for every $t \in I$, there exists an open set $J \subset I$ such that $\nu(J)$ is contained in vectorial hyperplane.

Proof The proof of this result is quite immediate and can be found in [Féj04, page 33] □

Proposition 2.1.1. *Let $\omega : B \subset \mathbb{R}^n \longrightarrow \mathbb{R}^m$ (where B is a domain) a real-analytic function; if ω is non-planar at $y_0 \in B$ then ω is non-degenerate in B.*

Proof By definition 2.1.1 if ω is non-planar at I_0 in B there exists an immersion $\gamma : I \subset \mathbb{R} \longrightarrow$ (where I is a compact interval) such that $\gamma(t_0) = y_0$ and $\omega \circ \gamma(t)$ is non-planar t_0. Therefore, denoting $\nu(t) = \omega \circ \gamma(t)$, we have

$$\det\left[\nu(t_0), \nu'(t_0), \dots, \nu^{(m-1)}(t_0)\right] \neq 0\,;$$

by continuity there exists $0 < T_0 <$ dist$(t_0, \partial I)$ such that

$$\det\left[\nu(t), \nu'(t), \dots, \nu^{(m-1)}(t)\right] \neq 0$$

for every $t \in I_0 = (t_0 - T_0, t_0 + T_0)$. This implies, by the preceding lemma, that $\nu(I_0)$ is not contained in any vectorial hyperplane, i.e., for all $c \in \mathbb{R}^m$ it results that $c \cdot \nu(t)$ is not constantly zero for $t \in I_0$. By the definition of ν we have that for every $c \in \mathbb{R}^m$ the function $c \cdot \omega(y)$ is never constantly vanishing for $y \in \gamma(I_0)$; it follows that for every $c \in \mathbb{R}^m$ $c \cdot \omega$ does not constantly vanish on B which the statement □

Another simple observation that will turn out to be fundamental is the following:

Remark 2.1.1. *Let $\omega : y \in B \subset \mathbb{R}^m \longrightarrow \mathbb{R}^m$ a non-degenerate (non-planar) function then $\bar{\omega} : (y, \rho) \in B \times V \subset \mathbb{R}^m \times \mathbb{R}^p \longrightarrow \mathbb{R}^m$ defined by $\bar{\omega}(y, \rho) = \omega(y)$ for every $\rho \in V$, is non-degenerate (non-planar) on $B \times V$.*

2.2 Preliminaries to Rüßmann's theorem

Definition 2.2.1 (Approximation function). *A continuous function $\Phi\colon [0, \infty) \to \mathbb{R}$ is called an approximation function if:*

1. $1 = \Phi(0) \geq \Phi(s) \geq \Phi(t) > 0$ *for* $0 \leq s < t < \infty$*;*
2. $\Phi(1) = 1$ *so that* $\Phi(s) = 1$ *for any* $0 \leq s \leq 1$*;*
3. $s^\lambda \Phi(s) \overset{s\to\infty}{\longrightarrow} 0$ *for any* $\lambda > 0$ *so that* $\sup_{s\geq 1} s^\lambda \Phi(s)^{\frac{1}{\mu}} < \infty$ *for all* $\mu > 0$ *and* $\lambda \geq 0$*;*
4. $\int_1^\infty \log \frac{1}{\Phi(s)} \frac{ds}{s^2} < \infty$.

In his work [Rüßm01], Rüßmann uses an approximation function to perform in a most general way the control on the small divisors; in fact, since invariant tori have to be constructed requiring a priori certain arithmetical properties that the frequencies of quasi-periodic motions lying on this tori should have, it is necessary to establish in advance a certain control on the small divisors appearing in the construction. This is done in the proof of Kolmogorov's theorem (see appendix A) by requiring for the frequency vector ω the diophantine condition $|k\cdot\omega| \geq \gamma|n|^{-\tau}$. However it would be clearer in the statement of theorem 2.3.1 the role played by the approximation function. As a possible choice of an approximation function we indicate

$$\Phi(x) = \begin{cases} 1 & 0 \leq x \leq 1 \\ e^{-(x-1)^\alpha} & x \geq 1 \end{cases}$$

for a chosen $\alpha < 1$[1].

Lemma 2.2.1. *Let $f_j : B \longrightarrow \mathbb{R}^{m_j}$ be real-analytic and non-degenerate functions defined on a domain $B \subseteq \mathbb{R}^d$, for each $j = 1 \ldots N$. Consider*

$$\mathcal{C} = \mathbb{R}^{m_1} \times \cdots \times \mathbb{R}^{m_N}$$

and let $c = (c_1, \ldots, c_N) \in \mathcal{C}$ be some parameters. If we define $f : \mathcal{C} \times B \longrightarrow \mathbb{R}$ as the real-analytic function (with respect to the y variables)

$$f(c, y) = \prod_{j=1}^{N} \langle c_j, f_j(y) \rangle \tag{2.2.1}$$

and $\mathcal{S}$ as the following subset of $\mathcal{C}$

$$\mathcal{S} = \{ c = (c_1, \ldots, c_N) \in \mathcal{C} \mid |c_j|_2 = 1 \;\; \forall\; j = 1 \ldots N \},$$

then for any non-void compact set $\mathcal{K} \subset B$ there exist numbers $\mu_0 = \mu_0(f, \mathcal{K}) \in \mathbb{N}$ and $\beta = \beta(f, \mathcal{K}) > 0$ such that

$$\max_{0 \leq \mu \leq \mu_0} |D^\mu f(c, y)| \geq \beta \quad \forall\; c \in \mathcal{S}, \;\; \forall\; y \in \mathcal{K} \tag{2.2.2}$$

(here the derivatives are obviously taken with respect to the y variables).

Proof The proof of this lemma can be found in [Rüßm01, page 185] □

Observe that if we consider a function f as in (2.2.1), then the function

$$(c, y) \in \mathcal{C} \times B \longrightarrow \max_{0 \leq \nu \leq \mu} |D^\nu f(c, y)|$$

[1] We remark in advance that when we will try to apply Rüßmann's theorem for our purposes, we will not be able to use an approximation function to control the small denominators but we will be forced to use classical diophantine inequalities. Details can be found in section 4.3.5.

is continuous in $\mathcal{C} \times B$ for every $\mu \in \mathbb{N}$. Therefore the number

$$\beta(f,\mu,\mathcal{K}) := \min_{y\in\mathcal{K},|c|_2=1} \max_{0\le\nu\le\mu} |D^\nu f(c,y)|$$

is well defined for any compact set $\mathcal{K} \subset B$ and verifies $\beta(f,\mu_1,\mathcal{K}) \le \beta(f,\mu_2,\mathcal{K})$ for every $0 \le \mu_1 \le \mu_2$. Then by Lemma 2.2.1 we can well define the numbers $\mu_0(f,\mathcal{K})$ and $\beta(f,\mu_0,\mathcal{K})$ as follows

Definition 2.2.2. *We call index of non-degeneracy of f with respect to $\mathcal{K}$ the first integer μ_0 such that $\beta(f,\mu_0,\mathcal{K}) > 0$ (while $\beta(f,\mu,\mathcal{K}) = 0$ for every $\mu < \mu_0$); we call the number $\beta(f,\mu_0(f,\mathcal{K}),\mathcal{K})$ amount of non-degeneracy of f with respect to $\mathcal{K}$.*

Now we give the definition of the index and amount of non-degeneracy of the real-analytic function ω. In Rüßmann's theorem such function will be the "frequency map" of the unperturbed Hamiltonian system considered.

Definition 2.2.3. *Let $\mathcal{K} \subset \mathbb{R}^n$ be a compact set, $B \subseteq \mathbb{R}^n$ a domain containing $\mathcal{K}$ and $\omega : y \in B \longrightarrow \mathbb{R}^m$ a real-analytic and non-degenerate function. Let $\mathcal{S} := \{c \in \mathbb{R}^n : |c|_2 = 1\}$, we define $\mu_0(\omega,\mathcal{K}) \in \mathbb{N}_+$, the* **index of non-degeneracy** *of ω with respect to $\mathcal{K}$, as the first integer such that*

$$\beta := \min_{y\in\mathcal{K},\, c\in\mathcal{S}} \max_{0\le\nu\le\mu_0} \left|D^\nu |\langle c,\omega(y)\rangle|^2\right| > 0 \tag{2.2.3}$$

where $\beta = \beta(\omega,\mathcal{K})$ is called **amount of non-degeneracy** *of ω with respect to $\mathcal{K}$.*

We just observe that this definition is well posed since $|\langle c,\omega(y)\rangle|^2$ is a function in the form considered in (2.2.1) with $N = 2$, $m_1 = m_2 = m$ and $f_1 = f_2 = \omega$, with the only difference that the parameters c_1, c_2 are not independently varying in $\mathcal{S} \times \mathcal{S} \subseteq \mathbb{R}^n \times \mathbb{R}^n$ but have been chosen to coincide.

2.3 Statement of Rüßmann's Theorem on maximal KAM tori

As far as we consider only the case of maximal tori ($p = q = 0$ in Rüßmann's notation), the main theorem contained in [Rüßm01] can be formulated as follows:

Theorem 2.3.1 (Rüßmann's theorem for maximal tori). *Let $\mathcal{Y}$ be an open connected set of $\mathbb{R}^n$ and*

$$H(x,y) = h(y) + P(x,y) \tag{2.3.1}$$

a real-analytic Hamiltonian defined for

$$(x, y) \in \mathbb{T}^n \times \mathcal{Y}$$

endowed with the standard symplectic form $dy \wedge dx$. *Let* D *a complex domain on which* H *can be holomorphically extended; let* $\mathcal{K}$ *be any non-empty compact subset of* $\mathcal{Y}$ *with positive* n*-dimensional Lebesgue measure* $\text{meas}_n\mathcal{K} > 0$ *and let* $\mathcal{A} \subseteq D$ *be an open set such that*

$$\mathbb{T}^d \times \mathcal{K} \subseteq \mathcal{A} \; . \tag{2.3.2}$$

Choose $\vartheta \in (0, 1)$ *such that*

$$\mathbb{T}^d(\vartheta) \times (\mathcal{K} + 4\vartheta) \subseteq \mathcal{A} \tag{2.3.3}$$

and $(\mathcal{K} +_{\mathbb{R}} 2\vartheta) \subseteq \mathcal{Y}$. *Define* $\omega(y) = \frac{\partial}{\partial y} h(y)$,

$$C_1 = |\omega|_{\mathcal{K}+3\vartheta} \tag{2.3.4}$$

and assume that the function

$$\omega : \mathcal{Y} \subset \mathbb{R}^n \longrightarrow \mathbb{R}^n$$

is non-degenerate. Then for any $\epsilon^\star$ *with* $0 < \epsilon^\star < \text{meas}_n\mathcal{K}$ *there exist positive numbers* ϵ_0 *and* γ *(see section 2.4 for details) depending on* $\omega, \mathcal{K}, \epsilon^\star, \mathcal{A}, \Phi$*, such that assuming*

$$|P|_{\mathcal{A}} \leq \frac{1}{2}\epsilon_0 \tag{2.3.5}$$

and taking real numbers σ_0, r_0 *verifying*

$$\sigma_0 = \vartheta \tag{2.3.6}$$

$$r_0 = \left(\frac{\epsilon_0 \vartheta}{C_1}\right)^{\frac{1}{2}} \tag{2.3.7}$$

there exists a compact subset $\mathcal{H} \subseteq \mathcal{K}$ *with* $\text{meas}_n\mathcal{H} > \text{meas}_n\mathcal{K} - \epsilon^\star$ *and a bi-lipschitz mapping*

$$X : (b, \xi, \eta) \in \mathcal{H} \times \mathbb{T}^n \times \mathcal{U} \longrightarrow D \tag{2.3.8}$$

where $\mathcal{U}$ *is an open neighborhood of the origin in* $\mathbb{R}^n$ *such that*

- *the mapping*

  $$(\xi, \eta) \longmapsto (x, y) = X(b, \xi, \eta) \tag{2.3.9}$$

  defines, for every $b \in \mathcal{H}$*, an holomorphic canonical transformation on* $\mathbb{T}^n\left(\frac{\sigma_0}{5}\right) \times \mathcal{U}_\rho$ *and*

  $$X\left(\mathcal{H} \times \mathbb{T}^n\left(\frac{\sigma_0}{5}\right) \times \mathcal{U}_\rho\right) \subseteq \mathbb{T}^n_{\sigma_0} \times \mathcal{Y}_{r_0}$$

  for sufficiently small $\rho > 0$*;*

- *the transformed Hamiltonian is in the form:*

$$H\left(X(b,\xi,\eta)\right) = h^\star(b) + \langle \omega^\star(b), \eta\rangle + O(|\eta|^2) \tag{2.3.10}$$

for every $b \in \mathcal{H}$ *and* $(\xi,\eta) \in \mathbb{T}^n\left(\frac{\sigma_0}{5}\right) \times \mathcal{U}_\rho$*;*

- *the new frequency vector* $\omega^\star$ *satisfies for all* b *in* $\mathcal{H}$ *the diophantine inequality*

$$|\langle k, \omega^\star(b)\rangle| \geq \gamma\Phi(|k|_2),\;\; \forall\; k \in \mathbb{Z}^n \smallsetminus \{0\}\,.$$

Finally, we observe that the transformed Hamiltonian system possesses the solutions

$$\begin{aligned} \xi &= \omega^\star(b)t + c \\ \eta &= 0 \end{aligned} \tag{2.3.11}$$

so that the system described by H *in (2.3.1) possesses the invariant torus*

$$(x,y) = X(b,\xi,0)$$

for ξ *in* $\mathbb{T}^d$*, with quasi-periodic flow (2.3.11) for all* b *in* $\mathcal{H}$ *.*

2.4 Estimate for the size of the perturbation in Rüßmann's theorem

We now display an admissible value of ϵ_0 (the size of the perturbation) in 2.3.1 remarking that the result we give here is obtained by Rüßmann's estimate in [Rüßm01, page 171] considering the case of maximal tori ($p = 0$) all along the procedure followed to reach such an estimate; moreover the presence of some numerical quantities (just absolute numbers) has been simplified by imposing more strict upper bounds.

To be as much clearer as possible we recall briefly the role played by the different quantities involved in the estimate, always referring to C.2.1 for definitions and further explanations:

- $\beta = \beta(\omega, \mathcal{K})$ is the *amount of non-degeneracy* of the real-analytic and non-degenerate "frequency map" ω, while $\mu_0 = \mu_0(\omega,\mathcal{K})$ is its *index of non-degeneracy*; observe that for a non-degenerate function the amount of degeneracy $\alpha(\omega,\mathcal{K})$ equals 1.

- $\epsilon^\star$, ϑ and C_1 are chosen as in the statement and d_0 is defined as the diameter of $\mathcal{K}$, i.e., $d_0 = \sup_{x,y\in\mathcal{K}} |x-y|$;
- let $\Phi : [0,\infty] \longrightarrow \mathbb{R}$ the chosen approximation function, in according with definition 2.2.1, we choose and fix $T_0 \geq 1$ such that $\Phi(T_0) \leq e^{-(n+1)}$ and the following inequality holds

$$\int_{T_0}^{\infty} \log \frac{1}{\Phi(T)} \frac{dT}{T^2} \leq \frac{\vartheta \log 2}{12\mu_0 n \log(3^{25} n)}\,; \tag{2.4.1}$$

  moreover we define

$$\Phi_{n\mu_0} = \sup_{s\geq 1} s^n \Phi(s)^{\frac{1}{\mu_0}} < \infty$$

  (that is bounded by hypothesis on Φ);
- We define

$$M^\star \;=\; 2^{\mu_0+2}\mu_0!\vartheta^{-\mu_0}(C_1+1) \tag{2.4.2}$$

$$C^\star \;=\; 2^{\mu_0+1}\frac{(\mu_0+1)^{\mu_0+2}}{\vartheta^{\mu_0+1}}\Phi^2_{n\mu_0}(C_1+1)\,; \tag{2.4.3}$$

  and

$$\gamma_1 = \left[3^{n+3}(2\pi e)^{\frac{n}{2}} d_0^{n-1}\left(n^{-\frac{1}{2}} + 2d_0 + \vartheta^{-1}d_0\right) C^\star\right]^{-\frac{\mu_0}{2}} \beta^{\frac{\mu_0+1}{2}} \epsilon^{\star\,\frac{\mu_0}{2}} \tag{2.4.4}$$

  and set

$$\gamma \;=\; \min\{2,\gamma_1\} \tag{2.4.5}$$

$$t_0 \;=\; \frac{\gamma\Phi(T_0)T_0^{-(n+1)}\vartheta}{8(C_1+1)} \tag{2.4.6}$$

- Let $\bar{g} : \mathbb{R}^d \longrightarrow \mathbb{R}$ defined by

$$\bar{g}(x) = \begin{cases} e^{-\frac{1}{1-|x|^2}} & |x| < 1 \\ 0 & |x| \geq 1 \end{cases}$$

  we consider

$$g(s) = \bar{g}(s)\left(\int_{\mathbb{R}^d} \bar{g}(t)\,dt\right)^{-1},$$

and then define for $N \in \mathbb{N}$

$$C(n, N) = 3^N N! \sup_{a \in \mathbb{R}^n \,,\, |a|_2 = 1} \int_{\mathbb{R}^d} \sum_{j=0}^{N} \frac{1}{j!} |D^j g(s)(a^j)| \, ds \tag{2.4.7}$$

with $C(n, 0) = 1$;

then a possible value for ϵ_0 is given by

$$\epsilon_0 := \frac{\vartheta}{C_1} (\min\{E_1, E_2, E_3\})^2 \tag{2.4.8}$$

where

$$\begin{aligned} E_1 &:= \frac{\gamma \Phi(T_0) T_0^{-(n+1)}}{3^{28+n}} \min\left\{\vartheta \,,\, \frac{1}{3^{17}}\right\} \\ E_2 &:= \frac{\gamma \Phi(T_0) T_0^{-(n+1)}}{3^{22}} \left(1 - \frac{1}{2^{\mu_0}}\right) \left(1 - \frac{1}{2^{5 - \frac{2}{\mu_0}}}\right) \\ E_3 &:= \frac{\beta {t_0}^{\mu_0}}{3^{19} 2 T_0 M^\star C(n, \mu_0)} \,. \end{aligned} \tag{2.4.9}$$

2.5 Existence of non-resonant frequencies

We now follow Rüßmann's work to show how maximal invariant tori for the nearly-integrable Hamiltonian system considered in 2.3.1 can be constructed. We always refer to [Rüßm01, pages 178-203] for further details and for the proofs of results we just cite here as well as for the complete result concerning the construction of lower dimensional invariant tori.

2.5.1 Description of the iteration process

The initial condition of the iteration process are:

1. $\omega : \mathcal{Y} \longrightarrow \mathbb{R}^n$ is a real-analytic function defined on a domain (an open and connected set) $\mathcal{Y} \subseteq \mathbb{R}^n$ that is non-degenerate in the sense of Rüßmann (see definition 2.1.2);

2. $\mathcal{K} \subset \mathcal{Y}$ is a chosen and fixed compact set with $\text{meas}_n \mathcal{K} > 0$;

3. $\vartheta \in (0,1)$ is chosen small enough to verify

$$\mathcal{K} +_{\mathbb{R}} 2\vartheta \subseteq \mathcal{Y}$$

and $\omega_\bullet \in \mathcal{H}(\mathcal{K}+3\vartheta)$ is the holomorphic extension of $\omega|_{\mathcal{K}+_{\mathbb{R}}\vartheta}$; it follows that $\omega_\bullet$ is bounded on $\mathcal{K}+2\vartheta$ so that we can well define

$$C_1 = |\omega_\bullet|_{\mathcal{K}+2\vartheta} > 0\,. \tag{2.5.1}$$

We will often use the same notation ω to identify both the real function and its holomorphic extension when there will be no ambiguity.

4. $\mathcal{K}_0 := \mathcal{K}$, $\mathcal{B}_0 := \mathcal{K}_0 + t_0$ with $t_0 \in (0,\vartheta)$ to be determined and $\omega_0 := \omega_\bullet|_{\mathcal{B}_0}$.

At the beginning of the iteration process we take $\delta := \left(\frac{12}{25}\right)^{25} < \frac{1}{4}$ and

$$\begin{aligned} &0 < L_0 \le \min\{L_1^\star, \gamma L_2^\star\}\ , \quad 0 < \tau < \tau^\star \le \tau_0 := \tfrac{1}{9} \\ &t_\nu := t_0\delta^{\tau\nu}\,, \qquad\qquad L_\nu = L_0\delta^{\tau\nu} \qquad \text{for } \nu = 0,1,\dots \end{aligned} \tag{2.5.2}$$

where positive numbers $L_1^\star, L_\star^2, \gamma^\star, \tau^\star$ will be later determined. Moreover Φ is a chosen approximation function (see definition 2.2.1), and $\Psi(T) = T^{-(n+1)}\Phi(T)$ for any $T \ge 1$; $T_0 \ge 1$ is a fixed real number and

$$T_\nu := \Psi^{-1}\left(\Psi(T_0)\delta^{\tau\nu}\right) \tag{2.5.3}$$

for $\nu \ge 0$ (observe that by the properties of Φ we obtain $T_0 < T_{\nu_1} < T_{\nu_2}$ for any $0 < \nu_1 < \nu_2$).

With the framework described the general step of the iteration process works out recursively as we shall see now; if $\mathcal{K}_\nu \ne \emptyset$, $\mathcal{B}_\nu = \mathcal{K}_\nu + t_\nu$ and $\omega_\nu \in \mathcal{H}(\mathcal{B}_\nu)$ are given for some $\nu \ge 0$ (where $\mathcal{K}_0$, $\mathcal{B}_0$ and ω_0 have been set before) then we determine $\mathcal{K}_{\nu+1}$, $\mathcal{B}_{\nu+1}$ and $\omega_{\nu+1} \in \mathcal{H}(B_{\nu+1})$ as follows: we choose an arbitrary function $\Delta\omega_\nu \in \mathcal{H}(B_\nu)$ such that

$$|\Delta\omega_\nu|_{\mathcal{B}_\nu} \le \frac{L_\nu}{T_0} \tag{2.5.4}$$

and define

$$\mathcal{K}_\nu^\star := \left\{ b \in \mathcal{K}_\nu \mid |\langle k, \omega_\nu(b)\rangle| \ge \frac{1}{2}\gamma\Phi(T_\nu) \quad \forall\ 0 < |k|_2 \le T_\nu \right\}\ ; \tag{2.5.5}$$

then we take $\mathcal{K}_{\nu+1} = \mathcal{K}_\nu^\star$ and if $\mathcal{K}_\nu^\star \ne \emptyset$ we set

$$\mathcal{B}_{\nu+1} := \mathcal{K}_{\nu+1} + t_{\nu+1}\ , \quad \omega_{\nu+1} = \omega_\nu + \Delta\omega_\nu|_{\mathcal{B}_{\nu+1}}\ . \tag{2.5.6}$$

2.5.2 Theorem about existence of non-resonant frequencies

The following theorem ([Rüßm01, theorem 16.7]) states the possibility to carry out the general step of the iteration process infinitely many times, assuring the existence of non-resonant frequency vectors at each step as well as the convergence of the scheme.

Theorem 2.5.1 (Existence of non-resonant frequency vectors). *For any*

$$\epsilon^\star \in (0, \text{meas}_n \mathcal{K}) \tag{2.5.7}$$

there exist positive real numbers $L_1^\star, L_2^\star, \gamma^\star, \tau^\star$ such that choosing

$$\begin{aligned} &0 < \tau < \tau^\star\,, \quad\ \ 0 < \gamma < \gamma^\star \\ &0 < L_0 \le L_1^\star\,, \quad L_0 \le \gamma L_2^\star \end{aligned} \tag{2.5.8}$$

the iteration process described above can be carried out for any $\nu \in \mathbb{N}$ with the resulting sets $\mathcal{K}_\nu$ and the functions ω_ν possessing the following properties:

1. *$\mathcal{K} = \mathcal{K}_0 \supseteq \mathcal{K}_1 \supseteq \cdots \supseteq \mathcal{K}_\infty$ with*

$$\mathcal{K}_\infty := \bigcap_{\nu=1}^{\infty} \mathcal{K}_\nu \ne \emptyset\,; \tag{2.5.9}$$

2. *the sequence $\{\omega_\nu\}_{\nu\in\mathbb{N}}$ converges uniformly on $\mathcal{K}_\infty$ to a function ω_∞ (that is therefore continuous on $\mathcal{K}_\infty$);*

3. *let*

$$\mathcal{H} = \{b \in \mathcal{K}_\infty \mid |\langle k, \omega_\infty(b)\rangle| \ge \gamma\Phi(|k|_2) \quad \forall\ k \in \mathbb{Z}^n \setminus \{0\}\} \tag{2.5.10}$$

   then

$$\text{meas}_n\mathcal{K}_\infty \ge\ \text{meas}_n\mathcal{H} \ge\ \text{meas}_n\mathcal{K} - \epsilon^\star\,. \tag{2.5.11}$$

Furthermore let $\mu_0 \in \mathbb{N}_+$ and $\beta > 0$ be such that

$$\min_{y\in\mathcal{K}}\ \max_{0\le\nu\le\mu_0} \left|D^\nu |\langle c, \omega\rangle|^2\right| \ge \beta \tag{2.5.12}$$

for any $c \in \mathbb{R}^n$ with $|c|_2 = 1$, C_1 as defined in (2.5.1), $d_0 = \text{diam}\,\mathcal{K}$ and $C(n, N)$

as defined in (2.4.7); then, we can take

$$\tau^\star = \frac{\tau_0}{\mu_0} \tag{2.5.13}$$

$$\gamma^\star = \left[3^{n+2}(2\pi e)^{\frac{n}{2}} d_0^{n-1}\left(n^{-\frac{1}{2}} + 2d_0 + \vartheta^{-1}d_0\right) C^\star\right]^{-\mu_0} \beta^{\frac{\mu_0}{2}} \epsilon^{\star\frac{\mu_0}{2}} \tag{2.5.14}$$

$$L_1^\star = \frac{\beta t_0^{\mu_0}}{2M^\star C(n,\mu_0)}\left(1 - \delta^{\tau_0 - \mu_0 \tau}\right) \tag{2.5.15}$$

$$L_2^\star = \frac{1}{2(C_1+1)}\Phi(T_0)(1-\delta^{\tau_0}) \tag{2.5.16}$$

with

$$M^\star = 2^{\mu_0+2}\mu_0!\vartheta^{-\mu_0}(C_1+1) \tag{2.5.17}$$

$$C^\star = 2^{\mu_0+1}\frac{(\mu_0+1)^{\mu_0+2}}{\vartheta^{\mu_0+1}}\Phi^2_{n\mu_0}\left(\frac{(\mu_0+1)!}{\vartheta^{\mu_0}}(C_1+1)\right)^2 \tag{2.5.18}$$

accordingly to section 2.4.

We dedicate the remaining part of this section to the explanation of the proof that Rüßmann gives to this theorem. We will focus our attention on some parts and underline the fact that we are searching for the existence of maximal tori performing some proofs of intermediate results in the case $p = 0$ (with refer to Rüßmann's notation).

2.5.3 Theorems on the measure of a set defined by small divisors

We start citing an important result given and proved by Rüßmann in [Rüßm01, pages 180-183]:

Theorem 2.5.2. *Let $\mathcal{K} \subseteq \mathbb{R}^n$ be a non-empty compact set with diameter $d_0 = \sup_{x,y\in\mathcal{K}}|x-y| > 0$ and define $\mathcal{B} = \mathcal{K} +_{\mathbb{R}} \vartheta$ for some $\vartheta > 0$. Let $g \in C^{\mu_0+1}(\mathcal{B},\mathbb{R})$ (for some $\mu_0 \in \mathbb{N}$) be a function such that*

$$\min_{y\in\mathcal{K}}\max_{0\le\nu\le\mu_0}|D^\nu g(y)| \ge \beta \tag{2.5.19}$$

with $\beta > 0$. Then for any function $\tilde{g} \in \mathbb{C}^{\mu_0}(\mathcal{B},\mathbb{R})$ satisfying $|\tilde{g} - g|_{\mathcal{B}}^{\mu_0} \le \frac{1}{2}\beta$ and any

$$0 < \varepsilon \le \frac{\beta}{2\mu_0+2} \tag{2.5.20}$$

we can estimate

$$\textit{meas}_n\,\{y \in \mathcal{K} \,:\, |\tilde{g}(y)| \le \varepsilon\} \le B d_0^{n-1}\left(n^{-\frac{1}{2}} + 2d_0 + \vartheta^{-1}d_0\right)\left(\frac{\varepsilon}{\beta}\right)^{\frac{1}{\mu_0}}\frac{1}{\beta}|g|_{\mathcal{B}}^{\mu_0+1} \tag{2.5.21}$$

with

$$B = 3(2\pi e)^{\frac{n}{2}}\frac{(\mu_0+1)^{\mu_0+2}}{(\mu_0+1)!}\,. \tag{2.5.22}$$

A proof of this theorem, in the case $\tilde{g} = g$, can be found in [Pya69] where it appeared for the first time.

Here we state a simple lemma that we are going to use in next theorem's proof:

Lemma 2.5.1. *Let $F, G \in C^{\mu}(\mathcal{B}, \mathbb{R}^{m\times m})$ for some open set $\mathcal{B} \subseteq \mathbb{R}^n$ and $\mu \in \mathbb{N}$, then it results*

$$|FG|_{\mathcal{B}}^{\mu} \le 2^{\mu}|F|_{\mathcal{B}}^{\mu}|G|_{\mathcal{B}}^{\mu}$$

where $|\cdot|_{\mathcal{B}}^{\mu}$ is defined in 0.1.4.

Proof The proof is immediate and can be easily obtained by Leibniz rule

$$D^{\nu}(FG)(b)(a^{\nu}) = \sum_{i+j=\nu}\frac{\nu!}{i!j!}D^iF(b)(a^i)D^jG(b)(a^j)$$

for any $a \in \mathbb{R}^n$ with $|a|_2 = 1$ and $b \in \mathcal{B}$. Therefore

$$|D^{\nu}(FG)(b)(a^{\nu})| \le \sum_{i+j=\nu}\frac{\nu!}{i!j!}\left|D^iF(b)(a^i)\right|\,\left|D^jG(b)(a^j)\right| \le 2^{\mu}|F|_{\mathcal{B}}^{\mu}|G|_{\mathcal{B}}^{\mu}$$

which gives the statement if we take the $\sup$ over a,b and ν.

Let $\omega \,:\, \mathcal{Y} \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^n$ a real-analytic and non-degenerate function, let $\mathcal{K} \subseteq \mathcal{Y}$ be a compact set with diameter $d_0 = \sup_{x,y\in\mathcal{K}}|x-y| > 0$ and take $\vartheta \in (0,1)$ such that $\tilde{\mathcal{B}} := \mathcal{K} +_{\mathbb{R}} \vartheta \subseteq \mathcal{Y}$. We recall that by lemma C.1.2 there exist numbers $\mu_0 \in \mathbb{N}$ and $\beta > 0$ such that

$$\min_{y\in\mathcal{K}}\max_{0\le\nu\le\mu_0}\left|D^{\nu}|\langle c, \omega(y)\rangle|^2\right| \ge \beta \tag{2.5.23}$$

for every $c \in \mathbb{R}^n$. Then we can state the following result:

Theorem 2.5.3 (Measure of a set defined by small divisors). *Let ω as considered above and $\tilde{\omega} \in C^{\mu_0}(\tilde{B}, \mathbb{R}^n)$, satisfying the estimates*

$$|\omega|_{\tilde{\mathcal{B}}}^{\mu_0}\,,\ |\tilde{\omega}|_{\tilde{\mathcal{B}}}^{\mu_0} \le M_0\,,\ |\omega|_{\tilde{\mathcal{B}}}^{\mu_0+1} \le M_1 \tag{2.5.24}$$

$$|\omega - \tilde{\omega}|_{\tilde{\mathcal{B}}}^{\mu_0} \le \frac{\beta}{2^{\mu_0+2}M_0} \tag{2.5.25}$$

for some $M_0, M_1 \geq 1$ *and let* Φ *the chosen approximation function (see definition 2.2.1). Moreover let* $0 < \epsilon^\star <$ *meas* ${}_n\mathcal{K}$ *and*

$$0 < \gamma \leq \gamma' := \left[3^{n+2}(2\pi e)^{\frac{n}{2}} d_0^{n-1}\left(n^{-\frac{1}{2}} + 2d_0 + \vartheta^{-1}d_0\right) C'\right]^{-\frac{\mu_0}{2}} \beta^{\frac{\mu_0+1}{2}} \epsilon^{\star\frac{\mu_0}{2}} \tag{2.5.26}$$

with

$$C' = 2^{\mu_0+1}\frac{(\mu_0+1)^{\mu_0+2}}{(\mu_0+1)!}(\Phi_{n\mu_0})^2 M_1^2\,; \tag{2.5.27}$$

then the measure of the set

$$\mathcal{H}(\tilde\omega) := \{b \in \mathcal{K}\,:\, |\langle k, \tilde\omega(b)\rangle| \geq \gamma\Phi(|k|_2)\,,\ \ \forall\ k \in \mathbb{Z}^n \setminus \{0\}\} \tag{2.5.28}$$

can be estimated by

$$\textit{meas}\,{}_n\mathcal{H}(\tilde\omega) \geq \textit{meas}\,{}_n\mathcal{K} - \epsilon^\star\,. \tag{2.5.29}$$

Furthermore if $|\omega|^{\mu_0}_{\tilde{\mathcal{B}}} \leq M'$ *we can take* $M_0 = 2M'$*.*

We remark that a complete proof of this result, even in the case $p \neq 0$ and $\omega : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ with $m \neq n$, can be found in [Rüßm01, pages 189-193].

Proof Let $F_k(y) := i\langle k, \omega(y)\rangle$ and $\tilde F_k(y) := i\langle k, \tilde\omega(y)\rangle$ by hypotheses we have $F_k, \tilde F_k \in C^{\mu_0}(\tilde B, \mathbb{R})$. Setting

$$\begin{aligned}
[\omega]_k(y) &:= |\langle k, \omega(y)\rangle|^2 |k|_2^{-2} = F_k^2(y)|k|_2^{-2} \\
[\tilde\omega]_k(y) &:= |\langle k, \tilde\omega(y)\rangle|^2 |k|_2^{-2} = \tilde F_k^2(y)|k|_2^{-2}
\end{aligned}$$

(accordingly to (C.1.7) in which we consider also the case of frequencies $\chi = (\omega, \Omega)$ of lower dimensional tori) it results

$$\begin{aligned}
|[\omega]_k - [\tilde\omega]_k|^{\mu_0}_{\tilde{\mathcal{B}}} &= \left|\frac{F_k}{|k|_2}\frac{F_{-k}}{|k|_2} - \frac{\tilde F_k}{|k|_2}\frac{\tilde F_{-k}}{|k|_2}\right|^{\mu_0}_{\tilde{\mathcal{B}}} \\
|[\omega]_k|^{\mu_0}_{\tilde{\mathcal{B}}} &= \left|\frac{F_k}{|k|_2}\frac{F_{-k}}{|k|_2}\right|^{\mu_0}_{\tilde{\mathcal{B}}}\,.
\end{aligned}$$

Now for every $0 \leq \nu \leq \mu_0$, $b \in \tilde{\mathcal{B}}$ and $a \in \mathbb{R}^n$ with $|a|_2 = 1$ we have

$$\left|\frac{D^\nu F_k(b)(a^\nu)}{|k|_2}\right| = \frac{|\langle k, D^\nu\omega(b)(a^\nu)\rangle|}{|k|_2} \leq |D^\nu\omega(b)(a^\nu)| \leq |\omega|^{\mu_0}_{\tilde{\mathcal{B}}}$$

having used the definition of $|\cdot|^{\mu_0}_{\tilde{\mathcal{B}}}$ in 0.1.4. Since analogous estimate can be done for $\tilde F_k$ and $F_k - \tilde F_k$ instead of F_k, taking the $\sup$ over a, b and ν we obtain

$$|k|_2^{-1}\,|F_k|^{\mu_0}_{\tilde{\mathcal{B}}} \leq |\omega|^{\mu_0}_{\tilde{\mathcal{B}}} \qquad |k|_2^{-1}\left|\tilde F_k\right|^{\mu_0}_{\tilde{\mathcal{B}}} \leq |\tilde\omega|^{\mu_0}_{\tilde{\mathcal{B}}} \qquad |k|_2^{-1}\left|F_k - \tilde F_k\right|^{\mu_0}_{\tilde{\mathcal{B}}} \leq |\omega - \tilde\omega|^{\mu_0}_{\tilde{\mathcal{B}}}\,;$$

observe that for what concerns the estimate regarding ω only, we can replace μ_0 with any integer (μ_0+1 for instance) since ω is real-analytic. We now write the identity

$$\frac{F_k}{|k|_2}\frac{F_{-k}}{|k|_2}-\frac{\tilde F_k}{|k|_2}\frac{\tilde F_{-k}}{|k|_2}=\left(\frac{F_k}{|k|_2}-\frac{\tilde F_k}{|k|_2}\right)\frac{F_{-k}}{|k|_2}+\frac{\tilde F_k}{|k|_2}\left(\frac{F_{-k}}{|k|_2}-\frac{\tilde F_{-k}}{|k|_2}\right)$$

obtaining by the preceding estimates, the definition of $[\omega]_k$ and $[\tilde\omega]_k$ and lemma 2.5.1 for the C^{μ_0}-norm of a product

$$\begin{aligned}
|[\omega]_k-[\tilde\omega]_k|^{\mu_0}_{\tilde{\mathcal B}} &\le 2^{\mu_0}\frac{\left|F_k-\tilde F_k\right|^{\mu_0}_{\tilde{\mathcal B}}}{|k|_2}\frac{|F_{-k}|^{\mu_0}_{\tilde{\mathcal B}}}{|k|_2}+2^{\mu_0}\frac{\left|F_{-k}-\tilde F_{-k}\right|^{\mu_0}_{\tilde{\mathcal B}}}{|k|_2}\frac{\left|\tilde F_k\right|^{\mu_0}_{\tilde{\mathcal B}}}{|k|_2}\\
&\le 2^{\mu_0}|\omega-\tilde\omega|^{\mu_0}_{\tilde{\mathcal B}}|\omega|^{\mu_0}_{\tilde{\mathcal B}}+2^{\mu_0}|\omega-\tilde\omega|^{\mu_0}_{\tilde{\mathcal B}}|\tilde\omega|^{\mu_0}_{\tilde{\mathcal B}}=2^{\mu_0}\left(|\omega|^{\mu_0}_{\tilde{\mathcal B}}+|\tilde\omega|^{\mu_0}_{\tilde{\mathcal B}}\right)|\omega-\tilde\omega|^{\mu_0}_{\tilde{\mathcal B}}\\
&\le 2^{\mu_0+1}M_0|\omega-\tilde\omega|^{\mu_0}_{\tilde{\mathcal B}}\le\frac{1}{2}\beta\,.
\end{aligned}$$

Furthermore if we consider the two cases $\omega=0$ and $\tilde\omega=0$, which imply respectively $F_k=0$ and $\tilde F_k=0$, we have

$$\begin{aligned}
|[\omega]_k|^{\mu_0}_{\tilde{\mathcal B}} &= \left|\frac{F_k}{|k|_2}\frac{F_{-k}}{|k|_2}\right|^{\mu_0}_{\tilde{\mathcal B}}\le 2^{\mu_0}\left(|\omega|^{\mu_0}_{\tilde{\mathcal B}}\right)^2\le 2^{\mu_0}M_0^2\\
|[\tilde\omega]_k|^{\mu_0}_{\tilde{\mathcal B}} &= \left|\frac{\tilde F_k}{|k|_2}\frac{\tilde F_{-k}}{|k|_2}\right|^{\mu_0}_{\tilde{\mathcal B}}\le 2^{\mu_0}\left(|\omega|^{\mu_0}_{\tilde{\mathcal B}}\right)^2\le 2^{\mu_0}M_0^2\\
|[\omega]_k|^{\mu_0+1}_{\tilde{\mathcal B}} &= \left|\frac{F_k}{|k|_2}\frac{F_{-k}}{|k|_2}\right|^{\mu_0+1}_{\tilde{\mathcal B}}\le 2^{\mu_0+1}\left(|\omega|^{\mu_0+1}_{\tilde{\mathcal B}}\right)^2\le 2^{\mu_0+1}M_1^2\,.
\end{aligned}$$

Now we are in a position to apply theorem 2.5.2 with $g=[\omega]_k|_{\tilde{\mathcal B}}\in\mathcal H(\tilde{\mathcal B},\mathbb R)\subset C^{\mu_0+1}(\tilde{\mathcal B},\mathbb R)$, $\tilde g=[\tilde\omega]_k\in C^{\mu_0}(\tilde{\mathcal B},\mathbb R)$ and $\tilde B=B$ obtaining

$$\operatorname{meas}_n\left\{b\in\tilde{\mathcal B}\,:\,[\tilde\omega]_k\le\varepsilon\right\}\le M\varepsilon^{\frac{1}{\mu_0}} \tag{2.5.30}$$

whenever

$$0<\varepsilon\le\frac{\beta}{2\mu_0+2} \tag{2.5.31}$$

and with M given by

$$M=3(2\pi e)^{\frac{n}{2}}\frac{(\mu_0+1)^{\mu_0+2}}{(\mu_0+1)!}d_0^{n-1}\left(n^{-\frac12}+2d_0+\vartheta^{-1}d_0\right)\left(\frac{1}{\beta}\right)^{1+\frac{1}{\mu_0}}2^{\mu_0+1}M_1^2 \tag{2.5.32}$$

(see (2.5.21), (2.5.22) and the estimate on $|[\omega]_k|_{\tilde{\mathcal{B}}}^{\mu_0+1}$). Now by the definition of $\mathcal{H}(\tilde{\omega})$ in (2.5.28) we have

$$\begin{aligned}
\mathcal{K} \setminus \mathcal{H}(\tilde{\omega}) &= \{b \in \mathcal{K} \,:\, |\langle k, \omega(b)\rangle| < \gamma\Phi(|k|_2)\,,\ \forall\ k \in \mathbb{Z}^n \setminus \{0\}\} = \\
&= \left\{b \in \mathcal{K} \,:\, |\langle k, \omega(b)\rangle|^2 < (\gamma\Phi(|k|_2))^2\,,\ \forall\ k \in \mathbb{Z}^n \setminus \{0\}\right\} = \\
&= \left\{b \in \mathcal{K} \,:\, \left|\tilde{F}_k\right|^2 < (\gamma\Phi(|k|_2))^2\,,\ \forall\ k \in \mathbb{Z}^n \setminus \{0\}\right\} = \\
&= \left\{b \in \mathcal{K} \,:\, |[\tilde{\omega}]_k|^2 < \left(\frac{\gamma\Phi(|k|_2)}{|k|_2}\right)^2,\ \forall\ k \in \mathbb{Z}^n \setminus \{0\}\right\}.
\end{aligned}$$

Then in view of (2.5.30) applied with

$$\varepsilon = \left(\frac{\gamma\Phi(|k|_2)}{|k|_2}\right)^2 \tag{2.5.33}$$

it results

$$\begin{aligned}
\text{meas}_n\,(\mathcal{K} \setminus \mathcal{H}(\tilde{\omega})) &\leq \text{meas}_n \bigcup_{k \in \mathbb{Z}^n \setminus \{0\}} \left\{b \in \mathcal{K} \,:\, |[\tilde{\omega}]_k|^2 < \left(\frac{\gamma\Phi(|k|_2)}{|k|_2}\right)^2\right\} \leq \\
&\leq \sum_{k \in \mathbb{Z}^n \setminus \{0\}} \text{meas}_n \left\{b \in \mathcal{K} \,:\, |[\tilde{\omega}]_k|^2 < \left(\frac{\gamma\Phi(|k|_2)}{|k|_2}\right)^2\right\} \leq \\
&\leq M\gamma^{\frac{2}{\mu_0}} \sum_{k \in \mathbb{Z}^n \setminus \{0\}} \left(\frac{\Phi(|k|_2)}{|k|_2}\right)^{\frac{2}{\mu_0}}.
\end{aligned}$$

To estimate this last sum we observe that the function $s \in (0, \infty) \longmapsto \Phi(s)s^{-1}$ is strictly decreasing because Φ is itself decreasing (see condition 1 in definition 2.2.1); therefore in view of $|k|_\infty \leq |k|_2$ for any $k \in \mathbb{Z}^n$ we have

$$\sum_{k \neq 0} \left(\frac{\Phi(|k|_2)}{|k|_2}\right)^{\frac{2}{\mu_0}} \leq \sum_{k \neq 0} \left(\frac{\Phi(|k|_\infty)}{|k|_\infty}\right)^{\frac{2}{\mu_0}} = \sum_{\nu=1}^{\infty} \left(\sum_{|k|_\infty = \nu} 1\right) \left(\frac{\Phi(\nu)}{\nu}\right)^{\frac{2}{\mu_0}}.$$

Denoting now $N_\nu := \sum_{0 < |k|_\infty \leq \nu} 1 = (2\nu + 1)^n - 1$ for $\nu \in \mathbb{N}_+$ and $N_0 = 0$, we can continue the above chain of equalities writing

$$\sum_{\nu=1}^{\infty}\left(\sum_{|k|_\infty=\nu}1\right)\left(\frac{\Phi(\nu)}{\nu}\right)^{\frac{2}{\mu_0}}=\sum_{\nu=1}^{\infty}(N_\nu-N_{\nu-1})\left(\frac{\Phi(\nu)}{\nu}\right)^{\frac{2}{\mu_0}}=$$
$$=\sum_{\nu=1}^{\infty}N_\nu\left(\frac{\Phi(\nu)}{\nu}\right)^{\frac{2}{\mu_0}}-\sum_{\nu=1}^{\infty}N_\nu\left(\frac{\Phi(\nu+1)}{\nu+1}\right)^{\frac{2}{\mu_0}}=$$
$$=\sum_{\nu=1}^{\infty}\left[(2\nu+1)^n-1\right]\int_\nu^{\nu+1}-\frac{d}{ds}\left[\left(\Phi(s)s^{-1}\right)^{\frac{2}{\mu_0}}\right]\,ds\le$$
$$\le-\int_1^\infty\left[(2s+1)^n-1\right]\frac{d}{ds}\left[\left(\Phi(s)s^{-1}\right)^{\frac{2}{\mu_0}}\right]\,ds=$$
$$=-\left[(2s+1)^n-1\right]\left(\Phi(s)s^{-1}\right)^{\frac{2}{\mu_0}}\Big|_1^\infty+2n\int_1^\infty(2s+1)^{n-1}\left(\frac{\Phi(s)}{s}\right)^{\frac{2}{\mu_0}}ds=$$
$$=3^n-1+2n\int_1^\infty(2s+1)^{n-1}\left(\frac{\Phi(s)}{s}\right)^{\frac{2}{\mu_0}}ds$$

having used in this last equality the property $s^\lambda\Phi(s)\overset{s\to\infty}{\longrightarrow}0$ for $\lambda\in\mathbb{R}$ and $\Phi(1)=1$ (recall once again definition 2.2.1). Now since $(2s+1)^n\le(3s)^n-1$ we may write

$$\begin{aligned}&3^n-1+2n\int_1^\infty(2s+1)^{n-1}\left(\frac{\Phi(s)}{s}\right)^{\frac{2}{\mu_0}}ds\le\\&\le3^n-1+2n3^{n-1}\int_1^\infty s^{-n-1-\frac{2}{\mu_0}}\left(s^n\Phi(s)^{\frac{1}{\mu_0}}\right)^2ds\le\\&\le3^n-1+2n3^{n-1}\left(\Phi_{n\mu_0}\right)^2\int_1^\infty s^{-n-1-\frac{2}{\mu_0}}\,ds=\\&=3^n+23^{n-1}\frac{n\mu}{n\mu+2}\left(\Phi_{n\mu_0}\right)^2\le3^{n+1}\left(\Phi_{n\mu_0}\right)^2.\end{aligned}\tag{2.5.34}$$

Finally we come back to the beginning of this chains of inequalities estimating

$$\text{meas}_n\left(\mathcal{K}\setminus\mathcal{H}(\tilde\omega)\right)=\text{meas}_n\mathcal{K}-\text{meas}_n\mathcal{H}(\tilde\omega)\le3^{n+1}M\gamma^{\frac{2}{\mu_0}}\left(\Phi_{n\mu_0}\right)^2.\tag{2.5.35}$$

Therefore in order to get (2.5.29), that is $\text{meas}_n\mathcal{H}(\tilde\omega)\ge\text{meas}_n\mathcal{K}-\epsilon^\star$, we must require

$$3^{n+1}M\gamma^{\frac{2}{\mu_0}}\left(\Phi_{n\mu_0}\right)^2\le\epsilon^\star$$

which is satisfied by means of $\gamma \le \gamma'$ as in (2.5.26) together with (2.5.27) and in view of M as taken in (2.5.32).

To complete the proof we still have to assure that condition (2.5.31) is satisfied with our choice of ε in (2.5.33). From the property verified by ω in (2.5.23) (which is a consequence of its non-degeneracy as we show in appendix C) we have

$$\beta \le |[\omega]_k|_{\tilde{\mathcal{B}}}^{\mu_0+1} \le 2^{\mu_0+1} M_1^2$$

$$\beta \le |[\omega]_k|_{\tilde{\mathcal{B}}}^{\mu_0} \le 2^{\mu_0} M_0^2 \tag{2.5.36}$$

having used the estimate done before during the proof. From this two inequalities, the definition of C' in (2.5.27) and $\Phi_{n\mu_0} \ge 1$ we derive

$$C' = 2^{\mu_0+1} \frac{(\mu_0+1)^{\mu_0+2}}{(\mu_0+1)!} (\Phi_{n\mu_0})^2 M_1^2 \ge (\mu_0+1) 2^{\mu_0+1} M_1^2 \ge (\mu_0+1)\beta\,. \tag{2.5.37}$$

and by hypothesis we have

$$\epsilon^\star < \text{ meas}_n \mathcal{K} \le (2d_0)^n\,. \tag{2.5.38}$$

Now observe that the first terms that constitute γ' in (2.5.26) verify

$$3^{n+2}(2\pi e)^{\frac{n}{2}} d_0^{n-1} \left(n^{-\frac{1}{2}} + 2d_0 + \vartheta^{-1} d_0 \right) C' \ge (2d_0)^n 2C'$$

so that we obtain, together with (2.5.33), $\Phi \le 1$, (2.5.26), (2.5.37) and (2.5.38),

$$\begin{aligned} \varepsilon &\le \gamma^2 \le (\gamma')^2 \le [(2d_0)^n 2C']^{-\mu_0} \beta^{\mu_0+1} (\epsilon^\star)^{\mu_0} = \\ &= \left[\frac{\epsilon^\star}{(2d_0)^n} \right]^{\mu_0} \left[\frac{(\mu_0+1)\beta}{C'} \right]^{\mu_0} \frac{\beta}{(2\mu_0+2)^{\mu_0}} \le \frac{\beta}{2\mu_0+2} \end{aligned}$$

To complete the proof we observe that if $|\omega|_{\tilde{\mathcal{B}}}^{\mu_0} \le M'$ from (2.5.25) and $\beta \ge 2^{\mu_0} M_0^2$ we obtain

$$|\omega - \tilde{\omega}| \le \frac{M_0}{4}$$

which gives

$$|\tilde{\omega}|_{\tilde{\mathcal{B}}}^{\mu_0} \le |\omega|_{\tilde{\mathcal{B}}}^{\mu_0} + |\omega - \tilde{\omega}|_{\tilde{\mathcal{B}}}^{\mu_0} \le M' + \frac{M_0}{4}\,. \tag{2.5.39}$$

Thus, taking $M^{(0)} = M_0 > rM'$, with $r := \frac{4}{3}$ we have that also $M^{(1)} = M' + \frac{1}{4}M_0 > rM'$ is a common upper bound for $|\omega|_{\tilde{\mathcal{B}}}^{\mu_0}$ and $|\tilde{\omega}|_{\tilde{\mathcal{B}}}^{\mu_0}$ as shown. Continuing this procedure we obtain a sequence of common upper bounds $M^{(j)} > rM'$; moreover, from the recursive relation $M^{(j+1)} = M' + \frac{1}{4}M^{(j)}$ we can infer that the sequence tends to rM' so that we can effectively take $2M'$ as a common upper bounds for the two norms considered $\square$

2.5.4 Links and Chains

In the following pages we complete the proof of theorem 2.5.1 displaying Rüßmann's theory of links and chain for the construction of non-resonant frequencies on an arbitrarily non-empty compact set contained in the domain of the action-variables. The full detailed construction of non-resonant frequencies (in the case of lower dimensional tori) can be found in [Rüßm01, pages 198-203].

Definition 2.5.1. *Let $\mathcal{K}_\nu \subseteq \mathbb{R}^n$ a non-empty set and $\omega_\nu \in \mathcal{H}(\mathcal{B}_\nu, \mathbb{R}^n)$ where $\mathcal{B}_\nu := \mathcal{K}_\nu + t_\nu$ (see (2.5.2) for t_ν), we call $\mathcal{L}_\nu = (\mathcal{K}_\nu, \omega_\nu)$ a* **link**.

The initial link in our iterative scheme is given by $\mathcal{L}_0 = (\mathcal{K}_0, \omega_0)$ with $\mathcal{K}_0$ and ω_0 as in subsection 2.5.1.

Definition 2.5.2. *A link $\mathcal{L}_\nu = (\mathcal{K}_\nu, \omega_\nu)$ is said to be* **open** *if $\mathcal{K}_\nu^\star \neq \emptyset$ where we recall*

$$K_\nu^\star := \left\{ b \in \mathcal{K}_\nu \mid |\langle k, \omega_\nu(b)\rangle| \geq \frac{1}{2}\gamma\Phi(T_\nu) \quad \forall\ 0 < |k|_2 \leq T_\nu \right\} \tag{2.5.40}$$

(see subsection 2.5.1 for the definition of T_ν).

Definition 2.5.3. *If a link $\mathcal{L}_\nu$ is open, we call a link $\mathcal{L}_{\nu+1} = (\mathcal{K}_{\nu+1}, \omega_{\nu+1})$ a* **successor** *of $\mathcal{L}_\nu$ if*

$$\mathcal{K}_{\nu+1} = \mathcal{K}_\nu^\star\ ,\ \ \omega_{\nu+1} = (\omega_\nu + \Delta\omega_\nu)|_{\mathcal{B}_{\nu+1}}\ ,\ \ \mathcal{B}_{\nu+1} = \mathcal{K}_{\nu+1} + t_{\nu+1} \tag{2.5.41}$$

where $\Delta\omega_\nu$ is an arbitrarily chosen function in $\mathcal{H}(\mathcal{B}_\nu, \mathbb{R}^n)$ satisfying the estimate

$$|\Delta\omega_\nu|_{\mathcal{B}} \leq \frac{L_\nu}{T_0} \tag{2.5.42}$$

(according to the framework described in subsection 2.5.1).

Finally we denote $\lambda \preceq \nu$ if $\lambda \leq \nu < \infty$ or $\lambda < \nu = \infty$ and give the following

Definition 2.5.4. *A collection of links*

$$(\mathcal{L}_\lambda)_{0\leq\lambda\preceq\nu} = \begin{cases} (\mathcal{L}_0, \mathcal{L}_1, \dots, \mathcal{L}_\nu) & \text{if } \nu < \infty \\ (\mathcal{L}_0, \mathcal{L}_1, \dots) & \text{if } \nu = \infty \end{cases}$$

is called a **chain** *if $\mathcal{L}_{\lambda+1}$ is a successor of $\mathcal{L}_\lambda$ for any $0 \leq \lambda < \nu$. The initial link $\mathcal{L}_0$ is itself considered as a chain. A chain is said to be* **maximal** *if $\nu < \infty$ and $\mathcal{L}_\nu$ is not open (i.e., $\mathcal{K}_\nu^\star = \emptyset$) or $\nu = \infty$*

Since we would like to work with C^∞-functions $\omega_0, \omega_1, \ldots$ defined on the same open set (say $\tilde{\mathcal{B}} := \mathcal{K} +_{\mathbb{R}} \vartheta$), we give here the following result:

Theorem 2.5.4. *Let $\mathcal{K} \subseteq \mathbb{R}^n$ be a non-empty set, $t > 0$ and $f : \mathcal{K} + t \longrightarrow \mathbb{C}^m$ be an holomorphic and bounded function assuming real values on $\mathbb{R}^n$ (we will always think $n = m$). Then there exists a C^∞-function $\tilde{f} : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ such that $\tilde{f}(x) = f(x)$ for every $x \in \mathcal{K}$ and the following estimates hold for every $\nu = 0, 1, \ldots$:*

$$\sup_{x \in \mathbb{R}^n} \sup_{a \in \mathbb{R}^n, |a|_2 = 1} \left| D^\nu \tilde{f}(x)(a^\nu) \right| \leq \frac{C(n,\nu)}{t^\nu} \sup_{x \in \mathcal{K}+t} |f(x)| \tag{2.5.43}$$

with $C(n,\nu)$ not depending on f and increasing in ν as defined in (2.4.7).

Let $\mathcal{L}_0 = (\mathcal{K}_0, \omega_0)$ the initial link, we define its **extension** $\tilde{\mathcal{L}}_0$ by

$$\tilde{\mathcal{L}}_0 := (\mathcal{K}_0, \omega_0, \tilde{\omega}_0) \quad \text{with} \quad \tilde{\omega}_0 = \omega|_{\tilde{\mathcal{B}}}$$

where $\tilde{\mathcal{B}} := \mathcal{K} +_{\mathbb{R}} \vartheta$. Thus we may attach to the chain $(\mathcal{L}_0) = \mathcal{L}_0$ the **extended chain** $(\tilde{\mathcal{L}}_0) = \tilde{\mathcal{L}}_0$.

We want now to define the extended chain of any assigned chain $(\mathcal{L}_\lambda)_{0 \leq \lambda \preceq \nu}$. Consider $\Delta\omega_\lambda \in \mathcal{H}(\mathcal{B}_\lambda, \mathbb{R}^n)$, for $0 \leq \lambda < \nu$, i.e., the functions defining the chain $(\mathcal{L}_\lambda)_{0 \leq \lambda \preceq \nu}$, we may apply to each of this functions theorem 2.5.4 with $f = \Delta\omega_\lambda$, $\mathcal{K} = \mathcal{K}_\lambda$, $t = t_\lambda$, obtaining $\widetilde{\Delta\omega_1}, \widetilde{\Delta\omega_2}, \cdots \in C^\infty(\tilde{\mathcal{B}}\mathbb{R}^n)$ such that

$$\widetilde{\Delta\omega_\lambda}\,|_{\mathcal{K}_\lambda} = \Delta\omega_\lambda|_{\mathcal{K}_\lambda} \quad \text{for} \quad 0 \leq \lambda < \nu \,.$$

Moreover by the estimate in (2.5.43) we have

$$|D^\mu \widetilde{\Delta\omega_\lambda}| \leq C(n,\mu) t_\lambda^{-\mu} \frac{L_\lambda}{T_0} \tag{2.5.44}$$

in view of (2.5.42). Now we are in position to define recursively C^∞-functions on $\tilde{\mathcal{B}} = \mathcal{K} +_{\mathbb{R}} \vartheta$ as follows

$$\tilde{\omega}_0 := \omega|_{\tilde{\mathcal{B}}} \ , \ \ \tilde{\omega}_{\lambda+1} := \tilde{\omega}_\lambda + \widetilde{\Delta\omega_\lambda} \tag{2.5.45}$$

for any $0 \leq \lambda < \nu$. Furthermore by induction we can easily obtain

$$\tilde{\omega_\lambda}|_{\mathcal{K}_\lambda} = \omega_\lambda|_{\mathcal{K}_\lambda} \quad \text{for} \quad 0 \leq \lambda \preceq \nu \,. \tag{2.5.46}$$

As this recursively construction shows, each link $\mathcal{L}_\lambda$ of the chain $(\mathcal{L}_\lambda)_{0 \leq \lambda \preceq \nu}$ possesses a well defined extension $\tilde{\mathcal{L}}_\lambda = (\mathcal{K}_\lambda, \omega_\lambda, \tilde{\omega}_\lambda)$. Therefore, in the sense described, we can call $\tilde{\mathcal{L}}_{\lambda+1} = (\mathcal{K}_{\lambda+1}, \omega_{\lambda+1}, \tilde{\omega}_{\lambda+1})$ a successor of $\tilde{\mathcal{L}}_\lambda$ accordingly to definitions, and define the extended chain

$$(\tilde{\mathcal{L}}_\lambda)_{0\le\lambda\preceq\nu} = \begin{cases} (\tilde{\mathcal{L}}_0, \tilde{\mathcal{L}}_1, \dots, \tilde{\mathcal{L}}_\nu) & \text{if } \nu < \infty \\ \\ (\tilde{\mathcal{L}}_0, \tilde{\mathcal{L}}_1, \dots) & \text{if } \nu = \infty \end{cases}$$

as an **extension** of the chain $(\mathcal{L}_\lambda)_{0\le\lambda\preceq\nu}$.

Lemma 2.5.2 (Estimates for extended chains 1). *Let $0 \le \mu < \frac{\tau_0}{\tau}$, (where τ is chosen accordingly to (2.5.2)) and $(\mathcal{L}_\lambda)_{0\le\lambda\preceq\nu}$ be a chain with extension $(\tilde{\mathcal{L}}_\lambda)_{0\le\lambda\preceq\nu}$.*

Then for every $0 \le \sigma \le \lambda \preceq \nu$ it holds

$$|D^\mu(\tilde{\omega}_\lambda - \tilde{\omega}_\sigma)|_{\tilde{\mathcal{B}}} \le C(n,\nu)\,\frac{L_0}{T_0 t_0^\mu}\,\frac{\delta^{(\tau_0-\mu\tau)\sigma} - \delta^{(\tau_0-\mu\tau)\lambda}}{1-\delta^{\tau_0-\mu\tau}} \tag{2.5.47}$$

and consequently

$$|\tilde{\omega}_\lambda - \tilde{\omega}_0|^\mu_{\tilde{\mathcal{B}}} \le C(n,\mu)\,\frac{L_0}{T_0 t_0^\mu}\,\frac{1}{1-\delta^{\tau_0-\mu\tau}} \tag{2.5.48}$$

where $C(n,\mu)$ is defined in (2.4.7). Moreover, for any $0 \le \lambda \preceq \nu$ it results

$$|\tilde{\omega}_\lambda|^\mu_{\tilde{\mathcal{B}}} \le \frac{\mu!}{\vartheta^\mu} C_1 + \le C(n,\mu)\,\frac{L_0}{t_0^\mu}\,\frac{1}{1-\delta^{\tau_0-\mu\tau}}$$

Proof From the recursive construction described before we get $\tilde{\omega}_\lambda - \tilde{\omega}_\sigma = \widetilde{\Delta\omega}_\sigma + \dots + \widetilde{\Delta\omega}_{\lambda-1}$ and with estimate (2.5.44) we have

$$\begin{aligned} |D^\mu(\tilde{\omega}_\lambda - \tilde{\omega}_\sigma)|_{\tilde{\mathcal{B}}} &\le \sum_{j=\sigma}^{\lambda-1} |D^\mu \widetilde{\Delta\omega}_j|_{\tilde{\mathcal{B}}} \le \frac{C(n,\mu)}{T_0} \sum_{j=\sigma}^{\lambda-1} \frac{L_j}{t_j^\mu} \\ \le\ & C(n,\mu)\frac{L_0}{T_0 t_0^\mu} \sum_{j=\sigma}^{\lambda-1} \delta^{(\tau_0-\mu\tau)j} \le C(n,\mu)\frac{L_0}{T_0 t_0^\mu}\,\frac{\delta^{(\tau_0-\mu\tau)\sigma} - \delta^{(\tau_0-\mu\tau)\lambda}}{1-\delta^{\tau_0-\mu\tau}}\,. \end{aligned}$$

As a consequence we get for every $j = 0, 1, \dots, \mu < \frac{\tau_0}{\tau}$

$$|D^j(\tilde{\omega}_\lambda - \tilde{\omega}_0)|_{\tilde{\mathcal{B}}} \le C(n,j)\frac{L_0}{T_0 t_0^j}\,\frac{1}{1-\delta^{\tau_0-j\tau}}\,;$$

now, since $\delta, t_0 < 1$ and $C(n,j)$ is increasing in j, the right member of the above inequality is increasing and therefore

$$|(\tilde{\omega}_\lambda - \tilde{\omega}_0)|^\mu_{\tilde{\mathcal{B}}} = \max_{0\le j\le\mu} |D^j(\tilde{\omega}_\lambda - \tilde{\omega}_0)|_{\tilde{\mathcal{B}}} \le C(n,\mu)\frac{L_0}{T_0 t_0^\mu}\,\frac{1}{1-\delta^{\tau_0-\mu\tau}}\,.$$

From this last inequality and Cauchy's estimate we infer

$$|\tilde{\omega}_\lambda|^\mu_{\tilde{\mathcal{B}}} \le |\tilde{\omega}_0|^\mu_{\tilde{\mathcal{B}}} - |\tilde{\omega}_\lambda - \tilde{\omega}_0|^\mu_{\tilde{\mathcal{B}}} \le \frac{\mu!}{\vartheta^\mu} C_1 + \le C(n,\mu)\, \frac{L_0}{t_0^\mu} \frac{1}{1-\delta^{\tau_0 - \mu\tau}} \quad \square$$

Lemma 2.5.3 (Estimates for extended chains 2). *Let* $(\tilde{\mathcal{L}}_\lambda)_{0\le\lambda\le\nu<\infty}$ *and extended chain with* $\tilde{\mathcal{L}}_\lambda = (K_\lambda, \omega_\lambda, \tilde{\omega}_\lambda)$ *and assume that* L_0 *in (2.5.2) satisfies*

$$L_0 \le \frac{\gamma}{2(C_1+1)} \Phi(T_0)(1-\delta^{\tau_0})\,. \tag{2.5.49}$$

Then we have

$$|\tilde{\omega}_\lambda|_{\tilde{\mathcal{B}}} \le C_1 + 1 \tag{2.5.50}$$

for any $0 \le \lambda \le \nu$*; furthermore for any function* $\hat{\omega} \in C^{\mu_0}(\tilde{\mathcal{B}}, \mathbb{R}^n)$ *verifying the estimates*

$$|\hat{\omega}|_{\tilde{\mathcal{B}}} \le C_1 + 1 \tag{2.5.51}$$

and

$$|\hat{\omega} - \tilde{\omega}_\lambda|_{\tilde{\mathcal{B}}} \le \frac{L_0 \delta^{\tau_0 \lambda}}{T_0(1-\delta^{\tau_0})} \tag{2.5.52}$$

for any $0 \le \lambda \le \nu$*, we have*

$$\mathcal{K} = \mathcal{K}_0 \supseteq \mathcal{K}_1 \supseteq \cdots \supseteq \mathcal{K}_\nu \supseteq \mathcal{K}^\star_\nu \supseteq \bigcap_{k\in\mathbb{Z}^n, |k|_2 \le T_\nu} \mathcal{H}_k(\hat{\omega}) \tag{2.5.53}$$

where

$$\mathcal{H}_k(\hat{\omega}) := \{ b \in \mathcal{K} \,:\, |\langle k, \hat{\omega}(b)\rangle| \ge \gamma \Phi(|k|_2)\}$$

Proof To prove (2.5.50) it is sufficient to observe that from (2.5.4) with $\mu = 0$ we have

$$|\tilde{\omega}_l|_{\tilde{\mathcal{B}}} \le C_1 + \frac{L_0}{1-\delta^{\tau_0}} \le C_1 + \frac{\gamma}{2(C_1+1)} \Phi(T_0) \le C_1 + 1$$

having used hypothesis in (2.5.49), $\gamma \in [0,2]$ and $\Phi(T) \le 1$.

For what concerns the remaining part of the statement we just have to prove

$$\bigcap_{j=1}^{\nu} \mathcal{H}_j(\hat{\omega}) \subseteq \mathcal{K}^\star_\nu$$

since the other inclusions in (2.5.53) are given by the construction of the compacts $\mathcal{K}_0, \mathcal{K}_1, \ldots, \mathcal{K}_\nu$ and the definition of $\mathcal{K}^\star_\nu$ in (2.5.40). We start showing

$$\mathcal{H}_k(\hat{\omega}) \subseteq \mathcal{H}_{k\lambda} := \left\{ b \in \mathcal{K} : |\langle k, \tilde{\omega}_\lambda(b)\rangle| \ge \frac{\gamma}{2} \Phi(T_\lambda) \right\} \tag{2.5.54}$$

for all $k \in \mathbb{Z}^n$ with $|k|_2 \leq T_\lambda$ and $l \in \mathbb{N}_0$ such that $\lambda \leq \nu$. If b belongs to $\mathcal{H}_k(\hat{\omega})$ then $|\langle k, \hat{\omega}(b)\rangle| \geq \gamma\Phi(|k|_2)$ so that it results

$$\begin{aligned}
|\langle k, \tilde{\omega}_\lambda(b)\rangle| \;&\geq\; |\langle k, \hat{\omega}(b)\rangle| - |\langle k, \tilde{\omega}_\lambda(b) - \hat{\omega}(b)\rangle| \geq \gamma\Phi(|k|_2) - |k|_2 |\tilde{\omega}_\lambda - \hat{\omega}|_{\tilde{\mathcal{B}}} \geq \\
&\geq\; \gamma\Phi(T_\lambda) - \frac{T_\lambda}{T_0}\frac{L_0\delta^{\tau_0\lambda}}{1-\delta^{\tau_0}} \geq \gamma\Phi(T_\lambda) - \frac{\gamma}{2(C_1+1)}\frac{T_\lambda}{T_0}\Phi(T_0)\delta^{\tau\lambda} = \\
&\geq\; \gamma\Phi(T_\lambda) - \frac{\gamma}{2(C_1+1)}\left(\frac{T_0}{T_\lambda}\right)^d \Phi(T_\lambda) \geq \frac{\gamma}{2}\Phi(T_\lambda)
\end{aligned}$$

having used of hypotheses (2.5.52) and (2.5.49), inequalities $1 \leq T_0 \leq T_\lambda$ and $\tau \leq \tau_0$ and $\Phi(T_0)\delta^{\tau\lambda}T_\lambda^{d+1} = \Phi(T_\lambda)T_0^{d+1}$. So, we have proved inclusion (2.5.54) which gives

$$\mathcal{K}^\star_\lambda = \mathcal{K}_\lambda \cap \bigcap_{|k|_2 \leq T_\lambda} \mathcal{H}_{k\lambda} \;\supseteq\; \mathcal{K}_\lambda \cap \bigcap_{|k|_2 \leq T_\lambda} \mathcal{H}_k(\hat{\omega})$$

for $\lambda \leq \nu$. Finally, from $\mathcal{K}_{\lambda+1} = \mathcal{K}_\lambda$, for any $0 \leq \lambda < \nu$, and induction we obtain

$$\mathcal{K}^\star_\nu \supseteq \mathcal{K}_0 \cap \bigcap_{|k|_2 \leq T_\nu} \mathcal{H}_k(\hat{\omega}) = \bigcap_{|k|_2 \leq T_\nu} \mathcal{H}_k(\hat{\omega}) \qquad \square$$

Theorem 2.5.5 (Theorem on chains). *Let Φ be the chosen approximation function according to definition (2.2.1) and assume that*

$$\epsilon^\star \in (0, \textit{meas}_n\mathcal{K}) \tag{2.5.55}$$

and

$$\begin{aligned} &0 < \tau < \tau^\star \quad\; 0 < \gamma \leq \gamma^\star \\ &0 < L_0 \leq L_1^\star \quad 0 < L_0 \leq \gamma L_2^\star \end{aligned} \tag{2.5.56}$$

for $\tau^\star, \gamma^\star, L_1^\star$ and $L_2^\star$ as in (2.5.13), (2.5.14), (2.5.15) and (2.5.16) respectively. Then the following is true

1. *Any maximal chain is infinite.*
2. *For any infinite chain $(\mathcal{L}_\lambda)_{0\leq\lambda\leq\infty}$, with $\mathcal{L}_\lambda = (\mathcal{K}_\lambda, \omega_\lambda)$, the limit $\omega_\nu \longrightarrow \omega_\infty$ exists uniformly on*

   $$\mathcal{K}_\infty := \bigcap_{\nu=0}^{\infty} \mathcal{K}_\nu \neq \emptyset$$

   so that $\omega_\infty : \mathcal{K}_\infty \longrightarrow \mathbb{R}^n$ is continuous. Furthermore, if we define

   $$\mathcal{H}_\infty := \{b \in \mathcal{K}_\infty \;:\; |\langle k, \omega_\infty(b)\rangle| \geq \gamma\Phi(|k|_2)\,,\;\; \forall\; k \in \mathbb{Z}^n\} \tag{2.5.57}$$

   we have the estimate

   $$\textit{meas}_n\mathcal{K}_\infty \geq\; \textit{meas}_n\mathcal{H}_\infty \geq\; \textit{meas}_n\mathcal{K} - \epsilon^\star \tag{2.5.58}$$

Proof Let $(\mathcal{L}_\lambda)_{0\le\lambda\le\nu<\infty}$ be a maximal chain (that is $\mathcal{K}^\star_\nu$ as in (2.5.40) is empty) and let $(\tilde{\mathcal{L}}_\lambda)_{0\le\lambda\le\nu}$ be its extension with $\tilde{L}_\lambda = (\mathcal{K}_\lambda, \omega_\lambda.\tilde{\omega}_\lambda)$. By Cauchy's estimate and $\tilde{\mathcal{B}} = \mathcal{K} +_{\mathbb{R}} \vartheta$ we have

$$|\omega|^{\mu_0}_{\tilde{\mathcal{B}}} \le \mu_0!\vartheta^{-\mu_0}|\omega|_{\mathcal{K}+2\vartheta} \le \mu_0!\vartheta^{-\mu_0}(C_1+1) := M' \tag{2.5.59}$$

and analogously

$$|\omega|^{\mu_0+1}_{\tilde{\mathcal{B}}} \le (\mu_0+1)!\vartheta^{-(\mu_0+1)}(C_1+1) := M_1\,; \tag{2.5.60}$$

then referring to estimates (2.5.24) and the last statement in theorem (2.5.3), we can take

$$M_0 =: 2M' = 2\mu_0!\vartheta^{-\mu_0}(C_1+1) \tag{2.5.61}$$

From $0<\tau<\tau_\star$ and the choice of $\tau_\star$ in (2.5.13) we have

$$\mu_0 < \frac{\tau_0}{\tau}$$

such that we can meet hypothesis in lemma 2.5.2 and obtain, with $\mu = \mu_0$ and $\lambda = \nu$ in (2.5.47),

$$\begin{aligned}
|\tilde{\omega}_\nu - \tilde{\omega}_0|^{\mu_0}_{\tilde{\mathcal{B}}} &\le C(n,\mu_0)\frac{L_0}{T_0 t_0^{\mu_0}}\frac{1}{1-\delta^{\tau_0-\mu\tau}} \le C(n,\mu_0)\frac{L_1^\star}{t_0^{\mu_0}}\frac{1}{1-\delta^{\tau_0-\mu\tau}}\\
&\le \frac{\beta}{2M^\star} = \frac{\beta}{2^{\mu_0+2}M_0}
\end{aligned}$$

where $M_\star$ has been taken from (2.5.17) and M_0 is defined in (2.5.61). Moreover $\tilde{\omega}_\nu$ is a C^∞-function on $\tilde{\mathcal{B}}$ since it is a member of an extended chain (it is the extension of ω_ν) and it holds

$$|\tilde{\omega}_\nu|^{\mu_0}_{\tilde{\mathcal{B}}} \le |\tilde{\omega}_0|^{\mu_0}_{\tilde{\mathcal{B}}} + |\tilde{\omega}_\nu - \tilde{\omega}_0|^{\mu_0}_{\tilde{\mathcal{B}}} \le M' + \frac{\beta}{2^{\mu_0+2}M_0} \le M' + \frac{M_0}{4} < M_0 \tag{2.5.62}$$

where we used $\beta < 2^{\mu_0}M_0^2$ as we proved in (2.5.36). Then to apply theorem 2.5.3 with $\tilde{\omega}_\nu = \tilde{\omega}$ it is sufficient to observe that γ' in (2.5.26), with C' in (2.5.27), equals $\gamma^\star$ in (2.5.14) with $C^\star$ in (2.5.18) once that M_1 has been inserted from (2.5.60). Therefore it results

$$\text{meas}_n\mathcal{H}(\tilde{\omega}_\nu) \ge \text{meas}_n\mathcal{K} - \epsilon^\star > 0 \tag{2.5.63}$$

where

$$\mathcal{H}(\tilde{\omega}_\nu) = \{b\in\mathcal{K} \,:\, |\langle k, \tilde{\omega}_\nu(b)\rangle| \ge \gamma\Phi(|k|_2)\,,\ \forall\, k\in\mathbb{Z}^n\setminus\{0\}\}\,. \tag{2.5.64}$$

Now by (2.5.56) and (2.5.16) we have

$$L_0 \leq \gamma L_2^\star \leq \frac{\gamma}{2(C_1+1)}\Phi(T_0)(1-\delta^{\tau_0}) \tag{2.5.65}$$

such that we can meet hypothesis (2.5.49) in lemma (2.5.3) with $\hat\omega = \tilde\omega_\nu$; moreover $|\tilde\omega_\nu|_{\tilde{\mathcal{B}}}^{\mu_0} \leq C_1 + 1$ by (2.5.50) with $\lambda = \nu$ and (2.5.52) follows (always with $\hat\omega = \tilde\omega_\nu$) from (2.5.47) with $\mu = 0$, $\lambda = \nu$ and $\sigma = \lambda$. Then lemma 2.5.3 and (2.5.63) yield

$$\begin{aligned}\mathcal{K}_\nu^\star \;&\supseteq\; \bigcup_{|k|_2\leq T_\nu} \mathcal{H}_k(\tilde\omega_\nu) = \{b\in\mathcal{K} : |\langle k, \tilde\omega_\nu(b)\rangle| \geq \gamma\Phi(|k|_2)\ \ \forall\ 0<|k|_2\leq T_\nu\} \supseteq \\ &\supseteq\; \mathcal{H}(\tilde\omega_\lambda)\neq\emptyset\,.\end{aligned}$$

So, we have reached a contradiction to the maximality of the considered chain $(\mathcal{L}_\lambda)_{0\leq\lambda\leq\nu<\infty}$.

Now $\mu_0 \leq \tau^0\tau^{-1}$ permits once again the application of lemma 2.5.2, with $\nu = \infty$ and $0 \leq \mu \leq \mu_0$, to the infinite chain $(\mathcal{L}_\lambda)_{0\leq\lambda<\infty}$; by inequality (2.5.47) we can infer that the extended chain $(\tilde{\mathcal{L}}_\lambda)$ is a Cauchy sequence in $C^{\mu_0}(\mathcal{B},\mathbb{R}^n)$ (since $\delta < 1$ and $\tau_0 - \mu\tau > 0$). Then there exists $\tilde\omega_\infty \in C^{\mu_0}(\mathcal{B},\mathbb{R}^n)$ such that

$$\tilde\omega_\lambda \overset{\lambda\to\infty}{\longrightarrow} \tilde\omega_\infty \tag{2.5.66}$$

uniformly in $C^{\mu_0}(\mathcal{B},\mathbb{R}^n)$. Thus, performing this limit in (2.5.62) and (2.5.62) we get

$$|\tilde\omega_\infty - \tilde\omega_0|_{\tilde{\mathcal{B}}}^{\mu_0} \leq \frac{\beta}{2^{\mu_0+2}M_0}\ , \qquad |\tilde\omega_\infty|_{\tilde{\mathcal{B}}}^{\mu_0} \leq M_0$$

so that we are able to apply theorem 2.5.3 with $\tilde\omega = \tilde\omega_\infty$ obtaining

$$\operatorname{meas}_n \mathcal{H}(\tilde\omega_\infty) \geq \operatorname{meas}_n \mathcal{K} - \epsilon^\star\,, \tag{2.5.67}$$

where as usual we denote

$$\mathcal{H}(\tilde\omega_\infty) = \{b\in\mathcal{K}\,:\,|\langle k,\tilde\omega_\infty(b)\rangle| \geq \gamma\Phi(|k|_2)\,,\ \forall\ k\in\mathbb{Z}^n\setminus\{0\}\}\,. \tag{2.5.68}$$

Now we want to apply once again lemma 2.5.3 with $\hat\omega = \tilde\omega_\infty$. Clearly the links $\mathcal{L}_0, \mathcal{L}_1, \ldots, \mathcal{L}_\nu$ of the considered chain form a finite chain; the two estimates (2.5.51) and (2.5.52) can be obtained by taking the limit for $\lambda\to\infty$ in the relations obtained before when we applied lemma 2.5.3 with $\hat\omega = \tilde\omega_\nu$. Then, as above, we have

$$\mathcal{K}_\nu \supseteq \mathcal{K}_\nu^\star \supseteq \bigcap_{|k|_2\leq T_\nu} \mathcal{H}_k(\tilde\omega_\infty) \supseteq \mathcal{H}(\tilde\omega_\infty)$$

for $\nu = 0, 1, \ldots$, where

$$\bigcap_{|k|_2 \le T_\nu} \mathcal{H}_k(\tilde{\omega}_\infty) = \{b \in \mathcal{K} \,:\, |\langle k, \tilde{\omega}_\infty(b)\rangle| \ge \gamma \Phi(|k|_2) \;\; \forall\; 0 < |k|_2 \le T_\nu\}$$

and $\mathcal{H}(\tilde{\omega}_\infty)$ is given by (2.5.64) with $\tilde{\omega}_\infty$ instead of $\tilde{\omega}_\nu$. Thus by means of the definition of $\mathcal{K}_\infty$ and (2.5.67) it results

$$\mathcal{K}_\infty = \bigcap_{\nu=1}^{\infty} K_\nu \supseteq \mathcal{H}(\tilde{\omega}_\infty) \neq 0\,;$$

Now, since $\tilde{\omega}_\lambda|_{\mathcal{K}_\lambda} = \omega_\lambda|_{\mathcal{K}_\lambda}$, for any $\lambda \in \mathbb{N}$, we have

$$\omega_\lambda|_{\mathcal{K}_\infty} = \tilde{\omega}_\lambda|_{\mathcal{K}_\infty} \stackrel{\lambda\to\infty}{\longrightarrow} \tilde{\omega}_\infty|_{\mathcal{K}_\infty} := \omega_\infty \tag{2.5.69}$$

where this convergence is uniform on $\mathcal{K}_\infty$. As a consequence $\mathcal{H}(\tilde{\omega}_\infty) = \mathcal{K}_\infty \cap \mathcal{H}(\tilde{\omega}_\infty) = \mathcal{H}_\infty$ and by (2.5.67) we finally obtain

$$\text{meas}_n \mathcal{K}_\infty \ge \text{meas}_n \mathcal{H}_\infty \ge \text{meas}_n \mathcal{K} - \epsilon^\star \quad \square$$

Chapter 3

Properly degenerate Hamiltonian systems

3.1 Statement of results

Let $\mathcal{B}$ be an open set in $\mathbb{R}^d$, $\mathcal{U}$ some open neighborhood of the origin in $\mathbb{R}^{2p}$ and ϵ a "small" real and positive parameter, we consider an Hamiltonian function in the form

$$H_\epsilon(\varphi, I, u, v) = h(I) + \epsilon f(\varphi, I, u, v) \tag{3.1.1}$$

real-analytic for

$$(\varphi, I, (u, v)) \in \mathbb{T}^d \times \mathcal{B} \times \mathcal{U} := \mathcal{M} \tag{3.1.2}$$

endowed with the standard symplectic form

$$dI \wedge d\varphi + du \wedge dv\,. \tag{3.1.3}$$

Let $\sigma, r_0, r_1 > 0$ be such that H possesses an holomorphic extension on

$$\mathbb{T}^d_\sigma \times \mathcal{B}_{r_0} \times \mathcal{U}_{r_1} := \mathcal{M}_\star\,.$$

We assume that f is in the form

$$f(\varphi, I, u, v) = f_0(I, u, v) + f_1(\varphi, I, u, v) \quad \text{with} \quad \int_{\mathbb{T}^d} f_1(\varphi, I, u, v)\, d\varphi = 0$$

where

$$f_0(I, u, v) = f_{00}(I) + \sum_{j=1}^{p} \Omega_j(I) \frac{u_j^2 + v_j^2}{2} + f_2(I, u, v) \tag{3.1.4}$$

with

$$\sup_{I \in \mathcal{B}_{r_0}} |f_2(I, u, v)| \le c_0 |(u, v)|^3\,, \quad \forall \ (u, v) \in \mathcal{U}_{r_1} \tag{3.1.5}$$

for some $c_0 > 0$.

Now, observe that the Hamiltonian $h + \epsilon f_0$ possesses for every $I_0 \in \mathcal{B}$ the invariant lower dimensional torus

$$\mathbb{T}^d_{I_0} := \mathbb{T}^d \times \{I_0\} \times \{0\} \subset \mathcal{M}$$

with corresponding quasi-periodic flow

$$\varphi(t) = \left[\frac{\partial h}{\partial I}(I_0) + \epsilon\frac{\partial f_{00}}{\partial I}(I_0)\right] t + \varphi_0 \qquad I(t) \equiv I_0 \qquad (u(t), v(t)) \equiv 0\,.$$

Disregarding the elliptic singularity in every single elliptic plane $u_j v_j$, we aim to find Lagrangian invariant tori for H_ϵ, i.e., maximal tori in the form

$$\mathbb{T}^{d+p}_{I_0,w} = \mathbb{T}^d \times \{I_0\} \times \{(u,v) \in \mathbb{R}^{2p}\ ,\ |(u_j, v_j)| = w_j\ ,\ \ \forall\ j = 1, \dots, p\}$$

for I_0 in $\mathcal{B}$ and $w \in \mathbb{R}^p$ with $w_j > 0$.

The main theorem we are going to prove in this and next chapter is the following.

Theorem 3.1.1. *Let $H_\epsilon(\varphi, I, u, v) = h(I) + \epsilon f(\varphi, I, u, v)$ be the real-analytic Hamiltonian described above. Assume that the "frequency map", i.e., the real-analytic application*

$$I \in \mathcal{B} \longrightarrow (\nabla h(I), \Omega_1(I), \dots, \Omega_p(I)) \in \mathbb{R}^d \times \mathbb{R}^p\ , \qquad (3.1.6)$$

is non-degenerate in the sense of Rüßmann (definition 2.1.2). Then, provided that ϵ is sufficiently small, in any neighborhood of $\mathbb{T}^d \times \{I_0\} \times \{0, 0\} \subset \mathcal{M}$ there exists a positive measure set of phase space point belonging to analytic maximal KAM tori for H_ϵ carrying quasi-periodic motions.

We observe that theorem 3.1.1 is an analogous, in analytic class, of M. Herman's KAM theorem in [Her98] (a proof can be also found in [Féj04]). This last theorem is based on a C^∞ local inversion theorem on "tame" Fréchet spaces due to F. Sergeraert and R. Hamilton which, in turn, is related to the Nash-Moser implicit function theorem (refer to [Ham74] for an elegant proof given by R. Hamilton). In [Féj04], J. Féjoz applies the cited KAM theorem by M. Herman to the planetary $(n + 1)$-body problem; analogously, at the end of chapter 5, we are going to apply theorem 3.1.1 to the results given by J. Féjoz on the non-degeneration of the planetary frequency application (a discussion of J. Féjoz' results is provided in chapter 5 but we always refer to [Féj04] for more detailed proofs).

The proof of theorem 3.1.1 is performed in three main steps. The first step, carried out in this chapter, consists in the proof of the following theorem.

Theorem 3.1.2. *Assume that the "frequency map"*

$$I \in \mathcal{B} \longrightarrow (\nabla h(I), \Omega_1(I), \ldots, \Omega_p(I)) \in \mathbb{R}^d \times \mathbb{R}^p$$

is non-degenerate in the sense of Rüßmann. Chose and fix an integer $N \geq 2$ and an open set $V \subset \mathcal{B}$. Then, there exists an open ball

$$B^d(I_0, r) := \left\{ I \in \mathbb{R}^d \ : \ |I - I_0| < r \right\} \subset V$$

so that, provided ϵ is small enough (see inequality ((4.4.5)) and section 4.4), there exists a canonical transformation

$$\Phi_\epsilon : (\vartheta, r, \zeta, \rho) \ \in \mathbb{T}^d \times B^d(0, r/5) \times \mathbb{T}^p \times B^p(0, \epsilon) \longrightarrow$$

$$\longrightarrow (\varphi, I, u, v) \in \mathbb{T}^d \times B^d(I_0, r) \times \mathcal{U} \tag{3.1.7}$$

such that $\hat{H}_\epsilon := H_\epsilon \circ \Phi_\epsilon$ assumes the form

$$\hat{H}_\epsilon(\vartheta, \zeta, r, \rho; \epsilon) = \frac{1}{\epsilon} \left[h(I_0 + \epsilon r) + \epsilon \hat{g}(I_0 + \epsilon r) \right] + \frac{1}{2} \hat{\Omega}(I_0 + \epsilon r) \cdot (\rho^0 + \epsilon \rho)$$

$$+ Q_{\epsilon, I_0 + \epsilon r}(\rho^0 + \epsilon \rho) + \epsilon^N \left[\hat{F}_1(\zeta, r, \rho; \epsilon, \rho^0) + \hat{F}_2(\vartheta, \zeta, r, \rho; \epsilon, \rho^0) \right] \tag{3.1.8}$$

where ρ^0 in $(\mathbb{R}_+)^p$ is a chosen point having euclidean norm 2ϵ, $Q_{\epsilon, I_0 + \epsilon r}$ is a polynomial of degree $N - 1$ depending on ϵ and $I_0 + \epsilon r$, $\hat{g}$, $\hat{\Omega}$, $\hat{F}_1$ and $\hat{F}_2$ are real-analytic functions.

Furthermore it results

$$\sup_{B^d(I_0, r/4)} |\hat{\Omega} - \Omega| = O(\epsilon) \tag{3.1.9}$$

(see (3.3.30)) and proposition 3.3.1 for details).

The proof of this theorem is performed in sections 3.2, 3.3, 3.4 and 3.5. Chapter 3 is then dedicated to the application of Rüßmann's theorem in the case of maximal (Lagrangian) tori to $\hat{H}_\epsilon$. In sections 4.1 and 4.2 we prove

Theorem 3.1.3. *For small enough ϵ (see inequality (4.4.8) and section 4.4), the frequency application of the torus $\mathbb{T}^{d+p}_{r,\rho}$ of the integrable part of $\hat{H}_\epsilon$, i.e., the real-analytic function*

$$\hat{\Psi}_\epsilon : (r, \rho) \in B^d(0, r/5) \times B^{2p}(0, \epsilon) \longrightarrow \left(\frac{\partial}{\partial r} F_\epsilon \frac{\partial}{\partial \rho} F_\epsilon \right) \tag{3.1.10}$$

where $F_\epsilon := \hat{H}_\epsilon - \epsilon^N(\hat{F}_1 + \hat{F}_2)$, is non-degenerate in the sense of Rüßmann.

Furthermore if $\bar{\mu}$ and $\bar{\beta}$ denote the index and amount of non-degeneracy of the initial frequency application in (3.1.6), with respect to a compact set $\overline{\mathcal{K}} \subset B^d(I_0, r/5)$, then $\hat{\mu}_\epsilon$ and $\hat{\beta}_\epsilon$, the index and amount of non-degeneration of $\hat{\Psi}_\epsilon$ with respect to a suitable compact set $\mathcal{S} \subset B^d(0, r/5) \times B^p(0, \epsilon)$, verify

$$\hat{\mu}_\epsilon \leq \bar{\mu} \qquad \textit{and} \qquad \hat{\beta}_\epsilon \geq \frac{\epsilon^{\bar{\mu}+2}\bar{\beta}}{8}\,. \tag{3.1.11}$$

To conclude, in section 4.3 we will control how the quantities involved in the estimate of the size of the perturbation in Rüßmann' theorem 2.3.1 change their order in ϵ when $\hat{H}_\epsilon$ is considered, obtaining

Theorem 3.1.4. *If we take*

$$N = 2\bar{\mu}(\bar{\mu}+1)^2 + p\bar{\mu}^2 + 2(2a + 2d + 2p + 5)\bar{\mu} + 10$$

in theorem 3.1.2 and ϵ is small enough (see (4.3.53) and subsection 4.3), then it is possible to apply Rüßmann's theorem for maximal KAM tori to $\hat{H}_\epsilon$ obtaining $\gamma > 0$ (whose final determination takes place in (4.3.39)) and a positive measure set of phase space points corresponding to quasi-periodic motions with (γ, a)-Diophantine frequencies.

Observe that in this last results the frequencies $\omega = (\omega_1, \omega_2, \ldots, \omega_d)$ of the KAM tori obtained are (γ, a)-Diophantine, i.e.,

$$|\omega \cdot k| \geq \frac{\gamma}{|k|^a} \ \ , \ \gamma, a > 0$$

while in Rüßmann's theorem 2.3.1 the control on the frequencies is performed in a more general way through an approximation function (definition 2.2.1); we display the details of this difference in subsection 4.3.5 and give a complete exposition of the application of Rüßmann's theorem in section 4.3.

3.2 Averaging theory

Let f be a real-analytic function for (φ, I, u, v) in $\mathbb{T}^d \times B \times E \times F$ where $B \subset \mathbb{R}^d$ is an open set and E and F are two open neighborhoods of the origin in $\mathbb{R}^p$. Assume that f admits an holomorphic continuation on $\mathbb{T}^d_\sigma \times B_r \times E_{r_u} \times F_{r_v}$ and possesses the following Fourier's expansion

$$f(\varphi, I, u, v) = \sum_{k \in Z^d} f_k(I, u, v) e^{k \cdot \varphi}\,,$$

then we denote with $\|f\|_{\sigma,r,r_u,r_v}$ the norm given by

$$\|f\|_{\sigma,r,r_u,r_v} := \sum_{k\in\mathbb{Z}^d}\left(\sup_{B_r\times E_{r_u}\times F_{r_v}}|f_k(I,u,v)|\right)e^{|k|_1\sigma}\,.$$

We now display a well known issue of classical theory of dynamical systems, providing a general formulation of what is known as "averaging theorem":

Theorem 3.2.1 (Averaging theorem). *Let $H(\varphi,I,u,v) := h(I)+f(\varphi,I,u,v)$ a real-analytic Hamiltonian function on*

$$D := \mathbb{T}^d_\sigma \times B_r \times E_{r_u} \times F_{r_v} \tag{3.2.1}$$

where E and F are two neighborhoods of the origin in $\mathbb{R}^p$, and denote $\omega(I) := h'(I)$. Let Λ be a sub-lattice of $\mathbb{Z}^d$ and suppose that ω satisfies the non-resonance condition

$$|\omega(I)\cdot k| \geq \alpha > 0$$

for all $|k|\leq K$, $k\notin\Lambda$ and for all $I\in B_r$; suppose also that $K\sigma\geq 6$ and

$$\|f\|_{\sigma,r,r_ur_v} := \epsilon \leq \frac{\alpha d'}{2^7c_pK\sigma} \tag{3.2.2}$$

where $d' := \min\{rs, r_ur_v\}$ and $c_p = e(1+ep)/2$. Then there exists a real-analytic symplectic transformation

$$\Psi : (\tilde\varphi,\tilde I,\tilde u,\tilde v) \in \mathbb{T}^d_{\frac{\sigma}{6}} \times B_{\frac{r}{2}} \times E_{\frac{r_u}{2}} \times F_{\frac{r_v}{2}} \longrightarrow (\varphi,I,u,v)\in D$$

such that

$$\tilde H := H\circ\Psi = h+g+\tilde f$$

where g is in normal form

$$g(\tilde\varphi,\tilde I,\tilde u,\tilde v) = \sum_{k\in\Lambda} g_k(\tilde I,\tilde u,\tilde v)e^{ik\cdot\tilde\varphi} \tag{3.2.3}$$

and the two following inequalities hold

$$\|g-\sum_{|k|\leq K, k\in\Lambda} f_ke^{ik\cdot\tilde\varphi}\|_{\frac{\sigma}{6},\frac{r}{2},\frac{r_u}{2},\frac{r_v}{2}} \;\leq\; \frac{2^8c_p}{\alpha d}\epsilon^2 \tag{3.2.4}$$

$$\|\tilde f\|_{\frac{\sigma}{6},\frac{r}{2},\frac{r_u}{2},\frac{r_v}{2}} \leq e^{-\frac{K\sigma}{6}}\epsilon\,. \tag{3.2.5}$$

Moreover the projections of $\Psi(\tilde\varphi,\tilde I,\tilde u,\tilde v)$ satisfy the estimates

$$\sigma|\tilde\varphi-\varphi|,\; r|\tilde I-I|,\; r_u|\tilde u-u|,\; r_v|\tilde v-v| \leq 9\alpha^{-1}\epsilon\,.$$

Proof A proof of this theorem can be found in [Val03, Appendix A] $\square$

Corollary 3.2.1. *Let $\epsilon < 1$ and $H_\epsilon(\varphi, I, u, v) = h(I) + \epsilon f(\varphi, I, u, v)$ a real-analytic Hamiltonian function on $\mathcal{D} := \mathbb{T}^d_\sigma \times B_r \times E_{r_u} \times F_{r_v}$ define*

$$\mu := \|f\|_{\sigma,r,r_u,r_v} .$$

Let $N \geq 2$ an arbitrarily fixed real number and

$$K = \frac{6}{\sigma}(N-1)\log\frac{1}{\epsilon\mu}, \tag{3.2.6}$$

suppose that $\omega = \nabla h$ satisfies

$$|\omega(I)\cdot k| \geq \alpha > 0\,, \ \ \forall\ k \in \mathbb{Z}^d,\ 0 < |k|_1 \leq K, \ \ \forall\ I \in B_r\,; \tag{3.2.7}$$

then under the condition

$$\epsilon\mu\log\frac{1}{\epsilon\mu} \leq \frac{\alpha d'}{C_p(N-1)} \tag{3.2.8}$$

where $C_p = 2^7 3e(1+ep)$ and $d' := \min\{r\sigma, r_u r_v\}$, there exists an analytic symplectic transformation

$$\Psi_\epsilon : (\tilde{\varphi}, \tilde{I}, \tilde{u}, \tilde{v}) \in \mathbb{T}^d_{\frac{\sigma}{6}} \times B_{\frac{r}{2}} \times E_{\frac{r_u}{2}} \times F_{\frac{r_v}{2}} \longrightarrow (\varphi, I, u, v) \in \mathcal{D} \tag{3.2.9}$$

such that

$$\tilde{H}_\epsilon := H_\epsilon \circ \Psi = h + g + \tilde{f}$$

with

$$g = g(\tilde{I}, \tilde{u}, \tilde{v}) \tag{3.2.10}$$

satisfying

$$\|g - \epsilon f_0\|_{\frac{r}{2},\frac{r_u}{2},\frac{r_v}{2}} \leq \frac{C_p}{\alpha d}(\epsilon\mu)^2 \tag{3.2.11}$$

where

$$f_0 = f_0(\tilde{I}, \tilde{u}, \tilde{v}) = \frac{1}{(2\pi)^d}\int_{\mathbb{T}^d} f(\tilde{\varphi}, \tilde{I}, \tilde{u}, \tilde{v})d\tilde{\varphi}$$

is the average (with respect to the angles variables) of f; besides

$$\|\tilde{f}\|_{\frac{\sigma}{6},\frac{r}{2},\frac{r_u}{2},\frac{r_v}{2}} \leq (\epsilon\mu)^N . \tag{3.2.12}$$

Proof The proof of this statement can be easily obtained by the preceding formulation of averaging theorem; taking $\Lambda = \{0\}$ in (3.2.3) we can derive the independence of g from the angle variables $\tilde{\varphi} \in \mathbb{T}^d_{\frac{\sigma}{6}}$ (equation (3.2.10)) and obtain inequality (3.2.11) directly from (3.2.4). Moreover with the choice of K in (3.2.6), estimate (3.2.12) holds by mean of (3.2.5); observe that hypothesis (3.2.8) is needed accordingly with this choice of K in order to meet hypothesis (3.2.2) in theorem 3.2.1 $\square$

We state now a lemma showing how this corollary can be applied for I in an open set (say an open ball) under the hypothesis of non-degeneration in the sense of Rüßmann of $\omega = \nabla h$.

Lemma 3.2.1. *Let $\mathcal{I}$ be a domain in $\mathbb{R}^d$ and $f : I \in \mathcal{I} \longrightarrow \mathbb{R}^m$ a real-analytic and non-degenerate function in the sense of Rüßmann (see definition 2.1.2). Then for any chosen and fixed $K > 0$ there exist a point $I_\star \in \mathcal{I}$, a radius $r_\star > 0$ and a real number $\alpha > 0$ such that for any $I \in D_{r_\star}(I_\star) \subset \mathbb{C}^d$ it results*

$$|f(I) \cdot k| \geq \alpha\,, \quad \forall\; k \in \mathbb{Z}^m \;\; \textit{with} \;\; 0 < |k|_1 \leq K\,. \tag{3.2.13}$$

Proof Enumerate as $k_1, k_2, \dots, k_n$ (with $n = n(m, K)$) all the vectors with integers coordinate in $\mathbb{Z}^m$ having norm less than $[K]$, i.e.,

$$\{k_j\}_{j=1}^n = \mathbb{Z}^m_{[K]}$$

and set

$$f_j(y) = f(y) \cdot k_j\,.$$

Under the hypothesis of non degeneration, f_1 is a non-vanishing analytic function on $\mathcal{I}$ (because f does not satisfy any linear relation on $\mathcal{I}$) then there exists $I_1 \in \mathcal{I}$ such that $|f_1(I_1)| > 0$ and by continuity there exists a radius r_1 (we may assume to be less than the analycity radius of f in I_0) such that $|f| > 0$ holds on $D_{r_1}(I_1)$. Now since f_2 is also a non-vanishing real-analytic function there exists I_2 in $D_{r_1}(I_1) \cap \mathcal{I}$ such that $|f_2(I_2)| > 0$ Once again by continuity there exists r_2, which we may assume to be less than dist$(I_2, \partial D_{r_1}(I_1) \cap \mathcal{I})$, such that $|f_2(I)| > 0$ for every I in $D_{r_2}(I_2)$. Then we have both $|f_1|$ and $|f_2|$ greater than 0 on the whole ball $D_{r_2}(I_2)$.

Applying recursively the scheme described we may find $I_1, I_2, \dots, I_n$ in $\mathcal{I}$ and $r_1 \geq r_2 \geq \cdots \geq r_n > 0$ such that

1. $I_j \in D_{r_{j-1}}(I_{j-1}) \cap \mathcal{I}\,, \quad \forall\; 2 \leq j \leq n$

2. $D_{r_j}(I_j) \subseteq D_{r_{j-1}}(I_{j-1})\,, \quad \forall\; 2 \leq j \leq n$

3. $|f_1(I)|, \dots, |f_j(I)| > 0\,, \quad \forall\; I \in D_{r_j}(I_j)\,, \quad \forall\; 2 \leq j \leq n\,.$

In particular for every $I \in D_{r_n}(I_n)$ we have $|f_1(I)|, |f_2(I)|, \ldots, |f_n(I)| > 0$. By the open map theorem this functions assume their minima on $\partial D_{r_n}(I_n)$ therefore the proof of the statement follows taking $I_\star := I_n \in \mathcal{I}$, $r_\star := \frac{r_n}{2}$ and

$$\alpha := \min_{j=1,\ldots,n} \inf_{I \in D_{r_\star}} |f_j(I)| \qquad \square$$

Lemma 3.2.2. *Let $(\omega, \Omega) : B \subset \mathbb{R}^d \longrightarrow \mathbb{R}^d \times \mathbb{R}^p$ a real-analytic and non-degenerate function; let $s > 0$ such that (ω, Ω) can be holomorphically extended on B_s, and let K_1 and K_2 two positive integers greater than 1; then there exists $r \leq s$, $\alpha > 0$ and $I_0 \in B$ such that*

$$\inf_{I \in D_r(I_0)} |\omega(I) \cdot k| \geq \alpha\,, \quad \forall\ 0 < |k|_1 \leq K_1$$

$$\inf_{I \in D_r(I_0)} |\Omega(I) \cdot k| \geq \alpha\,, \quad \forall\ 0 < |k|_1 \leq K_2\,. \tag{3.2.14}$$

where $D_r(I_0) = \{I \in \mathbb{C}^d\ :\ |I - I_0| < r\}$ as usual.

Proof Since (ω, Ω) is non-degenerate in the sense of Rüßmann on B obviously both ω and Ω are non-degenerate on B. Applying lemma 3.2.1 with $f = \omega$ and $\mathcal{I} = B$ we obtain $I_1 \in B$, $r_1 \leq s$ and $\alpha_1 > 0$ such that

$$\inf_{I \in D_{r_1}(I_1)} |\omega(I) \cdot k| \geq \alpha_1\,, \quad \forall\ 0 < |k|_1 \leq K_1$$

and we may assume $D_{r_1}(I_1) \subseteq B_s$. Now with $f = \Omega$ and $\mathcal{I} = B_{r_1}(I_1) \cap \mathbb{R}^d$ we may find $r \leq r_1$, $\alpha_2 > 0$ and $I_0 \in B_{r_1}(I_1) \cap \mathbb{R}^d$, with $D_r(I_0) \subseteq D_{r_1}(I_1)$, such that

$$\inf_{I \in D_r(I_0)} |\Omega(I) \cdot k| \geq \alpha_2\,, \quad \forall\ 0 < |k|_1 \leq K_2\,.$$

The proof ends taking $\alpha = \min\{\alpha_1, \alpha_2\}$ $\square$

Consider now $H_\epsilon = h + \epsilon f$ as in (3.1.1) real-analytic on $\mathcal{M}_\star = \mathbb{T}^d_\sigma \times \mathcal{B}_{r_0} \times \mathcal{U}_{r_1}$. Let V be an open set contained in $\mathcal{B}$ (with the same notation of theorem 3.1.2) Let $N_1, N_2 \geq 2$ two fixed integers to be later determined, we take

$$K_1 = \frac{6}{\sigma}(N_1 - 1) \log \frac{1}{\epsilon\mu}$$

and apply lemma 3.2.2 (with $B = V$ and $s = r_0$) in order to find $\alpha > 0$, $0 < r \leq 0$ and $I_0 \in V$ such that

$$|\omega(I) \cdot k| \geq \alpha > 0\,,\ \ \forall\ k \in \mathbb{Z}^d,\ 0 < |k|_1 \leq K_1, \quad \forall\ I \in D_r(I_0) \tag{3.2.15}$$

and

$$|\Omega(I)\cdot k|\geq\alpha>0\,,\ \ \forall\ k\in\mathbb{Z}^p,\ 0<|k|_1\leq N_2,\ \ \forall\ I\in D_r(I_0) \tag{3.2.16}$$

where (ω,Ω) is the frequency application of H_ϵ (see (3.1.6)). Notice that under the hypotheses made $D_r(I_0)\subseteq\mathcal{B}_{r_0}$. Then, assuming

$$\epsilon\mu\log\frac{1}{\epsilon\mu}\leq\frac{\alpha d'}{c_p(N_1-1)} \tag{3.2.17}$$

and in view of (3.2.15), we may apply corollary 3.2.1 to H_ϵ with $E\times F=D\subset\mathbb{R}^{2p}$, $r_u=r_v=r_1$ and $B=\{I_0\}$ (which implies $B_r=D_r(I_0)$). Thus, we obtain a real-analytic and symplectic transformation

$$\Phi^1_\epsilon:(\tilde\varphi,\tilde I,\tilde u,\tilde v)\in\mathcal{D}_1:=\mathbb{T}^d_{\frac{\sigma}{6}}\times B_{\frac{r}{2}}\times D_{\frac{r_1}{2}}\longrightarrow(\varphi,I,u,v)\in\mathcal{D}_0:=\mathbb{T}^d_\sigma\times B_r\times D_{r_1} \tag{3.2.18}$$

where we denote from now on

$$B_t:=\left\{I\in\mathbb{C}^d\ :\ |I-I_0|<t\right\}=D^d_t(I_0)\,, \tag{3.2.19}$$

that cast H_ϵ in the form

$$H^1_\epsilon:=H_\epsilon\circ\Phi^1_\epsilon=h+g+\tilde f\,.$$

By equation (3.2.12) we have for the initially fixed $N_1\geq 2$

$$\|\tilde f\|_{\mathcal{D}_1}\leq(\epsilon\mu)^{N_1} \tag{3.2.20}$$

where μ is the Fourier's norm of f

$$\mu:=\|f\|_{\mathcal{D}_0}\,. \tag{3.2.21}$$

Moreover g verifies (3.2.10) (i.e., it is independent from the new angles $\tilde\varphi$) and satisfies

$$|g-\epsilon f_0|_{\frac{r}{2},\frac{r_1}{2}}\leq\frac{c_p}{\alpha d'}\,\epsilon^2 \tag{3.2.22}$$

(since $|\cdot|:=|\cdot|_2\leq\|\cdot\|$) where

$$c_p:=2^7 3e(1+ep) \tag{3.2.23}$$

and

$$d':=\min\{\sigma r,r_1^2\}\,; \tag{3.2.24}$$

therefore, we may write

$$g(\tilde I,\tilde u,\tilde v)=\epsilon f_0(\tilde I,\tilde u,\tilde v)+\tilde g(\tilde I,\tilde u,\tilde v)\ \ \text{with}\ \ |\tilde g|_{\frac{r}{2},\frac{r_1}{2}}\leq\frac{c_p}{\alpha d'}\,(\epsilon\mu)^2 \tag{3.2.25}$$

obtaining

$$H_\epsilon^1(\tilde\varphi,\tilde I,\tilde u,\tilde v)=h(\tilde I)+\epsilon f_0(\tilde I,\tilde u,\tilde v)+\tilde g(\tilde I,\tilde u,\tilde v)+\tilde f(\tilde\varphi,\tilde I,\tilde u,\tilde v)\,.$$

Now recalling the form of f_0 in (3.1.4) and using estimates (3.2.20) and (3.2.25), we set $\epsilon^2\bar g=\tilde g$ and $\epsilon^{N_1}\bar f=\tilde f$, and rewrite H_ϵ^1 in the form

$$H_\epsilon^1(\tilde\varphi,\tilde I,\tilde u,\tilde v)=h(\tilde I)+\epsilon\left[f_0(\tilde I,\tilde u,\tilde v)+\epsilon\bar g(\tilde I,\tilde u,\tilde v)\right]+\epsilon^{N_1}\bar f(\tilde\varphi,\tilde I,\tilde u,\tilde v)\,. \quad (3.2.26)$$

where

$$f_0(\tilde I,\tilde u,\tilde v)=f_{00}(\tilde I)+\sum_{j=1}^{p}\Omega_j(\tilde I)\frac{\tilde u_j^2+\tilde v_j^2}{2}+f_2(\tilde I,\tilde u,\tilde v) \quad (3.2.27)$$

with

$$\sup_{\tilde I\in B_{\frac{r}{2}}}|f_2(\tilde I,\tilde u,\tilde v)|\le c_0|(\tilde u,\tilde v)|^3\,, \quad (3.2.28)$$

the functions $\bar g$ and $\bar f$ are real-analytic on $\mathcal{D}_1$ and the following bounds hold:

$$\begin{aligned}|f_0|_{\frac{r}{2},\frac{r_1}{2}}&\le\mu\\ |\bar g|_{\frac{r}{2},\frac{r_1}{2}}&\le\frac{c_p}{\alpha d'}\mu^2\\ |\bar f|_{\mathcal{D}_1}&\le\mu^{N_1}\end{aligned} \quad (3.2.29)$$

(having used again $|\cdot|:=|\cdot|_2\le\|\cdot\|$).

3.3 Elliptic equilibrium for $f_0+\epsilon\bar g$

We now consider the real-analytic Hamiltonian function on $\mathcal{D}_1$

$$F(\tilde I,\tilde u,\tilde v,\epsilon)=f_0(\tilde I,\tilde u,\tilde v)+\epsilon\bar g(\tilde I,\tilde u,\tilde v)$$

as it appears in (3.2.26). As equation (3.2.27) together with (3.2.28) show, $\tilde z=(\tilde u,\tilde v)=0$ is an elliptic equilibrium point for $\Phi^t_{f_0}$; since this equilibrium is perturbed by the presence of $\epsilon\bar g$ our aim is to find an analytic and symplectic transformation Φ^2_ϵ, $0(\epsilon)$ close to identity, such that $F\circ\Phi^2_\epsilon$ possesses an elliptic equilibrium in the origin.

Define

$$G(\tilde I,\tilde u,\tilde v,\epsilon)=\Big(\partial_{\tilde u}F(\tilde I,\tilde u,\tilde v,\epsilon),\partial_{\tilde v}F(\tilde I,\tilde u,\tilde v,\epsilon)\Big)$$

and recall that from the definition of f_0 in (3.2.27) we have

$$G(\tilde{I},0,0,0)=0$$

and

$$\det\partial_{(\tilde{u},\tilde{v})}G(\tilde{I},0,0,0)=\prod_{i=1}^{p}\Big(\Omega_j(\tilde{I})\Big)^2>0$$

where both equations are verified for every $\tilde{I}\in B_{\frac{r}{2}}$. Then we can apply the Implicit Function Theorem to find two real-analytic functions $\tilde{u}$ and $\tilde{v}$

$$(\tilde{u},\tilde{v}):(\tilde{I},\epsilon)\in B_{\frac{r}{2}}\times\{|\epsilon|<\epsilon_1\}\longrightarrow\Big(\tilde{u}(\tilde{I},\epsilon),\tilde{v}(\tilde{I},\epsilon)\Big)\in D_\rho \tag{3.3.1}$$

with ϵ_1 and ρ to be later determined, such that

$$\partial_{\tilde{u}}F\left(\tilde{I},\tilde{u}(\tilde{I},\epsilon),\tilde{v}(\tilde{I},\epsilon),\epsilon\right)=0=\partial_{\tilde{v}}F\left(\tilde{I},\tilde{u}(\tilde{I},\epsilon),\tilde{v}(\tilde{I},\epsilon),\epsilon\right)$$

for every $\tilde{I}\in B_{\frac{r}{2}}$ and $|\epsilon|<\epsilon_1$. To determine a possible value of ϵ_1, as well as estimate the codomain of $\tilde{u}$ and $\tilde{v}$ (which we expect to be $0(\epsilon)$), we are going to use the estimates given by a quantitative formulation of the Implicit Function Theorem (see for instance [Chi97, page 150]); setting $X_0=B_{\frac{r}{2}}\times\{|\epsilon|<\epsilon_1\}$ and $Y_0=D_\rho$ we have to choose ϵ_1 and ρ in order meet the following two inequalities

$$\sup_{X_0\times Y_0}\left|\mathbb{I}_{2p}-T(\tilde{I})\partial_{(\tilde{u},\tilde{v})}G(\tilde{I},\tilde{u},\tilde{v},\epsilon)\right|\le\frac{1}{2} \tag{3.3.2}$$

$$\sup_{X_0}\left|G(\tilde{I},0,0,\epsilon)\right|\le\frac{1}{2}\rho|T|^{-1} \tag{3.3.3}$$

where

$$\begin{aligned}T(\tilde{I}) &= \left[\partial_{(\tilde{u},\tilde{v})}G(\tilde{I},0,0,0)\right]^{-1}=\left[\partial^2_{(\tilde{u},\tilde{v})}F(\tilde{I},0,0,0)\right]^{-1}=\\ &= \text{diag}\left(\Omega_1{}^{-1}(\tilde{I}),\dots,\Omega_p{}^{-1}(\tilde{I}),\Omega_1{}^{-1}(\tilde{I}),\dots,\Omega_p{}^{-1}(\tilde{I})\right).\end{aligned}$$

We start by (3.3.3) to see how small ρ can be fixed finding a lower bound for the size of the image of the two implicit functions; since

$$G(\tilde{I},0,0,\epsilon)=\epsilon\partial_{(\tilde{u},\tilde{v})}\bar{g}(\tilde{I},0,0)$$

we have

$$\sup_{X_0}\left|G(\tilde{I},0,0,\epsilon)\right|\le\epsilon_1\sup_{\tilde{I}\in B_{\frac{r}{2}}}\left|\partial_{(\tilde{u},\tilde{v})}\bar{g}(\tilde{I},0,0)\right|\le\epsilon_1\frac{2}{r_1}|\bar{g}|_{\frac{r}{2},\frac{r_1}{2}}\le\epsilon_1 2\frac{c_p\mu^2}{\alpha d' r_1}$$

where we used Cauchy's estimate (with a loss of analycity $\frac{r_1}{2}$) and the estimate on $|\bar{g}|$ in (3.2.29). Now, to satisfy (3.3.3) we may take

$$\rho = \epsilon_1 4 \frac{c_p \mu^2}{\alpha^2 d' r_1} \geq \epsilon_1 4 \frac{c_p \mu^2}{\alpha d' r_1} \sup_{\tilde{I} \in B_{\frac{r}{2}}} |T(\tilde{I})| \tag{3.3.4}$$

since, in view of (3.2.16), we have

$$\sup_{\tilde{I} \in B_{\frac{r}{2}}} |T(\tilde{I})| = \max_{j=1,\ldots,p} \sup_{\tilde{I} \in B_{\frac{r}{2}}} \frac{1}{|\Omega_j(\tilde{I})|} = \left(\min_{j=1,\ldots,p} \inf_{\tilde{I} \in B_{\frac{r}{2}}} |\Omega_j(\tilde{I})| \right)^{-1} \leq \frac{1}{\alpha}. \tag{3.3.5}$$

This proves that ρ can be chosen of order ϵ. Let now verify under which conditions on ϵ_1 inequality (3.3.2) can be satisfied:

$$\begin{aligned}
& \sup_{X_0 \times Y_0} \left| \mathbb{I}_{2p} - T(\tilde{I}) \partial_{(\tilde{u},\tilde{v})} G(\tilde{I}, \tilde{u}, \tilde{v}, \epsilon) \right| = \\
= & \sup_{X_0 \times Y_0} \left| \mathbb{I}_{2p} - T(\tilde{I}) \partial^2_{(\tilde{u},\tilde{v})} \Big(f_0(\tilde{I}, \tilde{u}, \tilde{v}) + \epsilon \bar{g}(\tilde{I}, \tilde{u}, \tilde{v}) \Big) \right| = \\
= & \sup_{X_0 \times Y_0} \left| \mathbb{I}_{2p} - T(\tilde{I}) \left[T^{-1}(\tilde{I}) + \partial^2_{(\tilde{u},\tilde{v})} f_2(\tilde{I}, \tilde{u}, \tilde{v}) + \epsilon \partial^2_{(\tilde{u},\tilde{v})} \bar{g}(\tilde{I}, \tilde{u}, \tilde{v}) \right] \right| \leq \\
\leq & \frac{1}{\alpha} \left[\sup_{X_0 \times Y_0} \left| \partial^2_{(\tilde{u},\tilde{v})} f_2(\tilde{I}, \tilde{u}, \tilde{v}) \right| + \epsilon_1 \sup_{X_0 \times Y_0} \left| \partial^2_{(\tilde{u},\tilde{v})} \bar{g}(\tilde{I}, \tilde{u}, \tilde{v}) \right| \right]
\end{aligned}$$

having used the particular form of f_0 in (3.2.27). Then it is sufficient to require that both members in the last expression are smaller than $\frac{1}{4}$. By (3.2.28), the usual Cauchy's estimate and assuming $2\rho \leq \frac{r_1}{2}$, we obtain

$$\begin{aligned}
& \frac{1}{\alpha} \sup_{X_0 \times Y_0} \left| \partial^2_{(\tilde{u},\tilde{v})} f_2(\tilde{I}, \tilde{u}, \tilde{v}) \right| \leq \frac{2}{\alpha \rho^2} \sup_{B_{\frac{r}{2}} \times D_{2\rho}} \left| f_2(\tilde{I}, \tilde{u}, \tilde{v}) \right| \\
\leq & \frac{c_0}{\alpha \rho^2} \sup_{D_{2\rho}} |(\tilde{u}, \tilde{v})|^3 = 16 \frac{c_0 \rho}{\alpha} = 2^6 c_p c_0 \frac{\mu^2}{\alpha^3 d' r_1} \epsilon_1
\end{aligned}$$

and this is bounded by $\frac{1}{4}$ if we impose

$$\epsilon_1 \leq a_1 \frac{\alpha^3 d' r_1}{c_0 \mu^2} \qquad \text{with} \qquad a_1 := \left(2^8 c_p\right)^{-1}. \tag{3.3.6}$$

Now if we assume

$$\epsilon_1 \leq a_2 \frac{\alpha^2 d' r_1^2}{\mu^2} \qquad \text{with} \qquad a_2 := \left(2^7 c_p\right)^{-1} \tag{3.3.7}$$

with the definition of ρ in (3.3.4) we obtain $2\rho \leq \frac{r_1}{2}$ and applying Cauchy's estimate once again and (3.2.29) we have

$$\frac{\epsilon_1}{\alpha} \sup_{X_0 \times Y_0} \left| \partial^2_{(\tilde{u},\tilde{v})} \bar{g}(\tilde{I}, \tilde{u}, \tilde{v}) \right| \leq \frac{\epsilon_1}{\alpha} \sup_{B_{\frac{r}{2}} \times D_{\frac{r_1}{4}}} \left| \partial^2_{(\tilde{u},\tilde{v})} \bar{g}(\tilde{I}, \tilde{u}, \tilde{v}) \right| \leq$$

$$\leq \quad \frac{2^5 \epsilon_1}{\alpha r_1^2} |\bar{g}|_{\frac{r}{2}, \frac{r_1}{2}} \leq \epsilon_1 \frac{2^5 c_p \mu^2}{\alpha^2 d' r_1^2} \leq \frac{1}{4} \, .$$

For clearness and further references we display here what has just been proved:

Lemma 3.3.1. *Let $\epsilon \leq \epsilon_1$ with*

$$\epsilon_1 := \left(2^8 c_p\right)^{-1} \frac{\alpha^2 d' r_1^2}{\mu^2} \min\left\{ \frac{\alpha}{c_0 r_1}, 1 \right\} \tag{3.3.8}$$

(that is both conditions (3.3.6) and (3.3.7)) and define

$$\rho := \frac{4 c_p \mu^2}{\alpha^2 d' r_1^2} \epsilon \tag{3.3.9}$$

(with α as in (3.2.15), d' , c_p and μ defined respectively in (3.2.24), (3.2.23) and (3.2.21)). Then there exist two real-analytic functions

$$\tilde{u}, \tilde{v} : (\tilde{I}, \epsilon) \in B_{\frac{r}{2}} \times \{|\epsilon| < \epsilon_1\} \longrightarrow \tilde{u}(\tilde{I}, \epsilon), \tilde{v}(\tilde{I}, \epsilon) \in D_\rho \tag{3.3.10}$$

such that

$$\partial_{\tilde{u}} F \left(\tilde{I}, \tilde{u}(\tilde{I}, \epsilon), \tilde{v}(\tilde{I}, \epsilon), \epsilon \right) = 0 = \partial_{\tilde{v}} F \left(\tilde{I}, \tilde{u}(\tilde{I}, \epsilon), \tilde{v}(\tilde{I}, \epsilon), \epsilon \right) \, . \tag{3.3.11}$$

Consider now the real-analytic symplectic transformation Φ^2_ϵ generated by

$$y \cdot \tilde{\varphi} + (p + \tilde{u}(y, \epsilon)) \cdot (\tilde{v} - \tilde{v}(y, \epsilon)) \, .$$

We claim that Φ^2_ϵ is the desired transformation, that maps the $f_0 + \epsilon \bar{g}$ into an Hamiltonian which possesses an elliptic equilibrium in $(p, q) = 0$. The following lemma shows how the domain of Φ_ϵ can be well controlled for small enough ϵ.

Lemma 3.3.2. *Let $\epsilon \leq \min\{\epsilon_2, \epsilon_1\}$, with*

$$\epsilon_2 := \left(2^{10} c_p\right)^{-1} \frac{\alpha^2 d'^2}{\mu^2} \tag{3.3.12}$$

and ϵ_1 in (3.3.8), then we have

$$\Phi^2_\epsilon : (x, y, p, q) \in \mathcal{D}_2 := \mathbb{T}^d_{\frac{\sigma}{7}} \times B_{\frac{r}{4}} \times D_{\frac{r_1}{4}} \longrightarrow (\tilde{\varphi}, \tilde{I}, \tilde{u}, \tilde{v}) \in \mathcal{D}_1$$

where

$$\mathcal{D}_1 := \mathbb{T}^d_{\frac{\sigma}{6}} \times B_{\frac{r}{2}} \times D_{\frac{r_1}{2}} \, . \tag{3.3.13}$$

Proof As it can be seen by the definition of its generating function, Φ^2_ϵ is given by:

$$\begin{aligned} \tilde{I} &= y \\ \tilde{\varphi} &= x + \partial_{\tilde{I}}\tilde{v}(y,\epsilon)(p+\tilde{u}(y,\epsilon)) - \partial_{\tilde{I}}\tilde{u}(y,\epsilon)q \\ \tilde{u} &= p + \tilde{u}(y,\epsilon) \\ \tilde{v} &= q + \tilde{v}(y,\epsilon)\,. \end{aligned}$$

Now, since $(\tilde{u}(y,\epsilon),\tilde{v}(y,\epsilon))$ belong to D_ρ with $\rho \le \frac{r_1}{4}$, having assumed $\epsilon \le \epsilon_1$ (see the claim after (3.3.6)), it results

$$\begin{aligned} |\tilde{u}(y,p;\epsilon)| &\le |p| + |\tilde{u}(y,\epsilon)| \le \frac{r_1}{4} + \rho \le \frac{r_1}{2} \\ |\tilde{v}(y,p;\epsilon)| &\le |q| + |\tilde{v}(y,\epsilon)| \le \frac{r_1}{4} + \rho \le \frac{r_1}{2}\,. \end{aligned}$$

The estimate for the domain of the angles φ runs as follows :

$$|\tilde{\varphi}(x,y,p,q;\epsilon)| \le |x| + |\partial_{\tilde{I}}\tilde{v}(y,\epsilon)|\,(|p| + |\tilde{u}(y,\epsilon)|) + |\partial_{\tilde{I}}\tilde{u}(y,\epsilon)||q| \le$$

$$\begin{aligned} &\le \frac{\sigma}{7} + |\partial_{\tilde{I}}\tilde{v}(y,\epsilon)|\left(\frac{r_1}{4}+\rho\right) + |\partial_{\tilde{I}}\tilde{u}(y,\epsilon)|\frac{r_1}{4} \le \\ &\le \frac{\sigma}{7} + (|\partial_{\tilde{I}}\tilde{v}(y,\epsilon)| + |\partial_{\tilde{I}}\tilde{u}(y,\epsilon)|)\,\frac{r_1}{2} \end{aligned}$$

having used once again $\rho \le \frac{r_1}{4}$; then by Cauchy's estimate we obtain

$$\sup_{y\in B_{\frac{r}{4}},|\epsilon|\le\epsilon_2} |\partial_{\tilde{I}}\tilde{u}(y,\epsilon)| \le \sup_{y\in B_{\frac{r}{2}},|\epsilon|\le\epsilon_2} \frac{4}{r}|\tilde{u}(y,\epsilon)| \le \frac{4}{r}\sup_{|\epsilon|<\epsilon_2}\rho(\epsilon) \tag{3.3.14}$$

where the same estimates holds for $\tilde{v}(y;\epsilon)$; therefore, recalling the definition of $\rho = \rho(\epsilon)$ in (3.3.9) and using (3.3.12) we have

$$|\tilde{\varphi}(x,y,p,q;\epsilon)| \le \frac{\sigma}{7} + \frac{4r_1}{r}\rho(\epsilon) = \frac{\sigma}{7} + \frac{2^4 c_p \mu^2}{\alpha^2 d' r}\epsilon_2 \le$$

$$\le \frac{\sigma}{7} + 2^{-6}\frac{d'}{r} \le \frac{\sigma}{7} + \frac{\sigma}{42} = \frac{\sigma}{6}$$

having used $d' \le r\sigma$ $\square$

In view of this lemma and of the particular choice of Φ^2_ϵ the new Hamiltonian $H^2_\epsilon = H^1_\epsilon \circ \Phi^2_\epsilon$ has the form

$$H^2_\epsilon(x,y,p,q) = h(y) + \epsilon \hat{g}(y,p,q) + \epsilon^{N_1} \hat{f}(x,y,p,q) \tag{3.3.15}$$

where $\hat{g} = (f_0 + \epsilon \bar{g}) \circ \Phi^2_\epsilon$ and $\hat{f} = \tilde{f} \circ \Phi^2_\epsilon$ are real-analytic functions for

$$(x,y,p,q) \in \mathbb{T}^d_{\frac{\sigma}{7}} \times B_{\frac{r}{4}} \times D_{\frac{r_1}{4}} = \mathcal{D}_2 \,. \tag{3.3.16}$$

Moreover, denoting as done before $F = f_0 + \epsilon \bar{g}$, H^2_ϵ satisfies, in view of (3.3.11),

$$\partial_p \hat{g}(y,0,0) = \partial_{\tilde{u}} F \left(\tilde{I}, \tilde{u}(\tilde{I},\epsilon), \tilde{v}(\tilde{I},\epsilon), \epsilon\right) = 0$$

$$\partial_q \hat{g}(y,0,0) = \partial_{\tilde{v}} F \left(\tilde{I}, \tilde{u}(\tilde{I},\epsilon), \tilde{v}(\tilde{I},\epsilon), \epsilon\right) = 0 \,.$$

Therefore we can write

$$\hat{g}(y,p,q) = \hat{g}_0(y) + \frac{1}{2}\langle (p,q), \hat{A}(y)(p,q)\rangle + \hat{g}_3(y,p,q) \tag{3.3.17}$$

where

$$\hat{g}_0(y) = \hat{g}(y,0,0)$$

$$\hat{A}(y) = \partial^2_{(p,q)} \hat{g}(y,0,0)$$

$$\hat{g}_3(y,p,q) = \sum_{k=3}^{\infty} \frac{\partial^k_{(p,q)} \hat{g}(y,0,0)}{k!} \otimes (p,q)^k \,. \tag{3.3.18}$$

Furthermore by (3.2.29) we can estimate the norm of $\hat{g}$ with

$$|\hat{g}(y,p,q)|_{\mathcal{D}_2} \leq |f_0(\tilde{I},\tilde{u},\tilde{v}) + \epsilon \bar{g}(\tilde{I},\tilde{u},\tilde{v})|_{\mathcal{D}_1} \leq \mu + \frac{c_p}{\alpha d'}\mu^2 := \hat{M}_1 \tag{3.3.19}$$

so that we have

$$\begin{aligned} |\hat{g}_0(y)|_{\frac{r}{4}} &\leq \hat{M}_1 \\ |\hat{A}(y)|_{\frac{r}{4}} &\leq \frac{2^5}{r_1^2} |\hat{g}(y,p,q)|_{\mathcal{D}_2} \leq \frac{2^5}{r_1^2} \hat{M}_1 \,. \end{aligned}$$

By definition of $\hat{g}_3$ in (3.3.18) we have for (p,q) in $D_{\frac{r_1}{5}}$

$$
\begin{aligned}
\sup_{y\in B_{\frac{r}{4}}} |\hat{g}_3(y,p,q)| &\leq \sum_{k=3}^{\infty} \frac{|\partial^k_{(p,q)}\hat{g}(y,0,0)|_{\frac{r}{4}}}{k!} |(p,q)|^k \leq \\
&\leq \sum_{k=3}^{\infty} \left(\frac{4}{r_1}\right)^k |\hat{g}(y,p,q)|_{\mathcal{D}_2} |(p,q)|^{k-3} |(p,q)|^3 \leq \\
&\leq \frac{5^3}{r_1^3}\hat{M}_1 \sum_{k=3}^{\infty} \left(\frac{4}{5}\right)^k |(p,q)|^3 \leq \frac{5^3}{r_1^3} 3\hat{M}_1 |(p,q)|^3 . \qquad (3.3.20)
\end{aligned}
$$

From equation (3.3.17) we obtain

$$
|\hat{g}_3(y,p,q)|_{\mathcal{D}_2} \leq |\hat{g}(y,p,q)|_{\mathcal{D}_2} + |\hat{g}_0(y)|_{\frac{r}{4}} + \frac{r_1^2}{2^5} |\hat{A}(y)|_{\frac{r}{4}} \leq 3\hat{M}_1 . \qquad (3.3.21)
$$

We now focus on the $2p \times 2p$ matrix function $\hat{A}$ real-analytic on $B_{\frac{r}{4}}$ to show that it admits a diagonal form $O(\epsilon)$ close to $\operatorname{diag}(\Omega,\Omega)$ where $\Omega = (\Omega_1,\ldots,\Omega_p)$. By definition of $\hat{A}$, $\hat{g}$, $F = f_0 + \epsilon\bar{g}$ and Φ^2_ϵ it results

$$
\begin{aligned}
\hat{A}(y) &= \partial^2_{(p,q)}\hat{g}(y,0,0) = \partial^2_{(p,q)}\Big|_{(p,q)=0} F\left(\Phi_\epsilon(x,y,p,q)\right) = \\
&= \partial^2_{(p,q)}\Big|_{(p,q)=0} F\left(y, p+\tilde{u}(y,\epsilon), q+\tilde{v}(y,\epsilon)\right) = \\
&= \partial^2_{(\tilde{u},\tilde{v})}\left[f_0(y,\tilde{u}(y,\epsilon),\tilde{v}(y,\epsilon)) + \epsilon\bar{g}(y,\tilde{u}(y,\epsilon),\tilde{v}(y,\epsilon))\right] = \\
&= \operatorname{diag}(\Omega(y),\Omega(y)) + \partial^2_{(\tilde{u},\tilde{v})} f_2(y,\tilde{u}(y,\epsilon),\tilde{v}(y,\epsilon)) + \\
&+ \epsilon\partial^2_{(\tilde{u},\tilde{v})}\bar{g}(y,\tilde{u}(y,\epsilon),\tilde{v}(y,\epsilon)) .
\end{aligned}
$$

As already shown during the estimates needed to prove the existence of $\tilde{u}(y,\epsilon)$ and $\tilde{v}(y,\epsilon)$ and to find them suitable domains, the two following estimates are true

$$
\sup_{y\in B_{\frac{r}{2}}} \left|\partial^2_{(\tilde{u},\tilde{v})} f_2(y,\tilde{u}(y,\epsilon),\tilde{v}(y,\epsilon))\right| \leq 2^6 c_p \frac{c_0\mu^2}{\alpha^2 d' r_1}\epsilon
$$

$$
\epsilon \sup_{y\in B_{\frac{r}{2}}} \left|\partial^2_{(\tilde{u},\tilde{v})} \bar{g}(y,\tilde{u}(y,\epsilon),\tilde{v}(y,\epsilon))\right| \leq 2^5 c_p \frac{\mu^2}{\alpha d' r_1^2}\epsilon
$$

having used condition (3.3.6). Then, if we define the following $2p \times 2p$ matrices

$$\begin{aligned} A(y) &:= \hat{A}(y) J_{2p} \\ B(y) &:= \text{diag}\,(\Omega(y), \Omega(y)) J_{2p} \\ C(y) &:= \partial^2_{(\tilde{u},\tilde{v})}(f_0(y,0,0) + \epsilon \bar{g}(y,0,0)) J_{2p} \end{aligned} \tag{3.3.22}$$

it results

$$A(y) = B(y) + \epsilon C(y) \tag{3.3.23}$$

with

$$\begin{aligned} |A|_{\frac{r}{4}} &\leq |\hat{A}|_{\frac{r}{4}} \leq \frac{2^5}{r_1^2} \hat{M}_1 \\ |B|_{\frac{r}{4}} &\leq \max_{j=1,\dots,p} \sup_{y \in B_r} |\Omega_j(y)| := M_2 \\ |C|_{\frac{r}{4}} &\leq 2^7 c_p \frac{\mu^2}{\alpha d' r_1^2} \max\left\{ \frac{c_0 r_1}{\alpha}, 1 \right\} := M_3\,. \end{aligned} \tag{3.3.24}$$

Here is a preliminary lemma which will be useful also later:

Lemma 3.3.3. *Let $B, C : y \in D \subset \mathbb{R}^d \longrightarrow \mathbb{R}^{n \times n}$ be two matrix functions, $0 < \epsilon < 1$ a real parameter and assume $|\det B(y)| \geq m_a$ for every $y \in D$. Then if*

$$\epsilon \leq \frac{m_a}{n! 2^n n} \max\{\|B\|, \|C\|\}^{-n}\,, \tag{3.3.25}$$

where here we denote $\|M\| := \sup_{y \in D} \max_{i,j=1,\dots,n} M_{ij}(y)$, it results

$$|\det(B(y) + \epsilon C(y))| \geq \frac{m_a}{2}$$

for every $y \in D$.

Proof Denoting by S_n the group of permutation on n elements, we have

$$\det(B + \epsilon C) = \sum_{p \in S_n} \left(b_{1p(1)} + \epsilon c_{1p(1)} \right) \cdots \left(b_{np(n)} + \epsilon c_{np(n)} \right) =$$

$$\begin{aligned} &= \sum_{p \in S_n} b_{1p(1)} \cdots c_{np(n)} + \sum_{k=1}^{n} \sum_{p \in S_n} \left((1 - \delta_{1k}) b_{1p(1)} + \epsilon c_{1p(1)} \right) \cdots \\ &\cdots \left((1 - \delta_{nk}) b_{np(n)} + \epsilon c_{np(n)} \right) = \det B + d(\epsilon)\,. \end{aligned} \tag{3.3.26}$$

Calling $d_j(\epsilon)$ the j-th term in the sum that constitutes $d(\epsilon)$ and using $\epsilon < 1$,we obtain for every $j = 1, \ldots, n$

$$|d_j(\epsilon)| \leq \sum_{p \in S_n} \left(|b_{1p(1)}| + |c_{1p(1)}|\right) |\epsilon c_{jp(j)}| \left(|b_{np(n)}| + |c_{np(n)}|\right) \leq$$

$$\leq \quad \epsilon \|C\| \sum_{p \in S_n} (\|B\| + \|C\|)^{n-1} \leq \epsilon n! 2^{n-1} (\max\{\|B\|, \|C\|\})^n .$$

Therefore

$$|d(\epsilon)| \leq n|d_j(\epsilon)| = \epsilon n! 2^{n-1} n (\max\{\|B\|, \|C\|\})^n \tag{3.3.27}$$

so that with the hypothesis on ϵ

$$|\det(B + \epsilon C)| = |\det B + d(\epsilon)| \geq |\det B| - |d(\epsilon)| \geq \frac{m_a}{2} > 0 \quad \square$$

Observe that from equation (3.3.26) and inequality (3.3.27) we also get

$$|\det(A' + \epsilon B')| \leq |\det A'| + \epsilon n! 2^{n-1} n \max\{\|A'\|, \|B'\|\}^n \tag{3.3.28}$$

for every $n \times n$ matrices A' and B'.

We are now ready to state the following

Proposition 3.3.1. *Consider A, B, C as defined in (3.3.22) with $A = B + \epsilon C$ for $0 < \epsilon < 1$, M_1, M_2, M_3 as defined in (3.3.24) and assume*

$$\epsilon \leq \frac{\min\left\{\alpha(2M_2)^{-1}, \alpha M_3^{-1}, 1\right\}^{2p}}{((2p+1)!)^2 2^{2p}} . \tag{3.3.29}$$

Then the eigenvalues of $A(y)$ are $2p$ purely imaginary analytic functions , with

$$\sup_{y \in B_{\frac{r}{4}}} |\hat{\Omega}(y) - \Omega(y)| \leq 2^{2p} (2p)! 2p \min\{\alpha(2M_2)^{-1}, \alpha M_3^{-1}\}^{-2p} \alpha \epsilon \tag{3.3.30}$$

Proof First of all, from classical arguments about symplectic quadratic forms of Hamiltonian, we can infer that the symmetric matrix $A = \hat{A} J_{2p}$ has an elliptic equilibrium point at the origin or equivalently possesses all purely imaginary eigenvalues $\pm i\hat{\Omega}_j$ for $j = 1, \ldots, p$. What follows is the proof of estimate (3.3.30), that is $\hat{\Omega}_j$ is $O(\epsilon)$ close to Ω_j for every $j = 1, \ldots, p$. Let $\lambda_1, \ldots \lambda_{2p}$ be the eigenvalues of B, i.e.,

$$\lambda_k(y) = \begin{cases} i\Omega_k(y) & \text{if} \quad 1 \leq k \leq p \\ -i\Omega_{k-p}(y) & \text{if} \quad p+1 \leq k \leq 2p \end{cases}$$

as it can be easily seen by definition of $B = \operatorname{diag}(\Omega, \Omega) J_{2p}$. Consider the real-analytic function

$$f(y, \epsilon, \lambda) = \det\left(A(y) - \lambda \mathbb{I}_{2p}\right) = \det\left(B(y) + \epsilon C(y) - \lambda \mathbb{I}_{2p}\right) .$$

Observe that $f(y, 0, \lambda_k(y)) = \det(B(y) - \lambda_k(y)\mathbb{I}_{2p}) = 0$ and

$$\frac{\partial}{\partial\lambda} f(y, 0, \lambda_k(y)) = \frac{\partial}{\partial\lambda}\Big|_{\lambda=\lambda_k} \prod_{j=1}^{2p} (\lambda(y) - \lambda_j(y)) =$$

$$= \prod_{j=1, j\neq k}^{2p} (\lambda_k(y) - \lambda_j(y)) \neq 0$$

for every $j = 1, \ldots, p$ and $y \in B_{\frac{r}{4}}$ in view of (3.2.14). Therefore for every fixed k, it is possible to apply the Implicit Function Theorem in order to find two positive numbers ϵ_0, ρ and a real-analytic function

$$\hat{\lambda}_k : (y, \epsilon) \in B_{\frac{r}{4}} \times \{|\epsilon| < \epsilon_0\} := X_0 \longrightarrow \{|\lambda - \lambda_k| < \rho\} := Y_0$$

such that $f(y, \epsilon, \hat{\lambda}_k(y, \epsilon)) = 0$ for every $(y, \epsilon) \in X_0$. We now proceed as already done previously to determine possible values of ϵ_0 and ρ with the aim to prove that ρ can be taken of order ϵ obtain the statement. Let

$$T(y) = \left[\frac{\partial}{\partial\lambda} f(y, 0, \lambda_k(y))\right]^{-1} = \prod_{j=1, j\neq k}^{2p} (\lambda_k(y) - \lambda_j(y))^{-1} \qquad (3.3.31)$$

in view of (3.2.14) it results $|\lambda_k(y) - \lambda_j(y)| \geq \alpha > 0$ for every $j \neq k$ and $y \in B_{\frac{r}{4}}$ which leads to

$$\sup_{y\in B_{\frac{r}{4}}} |T(y)| \leq \alpha^{1-2p} .$$

Furthermore the definition of f and inequality (3.3.28) applied with $A' = B - \lambda_k \mathbb{I}_{2p}$ and $B' = C$ give

$$\begin{aligned} |f(y, \epsilon, \lambda_k)| &\leq |\det(B - \lambda_k \mathbb{I}_{2p})| + \epsilon(2p)!2p\, 2^{2p-1} \max\{\|B - \lambda_k \mathbb{I}_{2p}\|, \|C\|\}^{2p} \\ &\leq \epsilon(2p)!2p\, 2^{2p-1} \max\{2M_2, M_3\}^{2p} \end{aligned}$$

for every $(\epsilon, y) \in X_0$, since

$$M_2 \geq \max_{j=1,\ldots,2p} \sup_{y\in B_{\frac{r}{4}}} |\lambda_k(y)| .$$

Therefore the inequality $\sup_{X_0} |f(y,\epsilon,\lambda_k)| \leq \frac{1}{2}\rho|T|^{-1}$ can be satisfied by taking

$$\rho := (2p)!2^{2p}2p\max\{2M_2, M_3\}^{2p}\alpha^{1-2p}\epsilon$$

which proves, for sufficiently small ϵ_0 to be next determined, the estimate in (3.3.30).

The second inequality to met is

$$\sup_{X_0\times Y_0}\left|1 - T(y)\frac{\partial}{\partial\lambda}f(y,\epsilon,\lambda)\right| \leq \frac{1}{2}\,. \tag{3.3.32}$$

we start observing that by the formula in (3.3.26) we have $f(y,\epsilon,\lambda) = \det(B - \lambda\mathbb{I}_{2p}) + g(\epsilon, b_{ij}(y), c_{ij}(y))$, therefore

$$\frac{\partial}{\partial\lambda}f(y,\epsilon,\lambda) = \frac{\partial}{\partial\lambda}\det(B-\lambda\mathbb{I}_{2p}) = \frac{\partial}{\partial\lambda}\prod_{j=1}^{2p}(\lambda-\lambda_j) = \sum_{i=1}^{2p}\prod_{j\neq i}(\lambda-\lambda_j)$$

and it follows

$$\begin{aligned}
\left|1 - T(y)\frac{\partial}{\partial\lambda}f(y,\epsilon,\lambda)\right| &= \left|1 - \prod_{j\neq k}(\lambda_k(y)-\lambda_j(y))^{-1}\sum_{i=1}^{2p}\prod_{j\neq i}(\lambda-\lambda_j)\right| = \\
= &\left|1 - \prod_{j\neq k}\left(\frac{\lambda-\lambda_j}{\lambda_k-\lambda_j}\right) + \sum_{i\neq k}\prod_{j\neq k}(\lambda_k-\lambda_j)^{-1}\prod_{j\neq i}(\lambda-\lambda_j)\right| = \\
= &\left|1 - \prod_{j\neq k}\left(\frac{\lambda-\lambda_j}{\lambda_k-\lambda_j}\right) + \sum_{i\neq k}\prod_{j\neq k,i}\left(\frac{\lambda-\lambda_j}{\lambda_k-\lambda_j}\right)\left(\frac{\lambda-\lambda_k}{\lambda_k-\lambda_i}\right)\right|\,.
\end{aligned} \tag{3.3.33}$$

Define now

$$a_j(y) = \frac{\lambda(y)-\lambda_k(y)}{\lambda_k(y)-\lambda_j(y)}$$

and observe that from (3.2.14), definition of Y_0, definition of ρ and condition (3.3.29) it results

$$|a_j(y)| \leq \frac{\rho}{\alpha} = (2p)!(2p)2^{2p}\max\{2M_2, M_3\}^{2p}\alpha^{-2p}\epsilon \leq 1$$

for any $y \in B_{\frac{r}{4}}$; then

$$\prod_{j\neq k}\left(\frac{\lambda-\lambda_j}{\lambda_k-\lambda_j}\right) = \prod_{j\neq k}\left(\frac{\lambda-\lambda_k+\lambda_k-\lambda_j}{\lambda_k-\lambda_j}\right) = \prod_{j\neq k}(1+a_j) = 1+\delta(a_j)$$

with

$$|\delta(a_j)| \le ((2p-1)!-1)|a_j| \le (2p-1)!\frac{\rho}{\alpha}$$

and the following estimate holds for every $i \neq k$

$$\left|\prod_{j\neq k,i}\left(\frac{\lambda-\lambda_j}{\lambda_k-\lambda_j}\right)\right| = \left|\prod_{j\neq k,i}(1+a_j)\right| \le \prod_{j\neq k,i}(1+|a_j|)$$

$$\le 1+(2p-2)!|a_j| \le (2p-1)!.$$

We now come back to (3.3.33) and obtain

$$\left|1-\prod_{j\neq k}\left(\frac{\lambda-\lambda_j}{\lambda_k-\lambda_j}\right)+\sum_{i\neq k}\prod_{j\neq k,i}\left(\frac{\lambda-\lambda_j}{\lambda_k-\lambda_j}\right)\left(\frac{\lambda-\lambda_k}{\lambda_k-\lambda_i}\right)\right| \le$$

$$\le |\delta(a_j)| + \sum_{i\neq k}\prod_{j\neq k,i}\left|\frac{\lambda-\lambda_j}{\lambda_k-\lambda_j}\right|\left|\frac{\lambda-\lambda_k}{\lambda_k-\lambda_i}\right| \le$$

$$\le (2p-1)!\frac{\rho}{\alpha}+(2p-1)!\sum_{i\neq k}\left|\frac{\lambda-\lambda_k}{\lambda_k-\lambda_i}\right| \le$$

$$\le (2p-1)!\frac{\rho}{\alpha}+(2p-1)!(2p-1)\frac{\rho}{\alpha} = (2p)!\frac{\rho}{\alpha}$$

so that estimate (3.3.32) holds by means of the definition of ρ and hypothesis (3.3.29) □

This easy corollary runs as a consequence:

Corollary 3.3.1. *The eigenvalues of $\hat{A}J_{2p} = B + \epsilon C$ (see (3.3.22)) verify*

$$\inf_{y\in B}\hat{\Omega}_j(y) \ge \inf_{y\in B_{\frac{r}{4}}}|\hat{\Omega}_j(y)| \ge \frac{\alpha}{2} > 0\,, \tag{3.3.34}$$

$$\min_{j_1\neq j_2}\inf_{y\in B_{\frac{r}{4}}}|\hat{\Omega}_{j_1}(y)-\hat{\Omega}_{j_2}(y)| \ge \frac{\alpha}{2} > 0\,; \tag{3.3.35}$$

under hypothesis (3.3.29) on ϵ.

Proof The first inequality is given by

$$\inf_{y\in B_{\frac{r}{4}}}|\hat{\Omega}_j(y)| \ge \inf_{y\in B_{\frac{r}{4}}}|\Omega_j(y)| - \sup_{y\in B_{\frac{r}{4}}}|\hat{\Omega}_j(y)-\Omega_j(y)| \ge$$

$$\ge \alpha-(2p)!2^{2p}2p\min\{\alpha(2M_2)^{-1},\alpha M_3^{-1}\}^{-2p}\alpha\epsilon \ge \alpha-\frac{\alpha}{2}=\frac{\alpha}{2}$$

having used (3.2.14), (3.3.29) and (3.3.30). For what concerns the second property, let $j_1 \neq j_2$ then, using the same inequalities we just referred to, it results

$$\inf_{y \in B_{\frac{r}{4}}} |\hat{\Omega}_{j_1}(y) - \hat{\Omega}_{j_2}(y)| \geq \inf_{y \in B_{\frac{r}{4}}} \Big(|\Omega_{j_1}(y) - \Omega_{j_2}(y)| - |\hat{\Omega}_{j_1}(y) - \Omega_{j_1}(y)| +$$

$$- \quad |\hat{\Omega}_{j_2}(y) - \Omega_{j_2}(y)| \Big) \geq \alpha - 2 \max_{j=1,\dots,2p} \sup_{y \in B_{\frac{r}{4}}} |\hat{\Omega}_j(y) - \Omega_j(y)| \geq$$

$$\geq \quad \alpha - (2p)! 2^{2p+1} 2p \min\{\alpha (2M_2)^{-1}, \alpha M_3^{-1}\}^{2p} \alpha\epsilon \geq \alpha - \frac{\alpha}{2} = \frac{\alpha}{2} \qquad \square$$

Now, as already observed several times before, we may consider the quadratic form associated to the matrix $\hat{A}$ to be definite positive (up to establishing every result for $-\hat{A}$ and then change the sign). Therefore, we can now perform the symplectic diagonalization of the matrix $\hat{A}$, i.e., the quadratic part of Hamiltonian $\hat{g}$ in (3.3.17), through the following well-known classical result by K. Weierstraß

Proposition 3.3.2 (Weierstraß Diagonalization). *Let $\hat{g}_2$ be a real-analytic function on $\mathcal{D}_2$ (see (3.3.16)) in the form*

$$\hat{g}_2(y, p, q) = \frac{1}{2} \langle (p, q), \hat{A}(y)(p, q) \rangle$$

where $\hat{A}(y)$ is a $2p \times 2p$ symmetric and positive definite matrix for every $y \in B_{\frac{r}{4}}$, then there exists a real-analytic and linear symplectic transformation

$$\Phi^3_\epsilon : (\tilde{x}, \tilde{y}, \tilde{p}, \tilde{q}) \in \mathbb{T}^d_{\frac{\sigma}{8}} \times B_{\frac{r}{5}} \times D_{\frac{r_1}{5}} := \mathcal{D}_3 \longrightarrow (x, y, p, q) \in \mathcal{D}_2 \tag{3.3.36}$$

such that $\tilde{y} = y$ and

$$\hat{g}_2 \circ \Phi^3_\epsilon = \tilde{g}_2(\tilde{y}, \tilde{p}, \tilde{q}) = \frac{1}{2} \sum_{j=1}^{p} \hat{\Omega}_j(\tilde{y}) \left(\tilde{p}_j^2 + \tilde{q}_j^2 \right) \tag{3.3.37}$$

where $\hat{\Omega}_1, \dots, \hat{\Omega}_p$ are uniquely determined by $\hat{g}_2$; indeed $\pm i\hat{\Omega}_1, \dots, \pm i\hat{\Omega}_p$ are the $2p$ eigenvalues of the matrix $\hat{A} J_{2p}$.

With this results we obtain that the Hamiltonian function H^2_ϵ in (3.3.15) can be transformed into an Hamiltonian

$$H^3_\epsilon(\tilde{x}, \tilde{y}, \tilde{p}, \tilde{q}) = H^2_\epsilon \circ \Phi^3_\epsilon(\tilde{x}, \tilde{y}, \tilde{p}, \tilde{q}) = h(\tilde{y}) + \epsilon \hat{g}_0(\tilde{y}) +$$

$$+ \quad \frac{\epsilon}{2} \sum_{j=1}^{p} \hat{\Omega}_j(\tilde{y}) \left(\tilde{p}_j^2 + \tilde{q}_j^2 \right) + \epsilon \tilde{g}_3(\tilde{y}, \tilde{p}, \tilde{q}) + \epsilon^{N_1} \tilde{f}_3(\tilde{x}, \tilde{y}, \tilde{p}, \tilde{q}) \tag{3.3.38}$$

where, in view of (3.3.20) and (3.3.21), $\tilde{g}_3 = \hat{g}_3 \circ \Phi^3_\epsilon$ verifies

$$|\tilde{g}_3(\tilde{y}, \tilde{p}, \tilde{q})|_{\mathcal{D}_3} \leq 3\hat{M}_1 \tag{3.3.39}$$

$$\sup_{\tilde{y}\in B_{\frac{r}{5}}} |\tilde{g}_3(\tilde{y}, \tilde{p}, \tilde{q})| \leq \frac{5^3}{r_1^3} 3\hat{M}_1 |(\tilde{p}, \tilde{q})|^3 \quad \forall\ (\tilde{p}, \tilde{q}) \in D_{\frac{r_1}{5}} \tag{3.3.40}$$

(see (3.3.19) for $\hat{M}_1$) and $\tilde{f}_3 = \bar{f} \circ \Phi^3_\epsilon$ is bounded by

$$|\tilde{f}_3|_{\mathcal{D}_3} \leq \mu^{N_1}\,. \tag{3.3.41}$$

3.4 Birkhoff's normal form

Now, we want now to put the real-analytic Hamiltonian function $\tilde{g}_3$ appearing (3.3.38) into Birkhoff's normal form up to any chosen order $N_2 \geq 3$. More precisely we aim to find a real-analytic transformation Φ^4_ϵ such that $\tilde{g}_3 \circ \Phi^4_\epsilon$ is an even polynomial of degree $\left[\frac{N_2}{2}\right]$ in the new elliptic variables plus a remainder of order greater than N_2. Here is a preliminary lemma that proves how the first Birkhoff's invariant of $\tilde{g}_3$ are non-resonant up to order N_2 for sufficiently small ϵ:

Lemma 3.4.1. *Let $\pm\hat{\Omega}_1(\tilde{y}), \ldots, \pm\hat{\Omega}_p(\tilde{y})$ the eigenvalues of $\hat{A}(\tilde{y})J_{2p}$ for $\tilde{y} \in B_{\frac{r}{5}}$ as found in proposition 3.3.1, we define the real-analytic function*

$$\hat{\Omega}(\tilde{y}) = \left(\hat{\Omega}_1(\tilde{y}), \hat{\Omega}_2(\tilde{y}), \ldots, \hat{\Omega}_p(\tilde{y})\right)\ ; \tag{3.4.1}$$

if we assume

$$\epsilon \leq \frac{\min\left\{\alpha(2M_2)^{-1}, \alpha M_3^{-1}, 1\right\}^{2p}\alpha}{2^{2p}(2p+1)!^2 N_2} \tag{3.4.2}$$

then $\hat{\Omega}$ is non-resonant of order N_2 and in particular it results

$$|\hat{\Omega}(\tilde{y}) \cdot k| \geq \frac{\alpha}{2} \quad \forall\ k \in \mathbb{Z}^{2p}\ \textit{with}\ 0 < |k|_1 \leq N_2 \tag{3.4.3}$$

and for all $\tilde{y}$ in $B_{\frac{r}{5}}$.

Proof The proof runs as a consequence of the following chain of inequalities that hold for any k in $\mathbb{Z}^{2p}$ and $\tilde{y}$ in $B_{\frac{r}{5}}$ in view of (3.2.14), (3.3.30) and hypothesis (3.4.2):

$$|\hat{\Omega}(\tilde{y}) \cdot k| \geq |\Omega(\tilde{y}) \cdot k| - |\hat{\Omega}(\tilde{y}) - \Omega(\tilde{y}) \cdot k| \geq \frac{\alpha}{2} - |\hat{\Omega}(\tilde{y}) - \Omega(\tilde{y})||k|_2$$

$$\geq \frac{\alpha}{2} - 2p\, 2^{2p}(2p)! \min\left\{(2M_2)^{-1}\alpha, M_3^{-1}\alpha\right\}^{-2p} \alpha\epsilon|k|_1 \geq$$

$$\geq \frac{\alpha}{2} - \frac{\alpha}{(2p+1)^2 N_2}|k|_1 \geq \frac{\alpha}{2} - \frac{\alpha}{4} = \frac{\alpha}{4} \quad \square$$

We now provide a general formulation of Birkhoff's normal form theorem:

Theorem 3.4.1. *Let H be a real-analytic function on $D_{r_0} \subseteq \mathbb{C}^{2p}$ of the form*

$$H(\zeta) = \sum_{j=1}^{p} \frac{\Omega_j}{2} \left(\xi_j^2 + \eta_j^2\right) + g(\zeta) \tag{3.4.4}$$

near the elliptic equilibrium point $\zeta = 0$, where $\zeta = (\xi_1, \ldots, \xi_p, \eta_1, \ldots, \eta_p)$, with $|g(\zeta)| \leq C_3 |\zeta|^3$ for all $\zeta \in D_{r_0}^{2p}$ for some $C_3 > 0$. Let M be such that

$$\sup_{\zeta \in D_\rho} |H(\zeta)| \leq \rho^2 M \ , \quad M \geq 1 \tag{3.4.5}$$

for every $\rho \leq r_0$. If we assume that the linear invariants $\Omega = (\Omega_1, \ldots, \Omega_p)$ are non resonant of order s, that is

$$\langle \Omega, k \rangle \geq a > 0 \quad \forall \ j \in \mathbb{Z}^p \ \text{ with } \ 0 < |j|_1 \leq s \,, \tag{3.4.6}$$

for some $a \leq 1$, then there exist numbers $0 < r_s < r \leq r_0$, $C_s > 0$ and an analytic symplectic diffeomorphism

$$\tau : z = (p, q) \in D_{r_s} \longrightarrow \zeta = (\xi, \eta) \in D_r \,, \tag{3.4.7}$$

leaving the origin and the quadratic part of H invariant (i.e., $\tau(z) = z + O(|z|^2)$), such that

$$H \circ \tau = F(I_1, \ldots, I_p) + R_s(z) \tag{3.4.8}$$

where

- *F is a polynomial of degree $\left[\frac{s}{2}\right]$ in the variables $I = (I_1, \ldots, I_p)$, with*

  $$I_j = \frac{1}{2} \left(p_j^2 + q_j^2\right) \tag{3.4.9}$$

  for all $j = 1, \ldots, p$, having the form

  $$F(I) = \langle \Omega, I \rangle + \frac{1}{2} \langle TI, I \rangle + \cdots \tag{3.4.10}$$

  for some $p \times p$ matrix T;

- *$|R_s| \leq M_s |(p, q)|^{s+1}$.*

In addition the polynomial F is uniquely determined by H and does not depend on the choice of τ. From this we infer that the coefficients of F are local symplectic invariants of H usually called **Birkhoff invariants***.*

Furthermore r_s and r can be determined as follows:

$$r = \delta r_0 \qquad \textit{with} \qquad \delta := \frac{1}{2^6(2s+1)^p}\left(\frac{a}{5M}\right)^{\frac{1}{2}} \tag{3.4.11}$$

$$r_s = \left(\frac{3}{4}\delta\right)^{s-2} r_0 \tag{3.4.12}$$

Proof The proof we provide here follows [Zeh94] for what concerns the main idea and the algebraic part (existence and of τ) and consists of an additional analytic part for the determination of r and r' (the radii of the complex domains on which τ can be well defined). We proceed with an iterative scheme and assume that H is already in normal form up to order $m-1$, for some $m \geq 3$, that is

$$H = H_2 + H_3 + \cdots + H_{m-1} + H_m + \cdots$$

where H_j is an homogeneous polynomial of order j. We look for an analytic symplectic map τ, with $\tau(0) = 0$ and $d\tau(0) =$ id that puts H into normal form up to order m. Let P be an homogeneous polynomial (to be determined later) of degree $m \geq 3$, we take

$$\tau(z) = \exp tX_p|_{t=1} \tag{3.4.13}$$

as the 1-time map of the Hamiltonian vector field X_P; in other words, let $\varphi(t;z) \in \mathcal{H}([0,1])$ the solution of the Cauchy's problem

$$\begin{cases} \dot{u}(t) = J\nabla P(u(t)) \\ \\ u(0) = z \in D_{r'} \end{cases} \tag{3.4.14}$$

we have $\tau(z) = \varphi(1;z)$. As it can be immediately seen $\tau(z)$ is a symplectic transformation because it belongs to the flow of an Hamiltonian vector field, $\tau(0) = 0$ and $d\tau(0) =$ id since $\tau(z) = z + J\nabla P(z) + \cdots$. By Taylor's expansion of $H \circ t \exp X_P$ in t at $t = 0$ we obtain

$$\begin{aligned} H \circ \tau &= H \circ \exp tX_p|_{t=1} = H + \{H,P\} + \{\{H,P\},P\} \cdots = \\ \\ &= H_2 + \cdots H_{m-1} + (H_m + \{H_2,P\}) + \cdots \end{aligned} \tag{3.4.15}$$

where this last dots stand for terms of order higher than m and $\{\cdot,\cdot\}$ denotes the usual Poisson brackets. As it can be easily seen through the map τ it is possible to modify H, and in particular H_m, by terms of the form $\{H_2,P\}$ that are homogeneous polynomial of degree m.

Let $\mathcal{P}_m$ the vector space of homogeneous polynomial of degree m and let L the linear operator defined by

$$\begin{aligned} L: \ \mathcal{P}_m & \longrightarrow \mathcal{P}_m \\ P & \longmapsto \{H_2, P\}\,. \end{aligned}$$

We first infer that $K(L)$, the Kernel of L, and $R(L)$, the range of L, are supplementary, i.e., $K(L)+R(L)=\mathcal{P}_m$ and $K(L)\cap R(L)=\{0\}$; in addition, if $m\leq s$ we can described $K(L)$ by:

$$\begin{aligned} &K(L)=\{0\} && \text{for } m=2k\,, \\ &K(L)=\text{span}\{I^k=I_1^{k_1}\dots I_p^{k_p}\mid \textstyle\sum_{j=1}^p k_j=m\} && \text{for } m=2k+1\,. \end{aligned} \tag{3.4.16}$$

To prove this two sentences we start by diagonalizing L in $\mathcal{P}_m$; we make a change of symplectic coordinates going to complex variables

$$\begin{aligned} \alpha &= \frac{1}{\sqrt{2}}(\xi+i\eta) \\ \beta &= \frac{1}{\sqrt{2}}(\eta+i\xi)\,. \end{aligned} \tag{3.4.17}$$

In this new set of symplectic coordinates we have $\alpha_j\beta_j=\frac{i}{2}(\xi_j^2+\eta_j^2)$, $H_2=\frac{1}{i}\sum_{j=1}^p\Omega_j(\alpha_j\beta_j)$ and it results

$$\begin{aligned} L(\alpha^k\beta^l)=\{H_2,\alpha^k\beta^l\}=\frac{1}{i}\sum_{j=1}^p\Omega_j\{\alpha_j\beta_j,\alpha^k\beta^l\}= \\ = \frac{1}{i}\sum_{j=1}^p\Omega_j\left[\beta_j(\alpha^k l_j\beta^{l-e^{(j)}})-\alpha_j(k_j\alpha^{k-e^{(j)}}\beta^l)\right]= \\ = \frac{1}{i}\langle\Omega,l-k\rangle\alpha^k\beta^l \end{aligned} \tag{3.4.18}$$

for every $k,l\in\mathbb{Z}^p$, where $e^{(j)}\in\mathbb{R}^p$ is given by $e_k^{(j)}=\delta_{jk}$. By this last chain of equalities we obtain that $K(L)$ consists of monomials $\alpha^k\beta^l$ with $l=k$ (since Ω is non-resonant up to order $m\leq s$) that is, in terms of the old symplectic variables (ξ,η), what stated in (3.4.16) (with I as defined in (3.4.9)); this obviously proves also the complementarity of $K(L)$ and $R(L)$.

Let $F=\sum_{|k|+|l|=m}f_{kl}\alpha^k\beta^k\in\mathcal{P}_m$ then we have the decomposition $F=$

$F^K + F^R$, with $F^K \in K(L)$ and $F^R \in R(L)$ given by

$$F^R = \sum_{|k|+|l|=m,\, k\neq l} f_{kl}\alpha^k\beta^l \tag{3.4.19}$$
$$F^K = \sum_{|k|+|l|=m,\, k=l} f_{kl}\alpha^k\beta^l\,.$$

With $F = H_m$ (with refer to equation (3.4.15)) we have

$$H_m + \{H_2, P\} = H_m^K + H_m^R + \{H_2, P\}$$

therefore we just need to solve

$$H_m^R + \{H_2, P\} = 0 \tag{3.4.20}$$

since the elements of the Kernel of L have exactly the desired normal form (as stated in (3.4.16)). Using equalities in (3.4.18) we obtain

$$\begin{aligned}
\{H_2, P\} &= \{H_2, \sum_{|k|+|l|=m} p_{kl}\alpha^k\beta^l\} = \sum_{|k|+|l|=m} p_{kl}\{H_2, \alpha^k\beta^l\} = \\
&= \sum_{|k|+|l|=m} p_{kl}L(\alpha^k\beta^l) = \frac{1}{i}\sum_{|k|+|l|=m} p_{kl}\langle\Omega, l-k\rangle\alpha^k\beta^l = \\
&= \frac{1}{i}\sum_{|k|+|l|=m,\, k\neq l} p_{kl}\langle\Omega, l-k\rangle\alpha^k\beta^l\,;
\end{aligned}$$

then, if

$$H_m^R(\alpha,\beta) = \sum_{|k|+|l|=m,\, k\neq l} b_{kl}\alpha^k\beta^l \tag{3.4.21}$$

we can solve equation (3.4.20) by taking

$$P(\alpha,\beta) = \sum_{|k|+|l|=m,\, k\neq l} -\frac{ib_{kl}}{\langle\Omega, l-k\rangle}\alpha^k\beta^l\,. \tag{3.4.22}$$

We have so proved the existence of a map τ, defined by (3.4.13) and (3.4.22), which puts H_m into normal form; we shall call this map τ_m so that the wanted final map τ in (3.4.7) is

$$\tau := \tau_s \circ \tau_{s-1} \circ \cdots \circ \tau_3\,. \tag{3.4.23}$$

The proof of the uniqueness of the map τ is quite simple and can be found in [Zeh94].

We now focus our attention on one single step of the iteration process needed to obtain τ and on the analytic determination of the domain and codomain of τ_m. For easier notations we put $H^{(2)} := H$ and define $H^{(m-1)}$ as the Hamiltonian function after $m-3$ ($m \geq 3$) steps, which means

$$H^{(m-1)} := H^{(m-2)} \circ \tau_{m-1} = \cdots = H^{(2)} \circ \tau_m \circ \tau_{m-1} \circ \cdots \circ \tau_3$$

is already in normal form up to order $m-1$. Let

$$\tau_m : z = (p, q) \in D^{2p}_{r'_m}(0) \longrightarrow \zeta = (\xi, \eta) \in D^{2p}_{r_m}(0)\,; \tag{3.4.24}$$

the map that casts $H^{(m-1)}$ into $H^{(m)}$ in normal form up to order $m+1$, we aim to estimate r'_m and r_m and in particular their dependence on the coefficients of order m of $H^{(m-1)}$ and on a (as in (3.4.6)).

We have already shown before that τ_m can be obtain through the analytic solution of Cauchy's problem (3.4.14) at the time $t = 1$, where now r' and P have to be replaced respectively with r'_m and

$$P_m(\alpha, \beta) := \sum_{|k|+|l|=m,\, k\neq l} -\frac{ib^{(m-1)}_{kl}}{\langle \Omega, l-k\rangle}\alpha^k \beta^l \tag{3.4.25}$$

where (α, β) have to be considered as functions of (ξ, η) (see (3.4.17)) and the apex $m-1$ on b_{kl} indicates that $H^{(m-1)}$ is taken into consideration. Now, to obtain such a (relatively) wide time of existence for the solution $\Phi^t_P(z)$ as $t = 1$, we are forced to make some strong requirements on r_m and r'_m in order to assure that that the norm of P is sufficiently small as well as the domain of the initial data $D_{r'_m}$. We search $\Phi^t_P(z)$ (solution of (3.4.14)) as the solution of a fixed point problem. Let $E := \mathcal{H}\left(D^1_1(0), D^{2p}_{r_m}(0)\right)$ provided with $|u| := \sup_{t\in D^1_1(0)} |u(t)|$, consider the map

$$\begin{aligned} T:\quad E \quad &\longrightarrow E \\ u(t) \quad &\longmapsto z + \int_0^t J\nabla P^\star_m(u(s))\,ds \end{aligned}$$

where $P^\star_m$ is a suitable homogeneous polynomial of degree m such that

$$P^\star_m(\xi, \eta) = P_m(\alpha(\xi, \eta), \beta(\xi, \eta)) \tag{3.4.26}$$

We initially show that T is a contraction for sufficiently small r_m: let u, v be in E

and P as determined in (3.4.25)

$$|T(u) - T(v)| = \left| \int_0^t J[\nabla P_m^\star(u(s)) - \nabla P_m^\star(v(s))]\, ds \right| \leq$$

$$\leq \int_0^1 |J[\nabla P_m^\star(u(s)) - \nabla P_m^\star(v(s))]|\, ds \leq \sup_{D_1^1(0)} |\nabla P_m^\star(u(t)) - \nabla P_m^\star(v(t))| \leq$$

$$\leq \sup_{D_{r_m}^{2p}(0)} |D^2 P_m^\star||u - v|$$

where $D^2 P_m^\star$ stands for the $2p \times 2p$ matrix of the second derivatives of $P_m^\star$ and this last equality holds because of Lagrange's Theorem. In view of (3.4.26), to estimate the norm of $P_m^\star$ on $D_{2r_m}^{2p}$ it is sufficient to estimate the norm of P on $D_{2\sqrt{2}r_m}^{2p}$ (according to (3.4.17)); then we obtain

$$\begin{aligned} \sup_{D_{2r_m}} |P_m^\star(\xi, \eta)| &\leq \sup_{D_{2\sqrt{2}r_m}} |P_m(\alpha, \beta)| \leq \sum_{|k|+|l|=m,\, k\neq l} \frac{|b_{kl}^{(m-1)}|}{|\langle \Omega, l-k \rangle|} (8r_m)^m \leq \\ &\leq \frac{1}{a} \sup_{|k|+|l|=m} |b_{kl}^{(m)}| (8r_m)^m \sum_{|k|+|l|=m} 1 \leq \frac{1}{a} C_m\, r_m^m \end{aligned}$$

in view of the non-resonance condition (3.4.6) satisfied by Ω and having defined

$$C_m := 8^m (2m+1)^{2p} \sup_{|k|+|l|=m} |b_{kl}^{(m-1)}|\,. \tag{3.4.27}$$

By Cauchy's theorem it results (always with refer to the original variables $\zeta = (\xi, \eta)$)

$$\begin{aligned} |\nabla P_m^\star|_{r_m} &\leq |\nabla P_m^\star|_{\frac{3}{2}r_m} \leq \frac{2}{r} |P_m^\star|_{2r_m} \leq \frac{2}{a} C_m\, r_m^{m-1} \\ |D^2 P_m^\star|_{r_m} &\leq \frac{2}{r_m} |\nabla P_m^\star|_{\frac{3}{2}r_m} \leq \frac{4}{r_m^2} |P_m^\star|_{2r_m} \leq \frac{4}{a} C_m\, r_m^{m-2}\,; \end{aligned}$$

since we need $|D^2 P_m^\star|_{r_m} < 1$ to have a contraction it is sufficient to impose

$$r_m \leq \left(\frac{a}{5C_m} \right)^{\frac{1}{m-2}}. \tag{3.4.28}$$

Now we have to find sufficient conditions on the domain of the initial data (i.e conditions on r_m') under which $T(E)$ is contained in E; let $u \in E$, by the preceding estimates and using (3.4.28) it results

$$\begin{aligned} |T(u)| &= \left| z + \int_0^t J\nabla P_m^\star(u(s)) \right| \leq r'_m + |\nabla P_m^\star|_{r_m} \leq r'_m + \frac{|P_m^\star|_{2r_m}}{r_m} \\ &\leq r'_m + \frac{1}{a} C_{m,p}\, r_m^{m-1} \leq r'_m + \frac{r_m}{4} \end{aligned}$$

and therefore it is sufficient to take

$$r'_m \leq \frac{3}{4} r_m\,. \tag{3.4.29}$$

To conclude we need to show that the choices of r and r_s in (3.4.11) and (3.4.12) are possible. From what proved before, every single map τ_m for $m = 3, \ldots, s$ can be defined from $D^{2p}_{\rho_{m+1}}$ and have its image contained in $D^{2p}_{\tilde\rho_m}$ provided this radii verify

$$\tilde\rho_m \leq \left(\frac{a}{5C_m}\right)^{\frac{1}{m-2}} \qquad \text{and} \qquad \rho_{m+1} \leq \frac{3}{4}\tilde\rho_m \tag{3.4.30}$$

(that is inequality (3.4.28) and (3.4.29) with a simple change of notation) where C_m is defined in (3.4.27). If we set $H^{(2)} := H$ and define recursively $H^{(m)} = H^{(m-1)} \circ \tau_m$ for $m = 3, \ldots, s$, assuming (3.4.30) and

$$\tilde\rho_m \leq \rho_m \tag{3.4.31}$$

it is easy to see that the maps τ_m for $m = 3, \ldots, s$ can be well defined and composed as in (3.4.23). Now by the assumptions made each function $H^{(m)}$ is real-analytic on $D^{2p}_{\rho_{m+1}}$ where we define

$$\rho_3 := r_0\,. \tag{3.4.32}$$

Furthermore we infer that $H^{(m)}$ has its norm bounded by $M\rho^2_{m+1}$ (for M in (3.4.5)) on $D_{\rho_{m+1}}$. Indeed, every Hamiltonian function $H^{(m)}$ can be written in the form $H^{(m)}(z) = \tilde H^{(m)}(z) \cdot z \cdot z$ for some appropriate $\tilde H^{(m)}$ real-analytic on $D^{2p}_{\rho_{m+1}}$; moreover if $\tilde H$ is a real-analytic function on D_{r_0} such that $H(\zeta) = \tilde H(\zeta) \cdot \zeta \cdot \zeta$, it results $\tilde H^{(m)} = \tilde H \circ \tau_3 \circ \cdots \circ \tau_{m-1} \circ \tau_m$ and therefore

$$|H^{(m)}|_{\rho_{m+1}} \leq \rho^2_{m+1} |\tilde H^{(m)}|_{\rho_{m+1}} \leq \rho^2_{m+1} |\tilde H|_{\rho_3} \leq \rho^2_{m+1} M$$

having used (3.4.5). Now, by Cauchy's estimate, the definition of C_m in (3.4.27) and the preceding inequality we have

$$C_m = 8^m (2m+1)^{2p} \sup_{|k|+|l|=m} \frac{1}{k!l!} \left| \frac{\partial^m H^{(m-1)}(0,0)}{\partial\xi^k \partial\eta^l} \right| \leq$$

$$\leq 8^m (2m+1)^{2p} \frac{|H^{(m-1)}|_{\rho_m}}{\rho_m^m} \leq 8^m (2m+1)^{2p} \frac{M}{\rho_m^{m-2}}\,.$$

Thus, we may fulfill the first condition in (3.4.30) by requiring

$$\tilde\rho_m \le \widetilde{C}_m^{\frac{1}{m-2}} \frac{\rho_m}{8} \qquad \text{with} \qquad \widetilde{C}_m := \frac{1}{8^2 (2m+1)^{2p}} \frac{a}{5M} \,. \tag{3.4.33}$$

Now our aim is to find a sequence of $\tilde\rho_m$ and ρ_m verifying

$$\begin{cases} \tilde\rho_m \le \widetilde{C}_m^{\frac{1}{m-2}} \frac{1}{8} \rho_m & \rho_3 := r_0 \\ \rho_{m+1} \le \frac{3}{4} \tilde\rho_m & m = 3, \ldots, s \,. \end{cases} \tag{3.4.34}$$

Observe that $\widetilde{C}_m \le 1$ since we assumed $a \le 1$ and $M \ge 1$; thus the first condition in (3.4.34) as well as (3.4.31) can be satisfied for every $m = 3, \ldots, s$ taking

$$\tilde\rho_m := \frac{\delta_1^{\frac{1}{2}}}{8} \rho_m \qquad \text{with} \qquad \delta_1 := \frac{1}{8^2 (2s+1)^{2p}} \frac{a}{5M} \tag{3.4.35}$$

while the second condition in (3.4.34) can be met simply defining

$$\rho_{m+1} := \frac{3}{4} \tilde\rho_m = \delta_2 \rho_m \qquad \text{with} \qquad \delta_2 := \frac{3}{2^5} \delta_1^{\frac{1}{2}} \,.$$

This definitions permit the construction of the map τ as described before and give as a consequence the definitions of $r := \tilde\rho_3$ in (3.4.11) (with (3.4.32)) and

$$r_s := \rho_{s+1} = \delta_2 \rho_s = \delta_2^{s-2} \rho_3 = \delta_2^{s-2} r_0 \tag{3.4.36}$$

as it is in (3.4.12) since $\delta = \delta_1^{\frac{1}{2}} 8^{-1}$ □

We now apply theorem 3.4.1 with

$$H(\tilde y, \tilde p, \tilde q) = \frac{1}{2} \sum_{j=1}^{p} \hat\Omega_j(\tilde y) \left(\tilde p_j^2 + \tilde q_j^2 \right) + \tilde g_3(\tilde y, \tilde p, \tilde q) \tag{3.4.37}$$

where $(\tilde p, \tilde q)$ play the role of (ξ, η) and $\tilde g_3$ of g, the y-variables are considered as fixed parameters varying in $B_{\frac{r}{5}}$, $r_0 = \frac{r_1}{5}$, $s = N_2$ and $a = \frac{\alpha}{4}$ in view of lemma 3.4.1. Then there exists a real-analytic symplectic transformation

$$\Phi_\epsilon^4 : (\theta, r, u, v) \in \mathbb{T}^d_{\frac{\sigma}{8}} \times B_{\frac{r}{5}} \times D_{r'} := \mathcal{D}_4 \longrightarrow (\tilde x, \tilde y, \tilde p, \tilde q) \in \mathbb{T}^d_{\frac{\sigma}{8}} \times B_{\frac{r}{5}} \times D_{r_0'} := \mathcal{D}_3' \tag{3.4.38}$$

where (u, v) are the variables (p, q) in (3.4.7), $\theta = \tilde x$ and $r = \tilde y$ (i.e., if $f(u, \tilde q)$ is a generating function of the map τ then $f(u, \tilde q) + \theta \cdot \tilde y$ is a generating function of Φ_ϵ^4), which puts H into Birkhoff's normal form up to order N_2 (that is into the form (3.4.10)).

To determine the two radii of the elliptic variables' domains we need to find a suitable value for M in (3.4.5). We start estimating

$$|H(\tilde{y},\tilde{p},\tilde{q})|_{\mathcal{D}_3} \le \left[M_2\frac{r_1^2}{5^2}+3\hat{M}_1\right]$$

in view of (3.3.39) (see (3.3.19) for $\hat{M}_1$). Now, since H possesses only terms of order greater than two in $(\tilde{p},\tilde{q})$ (see also (3.3.40)), we may write for any $(\tilde{p},\tilde{q})$ in $D_{\frac{r_1}{6}}$

$$\sup_{y\in B_{\frac{r}{5}}}|H(\tilde{y},\tilde{p},\tilde{q})| \le \sum_{k=2}^{\infty}\frac{\left|\partial^k_{(\tilde{p},\tilde{q})}H(\tilde{y},0,0)\right|_{\frac{r}{5}}}{k!}|(\tilde{p},\tilde{q})|^k \le$$

$$\le \sum_{k=2}^{\infty}\left(\frac{5}{r_1}\right)^k |H(\tilde{y},\tilde{p},\tilde{q})|_{\mathcal{D}_3}|(p,q)|^{k-2}|(p,q)|^2 \le$$

$$\le \frac{6^2}{r_1^2}\left[pM_2\frac{r_1^2}{5^2}+3\hat{M}_1\right]\sum_{k=2}^{\infty}\left(\frac{5}{6}\right)^k|(p,q)|^2 \le 6\left[M_2+75\hat{M}_1r_1^{-2}\right]|(p,q)|^2\,;$$

since $\frac{r_1}{6}\ge r_0'$ for any choice of $M\ge 1$ (see (3.4.11) where r plays the role of r_0' and remind $a\le 1$), we may consider $\rho\le\frac{r_1}{6}$ (instead of $\rho\le r_0=\frac{r_1}{5}$) in (3.4.5), so that we are allowed to take

$$M := \max\left\{pM_2+75\hat{M}_1r_1^{-2},1\right\}\,. \tag{3.4.39}$$

Then, equations (3.4.11) and (3.4.12) give the following suitable definitions of r_0' and r' in (3.4.38) (respectively r and r_s in theorem (3.4.1))

$$r_0' := \delta\frac{r_1}{5} \qquad \text{and} \qquad r' := \left(\frac{3}{4}\delta\right)^{N_2-2}\frac{r_1}{5} \tag{3.4.40}$$

with

$$\delta := \frac{1}{2^7(2N_2+1)^p}\left(\frac{\alpha}{5M}\right)^{\frac{1}{2}} \tag{3.4.41}$$

in view of all the agreements previously made.

With the domains just defined, from theorem (3.4.1) we obtain that H can be transformed into

$$N(r,u,v) := H\circ\Phi_\epsilon^4(\theta,r,u,v) = Q(r,u,v)+R_{N_2}(r,u,v) \tag{3.4.42}$$

for any (θ,r,u,v) in $\mathcal{D}_4$ (see (3.4.38)), where Q is a polynomial of degree $k := \left[\frac{N_2}{2}\right]$ in the variables $I=(I_1,\dots,I_p)$ with

$$I_j = \frac{1}{2}(u_j^2+v_j^2)\,,$$

depending also on $r \in B_{\frac{r}{5}}$, in the form

$$Q(r,u,v) = \hat{\Omega}(r)\cdot I + A^{(2)}(r)\otimes(I)^2 + A^{(3)}(r)\otimes(I)^3 + \cdots + A^{(k)}(r)\otimes(I)^k \quad (3.4.43)$$

where $A^{(j)}(r)$ is a real-analytic tensor of order j in $\mathbb{R}^p$, and R_{N_2} is a real-analytic function on $B_{\frac{r}{5}} \times D_{r'}$ verifying

$$\sup_{r\in B_{\frac{r}{5}}} |R_{N_2}(r,u,v)| \leq M_{N_2}|(u,v)|^{N_2+1} \ , \quad \forall\ (u,v)\in D_{r'} \quad (3.4.44)$$

for some $M_{N_2} > 0$. Then the new Hamiltonian function $H^4_\epsilon := H^3_\epsilon \circ \Phi^4_\epsilon$ (see (3.3.38) for H^3_ϵ) assumes for every $(\theta,r,u,v)\in\mathcal{D}_4$ the form

$$\begin{aligned} H^4_\epsilon(\theta,r,u,v) &= h(r) + \epsilon\hat{g}_0(r) + \frac{\epsilon}{2}\sum_{j=1}^{p}\hat{\Omega}_j(r)(u_j^2+v_j^2) + \epsilon Q(r,u,v) + \\ &+ \epsilon R_{N_2}(r,u,v) + \epsilon^{N_1}\tilde{f}_4(\theta,r,u,v) \end{aligned} \quad (3.4.45)$$

where $\tilde{f}_4 = \tilde{f}_3 \circ \Phi^4_\epsilon$.

Now we aim to estimate the coefficients $A^{(j)}$ of Q and the constant M_R appearing in (3.4.44). For the uniqueness of the Taylor coefficients, for every $j = 2,3,\ldots,\left[\frac{N_2}{2}\right]$ we have

$$\frac{A^{(j)}(r)}{2^j} = \frac{\partial^{2j}_{(u,v)}Q(r,0,0)}{(2j)!} = \frac{\partial^{2j}_{(u,v)}N(r,0,0)}{(2j)!}$$

where this last equality holds because of (3.4.44) and the definition of I. Therefore Cauchy's estimate and easy calculations, together with (3.3.39), yield

$$\begin{aligned} \sup_{r\in B_{\frac{r}{5}}}|A^{(j)}(r)| &\leq \frac{2^j}{(2j)!}\sup_{r\in B_{\frac{r}{5}}}\left|\partial^{2j}_{(u,v)}N(r,0,0)\right| \leq \frac{2^j}{(r')^{2j}}|N(r,u,v)|_{\mathcal{D}_4} \leq \\ &\leq \frac{2^j}{(r')^{2j}}|H(\tilde{y},\tilde{p},\tilde{q})|_{\mathcal{D}'_3} \leq \frac{2^j}{(r')^{2j}}|H(\tilde{y},\tilde{p},\tilde{q})|_{\mathcal{D}_3} \leq \\ &\leq \frac{2^j}{(r')^{2j}}\left[5^{-2}pM_2r_1^2 + 3\hat{M}_1\right] \leq \frac{2^j}{(r')^{2j}}\frac{r_1^2M}{5^2} \end{aligned} \quad (3.4.46)$$

where M is defined in (3.4.39). Moreover, from equations (3.4.42), (3.4.43) and (3.4.44), we can write

$$R_{N_2}(r,u,v) = \sum_{j=N_2+1}^{\infty}\frac{\partial^j_{(u,v)}N(r,0,0)}{j!}\otimes(u,v)^j$$

whence for every (u,v) in $D_{\frac{r'}{2}}$ it results

$$\begin{aligned}
&\sup_{r\in B_{\frac{r}{5}}} |R_{N_2}(r,u,v)| \leq \sum_{j=N_2+1}^{\infty} \frac{\left|\partial^j_{(u,v)} N(r,0,0)\right|_{\frac{r}{5}}}{j!} |(u,v)|^j \leq \\
&\leq |(u,v)|^{N_2+1} \sum_{j=N_2+1}^{\infty} \frac{|N(r,u,v)|_{\mathcal{D}_4}}{(r')^j} \left(\frac{r'}{2}\right)^{j-(N_2+1)} \leq \\
&\leq \frac{|(u,v)|^{N_2+1}}{(r')^{N_2+1}} \sum_{j=N_2+1}^{\infty} |H(\tilde{y},\tilde{p},\tilde{q})|_{\mathcal{D}'_3} \frac{1}{2^{j-(N_2+1)}} \leq \\
&\leq \frac{|H(\tilde{y},\tilde{p},\tilde{q})|_{\mathcal{D}_3}}{(r')^{N_2+1}} |(u,v)|^{N_2+1} \leq \frac{2r_1^2 M}{5^2} \frac{|(u,v)|^{N_2+1}}{(r')^{N_2+1}}
\end{aligned}$$

so that we can take

$$M_{N_2} := \frac{2}{(r')^{N_2+1}} \frac{r_1^2 M}{5^2} \tag{3.4.47}$$

in (3.4.44).

3.5 Symplectic polar coordinates

We aim now to find a symplectic transformation casting H^4_ϵ in a simpler form than (3.4.45) and that serves three main purposes:

- the new variables are action-angle variables for the integrable part of the new Hamiltonian function;
- the domain of the new actions is a neighborhood of the origin in $\mathbb{R}^{d+p}$ (which will allow a further rescaling);
- under some hypotheses we are going to make, the elliptic singularity in every symplectic plane (u_j, v_j) can be avoided.

Let $r_0 := I_0$ the center of the ball $B_{\frac{r}{5}} = D^d_{\frac{r}{5}}(I_0)$ (see notation (3.2.19) and section 3.2) and take

$$\rho^0 = (\rho^0_1, \dots, \rho^0_j) \in (\mathbb{R}_+)^p \qquad \text{with} \qquad |\rho^0| \leq \frac{1}{12}(r')^2 \tag{3.5.1}$$

$$\upsilon := \frac{1}{4} \log \frac{(r')^2}{6|\rho^0|} > 0\,. \tag{3.5.2}$$

Now we consider the transformation

$$\Phi_\epsilon^5 : (\theta, r, \zeta, \rho) \in \mathbb{T}^d_{\frac{\sigma}{8}} \times D^d_{\frac{r}{5}} \times \mathbb{T}^p_v \times D^p_{\frac{|\rho^0|}{2}} := \mathcal{D}_5 \longrightarrow (\theta, r_0 + r, z) \in \mathcal{D}_4$$

where $z = u + iv$ with

$$z_j := \sqrt{\rho_j^0 + \rho_j}\, e^{-2i\zeta_j} \tag{3.5.3}$$

for every $j = 1, \ldots, p$ and D^n_t denotes for $t > 0$ and $n = d, p$ the open ball of radius t and center 0 in $\mathbb{C}^n$, i.e.,

$$D^n_t := \{x \in \mathbb{C}^n \ : \ |x| \le t\} \ . \tag{3.5.4}$$

As it can be easily seen Φ_ϵ^5 is a symplectic map with generating function

$$f(\theta, r, \zeta, v) = \theta \cdot (r_0 + r) - \rho^0 \cdot \zeta - \sum_{j=1}^p -\frac{v_j^2}{2} \cot(2\pi\zeta_j) \, ;$$

from its image is effectively contained in $\mathcal{D}_4$ (see (3.4.38)) since, from (3.5.1) and (3.5.2)

$$|z_j| \le \left|\rho_j^0 + \rho_j\right|^{\frac{1}{2}} e^{2v} \le \sqrt{\frac{3}{2}}\sqrt{|\rho^0|}\sqrt{\frac{1}{6}}\frac{r'}{\sqrt{|\rho^0|}} = \frac{r'}{2} \, .$$

The transformed Hamiltonian function, real-analytic on $\mathcal{D}_5$, is given by

$$\begin{aligned} H_\epsilon^5(\theta, r, \zeta, \rho) \ &:= \ H_\epsilon^4 \circ \Phi_\epsilon^5 = h(r_0 + r) + \epsilon \hat{g}_0(r_0 + r) + \\ &+ \ \frac{\epsilon}{2}\sum_{j=1}^p \hat{\Omega}_j(r^0 + r)(\rho_j^0 + \rho_j^0) + \epsilon Q'(r_0 + r, \rho^0 + \rho) + \\ &+ \ \epsilon R'_{N_2}(r_0 + r, \zeta, \rho; \rho^0) + \epsilon^{N_1} \tilde{f}_5(\theta, r, \zeta, \rho) \end{aligned} \tag{3.5.5}$$

where, in view of $2I_j = u_j^2 + v_j^2 = |z_j|^2 = \rho_j^0 + \rho_j$, we have

$$Q'(r_0 + r, \rho^0 + \rho) := Q \circ \Phi_\epsilon^5 = \sum_{j=2}^{\left[\frac{N_2}{2}\right]} \frac{A^{(j)}(r_0 + r)}{2^j} \otimes (\rho^0 + \rho)^j \tag{3.5.6}$$

(see (3.4.43) for $Q(r, u, v)$) and $R'_{N_2} := R_{N_2} \circ \Phi_\epsilon^5$ verifies (in view of (3.4.44))

$$\begin{aligned} &\sup_{r \in B_{\frac{r}{5}}} |R'_{N_2}(r_0 + r, \zeta, \rho; \rho^0)| \le \sup_{r \in B_{\frac{r}{5}}} |R_{N_2}(r, u, v)| \le M_{N_2} |(u, v)|^{N_2 + 1} \le \\ &\le M_{N_2} |\rho^0 + \rho|^{\frac{N_2+1}{2}} \le M_{N_2} \left(\frac{3}{2}\right)^{\frac{N_2+1}{2}} |\rho^0|^{\frac{N_2+1}{2}} \ ; \end{aligned} \tag{3.5.7}$$

moreover $\tilde{f}_5 = \tilde{f}_4 \circ \Phi^5_\epsilon$ is bounded on $\mathcal{D}_5$ by μ^{N_1}.

We now choose and fix ρ^0 in (3.5.1) in order to have

$$|\rho^0| = 2\epsilon \tag{3.5.8}$$

and consider the homothety A_ϵ on $\mathcal{D}_5 = \mathbb{T}^d_{\frac{\sigma}{8}} \times D^d_{\frac{r}{5}} \times \mathbb{T}^p_v \times D^p_{\frac{|\rho^0|}{2}}$ given by

$$A_\epsilon : (\theta, r, \zeta, \rho) \longrightarrow (\theta, \epsilon r, \zeta, \epsilon\rho)\,. \tag{3.5.9}$$

It is well known that though A_ϵ is not a symplectic change of coordinates (unless $\epsilon = 1$), it is conformally symplectic, i.e., it preserves the structure of Hamilton's equations (and therefore of the solutions of the dynamical system considered), if we take as new Hamiltonian function $H^6_\epsilon := \frac{1}{\epsilon} H^5_\epsilon \circ A_\epsilon$. Explicitly we have

$$\begin{aligned} H^6_\epsilon(\theta, r, \zeta, \rho) &= \frac{1}{\epsilon} h(r_0 + \epsilon r) + \hat{g}_0(r_0 + \epsilon r) + \frac{1}{2}\hat{\Omega}(r_0 + \epsilon r)\cdot(\rho^0 + \epsilon\rho) + \\ &+ Q'(r_0 + \epsilon r, \rho^0 + \epsilon\rho) + R'_{N_2}(r_0 + \epsilon r, \epsilon\rho, \zeta; \rho^0) + \\ &+ \epsilon^{N_1 - 1}\tilde{f}_6(\theta, r, \zeta, \rho) \end{aligned} \tag{3.5.10}$$

where $\tilde{f}_6 = \tilde{f}_5 \circ A_\epsilon$ and Q' is defined in (3.5.6). Besides inequality in (3.5.7) and the choice made in (3.5.8) give

$$|R'_{N_2}(r_0 + \epsilon r, \zeta, \epsilon\rho; \rho^0)| \le M'_{N_2}\epsilon^{\frac{N_2+1}{2}} \tag{3.5.11}$$

for every $(r, \zeta, \rho) \in D^d_{\frac{r}{5}} \times \mathbb{T}^p_v \times D^p_\epsilon$, having defined

$$M'_{N_2} := 3^{\frac{N_2+1}{2}} M_{N_2}\,. \tag{3.5.12}$$

Therefore we can move R'_{N_2} to the perturbative part of H^6_ϵ; more precisely we choose $N_2 = 2N_1 - 3$ and define N as

$$N := N_1 - 1 = \frac{N_2 + 1}{2}. \tag{3.5.13}$$

Then the perturbation of H^6_ϵ can be written as

$$P_\epsilon(\theta, r, \zeta, \rho; \epsilon) = R'_{N_2}(r_0 + \epsilon r, \zeta, \epsilon\rho; \rho^0) + \epsilon^N \tilde{f}_6(\theta, r, \zeta, \rho) \tag{3.5.14}$$

which can be estimated by

$$|P_\epsilon|_{\mathcal{D}_5} \le (M'_{N_2} + \mu^{N+1})\epsilon^N \tag{3.5.15}$$

where M'_{N_2} is defined in (3.5.12) (together with (3.4.47), (3.3.19) and (3.3.24)) and μ is defined in (3.2.21).

Chapter 4

Rüßmann's tori for properly degenerate systems

4.1 Non-degeneration of the frequency application

Define F_ϵ as the integrable part of the Hamiltonian function H_ϵ^6 in (3.5.10), i.e.,

$$F_\epsilon(r_0+\epsilon r,\rho^0+\epsilon\rho) := \frac{1}{\epsilon}h(r_0+\epsilon r)+\hat{g}_0(r_0+\epsilon r)+ \tag{4.1.1}$$

$$+\frac{1}{2}\sum_{j=1}^{p}\hat{\Omega}_j(r_0+\epsilon r)(\rho_j^0+\epsilon\rho_j)+\sum_{j=2}^{N-1}\frac{A^{(j)}(r_0+\epsilon r)}{2^j}\otimes(\rho^0+\epsilon\rho)^j$$

where we wrote the explicit expression for $Q'(r_0+\epsilon r,\rho^0+\epsilon\rho)$ in (3.5.6) using $\left[\frac{N_2}{2}\right]=\left[N-\frac{1}{2}\right]=N-1$ (in view of (3.5.13)); recall that in view of $\epsilon<1$ and the choice of $\rho^0\in(\mathbb{R}_+^p)$ with $|\rho^0|=2\epsilon$, F_ϵ is real-analytic on $D^d_{\frac{r}{5}}\times D^p_\epsilon$ where here and from now on, we are coherent with notation in (3.5.4). Our aim is now to show that the frequency application of the torus $\mathbb{T}^{d+p}_{(r,\rho)}$ of F_ϵ is non-degenerate in the sense of Rüßmann under the hypothesis of non-degeneration of the "unperturbed frequency application"

$$\Psi : I\in D_r^d(I_0)\cap\mathbb{R}^d \longrightarrow (\omega(I),\Omega_1(I),\ldots,\Omega_p(I))\in\mathbb{R}^d\times\mathbb{R}^p \tag{4.1.2}$$

where $\omega(I):=\nabla h(I)$ and $D_r^d(I_0)$ (denoted before with B_r) is the ball where we localized initially in order to have conditions in (3.2.14) satisfied simultaneously for any fixed $N_1,N_2\geq 2$.

We start trying to establish a relation between Ψ, the frequency application of the integrable part of the Hamiltonian H_ϵ considered initially, and the frequency application of F_ϵ that we shall call $\hat{\Psi}_\epsilon$; we will show that $\hat{\Psi}_\epsilon$ is $0(\epsilon)$-close to a slight modification of Ψ that is still non-degenerate, obtaining as a consequence the non-degeneration of $\hat{\Psi}_\epsilon$ for small enough ϵ.

The frequency application of F_ϵ is given by

$$\hat{\Psi}_\epsilon(r_0 + \epsilon r, \rho^0 + \epsilon\rho) := \left(\frac{\partial}{\partial r} F_\epsilon(r_0 + \epsilon r, \rho^0 + \epsilon\rho), \frac{\partial}{\partial \rho} F_\epsilon(r_0 + \epsilon r, \rho^0 + \epsilon\rho) \right) \tag{4.1.3}$$

for any $(r, \rho) \in D^d_{\frac{r}{5}} \times D^p_\epsilon$; more explicitly we can compute

$$\frac{\partial}{\partial r} F_\epsilon(r_0 + \epsilon r, \rho^0 + \epsilon\rho) = \omega(r_0 + \epsilon r) + \epsilon \frac{\partial \hat{g}_0}{\partial r}(r_0 + \epsilon r) +$$
$$+ \frac{\epsilon}{2} \sum_{j=1}^{p} \frac{\partial \hat{\Omega}_j}{\partial r}(r_0 + \epsilon r)(\rho^0_j + \epsilon\rho_j) + \epsilon \frac{\partial Q'}{\partial r}(r_0 + \epsilon r, \rho^0 + \epsilon\rho) \tag{4.1.4}$$

where $Q'(r_0 + \epsilon r, \rho^0 + \epsilon\rho) = \sum_{j=2}^{N-1} \frac{A^{(j)}(r_0+\epsilon r)}{2^j} \otimes (\rho^0 + \epsilon\rho)^j$ (see (3.5.6)), and

$$\frac{\partial}{\partial \rho} F_\epsilon(r_0 + \epsilon r, \rho^0 + \epsilon\rho) = \frac{\epsilon}{2} \hat{\Omega}(r_0 + \epsilon r) + \epsilon \frac{\partial Q'}{\partial \rho}(r_0 + \epsilon r, \rho^0 + \epsilon\rho) \tag{4.1.5}$$

with $\hat{\Omega} = (\hat{\Omega}_1, \ldots, \hat{\Omega}_p)$ as usual. We can immediately notice that $\partial_r F_\epsilon$ is $O(\epsilon)$-close to $\omega(r_0 + \epsilon r)$, that is a non-degenerate function on $D^d_{\frac{r}{5}} \cap \mathbb{R}^d$ since r_0 as been chosen to coincide with I_0 and therefore

$$\{r_0\} + \epsilon \left(D^d_{\frac{r}{5}} \cap \mathbb{R}^d \right) \subset D_r(I_0) \cap \mathbb{R}^d \,.$$

Analogously $2\epsilon^{-1}\partial_\rho F_\epsilon$ is $O(\epsilon)$-close to the non-degenerate function $\Omega(r_0 + \epsilon r) = (\Omega_1(r_0 + \epsilon r), \ldots, \Omega_p(r_0 + \epsilon r))$ since $\hat{\Omega}$ is $O(\epsilon)$-close to Ω and $\partial_\rho Q'$ is a function of order 1 in $\rho = O(\epsilon)$ (being Q' a function of order $|\rho|^2$). The following proposition displays the details of this observations:

Proposition 4.1.1. *Let*

$$\Psi_\epsilon(r_0 + \epsilon r, \rho^0 + \epsilon\rho) := \left(\frac{\partial}{\partial r} F_\epsilon(r_0 + \epsilon r, \rho^0 + \epsilon\rho), \frac{2}{\epsilon} \frac{\partial}{\partial \rho} F_\epsilon(r_0 + \epsilon r, \rho^0 + \epsilon\rho) \right) \tag{4.1.6}$$

then for any $(r, \rho) \in D^d_{\frac{r}{10}} \times D^p_{\frac{\epsilon}{2}}$ *it results*

$$\left| \Psi_\epsilon(r_0 + \epsilon r, \rho^0 + \epsilon\rho) - \Psi(r_0 + \epsilon r) \right| \leq \Delta\epsilon \,. \tag{4.1.7}$$

where

$$\Delta := B_{N,p} A_1^{2p} \left(\frac{M}{\alpha} \right)^\kappa \frac{\alpha}{d'} \tag{4.1.8}$$

with M defined in (3.4.39), d' in (3.2.24),

$$\left\{\begin{array}{rcl} B_{N,p} & := & (2p)!p[2^9(4N-1)^p]^{8N-10} \\ A_1 & := & \max\left\{\frac{c_0 r_1}{\alpha}, 1\right\} \\ \kappa & := & \{4N-5, 2p\} \,. \end{array}\right. \tag{4.1.9}$$

Proof First of all we translate the domain of Ψ in (4.1.2) considering $\Psi^0_\epsilon(r) = \Psi(r_0 + \epsilon r)$ given by

$$\Psi^0_\epsilon : r \in D^d_{\frac{r}{5}} \longrightarrow (\omega(r_0+\epsilon r), \Omega_1(r_0+\epsilon r), \ldots, \Omega_p(r_0+\epsilon r)) \in \mathbb{R}^d \times \mathbb{R}^p \tag{4.1.10}$$

real-analytic and non-degenerate on $D^d_{\frac{r}{5}}$ since $r_0 = I_0$ and $\epsilon < 1$. Now we define

$$\begin{aligned} G_1(r_0+\epsilon r, \rho^0+\epsilon\rho) &= \frac{\partial}{\partial r} F_\epsilon(r_0+\epsilon r, \rho^0+\epsilon\rho) - \omega(r_0+\epsilon r) \\ G_2(r_0+\epsilon r, \rho^0+\epsilon\rho) &= \frac{2}{\epsilon}\frac{\partial}{\partial \rho} F_\epsilon(r_0+\epsilon r, \rho^0+\epsilon\rho) - \Omega(r_0+\epsilon r) \end{aligned} \tag{4.1.11}$$

so that it results

$$\Psi_\epsilon(r_0+\epsilon r, \rho^0+\epsilon\rho) - \Psi^0_\epsilon(r) = \big(G_1(r_0+\epsilon r, \rho^0+\epsilon\rho), G_2(r_0+\epsilon r, \rho^0+\epsilon\rho)\big) \tag{4.1.12}$$

In view of (4.1.4) and the definition of G_1 we have

$$|G_1(r_0+\epsilon r, \rho^0+\epsilon\rho)| \le \epsilon|\partial_r \hat{g}_0(r_0+\epsilon r)| +$$

$$+\frac{\epsilon}{2}\sum_{i=1}^{p} |\partial_r \hat{\Omega}_j(r_0+\epsilon r)|\,\big|(\rho^0_j+\epsilon\rho_j)\big| + \epsilon|\partial_r Q'(r_0+\epsilon r, \rho^0+\epsilon\rho)|\,.$$

Then we can estimate

$$\sup_{r\in D^d_{\frac{r}{10}}} |\partial_r \hat{g}_0(r_0+\epsilon r)| \le \frac{10}{r}\sup_{r\in D^d_{\frac{r}{5}}} |\hat{g}_0(r_0+\epsilon r)| \le \frac{10\hat{M}_1}{r}$$

in view of (3.3.19), while from (3.3.24), (3.5.8) and (3.5.1) we obtain for every ρ in D^p_ϵ

$$\sup_{r\in D^d_{\frac{r}{10}}} |\partial_r \hat{\Omega}_j(r_0+\epsilon r)|\,\big|(\rho^0_j+\epsilon\rho_j)\big| \le \frac{10}{r}\sup_{r\in D^d_{\frac{r}{5}}} |\hat{\Omega}_j(r_0+\epsilon r)|\,\big(|\rho^0|+\epsilon|\rho|\big) \le$$

$$\le \frac{10M_2}{r}2|\rho^0| \le \frac{10M_2}{r}\frac{1}{6}(r')^2 \le \frac{5M_2}{3r}\left(\frac{r_1}{5}\right)^2 \le \frac{M_2 r_1^2}{15r}\,.$$

Furthermore the definition of Q' in (3.5.6) and the estimate of $A^{(j)}$ in (3.4.46) yield for every ρ in D^p_ϵ

$$\sup_{r\in D^d_{\frac{r}{10}}} |\partial_r Q'(r_0+\epsilon r,\rho^0+\epsilon\rho)| \leq \frac{10}{r}\sup_{r\in D^d_{\frac{r}{5}}} |Q'(r_0+\epsilon r,\rho^0+\epsilon\rho)| \leq$$

$$\leq \frac{10}{r}\sup_{r\in D^d_{\frac{r}{5}}}\sum_{j=2}^{N-1}\frac{\left|A^{(j)}(r_0+\epsilon r)\right|}{2^j}\left|\rho^0+\rho\right|^j \leq \frac{10}{r}\sum_{j=2}^{N-1}\frac{Mr_1^2}{5^2(r')^{2j}}(2|\rho^0|)^j \leq$$

$$\leq \frac{2(N-2)Mr_1^2}{5r}$$

having used $2|\rho^0| \leq (r')^2$ from (3.5.1). Putting together this three estimates we obtain

$$\sup_{r\in D^d_{\frac{r}{10}}} |G_1(r,\rho)| \leq \frac{1}{r}\left(10\hat{M}_1+\frac{M_2r_1^2}{15}+\frac{2(N-2)Mr_1^2}{5}\right)\epsilon \leq$$

$$\leq \frac{1}{r}\left(10\hat{M}_1+\frac{N-1}{2}Mr_1^2\right)\epsilon \leq \frac{(N+19)M}{2r}\epsilon \tag{4.1.13}$$

where we used $M > pM_2 \geq M_2$ and $Mr_1^2 > 75\hat{M}_1 > 10\hat{M}_1$ from the definition of M in (3.4.39) and $r_1^2 \leq 1$.

Analogously from (4.1.5) and the definition of G_2 it results

$$|G_2(r_0+\epsilon r,\rho^0+\epsilon\rho)| \leq |\hat{\Omega}(r_0+\epsilon r)-\Omega(r_0+\epsilon r)| + \frac{1}{2}|\partial_\rho Q'(r_0+\epsilon r,\rho^0+\epsilon\rho)|\,;$$

for what concerns the first term, from (3.3.30) we have

$$\sup_{r\in D^d_{\frac{r}{5}}} |\hat{\Omega}(r_0+\epsilon r)-\Omega(r_0+\epsilon r)| \leq 2^{2p}(2p)!2p\min\{(2M_2)^{-1}\alpha, M_3^{-1}\alpha\}^{-2p}\alpha\epsilon \tag{4.1.14}$$

while $\partial_\rho Q'$ can be estimated for every $r \in D^d_{\frac{r}{5}}$ as follows:

$$\frac{1}{2}\sup_{\rho\in D^p_{\frac{\epsilon}{2}}} \left|\partial_\rho Q'(r_0+\epsilon r,\rho^0+\epsilon\rho)\right| \leq \frac{1}{\epsilon}\sup_{\rho\in D^p_\epsilon} |Q'(r_0+\epsilon r,\rho^0+\epsilon\rho)| \leq$$

$$\leq \frac{1}{\epsilon}\sup_{\rho\in D^p_\epsilon}\sum_{j=2}^{N-1}\frac{|A^{(j)}(r_0+\epsilon r)|}{2^j}\left|\rho^0+\rho\right|^j \leq \frac{1}{\epsilon}\sum_{j=2}^{N-1}\frac{Mr_1^2}{5^2}\frac{(2|\rho^0|)^j}{(r')^{2j}} \leq$$

$$\leq \frac{(2|\rho^0|)^2}{\epsilon}\frac{Mr_1^2}{5^2(r')^4}\sum_{j=2}^{N-1}\frac{(2|\rho^0|)^{j-2}}{(r')^{2(j-2)}} \leq \frac{(N-2)Mr_1^2}{5^2(r')^4}\epsilon \tag{4.1.15}$$

where we used once again $2|\rho^0| = \epsilon \le (r')^2$.

Now we search for an easy expression as an upper bound for $|G_2|$. First of all from the definitions of M in (3.4.39) and $\hat{M}_1$ in (3.3.19) we obtain

$$M \ge \frac{5^2 6 \hat{M}_1}{r_1^2} \ge 5^2 6 c_p \frac{\mu^2}{\alpha d' r_1^2} \ge 2^7 c_p \frac{\mu^2}{\alpha d' r_1^2}\,; \qquad (4.1.16)$$

thus, from the definition of A_1 in (4.1.9), M_3 in (3.3.24) and $M \ge M_2$, we obtain

$$M A_1 \ge \max\{M_2, M_3\}\,. \qquad (4.1.17)$$

Therefore the second member in (4.1.14) can be estimated by

$$2^{4p}(2p)!2p \, \max\left\{\frac{M_2}{\alpha}, \frac{M_3}{\alpha}\right\}^{2p} \alpha\epsilon \le 2^{4p}(2p)!2p\left(\frac{M}{\alpha}\right)^{2p} A_1^{2p}\alpha\,\epsilon\,.$$

For what concerns the last member in (4.1.15) we recall the definition of r' in (3.4.40) and δ in (3.4.41) and compute

$$\begin{aligned} \frac{M r_1^2}{5^2 (r')^4} &= \frac{M 5^2}{r_1^2}\left(\frac{4}{3\delta}\right)^{4N_2-8} = \frac{M 5^2}{r_1^2}\left[\frac{2^9(2N_2+1)^p}{3}\right]^{4N_2-8}\left(\frac{5M}{\alpha}\right)^{2N_2-4} = \\ &\le \left[2^9(2N_2+1)^p\right]^{4N_2-7}\left(\frac{M}{\alpha}\right)^{2N_2-3}\frac{\alpha}{r_1^2}\,. \end{aligned}$$

Putting together this two last estimates with (4.1.14) and (4.1.15) and observing that $2^{4p+1} \le [2^9(2N_2+1)]^{p(4N_2-7)}$ (remind $N_2 \ge 3$), for any $(r,\rho) \in D^d_{\frac{r}{10}} \times D^p_{\frac{\epsilon}{2}}$ we obtain

$$\begin{aligned} &|G_2(r_0+\epsilon r, \rho^0+\epsilon\rho)| \le (2p)!p\left[2^9(2N_2+1)^p\right]^{4N_2-7}\left[\left(\frac{M}{\alpha}\right)^{2p} A_1^{2p} + \right. \\ &\left. + \frac{(N-2)}{r_1^2}\left(\frac{M}{\alpha}\right)^{2N_2-3}\right]\alpha\,\epsilon \le (2p)!p\left[2^9(2N_2+1)^p\right]^{4N_2-6} A_1^{2p}\left(\frac{M}{\alpha}\right)^{\kappa}\frac{\alpha}{r_1^2}\epsilon \le \\ &\le B A_1^{2p}\left(\frac{M}{\alpha}\right)^{\kappa}\frac{\alpha}{r_1^2}\epsilon \end{aligned}$$

having used $A_1 \ge 1$, $r_1^2 \le 1$, the definition of κ in (4.1.9) and taking

$$B := \frac{1}{2}(2p)!p\left[2^9(4N-1)^p\right]^{8N-10} = \frac{1}{2}B_{N,p} \qquad (4.1.18)$$

in view of $N_2 = 2N-1$ from (3.4.42). Now, combining this last result with (4.1.13) we have

$$\sup_{D^d_{\frac{r}{10}}\times D^p_{\frac{\epsilon}{2}}} |G(r_0+\epsilon r,\rho^0+\epsilon\rho)| \le \sup_{D^d_{\frac{r}{10}}\times D^p_{\frac{\epsilon}{2}}} \big(|G_1(r_0+\epsilon r,\rho^0+\epsilon\rho)|+$$

$$+|G_2(r_0+\epsilon r,\rho^0+\epsilon\rho)|\big) \le \left[\frac{(N+19)M}{2r} + BA_1^{2p}\left(\frac{M}{\alpha}\right)^{\kappa}\frac{\alpha}{r_1^2}\right] \le$$

$$\le BA_1^{2p}\left[\left(\frac{M}{\alpha}\right)\frac{\alpha}{r} + \left(\frac{M}{\alpha}\right)^{\kappa}\frac{\alpha}{r_1^2}\right] \le B_{N,p}A_1^{2p}\left(\frac{M}{\alpha}\right)^{\kappa}\frac{\alpha}{d'}$$

since $B \ge N+19$, $A_1 \ge 1$, $\alpha \le 1$ and $d' \le \min\{r\sigma, r_1^2\}$ with $\sigma < 1$ $\square$

We now cite a result by Rüßmann concerning the relation between non degeneracy and Taylor's coefficients of a real-analytic function:

Proposition 4.1.2. *Let $f : B \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m$ (where B is a domain) a real-analytic function. If the Taylor series*

$$f(y) = \sum_{j\in N^n} \frac{(y-b)^j}{j!} f^{(j)}(b)$$

of f in some point $b \in B$ contains m linearly independent coefficients, i.e., exist $j_1,\dots,j_m$ in $\mathbb{N}^n$ such that

$$\det\left[f^{(j_1)}(b),\dots,f^{(j_m)}(b)\right] \ne 0\,,$$

then f is non-degenerate. Conversely, if f is non-degenerate then in any point $b \in B$ there exists a set of m linearly independent coefficients (obviously depending on b).

We now use this proposition to prove that the non-degeneracy of Ψ implies the non-degeneracy of Ψ_ϵ for small enough ϵ

Proposition 4.1.3. *Assume Ψ in (4.1.2) is non-degenerate in the sense of Rüßmann then there exist $\delta = \delta(\Psi, r_0) > 0$ and $\nu = \nu(\Psi, r_0) \in \mathbb{N}_+$ such that if*

$$\epsilon \le \frac{\delta}{C(\nu!)^{d+p}}\left(\frac{r}{5}\right)^{(d+p)\nu}\min\left\{\frac{1}{M_0+M_2},\frac{1}{\Delta}\right\}^{(d+p)} \tag{4.1.19}$$

with

$$C = 2^{d+p}(d+p+1)!$$

then Ψ_ϵ in (4.1.6) is non-degenerate.

Proof Let Ψ_ϵ^0 be the real-analytic and non-degenerate function for r in $D^d_{\frac{r}{5}}$ considered in (4.1.10), then from (4.1.12) we have

$$\Psi_\epsilon(r_0 + \epsilon r, \rho^0 + \epsilon\rho) = \Psi_\epsilon^0(r) + \big(G_1(r_0 + \epsilon r, \rho^0 + \epsilon\rho), G_2(r_0 + \epsilon r, \rho^0 + \epsilon\rho)\big)$$

with G_1, G_2 defined in (4.1.11) and verifying

$$|(G_1, G_2)| \leq \Delta\epsilon$$

for Δ in (4.1.8) with (4.1.9). Now, from proposition 4.1.2 there exist $d+p$ vectors $j^{(1)}, \ldots, j^{(d+p)} \in \mathbb{N}^d$ such that

$$\det\left[\partial_r^{j^{(1)}}\Psi_\epsilon^0(0), \ldots, \partial_r^{j^{(d+p)}}\Psi_\epsilon^0(0)\right] \neq 0$$

and we may additionally assume without loss of generality

$$|j^{(1)}|_1 \leq |j^{(2)}|_1 \leq \cdots \leq |j^{(d+p)}|_1 \,. \tag{4.1.20}$$

Recalling the definition of Ψ_ϵ^0, we define for ρ in $D^p_{\frac{\epsilon}{2}}$ the following matrices:

$$\begin{aligned}
A &:= \left[\partial_r^{j^{(1)}}\Psi_\epsilon(r_0, \rho^0 + \epsilon\rho), \ldots, \partial_r^{j^{(d+p)}}\Psi_\epsilon(r_0, \rho^0 + \epsilon\rho)\right] \\
B &:= \left[\partial_r^{j^{(1)}}\Psi(r_0), \ldots, \partial_r^{j^{(d+p)}}\Psi(r_0)\right] \\
C &:= \frac{1}{\epsilon}\left[\partial_r^{j^{(1)}}(G_1, G_2)(r_0, \rho^0 + \epsilon\rho), \ldots, \partial_r^{j^{(d+p)}}(G_1, G_2)(r_0, \rho^0 + \epsilon\rho)\right] .
\end{aligned}$$

From what initially observed, it results $A = B + \epsilon C$; we may then apply lemma 3.3.3 obtaining $|\det A| \geq \frac{1}{2}|\det B| \neq 0$ for small enough ϵ, and as a consequence the non-degeneration of $\Psi_\epsilon(r, \rho)$ in view of proposition 4.1.2. By lemma 3.3.3 a sufficient condition to impose on ϵ is

$$\epsilon \leq \frac{|\det B|}{2^{d+p}(d+p+1)!} \max\{\|B\|, \|C\|\}^{-(d+p)} \,. \tag{4.1.21}$$

Observe now that each element b_{nk} of the matrix B verifies, by Cauchy's estimate and (4.1.20),

$$|b_{nk}| \leq \frac{5^{|v|_1} v!}{r^{|v|_1}} \sup_{r \in D^d_{\frac{r}{5}}} |\Psi_\epsilon^0(r)| \leq \frac{5^{|v|_1}|v|_1!}{r^{|v|_1}} \sup_{r \in D^d_{\frac{r}{5}}(r_0)} |\Psi(r)| \leq \frac{5^{|v|_1}|v|_1!}{r^{|v|_1}}(M_0 + M_2)$$

where $v := j^{(d+p)}$ (the multi-index with greatest norm), $M_0 := |\omega|_r$ and $M_2 := |\Omega|_r$. Furthermore we can estimate the norm of $C = (c_{nk})$ observing that for every ρ in $D_{\frac{\epsilon}{2}}$

$$|c_{nk}| \leq \frac{1}{\epsilon} \frac{5^{|v|_1} v!}{r^{|v|_1}} \sup_{r \in D^d_{\frac{r}{5}}} |G_1(r_0 + \epsilon r, \rho^0 + \epsilon\rho), G_2(r_0 + \epsilon r, \rho^0 + \epsilon\rho)| \leq \frac{5^{|v|_1} v!}{r^{|v|_1}} \Delta$$

in view of (4.1.6), (4.1.11) and (4.1.12). From this two estimates and the definition of $\|M\|$ given in lemma 3.3.3, we obtain

$$\begin{aligned} \|B\| &\leq \sup_{n,k=1,\ldots,d+p} |b_{nk}| \leq \frac{5^{|v|_1} |v|_1!}{r^{|v|_1}} (M_0 + M_2) \\ \|C\| &\leq \sup_{n,k=1,\ldots,d+p} |c_{nk}| \leq \frac{5^{|v|_1} |v|_1!}{r^{|v|_1}} \Delta\,. \end{aligned}$$

The statement follows taking in (4.1.19)

$$\nu(\Psi, r_0) := \left| j^{(d+p)} \right|_1 \tag{4.1.22}$$

$$\delta(\Psi, r_0) := \left| \det[\partial_r^{j^{(1)}} \Psi(r_0), \ldots, \partial_r^{j^{(d+p)}} \Psi(r_0)] \right| \tag{4.1.23}$$

and using (4.1.21) □

In view of this last result we prove, through an immediate corollary of proposition 4.1.2, the non-degeneracy of $\hat{\Psi}_\epsilon$ in (4.1.3):

Lemma 4.1.1. *Let $F = (F_1, F_2) \in \mathbb{R}^d \times \mathbb{R}^p$ a non-degenerate function on a domain $B \subseteq \mathbb{R}^d \times \mathbb{R}^p$, then for any fixed ϵ in $\mathbb{R} \setminus \{0\}$, $(F_1, \epsilon F_2)$ is non-degenerate on B.*

Proof Let b be any chosen point in B, since F is non-degenerate, by proposition 4.1.2 there exist $j_1, j_2, \ldots, j_{d+p} \in \mathbb{N}^{d+p}$ such that

$$\det\left[F^{(j_1)}(b), \ldots, F^{(j_{d+p})}(b)\right] \neq 0$$

that is

$$\det \begin{pmatrix} F_1^{(j_1)} & F_2^{(j_1)} \\ \vdots & \vdots \\ F_1^{(j_{d+p})} & F_2^{(j_{d+p})} \end{pmatrix} \neq 0\,.$$

Now observe that for every $\epsilon \in \mathbb{R}$ it results

$$\det \begin{pmatrix} F_1^{(j_1)} & \epsilon F_2^{(j_1)} \\ \vdots & \vdots \\ F_1^{(j_{d+p})} & \epsilon F_2^{(j_{d+p})} \end{pmatrix} = \epsilon^p \det \begin{pmatrix} F_1^{(j_1)} & F_2^{(j_1)} \\ \vdots & \vdots \\ F_1^{(j_{d+p})} & F_2^{(j_{d+p})} \end{pmatrix}$$

which proves, together with proposition 4.1.2 and $\epsilon \neq 0$ the statement $\square$

Proposition 4.1.4. *Under condition (4.1.19) the frequency application $\hat{\Psi}_\epsilon$ (see (4.1.3)) of the integrable part of H_ϵ^6 (that is F_ϵ in (4.1.1)) is non-degenerate in the sense of Rüßmann.*

Proof The proof is immediately obtained by proposition 4.1.3, which gives the non degeneracy of Ψ_ϵ, and lemma 4.1.1 applied with $F_j = \Psi_\epsilon^{(j)}$, for $j = 1, 2$, together with definitions (4.1.6) and (4.1.3) $\square$

4.2 Index and amount of non-degeneracy

In the preceding section we proved the non-degeneracy of $\hat{\Psi}_\epsilon$ (defined in (4.1.6) with F_ϵ in (4.1.1)), i.e., the frequency application of the integrable part of the Hamiltonian function H_ϵ^6. This means that H_ϵ^6 meets the main hypothesis in theorem 2.3.1 but since $\hat{\Psi}_\epsilon$ depends on ϵ (for instance it possesses p components of order 1 in ϵ), a direct application of Rüßmann's theorem for maximal tori is not possible unless the perturbative part of H_ϵ^6 is of a certain higher order in ϵ than simply 1. Actually, we showed that the perturbation of H_ϵ^6 (see (3.5.10)) can be moved to order $N \geq 2$ in ϵ, where N can be chosen arbitrarily and fixed at the beginning of the process that conjugates the initial Hamiltonian H_ϵ to H_ϵ^6. Therefore we are now concerned in establishing a suitable value for N so that it is possible to apply Rüßmann's theorem to H_ϵ^6 with $y = (r, \rho)$ and $x = (\theta, \zeta)$, finding maximal tori for H_ϵ as a consequence. We shall see how each quantity involved in the estimate of ϵ_0, the size of the perturbation in Rüßmann's theorem for Lagrangian tori, changes order in ϵ when $\hat{\Psi}_\epsilon$ is considered instead of a frequency application independent from ϵ so to be able to determine a priori a lower bound to impose on N (or equivalently on N_1 and N_2). In this section we focus on the index and amount of non-degeneracy.

Let $\mathcal{K} \subset \mathbb{R}^n$ be a compact set, $B \subseteq \mathbb{R}^n$ a domain containing $\mathcal{K}$ and $f : y \in B \longrightarrow \mathbb{R}^m$ a real-analytic and non-degenerate function. Moreover, let $\mathcal{S} = \{c \in \mathbb{R}^n \,:\, |c|_2 = 1\}$ and $F(c, y) : \mathcal{S} \times B \longrightarrow \mathbb{R}$ the following function

$$F(c, y) = |\langle c, f \rangle|^2 . \tag{4.2.1}$$

We observe that $F(c, y)$ is a function in the form considered in lemma 2.2.1 (equation (2.2.1)) with $N = 2$, $m_1 = m_2 = m$ and $f_1 = f_2 = f$, with the only difference that the parameters c_1, c_2 are not independently varying in $\mathcal{S} \times \mathcal{S} \subseteq \mathbb{R}^n \times \mathbb{R}^n$ but have been chosen to coincide; however this will not influence our purposes as it might be easily noticed. If we define now numbers

$$\beta(f, \mu, \mathcal{K}) := \min_{y \in \mathcal{K}, |c|_2 = 1} \max_{0 \leq \nu \leq \mu} |D^\nu F(c, y)| \tag{4.2.2}$$

verifying obviously

$$\beta(f,0,\mathcal{K}) \leq \beta(f,1,\mathcal{K}) \leq \dots\,, \tag{4.2.3}$$

we stated in lemma 2.2.1 (actually for a much wider class of functions) that there exists a first integer $\mu_0 = \mu_0(f,\mathcal{K})$ such that

$$\beta(f,\mu_0(f,\mathcal{K}),\mathcal{K}) > 0\,; \tag{4.2.4}$$

we called this integer the **index of non-degeneracy** of f with respect to $\mathcal{K}$ while $\beta(f,\mu_0(f,\mathcal{K}),\mathcal{K})$, denoted with an abuse of notation $\beta(f,\mathcal{K})$, is the **amount of non-degeneracy** of f with respect to $\mathcal{K}$.

We now consider the real-analytic application Ψ in (4.1.2) and fix a compact and convex set $\overline{\mathcal{K}} \subset \mathbb{R}^d$ with positive d-dimensional Lebesgue measure and a number $\vartheta_1 \in (0,1)$ such that

$$r_0 = I_0 \in \overline{\mathcal{K}} \qquad \text{and} \qquad \overline{\mathcal{K}} + \vartheta_1 \subset D^d_{\frac{r}{10}}(I_0) \tag{4.2.5}$$

(which implies $\overline{\mathcal{K}} +_{\mathbb{R}} \vartheta_1 \subset D^d_{\frac{r}{10}}(I_0) \cap \mathbb{R}^d$). Then we define

$$\begin{aligned} \bar\mu &= \mu_0(\Psi,\overline{\mathcal{K}}) = \textit{index of non-degeneracy of } \Psi \textit{ with respect to } \overline{\mathcal{K}} \\ \bar\beta &= \beta(\Psi,\overline{\mathcal{K}}) = \textit{amount of non-degeneracy of } \Psi \textit{ with respect to } \overline{\mathcal{K}}\,. \end{aligned}$$

In this section our aim is to see how the index and amount of non-degeneracy of $\hat\Psi_\epsilon$ are related to $\bar\mu$ and $\bar\beta$, with respect to suitable compact sets in their domains of definition. For further details on the index and amount of non-degeneracy refer to lemma C.1.2, definition C.1.3 and proposition C.1.2, where in this last issue only the condition with $j = 1$ has to be considered since we take $\omega = \Psi$ and $A = 0$ establishing the correspondence

$$[\chi]^{(1)}_k = |k|_2^{-2} |\langle k, \Psi\rangle|^2\,.$$

Let $\Psi^0_\epsilon(r) = \Psi(r_0 + \epsilon r)$ as in (4.1.10) (we remind $r_0 = I_\star$ and $\epsilon < 1$) and

$$\mathcal{K}_0 := \left\{ r \in \mathbb{R}^d \;:\; r_0 + r \in \overline{\mathcal{K}} \right\} \tag{4.2.6}$$

it results

$$\begin{aligned} &\min_{r\in\mathcal{K}_0} \max_{0\leq\nu\leq\bar\mu} \left| \frac{\partial^\nu}{\partial r} |\langle c, \Psi^0_\epsilon(r)\rangle|^2 \right| = \min_{r\in\mathcal{K}_0} \max_{0\leq\nu\leq\bar\mu} \left| \frac{\partial^\nu}{\partial r} |\langle c, \Psi(r_0+\epsilon r)\rangle|^2 \right| = \\ &= \min_{r\in\mathcal{K}_0} \max_{0\leq\nu\leq\bar\mu} \epsilon^\nu \left| \frac{\partial^\nu |\langle c, \Psi(r_0+\epsilon r)\rangle|^2}{\partial r} \right| \geq \min_{r\in\overline{\mathcal{K}}} \max_{0\leq\nu\leq\bar\mu} \epsilon^\nu \left| \frac{\partial^\nu |\langle c, \Psi(r)\rangle|^2}{\partial r} \right| \geq \\ &\geq \epsilon^{\bar\mu} \min_{r\in\overline{\mathcal{K}}} \max_{0\leq\nu\leq\bar\mu} \left| \frac{\partial^\nu |\langle c, \Psi(r)\rangle|^2}{\partial r} \right| = \epsilon^{\bar\mu}\bar\beta > 0 \end{aligned}$$

where we used the convexity of $\mathcal{K}_0$ (deriving from the convexity of $\overline{\mathcal{K}}$) and $0 \in \mathcal{K}_0$ (since $r_0 \in \overline{\mathcal{K}}$) to obtain $r \in \mathcal{K}_0$ implies $r_0 + \epsilon r \in \overline{\mathcal{K}}$. Therefore, denoting

$$\mu_\epsilon^0 = \mu_0(\Psi_\epsilon^0, \mathcal{K}_0) = \textit{ the index of non-degeneracy of } \ \Psi_\epsilon^0 \textit{ with respect to } \mathcal{K}_0$$

we have just proved

$$\mu_\epsilon^0 \leq \bar{\mu} \qquad \text{and} \qquad \beta(\Psi_\epsilon^0, \bar{\mu}, \overline{\mathcal{K}}) \geq \epsilon^{\bar{\mu}} \bar{\beta} \tag{4.2.7}$$

according to the definition in (4.2.2).

Consider now Ψ_ϵ as in (4.1.6), with F_ϵ in (4.1.1) and its derivatives computed in (4.1.4) and (4.1.5). Let $\mathcal{K}_1 \subset \mathbb{R}^p$ be an arbitrarily chosen and fixed compact set such that

$$0 \in \mathcal{K}_1 \qquad \text{and} \qquad \mathcal{K}_1 \subset B^p_{\frac{\epsilon}{4}} \, ; \tag{4.2.8}$$

then we define

$$\mathcal{S} := \mathcal{K}_0 \times \mathcal{K}_1 \tag{4.2.9}$$

where $\mathcal{K}_0$ is in (4.2.6) for $\overline{\mathcal{K}}$ verifying (4.2.5). Notice that his compact set $\mathcal{S}$ has nothing to do with the domain of the parameters c in the previously reminded definition of index and amount of non-degeneracy (in fact, from now on we will explicitly write the space of parameters each time they appear).

Now, we denote

$$\mu_\epsilon = \mu_0(\Psi_\epsilon, \mathcal{S}) = \textit{index of non-degeneracy of } \ \Psi_\epsilon \textit{ with respect to } \mathcal{S} \, .$$

and state a preliminary lemma:

Lemma 4.2.1. *Let*

$$B(x) = (b_{i_1 \dots i_m}(x))_{1 \leq i_1, \dots, i_m \leq n_1+n_2} \in \left(\mathbb{R}^{n_1+n_2}\right)^m$$

a tensor of order m in $\mathbb{R}^{n_1+n_2}$ defined for x in a compact subset $\mathcal{S}$ of $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ (for instance B is an m-th derivative of an m-times continuously differentiable function from $\mathcal{S}$ to $\mathbb{R}$). Consider

$$B'(x) = (b_{i_1 \dots i_m}(x))_{1 \leq i_1, \dots, i_m \leq n_1} \in \left(\mathbb{R}^d\right)^m$$

then

$$|B(x)| \geq |B'(x)|$$

for every $x \in \mathcal{S}$, where the norms of the tensors are the same defined in 0.1.4 for the m-th derivatives of m-times continuously differentiable functions.

Proof Let $v \in \mathbb{R}^{n_1}$ with $|v|_2 = 1$ such that

$$|B'(x)| := \max_{a \in \mathbb{R}^d, |a|_2 = 1} |B'(x)\,(a^m)\,| = |B'(x)\,(v^m)\,|$$

set $w=(v,0,\ldots,0)\in\mathbb{R}^{n_1}\times\mathbb{R}^{n_2}$; obviously $|w|_2=1$ therefore

$$|B(x)|:=\max_{a\in\mathbb{R}^{n_1+n_2},|a|_2=1}|B(x)\,(a^m)\,|\geq|B(x)\,(w^m)\,|=$$

$$=\left|\sum_{1\leq i_1,\ldots,i_m\leq n_1+n_2} b_{i_1 i_2\ldots i_m}w_{i_1 i_2\ldots i_m}\right|=\left|\sum_{1\leq i_1,\ldots,i_m\leq n_1} b_{i_1 i_2\ldots i_m}w_{i_1 i_2\ldots i_m}\right|=$$

$$=|B'(x)\,(v^m)\,|=|B'(x)|$$

where we used $w_{i_j}=0$ for any $i_j\geq n_1+1$ $\square$

We are now ready to obtain the following result

Proposition 4.2.1. *If we assume*

$$\epsilon\leq\frac{\bar\beta}{4\Delta}\min\left\{\frac{1}{2(M_0+M_2)},\frac{1}{\Delta}\right\}\frac{\vartheta_1^{\bar\mu}}{\bar\mu!} \tag{4.2.10}$$

then it results

$$\mu_\epsilon\leq\bar\mu\qquad\textit{and}\qquad\beta(\Psi_\epsilon,\bar\mu,\mathcal{S})\geq\frac{\epsilon^{\bar\mu}\bar\beta}{2}\,. \tag{4.2.11}$$

Proof By inequality in (4.1.7) we have $\Psi_\epsilon=\Psi+\epsilon G$ with

$$|G(r_0+\epsilon r,\rho^0+\epsilon\rho)|\leq\Delta$$

for any $(r,\rho)\in D^d_{\frac{r}{10}}\times D^p_{\frac{\epsilon}{2}}$. Thus, for any $0\leq\nu\leq\bar\mu$ it results

$$\left|\frac{\partial^\nu}{\partial r}\left|\langle c,\Psi_\epsilon(r+\epsilon r,\rho^0+\epsilon\rho)\rangle\right|^2\right|=\epsilon^\nu\left|\frac{\partial^\nu|\langle c,\Psi+\epsilon G\rangle|^2}{\partial r}\right|=$$

$$=\epsilon^\nu\left|\frac{\partial^\nu|\langle c,\Psi\rangle|^2}{\partial r}+\epsilon\frac{\partial^\nu\,|\langle c,\Psi\rangle\langle c,G\rangle|}{\partial r}+\epsilon^2\frac{\partial^\nu|\langle c,G\rangle|^2}{\partial r}\right|\geq$$

$$\geq\epsilon^\nu\left|\frac{\partial^\nu|\langle c,\Psi\rangle|^2}{\partial r}\right|-2\epsilon^{\nu+1}\max_{(r,\rho)\in\mathcal{S}}\left|\frac{\partial^\nu\,|\langle c,\Psi\rangle\langle c,G\rangle|}{\partial r}\right|+$$

$$-\epsilon^{\nu+2}\max_{(r,\rho)\in\mathcal{S}}\left|\frac{\partial^\nu|\langle c,G\rangle|^2}{\partial r}\right|$$

where the omitted arguments of Ψ and G are $(r_0+\epsilon r)$ and $(r_0+\epsilon r,\rho^0+\epsilon\rho)$ respectively, and we agree to use this notation during this proof. Now using Cauchy's estimate and the uniform estimates on the norms of Ψ and G we obtain for any $c \in \mathbb{R}^{d+p}$ with $|c|_2 = 1$

$$\max_{(r,\rho)\in\mathcal{S}} \left| \frac{\partial^\nu \left|\langle c, \Psi\rangle\langle c, G\rangle\right|}{\partial r} \right| \leq \frac{\nu!}{\vartheta_1^\nu} \sup_{(\mathcal{K}_0+\vartheta_1)\times\mathcal{K}_1} |\langle c, \Psi\rangle\langle c, G\rangle| \leq$$

$$\leq \frac{\nu!}{\vartheta_1^\nu} \sup_{\mathcal{K}_0+\vartheta_1} |\Psi| \sup_{(\mathcal{K}_0+\vartheta_1)\times\mathcal{K}_1} |G| \leq \frac{\nu!}{\vartheta_1^\nu}(M_0+M_2)\Delta \leq \frac{\bar{\mu}!}{\vartheta_1^{\bar{\mu}}}(M_0+M_2)\Delta$$

and analogously

$$\max_{(r,\rho)\in\mathcal{S}} \left| \frac{\partial^\nu |\langle c, G\rangle|^2}{\partial r} \right| \leq \frac{\nu!}{\vartheta_1^\nu} \sup_{(\mathcal{K}_0+\vartheta_1)\times\mathcal{K}_1} |\langle c, G\rangle|^2 \leq \frac{\nu!}{\vartheta_1^\nu} \sup_{(\mathcal{K}_0+\vartheta_1)\times\mathcal{K}_1} |G|^2 \leq \frac{\bar{\mu}!}{\vartheta_1^{\bar{\mu}}}\Delta^2$$

where we have used $0 \leq \nu \leq \bar{\mu}$ and $\vartheta_1 < 1$. Substituting this last two estimates at the end of the first chain of inequalities and using hypothesis (4.2.10), it results

$$\left| \frac{\partial^\nu}{\partial r} \left|\langle c, \Psi_\epsilon(r_0+\epsilon r, \rho^0+\epsilon\rho)\rangle\right|^2 \right| \geq \epsilon^\nu \left| \frac{\partial^\nu |\langle c, \Psi\rangle|^2}{\partial r} \right| +$$

$$-2\epsilon^{\nu+1}\frac{\bar{\mu}!}{\vartheta_1^{\bar{\mu}}}(M_0+M_2)\Delta - \epsilon^{\nu+2}\frac{\bar{\mu}!}{\vartheta_1^{\bar{\mu}}}\Delta^2 \geq \epsilon^\nu \left| \frac{\partial^\nu |\langle c, \Psi\rangle|^2}{\partial r} \right| - \frac{\epsilon^\nu \bar{\beta}}{2}$$

for any $0 \leq \nu \leq \bar{\mu}$, $(r,\rho) \in \mathcal{S}$ and $c \in \mathbb{R}^{d+p}$ with $|c|_2 = 1$.

Therefore we obtain

$$\min_{(r,\rho)\in\mathcal{S}} \max_{0\leq\nu\leq\bar{\mu}} \left| \frac{\partial^\nu |\langle c, \Psi_\epsilon(r_0+\epsilon r, \rho^0+\epsilon\rho)\rangle|^2}{\partial r} \right| \geq$$

$$\geq \min_{(r,\rho)\in\mathcal{S}} \max_{0\leq\nu\leq\bar{\mu}} \epsilon^\nu \left\{ \left| \frac{\partial^\nu |\langle c, \Psi\rangle|^2}{\partial r} \right| - \frac{\bar{\beta}}{2} \right\} \geq$$

$$\geq \epsilon^{\bar{\mu}} \left\{ \min_{r\in\mathcal{K}_0} \max_{0\leq\nu\leq\bar{\mu}} \left| \frac{\partial^\nu |\langle c, \Psi\rangle|^2}{\partial r} \right| - \frac{\bar{\beta}}{2} \right\} \geq \frac{\epsilon^{\bar{\mu}}\bar{\beta}}{2} > 0 \qquad (4.2.12)$$

for any $c \in \mathbb{R}^{d+p}$ with $|c|_2 = 1$. We now apply lemma 4.2.1 with $x = (r,\rho)$, $n_1 = d, n_2 = p$, $m = \nu = 1, 2, \ldots, \bar{\mu}$ (in the case $\nu = 0$ there nothing to prove since there are no derivatives) and

$$B_\nu = \frac{\partial^\nu |\langle c, \Psi_\epsilon(r,\rho)\rangle|^2}{\partial r \partial \rho} \in \left(\mathbb{R}^{d+p}\right)^\nu$$

$$B'_\nu = \frac{\partial^\nu |\langle c, \Psi_\epsilon(r,\rho)\rangle|^2}{\partial r} \in \left(\mathbb{R}^{d}\right)^\nu$$

in order to obtain

$$\left|\frac{\partial^\nu |\langle c, \Psi_\epsilon(r,\rho)\rangle|^2}{\partial r \partial \rho}\right| \geq \left|\frac{\partial^\nu |\langle c, \Psi_\epsilon(r,\rho)\rangle|^2}{\partial r}\right|$$

for any $0 \leq \nu \leq \bar{\mu}$, $(r,\rho) \in \mathcal{S}$ and $c \in \mathbb{R}^{d+p}$ with unitary Euclidean norm. This inequality, together with (4.2.12), yields

$$\min_{(r,\rho)\in\mathcal{S}} \max_{0\leq\nu\leq\bar{\mu}} \left|\frac{\partial^\nu |\langle c, \Psi_\epsilon(r_0+\epsilon r, \rho^0+\epsilon\rho)\rangle|^2}{\partial r \partial \rho}\right| \geq$$

$$\geq \min_{(r,\rho)\in\mathcal{S}} \max_{0\leq\nu\leq\bar{\mu}} \left|\frac{\partial^\nu |\langle c, \Psi_\epsilon(r_0+\epsilon r, \rho^0+\epsilon\rho)\rangle|^2}{\partial r}\right| \geq \frac{\epsilon^{\bar{\mu}}\bar{\beta}}{2} > 0$$

which implies the statement □

Now recall that we defined (in equation (4.1.6))

$$\Psi_\epsilon = \left(\Psi_\epsilon^{(1)}, \Psi_\epsilon^{(2)}\right) = \left(\frac{\partial}{\partial r}F_\epsilon, \frac{2}{\epsilon}\frac{\partial}{\partial \rho}F_\epsilon\right) \in \mathbb{R}^d \times \mathbb{R}^p \tag{4.2.13}$$

(where the argument of the functions involved is $(r_0+\epsilon r, \rho^0+\epsilon\rho)$ as usual) and consider

$$\hat{\Psi}_\epsilon = \left(\hat{\Psi}_\epsilon^{(1)}, \hat{\Psi}_\epsilon^{(2)}\right) = \left(\frac{\partial}{\partial r}F_\epsilon, \frac{\partial}{\partial \rho}F_\epsilon\right) \tag{4.2.14}$$

the frequency application of the integrable part of H_ϵ^6. We can trivially observe that $\hat{\Psi}_\epsilon$ is obtain from Ψ_ϵ by rescaling its last p components by a factor $\frac{\epsilon}{2}$; this fact allows to formulate the following results

Proposition 4.2.2. *Let Ψ_ϵ be the real-analytic function for $(r,\rho) \in D^d_{\frac{r}{5}} \times D^p_{\frac{\epsilon}{2}}$ considered before and define*

$$\hat{\mu}_\epsilon = \textit{the index of non-degeneracy of } \hat{\Psi}_\epsilon \textit{ with respect to } \mathcal{S}\,; \tag{4.2.15}$$

then, for any $0<\epsilon<1$ we have

$$\hat{\mu}_\epsilon \leq \bar{\mu} \qquad \textit{and} \qquad \beta(\hat{\Psi}_\epsilon, \bar{\mu}, \mathcal{S}) \geq \frac{\epsilon^{\bar{\mu}+2}\bar{\beta}}{8} \tag{4.2.16}$$

Proof Let $c = (c_1, c_2) \in \mathbb{R}^d \times \mathbb{R}^p$ with $|c|_2 = 1$ define

$$f(r,\rho,c_1,c_2) = \max_{0\leq\nu\leq\bar{\mu}} \left|\frac{\partial^\nu |\langle c, \Psi_\epsilon\rangle|^2}{\partial r\partial\rho}\right| = \max_{0\leq\nu\leq\bar{\mu}} \left|\frac{\partial^\nu \left|\langle c_1, \Psi_\epsilon^{(1)}\rangle + \langle c_2, \Psi_\epsilon^{(2)}\rangle\right|^2}{\partial r\partial\rho}\right|$$

for $(r,\rho) \in D^d_{\frac{r}{5}} \times D^p_{\frac{\epsilon}{2}}$; by definition of $\hat{\Psi}_\epsilon$ we obtain

$$f\left(r,\rho,c_1,\frac{\epsilon}{2}c_2\right) = \max_{0\le\nu\le\bar{\mu}} \left| \frac{\partial^\nu \left|\langle c, \hat{\Psi}_\epsilon(r,\rho)\rangle\right|^2}{\partial r \partial \rho} \right| .$$

Observe now that for any (r,ρ,c_1,c_2) and $t \in \mathbb{R}$ the following equality holds

$$f(r,\rho,tc_1,tc_2) = t^2 f(r,\rho,c_1,c_2) . \tag{4.2.17}$$

Set now for any given $c = (c_1, c_2)$ with $|c|_2 = 1$

$$t_\epsilon := \sqrt{|c_1|_2^2 + \frac{\epsilon^2}{4}|c_2|_2^2} \ \ge\ \frac{\epsilon}{2}$$

and

$$\bar{c}_1 := \frac{c_1}{t_\epsilon} \ , \quad \bar{c}_2 := \frac{\epsilon c_2}{2t_\epsilon}$$

so that it results

$$|\bar{c}_1|_2^2 + |\bar{c}_2|_2^2 = \left(|c_1|^2 + \frac{\epsilon^2}{4}|c_2|^2\right)(t_\epsilon^2)^{-1} = 1 .$$

Then by the definition of f we obtain for any $(r,\rho) \in \mathcal{S}$

$$\max_{0\le\nu\le\bar{\mu}} \left| \frac{\partial^\nu |\langle c, \hat{\Psi}_\epsilon\rangle|^2}{\partial r\partial\rho} \right| = f\left(r,\rho,c_1,\frac{\epsilon}{2}c_2\right) = f\left(r,\rho,t_\epsilon\bar{c}_1,t_\epsilon\bar{c}_2\right) =$$

$$= \ \ t_\epsilon f\left(r,\rho,\bar{c}_1,\bar{c}_2\right) = t_\epsilon^2 \max_{0\le\nu\le\bar{\mu}} \left| \frac{\partial^\nu |\langle(\bar{c}_1,\bar{c}_2), \Psi_\epsilon\rangle|^2}{\partial r\partial\rho} \right| \ge t_\epsilon^2 \frac{\epsilon^{\bar{\mu}}\bar{\beta}}{2} \ge \frac{\epsilon^{\bar{\mu}+2}\bar{\beta}}{8} > 0$$

having used $|(\bar{c}_1,\bar{c}_2)|_2 = 1$; this proves (4.2.16).

4.3 Determination of N and sufficient conditions to apply Rüßmann's theorem

We now refer to the estimate of ϵ_0 (the admissible size of the perturbation in Rüßmann theorem) given in section 2.4 and analyze each quantity involved, focusing in particular on how they change order in ϵ when H^6_ϵ is considered. Our aim is to apply the above mentioned estimate to the perturbative part of H^6_ϵ, that is P_ϵ as it appears in (3.5.14). Recall that we estimated

$$|P_\epsilon|_{\mathcal{D}_5} \le (M'_{N_2} + \mu^{N+1})\epsilon^N \tag{4.3.1}$$

where

$$\mathcal{D}_5 = \mathbb{T}^d_{\frac{\sigma}{8}} \times D^d_{\frac{r}{5}} \times \mathbb{T}^p_{\upsilon} \times D^p_{\frac{|\rho^0|}{2}}$$

with υ defined in (3.5.2), M'_{N_2} in (3.5.12) (together with (3.4.47), (3.3.19) and (3.3.24)), μ in (3.2.21), $N \geq 2$ is an arbitrarily fixed integer and $0 < \epsilon < 1$. Thus, understand how ϵ_0 depends on ϵ will make possible a suitable choice of N imposing

$$(M'_{N_2} + \mu^{N+1})\epsilon^N \leq \frac{1}{2}\epsilon_0 \,. \tag{4.3.2}$$

We take into consideration the Hamiltonian function H^6_ϵ (which plays the role of H in theorem 2.3.1) real-analytic for $(\theta, r, \zeta, \rho) \in \mathbb{T}^d \times B^d_{\frac{r}{5}} \times \mathbb{T}^p \times B^p_\epsilon$, where

$$B^n_t := D^n_t \cap \mathbb{R}^n = \{x \in \mathbb{R}^n \,:\, |x| < t\} \tag{4.3.3}$$

(see notation (3.5.4)), having holomorphic expansion on $\mathcal{D}_5$. As a consequence, with refer to the statement of theorem 2.3.1, we have the following correspondences

$$\begin{aligned} &\mathbb{T}^n = \mathbb{T}^d \times \mathbb{T}^p \,, \quad B = B^d_{\frac{r}{5}} \times B^p_\epsilon \,, \quad D = \mathcal{D}_5 \\ &\mathcal{A} = \mathcal{D}_5 \,, \qquad x = (r, \rho) \,, \qquad y = (\theta, \zeta) \,, \\ &N = F_\epsilon \,, \qquad P = P_\epsilon \,. \end{aligned} \tag{4.3.4}$$

4.3.1 Choice of the initial compact set

First of all, we fix

$$\mathcal{K}_1 := B^p_{\frac{\epsilon}{5}} = \left\{x \in \mathbb{R}^p \,:\, |x| < \frac{\epsilon}{5}\right\} \tag{4.3.5}$$

accordingly to condition 4.2.8. Now, we need to fix a compact set $\mathcal{K} \subset B^d_{\frac{r}{5}} \times B^p_\epsilon$, with positive $(d+p)$-dimensional Lebesgue measure, and a number $\vartheta \in (0, 1)$ satisfying (2.3.3). We take $\mathcal{K} = \mathcal{S}$, where

$$\mathcal{S} := \mathcal{K}_0 \times \mathcal{K}_1$$

is taken as in definition (4.2.9), with $\mathcal{K}_0$ verifying

$$\mathcal{K}_0 + \vartheta_1 \subset D^d_{\frac{r}{10}} \tag{4.3.6}$$

(see (4.2.6) and (4.2.5)) and the arbitrary choice of $\mathcal{K}_1$ is replaced by the above definition. Then a suitable value for ϑ could be

$$\vartheta := \frac{\epsilon}{16} \,. \tag{4.3.7}$$

In fact, assuming $\epsilon \le 4\vartheta_1$ (which is implied for sufficiently small $\bar{\beta}$ by inequality (4.4.8)), it results

$$\begin{aligned}\mathcal{S} + 4\vartheta &= (\mathcal{K}_0 \times \mathcal{K}_1) + 4\vartheta \subseteq (\mathcal{K}_0 + 4\vartheta) \times (\mathcal{K}_1 + 4\vartheta) = \\ &= \left(\mathcal{K}_0 + \frac{\epsilon}{4}\right) \times \left(\mathcal{K}_1 + \frac{\epsilon}{4}\right) \subset (\mathcal{K}_0 + \vartheta_1) \times \left(\mathcal{K}_1 + \frac{\epsilon}{4}\right) \subset \\ &\subset D^d_{\frac{r}{10}} \times D^p_{\frac{\epsilon}{2}} \,. \end{aligned} \tag{4.3.8}$$

Moreover, to completely satisfy inclusion (2.3.3), we need to verify $\mathbb{T}^{d+p} + \vartheta \subseteq \mathbb{T}^d_{\frac{\sigma}{8}} \times \mathbb{T}^p_{\upsilon}$. For this purpose we assume $\epsilon \le \sigma$ and observe that by (3.5.1) and (3.5.8) we have $2\epsilon \le \frac{1}{3}(r')^2$ which trivially implies $2 \le \frac{1}{3\epsilon}(r')^2$; thus, using the definition of υ in (3.5.2) with (3.5.8) and $\epsilon \le 1 \le 4\log 2$, we obtain

$$\frac{\epsilon}{16} \le \frac{1}{4}\log 2 \le \frac{1}{4}\log\frac{(r')^2}{3\epsilon} = \frac{1}{4}\log\frac{2(r')^2}{3|\rho^0|} = \upsilon \,.$$

Therefore we can conclude

$$\mathbb{T}^{d+p} + \vartheta \subseteq (\mathbb{T}^d + \vartheta) \times (\mathbb{T}^p + \vartheta) = \left(\mathbb{T}^d + \frac{\epsilon}{16}\right) \times \left(\mathbb{T}^p + \frac{\epsilon}{16}\right) \subseteq \mathbb{T}^d_{\frac{\sigma}{16}} \times \mathbb{T}^p_{\upsilon} \tag{4.3.9}$$

which gives, together with the inclusion obtained before,

$$(\mathbb{T}^{d+p} + \vartheta) \times (\mathcal{S} + 4\vartheta) \subseteq \mathcal{A} \,. \tag{4.3.10}$$

where the choice of ϑ has been made in (4.3.7).

4.3.2 Parameters related to the initial compact set

The diameter of $\mathcal{S} = \mathcal{K}_0 \times \mathcal{K}_1$ is given by $d_{\mathcal{S}} := \sup_{x,y\in\mathcal{S}} |x - y|$ and verifies

$$d_{\mathcal{S}} \le \text{diam}\,\mathcal{K}_0 + \text{diam}\,\mathcal{K}_1 \le \frac{1}{5}r + \frac{2}{5}\epsilon \,.$$

By (2.4.5) and (2.4.4) we notice that the estimate given in (2.4.8) is decreasing in d_0 (that is the diameter of the compact set $\mathcal{K}$ in theorem 2.3.1); therefore we may take

$$d_0 := r + 1 \tag{4.3.11}$$

in the definition of ϵ_0.

Now, referring to theorem 2.3.1 we have to choose and fix a parameter $0 < \epsilon^\star <$ meas$_{d+p}\mathcal{S}$. A well-known formula for the volume of a p-ball with fixed radius gives

$$\text{meas}_p\mathcal{K}_1 = \text{meas}_p B^p_{\frac{\epsilon}{5}} = c^\star_p \epsilon^p \qquad \text{with} \qquad c^\star_p := \frac{2\pi^{\frac{p}{2}}}{5^p p}\,\Gamma\left(\frac{p}{2}\right)^{-1} . \tag{4.3.12}$$

Therefore, if we consider the compact set $\mathcal{K}_0$ (which is directly determined by the choice of the initial compact and convex set $\overline{\mathcal{K}}$ in (4.2.6)), we have to fix $\epsilon_0^\star$ such that

$$0 < \epsilon_0^\star < c_p^\star \operatorname{meas} \mathcal{K}_0 = c_p^\star \operatorname{meas} \overline{\mathcal{K}} \tag{4.3.13}$$

and consider

$$\epsilon^\star = \epsilon_0^\star \epsilon^p \tag{4.3.14}$$

in the estimate for the size of the perturbation of H_ϵ^6.

4.3.3 Choice of index and amount of non-degeneracy

The frequency application of the integrable part of H_ϵ^6 is the real-analytic function $\hat{\Psi}_\epsilon(r_0 + \epsilon r, \rho^0 + \epsilon\rho)$ for

$$(r, \rho) \in D^d_{\frac{r}{5}} \times D^p_\epsilon \tag{4.3.15}$$

defined in (4.1.3) (see also (4.1.1), (4.1.4) and (4.1.5)). In section 4.1 we proved that $\hat{\Psi}_\epsilon$ is non-degenerate, for small enough ϵ (condition (4.1.19)). Moreover, let $\hat{\mu}_\epsilon$ denote the index of non-degeneracy of Ψ_ϵ with respect to $\mathcal{S}$, $\bar{\mu}$ and $\bar{\beta}$ denote respectively the index and amount of non-degeneracy of Ψ (defined in (4.1.2)) with respect to the compact set $\overline{\mathcal{K}} \subseteq B_{\frac{r}{5}}(I_\star)$ initially considered, we proved

$$\hat{\mu}_\epsilon \leq \bar{\mu} \qquad \text{and} \qquad \beta(\hat{\Psi}_\epsilon, \bar{\mu}, \mathcal{S}) \geq \frac{\epsilon^{\bar{\mu}+2}\bar{\beta}}{8} \tag{4.3.16}$$

(see section 4.2 for definitions and details).

Remark 4.3.1. *In section 2.5 we showed how it is possible to construct an iterative process to obtain non-resonant frequencies as stated in theorem 2.5.1. Now observe that the numbers $\tau^\star, g^\star, L_1^\star, L_2^\star$ appearing in the cited theorem can be found as function of $\mu_0 \in \mathbb{N}_+$ and $\beta > 0$ verifying (2.5.12). Two such numbers can be obviously the index and amount of non-degeneracy of $\hat{\Psi}_\epsilon$. However this is not necessary: in fact, as section 2.5 clearly explains, it is possible to choose any integer $\bar{\mu} \geq \mu_0(\hat{\Psi}_\epsilon, \mathcal{S})$ (= index of non-degeneracy of $\hat{\Psi}_\epsilon$ with respect to $\mathcal{S}$) and and any $0 < \bar{\beta} \leq \beta(\hat{\Psi}_\epsilon, \bar{\mu}, \mathcal{S})$ to obtain the same results. Therefore, accordingly to (4.3.16), we can take*

$$\mu_0 = \bar{\mu} \qquad \textit{and} \qquad \beta = \frac{\epsilon^{\bar{\mu}+2}\bar{\beta}}{8} \tag{4.3.17}$$

in the estimate for the size of the perturbation.

Remark 4.3.2. *Apparently the choice of $\bar{\mu} \geq \hat{\mu}_\epsilon$ and the associated value $\beta \leq \beta(\hat{\Psi}_\epsilon, \bar{\mu}, \mathcal{S})$ could increase the estimated value of ϵ_0; however, this advantage brings the necessity to compute a greater number of derivatives when an explicit value for ϵ_0 is searched.*

4.3.4 Norm of the frequency application

Since we defined

$$\begin{aligned} \Psi_\epsilon &= \left(\Psi_\epsilon^{(1)}, \Psi_\epsilon^{(2)}\right) = \left(\frac{\partial}{\partial r} F_\epsilon, \frac{2}{\epsilon}\frac{\partial}{\partial \rho} F_\epsilon\right) \\ \hat{\Psi}_\epsilon &= \left(\hat{\Psi}_\epsilon^{(1)}, \hat{\Psi}_\epsilon^{(2)}\right) = \left(\frac{\partial}{\partial r} F_\epsilon, \frac{\partial}{\partial \rho} F_\epsilon\right) \end{aligned}$$

(where the argument of all the functions considered is $(r_0 + \epsilon r, \rho^0 + \epsilon\rho)$) it results

$$\sup_{\mathcal{S}+3\vartheta} |\hat{\Psi}_\epsilon(r + \epsilon r, \rho^0 + \epsilon\rho)| \leq \sup_{\mathcal{S}+3\vartheta} |\Psi_\epsilon(r + \epsilon r, \rho^0 + \epsilon\rho)| \,. \tag{4.3.18}$$

Furthermore by (4.3.8), lemma 4.1.1 and the definition of $\mathcal{K}_0$ in (4.2.6) we have

$$\sup_{\mathcal{S}+3\vartheta} |\Psi_\epsilon(r + \epsilon r, \rho^0 + \epsilon\rho)| \leq \sup_{D^d_{\frac{r}{10}} \times D^p_{\frac{\epsilon}{2}}} |\Psi_\epsilon(r + \epsilon r, \rho^0 + \epsilon\rho)| \leq$$

$$\leq \sup_{D^d_{\frac{r}{10}}} |\Psi(r_0 + \epsilon r)| + \Delta\epsilon \leq M_0 + M_2 + \epsilon\Delta \,.$$

Therefore, in the estimate for ϵ_0 we may take

$$C_1 := M_0 + M_2 + \Delta \tag{4.3.19}$$

as upper bound for $|\hat{\Psi}_\epsilon|_{\mathcal{S}+3\vartheta}$.

4.3.5 Approximation function and control on the small divisors

In his work [Rüßm01] Rüßmann uses what he calls an “approximation function” to control the small denominators appearing in the problem with diophantine inequalities of the form

$$|\omega \cdot k| \geq \gamma\Phi(|k|_2)$$

where $\omega = (\omega_1, \omega_2, \cdot, \omega_n)$ is the “frequency map”. We remark that $\omega \cdot k$ are the only kind of small divisors appearing in the problem when we consider the maximal case, i.e $p = 0 = q$ with refer to Rüßmann’s notation at the beginning of his work [Rüßm01, page 123].

On page 125 of his work, Rüßmann defines an approximation function as a continuous function $\Phi : [0, \infty] \longrightarrow \mathbb{R}$ verifying the following four properties

1. $1 = \Phi(0) \geq \Phi(s) \geq \Phi(t) > 0$ for $0 \leq s < t < \infty$;

2. $\Phi(1) = 1$ so that $\Phi(s) = 1$ for any $0 \leq s \leq 1$;

3. $s^\lambda \Phi(s) \overset{s\to\infty}{\longrightarrow} 0$ for any $\lambda \geq 0$ so that

$$\Phi_{\lambda\mu} := \sup_{s\geq 1} s^\lambda \Phi(s)^{\frac{1}{\mu}} < \infty \tag{4.3.20}$$

   for all $\mu > 0$ and $\lambda \geq 0$;

4. the following integral is finite

$$\int_1^\infty \log \frac{1}{\Phi(s)} \frac{ds}{s^2} < \infty\,. \tag{4.3.21}$$

However this definition of approximation function would create an unsurmountable problem in the application of Rüßmann's theorem to H^6_ϵ. Without going into details, we just observe that the choice of $\vartheta = O(\epsilon)$ as in (4.3.7) would force the choice of $T_0 = O(\epsilon^{-\alpha})$ for some $\alpha > 0$ in order fulfill the second condition in (2.4.1). Then, in view of condition 4.3.5.3, $\Phi(T_0)$ would be of order greater than ϵ^b for any $b \in \mathbb{N}$. Now, since we have a perturbation of order ϵ^N for a fixed $N \in \mathbb{N}$ (see (4.3.1)) and the definition of ϵ_0 in (2.4.8) together with (2.4.9) requires $\epsilon_0 \leq C\Phi(T_0)$, we reach a contradiction.

To overcome this problem we claim, and prove next, that the control on the small divisors can be performed with a function in the form

$$\widetilde{\Phi}(s) = \begin{cases} 1 & 0 \leq s \leq 1 \\ \\ s^{-a} & s \geq 1 \end{cases} \tag{4.3.22}$$

with

$$a > (d+p)\bar{\mu} \tag{4.3.23}$$

instead that with an approximation function as defined by Rüßmann.

First of all observe that $\widetilde{\Phi}$ is continuous and trivially verifies properties 4.3.5.1, 4.3.5.2 and 4.3.5.3 while it does not fulfill condition 4.3.5.3.

Remark 4.3.3. *In Rüßmann's work the property 4.3.5.3 of the approximation function Φ is used only in the second inequality on top of page 174 and in the chain of inequalities at the beginning of page 193 (inequality (2.5.34) in this thesis).*

Lemma 4.3.1. *The two inequalities just cited can be obtained with $\widetilde{\Phi}$ in (4.3.22) (and a in (4.3.23)) instead of Φ.*

Proof On page 174 of [Rüßm01] we find the following inequalities

$$T_\nu^2 \delta^{(\kappa-2\psi)\nu} \le T_\nu^{(\lambda+1)\frac{\kappa-2\psi}{\tau}} \delta^{(\kappa-2\psi)\nu} \le \left(\frac{\Phi_{\lambda 1}}{\Psi(T_0)}\right)^{\frac{\kappa-2\psi}{\tau}} \tag{4.3.24}$$

where δ, Ψ and T_ν, with $\nu \in \mathbb{N}$, are defined in subsection 2.5.1, $\tau := \frac{1}{18\bar{\mu}}$ (see [Rüßm01, 14.10.6] accordingly to (2.5.2) with $\tau^\star$ in (2.5.13) and $\tau_0 = \frac{1}{9}$ as fixed in (13.1) by Rüßmann), $\psi := \frac{4}{9}$, $\kappa := \frac{24}{25}$, $T_0 \ge 1$ verifies (4.3.27), $\Phi_{\lambda\mu}$ is defined in (4.3.20) and λ is sufficiently large. The inequality formed by the first and the last member in (4.3.24) is used immediately after to obtain

$$\sum_{\nu=0}^{\infty} T_\nu \delta^{(\kappa-2\psi)\nu} \le \left(\frac{\Phi_{\lambda 1}}{\Psi(T_0)}\right)^{\frac{\kappa-2\psi}{\tau}} \sum_{\nu=0}^{\infty} \frac{1}{T_\nu} < \infty \tag{4.3.25}$$

where the convergence of this last sum is proved in lemma 13.2 on page 156 in Rüßmann's paper. First of all observe that in the first inequality in (4.3.24) it is sufficient to take

$$\lambda = \frac{2\tau}{\kappa - 2\psi} - 1 = \frac{225}{144\bar{\mu}} - 1 \tag{4.3.26}$$

in view of $T_\nu > T_0 \ge 1$. Now we consider $\widetilde{\Phi}$ instead of Φ and consequently define

$$\begin{aligned} \widetilde{\Psi}(T) &:= T^{-(d+p+1)}\widetilde{\Phi}(T) = T^{-(d+p+1+a)} \\ \widetilde{T}_\nu &:= \widetilde{\Psi}^{-1}\left(\widetilde{\Psi}(\widetilde{T}_0)\delta^{\tau\nu}\right) = \widetilde{T}_0 \delta^{-\frac{\tau\nu}{a+d+p+1}} \end{aligned}$$

where $\widetilde{T}_0$ verifies (4.3.27) (with $\widetilde{\Phi}$ instead of Φ). We claim that an analogous estimate as in (4.3.24) can be obtain to make the series in (4.3.25) converge (with $\widetilde{U}$ instead of U for $U = \Psi, T_\nu$). Indeed

$$\widetilde{T}_\nu^2 \delta^{(\kappa-2\psi)\nu} \le \widetilde{T}_\nu^{(\lambda+1)\frac{\kappa-2\psi}{\tau}} \delta^{(\kappa-2\psi)\nu} = \left(\widetilde{T}_\nu^{\lambda+1}\widetilde{\Phi}(T_\nu)\frac{\delta^{\tau\nu}}{\widetilde{\Phi}(T_\nu)}\right)^{\frac{\kappa-2\psi}{\tau}} \le$$

$$\le \left(\frac{\widetilde{T}_\nu^{\lambda+1-a}}{\widetilde{\Phi}(T_\nu)\widetilde{T}_\nu^{-(d+p+1)}\delta^{-\tau\nu}}\right)^{\frac{\kappa-2\psi}{\tau}} = \left(\frac{\widetilde{T}_\nu^{\lambda+1-a}}{\widetilde{\Psi}(T_\nu)\delta^{-\tau\nu}}\right)^{\frac{\kappa-2\psi}{\tau}} \le \left(\frac{1}{\Psi(T_0)}\right)^{\frac{\kappa-2\psi}{\tau}}$$

having used the definitions of $\widetilde{\Psi}$ and $\widetilde{T}_\nu \ge 1$ and

$$a > (d+p)\bar{\mu} \ge 2 > \frac{225}{144} \ge \frac{225}{144\bar{\mu}} = \lambda + 1\,.$$

Thus we obtain

$$\sum_{\nu=0}^{\infty} \widetilde{T}_\nu \delta^{(\kappa-2\psi)\nu} \leq \left(\frac{1}{\Psi(\widetilde{T}_0)}\right)^{\frac{\kappa-2\psi}{\tau}} \sum_{\nu=0}^{\infty} \frac{1}{\widetilde{T}_\nu} =$$

$$= \Psi(\widetilde{T}_0)^{-\frac{\kappa-2\psi}{\tau}} \widetilde{T}_0^{-1} \sum_{\nu=0}^{\infty} \delta^{\frac{\tau\nu}{a+d+p+1}} < \infty$$

in view of $\delta < 1$ and $\tau > 0$. It should be noted that we did not need lemma 13.2 on [Rüßm01, page 156] to prove the convergence of the considered sum since our choice of $\widetilde{\Phi}$ as approximation function permitted an explicit expression for $\widetilde{T}_\nu$.

The second inequality in which Rüßmann uses property 4.3.5.3 (inequality (2.5.34) in this thesis, on page 193 in [Rüßm01] with $m = n$) is the following

$$\int_1^{\infty} s^{-(d+p+1+\frac{2}{\bar{\mu}})} \left(s^{d+p}\Phi(s)^{\frac{1}{\bar{\mu}}}\right)^2 ds \leq \left(\Phi_{(d+p)\bar{\mu}}\right)^2 \int_1^{\infty} s^{-(d+p+1+\frac{2}{\bar{\mu}})}\, ds$$

$$\leq \frac{\bar{\mu}}{2+(d+p)\bar{\mu}} \left(\Phi_{(d+p)\bar{\mu}}\right)^2$$

once we have replaced μ_0 with $\bar{\mu}$ accordingly to (4.3.17) (see (2.5.12) in theorem 2.5.1 for μ_0), n with $d+p$ and where $\Phi_{\lambda\bar{\mu}}$ is defined in (4.3.20). Now we consider, as already done before, $\widetilde{\Phi}$ in (4.3.22) as approximation function in the place of Φ; then, in view of (4.3.23), it results

$$\int_1^{\infty} s^{-(d+p+1+\frac{2}{\bar{\mu}})} \left(s^{d+p}\widetilde{\Phi}(s)^{\frac{1}{\bar{\mu}}}\right)^2 \leq \int_1^{\infty} s^{-(d+p+1+\frac{2}{\bar{\mu}})} \left(s^{d+p-\frac{a}{\bar{\mu}}}\right)^2 ds \leq$$

$$\leq \int_1^{\infty} s^{-(d+p+1+\frac{2}{\bar{\mu}})} = \frac{\bar{\mu}}{2+(d+p)\bar{\mu}}\,.$$

We observe that a different estimate for this integral could have been obtained if we were not searching for an estimate similar to Rüßmann's one □

Remark 4.3.4. *The second inequality considered in lemma 4.3.1 is fundamental for the choice of γ in theorem 2.5.3. In fact the term ($n = d+p$ and $\mu_0 = \bar{\mu}$ in our notation) that exists in Rüßmann's determination of γ (see for instance [Rüßm01, 14.10.8]) does not appear in our choice of γ (see section 4.3.7). and has to be replaced by 1 in theorem 2.5.3 with the choice of $\widetilde{\Phi}$ as approximation function.*

Remark 4.3.5. *We conclude this part observing that if $\widetilde{\Phi}$ is defined as in (4.3.22) for some $a > (d+p)\bar{\mu}$ we have*

$$\lim_{s\to\infty} s^\lambda \widetilde{\Phi}(s) = 0 \qquad \textit{for any} \qquad 0 \leq \lambda < a$$

and

$$\sup_{s\geq 1} s^{\lambda}\widetilde{\Phi}^{\frac{1}{\mu}}(s) = 1 \qquad \textit{for any} \qquad 0 \leq \lambda < \frac{a}{\bar{\mu}}\,, \quad 0 \leq \mu \leq \bar{\mu}$$

(confront with property 4.3.5.3).

4.3.6 Determination of T_0

Let $\widetilde{\Phi}$ be a given approximation function as in (4.3.22) then, according to Rüßmann's theorem, we have to choose and fix a number $\widetilde{T}_0 \geq 1$ such that

$$\widetilde{\Phi}(\widetilde{T}_0) \leq e^{-(d+p+1)} \qquad \text{and} \qquad \int_{\widetilde{T}_0}^{\infty} \log\frac{1}{\widetilde{\Phi}(T)}\frac{dT}{T^2} \leq \frac{\vartheta \log 2}{12\bar{\mu}(d+p)\log(3^{25}(d+p))} \tag{4.3.27}$$

(see for instance section 2.4 where n plays the role of $d+p$).

With our definition of $\widetilde{\Phi}$ the first condition on $\widetilde{T}_0$ becomes

$$\frac{1}{\widetilde{T}_0^a} \leq e^{-(d+p+1)} \Longleftrightarrow \widetilde{T}_0 \geq e^{\frac{d+p+1}{a}}\,. \tag{4.3.28}$$

For what concerns the second condition, we set

$$c_{d+p} := \frac{\log 2}{192(d+p)\log(3^{25}(d+p))} \tag{4.3.29}$$

so that we have to fulfill

$$\int_{\widetilde{T}_0}^{\infty} \log\frac{1}{\widetilde{\Phi}(T)}\frac{dT}{T^2} \leq c_{d+p}\frac{\epsilon}{\bar{\mu}} \tag{4.3.30}$$

in view of $\vartheta := \frac{\epsilon}{16}$ as in (4.3.7). Now we assume $\widetilde{T}_0 \geq 4$ so that $\log T \leq T^{\frac{1}{2}}$ for any $T \geq \widetilde{T}_0$; then, by definition of $\widetilde{\Phi}$ it results

$$\int_{\widetilde{T}_0}^{\infty} \log\frac{1}{\widetilde{\Phi}(T)}\frac{dT}{T^2} = a\int_{\widetilde{T}_0}^{\infty}\frac{\log T}{T^2}\,dT \leq a\int_{\widetilde{T}_0}^{\infty}\frac{dT}{T^{\frac{3}{2}}} = \frac{2a}{\widetilde{T}_0^{\frac{1}{2}}}$$

so that we may fulfill both conditions in (4.3.27) with

$$\widetilde{T}_0 = \left(\frac{2a\bar{\mu}}{c_{d+p}}\right)^2\frac{1}{\epsilon^2}\,. \tag{4.3.31}$$

Furthermore, also $\widetilde{T}_0 \geq 4$ is satisfied as the following chain of inequalities shows:

$$\left(\frac{2a\bar{\mu}}{c_{d+p}}\right)^2\frac{1}{\epsilon^2} \geq (2a\bar{\mu})^2 \geq 4a^2 > 16 > e^{\frac{3}{2}} \geq e^{\frac{d+p+1}{d+p}} > e^{\frac{d+p+1}{a}}$$

having used $\epsilon, c_{d+p} < 1$, $\bar{\mu}, d, p \geq 1$ and (4.3.23) which gives $2 \leq d+p \leq (d+p)\bar{\mu} < a$. We conclude this subsection defining

$$\overline{T}_0 := \left(\frac{2a\bar{\mu}}{c_{d+p}}\right)^2 \tag{4.3.32}$$

and observing that with our choice of $\widetilde{\Phi}$ and $\widetilde{T}_0$ it results

$$\widetilde{\Phi}(\widetilde{T}_0)\widetilde{T}_0^{-(d+p+1)} = \frac{\epsilon^{2(a+d+p+1)}}{\overline{T}_0^{\,a+d+p+1}} \qquad \text{with} \qquad a > (d+p)\bar{\mu}\,. \tag{4.3.33}$$

4.3.7 Determination of γ

Let C_1 be the constant defined in (4.3.19) and $\bar{\mu}$ the index of non-degeneracy of Ψ (see (4.1.2)), we set

$$C_0^\star := 2^{\bar{\mu}+1}(\bar{\mu}+1)^{\bar{\mu}+2}(C_1+1) \tag{4.3.34}$$

(confront with $C^\star$ in (2.4.3) taking into consideration remark 4.3.4 and the absence of the term $\vartheta^{-(\bar{\mu}+1)}$ because of its dependence on ϵ that will be considered next). Now refer to (2.4.4), replace d_0 and ϑ respectively by equation (4.3.11) and (4.3.7) and observe

$$(d+p)^{-\frac{1}{2}} + 2(r+2) + \frac{16}{\epsilon}(r+2) \geq (d+p)^{-\frac{1}{2}} + 18(r+2)\,. \tag{4.3.35}$$

Thus, considering

$$\beta = \frac{\epsilon^{\bar{\mu}+2}}{8}\bar{\beta}\,, \qquad \vartheta = \frac{\epsilon}{16}\,, \qquad \epsilon^\star = \epsilon_0^\star \epsilon^p \tag{4.3.36}$$

(as in (4.3.16), (4.3.7) and (4.3.14)) and referring to the definition of γ in (2.4.5) with γ_1 in (2.4.4), we take

$$\begin{aligned} \bar{\gamma}_0 \;&:=\; \left[3^{d+p+3}(2\pi e)^{\frac{d+p}{2}}(r+2)^{d+p-1}\left((d+p)^{-\frac{1}{2}} + 18(r+2)\right)C_0^\star\right]^{-\frac{\bar{\mu}}{2}} \\ &* \;\; 2^{-\left(\frac{7}{2}\bar{\mu}+3\right)(\bar{\mu}+1)}\epsilon_0^{\star\frac{\bar{\mu}}{2}}\bar{\beta}^{\frac{\bar{\mu}+1}{2}}\,, \end{aligned} \tag{4.3.37}$$

set

$$\bar{\gamma}_1 := \min\{\bar{\gamma}_0, 2\} \tag{4.3.38}$$

and finally define

$$\gamma := \epsilon^{(\bar{\mu}+1)^2 + \frac{p\bar{\mu}}{2}}\bar{\gamma}_1\,. \tag{4.3.39}$$

For clearness we remark that the exponent of ϵ in this last expression, is given by the three contributions

$$\frac{(\bar{\mu}+2)(\bar{\mu}+1)}{2}\,, \qquad \frac{(\bar{\mu}+1)\bar{\mu}}{2}\,, \qquad \frac{p\bar{\mu}}{2}$$

given respectively by the presence of β, ϑ and $\epsilon^\star$ as in (4.3.36).

4.3.8 Definition of $M^\star$

Accordingly to the definition of $M_\star$ in (2.4.2), our choice of ϑ, μ_0 and C_1 (respectively in (4.3.7), (4.3.17) and (4.3.19)) we define

$$M_0^\star := 2^{5\bar{\mu}+2}\bar{\mu}!(M_0 + M_2 + \Delta + 1) \tag{4.3.40}$$

and consequently

$$M^\star = M_0^\star \epsilon^{-\bar{\mu}} \,. \tag{4.3.41}$$

4.3.9 Determination of t_0

Let γ be as in (4.3.39) (with $\bar{\gamma}_1$ and $\bar{\gamma}_0$ in (4.3.38) and (4.3.37)) and let $\widetilde{\Phi}$ be the chosen approximation function in (4.3.22) with a in (4.3.23). Recalling the expression found for $\widetilde{\Phi}(\widetilde{T}_0)\widetilde{T}_0^{-(d+p+1)}$ in (4.3.33) (where $\widetilde{T}_0$ and $\overline{T}_0$ are defined in (4.3.31) and (4.3.32)) and accordingly to the definition of t_0 in (2.4.6) we may put

$$t_0 := \epsilon^{n_1}\bar{t}_0 \qquad \text{with} \qquad n_1 := (\bar{\mu}+1)^2 + \frac{p\bar{\mu}}{2} + 2(a+d+p) + 3 \tag{4.3.42}$$

and

$$\bar{t}_0 := \frac{\bar{\gamma}_1 \overline{T}_0^{\,-(a+d+p+1)}}{2^7(M_0 + M_2 + \Delta + 1)} \tag{4.3.43}$$

having used also (4.3.19) and (4.3.7).

4.3.10 The quantities E_1, E_2, E_3

In the preceding subsection we have analyzed the behavior of all quantities involved in the estimate for the size of the perturbation in Rüßmann's theorem (see section 2.4) when we consider H_ϵ^6 as Hamiltonian function (see (3.5.10)) and $\widetilde{\Phi}$ as approximation function (see (4.3.22)). Let now see how what happens to E_1,E_2 and E_3 in (2.4.9). A suitable choice for the first of this three quantities may be

$$E_1 := \frac{\bar{\gamma}_1 \overline{T}_0^{\,-(a+d+p+1)}}{3^{35+d+p}}\,\epsilon^{n_1} \qquad \text{(with } n_1 \text{ in (4.3.42))} \tag{4.3.44}$$

in view of the definition of γ in (4.3.39) (together with (4.3.38) and (4.3.37)), the choice of ϑ in (4.3.7), equation (4.3.33) and $\min\{\epsilon 2^{-4}, 3^{-17}\} \geq \epsilon 3^{-17}$.

For what concerns E_2 as it appears in (2.4.9), we have to substitute γ, Φ and T_0 as taken in section 2.4 with γ in (4.3.39), $\widetilde{\Phi}$ in (4.3.22), $\widetilde{T}_0$ in (4.3.31) and consequently define

$$E_2 := \frac{\bar{\gamma}_1 \overline{T}_0^{\,-(a+d+p+1)}}{3^{22}} \left(1 - \frac{1}{2^{\bar{\mu}}}\right)\left(1 - \frac{1}{2^{5-\frac{2}{\bar{\mu}}}}\right)\epsilon^{n_2} \qquad \text{with} \quad n_2 := n_1 - 1 \tag{4.3.45}$$

in a totally analogous way to what done before for E_1.

Finally, accordingly to the definition given in (2.4.9), the choice of β and μ_0 in (4.3.17), the definitions of t_0 in (4.3.42), $\widetilde{T}_0$ in (4.3.31) and $M^\star$ in (4.3.41), we define

$$E_3 := \frac{\bar{\beta}\,\bar{t}_0^{\bar{\mu}}}{3^{19}2^4\overline{T}_0 M_0^\star C(d+p,\bar{\mu})}\,\epsilon^{n_3} \qquad \text{with} \quad n_3 := n_1\bar{\mu} + 2(\bar{\mu}+2) \tag{4.3.46}$$

where $\bar{t}_0$ is in (4.3.43), $\overline{T}_0$ in (4.3.32), $M_0^\star$ in (4.3.40) and $C(d+p,\bar{\mu})$ can be found in (2.4.7).

4.3.11 Determination of N and conditions on ϵ

We recall that in (2.4.8) we defined a possible value for ϵ_0 as it appears in (2.3.5) (and analogously in (C.2.8)) by

$$\epsilon_0 = \frac{\vartheta}{C_1}\min\{E_1, E_2, E_3\}^2$$

where E_1, E_2 and E_3 may be defined as in (4.3.44), (4.3.45) and (4.3.45). We remark once again that the values we have given are a simplification of Rüßmann's estimates in [Rüßm01, page 171] concerning the size of the perturbation in lower dimensional tori theorem (see theorem C.2.1 with estimate for ϵ_0 in C.2.1 in this thesis). On the other hand in the maximal case the simplifications are made through his whole work considering the case $p = 0 = q$ (in Rüßmann's notation) in parts III (Construction of Invariant Tori) and IV (Existence of Non-resonant Frequency Vectors), where this last is also discussed in its major aspects in section 2.5.

Now accordingly to the preceding subsections, the choice of ϑ in (4.3.7) and C_1 in (4.3.19), we take

$$\epsilon_0 = \frac{\epsilon}{2^4(M_0 + M_2 + \Delta)}\min\{E_1, E_2, E_3\}^2\,. \tag{4.3.47}$$

Moreover, from the definitions of E_1, E_2 and E_3 in subsection (4.3.10) and $\epsilon < 1$ we infer

$$\epsilon^{N^{(0)}}\min\left\{E_1^{(0)}, E_2^{(0)}, E_3^{(0)}\right\}^2 \leq \min\{E_1, E_2, E_3\}^2 \tag{4.3.48}$$

if we put $N^{(0)} := 2\max\{n_1, n_2, n_3\} = 2n_3$, that is

$$N^{(0)} = 2\bar{\mu}(\bar{\mu}+1)^2 + p\bar{\mu}^2 + 2(2a + 2d + 2p + 5)\bar{\mu} + 8 \tag{4.3.49}$$

and

$$\begin{aligned}
E_1^{(0)} &= \frac{\bar{\gamma}_1 \overline{T}_0{}^{-(a+d+p+1)}}{3^{35+d+p}} \\
E_2^{(0)} &= \frac{\bar{\gamma}_1 \overline{T}_0^{-(a+d+p+1)}}{3^{22}} \left(1 - \frac{1}{2^{\bar{\mu}}}\right) \left(1 - \frac{1}{2^{5-\frac{2}{\bar{\mu}}}}\right) \\
E_3^{(0)} &= \frac{\bar{\beta}\, \bar{t}_0^{\bar{\mu}}}{3^{19} 2^4 \overline{T}_0 M_0^\star C(n, \bar{\mu})}
\end{aligned} \tag{4.3.50}$$

accordingly to (4.3.44), (4.3.45) and (4.3.46).

Now, in view of (4.3.1) with correspondences (4.3.4), (4.3.47) and (4.3.48), we may fulfill inequality (2.3.5) by

$$(M'_{N_2} + \mu^{N+1})\epsilon^N \leq \frac{\epsilon^{N^{(0)}+1}}{2^4(M_0 + M_2 + \Delta)} \min\left\{E_1^{(0)}, E_2^{(0)}, E_3^{(0)}\right\}^2 \tag{4.3.51}$$

which requires $N \geq N^{(0)} + 2$ and ϵ sufficiently small as it will be described later. Thus, a suitable value for N is

$$N = 2\bar{\mu}(\bar{\mu} + 1)^2 + p\bar{\mu}^2 + 2(2a + 2d + 2p + 5)\bar{\mu} + 10 \tag{4.3.52}$$

where $d + p$ is half of the dimension of the phase space of the Hamiltonian system considered ($2p$ denotes the number of the initial elliptic variables), $\bar{\mu}$ is the index of non-degeneracy of the frequency application and $a > (d + p)\bar{\mu}$ is the exponent of the chosen approximation function.

Finally, in view of inequality (4.3.51) (which implies (2.3.5)) and equations (4.3.52) and (4.3.49) which give $N = N^{(0)} + 2$, we obtain the following sufficient condition to impose on ϵ to be able to apply Rüßmann's theorem:

$$\epsilon \leq \frac{\min\left\{E_1^{(0)}, E_2^{(0)}, E_3^{(0)}\right\}^2}{2^4(M_N + \mu^{N+1})(M_0 + M_2 + \Delta)} \tag{4.3.53}$$

where $E_1^{(0)}, E_2^{(0)}, E_3^{(0)}$ are defined in (4.3.50) (accordingly to all definitions given previously in this section), μ is defined in (3.2.21) and M_N is a positive number greater than M'_{N_2}. We conclude observing that from the definition of M'_{N_2} in (3.5.12) together with (3.4.47), (3.4.39) and (3.3.19), (3.4.40) and (3.4.41), we have

$$M'_{N_2} = \frac{2\, 5^{2N^2-N-2}}{r_1^{2N-2} 3^{4N^2-5N}} \left[2^9(4N-1)^p)\right]^{4N^2-6N} \left(\frac{M}{\alpha}\right)^{2N^2-3N} M \tag{4.3.54}$$

(see (3.4.39) for M) so that we may define

$$M_N := \frac{M}{r_1^{2N-2}} B_{N,p}^N \left(\frac{M}{\alpha}\right)^{N\kappa} M \tag{4.3.55}$$

with $B_{N,p}$ and κ in (4.1.9).

4.4 Conditions on ϵ in theorems 3.1.2 and 3.1.3

In this section we initially gather all the conditions on ϵ contained in chapter 3 and needed to conjugate the initial Hamiltonian H to H_ϵ^6 in (3.5.10); then we will synthesize the conditions imposed in sections 4.1 and 4.2 to prove the non-degeneracy in the sense Rüßmann of the frequency application and to control its index and amount of non degeneracy.

The first condition is formulated in (3.2.17) in order to apply averaging theorem and in particular corollary (3.2.1). Observe that if $\mu \geq \frac{1}{e}$ then we have

$$\epsilon \log \frac{1}{\epsilon\mu} \leq \epsilon(1 - \log\epsilon)\,.$$

Moreover we infer that $\epsilon(1 - \log\epsilon) \leq \sqrt{\epsilon}$ for every $0 \leq \epsilon \leq e^{-4}$; indeed the function $f(\epsilon) = \sqrt{\epsilon}(1 - \log\epsilon)$ equals zero in $\epsilon = 0$ and is increasing for $0 \leq \epsilon \leq \epsilon^{-1}$ so that $f(\epsilon) \leq f(e^{-4}) = 5e^{-2} \leq 1$ for every $0 \leq \epsilon \leq e^{-4}$. Then we can fulfill (3.2.17) by requiring

$$\epsilon \leq \frac{\alpha^2 (d')^2}{\mu^2 c_p^2 N^2} \tag{4.4.1}$$

in view of $N_1 - 1 = N$ from (3.4.42). Now notice that the condition $\epsilon \leq e^{-4}$ is superfluous in view of (3.3.29) and $p \geq 1$ which give $\epsilon \leq 12^{-2}$;

For what concerns condition (3.3.29), it can be immediately seen that it is implied by

$$\epsilon \leq \frac{\min\left\{\alpha(2M_2)^{-1}, \alpha M_3^{-1}, 1\right\}^{2p} \alpha}{2^{2p}(2p+1)!^2(2N-1)} \tag{4.4.2}$$

that is condition (3.4.2) once that N_2 has been replaced by means of (3.4.42). Together with $\alpha \leq 1$ the definition of M_3 in (3.3.24) and $N \geq 2$, this last inequality implies

$$\begin{aligned}\epsilon &\leq \frac{\alpha M_3^{-1}}{2^{2p}(2p+1)!^2(2N-1)} \leq \frac{1}{2^3(2p+1)!^2} \frac{\alpha^2 (d')^2}{2^7 c_p \mu^2} \min\left\{\frac{\alpha}{c_0 r_1}, 1\right\} \leq \\ &\leq \frac{\alpha^2 (d')^2}{2^{10} c_p \mu^2} \min\left\{\frac{\alpha}{c_0 r_1}, 1\right\} \end{aligned} \tag{4.4.3}$$

that $\epsilon \leq \min\{\epsilon_1, \epsilon_2\}$ with ϵ_1 in (3.3.8) and ϵ_2 in (3.3.12) as required in lemma 3.3.2.

Now, from (3.2.23) and $p \geq 1$ we have

$$c_p^2 = c_p 2^7 3e(1+ep) \leq c_p 2^7 3e^2(p+1) \leq c_p 2^{13}(p+1) \leq c_p 2^{10}(2p+1)!^2$$

so that condition (4.4.1) can be fulfilled taking

$$\epsilon \leq \frac{\alpha M_3^{-1}}{2^{2p+1}(2p+1)!^2 N^2} \tag{4.4.4}$$

using the definition of M_3 in (3.3.24) as in the first inequality in (4.4.3). Thus, we may finally satisfy conditions (4.4.1) and (4.4.2), that is all conditions in sections 3.2, 3.3 and 3.4, assuming

$$\epsilon \leq \frac{\min\left\{\alpha(2M_2)^{-1}, \alpha M_3^{-1}, 1\right\}^{2p} \alpha}{2^{2p+1}(2p+1)!^2 N^2} \tag{4.4.5}$$

where M_2 and M_3 are defined in (3.3.24) (together with (3.2.24), (3.2.23), (3.2.21) and (3.1.5)), $\alpha \leq 1$ is characterized by inequalities (3.2.15) and (3.2.16) and N is defined in (4.3.52).

Another upper bound on the parameter ϵ is given by the choice of ρ^0 in (3.5.1) verifying $|\rho^0| = 2\epsilon$ (equation (3.5.8)) which leads to require

$$\epsilon \leq \frac{1}{24}(r')^2 . \tag{4.4.6}$$

By the definition of r' in (3.4.40), δ in (3.4.41) and substituting N_2 with $2N-1$ (see (3.4.42)) we can compute

$$\begin{aligned}
\frac{1}{24}(r')^2 &= \frac{r_1^2}{600}\left(\frac{3}{4}\delta\right)^{4N-6} = \frac{r_1^2}{600}\left(\frac{3}{2^9(4N-1)^p}\right)^{4N-6}\left(\frac{\alpha}{5M}\right)^{2N-3} \geq \\
&\geq \frac{(d')^2}{600}\left(\frac{1}{2^9(4N-1)^p}\right)^{4N-6}\left(\frac{\alpha}{M}\right)^{2N-3} \geq \\
&\geq (d')^2 \frac{1}{[2^9(4N-1)^p]^{4N-5}}\left(\frac{\alpha}{M}\right)^{2N-3} \geq \frac{(d')^2}{\sqrt{B_{N,p}}}\left(\frac{\alpha}{M}\right)^{2N-3}
\end{aligned}$$

where $B_{N,p}$ is defined in (4.1.9). Then we can fulfill (4.4.6) taking

$$\epsilon \leq \frac{(d')^2}{\sqrt{B_{N,p}}}\left(\frac{\alpha}{M}\right)^{\kappa} \tag{4.4.7}$$

since $\alpha \leq 1$, $M \geq 1$ (see (3.4.39)) and from (4.1.9) $\kappa \geq 4N-5 \geq 2N-3$.

The two last conditions on ϵ are those needed to control the index and amount of non-degeneracy of the frequency application $\hat{\Psi}$ (see (4.1.3)) in terms of the index and amount of non-degeneracy of Ψ (see (4.1.2)); namely this are inequalities (4.1.19) and (4.2.10). Since $\Delta > 1$ (see (4.1.8) with (4.1.9)) the two minima appearing in this two estimates are less than 1 and this means that we may replace both of them with

$$\min\left\{\frac{1}{M_0+M_2}, \frac{1}{\Delta}\right\}^{d+p}$$

putting a factor $2^{-(d+p+1)}$ in second member of (4.2.10) instead of 4^{-1}. Thus, both inequalities hold if

$$\epsilon \leq \frac{1}{2^{d+p+1}} \min\left\{\frac{1}{M_0+M_2}, \frac{1}{\Delta}\right\}^{d+p} \min\left\{\frac{\vartheta_1^{\bar{\mu}}\bar{\beta}}{\bar{\mu}!\Delta}, \frac{\delta^\star}{(\nu!)^{d+p}}\left(\frac{r}{5}\right)^{(d+p)\nu}\right\} \tag{4.4.8}$$

where recall $M_0 = \sup_{B_r}|\omega|$, $M_2 = \sup_{j=1,\ldots,p}\sup_{B_r}|\Omega_j|$ (with $B_r = D_r^d(I_0) \subset \mathbb{R}^d$), $\bar{\mu}$ and $\bar{\beta}$ are respectively the index and amount of non-degeneracy of $\Psi = (\omega, \Omega)$ with respect to an initially chosen compact set $\overline{\mathcal{K}}$ (see (4.2.5)), Δ is defined in (4.1.8) with (4.1.9) and

$$\delta^\star = \delta^\star(\Psi, r_0) := \left|\det[\partial_r^{j^{(1)}}\Psi_\epsilon^0(0), \ldots, \partial_r^{j^{(d+p)}}\Psi_\epsilon^0(0)]\right| \neq 0 \tag{4.4.9}$$

$$\nu = \nu(\Psi, r_0) := \max_{k=1,\ldots,d+p}\left|j^{(k)}\right|_1 \tag{4.4.10}$$

accordingly to definitions (4.1.22), (4.1.23) and (4.1.10), $r_0 = I_0$ and propositions 4.1.2 and 4.1.3 .

Chapter 5

The spatial planetary $(n+1)$-body problem

In this chapter we expose the results contained in [Féj04, pages 45-62] about the non-planarity (definition 2.1.1) of the frequency application of the spatial planetary $(n+1)$-body problem Hamiltonian. The cited work in his entirety, is one of the highest achievement KAM theory applied to celestial mechanics. In fact, in [Féj04] a complete proof of Arnold's theorem on planetary motions (contained in [Arn63b]) is provided, more than 40 years after Arnold's statement. In that paper Jacques Féjoz has completed, with the help of other mathematicians in Paris, Michel Herman's work whose untimely death in 2001 had interrupted.

With the results we are going to show, we will be able to apply theorem 3.1.1 to the spatial planetary $(n+1)$-body problem giving an analogous result, in analytic class, to Herman and Féjoz' theorem formulated in class C^∞ (theorem 63 in [Féj04]).

As a remark, we say that the first section is the only one in this chapter where we do not follow [Féj04]. In fact, for the classical and well-known description of the Hamiltonian model for the planetary $(n+1)$-body problem, we use notations from [Chi05a]. There, beyond more recent numerical results from the authors, the reader may also find a precise historic description of how KAM theory and celestial mechanics have interacted through the years, from Kolmogorov's theorem in 1954 (see appendix A in this thesis) until Féjoz' 2004 paper. For more detailed historical remarks about KAM theory and celestial mechanics see also [Chi06].

5.1 Hamiltonian models for the planetary $(n + 1)$-body problem

The movements of $n + 1$ bodies (point masses) interacting only through gravitational attraction are ruled by Newton's equations

$$\ddot{u}^{(i)} = \sum_{\substack{0\le j\le n \\ j\ne i}} \bar{m}_j \frac{u^{(j)} - u^{(i)}}{\left|u^{(i)} - u^{(j)}\right|^3} \ , \qquad i = 0, \ldots , n \tag{5.1.1}$$

where $u^{(i)} = (u_1^{(i)}, u_2^{(i)}, u_3^{(i)}) \in \mathbb{R}^3$ are the cartesian coordinates of the i^{th}-body of mass $\bar{m}_i$, (once that the physical space has been identified with the euclidean space $\mathbb{R}^3$ through the choice of an inertial frame), $|\cdot|$ is the standard euclidean norm and the gravitational constant has been renormalized to one by rescaling the time t. As it can be easily seen, equations (5.1.1) expressing the universal gravitational attraction law, are invariant under change of inertial frames, that is under change of variables of the form $u^{(i)} \to u^{(i)} - (a + ct)$ for any chosen $a, c \in \mathbb{R}$. Therefore, taking

$$a := \frac{1}{m_{\text{tot}}} \sum_{i=0}^{n} \bar{m}_i u^{(i)}(0) \qquad \text{and} \qquad c := \frac{1}{m_{\text{tot}}} \sum_{i=0}^{n} \bar{m}_i \dot{u}^{(i)}(0)$$

with $m_{\text{tot}} := \sum_{i=0}^{n} \bar{m}_i$, we may restrict our attention to the manifold of initial data given by

$$\sum_{i=0}^{n} \bar{m}_i u^{(i)}(0) = 0 \ , \quad \sum_{i=0}^{n} \bar{m}_i \dot{u}^{(i)}(0) = 0 \, . \tag{5.1.2}$$

Now observe that the total linear momentum $M_{\text{tot}}(t) := \sum_{i=0}^{n} \bar{m}_i \dot{u}^{(i)}(t)$ is constant along the flow of (5.1.1), since it has vanishing derivative), so that we have $M_{\text{tot}}(t) = 0$ along trajectories using (5.1.2). This implies that also the position of the barycenter $B(t) := \sum_{i=0}^{n} \bar{m}_i u^{(i)}(t)$ is constant and equals 0, again by (5.1.2). This means that the manifold of initial data given by (5.1.2) is invariant under the flow newton's equation (5.1.1).

As it is well know, the integral curves of equations (5.1.1) are the integral curves of the Hamiltonian vector field generated by the Hamiltonian function

$$\widetilde{H}_{\text{New}} := \sum_{i=0}^{n} \frac{\left|U^{(i)}\right|^2}{2\bar{m}_i} - \sum_{0\le i<j\le n} \frac{\bar{m}_i \bar{m}_j}{\left|u^{(i)} - u^{(j)}\right|}$$

where $U^{(i)} = \bar{m}_i u^{(i)}$ is the momentum conjugated to $u^{(i)}$, $(U^{(i)}, u^{(i)})$ are standard symplectic variables and the phase space considered is the open domain in $\mathbb{R}^{6(n+1)}$

given by

$$\widetilde{\mathcal{M}} := \big\{ U^{(i)}, u^{(i)} \in \mathbb{R}^3 \ : \ u^{(i)} \neq u^{(j)} \, , \, 0 \le i \neq j \le n \big\} \tag{5.1.3}$$

endowed with the standard symplectic form

$$\sum_{i=0}^{n} dU^{(i)} \wedge du^{(i)} := \sum_{\substack{0 \le i \le n \\ j=1,2,3}} dU_j^{(i)} \wedge du_j^{(i)} \, . \tag{5.1.4}$$

Notice that considering $\widetilde{\mathcal{M}}$ as phase space we do not only exclude collisions between the bodies but also intersection between their orbits. As shown before we may assume that the motions governed by $\widetilde{H}_{\text{New}}$ lie on the symplectic submanifold

$$\widetilde{\mathcal{M}}_0 := \left\{ (U,u) \in \widetilde{M} \ : \ \sum_{i_0}^{n} \bar{m}_i u^{(i)} = 0 = \sum_{i=0}^{n} \bar{m}_i U^{(i)} \right\} \tag{5.1.5}$$

which corresponds to equations (5.1.2).

Now, the flow of $\widetilde{H}_{\text{New}}$ on $\widetilde{M}_0$ can be best described in terms of heliocentric coordinates. Let $\Phi_{\text{hel}} : (R,r) \longrightarrow (U,u)$ be the linear symplectic transformation given by

$$\Phi_{\text{hel}} : \begin{cases} u^{(0)} = r^{(0)} \, , & u^{(i)} = r^{(0)} + r^{(i)} \, , \quad \forall \ i = 1, \dots, n \\ U^{(0)} = R^{(0)} - \sum_{i=1}^{n} R^{(i)} \, , & U^{(i)} = R^{(i)} \, , \quad \forall \ i = 1, \dots, n \, ; \end{cases} \tag{5.1.6}$$

in the new variables (R,r), $\widetilde{\mathcal{M}}_0$ becomes

$$\begin{aligned} \Big\{ \ R^{(i)}, r^{(i)} \in \mathbb{R}^3 : R^{(0)} = 0, \, r^{(0)} = -m_{\text{tot}}^{-1} \sum_{i=1}^{n} \bar{m}_i r^{(i)}, \\ 0 \neq r^{(i)} \neq r^{(j)} \, , \ \forall \ 1 \le i \neq j \Big\} \ , \end{aligned}$$

the restriction on it of the 2-form (5.1.4) is $\sum_{i=i}^{n} R^{(i)} \wedge r^{(i)}$ and the new Hamiltonian function $H_{\text{New}} := \widetilde{H}_{\text{New}} \circ \Phi_{\text{hel}}|_{\widetilde{\mathcal{M}}_0}$ is given by

$$H_{\text{New}} = \sum_{i=1}^{n} \left(\frac{\big|R^{(i)}\big|^2}{2 \frac{\bar{m}_0 \bar{m}_i}{\bar{m}_0 + \bar{m}_i}} - \frac{\bar{m}_0 \bar{m}_i}{|r^{(i)}|} \right) + \sum_{1 \le i < j \le n} \left(\frac{R^{(i)} R^{(j)}}{\bar{m}_0} - \frac{\bar{m}_i \bar{m}_j}{|r^{(i)} - r^{(j)}|} \right) \, . \tag{5.1.7}$$

Thus, since $R^{(0)} = 0$ and $r^{(0)}$ does not appear, the dynamics generated by H_{New} on $\widetilde{\mathcal{M}}_0$ are equivalent to the dynamics on the phase space

$$\begin{aligned} \mathcal{M}_0 := \big\{ (R,r) = (R^{(1)}, \dots, R^{(n)}, r^{(1)}, \dots, r^{(n)}) \in R^{6n} \ : \\ 0 \neq r^{(i)} \neq r^{(j)} \, , \ \forall \ 1 \le i \neq j \le n \big\} \end{aligned}$$

endowed with the standard symplectic form $\sum_{i=i}^n R^{(i)} \wedge r^{(i)}$; to recover the whole dynamic on the manifold $\widetilde{\mathcal{M}}_0$ it is sufficient to take $R^{(0)}(t) := 0$ for all t and $r^{(0)}(t) := m_{\text{tot}}^{-1} \sum_{i=1}^n r^{(i)}(t)$ (see the definition of Φ_{hel} in (5.1.6) and the equations of $\widetilde{M}_0$ in (5.1.5)).

Now, motivated by the planetary case, in which one mass is much bigger than the others (for instance in our planetary system the biggest i^{th}-planet/Sun ratio is 10^{-3} in the case of Jupiter), we perform a simple rescaling of masses by a positive small parameter ϵ:

$$\begin{cases} m_0 = \bar{m}_0\,, & m_i = \frac{1}{\epsilon}\bar{m}_i\,(i \geq 1) \\ X^{(i)} := \frac{1}{\epsilon} R^{(i)}\,, & x^{(i)} := r^{(i)}\,. \end{cases} \tag{5.1.8}$$

If we take $H_{\text{plt}}(X,x) := \frac{1}{\epsilon} H_{\text{New}}(\epsilon X, x)$ as new Hamiltonian function, this rescaling clearly leaves unchanged Hamilton's equations; thus, denoting

$$\mu_i := \frac{m_0 m_i}{m_0 + \epsilon m_i}\,, \quad M_i := m_0 + \epsilon m_i\,. \tag{5.1.9}$$

we may write explicitly

$$H_{\text{plt}}(X,x) = \sum_{i=1}^n \left(\frac{\left|X^{(i)}\right|^2}{2\mu_i} - \frac{\mu_i M_i}{|x^{(i)}|} \right) + \epsilon \sum_{1 \leq i < j \leq n} \left(-\frac{m_i m_j}{|x^{(i)} - x^{(j)}|} + \frac{X^{(i)} \cdot X^{(j)}}{m_0} \right) \tag{5.1.10}$$

whose phase space is

$$\mathcal{M} := \big\{ (X,x) = (X^{(1)}, \dots X^{(n)}, x^{(1)}, \dots, x^{(n)}) \in R^{6n} \ : \\ 0 \neq x^{(i)} \neq x^{(j)}\,,\ \forall\ 1 \leq i \neq j \leq n \big\} \tag{5.1.11}$$

with respect to the standard symplectic form $\sum_{i=1}^n dX^{(i)} \wedge dx^{(i)}$.

When $\epsilon = 0$ the Hamiltonian function H_{plt} becomes simply the *Keplerian Hamiltonian*

$$F_{\text{Kep}} := \sum_{i=1}^n \left(\frac{\left|X^{(i)}\right|^2}{2\mu_i} - \frac{\mu_i M_i}{|x^{(i)}|} \right) \tag{5.1.12}$$

that is the Hamiltonian of n disjoint 2-body problems; we shall call the first order of H_{plt} in ϵ the *perturbative function* denoted by

$$F_{\text{per}} := \sum_{1 \leq i < j \leq n} \left(-\frac{m_i m_j}{|x^{(i)} - x^{(j)}|} + \frac{X^{(i)} \cdot X^{(j)}}{m_0} \right) \tag{5.1.13}$$

and decomposed in the sum of a *principal Hamiltonian*

$$F_{\text{princ}} := -\sum_{1 \leq i < j \leq n} \frac{m_i m_j}{|x^{(i)} - x^{(j)}|} \tag{5.1.14}$$

and a *complementary Hamiltonian*

$$F_{\text{comp}} := \sum_{1 \le i < j \le n} \frac{X^{(i)} \cdot X^{(j)}}{m_0} \,. \tag{5.1.15}$$

Now we introduce *Poincaré variables* denoted by

$$(\lambda, \Lambda, \xi, \eta, p, q) \in \mathbb{T}^n \times (0, \infty)^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \,.$$

If we define the elliptic elements of the $j - th$ orbit by

- l_j is the average anomaly,
- g_j is the argument of the perihelion,
- θ_j is the longitude of the node,
- a_j is the semi major axis,
- ε_j is the eccentricity,
- i_j is the inclination,
- λ_j is the average longitude,
- $\vec{C}_j$ is the vector of the kinetic momentum,
- Λ_j is the circular momentum,
- and $g_j + \theta_j$ is the longitude of the perihelion,

then the Poincaré variables are related to them by the following formulas:

$$\begin{cases} \lambda_j := l_j + g_j + \theta_j \\ \Lambda_j := \mu_j \sqrt{M_j a_j} \\ G_j := \Lambda_j \sqrt{1 - \varepsilon_j^2} = |\vec{C}_j| \\ r_j := \xi_j + i\eta_j := \sqrt{\Lambda_j} \sqrt{1 - \sqrt{1 - \varepsilon_j^2}} e^{i(g_j + \theta_j)} \\ z_j := p_j + iq_j := \sqrt{G_j} \sqrt{1 - \cos i_j} e^{i\theta} \,. \end{cases} \tag{5.1.16}$$

In the limit we are interested in, i.e., small eccentricities and inclinations ($\varepsilon_j, i_j \to 0$), it results

$$|r_j| = \sqrt{\Lambda_j/2}\epsilon_j \left(1 + O(\epsilon_j^2)\right) \qquad \text{and} \qquad |z_j| = \sqrt{\Lambda_j/2} \left(1 + O(\epsilon_j^2) + O(i_j^2)\right) \, . \tag{5.1.17}$$

In [Poi07, chap. III] the following theorem is stated and proved:

Theorem 5.1.1 (Poincaré). *Poincaré coordinates are analytic and symplectic, with respect to the standard form* $\sum_{1\le j\le n} d\lambda_j \wedge d\Lambda_j + d\xi_j \wedge d\eta_j + dp_j \wedge dq_j$*, in a neighborhood, diffeomorphic to* $\mathbb{T}^n \times \mathbb{R}^{5n}$*, of the union of Keplerian circular direct and coplanar orbits. Moreover, they are angle-action variables for the Keplerian Hamiltonian (5.1.12) that assumes the form*

$$F_{Kep} = \sum_{1\le j\le n} -\frac{\mu_j^3 M_j^2}{2\Lambda_j^2} \tag{5.1.18}$$

(we still denote F_{Kep} *the new Hamiltonian function).*

We define the *average movements* by

$$\nu_j := \frac{\partial F_{\text{Kep}}}{\partial \Lambda_j} = \frac{\mu_j^3 M_j^2}{\Lambda^3} = \frac{\sqrt{M_j}}{a_j^{\frac{3}{2}}} \tag{5.1.19}$$

having used the equation for Λ_j in (5.1.16), where this expression implies the Kepler's third law: the square of the revolution period of a planet is directly proportional to the cube of its semi major axis.

We are now ready to state Arnold's theorem contained in [Arn63b], which initially turned out to be true only in the case of the spatial three body problem until Herman and Féjoz' proof in [Féj04].

Theorem 5.1.2 (Arnold's theorem on planetary motions (real-analytic case)). *For all values of masses* $m_0, m_1, \ldots, m_n$ *and semi major axes* $a_1 > a_2 > \cdots > a_n > 0$*, there exists a real number* $\epsilon_0 > 0$ *such that, for all* $0 < \epsilon < \epsilon_0$*, the flow of the Hamiltonian function* H_{plt} *in (5.1.10) possesses a strictly positive measure set of phase space points, in a neighborhood of circular and coplanar Keplerian tori with semi major axes* $(a_1, a_2, \ldots, a_n)$*, leading to quasi-periodic motions with* $3n-1$ *frequencies. Furthermore, such quasi-periodic motions lay on* $(3n-1)$*-dimensional real-analytic Lagrangian tori.*

5.2 The secular Hamiltonian and its elliptic singularity

To prove Arnold's theorem our aim is to apply theorem 3.1.1 with $H_\epsilon = H_{\text{plt}}$,

$$f_0 = \langle F_{\text{per}} \rangle := \frac{1}{(2\pi)^n} \int_{\mathbb{T}^n} F_{\text{per}}\, d\lambda_1, \dots d\lambda_n \tag{5.2.1}$$

and $f_1 = F_{\text{per}} - \langle F_{\text{per}} \rangle$ (with refer to the first part of section 3.1). The averaged Hamiltonian $\langle F_{\text{per}} \rangle$ is well defined on the "collisionless" manifold $\mathcal{M}$ (see (5.1.11)) and, unless to rearrange the planets, we may assume that the semi major axes belong to the open subset of $\mathbb{R}^n$

$$\mathcal{A} := \{(a_1, a_2, \dots, a_n) \in \mathbb{R}^n \;:\; 0 < a_n < a_{n-1} < \dots < a_1\}\,. \tag{5.2.2}$$

Since λ and Λ are standard symplectic conjugate variables, the Hamiltonian function $\langle F_{\text{per}} \rangle$ possesses the n first integrals $\Lambda_1, \Lambda_2, \dots, \Lambda_n$ (that is Laplace's first theorem of stability). Thus, we may consider $\langle F_{\text{per}} \rangle$ parametrized by $(\Lambda_1, \Lambda_2, \dots, \Lambda_n)$ and defined on the space (diffeomorphic to $\mathbb{R}^{4n}$) of Keplerian tori with fixed semi major axes and coordinates (ξ, η, p, q). We still denote this Hamiltonian function by $\langle F_{\text{per}} \rangle$ and call it the *secular Hamiltonian* of the first order of the planetary system and $\mathcal{A}$ is called the *secular space*. This system describes (at the first order in ϵ) the slow variations along the centuries of Keplerian ellipses which change their shapes under perturbations due to the other planets.

Lemma 5.2.1. *Each term of F_{comp} in (5.1.15) has vanishing average along Keplerian tori, i.e.,*

$$\int_{\mathbb{T}^n} X^{(j)} \cdot X^{(k)}\, d\lambda_1 \dots d\lambda_n = 0 \tag{5.2.3}$$

for every $1 \le j < k \le n$. This implies in particular, that F_{comp} does not give any contribution to the secular Hamiltonian: $\langle F_{per} \rangle = \langle F_{princ} \rangle$.

Proof By (5.1.18) and Hamilton's equations we obtain

$$\dot{\lambda}_j = \frac{\partial F_{\text{Kep}}(\Lambda)}{\partial \Lambda_j} = \frac{\mu_j^3 M_j^2}{\Lambda_j^3}$$

and $\dot{\Lambda} = 0 = \dot{\xi} = \dot{\eta} = \dot{p} = \dot{q}$, while from (5.1.12) it results

$$\dot{x}^{(j)} = \frac{\partial F_{\text{Kep}}(X, x)}{\partial X^{(j)}} = \frac{X^{(j)}}{\mu_j}\,.$$

Thus, considering $X^{(j)}(\lambda_j, \Lambda_j, \xi_j, \eta_j, p_j, q_j)$ and $x^{(j)}(\lambda_j, \Lambda_j, \xi_j, \eta_j, p_j, q_j)$, from the chain rule we get

$$X^{(j)} = \mu_j \dot{x}^{(j)} = \mu_j \frac{\partial x^{(j)}}{\partial \lambda_j} \dot{\lambda}_j = \frac{\mu_j^4 M_j^2}{\Lambda_j^3} \frac{\partial x^{(j)}}{\partial \lambda_j} \tag{5.2.4}$$

and since $x^{(k)}$ does not depend on λ_j ($j \neq k$), we obtain the statement with the help of Fubini's theorem $\square$

The Hamiltonian $\langle F_{\text{per}} \rangle = \langle F_{\text{princ}} \rangle$ is an even function of secular coordinates $(r, z) = (r_1, \ldots, r_n, z_1, \dot{;} z_n)$. As a consequence the origin of the secular system is a critic point for the secular Hamiltonian $\langle F_{\text{per}} \rangle$ corresponding to direct, circular and coplanar movements. The following lemma explains this fact from the point of view of the dynamics generated by the secular Hamiltonian:

Lemma 5.2.2. *The Keplerian torus having vanishing eccentricity and inclination, i.e.,* $r_1 = \cdots = r_n = z_1 = \ldots z_n = 0$ *(see formulas (5.1.17)), is a fixed point for the flow of the secular Hamiltonian function* $\langle F_{per} \rangle$*.*

In view of this result we are going to study Birkhoff's normal form of $\langle F_{\text{per}} \rangle$ in $(r, z) = 0$ at the first order (the first Birkhoff invariants), i.e., the quadratic part of $\langle F_{\text{per}} \rangle$ in Poincaré coordinates r_j and z_j for $j = 1, \ldots, n$. For this reason we introduce a key aspect of the planetary many-body problem that is the coefficients of Laplace:

Definition 5.2.1. *The* **coefficients of Laplace** $b_s^{(k)}(\alpha)$ *are Fourier's coefficients of the function*

$$\frac{1}{(1 + \alpha^2 - 2\alpha \cos \vartheta)^s} = \sum_{k=0}^{\infty} b_s^{(k)}(\alpha) e^{ik\vartheta} \tag{5.2.5}$$

for $\alpha \in [0, 1)$, $\vartheta \in \mathbb{R}$ *and* $s > 0$.

Lagrange and Laplace proved that the quadratic part of the secular Hamiltonian can be written in a remarkable form through the coefficients $b_s^{(k)}(\alpha)$ where α is a function of the semi major axes. The calculations are quite long and difficult and we refer to [Rob95] for a complete and detailed proof of the results provided next.

Lemma 5.2.3. *Let* $m := (m_1, \ldots, m_n)$, $a := (a_1, \ldots, a_n)$, $\xi = (\xi_1, \ldots, \xi_n)$, $\eta = (\eta_1, \ldots, \eta_n)$, $p = (p_1, \ldots, p_n)$ *and* $q = (q_1, \ldots, q_n)$*. There exist two bilinear and symmetric forms* $\mathcal{D}_h = \mathcal{D}_h(m, a)$ *and* $\mathcal{D}_v = D_v(m, a)$ *defined on the tangent space in the origin of the secular space, called respectively horizontal and vertical, depending analytically on the masses and the semi major axes, such that*

$$\langle F_{per} \rangle(\xi, \eta, p, q) = C_0(m, a) + \mathcal{D}_h \cdot (\xi^2 + \eta^2) + \mathcal{D}_v \cdot (p^2 + q^2) + 0(4)$$

where

$$\mathcal{D}_h \cdot \xi^2 = \sum_{1\le j<k\le n} m_j m_k \left(C_1(a_j,a_k) \left(\frac{\xi_j^2}{\Lambda_j} + \frac{\xi_k^2}{\Lambda_k} \right) + 2C_2(a_j,a_k) \frac{\xi_j \xi_k}{\sqrt{\Lambda_j \Lambda_k}} \right)$$

$$\mathcal{D}_v \cdot p^2 = \sum_{1\le j<k\le n} -m_j m_k C_1(a_j,a_k) \left(\frac{p_j}{\sqrt{\Lambda_j}} - \frac{p_k}{\sqrt{\Lambda_k}} \right)^2 \tag{5.2.6}$$

with

$$\begin{aligned} C_0(m,a) &:= -\sum_{1\le j<k\le n} \frac{m_j m_k}{a_j} b_{\frac{1}{2}}^{(0)}(a_k/a_j) \\ C_1(a_j,a_k) &:= -\frac{a_k}{2a_j^2} b_{\frac{3}{2}}^{(1)}(a_k/a_j) \\ C_2(a_j,a_k) &:= \frac{a_k}{2a_j^2} b_{\frac{3}{2}}^{(2)}(a_k/a_j) \ . \end{aligned} \tag{5.2.7}$$

From now on we will identify the two bilinear forms $\mathcal{D}_h$ and $\mathcal{D}_v$ with the matrices representing them with respect to the canonic bases $(d\xi_1, \ldots, \delta\xi_n)$ and $(dp_1, \ldots, \delta p_n)$. Let the masses and the semi major axes be fixed, then there exist two matrices ρ_h and ρ_v in $SO(n)$ that put respectively D_h and D_v into diagonal form, that is

$$\rho_h^t \mathcal{D}_h \rho_h = \sum_{1\le j\le n} \sigma_j \xi_j^2 \qquad \text{and} \qquad \rho_v^t \mathcal{D}_v \rho_h = \sum_{1\le j\le n} \varsigma_j p_j^2 \tag{5.2.8}$$

for $\sigma_1, \ldots, \sigma_n$ and $\varsigma_1, \ldots, \varsigma_n$ in $\mathbb{R}$. The application

$$\rho : (\xi, \eta, p, q) \longmapsto (\rho_h \cdot \xi, \rho_h \cdot \eta, \rho_v \cdot p, \rho_v \cdot q)$$

is symplectic and it results

$$\langle F_{\text{per}} \rangle \circ \rho = C_0(m,a) + \sum_{1\le j\le n} \sigma_j (\xi_j^2 + \eta_j^2) + \sum_{1\le j\le n} \varsigma_j (p_j^2 + q_j^2) + O(4)$$

Definition 5.2.2. *Let $\mathcal{A}$ be the open set defined in (5.2.2), ν_j the average movements in (5.1.19), σ_j and ς_j the eigenvalues of the matrices $\mathcal{D}_h$ and $\mathcal{D}_v$ respectively, we will denote with α the multivalued application*

$$\alpha : a \in \mathcal{A} \longmapsto \{\sigma_1, \ldots, \sigma_n, \varsigma_1, \ldots, \varsigma_n, \nu_1, \ldots, \nu_n\} \subset \mathbb{R}^n$$

and call it the "frequency application".

In the following pages we will prove the following facts

1. for all values of masses and in a simply connected neighborhood of almost every value of the semi major axes, there exists an analytic determination of the frequency application that we will denote by

$$\alpha : a \in \mathcal{A} \longmapsto (\sigma_1, \ldots, \sigma_n, \varsigma_1, \ldots, \varsigma_n, \nu_1, \ldots, \nu_n) \subset \mathbb{R}^{3n} \tag{5.2.9}$$

2. this newly defined application does not meet the non-degeneracy hypothesis in theorem 3.1.1

3. it is not even possible to obtain the non-degeneracy condition considering an auxiliary Hamiltonian function.

We observe that the development of the secular Hamiltonian with respect to the ratio of the semi major axes of the different orbits, will be a key passage. Then, the analycity will permit the generalization of results obtained only for small ratios of semi major axes.

We first recall some well known facts about the ellipse: let u_j and v_j be respectively the eccentric and the true anomaly of the ellipse described by the motion of the j-planet, ε_j be the eccentricity of this ellipse, a_j its semi major axis and denote with $r_j := |x_j|$ the distance between the j-th planet and the "Sun" (one of the foci of the ellipse). Then the following equalities hold

$$\begin{cases} r_j = a_j(1+\varepsilon_j\cos v_j)^{-1}(1-\varepsilon_j^2) & \text{(definition of } v_j) \\ r_j = a_j(1-\varepsilon_j\cos u_j) & \text{(definition of } u_j) \\ d\lambda_j = (1+\varepsilon_j\cos v_j)^{-1}(1-\varepsilon_j^2)^{\frac{3}{2}} dv_j & \text{(Kepler's second law)} \\ d\lambda_j = (1-\varepsilon_j\cos u_j) & \text{(Kepler's equation)} \\ r_j\cos v_j = a_j(\cos u_j - \varepsilon_j) & \\ r_j \sin v_k = a_j\sqrt{1-\varepsilon_j^2}\sin u_j & \end{cases} \tag{5.2.10}$$

where

$$\lambda_j(v_j) = \frac{2}{a_j^2\sqrt{1-\epsilon_j^2}} \,\text{meas}_2\mathcal{E}(v_j) \tag{5.2.11}$$

with $\mathcal{E}(v_j)$ representing the following region of the ellipse

$$\mathcal{E}(v_j) := \left\{ x = x(r'_j, v'_j) \,:\, 0 \le r'_j \le r(v_j),\, 0 \le v'_j \le v_j \right\} .$$

We now state a lemma concerning the development of the secular Hamiltonian for small ratios of the semi major axes:

Lemma 5.2.4. *When the ratio a_k/a_j tends to 0 we obtain following formulas*

$$\begin{aligned} C_0(m,a) &= -\sum_{1\le j<k\le n} \frac{m_j m_k}{a_j}\left(1+\frac{1}{4}\left(\frac{a_k}{a_j}\right)^2 + O\left(\frac{a_k}{a_j}\right)^4\right) \\ C_1(a_j,a_k) &= -\frac{1}{a_j}\left(\frac{3}{4}\left(\frac{a_k}{a_j}\right)^2 + O\left(\frac{a_k}{a_j}\right)^4\right) \\ C_2(a_j,a_k) &= \frac{1}{a_j} O\left(\frac{a_k}{a_j}\right)^3. \end{aligned}$$

(compare with definitions in (5.2.7))

A complete proof of this result can be found in [Las91].

5.3 Verifi cation of Arnold-Pyartli condition

We start this section proving some technical results that we will later apply to the matrices of the bilinear symmetric forms D_h and D_v (defined in (5.2.6)) and their eigenvalues.

Lemma 5.3.1. *Let $\delta_1, \dots \delta_n \in \mathbb{R}$, $D =$ diag $(\delta_1, \dots \delta_n)$ and A an $n \times n$ matrix; then, if σ is an eigenvalue of the matrix $D + A$ it results*

$$\min_{1\le j\le n} |\sigma - \delta_j| \le \|A\|$$

where $\|A\| := \max_{|x|_2=1} |Ax|$.

Proof If $\sigma = \delta_j$ for some $j = 1, \dots, n$ then the statement is trivially verified. Assume $\sigma \neq \delta_j$ for every $j = 1, \dots, n$, then $D + A - \sigma\mathbb{I}_n$ is not invertible and this implies

$$1 \le \left\|(D-\sigma\mathbb{I}_n)^{-1}A\right\| \le \left\|(D-\sigma\mathbb{I}_n)^{-1}\right\| \, \|A\| = \max_{1\le j\le n} \frac{1}{|\sigma-\delta_j|}\|A\|$$

which proves the lemma □

The following lemma shows that the preceding estimate can be improved if the eigenvalues of D are distinct and the terms on the diagonal of A are zero.

Lemma 5.3.2. *Let $\delta_1 < \delta_2 < \dots < \delta_n \in \mathbb{R}$, $D =$ diag $(\delta_1, \dots \delta_n)$, $\sigma := \min_{1\le j\neq k\le n} |\delta_j - \delta_k|$ and $A \in$ mat$_{\mathbb{R}}(n)$ with diag $A := (A_{11}, \dots, A_{nn}) = 0$. Then, if ϵ is sufficiently small, the eigenvalues $\delta'_1 < \delta'_2 \cdots < \delta'_n$ of the matrix $D + \epsilon A$ satisfy*

$$|\delta'_j - \delta_j| \le \frac{3}{\sigma}\|A\|^2 \left(1\frac{\|D\|}{\sigma}\right)\epsilon^2 \tag{5.3.1}$$

Proof Consider the matrix $M = e^{-\epsilon U}(D + \epsilon A)e^{\epsilon U} = D + 0(\epsilon^2)$ for some $n \times n$ matrix U. Developing M in power series of ϵ, one finds that there exists a unique matrix $U \in \ \mathrm{mat}_{\mathbb{R}}(n)$ with null diagonal such that

$$e^{-\epsilon U}(D + \epsilon A)e^{\epsilon U} = D + 0(\epsilon^2) = D + O(\epsilon^2)\,; \tag{5.3.2}$$

U is the solution of $A = UD - DU$ (i.e., $u_{ij} = \frac{a_{ij}}{\delta_j - \delta_i}$ for $i \neq j$) and verifies $\|U\| \le \|A\|/\sigma$. Developing $e^{\pm\epsilon U}$ up to the second order it results

$$e^{-}\epsilon U(D + \epsilon A)e^{\epsilon U} = D + \epsilon^2 A_1 + O(\epsilon^3)$$

where

$$A_1 := \frac{1}{2}(U^2 D + DU^2) + AU - UA - UDU\,.$$

Therefore we may apply the preceding lemma with D as diagonal matrix and $\epsilon^2 A_1 + O(\epsilon^3)$ instead of A obtaining

$$\begin{aligned} |\delta'_j - \delta_j| \;&\le\; \|\epsilon^2 A_1 + 0(\epsilon^3)\| \le \epsilon^2(1 + O(\epsilon))\|A_1\| \le (2\|U\|^2\|D\| + \\ &+\; 2\|A\|\,\|U\|)\epsilon^2(1 + O(\epsilon)) \le \frac{2}{\sigma}\|A\|^2\left(1 + \frac{\|D\|}{\sigma}\right)\epsilon^2(1 + O(\epsilon)) \end{aligned}$$

which proves the statement for sufficiently small ϵ □

Proposition 5.3.1. *Let $\delta_1, \delta_2, \ldots, \delta_{n-1} \in \mathbb{R}$, $\delta_n = 0$ and $\sigma := \min_{1\le j\neq k\le n} |\delta_j - \delta_k| \neq 0$; let $\hat{D}$ be a $(n-1) \times (n-1)$ symmetric matrix with eigenvalues $\delta_1, \ldots \delta_{n-1}$, consider the symmetric matrix D given by*

$$D := \begin{pmatrix} \hat{D} & \mathbb{O}_{n\times 1} \\ & \\ \mathbb{O}_{1\times n} & 0 \end{pmatrix}$$

where $\mathbb{O}_{k\times m}$ represent a null matrix $k \times m$. Moreover let A_ϵ be an $n \times n$ symmetric matrix whose last coefficients is in the form

$$(A_\epsilon)_{nn} = c_1 + c_2\epsilon^\beta$$

with $c_1, c_2 \in \mathbb{R}$ and $0 \le \beta < 2$. Then, when ϵ tends to zero, the matrix $D + \epsilon A_\epsilon$ possesses an eigenvalue in the form

$$\sigma_n(\epsilon) = \epsilon(c_1 + c_2\epsilon^\beta) + O(\epsilon^2)\,. \tag{5.3.3}$$

Furthermore, if $\hat{D}$ is diagonal, then $D + \epsilon A_\epsilon$ is conjugated to a diagonal matrix through a matrix $\rho \in SO_n(\mathbb{R})$ verifying $\rho = \mathbb{I}_n + O(\epsilon)$.

Proof Let $\hat{\rho} \in \mathrm{SO}_{n-1}(\mathbb{R})$ such that

$$\hat{\rho}^t \hat{D} \hat{\rho} = \text{ diag} \left(\delta_1, \dots, \delta_{n-1}\right)$$

we define $\rho \in \mathrm{SO}_n(\mathbb{R})$ by

$$\rho := \begin{pmatrix} \hat{\rho} & \mathbb{O}_{n\times 1} \\ & \\ \mathbb{O}_{1\times n} & 1 \end{pmatrix} .$$

Now observe that $\rho^t(D + \epsilon A_\epsilon)\rho$ possesses the same eigenvalues of $D + \epsilon A_\epsilon$ (they have the same characteristic polynomial) and $\rho^t(D + \epsilon \,\text{diag}\, A_\epsilon)\rho$ possesses $\epsilon (A_\epsilon)_{nn} = \epsilon(c_1 + c_2 e^\beta)$ has last eigenvalues. Thus, for sufficiently small ϵ, we may apply lemma 5.3.2 together with the identity

$$\rho^t(D + \epsilon A_\epsilon)\rho = \rho^t(D + \epsilon \,\text{diag}\, A_\epsilon)\rho + \epsilon\rho^t(A_\epsilon - \text{ diag}\, A_\epsilon)\rho$$

obtaining, by inequality (5.3.1), that $D + \epsilon A_\epsilon$ possesses an eigenvalue $\sigma_n(\epsilon)$ verifying

$$|\sigma_n(\epsilon) - \epsilon(c_1 + c_2 e^\beta)| \le C\epsilon^2 \quad \square \tag{5.3.4}$$

5.3.1 Arnold-Pyartli condition in the plane

We now provide some preliminary results we will use to prove that the bilinear form $\mathcal{D}_h$ in (5.2.6) is definite negative

Lemma 5.3.3. *Let $b_s^{(j)}(\alpha)$ be Laplace's coefficients defined in (5.2.5); then, for any $\alpha \in (0,1)$, $s \in \mathbb{R}_+ \setminus \mathbb{N}$ and $j \in \mathbb{N}$ we have*

$$b_s^{(j)}(\alpha) > b_s^{(j+1)}(\alpha) \tag{5.3.5}$$

Proof We proceed by recurrence on s. Let $s \in (0,1)$, denote

$$s_j := s(s+1)\dots(s+j-1) = \frac{(s+j-1)!}{(s-1)!} ;$$

in [Poi07] the following development of Laplace's coefficients is given

$$b_s^{(j)}(\alpha) = \frac{s_j}{j!}\alpha^j \left(1 + \frac{s}{1!}\frac{s+j}{j+1}\alpha^2 + \frac{s(s+1)}{2!}\frac{(s+j)(s+j+1)}{(j+1)(j+2)}\alpha^4 + \dots\right) ;$$

from this formula we get

$$b_s^{(j)}(\alpha) - b_s^{(j+1)}(\alpha) = \frac{s_j}{j!}\alpha^j \left[\left(1 - \frac{s+j}{j+1}\alpha\right) + \right.$$

$$\left. + \frac{s}{1}\frac{s+j}{j+1}\alpha^2 \left(1 - \frac{s+j+1}{j+2}\alpha\right) + \dots\right] > 0 \, . \tag{5.3.6}$$

Furthermore the two following relation (to be found always in [Poi07]) hold

$$
\begin{aligned}
b_{s+1}^{(j)} &= \frac{(s+j)(1+\alpha^2)}{s(1-\alpha^2)} b_s^{(j)} - 2\frac{j-s+1}{s}\frac{\alpha}{(1-\alpha^2)^2} b_s^{(j+1)} \\
b_s^{(j+2)} &= \frac{j+1}{j-s+2}\left(\alpha+\frac{1}{\alpha}\right) b_s^{(j+1)} - \frac{j+s}{j-s+2} b_s^{(j)}
\end{aligned}
$$

so that we may infer

$$
b_{s+1}^{(j)} - b_{s+1}^{(j+1)} = \frac{jb_s^{(j)} + (j+1)b_s^{(j+1)}}{s(1+\alpha)^2} + \frac{b_s^{(j)} - b_s^{(j+1)}}{(1+\alpha)^2} \, ;
$$

this last equality shows that (5.3.6) holds for every $s \in \mathbb{R}_+ \setminus \mathbb{N}$ as stated $\square$

Lemma 5.3.4 (Hadamard). *Let Q be the matrix of a bilinear symmetric form in $\mathbb{R}^2$, whose coefficients (q_{jk}) satisfy the following proprieties: $q_{jj} < 0$ for all $1 \leq j \leq n$, $q_{jk} > 0$ for all $1 \leq j \neq k \leq n$ and the quantity*

$$
\delta q_j := -q_{jj} - \sum_{l \neq j} q_{jl} > 0
$$

(that is Q is a matrix with strictly dominant diagonal). Then the bilinear form associated to the matrix Q is negative.

Proof Let $v = (v_1, \ldots, v_n)$ a vector in $\mathbb{R}^n$, by the hypotheses made it results

$$
\langle v, Qv \rangle = \sum_{j,k=1}^{n} q_{jk} v_j v_k = -\sum_{j=1}^{n} \delta q_j v_j^2 - \sum_{j<k} q_{jk}(v_j - v_k)^2 \leq 0 \quad \square
$$

We are now ready to state

Proposition 5.3.2. *The bilinear form $\mathcal{D}_h$ in (5.2.6) is definite negative*

Proof Define $\Delta := \operatorname{diag}(\sqrt{\Lambda_1}, \ldots, \sqrt{\Lambda_n})$ and consider the matrix $\hat{\mathcal{D}}_h := \Delta^t \mathcal{D}_h \Delta$. Let q_{jk} for $j, k = 1, \leq n$ the coefficients of $\hat{\mathcal{D}}_h$, from the first equation in (5.2.6) and the definitions in (5.2.7) we get

$$
\begin{aligned}
q_{jj} &= -\sum_{1 \leq k < j} m_j m_k \frac{a_j}{2a_k^2} b_{\frac{3}{2}}^{(1)}(a_j/a_k) - \sum_{j<k\leq n} m_k m_k \frac{a_k}{a_j^2} b_{\frac{3}{2}}^{(1)}(a_k/a_j) \\
q_{jk} &= m_k m_k \frac{a_k}{2a_j^2} b_{\frac{3}{2}}^{(2)}(a_k/a_j) \, , \quad \text{if} \quad j < k \\
q_{jk} &= m_j m_k \frac{a_j}{2a_k^2} b_{\frac{3}{2}}^{(2)}(a_j/a_k) \, , \quad \text{if} \quad j > k \, .
\end{aligned}
$$

Thus, $q_{jj} < 0$ for every $j = 1, \dots, n$, while $q_{jk} > 0$ for every $1 \leq j \neq k \leq n$; now denoting $\delta q_j = -q_{jj} - \sum_{1 \leq k \neq j \leq n} q_{jk}$, by the preceding formulas and inequality (5.3.5) with $s = \frac{3}{2}$, $j = 1$ and $\alpha = \frac{a_k}{a_j}, \frac{a_j}{a_k}$ it results $\delta q_j > 0$. Therefore $\hat{\mathcal{D}}_h$ is a matrix with strictly dominant diagonal from which we infer (using lemma 5.3.4) that $\hat{\mathcal{D}}_h$ negative. To conclude we need to prove that $\hat{\mathcal{D}}_h$ is definite. Let $v \in R^n$ any vector such that $\hat{\mathcal{D}}_h \cdot v = 0$ and let $j \in \{1, \dots, n\}$ such that $|v_j| = \max_{k=1,\dots,n} |v_k|$. Observe that we may suppose $v_j > 0$, eventually considering $-v$ instead of v; then the j-th component of the vector $\hat{\mathcal{D}}_h \cdot v$ is given by

$$0 = q_{jj} v_j + \sum_{k \neq j} q_{jk} v_k \leq \left(q_{jj} + \sum_{k \neq j} q_{jk} \right) v_j = -\delta q_j v_j \tag{5.3.7}$$

which yields a contradiction to $v_j > 0$ in view of $\delta q_j > 0$. Therefore $v_j = 0$ which proves $v = 0$. As a consequence $\hat{\mathcal{D}}_h$ is definite negative and so is $\mathcal{D}_h$ □

Now let $m_0, m_1, \dots, m_n$ be fixed and consider the open set $\mathcal{A}$ in (5.2.2); we need to prove that the subset of $\mathcal{A}$ where the horizontal form $\mathcal{D}_h$ does not possess double eigenvalues is connected, and for this purpose we are going to complexify the application $a \in \mathcal{A} \longmapsto \mathcal{D}_h(m, a)$.

Lemma 5.3.5. *Laplace's coefficient $b_s^{(k)}(\alpha)$ can be extended to meromorphic function on Riemann's sphere possessing the only four poles $0, \pm 1$ and ∞.*

Proof This lemma runs as a consequence of the differential equation satisfied by Laplace's coefficients that can be found in [Poi07, page 252]:

$$\alpha^2(1-\alpha^2) \frac{d^2 b_s^{(k)}}{d\alpha^2} + \left(\alpha - (4s+1)\alpha^3\right) \frac{d b_s^{(k)}}{d\alpha} - \left(4s^2\alpha^2 + k^2(1-\alpha^2)\right) b_s^{(k)} = 0 \ \Box$$

From this lemma we obtain that Laplace's coefficients $b_s^{(k)}$ can be holomorphically extended for on the complex set $\{\alpha \in \mathbb{C} \ : \ |\alpha| < 1, \ \ \mathrm{Re}\ \alpha > 0\}$. Then, from formulas (5.2.6) and (5.2.7) we get that the quadratic form $\mathcal{D}_h$ can be holomorphically extended on the following connected subset of $\mathcal{A}$:

$$\mathbb{A} := \left\{ (a_1, \dots, a_n) \in (\mathbb{C} \setminus \{0\})^n \ : \ |a_k/a_j| < 1, \ \ \mathrm{Re}\ (a_k/a_j) > 0, \ \ \forall\ j < k \right\} . \tag{5.3.8}$$

Proposition 5.3.3 (Arnold-Pyartli condition in the plane). *For all $n \geq 2$ there exists an open set of $U \subset \mathcal{A}$ (defined in (5.2.2)) with full Lebesgue measure on which the matrix $\mathcal{D}_h$ verifies the following property: for any open and simply connected set $V \subset U$, the eigenvalues of $\mathcal{D}_h$ define n holomorphic function $\sigma_1, \dots, \sigma_n : V \longrightarrow \mathbb{C}$ such that the frequency application of the planetary system*

$$a \in V \longmapsto (\sigma_1, \dots, \sigma_n, \nu_1, \dots, \nu_n) , \tag{5.3.9}$$

where the average movements ν_j are defined in (5.1.19), is non-planar on V.

Proof We proceed by induction on the number of planets. For $n = 2$ the first formula in (5.2.6) gives

$$\mathcal{D}_h(m,a) = m_1 m_2 \begin{pmatrix} C_1(a_1,a_2)\Lambda_1^{-1} & C_2(a_1,a_2)(\Lambda_1\Lambda_2)^{-\frac{1}{2}} \\ C_2(a_1,a_2)(\Lambda_1\Lambda_2)^{-\frac{1}{2}} & C_1(a_1,a_2)\Lambda_2^{-1} \end{pmatrix};$$

thus, for $a_1 = O(1)$ and $a_2 \longrightarrow 0$, equations (5.2.12), together with $\Lambda_2 = O(\sqrt{a_2})$ from (5.1.16), yield

$$\mathcal{D}_h(m,a) = -m_1 m_2 \frac{3}{8}\frac{a_2^{3/2}}{a_1^3}\left[\begin{pmatrix} \sqrt{a_2}\Lambda_1^{-1} & 0 \\ 0 & \sqrt{a_2}\Lambda_2^{-1} \end{pmatrix} + \begin{pmatrix} a_2 & a_2^{3/4} \\ a_2^{3/4} & \sqrt{a_2} \end{pmatrix}\right].$$

If we write this last expression as follows

$$-m_1 m_2 \frac{3}{8}\frac{a_2^{3/2}}{a_1^3}\left[\begin{pmatrix} \sqrt{a_2}\Lambda_1^{-1} + O(a_2) & 0 \\ 0 & \sqrt{a_2}\Lambda_2^{-1} + O(\sqrt{a_2}) \end{pmatrix} + \begin{pmatrix} 0 & a_2^{3/4} \\ a_2^{3/4} & 0 \end{pmatrix}\right].$$

we are in a position to apply lemma 5.3.2 where the role of A is played by the second matrix in the sum, (symmetric with all zeros on its diagonal), while the first matrix obviously possesses distinct eigenvalues. Thus, if a_2 is sufficiently small, we obtain that also $\mathcal{D}_h$ possesses distinct eigenvalues in the form

$$\begin{aligned} \sigma_1 &= -m_1 m_2 \frac{3}{8}\frac{a_2^2}{a_1^3\Lambda_1}(1 + O(\sqrt{a_2})) \\ \sigma_2 &= -m_1 m_2 \frac{3}{8}\frac{a_2^2}{a_1^3\Lambda_2}(1 + O(\sqrt{a_2}))\,. \end{aligned} \qquad (5.3.10)$$

In particular the discriminant of the characteristic polynomial of $\mathcal{D}_h$ is not a constant function as holomorphic function on $\mathbb{A}$ (defined in (5.3.8)). This implies that the subset of $\mathbb{A}$ where $\mathcal{D}_h$ possesses a double eigenvalue has strictly positive codimension (in the complex plane). Denote with $\mathbb{A}'$ its complementary in $\mathbb{A}$. Then, $\mathbb{A}'$ is connected and contains, in view of what we have proved before with $a_2 \longrightarrow 0$ and $a_1 = O(1)$, a subset having the form $\{(a_1,a_2) \in (\mathbb{R}_+)^2 \;:\; 0 < a_2/a_1 < \epsilon\}$ for some $0 < \epsilon < 1$. The eigenvalues of $\mathcal{D}_h$ define two holomorphic functions $\sigma_1, \sigma_2 : \mathbb{A}'_\star \longrightarrow \mathbb{C}$ where $\mathbb{A}'_\star$ denotes the universal recovering of $\mathbb{A}'$.

Now we consider the frequency application

$$\alpha : a \in \mathbb{A}'_\star \longmapsto (\nu_1, \nu_2, \sigma_1, \sigma_2) \in \mathbb{R}^4$$

where ν_j is defined in (5.1.19). Suppose that there exists an open subset V of $\mathbb{A}'_\star$ where the frequencies satisfy a linear relation

$$c_1\nu_1 + c_2\nu_2 + c_3\sigma_1 + c_4\sigma_2 = 0$$

for some $c_j \in \mathbb{R}$. Since $\mathbb{A}'_\star$ is connected, the principle of analytic continuation states that this linear relation is satisfied on the whole $\mathbb{A}'_\star$. Now, from the definition of ν_j in (5.1.19) and the form of σ_j in (5.3.10), if a_2 tends to zero, we see that $c_1 = c_2 = c_3 = c_4 = 0$. This means that the frequency application α does not satisfy any linear relation on any open set $V \subset \mathbb{A}'_\star$; therefore, from lemma 2.1.1 (negation of the second part) we obtain that the holomorphic function

$$\det\left[\alpha, \frac{d\alpha}{da_2}, \frac{d^2\alpha}{da_2^2}, \frac{d^3\alpha}{da_2^3}\right]$$

is not constantly vanishing on any open set contained $\mathbb{A}'_\star$. This implies the existence of a dense and full Lebesgue measure set $U \subset \mathcal{A}$, where the frequency application defined almost every and mapping continuously $a \in U$ in $\alpha \in \mathbb{R}^4$ is non-planar in the sense of definition 2.1.1.

Now consider $n \geq 3$ and suppose that the statement is verified up to order $n-1$. Consider the case $a_1, \ldots, a_{n-1} = O(1)$ and $a_n \longrightarrow 0$; using the formulas in (5.2.12) and denoting $\hat{\mathcal{D}}_h$ the matrix of the horizontal form at the order $n-1$, we have

$$\mathcal{D}_h = \begin{pmatrix} \hat{\mathcal{D}}_h + O(a_n^{5/2}) & O(a_n^{5/2}) \\ \\ O(a_n^{5/2}) & b_{nn} \end{pmatrix}$$

for some $b_{nn} \in \mathbb{R}$. Applying proposition 5.3.1 we obtain that $\mathcal{D}_h$ possesses an eigenvalue $\sigma_n \neq 0$ that tends to zero with a_n while the other eigenvalues are in the form $\sigma_j = \hat{\sigma}_j + O(a_n^{5/2})$ where $\hat{\sigma}_j$ denote the eigenvalue of $\hat{\mathcal{D}}_h$ for $j = 1, \ldots, n-1$. Proceeding analogously to what done before in the case $n = 2$, we will obtain that the eigenvalues $\sigma_1, \ldots, \sigma_n$ are distinct on an open and dense subset U of $\mathbb{A}$ with full Lebesgue measure; moreover $\sigma_1, \ldots, s_n$ together with the average movements $\nu_1, \ldots, \nu_n$ define a frequency application that is non-planar on every open set V contained in U $\square$

5.3.2 Arnold-Pyartli condition in the space

We now show that the frequency application in (5.2.9) satisfies only two linear relations. Before stating this result we need a bit of preparation.

Lemma 5.3.6. *The bilinear form $\mathcal{D}_v$ is positive.*

Proof It is a direct consequence of formula (5.2.6) together with (5.2.7) $\square$

Lemma 5.3.7. *The bilinear form $\mathcal{D}_v$ possesses a null eigenvalue we will denote $\varsigma_n = 0$*

Proof As it can be easily seen by the formula in (5.2.6), the vector $v = (\sqrt{\Lambda_1}, \ldots, \sqrt{\Lambda_n})$ belongs to the Kernel of $\mathcal{D}_v$, indeed

$$\mathcal{D}_h \cdot v = \sum_{j<k} -m_j m_k C_1(a_j, a_k) \begin{pmatrix} \Lambda_j^{-1} & (\Lambda_j \Lambda_k)^{-1/2} \\ (\Lambda_j \Lambda_k)^{-1/2} & \Lambda_j^{-1} \end{pmatrix} \cdot \begin{pmatrix} \sqrt{\Lambda_j} \\ \sqrt{\Lambda_k} \end{pmatrix} = 0 \ \square$$

Lemma 5.3.8 (Herman). *The trace of the matrix $\mathcal{D}_h + \mathcal{D}_v$ is null.*

Proof This property can be immediately verified through the formulas for $\mathcal{D}_h$ and $\mathcal{D}_v$ in (5.2.6). Curiously this fact has not been remarked in its generality before Herman $\square$

Proposition 5.3.4. *For all $n \geq 2$ there exists an open and dense set with full Lebesgue measure $U \subset \mathcal{A}$ where the eigenvalues of $\mathcal{D}_h$ and $\mathcal{D}_v$ are distinct two by two and satisfy the following property: for any open and simply connected set $V \subset U$, the eigenvalues of $\mathcal{D}_h$ and $\mathcal{D}_v$ define $2n$ holomorphic functions $\sigma_1, \ldots, \sigma_n, \varsigma_1, \ldots, \varsigma_n : V \longrightarrow \mathbb{C}$ which together with the average movements $\nu_1, \ldots, \nu_n$ satisfy this linear relations only:*

$$\sum_{j=1}^{n} (\sigma_j + \varsigma_j) = 0 \qquad \textit{and} \qquad \varsigma_n = 0\,. \tag{5.3.11}$$

Proof We start with the case $n = 2$ observing that the sum of the eigenvalues of a 2×2 matrix $A = (a_{ij})$ is given by its trace $a_{11} + a_{22}$. This fact, together with lemmata 5.3.8 and 5.3.7 shows that both relations in (5.3.11) are satisfied. Besides, from proposition 5.3.3 we may infer that this are the only linear relations satisfied by $\sigma_1, \sigma_2, \varsigma_1, \varsigma_2, \nu_1, \nu_2$ (an any open set V contained in an open and dense subset $U \subset \mathcal{A}$ with full measure), since $\sigma_1, \sigma_2, \nu_1, \nu_2$ do not verify any linear relation, $\varsigma_2 = 0$ and ς_1 equals the trace of $\mathcal{D}_v$. Now let $n \geq 3$ and suppose that the statement is verified up to order $n-1$; as done in the induction in proposition 5.3.3 we consider the semi major axis a_n tending to 0 and denote $\hat{\mathcal{D}}_v$ the matrix of the form $\mathcal{D}_v$ at the order $n-1$; thus, with the application of proposition 5.3.1 we obtain what stated $\square$

If we consider the submanifold of the phase space obtained by choosing the total angular momentum to be vertical, we lose the frequency ς_n related to the invariance under rotation. Arnold in [Arn63b] infers that the frequency application

$(\sigma_1, \ldots, \sigma_n, \varsigma_1, \ldots, \varsigma_{n-1})$, considered as local function on the phase space with values in $\mathbb{R}^{3n-1}$ is a submersion; the preceding proposition shows that this is not true, since this frequency application verifies a linear relation and is therefore contained in a linear subspace of $\mathbb{R}^{3n-1}$ with codimension 1.

As proposition 5.3.4 shows we are not in a position to apply theorem directly 3.1.1 since the frequency application is degenerate in the sense of definition 2.1.2. Denote now $C = (C_x, C_y, C_z) \in \mathbb{R}^3$ the total angular momentum of the planetary system; in [Poi07] it is shown how the components of C can be expressed in Poincaré variables (see (5.2.10)) as follows

$$C_x + iC_y = \sum_{1\leq j\leq n} z_j \sqrt{2\Lambda_j - |r_j|^2 - \frac{1}{2}|z_j|^2}$$
$$C_z = \sum_{1\leq j\leq n} \left[\Lambda_j - \frac{1}{2}\left(|r_j|^2 + |z_j|^2\right)\right] . \tag{5.3.12}$$

To avoid the degeneracy of the frequency application shown before, Arnold, in the case of the three-body problem, choses to consider the symplectic submanifold with vertical total angular momentum. On the other hand, Herman seems to add a linear combination of C_x^2 , C_y^2 and C_z^2 to avoid the null trace relation (first equation in (5.3.11)). Féjoz, in the paper we are reviewing, choses an intermediate strategy which consists in considering first a spinning reference and then fixing vertically the total angular momentum ($C_x = C_y = 0$).

Consider the Hamiltonian function

$$H_\delta := H_{\text{plt}} - \delta C_z \tag{5.3.13}$$

where δ is a real parameter and H_{plt} is defined in (5.1.10). H_δ is the Hamiltonian function of the planetary problem in a reference spinning with speed δ with respect to the initial Galilean reference. The origin of the secular space is an elliptic critic point for C_z (as it is for H_{plt}) whose quadratic part in this point is

$$\mathcal{D}_C = -\frac{1}{2}\sum_{1\leq j\leq n}\left(|r_j|^2 + |z_j|^2\right) \tag{5.3.14}$$

as it can be easily seen by (5.3.12). Denoting by $\mathcal{D}$ the quadratic part of $\langle F_{\text{per}}\rangle$, we have that the quadratic part of H_δ is given by

$$\mathcal{D}_\delta = \mathcal{D} - \delta\mathcal{D}_C . \tag{5.3.15}$$

Similarly to the quadratic part of $\langle F_{\text{per}}\rangle$, $\mathcal{D}_\delta$ possesses $4n$ eigenvalues with double multiplicity, corresponding to $2n$ frequencies that form an "extended frequency application" $(\delta, a) \longmapsto \alpha(\delta, a)$ extending α in (5.2.9) for every $\delta \neq 0$.

Lemma 5.3.9. *The image of the extended frequency application* $\alpha(\delta, a)$ *satisfies the only linear relation*

$$\varsigma_n = 0 \tag{5.3.16}$$

Proof The expression of $\mathcal{D}_C$ in (5.3.14) is equivalent to $(\mathcal{D}_C)_{ij} = -\frac{1}{2}\delta_{ij}$ (where δ_{ij} are the classical Kronecker symbols); this implies that the trace of $\mathcal{D}_C$ equals $-2n$, which together with proposition 5.3.4 proves the lemma $\square$

Now denote by $\mathcal{M}_{\text{vert}}$ the symplectic submanifold of the secular space described by the equations $C_x = 0$ and $C_y = 0$, locally diffeomorphic to $\mathbb{R}^{4n-2}$. Let $\widehat{\langle H_\delta \rangle}$ be the restriction to $\mathcal{M}_{\text{vert}}$ of the averaged Hamiltonian $\langle H_\delta \rangle$ and $\hat{\mathcal{D}}_\delta := \frac{1}{2} D^2 \widehat{\langle H_\delta \rangle}(0,0)$ its quadratic part in $(r,z) = (0,0)$. The bilinear form $\hat{\mathcal{D}}_\delta$ define $4n-2$ eigenvalues which corresponds to $2n-1$ frequencies described by the following lemma:

Lemma 5.3.10. *For all* $n \geq 2$ *the eigenvalues of* $\hat{\mathcal{D}}_\delta$ *define* $2n-1$ *frequencies given by* $\sigma_1, \ldots, \sigma_n, \varsigma_1, \ldots, \varsigma_{n-1}$*, with the same notation of proposition 5.3.4. In particular they do not locally satisfy any linear relation.*

Proof In view of proposition 5.3.4 and lemma 5.3.7, the eigenspace associated to the eigenvalue $\sigma_n = 0$ is generated by the values of the Hamiltonian vector fields $\vec{C}_x$ and $\vec{C}_y$ in the origin $(r,z) = (0,0)$ of the secular space, i.e.,

$$\vec{C}_x(0,0) = \sum_{1 \leq j \leq n} \sqrt{2\Lambda} \frac{\partial}{\partial q_j} \qquad \text{and} \qquad \vec{C}_y(0,0) = -\sum_{1 \leq j \leq n} \sqrt{2\Lambda} \frac{\partial}{\partial p_j}\,.$$

Moreover, the tangent space in the origin of symplectic submanifold $\mathcal{M}_{\text{vert}}$ is described by the equation

$$D(C_x + iC_y)(0,0) = \sum_{1 \leq j \leq n} \sqrt{2\Lambda} \partial z_j \tag{5.3.17}$$

and is therefore orthogonal to the eigenspace associated to the null eigenvalue of the form $\mathcal{D}$. Thus, the eigenvalues of $\hat{\mathcal{D}}_\delta$ coincide with the eigenvalues of $\mathcal{D}$ except zero $\square$

From this last result, the facts that H_{plt} is real-analytic and the space of parameters is connected we can infer

Proposition 5.3.5. *For all* $n \geq 2$ *there exists an open and dense set* $U \subset \mathcal{A} \times \mathbb{R}$ *with full Lebesgue measure such that the* $2n-1$ *frequencies associated to the eigenvalues of* $\hat{\mathcal{D}}_\delta$*, regarded as function of* $a \in \mathcal{A}$ *and* $\delta \in \mathbb{R}$*, are all distinct and satisfy the following property: for every open and simply connected* $V \subset U$ *this frequencies define* $2n-1$ *holomorphic functions* $\sigma_1, \ldots, \sigma_n, \varsigma_1, \ldots, \varsigma_{n-1}$:

$V \longrightarrow \mathbb{C}$ which, together with the average movements $\nu_1, \ldots, \nu_n$, do not satisfy any linear relation. In particular the frequency application

$$\hat{\alpha} : (\Lambda, \delta) \in (\mathbb{R}_+)^n \times \mathbb{R} \longrightarrow (\sigma_1, \ldots, \sigma_n, \nu_1, \ldots, \nu_n, \varsigma_1, \ldots, \varsigma_{n-1}) \in \mathbb{R}^{3n-1} \tag{5.3.18}$$

is non-degenerate in the sense of Rüßmann on an open and dense subset of $\mathcal{A} \times \mathbb{R}$ having full Lebesgue measure.

5.4 Proof of Arnold's theorem

Now we fix the masses $m = (m_0, \ldots, m_n)$, the semi major axes $a = (a_1, \ldots, a_n)$ and the parameter δ so that the frequency application of the planetary $(n+1)$-body problem defined in (5.3.18), regarded as function of $\Lambda \in (\mathbb{R}_+)^n$ and depending on parameters $a, m \in \mathbb{R}^n$ and $\delta \in \mathbb{R}$, is non-degenerate (it is sufficient to avoid a closed set in $\mathcal{A} \times \mathbb{R}$ having null Lebesgue measure). We aim to apply theorem 3.1.1 to H_δ, in particular with $h = F_{\text{Kep}}$, $f = F_{\text{per}}$, $f_0 = \langle F_{\text{per}} \rangle$ and $f_1 = F_{\text{per}} - \langle F_{\text{per}} \rangle$, under the notations of section 3.1. We first need two preliminary results. This first lemma is an extension of theorem 3.1.1 in the case H depends on additional parameters.

Lemma 5.4.1. *Suppose that the Hamiltonian function H described at the beginning of section 3.1 depends on an additional parameter $\delta \in \mathbb{R}$ and assume that the frequency application*

$$(I, \delta) \in \mathcal{B} \times \mathbb{R} \longrightarrow (\omega(I; \delta), \Omega_1(I; \delta), \ldots, \Omega_p(I; \delta)) \in \mathbb{R}^d \times \mathbb{R}^p$$

is non-degenerate in the sense of Rüßmann. Then, if ϵ is sufficiently small, there exists a subset $\Delta \subset \mathbb{R}$ with strictly positive Lebesgue measure such that for every fixed $\delta \in \Delta$ the set of phase space points leading to quasi-periodic motions, laying on analytic Lagrangian KAM tori, has strictly positive Lebesgue measure.

Proof Under the hypothesis of non-degeneracy formulated here, theorem 3.1.1 assures the existence of a strictly positive Lebesgue measure set $B \times J \subset \mathbb{R}^{d+p} \times \mathbb{R}$ giving rise to quasi-periodic motions laying on analytic Lagrangian tori. Therefore, by Fubini's theorem, there exists a subset $\Delta \subset J$, with strictly positive Lebesgue measure, such that for any $\delta \in \Delta$ the measure of the points in B leading to quasi-periodic motions is strictly positive. □

Lemma 5.4.2. *Since the two Hamiltonian function H_{plt} and H_δ commute (because C_z is an integral for the system described by H_{plt}), a Lagrangian ergodic invariant torus for the flow of H_δ is automatically invariant for the flow of H_{plt}*

Proof Let $\mathbb{T}$ be an H_δ-invariant ergodic torus and denote Φ_t and Ψ_t the flows of H_{plt} and H_δ respectively. Let $t > 0$ be a sufficiently small fixed time and consider $\widetilde{\mathbb{T}} := \Phi_t(\mathbb{T})$. Since $\mathbb{T}$ is a Lagrangian (maximal) torus and Φ_t is symplectic, then $\widetilde{\mathbb{T}}$ is Lagrangian too. Moreover, since the flows Φ_t and Ψ_t commute, $\widetilde{\mathbb{T}}$ is Ψ_t-invariant. A classical argument of Lagrangian intersections shows that $\widetilde{\mathbb{T}} \cap \mathbb{T} \neq \emptyset$. Finally, since $\Psi_t|_{\mathbb{T}}$ is ergodic it results $\widetilde{\mathbb{T}} = \mathbb{T}$ which implies that $\mathbb{T}$ is invariant for the flow of H_{plt} $\square$

Let $\mathcal{M}_{\text{vert}}$ the symplectic submanifold of the phase space of H_{plt} corresponding to vertical total angular momentum. From proposition 5.3.5 and lemma 5.4.1 the Hamiltonian function H_δ (see (5.3.13)) possesses a strictly positive set of points in $\mathcal{M}_{\text{vert}}$ leading to quasi-periodic motions with $3n-1$ frequencies. By lemma 5.4.2 we obtain that H_{plt} (see (5.1.10)) possesses the same positive measure subset of points in $\mathcal{M}_{\text{vert}}$ belonging to quasi-periodic motions laying on real-analytic maximal invariant tori with $3n-1$ frequencies.

Appendix A

Proof of Kolmogorov's 1954 theorem on persistence of quasi-periodic motions

A.1 Introduction

In 1954 A.N. Kolmogorov showed evidence of the following theorem:

Theorem A.1.1 (Kolmogorov). *Let H be an Hamiltonian in the form $H(y,x) = E + \omega \cdot y + Q(y,x) + \epsilon P(y,x)$ where Q and P are real-analytic functions over $B^d \times \mathbb{T}^d$ (here B^d is an euclidean ball in $\mathbb{R}^d$) with $\partial_y^\alpha Q(0,x) = 0$ for $|\alpha| \leq 1$, $\omega \in \mathbb{R}^d$, $E \in \mathbb{R}$.*
Assume that

$$\det\langle Q_{yy}(0,\cdot)\rangle = \det \int_{\mathbb{T}^d} Q_{yy}(0,x)\,\frac{dx}{(2\pi)^d} \neq 0$$

then for almost all $\omega \in \mathbb{R}^d$ there exists ϵ_0 such that for all $|\epsilon| \leq \epsilon_0$ there exists Φ_ϵ symplectic diffeomorphism which maps H into the Hamiltonian $N_\epsilon = E_\epsilon + \omega \cdot y' + Q_\epsilon(y',x')$, with $\partial_{y'}^\alpha Q_\epsilon(0,x') = 0$ for $|\alpha| \leq 1$ and where we have denoted $(x,y) = \Phi_\epsilon(y',x')$.
Besides we have that $|E_\epsilon - E|$, $\|Q_\epsilon - Q\|_{C^1}$ and $\|\Phi_\epsilon -\ id\,\|_{C^1}$ are all $O(\epsilon)$.

Our aim is to give a proof of this theorem following the original ideas gave by Kolmogorov itself and focusing our attention on the estimate, in terms of some constants depending on different parameters, of the size of ϵ_0. We are interested in particular in the dependence of ϵ_0 from the diophantine constant γ because it is strictly related to the dimension of invariant tori in the phase space for the perturbed Hamiltonian H (we will discuss this matter in appendix B). For an elegant and extremely authoritative proof performed adopting a slightly different scheme refer to [Arn63a]; our proof is instead inspired by the original scheme suggested by Kolmogorov and is based on [Chi05b].

In order to explain how we are going to proceed, we want now to give an equivalent, but in some way more "quantitative" version of Kolmogorov's theorem. Let $\Omega \in \mathbb{C}^d$ we define the following sets:

$$
\begin{aligned}
\Omega_r &:= \bigcup_{x_0 \in \Omega} \{x \in \mathbb{C}^d : |x - x_0| < r\}\,, \\
\mathbb{T}^d_\sigma &:= \{x \in \mathbb{C}^d : |\text{ Im } x_j| < \sigma,\ \text{Re } x_j \in \mathbb{T}\ \ \forall\ j = 1 \ldots d\}\,, \\
\mathcal{D}^d_{\gamma,\tau} &:= \{\omega \in \mathbb{R}^d : |\omega \cdot n| > \frac{\gamma}{|n|^\tau},\ \ \forall\ n \in \mathbb{Z}\}\,;
\end{aligned}
$$

we shall refer to an element $\omega \in \mathcal{D}^d_{\gamma,\tau}$ as a Diophantine γ, τ vector. Let $f : \Omega \to \mathbb{R}$ be a real-analytic function on an open set $\Omega \subseteq \mathbb{R}^d$ with analytic complex extension on

$$
\Omega_r = \bigcup_{x_0 \in \Omega} \{x \in \mathbb{C}^d : |x - x_0| < r\}
$$

we put

$$
|f|_r = \sup_{\Omega_r} |f|;
$$

if $f : \mathbb{T}^d \to \mathbb{R}$ is real-analytic with complex extension on $\mathbb{T}^d_\sigma$ we define

$$
|f|_\sigma = \sup_{\mathbb{T}^d_\sigma} |f|;
$$

if $f : \Omega \times \mathbb{T}^d \to \mathbb{R}$ is real-analytic with complex extension on the cartesian product $\Omega_r \times \mathbb{T}^d_\sigma$ we naturally put

$$
|f|_{r,\sigma} = \sup_{\Omega_r \times \mathbb{T}^d_\sigma} |f|.
$$

The same definitions can be obviously given if f is a function whose analytic extension assumes values in $\mathbb{C}^n$ or $\text{mat}_{\mathbb{C}}(n \times n)$, where in this case $|\cdot|$ is some appropriate norm in the space considered. The theorem we are going to prove is the following:

Theorem A.1.2. *Let $H(y, x) = E + \omega \cdot y + Q(y, x) + \epsilon P(y, x)$ be a real-analytic Hamiltonian over $B^d \times \mathbb{T}^d$ with analytic extension for P and Q on the complex domain $B^d_r \times \mathbb{T}^d_\sigma$, for some $r > 0$ and $0 < \sigma \leq 1$ and $\omega \in \mathcal{D}^d_{\gamma,\tau}$. Suppose $Q(0, x) = \partial_y Q(0, x) = 0$ and*

$$
\det\langle Q_{yy}(0, \cdot)\rangle \neq 0\ .
$$

Let $\sigma_\infty < \sigma, r_\infty < r$ take

$$\begin{aligned}
\mu &= |P|_{r,\sigma} \\
M &= \max\left\{\frac{1}{r}|Q|_{r,\sigma}, |Q_y|_{r,\sigma}, r|Q_{yy}|_{r,\sigma}\right\} \\
\lambda &= \max\left\{\frac{1}{\sigma_\infty}, \frac{1}{\sigma-\sigma_\infty}\right\} \\
\nu &= \max\left\{\frac{r}{r_\infty}, \frac{r}{r-r_\infty}\right\} \\
S &= \frac{1}{r}|\langle Q_{yy}(0,\cdot)\rangle^{-1}| \\
Z &= |\omega| \ ;
\end{aligned}$$

and define

$$\begin{aligned}
\Gamma_1 &= \max\left\{M\gamma^{-1}, 1\right\} \\
\Gamma_2 &= \max\{MS, 1\} \\
\Gamma_3 &= \max\{ZS, 1\} \\
\Gamma_4 &= \max\left\{M^{-1}, S\right\}
\end{aligned}$$

there exist a positive constant $c(\tau, d) \geq 1$ such that if

$$\epsilon CD\mu < 1$$

where

$$C = c\nu^{14}\lambda^{4q}r^{-1}\Gamma_1^4\Gamma_2^4\Gamma_3\Gamma_4$$

and $D = 2^{12(\tau+d)+30}$, then there exists a symplectic diffeomorphism

$$\Phi : (y', x') \in B^d_{r_\infty} \times \mathbb{T}^d_{\sigma_\infty} \to (y, x) \in B^d_r \times \mathbb{T}^d_\sigma$$

which puts the Hamiltonian H into Kolmogorov's normal form

$$N'(y', x'; \epsilon) = E'(\epsilon) + \omega \cdot y' + Q'(y', x'; \epsilon) = H \circ \Phi;$$

we also have that $|E' - E|$, $\|Q' - Q\|_{C^1} \leq \epsilon CD\mu Mr$ and $\|\Phi - id\|_{C^1} \leq \epsilon CD\mu r$.

A.1.1 Some useful estimates

We now define $\mathcal{H}(\Omega_r), \mathcal{H}(\mathbb{T}^d_\sigma), \mathcal{H}(\Omega_r \times \mathbb{T}^d_\sigma)$ as the spaces of real-analytic functions having holomorphic extension on the prescribed domain and finite norm (respectively $|f|_r, |f|_\sigma$ or $|f|_{r,\sigma} < \infty$). We now state the following lemma:

Lemma A.1.1 (Cauchy's estimate). *Let* $f \in \mathcal{H}(\Omega_r)$, $\forall\ p \in \mathbb{N}^d$ *and* $\forall\ 0 < \rho < r$ *we have:*

$$\left|\partial_y^p f(y)\right|_\rho \le \frac{p!}{(r-\rho)^{|p|_1}}|f|_r$$

Proof The proof of this lemma can be easily obtained by Cauchy's integral formula for analytic functions □

Observe that Lemma A.1.1 can be immediately generalized to $f \in \mathcal{H}(\mathbb{T}^d_\sigma)$ or $\mathcal{H}(\Omega_r \times \mathbb{T}^d_\sigma)$.

We now define , for $\omega \in \mathcal{D}^d_{\gamma,\tau}$, the operator

$$\mathcal{D}_\omega = \sum_{i=1}^d \omega_i \partial_{x_i}.$$

Suppose to have a function $f \in \mathcal{H}(\mathbb{T}^d_\sigma)$, we are interested in solving the equation

$$\mathcal{D}_\omega u = f\ . \tag{A.1.1}$$

First recall that $f(y,x) \in C(B^p(\bar{y}) \times \mathbb{R}^d, \mathbb{R}^m)$, 2π-periodic in the second variables, is analytic if and only if there exist positive numbers M, r and ξ such that its Fourier's coefficients $f_{k,n}$ satisfy

$$\|f_{k,n}\|_\infty \le M r^{-|k|_1} e^{-|n|_1 \xi}. \tag{A.1.2}$$

Now observe that if $u(x) = \sum_{n \in \mathbb{Z}^d} \hat{u}_n e^{in\cdot x}$ is the Fourier series for u, then

$$\mathcal{D}_\omega u = \sum_{n \in \mathbb{Z}^d} in \cdot \omega \hat{u}_n e^{in\cdot x}$$

so it is easily verified that $\langle \mathcal{D}_\omega u \rangle = 0$ (the Fourier coefficient corresponding to $n = 0$ is zero). So to solve equation (A.1.1) we must necessarily require $\langle f \rangle = 0$. We can now expand f in its Fourier series obtaining $f(x) = \sum_{n \in \mathbb{Z}^d \setminus \{0\}} \hat{f}_n e^{in\cdot x}$ so that equation (A.1.1) becomes

$$\sum_{n \in \mathbb{Z}^d \setminus \{0\}} in \cdot \omega \hat{u}_n e^{in\cdot x} = \sum_{n \in \mathbb{Z}^d \setminus \{0\}} \hat{f}_n e^{in\cdot x}$$

and hence

$$\hat{u}_n = \frac{\hat{f}_n}{in \cdot \omega} \ .$$

We observe now that

$$u(x) = \sum_{n \in \mathbb{Z}^d \setminus \{0\}} \frac{\hat{f}_n}{in \cdot \omega} \ e^{in \cdot x}$$

converges absolutely by equation (A.1.2) (here $p = 0$ so that k does not appear) and by the diophantine estimate satisfied by ω. We can now state

Lemma A.1.2. *Let $f \in \mathcal{H}(\mathbb{T}^d_\sigma)$ and $\omega \in \mathcal{D}^d_{\gamma,\tau}$; if u is the only solution to $\mathcal{D}_\omega u = f$ with $\langle u \rangle = 0$, then there exists $c = c(\tau, d)$ such that*

$$|u|_{\sigma-\delta} \leq \frac{c}{\gamma} \frac{|f|_\sigma}{\delta^{d+\tau}}$$

Proof We have the following inequalities:

$$\begin{aligned} |u|_{\sigma-\delta} &\leq \left| \sum_{n \neq 0} \frac{\hat{f}_n}{in \cdot \omega} \ e^{in \cdot x} \right|_{\sigma-\delta} \leq \sum_{n \neq 0} \frac{|f|_\sigma}{|n \cdot \omega|} e^{-|n|\sigma} |e^{in \cdot x}|_{\sigma-\delta} \\ &\leq \sum_{n \neq 0} \frac{|n|^\tau}{\gamma} |f|_\sigma e^{-|n|\sigma} e^{|n|(\sigma-\delta)} = \frac{|f|_\sigma}{\gamma} \sum_{n \neq 0} e^{-|n|\delta} |n|^\tau \end{aligned}$$

where we have used equation (A.1.2) for f with $p = 0$ and $\xi = \delta$, while it effectively results by calculus that we can chose $M = |f|_\sigma$. We want now to estimate $\sum_{n \neq 0} e^{-|n|\delta} |n|^\tau$. Approximating the sum with an integral we have

$$\begin{aligned} \sum_{n \neq 0} e^{-|n|\delta} |n|^\tau &= c' \int_{\mathbb{R}^d} e^{-|x|\delta} |x|^\tau \ dx = \frac{c'}{\delta^\tau} \int_{\mathbb{R}^d} e^{-|\delta x|} |\delta x|^\tau \ dx = \\ &= \frac{c'}{\delta^{\tau+d}} \int_{\mathbb{R}^d} e^{-|y|} |y|^\tau \ dy = \frac{c(\tau, d)}{\delta^{\tau+d}} \end{aligned}$$

and the lemma is proved □

Combining this two preceding lemmata and simultaneously generalizing the result in Lemma A.1.2 to further inversions of the operator $\mathcal{D}_\omega$, we obtain

Lemma A.1.3. *Let $f \in \mathcal{H}(\mathbb{T}^d_\sigma)$ with $\langle f \rangle = 0$; for every choice of $\alpha \in \mathbb{N}^d$, $p \in \mathbb{N}$ we have :*

$$|\partial^\alpha \mathcal{D}_\omega{}^{-p} f|_{\sigma-\delta} \leq c(\tau, d, p, \alpha) \frac{|f|_\sigma}{\gamma^p \delta^{p\tau + d + |\alpha|_1}} \ .$$

A.1.2 Diffeomorphisms on $\mathbb{T}^d$

Consider $a \in \mathcal{H}(\mathbb{T}^d_\sigma)$ and the following analytic function on $\mathbb{T}^d$:

$$\phi : x \in \mathbb{T}^d \to \phi(x) = x + a(x) \in \mathbb{T}^d \,;$$

we want to give sufficient conditions on a in order to obtain that ϕ is an analytic diffeomorphism on $\mathbb{T}^d$. Our aim is to provide an inverse analytic function for ϕ, that is to say, $\tilde{\phi}(x') = x' + \tilde{a}(x')$ such that $\phi \circ \tilde{\phi} = \text{ id } = \tilde{\phi} \circ \phi$. Let's see what does this mean in terms of a and $\tilde{a}$:

$$\phi \circ \tilde{\phi}(x') = x' \iff \tilde{\phi}(x') + a(\tilde{\phi}(x')) = x' \iff$$

$$\iff x' + \tilde{a}(x') + a(\tilde{\phi}(x')) = x' \iff \tilde{a}(x') = -a(x' + \tilde{a}(x'))\,.$$

We now state the following lemma:

Lemma A.1.4. *Let $a \in \mathcal{H}(\mathbb{T}^d_\xi)$ and take $\xi' < \xi$ such that $|a|_\xi \le \xi - \xi'$ and $|a_x|_\xi < 1$; then $\exists!\ \tilde{a} \in \mathcal{H}(\mathbb{T}^d_{\xi'})$ with $|\tilde{a}|_{\xi'} \le \xi - \xi'$ such that:*

$$-a(x' + \tilde{a}) = \tilde{a}.$$

Proof We initially define the following space

$$\mathcal{X} = \{b \in \mathcal{H}(\mathbb{T}^d_{\xi'}) : |b|_{\xi'} \le \xi - \xi'\} \;;$$

$\mathcal{X}$ is a closed non-empty subset of the Banach space $\mathcal{H}(\mathbb{T}^d_{\xi'})$ and therefore is a Banach space itself. Let $\Phi : b(x') \in \mathcal{X} \to -a(x' + b(x')) \in \mathcal{H}(\mathbb{T}^d_{\xi'})$ we state that Φ is a contraction in $\mathcal{X}$; in fact for every choice of b and $c \in \mathcal{X}$ we have:

1. $|\text{ Im } x' + b(x')| \le |\text{ Im } x'| + |b(x')| < \xi' + |b(x')| \le \xi$ and this implies, by the hypotheses done on a, that $|\Phi(b)|_{\xi'} = |a(x' + b(x'))|_{\xi'} < \xi - \xi'$.
2. $|\Phi(b) - \Phi(c)|_{\xi'} = |a(x' + b(x')) - a(x' + c(x'))|_{\xi'} \le |a_x|_\xi |b - c|_{\xi'} < |b - c|_{\xi'}$ by the fundamental calculus Theorem applied on a .

Thesis follows from Banach fixed point Theorem $\square$

Observe that for any $a \in \mathcal{H}(\mathbb{T}^d_{\bar{\xi}})$, with $\bar{\xi} > \xi$, by Lemma A.1.1 we can estimate $|a_x|_\xi$ as follows:

$$|a_x|_\xi = \sup_{\mathbb{T}^d_\xi} |a_x| = \sup_{\mathbb{T}^d_\xi} \sup_i \sum_{j=1}^d \left| \frac{\partial a_i}{\partial x_j} \right| \le \sup_i \sum_{j=1}^d \frac{|a|_{\bar{\xi}}}{\bar{\xi} - \xi} = d \frac{|a|_{\bar{\xi}}}{\bar{\xi} - \xi}.$$

Now combining this last estimate and Lemma A.1.4 taking $\xi = \bar{\xi}$, we have

Proposition A.1.1. *Let $a \in \mathcal{H}(\mathbb{T}^d_\xi)$ and let $\xi' < \xi$ such that $|a|_\xi < \frac{\xi - \xi'}{d+1}$ then $\exists!\ \tilde{a} \in \mathcal{H}(\mathbb{T}^d_{\xi'})$ with $|\tilde{a}|_{\xi'} \le |a|_\xi$ and $-a(x' + \tilde{a}) = \tilde{a}$; therefore $\phi(x) = x + a(x)$ is an analytic diffeomorphism on $\mathbb{T}^d$.*

A.2 Kolmogorov's idea and first step of the proof

A.2.1 Reduction of the perturbation to order ϵ^2

Let $H^{(0)} = N^{(0)} + \epsilon P^{(0)} = E + \omega \cdot y + Q^{(0)}(y,x) + \epsilon P^{(0)}$ the analytic Hamiltonian in Kolmogorov's theorem on the phase space $\mathcal{U} := B^d \times \mathbb{T}^d$, with refer to the standard symplectic form

$$dy \wedge dx := \sum_{i=1}^{d} dy_i \wedge dx_i$$

(that is to say that Hamilton's equations are $\dot{x} = H_y, \dot{y} = -H_x$). Recall that $\omega \in \mathcal{D}^d_{\gamma,\tau}$ and $P, Q \in \mathcal{H}(B^d_r \times \mathbb{T}^d_\sigma)$ with Q quadratic in y. The first step (and main idea) to prove the theorem, is to find a symplectic transformation Φ which maps $H^{(0)}$ into $H^{(1)}$ that is still the sum of an Hamiltonian in Kolmogorov's normal form and a perturbation, but whose perturbative part is of order ϵ^2.

Proposition A.2.1. *Consider $H^{(0)}$ as previously defined and suppose to have*

$$\det\langle Q_{yy}(0,\cdot)\rangle \neq 0 \tag{A.2.1}$$

There exists a symplectic transformation $\Phi : (y', x') \to (y, x)$ generated by the second species function $F(y', x) = y' \cdot x + \epsilon g(y', x)$ where

$$g(y', x) = b \cdot x + s(x) + a(x) \cdot y'$$

for some $b \in \mathbb{R}^d$, $s : \mathbb{T}^d \to \mathbb{R}$ and $a : \mathbb{T}^d \to \mathbb{R}^d$ both analytic functions, such that

$$H^{(0)} \circ \Phi = H^{(1)} = E^{(1)} + \omega \cdot y' + Q^{(1)}(y', x') + \epsilon^2 P^{(1)}(y', x')$$

with $Q^{(1)}$ quadratic in y' and $Q^{(1)}, P^{(1)}$ real-analytic functions.

Proof By the definition of $F(y', x)$ we have the implicit definition of Φ given by:

$$\begin{cases} x' &= \frac{\partial F}{\partial y'} &= x + \epsilon a(x) \\ \\ y &= \frac{\partial F}{\partial x} &= y' + \epsilon(b + s_x(x) + (a_x(x))^T \cdot y') \end{cases}$$

Assume that $\varphi(x) = x' = x + \epsilon a(x)$ is a diffeomorphism on $\mathbb{T}^d$ with inverse $\tilde{\varphi}(x') = x = x' + \epsilon \tilde{a}(x')$. Following the Hamilton-Jacobi proceeding we aim to express $H^{(0)}(y, x)$ in the new variables (y', x'); notice that we will often leave x

instead of $\tilde{\varphi}(x')$ for simplicity, and we will not sometime use the apex 0 since there's no ambiguity for the moment. By Taylor's formula we have:

$$\begin{aligned}
H(y,x) &= H(y'+\epsilon g_x, x) = H(y',x) + \epsilon H_y(y',x)\cdot g_x + \epsilon^2 \tilde{P}_1(y',x) = \\
&= H(y',x) + \epsilon\left[N_y(y',x) + \epsilon P_y(y',x)\right]\cdot g_x + \epsilon^2 \tilde{P}_1(y',x) = \\
&= H(y',x) + \epsilon N_y(y',x)\cdot g_x + \epsilon^2 \tilde{P}_2(y',x) = \\
&= N(y',x) + \epsilon\left[P(y',x) + N_y(y',x)\cdot g_x\right] + \epsilon^2 \tilde{P}_2(y',x) \qquad \text{(A.2.2)}
\end{aligned}$$

where we have put $\tilde{P}_1(y',x) = \int_0^1 (1-t)H_{yy}(y'+t\epsilon g_x, x)\langle g_x, g_x\rangle\, dt$ and obviously $\tilde{P}_2(y',x) = P_y(y',x)\cdot g_x + \tilde{P}_1(y',x)$. We now focus our attention on $H(y',x) + \epsilon N_y(y',x)\cdot g_x$ in order to put it into the desired Kolmogorov's normal form with at least a perturbative part of order ϵ^2. Recalling that for an analytic function f we have $\mathcal{D}_\omega f = \omega\cdot f_x$ we obtain:

$$\begin{aligned}
& N_y(y',x)\cdot g_x = (\omega + Q_y)\cdot(b + s_x + (a_x)^T\cdot y') = \\
&= \omega\cdot b + \omega\cdot s_x + \omega\cdot(a_x)^T\cdot y' + Q_y\cdot(b+s_x) + Q_y\cdot(a_x)^T\cdot y' = \\
&= \omega\cdot b + \mathcal{D}_\omega s + \mathcal{D}_\omega a\cdot y' + Q_y\cdot(b+s_x) + \tilde{Q}_1(y',x)\,.
\end{aligned}$$

with

$$\tilde{Q}_1(y',x) = Q_y\cdot(a_x)^T\cdot y'\,.$$

Now by Taylor's formula applied on $Q_y(y',x)$, and recalling that $Q_y(0,x) = 0$, we have

$$N_y(y,x)\cdot g_x = \omega\cdot b + \mathcal{D}_\omega s + \mathcal{D}_\omega a\cdot y' + Q_{yy}(0,x)\cdot y'\cdot(b+s_x) + \tilde{Q}_2(y',x)$$

where we have naturally put

$$\tilde{Q}_2(y',x) = \tilde{Q}_1(y',x) + \left(\int_0^1 (1-t)Q_{yyy}(ty',x)\,dt\right)\langle y', y', b+s_x\rangle\,.$$

Combining the expression found for $N_y(y,x) \cdot g_x$ and equation (A.2.2), reorganizing the terms and applying Taylor's formula on $P(y',x)$, we obtain:

$$H(y,x) = N(y',x) + \epsilon[P(y',x) + \omega \cdot b + \mathcal{D}_\omega s + \mathcal{D}_\omega a \cdot y' +$$

$$+Q_{yy}(0,x) \cdot y' \cdot (b+s_x) + \tilde{Q}_2(y',x)] + \epsilon^2 \tilde{P}_2(y',x) =$$

$$= E + \epsilon(\omega \cdot b) + \omega \cdot y' + Q(y',x) + \epsilon[P(0,x) + P_y(0,x) \cdot y' +$$

$$+\tilde{Q}_3(y',x) + \mathcal{D}_\omega s + \mathcal{D}_\omega a \cdot y' + Q_{yy}(0,x) \cdot y' \cdot (b+s_x)]$$

$$+\epsilon^2 \tilde{P}_2(y',x) \tag{A.2.3}$$

having defined

$$\tilde{Q}_3(y',x) = \tilde{Q}_2(y',x) + \left(\int_0^1 (1-t) P_{yy}(ty',x)\, dt \right) \langle y', y' \rangle \, .$$

Starting from the equation (A.2.3) we want now to determine b , s and a. Observe that since $\tilde{Q}_3(0,x) = 0$ we have:

$$[\ldots]_{y'=0} = P(0,x) + \mathcal{D}_\omega s = (P(0,x) - \langle P(0,x) \rangle + \mathcal{D}_\omega s) + \langle P(0,x) \rangle$$

so taking

$$s(x) = \mathcal{D}_\omega{}^{-1}(P(0,x) - \langle P(0,x) \rangle) \tag{A.2.4}$$

it results $[\ldots]_{y'=0} = \langle P(0,x) \rangle$.

For what concerns the linear part in y' we want to maintain the same frequency ω of $H^{(0)}$. Since the term $\omega \cdot y'$ is already given by $N(y',x)$ we have to require

$$P_y(0,x) + \mathcal{D}_\omega a \cdot + Q_{yy}(0,x) \cdot (b + s_x) = 0. \tag{A.2.5}$$

By averaging we have

$$\langle P_y(0,\cdot) \rangle + \langle Q_{yy}(0,\cdot) \cdot b \rangle + \langle Q_{yy}(0,\cdot) \cdot s_x(\cdot) \rangle = 0$$

and by hypotheses $\langle Q_{yy}(0,\cdot) \rangle$ is invertible so that we can take

$$b = -\langle Q_{yy}(0,\cdot) \rangle^{-1} \langle P_y(0,\cdot) + Q_{yy}(0,\cdot) \cdot s_x(\cdot) \rangle \tag{A.2.6}$$

in order to have the average of the left member in (A.2.5) to be 0. We are now able to solve equation (A.2.5) taking

$$a = -\mathcal{D}_\omega{}^{-1} \left(P_y(0,x) + Q_{yy}(0,x) \cdot (b + s_x) \right) \tag{A.2.7}$$

In conclusion by (A.2.3) , (A.2.4) , (A.2.6) and (A.2.7) we have:

$$H(y,x) = H \circ \Phi(y',x') = H^{(1)}(y',\tilde{\varphi}(x')) = N^{(1)}(y',\tilde{\varphi}(x'))$$

$$+\epsilon^2 P^{(1)}(y',\tilde{\varphi}(x')) = E^{(1)} + \omega \cdot y' + Q^{(1)}(y',\tilde{\varphi}(x')) + \epsilon^2 P^{(1)}(y',\tilde{\varphi}(x')) \, .$$

where:

$$E^{(1)} = E + \epsilon(\omega \cdot b + \langle P(0,\cdot)\rangle) \, ; \tag{A.2.8}$$

$$Q^{(1)}(y',\tilde{\varphi}(x')) = Q(y',\tilde{\varphi}(x')) + \epsilon \tilde{Q}_3(y',\tilde{\varphi}(x')) \, ; \tag{A.2.9}$$

$$P^{(1)}(y',\tilde{\varphi}(x')) = \tilde{P}_2(y',\tilde{\varphi}(x')) = P_y(y',\tilde{\varphi}(x')) \cdot g_x(\tilde{\varphi}(x')) +$$

$$+ \quad \int_0^1 (1-t) H_{yy}(y' + t\epsilon g_x, \tilde{\varphi}(x')) \langle g_x(\tilde{\varphi}(x')), g_x(\tilde{\varphi}(x'))\rangle \, dt \, . \tag{A.2.10}$$

More expressly we recall that $\tilde{Q}_3 = Q_1 + Q_2 + Q_3$ with

$$\begin{aligned}
Q_1(y',\tilde{\varphi}(x')) &= \tilde{Q}_1(y',\tilde{\varphi}(x')) = Q_y(y',\tilde{\varphi}(x')) \cdot (a_x)^T(\tilde{\varphi}(x')) \cdot y' \\
Q_2(y',\tilde{\varphi}(x')) &= \left(\int_0^1 (1-t) Q_{yyy}(ty',\tilde{\varphi}(x')) \, dt\right) \langle y', y', b + s_x(\tilde{\varphi}(x'))\rangle \\
Q_3(y',\tilde{\varphi}(x')) &= \left(\int_0^1 (1-t) P_{yy}(ty',\tilde{\varphi}(x')) \, dt\right) \langle y', y'\rangle.
\end{aligned} \tag{A.2.11}$$

To end the proof we observe that $Q^{(1)}(y',\tilde{\varphi}(x'))$ is quadratic in y' so that $N^{(1)}$ is effectively in the desired Kolmogorov's normal form $\square$

Lemma A.2.1. *The non-degeneracy condition holds for $N^{(1)}(y',\tilde{\varphi}(x'))$ as found in proposition A.2.1, that is:*

$$\det\langle Q^{(1)}_{yy}(0,\cdot)\rangle \neq 0$$

Proof $Q^{(1)}(y',\tilde{\varphi}(x')) = Q(y',\tilde{\varphi}(x')) + \epsilon\tilde{Q}_3(y',\tilde{\varphi}(x'))$ so by derivation and averaging we have

$$\langle Q^{(1)}_{yy}(y',\cdot)\rangle = \langle Q_{yy}(y',\cdot)\rangle + \epsilon\langle \partial^2_y \tilde{Q}_3(y',\cdot)\rangle = \langle Q_{yy}(y',\cdot)\rangle + o(\epsilon)$$

Thesis follows for small enough ϵ , since $\det\langle Q_{yy}(y',\cdot)\rangle \neq 0$ by hypotheses. We postpone for the moment the discussion with full details on the estimate of how small must ϵ be in order to have $\langle Q^{(1)}(y',\tilde{\varphi}(x'))\rangle$ invertible $\square$

A.2.2 Control on the domain of Φ

Recall that $H^{(0)} = N^{(0)} + \epsilon P^{(0)} = E + \omega \cdot y + Q^{(0)}(y,x) + \epsilon P^{(0)}$ with $\omega \in \mathcal{D}^d_{\gamma,\tau}$ for some fixed $\gamma \in \mathbb{R}$, and $P, Q \in \mathcal{H}(B^d_r \times \mathbb{T}^d_\sigma)$. Let $\sigma_\infty < \sigma < 1$ and $r_\infty < r$ we define

$$\begin{aligned}
M &:= \max\left\{\frac{1}{r}|Q|_{r,\sigma},\ |Q_y|_{r,\sigma},\ r|Q_{yy}|_{r,\sigma}\right\} \\
S &:= \frac{1}{r}\left|\langle Q_{yy}(0,\cdot)^{-1}\rangle\right| \\
\lambda &:= \max\left\{\frac{1}{\sigma_\infty}\ ,\ \frac{1}{\sigma-\sigma_\infty}\right\} \\
\nu &:= \max\left\{\frac{r}{r_\infty}\ ,\ \frac{r}{r-r_\infty}\right\} \\
Z &:= |\omega| \\
\mu &:= |P|_{r,\sigma}
\end{aligned}$$

We want now to give estimates on $|g|$ in order to apply proposition A.1.1 to $g(y',x) = b\cdot x + s(x) + a(x)\cdot y'$ obtaining that the application

$$\varphi(x) \longmapsto \frac{\partial F}{\partial y'} = x + \epsilon a(x) = x'$$

is effectively a diffeomorphism on $\mathbb{T}^d$ and by consequence so is $\tilde{\varphi}(x') \longmapsto x' + \epsilon\tilde{a}(x') = x$, the first component of Φ . Recall that we have $b \in \mathbb{R}^d$ and by definition of s and a in equations (A.2.4) , (A.2.6) , (A.2.7) and lemma A.1.2 there exists $0 < \delta < \sigma$ such that $s \in \mathcal{H}(\mathbb{T}^d_{\sigma-\delta})$ and $a \in \mathcal{H}(\mathbb{T}^d_{\sigma-2\delta})$; here δ is the loss of analycity due to the inversion of the operator $\mathcal{D}_\omega$.

Remark A.2.1. *Let $\rho < r$ and $\delta < \sigma$ respectively the losses of analycity in the action and angles variables; combining lemmata A.1.1 and A.1.2, for any $f \in \mathcal{H}(\mathbb{T}^d_\sigma)$ or $\mathcal{H}(B^d_r \times \mathbb{T}^d_\sigma)$ this two estimate hold:*

$$\left|\partial^l_x \mathcal{D}_\omega{}^{-1} f(x)\right|_{\sigma-\delta} \ \leq\ \frac{c}{\gamma}\,\frac{|f|_\sigma}{\delta^q} \tag{A.2.12}$$

$$\left|\partial^p_y \partial^l_x f(y,x)\right|_{r-\rho,\sigma-\delta} \ \leq\ c\,\frac{|f|_{r,\sigma}}{\rho^{|p|_1}\delta^q} \tag{A.2.13}$$

where we take the same constant $c \geq 1$ for both inequalities and for any f scalar or vectorial function, matrix or tensor and where $q = d + \tau + 2$ (since we will have at the most $|l|_1 = 2$).

Lemma A.2.2. *There exists a constant $c_1 \geq 1$ depending on $q = \tau + d$, and $B_1 \geq 1$ depending on Q and P, such that for all $0 < \delta < \sigma - \sigma_\infty$*

$$\max\left\{|s|_{\sigma-\frac{\delta}{2}}, |s_x|_{\sigma-\frac{\delta}{2}}, |b|, |a|_{\sigma-\delta}\, r, |a_x|_{\sigma-\delta}\, r, |g_x|_{r,\sigma-\delta}\right\} \leq c_1 B_1 \delta^{-2q} r\,.$$

Proof Using inequalities (A.2.12) and (A.2.13) we estimate separately all terms, reminding the definitions of s , b and a in (A.2.4), (A.2.6) and (A.2.7). First of all we have

$$|s|_{\sigma-\frac{\delta}{2}}\,,\ |s_x|_{\sigma-\frac{\delta}{2}} \leq \frac{c}{\gamma} 2^q \delta^{-q} |P(0,x) - \langle P(0,x)\rangle|_\sigma \leq$$

$$\leq \frac{c}{\gamma} 2^{q+1} \delta^{-q}\, |P(0,x)|_\sigma \leq \frac{\bar{c}}{\gamma} \delta^{-q} |P(y,x)|_{r,\sigma} \leq \bar{c}\delta^{-q}\mu\gamma^{-1}$$

with $\bar{c} = c2^{q+1}$.

Furthermore we may estimate

$$|b| = |\langle Q_{yy}(0,\cdot)\rangle^{-1} \langle P_y(0,\cdot) + Q_{yy}(0,\cdot)\cdot s_x(\cdot)\rangle| \leq$$

$$\leq Sr \sup_{\mathbb{T}^d} \left(|P_y(0,x)| + |Q_{yy}(0,x)\cdot s_x(x)|\right) \leq$$

$$\leq Sr \left(\sup_{B^d\times\mathbb{T}^d_\sigma} |P_y(y,x)| + \sup_{B^d\times\mathbb{T}^d_\sigma} |Q_{yy}(y,x)||s_x(x)|_{\sigma-\frac{\delta}{2}} \right) \leq$$

$$\leq Sr \left(c\,\frac{\mu}{r} + c\frac{M}{r}\,\bar{c}\delta^{-q}\mu\gamma^{-1}\right) \leq c\bar{c}S\mu r^{-1}\left(1 + Mr^{-1}\,\delta^{-q}\gamma^{-1}\right) \leq$$

$$\leq c\bar{c}S\mu\delta^{-q}\left(1 + M\gamma^{-1}\right) \leq c'S\mu\delta^{-q}A_1$$

where we define the first auxiliary constant

$$A_1 := \max\left\{M\gamma^{-1},\, 1\right\}$$

and $c' := 2c\bar{c} = c^2 2^{q+2}$.

Using (A.2.12) and (A.2.13) once again, from the definition of a in (A.2.7) we get

$$
\begin{aligned}
&|a|_{\sigma-\delta}, |a_x|_{\sigma-\delta} \leq \frac{c}{\gamma}\frac{2^q}{\delta^q}|P_y(0,x) + Q_{yy}(0,x)\cdot(b+s_x)|_{\sigma-\frac{\delta}{2}} \leq \\
&\leq \frac{\bar{c}}{\gamma}\delta^{-q}\left[\sup_{B^d\times\mathbb{T}^d_\sigma}|P_y(y,x)| + \sup_{B^d\times\mathbb{T}^d_\sigma}|Q_{yy}(y,x)|\left(|b| + |s_x(x)|_{\sigma-\frac{\delta}{2}}\right)\right] \leq \\
&\leq \frac{\bar{c}}{\gamma}\delta^{-q}\left[c\frac{\mu}{r} + c\frac{M}{r}\left(|b| + |s_x|_{\sigma-\frac{\delta}{2}}\right)\right] \leq \\
&\leq \frac{\bar{c}}{\gamma}\delta^{-q}\left[c\mu r^{-1} + cMr^{-1}\left(c'S\mu\delta^{-q}A_1 + \bar{c}\delta^{-q}\mu\gamma^{-1}\right)\right] \leq \\
&\leq c\bar{c}c'\delta^{-2q}\gamma^{-1}\mu r^{-1}\left[1 + \left(MSA_1 + M\gamma^{-1}\right)\right] \leq \\
&\leq c\bar{c}c'\delta^{-2q}\gamma^{-1}\mu r^{-1}\left[1 + A_1\left(MS+1\right)\right] \leq \\
&\leq \hat{c}\delta^{-2q}\gamma^{-1}\mu r^{-1}A_1A_2
\end{aligned}
$$

where

$$A_2 := \max\{MS,\, 1\}$$

and we take $\hat{c} := 2c\bar{c}c' = c^3 2^{2q+4}$. By using the preceding estimates we have

$$
\begin{aligned}
&|g_x(y',x)|_{r,\sigma-\delta} = |b + s_x(x) + a_x(x)^T\cdot y'|_{r,\sigma-\delta} \leq \\
&\leq |b| + |s_x|_{\sigma-\delta} + |a_x|_{\sigma-\delta}|y'| \leq \\
&\leq c'S\mu\delta^{-q}A_1 * \bar{c}\delta^{-q}\mu\gamma^{-1} + \hat{c}\delta^{-2q}\gamma^{-1}\mu r^{-1}A_1A_2 r \leq \\
&\leq \hat{c}\delta^{-2q}\left[S\mu A_1 + \mu\gamma^{-1} + \mu\gamma^{-1}A_1A_2\right] \leq \\
&\leq 2\hat{c}\delta^{-2q}rA_1A_2\left[S\mu r^{-1} + \mu r^{-1}\gamma^{-1}\right] \leq \hat{c}\delta^{-2q}rA_1A_2A_3
\end{aligned}
$$

where

$$A_3 = \mu r^{-1}\max\left\{S,\, \gamma^{-1}\right\};$$

observe that A_3 is linear in μ and so is the final estimate that proves the lemma with $c_1 = 4\hat{c} = c^3 2^{2q+6}$ and $B_1 = A_1A_2A_3$ □

We can now state the following

Proposition A.2.2. *There exists $c_2 \geq c_1$ such that if*

$$\epsilon c_2 B_1 \rho^{-1} \delta^{-2q} r < 1 \tag{A.2.14}$$

then

1. *$\varphi(x) = x + \epsilon a(x)$, with a as in (A.2.7), is an analytic diffeomorphism on $\mathbb{T}^d$.*

2. *If $\tilde{\varphi}(x') = x' + \epsilon \tilde{a}(x'; \epsilon)$ is its inverse, we have*

$$|\tilde{a}|_{\sigma - \frac{3}{2}\delta} \leq |a|_{\sigma-\delta} \leq c_1 \delta^{-2q} B_1 \,.$$

3. *$\tilde{\varphi} : \mathbb{T}^d_{\sigma-\frac{3}{2}\delta} \longmapsto \mathbb{T}^d_{\sigma-\delta}$, $\varphi : \mathbb{T}^d_{\sigma-2\delta} \longmapsto \mathbb{T}^d_{\sigma-\frac{3}{2}\delta}$ and $\tilde{\varphi} \circ \varphi = id = \varphi \circ \tilde{\varphi}$ on $\mathbb{T}^d_{\sigma-2\delta}$.*

4. *Let $\rho < r$ then $\forall\, y' \in B_{r-\rho}$, $x \in \mathbb{T}^d_{\sigma-2\delta}$ we have $y' + t\epsilon g_x(y', x) \in B_{r-\frac{\rho}{2}}$, $\forall\, t \in [0,1]$; in particular $y = y' + \epsilon g_x(y', x) \in B_{r-\frac{\rho}{2}}$*

Proof The first three statements follow directly from proposition A.1.1 with $\xi = \sigma - \delta$ and $\xi' = \sigma - \frac{3}{2}\delta$ and by taking $c_2 = 2c_1(d+1)$ so that the condition $|a|_{\sigma-\delta} < \frac{\xi-\xi'}{d+1} = \frac{\delta}{2(d+1)}$ holds by the estimate in the preceding lemma. Again by lemma A.2.2 and by hypotheses we obtain $\epsilon |g_x(y', x)|_{r-\rho, \sigma-2\delta} \leq \frac{\rho}{2}$ so that the last statement is also proved $\square$

By this proposition we are now able to control domain and codomain of Φ; therefore we may use the following estimates:

$$|P^{(1)} \circ \Phi|_{r-\rho, \sigma-2\delta} \leq |P|_{r-\frac{\rho}{2}, \sigma-\delta}$$

$$|Q_i \circ \Phi|_{r-\rho, \sigma-2\delta} \leq |Q_i|_{r-\frac{\rho}{2}, \sigma-\delta} \qquad \text{for } i = 1, 2, 3$$

A.2.3 Estimates on $E^{(1)} - E^{(0)}$, $Q^{(1)} - Q^{(0)}$ and $P^{(1)}$

To complete the first step of the proof of Kolmogorov's theorem we want now to estimate the difference between the energies and the quadratic parts of $N^{(0)}$ and $N^{(1)}$, and the size of the new perturbation $P^{(1)}$.

Lemma A.2.3. *There exists $c_3 \geq c_2$, constant depending on $q = \tau + d$, and $B_2 \geq B_1$ such that:*

$$\max \Big\{ \left| E^{(1)} - E^{(0)} \right| , \epsilon \left| P^{(1)}(y', \tilde{\varphi}(x')) \right|_{r-\rho, \sigma-2\delta} ,$$

$$(\rho/2)^{|\alpha|_1} \left| \partial_y^\alpha \left(Q^{(1)}(y', \tilde{\varphi}(x')) - Q^{(0)}(y', \tilde{\varphi}(x') \right) \right|_{r-\frac{3}{2}\rho, \sigma-2\delta} \Big\} \leq$$

$$\leq \epsilon c_3 \rho^{-3} \delta^{-4q} r^3 B_2 \mu$$

for any $|\alpha|_1 \leq 2$.

Proof Identity (A.2.8) and lemma A.2.2 yield:

$$\left|E^{(1)} - E^{(0)}\right| = \epsilon|\omega \cdot b + \langle P(0,\cdot)\rangle| \leq \epsilon(|\omega|\,|b| + |P|_{r.\sigma}) \leq$$

$$\leq \epsilon(Zc'S\delta^{-q}\mu A_1 + \mu) \leq \epsilon c'\delta^{-q}\mu A_1\,(ZS+1) \leq \epsilon c'\delta^{-q}\mu A_1 A_4$$

with

$$A_4 := \max\{ZS,\, 1\}$$

Moreover, by identity (A.2.9) we have $Q^{(1)} - Q^{(0)} = \epsilon\tilde{Q}^{(3)} = \epsilon(Q_1 + Q_2 + Q_3)$; thus, we may estimate separately the three terms using definition in (A.2.11) and the inequality proved in lemma A.2.2; it result

$$|Q_1(y', \tilde{\varphi}(x'))|_{r-\frac{3}{2}\rho,\sigma-2\delta} \leq |Q_1(y', x))|_{r-\frac{3}{2}\rho,\sigma-\delta} \leq$$

$$\leq |Q_y(y', x)|_{r-\rho,\sigma-\delta}|a_x(x)|_{\sigma-\delta}|y'| \leq cc_1\rho^{-1}\delta^{-2q}\gamma^{-1}\mu M A_1 A_2 r$$

and

$$|Q_2(y', \tilde{\varphi}(x'))|_{r-\frac{3}{2}\rho,\sigma-2\delta} \leq |Q_2(y', x)|_{r-\frac{3}{2}\rho,\sigma-\delta} \leq$$

$$\leq c\frac{8}{27}M\rho^{-3}r^3(|b| + |s_x|_{\sigma-\delta}) \leq$$

$$\leq cM\rho^{-3}r^3\left(c_1 S\mu\delta^{-q}A_1 + c_1\delta^{-q}\mu\gamma^{-1}\right) \leq$$

$$\leq cc_1\rho^{-3}\delta^{-2q}r^3 M\left(S\mu A_1 + \mu\gamma^{-1}\right) \leq cc_1\rho^{-3}\delta^{-2q}MA_1A_3r^4\,;$$

analogously, for what concerns Q_3 we have:

$$|Q_3(y', \tilde{\varphi}(x'))|_{r-\frac{3}{2}\rho,\sigma-2\delta} \leq |Q_3(y', x)|_{r-\frac{3}{2}\rho,\sigma-\delta} \leq$$

$$\leq |P_y(y', x)|_{r-\rho,\sigma}|y'|^2 \leq c\mu\rho^{-2}r^2\,.$$

Now recall that $A_3 = \mu r^{-1}\max\{S, \gamma^{-1}\}$ and then

$$MA_3 = \mu r^{-1}\max\left\{MS, M\gamma^{-1}\right\} \leq \mu r^{-1}\max\{A_2, A_1\} \leq \mu r^{-1}A_1A_2\,;$$

besides observe that obviously $\rho r^{-1} < 1$ and therefore we have:

$$(\rho/2)^{|\alpha|_1}\left|\partial_y^\alpha\left(Q^{(1)}(y',\tilde{\varphi}(x')) - Q^{(0)}(y',\tilde{\varphi}(x'))\right)\right|_{r-2\rho,\sigma-2\delta} \leq$$

$$\leq \epsilon\left(|Q_1|_{r-\frac{3}{2}\rho,\sigma-2\delta} + |Q_2|_{r-\frac{3}{2}\rho,\sigma-2\delta} + |Q_3|_{r-\frac{3}{2}\rho,\sigma-2\delta}\right) \leq$$

$$\leq \epsilon\left(cc_1\rho^{-1}\delta^{-2q}\gamma^{-1}\mu M A_1 A_2 r + cc_1\rho^{-3}\delta^{-2q} M A_1 A_3 r^4 + c\mu\rho^{-2}r^2\right) \leq$$

$$\leq \epsilon cc_1\rho^{-3}\delta^{-2q}r^3\left(\gamma^{-1}\mu M A_1 A_2 + M A_1 A_3 r + \mu\right) \leq$$

$$\leq \epsilon cc_1\rho^{-3}\delta^{-2q}r^3\left(\mu {A_1}^2 A_2 + \mu {A_1}^2 A_2 + \mu\right) \leq$$

$$\leq \epsilon cc_1\rho^{-3}\delta^{-2q}r^3 {A_1}^2 {A_2}^2\mu$$

It remains now to be proved the estimate for P; by identity (A.2.10) we have

$$|P^{(1)}(y',\tilde{\varphi}(x'))|_{r-\rho,\sigma-2\delta} \leq |P^{(1)}(y',x)|_{r-\rho,\sigma-\delta} \leq$$

$$\leq \left|P_y^{(0)}(y',x')\right|_{r-\rho,\sigma-\delta}|g_x(x)|_{\sigma-\delta} + \left|H_{yy}^{(0)}(y',x)\right|_{r-\frac{\delta}{2},\sigma-\delta}|g_x(x)|^2_{\sigma-\delta} \leq$$

$$\leq c\rho^{-1}\mu c_1\delta^{-2q}B_1 r + \left(\left|Q_{yy}^{(0)}(y',x)\right|_{r-\frac{\rho}{2},\sigma-\delta}\right.$$

$$\left. + \epsilon\left|P_{yy}^{(0)}(y',x)\right|_{r-\frac{\rho}{2},\sigma-\delta}\right)\ |g_x(x)|^2_{\sigma-\delta} \leq$$

$$\leq cc_1\rho^{-1}\delta^{-2q}\mu B_1 r + \left(4c\rho^{-2}Mr + \epsilon 4c\rho^{-2}\mu\right)\left(c_1\delta^{-2q}B_1 r\right)^2 \leq$$

$$\leq cc_1\rho^{-1}\delta^{-2q}\mu B_1 r + 4cc_1^2\rho^{-2}\delta^{-4q}{B_1}^2 r^3 M\left(1 + \epsilon\frac{\mu}{Mr}\right) \leq$$

$$\leq 4cc_1^2\rho^{-2}\delta^{-4q}r^2 B_1\left[\mu + MrB_1\left(1 + \epsilon\frac{\mu}{Mr}\right)\right] \leq$$

$$\leq 4cc_1^2\rho^{-2}\delta^{-4q}r^2 B_1\mu\left[1 + {A_1}^2{A_2}^2\left(1 + \epsilon\frac{\mu}{Mr}\right)\right] \leq$$

$$\leq 12cc_1^2\rho^{-2}\delta^{-4q}r^2{A_1}^2{A_2}^2 B_1\mu$$

if we impose on ϵ the condition

$$\epsilon\frac{\mu}{Mr} \leq 1\,.$$

The lemma is so proved taking

$$c_3 = 12cc_1^2 = 3c^7 2^{4q+14} \tag{A.2.15}$$

and

$$B_2 = {A_1}^3 {A_2}^3 A_3 A_4 \quad \square \tag{A.2.16}$$

A.3 Iteration and conclusion

A.3.1 Inductive step and convergence of the scheme

In lemma A.2.1 we have proved that Kolmogorov's non-degeneracy condition holds for $Q^{(1)} = Q^{(0)} \circ \Phi$ and hence we can iterate proposition A.2.1 obtaining by consecutive symplectic transformations the following scheme:

$$\begin{aligned} H &= H^{(0)} = N^{(0)} + \epsilon P^{(0)} \overset{\Phi^{(0)}}{\longmapsto} H^{(1)} = N^{(1)} + \epsilon^2 P^{(1)} \overset{\Phi^{(1)}}{\longmapsto} H^{(2)} = \\ &= N^{(2)} + \epsilon^4 P^{(2)} \cdots \overset{\Phi^{(j-1)}}{\longmapsto} H^{(j)} = N^{(j)} + \epsilon^{2^j} P^{(j)} \ldots \end{aligned} \tag{A.3.1}$$

(notice that here $\Phi^{(0)} = \Phi$ in proposition A.2.1); to prove theorem A.1.2 we must therefore provide in some way the convergence of the scheme.

With proposition A.2.1 we have reduced the analycity domain from $B_r^d \times \mathbb{T}_\sigma^d$ to $B_{r-2\rho}^d \times \mathbb{T}_{\sigma-2\delta}^d$, where this loss is due to the inversion of the operator $\mathcal{D}_\omega$ and to the necessity of estimating the derivatives of some analytic functions (see lemmata A.1.1 and A.1.2) . Let r_j and δ_j be the losses of analycity at each step and $B_{r_j}^d \times \mathbb{T}_{\sigma_j}^d$ the analycity domain after j iterations; to iterate infinitely times the proceeding shown, obtaining a non-empty analycity domain, we must then require that the sequences $\sigma_0 = \sigma$, $\sigma_1 = \sigma_0 - 2\delta_0$, $\sigma_2 = \sigma_1 - 2\delta_1 \ldots \sigma_{j+1} = \sigma_j - 2\delta_j = \sigma_0 - 2\sum_{k=1}^j \delta_k$ and $r_0 = r$, $r_1 = r_0 - 2\rho_0$, $r_2 = r_1 - 2\rho_1 \ldots r_{j+1} = r_j - 2\rho_j = r_0 - 2\sum_{k=1}^j \rho_k$ admit a strictly positive limit. For any $\sigma_\infty < \sigma_0$ and $r_\infty < r_0$ we put

$$\delta_j = \frac{1}{2^j}\frac{\sigma_0 - \sigma_\infty}{2} \qquad \rho_j = \frac{1}{2^j}\frac{r_0 - r_\infty}{2} \tag{A.3.2}$$

in order to have a final analycity domain $B_{r_\infty}^d \times \mathbb{T}_{\sigma_\infty}^d$.

Recall that in lemmata A.2.2 and A.2.3 we defined

$$\begin{aligned} A_1 &= \max\left\{M\gamma^{-1},\, 1\right\} \\ A_2 &= \max\left\{MS,\, 1\right\} \\ A_3 &= \mu \max\left\{S,\, \gamma^{-1}\right\} := \mu \hat{A}_3 \\ A_4 &= \max\left\{ZS,\, 1\right\} \end{aligned}$$

and took $B_1 = A_1 A_2 A_3$ and $B_2 = {A_1}^3 {A_2}^3 A_3 A_4$. We now define iteratively the following quantities

$$M_j := \frac{1}{r_j} |Q^{(j)}|_{r_j,\sigma_j} \ , \quad S_j := \frac{1}{r_j} \left| \langle Q^{(j)}_{yy}(0,\cdot)^{-1} \rangle \right| \ , \quad \mu_j := |P^{(j)}|_{r_j,\sigma_j}$$

and the following real numbers

$$\lambda_j := \max\left\{ \tfrac{1}{\sigma_\infty} \, , \, \tfrac{1}{\sigma_j - \sigma_\infty} \right\} \qquad \nu_j := \max\left\{ \tfrac{r}{r_\infty} \, , \, \tfrac{r}{r_j - r_\infty} \right\}$$

$$A_1^{(j)} := \max\{M_j \gamma^{-1}, \, 1\} \qquad A_2^{(j)} := \max\{M_j S_j, \, 1\}$$

$$A_3^{(j)} := \mu_j \hat{A}_3^{(j)} = \mu_j r_j^{-1} \max\{S_j, \, \gamma^{-1}\} \quad A_4^{(j)} := \max\{Z S_j, \, 1\}$$

$$B_1^{(j)} := A_1^{(j)} A_2^{(j)} A_3^{(j)} \qquad B_2^{(j)} := {A_1^{(j)}}^3 {A_2^{(j)}}^3 A_3^{(j)} A_4^{(j)}$$

with the notation $M_0 = M$, $S_0 = S$, $\lambda_0 = \lambda$, $\nu_0 = \nu$, $\mu_0 = \mu$. We are now ready to state

Lemma A.3.1. *There exist positive constants $c_4 \geq c_3$ and q_4, depending on $q = \tau + d$, such that if*

$$\epsilon \, C D \mu < 1 \tag{A.3.3}$$

with $C = c_4 \nu^{14} \lambda^{4q} r^{-1} A_1^4 A_2^4 A_4 \max\{M^{-1}, S\}$, $D = 2^{q_4}$, then it is possible to define iteratively (by the scheme described) Hamiltonians $H^{(j)} = N^{(j)} + \epsilon^{2^j} P^{(j)}$ analytic on $B^d_{r_j} \times \mathbb{T}^d_{\sigma_j}$ and $\Phi^{(j)}$ symplectic transformations such that $H^{(j+1)} = H^{(j)} \circ \Phi^{(j)}$.

Besides, referring to the previously defined quantities, for every $j \in \mathbb{N}$ we have

$$Mr \leq M_j r_j \leq 2Mr \tag{A.3.4}$$

$$S_j r_j \leq 2Sr \tag{A.3.5}$$

$$\epsilon^{2^j} \mu_j \leq \frac{(\epsilon C D \mu)^{2^j}}{C D^{j+1}} \tag{A.3.6}$$

and by mere consequence

$$\begin{aligned} A_1^{(j)} &\leq 2A_1\nu \\ A_2^{(j)} &\leq 4A_2\nu^2 \\ \hat{A}_3^{(j)} &\leq 2\hat{A}_3\nu^2 \\ \epsilon^{2^j} A_3^{(j)} &\leq A_3\nu^2 \\ A_4^{(j)} &\leq 2A_4\nu \\ \epsilon^{2^j} \frac{\mu_j r_j}{M_j} &\leq 1 \end{aligned}$$

and

$$\begin{aligned}\epsilon^{2^j} B_1^{(j)} &\leq B_1 \\ \epsilon^{2^j} B_2^{(j)} &\leq B_2 \,.\end{aligned}$$

Furthermore, it results that the symplectic transformation $\Phi^{(j)} : B^d_{r_{j+1}} \times \mathbb{T}^d_{\sigma_{j+1}} \to B^d_{r_j} \times \mathbb{T}^d_{\sigma_j}$ *generated by* $F_j(y', x) = y' \cdot x + \epsilon^{2^j} g_j(y', x)$ *(we denote* $F_0 = F$ *), where* $g_j(y', x) = b_j \cdot x + s_j(x) + a_j(x) \cdot y'$ *, is a symplectic diffeomorphism since*

$$\epsilon^{2^j} c_2 B_1^{(j)} \rho_j{}^{-1} \delta_j{}^{-2q} r_j < 1 \tag{A.3.7}$$

for all $j \in \mathbb{N}$*.*

Proof We want now to prove by induction inequalities (A.3.4) to (A.3.7) . For $j = 0$ condition (A.3.6) is trivial and (A.3.4) and (A.3.5) are obviously satisfied. For what concerns (A.3.7) observe that

$$\begin{aligned}\delta_0{}^{-m} &= \left(\frac{\sigma_0 - \sigma_\infty}{2}\right)^{-m} \leq 2^m \lambda^m \\ \rho_0{}^{-m} &= \left(\frac{r_0 - r_\infty}{2}\right)^{-m} \leq 2^m \left(\frac{\nu}{r}\right)^m\end{aligned}$$

and therefore we have

$$\epsilon c_2 B_1 \rho_0^{-1} \delta_0^{-2q} r \leq \epsilon c_2 A_1 A_2 A_3 2 \nu r^{-1} 2^{2q} \lambda^{2q} r =$$

$$= \epsilon \mu c_2 2^{2q+1} M^{-1} A_1^2 A_2^2 \nu \lambda^{2q} r^{-1} \leq \epsilon C D \mu < 1$$

by hypotheses, taking $c_4 \geq c_2$, $C \geq c_4 M^{-1} A_1^2 A_2^2 \nu \lambda^{2q} r^{-1}$ and $q_4 \geq 2q + 1$ so that (A.3.7) holds for $j = 0$. Notice that during the proof we will come across several lower bounds on c_4, q_4 and C and in the end we will take the worst in order to have all conditions required satisfied simultaneously.

Assume now by induction that conditions from (A.3.4) to (A.3.7) hold for $i = 0 \ldots j-1$. Recall that by consequence of lemma A.2.3 we have $\forall\ |p| \leq 1$

$$\left|P^{(j+1)}\right|_{r_j,\sigma_j} \leq c_3 \rho_j^{-2} \delta_j^{-4q} r_j^2 A_1^{(j)^3} A_2^{(j)^3} \hat{A}_3{}^{(j)} A_4^{(j)} \mu_j^2 := \mu_{j+1}$$

$$\left|E^{(j+1)} - E^{(j)}\right| \leq \epsilon^{2^j} \mu_{j+1}$$

$$(r - r_j)^{|p|_1} \left|\partial_y^p \left(Q^{(j+1)} - Q^{(j)}\right)\right|_{r_j,\sigma_j} \leq \epsilon^{2^j} \mu_{j+1}$$

where we have denoted $\mu_0 = \mu$ and hence

$$\mu_1 = c_3 \rho^{-2} \delta^{-4q} r^2 A_1^3 A_2^3 \hat{A}_3 A_4 \mu^2 \ .$$

We now verify (A.3.6): $\forall\ 1 \leq i \leq j$ we have

$$\mu_i = c_3 \rho_{i-1}^{-2} \delta_{i-1}^{-4q} r_{i-1}^2 A_1^{(i-1)^3} A_2^{(i-1)^3} \hat{A}_3^{\ (i-1)} A_4^{(i-1)} \mu_{i-1}^2 \leq$$

$$\leq c_3 \left(\frac{r_o - r_\infty}{2^i} \right)^{-2} \left(\frac{\sigma_o - \sigma_\infty}{2^i} \right)^{-4q} r^2 2^{11} \nu^{12} A_1^3 A_2^3 \hat{A}_3 A_4 \mu_{i-1}^2 \leq$$

$$\leq c_3 2^{4q+13} 2^{(4q+2)(i-1)} \nu^{14} \lambda^{4q} A_1^4 A_2^4 A_4 M^{-1} r^{-1} \mu_{i-1}^2 \leq C_0 D_0^{i-1} \mu_{i-1}^2$$

taking $C_0 \geq c_4 \nu^{14} \lambda^{4q} A_1^4 A_2^4 A_4 M^{-1} r^{-1}$ with $c_4 \geq c_3 2^{4q+13}$ and $D_0 \geq 2^{4q+2}$ (that is $q_4 \geq 4q + 2$). Now let $\hat{\mu}_i = C_0 D_0^{i+1} \mu_i$ we have

$$\hat{\mu}_i \leq \left(C_0 D_0^{i+1} \right) \left(C_0 D_0^{i-1} \mu_{i-1}^2 \right) = C_0^2 D_0^{2i} \mu_{i-1}^2 = \hat{\mu}_{i-1}^2$$

therefore by iteration we obtain $\forall\ i \leq j$

$$\hat{\mu}_i \leq \hat{\mu}_0^{2^i}$$

that is $\forall\ C \geq C_0, D \geq D_0$ it results (taking $i = j$)

$$C D^{j+1} \mu_j \leq (CD\mu)^{2^j} \qquad \Rightarrow \qquad \epsilon^{2^j} \mu_j \leq \frac{(\epsilon C D \mu)^{2^j}}{C D^{j+1}}$$

and hence condition (A.3.6) holds $\forall\ j \in \mathbb{N}$.

Using (A.3.6) and hypothesis (A.3.3) we can obtain

$$\begin{aligned}
|Q^{(j)}|_{r_j, \sigma_j} &= \left| Q^{(0)} + \sum_{i=1}^{j} Q^{(i)} - Q^{(i-1)} \right|_{r_j, \sigma_j} \leq \\
&\leq |Q^{(0)}|_{r_1, \sigma_1} + \sum_{i=1}^{j} |Q^{(i)} - Q^{(i-1)}|_{r_i, \sigma_i} \leq \\
&\leq rM + \sum_{i=1}^{j} \epsilon^{2^{i-1}} \mu_{i-1} \leq rM + \sum_{i=1}^{j} \frac{(\epsilon C D \mu)^{2^{i-1}}}{C D^i} \leq \\
&\leq rM + \sum_{i=1}^{j} \frac{1}{C D^i} \leq rM + \frac{1}{C} \sum_{i=1}^{+\infty} D^{-i} = rM + \frac{1}{C(D-1)} \leq \\
&\leq rM + rM \leq 2rM \ .
\end{aligned}$$

since $C^{-1} \leq C_0^{-1} \leq Mr$, so that (A.3.4) is verified.

We now verify (A.3.5). Let $B_i := \langle Q_{yy}^{(j)}(0,\cdot)\rangle$ for $i = 0 \dots j-1$, we want to prove $|B_j^{-1}| \leq 2S$. Recall that if $A \in \mathrm{mat}(d \times d)$ then $(I+A)^{-1} = \sum_{k=0}^{\infty}(-1)^k A^k$ and $|(I+A)^{-1}| \leq \frac{1}{1-|A|}$. So

$$B_j = B_0 + \sum_{i=1}^{j} B_i - B_{i-1} = B_0 + \hat{B} = B_0\left(I + B_0^{-1}\hat{B}\right)$$

where obviously we took $\hat{B} = \sum_{i=1}^{j}(B_i - B_{i-1})$. By hypothesis B_0 is invertible, so to invert B_j we have to invert $I + B_0^{-1}\hat{B}$, that is we want to prove $|B_0^{-1}\hat{B}| < 1$:

$$\begin{aligned}
|B_0^{-1}\hat{B}| &\leq |B_0^{-1}|\sum_{i=1}^{j}|B_i - B_{i-1}| \leq \\
&\leq Sr\sum_{i=1}^{j}\left|\langle\partial_y^2 Q^{(i)}(0,\,\cdot) - Q^{(i-1)}(0,\,\cdot)\rangle\right|_{r_i,\sigma_i} \leq \\
&\leq Sr\sum_{i=1}^{j}\frac{c}{(r-r_i)^2}\left|Q^{(i)} - Q^{(i-1)}\right|_{r_i,\sigma_i} \leq \\
&\leq cS\frac{r}{(r-r_i)^2}\sum_{i=1}^{j}\epsilon^{2^i}\mu_{i-1} \leq Sc\frac{\nu}{r}\sum_{i=1}^{j}\frac{(\epsilon C D_0\mu)^{2^i}}{CD_0^{i+1}} \leq \\
&\leq \frac{cS\nu}{rC}\sum_{i=1}^{\infty}\frac{1}{D_0^{i+1}} = \frac{cS\nu}{rCD_0(D_0-1)} \leq \frac{cS\nu}{rC} \leq \frac{1}{2}\ .
\end{aligned}$$

if we assume $C \geq 2cS\nu r^{-1}$; the new condition on C_0 is

$$C_0 \geq c_4\nu^{14}\lambda^{4q}A_1^4A_2^4A_4r^{-1}\max\{M^{-1}, S\}\ .$$

We just proved that B_j is invertible and

$$|B_j^{-1}| = r_jS_j = |B_0^{-1}|\left|\left(I + B_0^{-1}\hat{B}\right)^{-1}\right| \leq |B_0^{-1}|\frac{1}{1-|B_0^{-1}\hat{B}|} \leq 2Sr.$$

so that $S_jr_j \leq 2Sr$ for every $j \in \mathbb{N}$.

To end the proof of this lemma we have to verify (A.3.7) for $i = j$. Using

(A.3.4) to (A.3.6) and hypothesis (A.3.3) we have

$$\epsilon^{2^j} c_2 B_1^{(j)} \rho_j{}^{-1} \delta_j{}^{-2q} r_j = c_2 \epsilon^{2^j} \mu_j A_1^{(j)} A_2^{(j)} \hat{A}_3{}^{(j)} \frac{2^{j+1}}{r_o - r_\infty} \left(\frac{2^{j+1}}{\sigma_o - \sigma_\infty} \right)^{2q} r_j \leq$$

$$\leq c_2 \epsilon^{2^j} \mu_j 2^4 A_1 A_2 \hat{A}_3 2^{(2q+1)(j+1)} \nu^6 \lambda^{2q} \leq$$

$$\leq c_2 \frac{(\epsilon C D \mu)^{2^j}}{C D^{j+1}} 2^4 A_1 A_2 \hat{A}_3 2^{(2q+1)(j+1)} \nu^6 \lambda^{2q} \leq$$

$$\leq \frac{c_2}{C D^{j+1}} 2^4 M^{-1} r^{-1} A_1^2 A_2^2 2^{(2q+1)(j+1)} \nu^6 \lambda^{2q} < 1$$

if we take $C \geq c_4 M^{-1} r^{-1} A_1^2 A_2^2 \nu^6 \lambda^{2q}$ with $c_4 \geq c_2 2^4$ and $D \geq 2^{2q+1}$.

Hence the lemma is proved by taking

$$\begin{aligned} c_4 &= c_3 2^{4q+13} \\ q_4 &= 4q+2 \qquad \square \end{aligned} \tag{A.3.8}$$

We are now ready to prove the convergence of the scheme described in (A.3.1) with the following

Proposition A.3.1. *Let $\Phi = \Phi^{(0)}, \Phi^{(1)} \ldots \Phi^{(j)}$ the sequence of symplectic diffeomorphisms obtained iterating lemma A.2.1; if we define*

$$\Psi^{(j)} = \Phi^{(0)} \circ \Phi^{(1)} \circ \cdots \circ \Phi^{(j)} : B^d_{r_{j+1}} \times \mathbb{T}^d_{\sigma_{j+1}} \to B^d_r \times \mathbb{T}^d_\sigma$$

then the sequence $\Psi^{(j)}$ converges (uniformly) to a symplectic diffeomorphism $\Psi := \lim_{j\to\infty} \Psi^{(j)}$ such that

1. $\Psi = id + O(\epsilon)$
2. $H^{(0)} \circ \Psi = N^{(\infty)} = E^{(\infty)} + \omega \cdot y' + Q^{(\infty)}(y', x')$

with $N^{(\infty)}$ (that is N' in theorem A.1.2) analytic on $B^d_{r_\infty} \times \mathbb{T}^d_{\sigma_\infty}$.

Proof We prove the uniform convergence of $\Psi^{(j)}$ which also guarantees the analycity of $N^{(\infty)}$. Let's write $\Psi^{(j)}$ through a telescopic series:

$$\Psi^{(j)} = \Psi^{(0)} + \sum_{i=1}^{j} \Psi^{(i)} - \Psi^{(i-1)} = \Phi + \sum_{i=1}^{j} \Psi^{(i)} - \Psi^{(i-1)}\,.$$

By lemma A.2.2 we have obtained that

$$|\Phi - \mathrm{id}|_{r_1,\sigma_1} \leq \epsilon c_2 B_1 \delta_0^{-2q} r$$

since

$$(\Phi - \mathrm{id})(y', x') = \epsilon \left(b + s_x(x) + a_x^T(x) \cdot y',\, \tilde{a}(x')\right)_{x = \tilde{\varphi}(x')}$$

and each term was estimated with $c_1 B_1 \delta^{-2q} r$ and $c_2 \geq 4c_1$. By induction we can therefore assume

$$|\Phi^{(i)} - \mathrm{id}|_{r_i, \sigma_i} \leq \epsilon^{2^i} c_2 B_1^{(j)} \delta_j^{-2q} r_j$$

which implies, together with lemma A.3.1 ,

$$\left|\Psi^{(i)} - \Psi^{(i-1)}\right|_{r_{i+1}, \sigma_{i+1}} = \left|\Phi^{(i)} \circ \Psi^{(i-1)} - \Psi^{(i-1)}\right|_{r_{i+1}, \sigma_{i+1}} \leq$$

$$\leq \epsilon^{2^i} c_2 B_1^{(j)} \delta_j^{-2q} r_j = c_2 \epsilon^{2^i} \mu_j A_1^{(j)} A_2^{(j)} \hat{A}_3{}^{(j)} \delta_j^{-2q} r_j \leq$$

$$\leq c_2 \epsilon^{2^i} \mu_j 2^4 \nu^5 A_1 A_2 \hat{A}_3 \lambda^{2q} 2^{2q(j+1)} r \leq$$

$$\leq c_2 \frac{(\epsilon C_0 D_0 \mu)^{2^i}}{C_0 D_0^{i+1}} 2^4 \nu^5 A_1^2 A_2^2 M^{-1} r^{-1} \lambda^{2q} 2^{2q(i+1)} r \leq (\epsilon C_0 D_0 \mu)^{2^i} r$$

since in lemma A.3.1 we took $C_0 \geq c_4 \nu^{16} \lambda^{4q} A_1^4 A_2^4 A_4 M^{-1} r^{-1}$ and $D_0 \geq 2^{4q+2}$ (notice that $\nu > 1$) . Therefore we can estimate $|\Psi - \mathrm{id}|$ as follows :

$$|\Psi - \mathrm{id}|_{r_\infty, \sigma_\infty} \leq |\Phi - \mathrm{id}|_{r_\infty, \sigma_\infty} + \sum_{i=1}^{\infty} \left|\Psi^{(i)} - \Psi^{(i-1)}\right|_{r_\infty, \sigma_\infty} \leq$$

$$\leq |\Phi - \mathrm{id}|_{r_1, \sigma_1} + \sum_{i=1}^{\infty} \left|\Psi^{(i)} - \Psi^{(i-1)}\right|_{r_{i+1}, \sigma_{i+1}} \leq$$

$$\leq \epsilon c_2 B_1 \delta_0^{-2q} r + \sum_{i=1}^{\infty} (\epsilon C_0 D_0 \mu)^{2^i} r \leq$$

$$\leq \epsilon \mu c_2 A_1 A_2 \hat{A}_3 \delta_0^{-2q} r + \sum_{i=1}^{\infty} (\epsilon C_0 D_0 \mu)^{2^i} r \leq$$

$$\leq \epsilon \mu c_2 M^{-1} r^{-1} A_1^2 A_2^2 2^{2q(i+1)} \lambda^{2q} r + \sum_{i=2}^{\infty} (\epsilon C_0 D_0 \mu)^i r \leq$$

$$\leq \epsilon C_0 D_0 \mu r + \sum_{i=2}^{\infty} (\epsilon C_0 D_0 \mu)^i r$$

since always by lemma A.3.1 it results $C_0 \geq c_2 M^{-1} r^{-1} \lambda^{2q}$ and $D_0 \geq 2^{2q}$. Then, taking $D \geq 2D_0$, that is to say the new hypothesis is $\epsilon C_0 D \mu < 1$ and hence

$\epsilon C_0 D_0 \mu < \frac{1}{2}$, we obtain

$$|\Psi - \text{ id}|_{r_\infty,\sigma_\infty} \leq \epsilon C_0 D_0 \mu r + \sum_{i=2}^{\infty} (\epsilon C_0 D_0 \mu)^i r \leq$$

$$\leq \epsilon C_0 D_0 \mu r + \frac{(\epsilon C_0 D_0 \mu)^2}{1 - \epsilon C_0 D_0 \mu} r \leq \epsilon C_0 D_0 \mu r + 2(\epsilon C_0 D_0 \mu)^2 r \leq \epsilon C_0 D \mu r.$$

Thus $\Psi^{(j)}$ converges uniformly to Ψ and $N^{(\infty)} = H^{(0)} \circ \Psi$ is analytic. To conclude we trivially observe that

$$\epsilon^{2^j} \left|P^{(j)}\right|_{\sigma_j,\ \sigma_j} \leq \epsilon^{2^j} \mu_j \leq (\epsilon C D \mu)^{2^j} \xrightarrow{j \to \infty} 0$$

so that $N^{(\infty)}$ is effectively in Kolmogorov's normal form $\square$

A.3.2 Final estimates

To completely prove theorem A.1.2 we estimate $|E^{(\infty)} - E^{(0)}|$ and $\left|Q^{(\infty)} - Q^{(0)}\right|$. Recall first that in order to have all the inductive conditions satisfied we must take $\epsilon C D \mu < 1$ for any

$$C \geq C_0 \quad = \quad c_4 \nu^{14} \lambda^{4q} r^{-1} \max\{M^{-1}, S\} A_1^4 A_2^4 A_4 \tag{A.3.9}$$

$$D \geq D_0 \quad = \quad 2^{4q+2} \tag{A.3.10}$$

with $c_4 = c_3 2^{4q+13}$. Now using the estimates done in the proof of lemma A.2.3 we have:

$$\left|E^{(1)} - E^{(0)}\right| \leq c_1 \epsilon \delta_0^{-q} \mu A_1 A_2$$

therefore by inductive hypotheses and lemma A.3.1 we obtain

$$\left|E^{(j+1)} - E^{(j)}\right| \leq c_1 \epsilon^{2^j} \delta_j^{-2q} \mu_j A_1^{(j)} A_2^{(j)} \leq$$

$$\leq c_1 \epsilon^{2^j} \mu_j \lambda^{2q} 2^{2q(j+1)} 2^3 \nu^3 A_1 A_2 \leq$$

$$\leq \frac{(\epsilon C D_0 \mu)^{2^j}}{C D_0^{j+1}} 2^{2q(j+1)} c_1 2^3 \nu^3 \lambda^{2q} A_1 A_2 \leq (\epsilon C D_0 \mu)^{2^j} M r$$

for any $C \geq C_0$. Now writing $E^{(\infty)}$ as a telescopic series and taking $D \geq 2D_0$, in order to have $\epsilon C D_0 \mu < \frac{1}{2}$, it results

$$\left|E^{(\infty)} - E^{(0)}\right| \leq \left|E^{(1)} - E^{(0)}\right| + \sum_{j=1}^{\infty} \left|E^{(j+1)} - E^{(j)}\right| \leq$$

$$\leq c_1 \epsilon \delta_0^{-q} \mu A_1 A_2 + Mr \sum_{j=1}^{\infty} (\epsilon C D_0 \mu)^{2^j} \leq$$

$$\leq c_1 \epsilon \lambda^q 2^q \mu A_1 A_2 + Mr \sum_{j=2}^{\infty} (\epsilon C D_0 \mu)^j \leq$$

$$\leq \epsilon C D_0 \mu M r + Mr \frac{(\epsilon C D_0 \mu)^2}{1 - \epsilon C D_0 \mu} \leq \epsilon C D \mu M r \ .$$

In a completely analogous way we can estimate $\left|\partial_y^p (Q^{(\infty)} - Q^{(0)})\right|_{r_\infty, \sigma_\infty}$; in lemma A.2.3 we obtained

$$\rho_0^{|p|_1} \left|\partial_y^p (Q^{(1)} - Q^{(0)})\right|_{r_1, \sigma_1} \leq c_3 \epsilon \rho_0^{-3} \delta_0^{-2q} r^3 A_1^2 A_2^2 \mu$$

thus, by induction,

$$(r - r_{j+1})^{|p|_1} \left|\partial_y^p (Q^{(j+1)} - Q^{(j)})\right|_{r_{j+1}, \sigma_{j+1}} \leq$$

$$\leq c_3 \epsilon^{2^j} \rho_j^{-3} \delta_j^{-2q} r_j^3 {A_1^{(j)}}^2 {A_2^{(j)}}^2 \mu_j \leq c_3 \epsilon^{2^j} \mu_j 2^{(3+2q)(j+1)} \nu^3 \lambda^{2q} 2^6 \nu^6 A_1^2 A_2^2 \leq$$

$$\leq \frac{(\epsilon C D_0 \mu)^{2^j}}{C D_0^{j+1}} 2^{(3+2q)(j+1)} \nu^9 \lambda^{2q} c_3 2^6 A_1^2 A_2^2 \leq (\epsilon C D_0 \mu)^{2^j} M r \, ;$$

writing as usual $Q^{(\infty)}$ as a telescopic series we obtain for $|p|_1 \le 2$

$$
\begin{aligned}
& (r - r_\infty)^{|p|_1} \left| \partial_y^p (Q^{(\infty)} - Q^{(0)}) \right|_{r_\infty, \sigma_\infty} \le (r - r_1)^{|p|_1} \left| \partial_y^p (Q^{(1)} - Q^{(0)}) \right|_{r_1, \sigma_1} + \\
& + \sum_{j=1}^{\infty} (r - r_{j+1})^{|p|_1} \left| \partial_y^p (Q^{(j+1)} - Q^{(j)}) \right|_{r_{j+1}, \sigma_{j+1}} \le \\
& \le c_3 \epsilon \rho_0^{-3} \delta_0^{-2q} r^3 A_1^2 A_2^2 \mu + Mr \sum_{j=1}^{\infty} (\epsilon C D_0 \mu)^{2^j} \le \\
& \le \epsilon c_3 2^{3+2q} \nu^3 \lambda^{2q} A_1^2 A_2^2 \mu + Mr \sum_{j=2}^{\infty} (\epsilon C D_0 \mu)^j \le \\
& \le \epsilon C D_0 \mu M r + Mr \frac{(\epsilon C D_0 \mu)^2}{1 - \epsilon C D_0 \mu} \le \epsilon C D \mu M r
\end{aligned}
$$

having imposed the same previous condition $D \ge 2D_0$.

We now conclude remarking that by the estimates done we can take $\epsilon C D \mu < 1$ with (see (A.3.9), (A.3.10), (A.2.15), (A.3.8))

$$
\begin{aligned}
c_4 &= 3c^7 2^{8q+27} \\
C &= c_4 \nu^{14} \lambda^{4q} A_1^4 A_2^4 A_4 r^{-1} \max\{M^{-1}, S\} \\
D &= 2^{4q+3}
\end{aligned}
$$

(where $c = c(\tau, d)$ in lemma A.1.2) that is

$$
\epsilon < \epsilon_0 := \frac{r}{3\mu} c^{-7} 2^{-(12(\tau+d)+30)} \nu^{-3} \lambda^{-4(\tau+d)} (A_1 A_2)^{-4} A_4^{-1} \min\{M, S^{-1}\} . \tag{A.3.11}
$$

A.4 Dependence on additional parameters

In this last section we are going to analyze what happens when the perturbed Hamiltonian depends, in addiction to the action-angles variables, on some parameters belonging to a compact set of $\mathbb{R}^m$. The result is that Kolmogorov's theorem applies easily to such Hamiltonians if we assume some uniform estimates on the norms of P and Q. We can formulate our statement as follows:

Theorem A.4.1. *Let $H(y,x;\beta) = E(\beta) + \omega \cdot y + Q(y,x;\beta) + P(y,x;\beta)$ be a real-analytic Hamiltonian over $B^d \times \mathbb{T}^d$ with $\omega \in \mathcal{D}^d_{\gamma,\tau}$. Assume that H has a C^k or C^k_w for $0 \le k \le \infty$, Lipschitz or analytic dependence on the parameters β in a compact subset $\mathcal{B}$ of $\mathbb{R}^m$. Suppose that P and Q have analytic extension on the complex domain $B^d_r \times \mathbb{T}^d_\sigma$, for some $0 < \sigma \le 1$. Suppose to have $Q(0,x;\beta) = \partial_y Q(0,x;\beta) = 0$ and*

$$\det\langle Q_{yy}(0,\cdot;\beta)\rangle \neq 0$$

for all $\beta \in \mathcal{B}$.

Let $\sigma_\infty < \sigma$ and $r_\infty < r$, take

$$\begin{aligned}
\mu &= \sup_{\beta\in\mathcal{B}} |P(y,x;\beta)|_{r,\sigma} \\
M &= \sup_{\beta\in\mathcal{B}} \max\left\{\frac{1}{r}|Q(y,x;\beta)|_{r,\sigma}, |Q_y(y,x;\beta)|_{r,\sigma}, r|Q_{yy}(y,x;\beta)|_{r,\sigma}\right\} \\
\lambda &= \max\left\{\frac{1}{\sigma_\infty}, \frac{1}{\sigma-\sigma_\infty}\right\} \\
\nu &= \max\left\{\frac{r}{r_\infty}, \frac{r}{r-r_\infty}\right\} \\
S &\ge \frac{1}{r}\sup_{\beta\in\mathcal{B}} \left|\langle Q_{yy}(0,\cdot;\beta)\rangle^{-1}\right| \ ;
\end{aligned}$$

and define

$$\begin{aligned}
\Gamma_1 &= \max\left\{M\gamma^{-1}, 1\right\} \\
\Gamma_2 &= \max\left\{MS, 1\right\} \\
\Gamma_3 &= \max\left\{ZS, 1\right\} \\
\Gamma_4 &= \max\left\{M^{-1}, S\right\};
\end{aligned}$$

there exist a positive constant $c(\tau,d) \ge 1$ such that if

$$CD\mu < 1$$

where

$$C = c\nu^{14}\lambda^{4q}r^{-1}\Gamma_1^4\Gamma_2^4\Gamma_3\Gamma_4 \tag{A.4.1}$$

and

$$D = 2^{12(\tau+d)+30}, \tag{A.4.2}$$

then there exists a symplectic diffeomorphism

$$\Phi : (y',x';\beta) \in B^d_{r_\infty} \times \mathbb{T}^d_{\sigma_\infty} \times \mathcal{B} \to (y,x;\beta) \in B^d_r \times \mathbb{T}^d_\sigma \times \mathcal{B}$$

with the same dependence of H on the parameters β, which puts the Hamiltonian H into the form

$$N'(y', x'; \beta) = E'(\beta) + \omega \cdot y' + Q'(y', x'; \beta) = H \circ \Phi$$

and we also have that $|E' - E|$, $\|Q' - Q\|_{C^1} \leq CD\mu Mr$ *and* $\|\Phi - id\|_{C^1} \leq CD\mu r$.

Remark A.4.1. *Notice that this is theorem A.1.2 if we take* $\beta = \epsilon$, $\mathcal{B} = \{|x| < \epsilon_0\}$ *(with* ϵ_0 *as found in (A.3.11)) and we make the substitution* $P \to \epsilon P$.

Proof The proof of this theorem is totally equivalent to the proof given in the previous sections. The only difference is that we assumed uniform estimates on the additional parameters β so that we are allowed to find, as done before, a symplectic map, obviously depending on the parameters (i.e., a family of symplectic maps parametrized by $\beta \in \mathcal{B}$) that puts the Hamiltonian H into Kolmogorov's normal form.

The only aspect we need to discuss briefly is the kind of dependence Φ has on the parameters. Recall first that in the first step of the proof of Kolmogorov's theorem we took

$$\begin{aligned} s(x;\beta) &= \mathcal{D}_\omega{}^{-1}(P(0,x;\beta) - \langle P(0,x;\beta)\rangle) \\ b(\beta) &= -\langle Q_{yy}(0,\cdot;\beta)\rangle^{-1}\langle P_y(0,\cdot;\beta) + Q_{yy}(0,\cdot;\beta)\cdot s_x(\cdot;\beta)\rangle \\ a(x;\beta) &= -\mathcal{D}_\omega{}^{-1}\left(P_y(0,x;\beta) + Q_{yy}(0,x;\beta)\cdot(b + s_x)\right) \ ; \end{aligned}$$

it can be immediately seen that s , b and a have the same dependence of P and Q on β . Observe that equivalent formulations of lemma A.1.4 and proposition A.1.1 can be given in the case $a \in \mathcal{H}(\mathbb{T}^d_\sigma)$ has an arbitrarily dependence on the parameters β. Therefore $\varphi(x;\beta) = x + a(x;\beta)$ is a diffeomorphism in the angles (since the estimates of lemma A.2.2 naturally hold also in this case by the assumption of uniformity made on the norms) and so is its inverse $\tilde{\varphi}$, while they have the same dependence of H on β . Hence the first symplectic transformation $\Phi^{(0)}$ generated by

$$g(y', x; \beta) = b(\beta)\cdot x + s(x;\beta) + a(x;\beta)\cdot y'$$

also depends in the same way on the parameters; this proves that $H^{(1)} = H^{(0)} \circ \Phi^{(0)}$ has the dependence of $H^{(0)}$ on β and so has $H^{(\infty)}$ since proposition A.3.1 still holds if we add the dependence on some parameters $\square$

Appendix B

Measure of Kolmogorov's invariant tori

B.1 Introduction

Let $H'(y',x') = E' + \omega \cdot y' + Q(y',x')$ a real-analytic Hamiltonian in Kolmogorov's normal form over $B^d \times \mathbb{T}^d$, with $\omega \in \mathcal{D}^d_{\gamma,\tau}$ fixed diophantine vector, and $Q(0,x') = \partial_y Q(0,x') = 0$. Then, we have that the torus

$$\mathcal{T} = \{(0,x';\omega);\ x' \in \mathbb{T}^d\}$$

is invariant for the flow $\Phi^t_{H'}$ of H' and the flow on $\mathcal{T}$ is given by $\Phi^t_{H'}(0,x) = (0, x+\omega t)$. If we consider now an Hamiltonian $H(y,x) = E + \omega \cdot y + Q(y,x) + \epsilon P(y,x)$ conjugated to H by the symplectic transformation

$$\Phi_\epsilon : (y',x';\omega) \in B^d \times \mathbb{T}^d \to (y,x;\omega) \in B^d \times \mathbb{T}^d\,,$$

as in Kolmogorov's theorem, we naturally have that the torus

$$\mathcal{T}_\omega = \Phi_\epsilon(\mathcal{T}) = \{\Phi_\epsilon\,(0,x';\omega)\,;\ x' \in \mathbb{T}^d\} \tag{B.1.1}$$

is invariant for the flow Φ^t_H of H and the flow on $\mathcal{T}$ is given by $\Phi^t_H(0,x) = \Phi_\epsilon((0, x+\omega t))$.

We now consider an Hamiltonian in the form

$$H_\epsilon(I,\varphi) = h(I) + \epsilon f(I,\varphi)$$

real-analytic on a certain domain $D \times \mathbb{T}^d$, D being an open bounded set in $\mathbb{R}^d$; we make the assumption that the "*frequency map*"

$$h' := \omega_0\ :\ I\ \in D\ \longrightarrow\ \omega_0(I) = \frac{\partial h(I)}{\partial I} \tag{B.1.2}$$

is a diffeomorphism of class C^∞ on D (up to restricting the domain this can be made without loss of generality). Hamiltonians in this form are called "nearly-integrable" and are of great interest since they often appear in problems dealing with celestial mechanics. Consider now for a fixed $\delta > 0$ the open subset of D

$$B = B(\delta) = \{I \in D \mid \text{dist}\,(I, \partial D) > \delta\} \tag{B.1.3}$$

Let

$$\Omega = \omega_0(B) \tag{B.1.4}$$

the mapping of B through the frequency map and let

$$\Omega_\star = \Omega(\gamma, \tau) = \{\omega \in \Omega \mid \omega \in \mathcal{D}_{\gamma,\tau}\} \tag{B.1.5}$$

for fixed γ and $\tau > d - 1$, we define

$$B_\star = {\omega_0}^{-1}(\Omega_\star).$$

With an appropriate change of variables we will reduce the Hamiltonian H_ϵ into a perturbed Hamiltonian H as in theorem A.4.1 in order to obtain for each $\omega \in B_\star$, as previously observed, an invariant torus for H and hence, coming back to the original variables, an invariant torus $\mathcal{T}_\omega$ for H_ϵ. In this chapter our aim is to give a result concerning the measure of this KAM tori, i.e., maximal tori possessing quasi-periodic motions with diophantine frequencies, as consequence of the results obtained in the previous chapter. More precisely we are going to estimate

$$\text{meas}_{\,2d} \bigcup_{\omega \in B_\star} \mathcal{T}_\omega$$

(here " meas $_n$" stands for the Lebesgue measure in $\mathbb{R}^n$) showing that the measure of this union of invariant tori for H_ϵ is at least

$$\inf_{\omega \in \Omega} \left| \frac{\partial}{\partial \omega} {\omega_0}^{-1}(\omega) \right| \left(1 - O(\epsilon^{\frac{1}{4}})\right) (2\pi)^{2d} (\,\text{meas}_{\,d} \Omega) \ .$$

We remark that a better result, more precisely $1 - O(\epsilon^{\frac{1}{2}})$ instead of $1 - O(\epsilon^{\frac{1}{4}})$, can be obtained using V. I. Arnold proof of Kolmogorov's theorem that is slightly different from the one we gave in the previous chapter which is based on Kolmogorov's original idea. Arnold's formulation and proof of Kolmogorov's theorem can be found in [Arn63a]. We also remark that a famous generalization of the first results on the measure of invariant tori obtained by V. I. Arnold and A. I. Neistadt, was given by J. Pöeschel who established, in [Pös82], similar estimates in the case of finitely many times differentiable perturbations.

B.2 Kolmogorov's normal form for $H_\epsilon(I,\varphi)$

Let $H_\epsilon(I,\varphi) = h(I) + \epsilon f(I,\varphi)$ the analytic Hamiltonian we are considering, ω_0 the frequency map and $B_\star$ as defined in the introduction with the same assumption already made. We now operate the following elementary change of variables:

$$\begin{cases} I &= J + y \\ \varphi &= x \ . \end{cases}$$

So, by Taylor's development we have

$$\begin{aligned} H_\epsilon(I,\varphi) &= H_\epsilon(J+y,x) = h(J+y) + \epsilon f(J+y,x) = \\ &= h(J) + \omega_0(J)\cdot y + Q(y,x;J) + \epsilon f(J+y,x) \end{aligned}$$

having defined $h' := \omega_0$ and

$$Q(y,x;J) = \frac{1}{2}\int_0^1 h''(J+ty)\,dt\ \langle y,y\rangle.$$

Let $\omega = \omega_0(J)$, that is $J = {\omega_0}^{-1}(\omega)$, it results

$$H_\epsilon(I,\varphi) = \hat{H}_\epsilon(y,x;\omega) = E(\omega) + \omega\cdot y + \hat{Q}(y,x;\omega) + \epsilon \hat{f}(y,x;\omega) \qquad \text{(B.2.1)}$$

with

$$\begin{aligned} E(\omega) &= h\left({\omega_0}^{-1}(\omega)\right) \\ \hat{Q}(y,x;\omega) &= Q\left(y,x;{\omega_0}^{-1}(\omega)\right) = \frac{1}{2}\int_0^1 h''\left({\omega_0}^{-1}(\omega)+ty\right)\,dt\ \langle y,y\rangle \\ \hat{f}(y,x;\omega) &= f\left({\omega_0}^{-1}(\omega)+y,x\right)\ . \end{aligned}$$

We now want to apply Kolmogorov's theorem on the persistence of quasi-periodic motions to $\hat{H}_\epsilon(y,x;\omega)$. $\hat{H}_\epsilon$ depends on parameter $(\omega,\ e) \in \Omega_\star \times \{|\epsilon| < \epsilon_0\}$ (ϵ_0 as in (A.3.11)). Let us verify that the hypotheses made in theorem A.1.2 hold for $\hat{H}_\epsilon(y,x;\omega)$ uniformly upon the parameters. By definition of $\hat{Q}(y,x;\omega)$ we have that

$$\hat{Q}(0,x;\omega) = 0 = \partial_y \hat{Q}(0,x;\omega);$$

moreover, it results

$$\det \langle \hat{Q}_{yy}(0,x;\omega)\rangle_{\mathbb{T}^d} = \det \langle h''({\omega_0}^{-1}(\omega))\rangle_{T^d} = \det h''(J)$$

where the average over $\mathbb{T}^d$ vanishes since $\hat{Q}$ does not depend on the angles $x = \varphi$; observe now that $J = {\omega_0}^{-1}(\omega)$, where $\omega \in \Omega = \omega_0(B)$, hence $J \in B$ and therefore , being h' a diffeomorphism on B by hypotheses, we have

$$\det h''(J) \neq 0$$

so that we can effectively apply theorem A.4.1 to $\hat{H}_\epsilon(y, x; \omega)$. Then for every $\epsilon < \epsilon_0(\omega) = (CD\mu)^{-1}$ and for every $\omega \in \Omega_\star$ there exist a symplectic map

$$\Phi_\epsilon : (y', x'; \omega) \to (y_0(y', x'; \omega), x_0(y', x'; \omega)) \tag{B.2.2}$$

having domain on $B^d_{r_\infty} \times \mathbb{T}^d_{\sigma_\infty}$ and codomain in $B^d_r \times \mathbb{T}^d_\sigma$ (where $B^d \subset B$ is an open ball in R^d), such that

$$\hat{H}_\epsilon \circ \Phi_\epsilon(y', x'; \omega) = N'(y', x'; \omega, \epsilon) = E'(\omega; \epsilon) + \omega \cdot y' + Q'(y', x'; \omega, \epsilon)$$

Then it results that

$$\hat{\mathcal{T}}_\omega = \{(y_0(0, x'; \omega), x_0(0, x'; \omega)) \mid x' \in \mathbb{T}^d,\ \omega \in \Omega_\star\}$$

is an invariant analytic torus for $\hat{H}_\epsilon$ and therefore

$$\mathcal{T}_\omega = \{I = \omega_0^{-1}(\omega) + y_0(0, x'; \omega), \varphi = x_0(0, x'; \omega) \mid x' \in \mathbb{T}^d,\ \omega \in \Omega_\star\}$$

is an invariant torus for H_ϵ.

B.3 Extension of Φ_ϵ

By Kolmogorov's theorem we obtained the symplectic transformation Φ_ϵ, as in (B.2.2), which is real-analytic for $(y', x') \in B^d_r \times \mathbb{T}^d_\sigma$ as long as $|\epsilon| < \epsilon_0$ (see (A.3.11) for the estimate on ϵ_0). To estimate the measure of invariant tori for H_ϵ we need that Φ_ϵ possesses sufficient regularity as a map on $\{0\} \times \mathbb{T}^d \times \Omega$. More precisely we want to extend $\Phi_\epsilon|_{y=0}$, defined on $\mathbb{T}^d \times \Omega_\star$ to a C^1 function on the whole space $\mathbb{T}^d \times \Omega$ (see (B.1.4) , (B.1.5) and (B.1.2) for the definitions of the mentioned sets).

Let $A_0 \subseteq \mathbb{R}^p$ a closed set and E a Fréchet space.

Definition B.3.1. *A function $f : A_0 \to E$ is said to be* C^m in the sense of Whitney *(for $m \in \mathbb{N}$) if there exist $m + 1$ applications $f_0, f_1, \dots f_m$, with $f_l : A_0 \to L^l_s(\mathbb{R}^p, E)$ (that is to say l-linear and symmetric applications on A_0) such that if $R_l : A_0 \times A_0 \to L^l_s(\mathbb{R}^p, E)$ is defined by*

$$f_l(y) = \sum_{i \le m-l} \frac{1}{i!} f_{l+i}(x)(y - x)^i + R_l(x, y) \tag{B.3.1}$$

then for all $l = 0 \dots m$ and for all x, y belonging to a compact subset of $\mathcal{K}A_0$, it results

$$\sup\left\{\frac{R_l(x,y)}{\|x-y\|^{m-l}};\|x-y\|\le\delta\right\}\overset{\delta\to 0}{\longrightarrow}0$$

(that is to say $R_l(x,y)$ is uniformly $o\left(\|x-y\|^{m-l}\right)$ on compact subsets of A_0). We will indicate with C_w^m the space of such functions and with f_l the derivatives in the sense of Whitney of f . Besides f is said to be C^∞ in the sense of Whitney if it is C^m in the sense of Whitney for all $m \in \mathbb{N}$.

The fundamental theorem we are going to use to obtain our result is

Theorem B.3.1 (Whitney's extension theorem). *If f is of class $C^m(A_0, E)$ in the sense of Whitney, then there exists a function $\tilde{f} : \mathbb{R}^p \to E$ such that:*

1. *$\tilde{f}$ is C^m (in the classical sense) ;*
2. *$\tilde{f}(x) = f(x) \;\; \forall\; x \in A_0$;*
3. *$\tilde{f}^{(l)}(x) = f_l(x) \;\; \forall\; x \in A_0, \;\; \forall\; l \in \mathbb{R}^p$ with $|l|_1 \le m$.*

We refer to [Whi34] for a complete proof of this theorem.

Remark B.3.1. *If f is a function of $C_w^m(A_0)$ (we are merely interested in the case $m = 1$) then by Whitney's theorem it can be extended in particular to a $C^m(A)$ function where A is an arbitrarily chosen open set in $\mathbb{R}^p$ containing A_0. Since the derivatives of the extension $\tilde{f}$ coincide on A_0 with the Whitney's derivatives of f we will always indicate with the classical notation $f^{(l)}$, for $|l|_1 \le m$, the derivatives of a $C_w^m(A_0)$ function even if A_0 is a closed set, referring to them as the restriction of $\tilde{f}^{(m)}$ to A_0.*

Remark B.3.2. *In the proof of Whitney's extension theorem it can be observed that the norm of the extended function can be controlled as follows as far as a bounded set B containing A_0 is considered:*

$$\sup_{x\in B}|\tilde{f}^{(l)}(x)|\le c\sum_{l=0}^{m}\sup_{x\in A_0}|f_l(x)|$$

for some $c \in \mathbb{R}$.

Definition B.3.2. *We define $\mathcal{C}_a^m(\mathbb{T}^d \times \Omega\,;\,\sigma)$, with Ω open set of $\mathbb{R}^d$ and $\sigma > 0$ in $\mathbb{R}$ as the set of functions having values in $\mathbb{R}^p$ for some $p \in \mathbb{N}$, such that*

1. *$f \in C^m(\mathbb{T}^d \times \Omega)$*
2. *$f^{(l)}(\cdot, \omega) \in \mathcal{H}(\mathbb{T}^d_\sigma) \;\; \forall\; \omega \in \Omega, \;\; \forall\; |l|_1 = 0 \dots m$;*

$\mathcal{C}_a^m(\mathbb{T}_0^d \times \Omega\,;\,\sigma)$ *will indicate the linear subspace of* $\mathcal{C}_a^m$ *of function having vanishing average (* $\langle f(\cdot\,;\omega)\rangle = 0$ *) for all* $\omega \in \Omega$.

Our aim is to prove that the canonical transformation $\Phi_\epsilon(x', y'; \omega)$ in theorem A.4.1, considered for $y' = 0$, can be extended to a function belonging to $\mathcal{C}_a^1(\mathbb{T}^d \times \Omega\,;\sigma_\infty)$.

We start stating the following lemmata:

Lemma B.3.1. *Let* $f(x,\omega) \in \mathcal{H}(\mathbb{T}_\sigma^d)\ \ \forall\ \omega \in \Omega$ *and suppose that*

$$\sup_{\mathbb{T}_\sigma^d \times \Omega} |f(x,\omega)| \le M\ . \tag{B.3.2}$$

If $f(x,\omega) = \sum_{n\in\mathbb{Z}^d} f_n(\omega) e^{in\cdot x}$ *is its Fourier's expansion, that is*

$$f_n(\omega) = \frac{1}{2\pi}\int_0^{2\pi} f(x,\omega)e^{-in\cdot x}\ ,$$

then it results

$$\sup_{\Omega} |f_n(\omega)| \ \ \le\ \ Me^{-|n|\sigma}$$

Proof The statement can easily be obtained by Cauchy's integral formula for holomorphic function □

Lemma B.3.2. *Let* $\omega \in \Omega_\star = \{\omega \in \Omega\,|\,\omega \in \mathcal{D}_{\gamma,\tau}\}$ *and* $x \in \mathbb{T}_{\sigma-\delta}^d$*, consider the function*

$$g_n(\omega) = \frac{1}{n\cdot\omega}\,;$$

for $n \in \mathbb{Z}^d \smallsetminus \{0\}$*; fix* $\alpha \ge 1$ *in* $\mathbb{N}$ *and define*

$$g_n'(\omega;\alpha) = -\alpha\frac{n}{(n\cdot\omega)^{\alpha+1}}\,;$$

then

$$|(g_n(\omega+h))^\alpha - (g_n(\omega))^\alpha - g_n'(\omega;\alpha)\cdot h| \le |h|^2\frac{c(\alpha,\Omega_\star)}{\gamma^{2\alpha+1}}|n|^{\alpha+1+\tau(2\alpha+1)}\ ; \tag{B.3.3}$$

that is $(g_n(\omega))^\alpha \in C_w^1(\Omega_\star)$ *and* $g_n'(\omega;\alpha)$ *is its derivative in the sense of Whitney.*

Proof We first verify (B.3.3) in the case $\alpha = 1$:

$$
\begin{aligned}
&|g_n(\omega + h) - g_n(\omega) - g_n'(\omega; 1) \cdot h| = \\
= &\left| \frac{1}{n \cdot (\omega + h)} - \frac{1}{n \cdot \omega} + \alpha \frac{n \cdot h}{(n \cdot \omega)^2} \right| = \\
= &\frac{\left|(n \cdot \omega)^2 - n \cdot (\omega + h)(n \cdot \omega) + n \cdot (\omega + h)(n \cdot h)\right|}{|n \cdot (\omega + h)|\, |n \cdot \omega|^2} \leq \\
\leq &\left|(n \cdot \omega)^2 - n \cdot (\omega + h)(n \cdot \omega) + n \cdot (\omega + h)(n \cdot h)\right| \frac{|n|^{3\tau}}{\gamma^3}
\end{aligned}
$$

where we used the diophantine estimate satisfied by ω and $\omega + h$ in $\Omega_\star \subset \mathcal{D}_{\gamma,\tau}$. Then, simplifying the terms in the last expression we obtain

$$
|g_n(\omega + h) - g_n(\omega) - g_n'(\omega; \alpha) \cdot h| \leq |n \cdot h|^2 |n|^{3\tau} \gamma^{-3} \leq |h|^2 |n|^{3\tau+2} \gamma^{-3} \ .
$$

Now let $\alpha \geq 2$ we have

$$
\begin{aligned}
&|(g_n(\omega + h))^\alpha - (g_n(\omega))^\alpha - g_n'(\omega; \alpha) \cdot h| = \\
= &\left| \frac{1}{[n \cdot (\omega + h)]^\alpha} - \frac{1}{[n \cdot \omega]^\alpha} + \alpha \frac{n \cdot h}{(n \cdot \omega)^{\alpha+1}} \right| \leq \\
= &\frac{\left|(n \cdot \omega)^{\alpha+1} - [n \cdot (\omega + h)]^\alpha (n \cdot \omega)^{\alpha+1} + \alpha (n \cdot h)[n \cdot (\omega + h)]^\alpha\right|}{|n \cdot (\omega + h)|^\alpha |n \cdot \omega|^{\alpha+1}} \\
\leq &\left|(n \cdot \omega)^{\alpha+1} - [n \cdot (\omega + h)]^\alpha (n \cdot \omega)^{\alpha+1} + \alpha (n \cdot h)[n \cdot (\omega + h)]^\alpha\right| c_n
\end{aligned}
$$

where we define $c_n := |n|^{(2\alpha+1)\tau} \gamma^{-(2\alpha+1)}$. We now set $a = n \cdot \omega$ and $b = n \cdot h$ so that it results :

$$
\begin{aligned}
& (n\cdot\omega)^{\alpha+1} - [n\cdot(\omega+h)]^{\alpha}(n\cdot\omega)^{\alpha+1} + \alpha(n\cdot h)[n\cdot(\omega+h)]^{\alpha} = \\
= \;& a^{\alpha+1} - a(a+b)^{\alpha} + \alpha\, b(a+b)^{\alpha} = \\
= \;& a^{\alpha+1} - a\sum_{k=0}^{\alpha}\begin{pmatrix}\alpha\\ k\end{pmatrix} a^k b^{\alpha-k} + \alpha\, b\sum_{k=0}^{\alpha}\begin{pmatrix}\alpha\\ k\end{pmatrix} a^k b^{\alpha-k} = \\
= \;& a^{\alpha+1} - a\left[\sum_{k=0}^{\alpha-2}\begin{pmatrix}\alpha\\ k\end{pmatrix} a^k b^{\alpha-k} + \alpha\, a^{\alpha-1}b + a^{\alpha}\right] + \\
+ \;& \alpha\, b\left[\sum_{k=0}^{\alpha-1}\begin{pmatrix}\alpha\\ k\end{pmatrix} a^k b^{\alpha-k} + a^{\alpha}\right] = \\
= \;& -a\sum_{k=0}^{\alpha-2}\begin{pmatrix}\alpha\\ k\end{pmatrix} a^k b^{\alpha-k} + \alpha\, b\sum_{k=0}^{\alpha-1}\begin{pmatrix}\alpha\\ k\end{pmatrix} a^k b^{\alpha-k}\ .
\end{aligned}
\tag{B.3.4}
$$

Therefore, considering the absolute value and taking $j = k+2$, the last expression becomes

$$
\begin{aligned}
& \left|\sum_{j=2}^{\alpha}\begin{pmatrix}\alpha\\ j-2\end{pmatrix} a^{j-1}b^{\alpha-j+2} - \sum_{j=2}^{\alpha+1}\alpha\begin{pmatrix}\alpha\\ j-2\end{pmatrix} a^{j-2}b^{\alpha-j+3}\right| \leq \\
\leq \;& |b|^2\sum_{j=2}^{\alpha+1}(\alpha+1)!|a|^{j-2}|b|^{\alpha-j}\left(|a|+|b|\right)\ .
\end{aligned}
$$

We now take $s > 0$ such that $\Omega_\star \subset B_{\frac{s}{2}}(0)$ in order to obtain $|\omega|, |h| < s$. Thus, coming back to (B.3.4) considering the first and last member of the chain of equalities, we have:

$$
\begin{aligned}
& \left|(n\cdot\omega)^{\alpha+1} - [n\cdot(\omega+h)]^{\alpha}(n\cdot\omega)^{\alpha+1} + \alpha(n\cdot h)[n\cdot(\omega+h)]^{\alpha}\right| \leq \\
\leq \;& |n\cdot h|^2\sum_{j=2}^{\alpha+1}(\alpha+1)!|n\cdot\omega|^{j-2}|n\cdot h|^{\alpha-j}\left(|n\cdot\omega|+|n\cdot h|\right) \leq \\
\leq \;& |h|^2|n|^2\sum_{j=2}^{\alpha+1}(\alpha+1)!|n|^{\alpha-2}|\omega|^{j-2}|h|^{\alpha-j}|n|\left(|\omega|+|h|\right) \leq \\
\leq \;& |h|^2|n|^{\alpha+1}(\alpha+3)!s^{\alpha-1} = c(\alpha,\Omega_\star)|h|^2|n|^{\alpha+1}
\end{aligned}
$$

so that we can finally obtain

$$
\begin{aligned}
& |(g_n(\omega+h))^\alpha - (g_n(\omega))^\alpha - g_n'(\omega;\alpha)\cdot h| \leq \\
\leq\ & \left|(n\cdot\omega)^{\alpha+1} - [n\cdot(\omega+h)]^\alpha (n\cdot\omega)^{\alpha+1} + \alpha(n\cdot h)[n\cdot(\omega+h)]^\alpha\right| c_n \leq \\
\leq\ & |h|^2 c(\alpha,\Omega_\star)|n|^{\alpha+1+\tau(2\alpha+1)}\gamma^{-(2\alpha+1)} \qquad \square
\end{aligned}
$$

Corollary B.3.1. *Let $g_n(\omega)$ as in lemma B.3.2 for $\omega \in \mathcal{D}^d_{\gamma,\tau}$, then $g_n(\omega)$ can be extended to $\tilde{g}_n \in C^\infty(\mathbb{R}^d)$; besides for every B open set in $\mathbb{R}^d$ the following estimate hold*

$$\sup_B |\tilde{g}_n^{(p)}(\omega)| \leq c\gamma^{-(|p|_1+1)}|n|^{\tau(|p|_1+1)+|p|_1} \tag{B.3.5}$$

Proof The proof follows immediately by the formula

$$g_n^{(p)}(\omega) = (-1)^{|p|_1}\frac{n^p}{(n\cdot\omega)^{|p|_1+1}} = (-1)^{|p|_1} n^p (g_n(\omega))^{|p|_1+1}$$

for $p \in \mathbb{N}^d$ and where $n^p = (n_1^{p_1},\ldots,n_d^{p_d})$, by lemma B.3.2 and theorem B.3.1. The estimate (B.3.5) on the norm follows from remark B.3.2 and the diophantine estimate satisfied by ω $\square$

We now come back considering the Hamiltonian

$$\hat{H}_\epsilon(y,x;\omega) = E(\omega) + \omega\cdot y + \hat{Q}(y,x;\omega) + \epsilon\hat{f}(y,x;\omega) \tag{B.3.6}$$

consistently with all the notations of the preceding section; in particular we recall that $\hat{H}_\epsilon(y,x;\omega) = H_\epsilon(\omega_0^{-1}(\omega)+y,x)$ where $H_\epsilon(I,\varphi) = h(I) + \epsilon f(I,\varphi) \in \mathcal{H}(D\times\mathbb{T}^d)$ and

$$
\begin{aligned}
E(\omega) &= h\left({\omega_0}^{-1}(\omega)\right) \\
\hat{Q}(y,x;\omega) &= Q\left(y,x;{\omega_0}^{-1}(\omega)\right) = \frac{1}{2}\int_0^1 h''\left({\omega_0}^{-1}(\omega)+ty\right)\, dt\ \langle y,y\rangle \\
\hat{f}(y,x;\omega) &= f\left({\omega_0}^{-1}(\omega)+y,x\right)
\end{aligned}
$$

where $\omega_0 = h' : I \in B \subseteq D \longrightarrow \Omega$ is, by assumption, a diffeomorphism between the open sets B and Ω of $\mathbb{R}^d$. It can be immediately observed, by the assumptions made, that $\hat{H}_\epsilon(0,x;\omega) \in \mathcal{C}^1_a$. We remark that we are interested to analyze the case $y=0$ since we want to apply our following results to $f(x;\omega) = \Phi_\epsilon(0,x;\omega)$.

Proposition B.3.1. *For any* $f \in \mathcal{C}^1_a(\mathbb{T}^d_0 \times \Omega; \sigma_0)$ *(see definition B.3.2) there exists* $\tilde{f} \in \mathcal{C}^1_a(\mathbb{T}^d \times \Omega; \sigma_1)$ *such that*

$$\tilde{f}\Big|_{\mathbb{T}^d \times \Omega_\star} = \mathcal{D}_\omega{}^{-1} f$$

for any $\sigma_1 < \sigma_0$ *.*

Proof Let $f \in \mathcal{C}^1_a$ then by definition, f and $f^{(l)}$ are analytic over $\mathbb{T}^d$ for any $l \in \mathbb{N}^d$ with $|l|_1 = 1$; this implies, as seen in lemma B.3.1, that they can be expanded in Fourier's series

$$\begin{aligned} f(x,\omega) &= \sum_{n \in \mathbb{Z}^d \setminus \{0\}} f_n(\omega) e^{in\cdot x} \\ f^{(l)}(x,\omega) &= \sum_{n \in \mathbb{Z}^d \setminus \{0\}} f_{n,l}(\omega) e^{in\cdot x} \end{aligned}$$

where

$$\begin{aligned} f_n(\omega) &= \frac{1}{2\pi}\int_0^{2\pi} f(x,\omega) e^{-in\cdot x}\, dx \\ f_{n,l}(\omega) &= \frac{1}{2\pi}\int_0^{2\pi} f^{(l)}(x,\omega) e^{-in\cdot x}\, dx = f^{(l)}_n(\omega) \end{aligned}$$

and the following estimates hold

$$\sup_{\Omega} |f^{(l)}_n(\omega)| \le M_{|l|_1} e^{-|n|\sigma_0} \tag{B.3.7}$$

with $M_k = \max_{|l|_1 = k} \sup_{\mathbb{T}^d_\sigma \times \Omega} |f^{(l)}(x,\omega)|$.

Consider now for $\omega \in \Omega_\star$

$$\mathcal{D}_\omega{}^{-1} f(x;\omega) = \sum_{n \in \mathbb{Z}^d \setminus \{0\}} \frac{f_n(\omega)}{in\cdot\omega} e^{in\cdot x} = \sum_{n \in \mathbb{Z}^d \setminus \{0\}} -i f_n(\omega) g_n(\omega) e^{in\cdot x} \ ;$$

By lemma B.3.2 g_n can be extended to a $C^1(\Omega)$ function $\tilde{g}_n$whose derivatives coincide on $\Omega_\star$ with the derivatives in the sense of Whitney of g_n and verifies the estimate in (B.3.5). Now define $d_n(\omega) = -i f_n(\omega)\tilde{g}_n(\omega) \in C^1(\Omega)$ obtaining

$$D(x,\omega) = \sum_{n \in \mathbb{Z}^d \setminus \{0\}} d_n(\omega) e^{in\cdot x}$$

that extends ${\mathcal{D}_\omega}^{-1} f(x;\omega)$. We now prove that D belongs to $\mathcal{C}^1_a(\mathbb{T}^d \times \Omega;\, \sigma_1)$. We start observing that by lemma B.3.1 we obtain for $x \in \mathbb{T}^d_{\sigma_1}$

$$\sum_{n\in\mathbb{Z}^d\setminus\{0\}} \sup_\Omega |d_n(\omega)||e^{in\cdot x}| \le \sum_{n\in\mathbb{Z}^d\setminus\{0\}} \sup_\Omega |f_n(\omega)| \sup_\Omega |\tilde{g}_n(\omega)| e^{|n|\sigma_1} \le$$

$$\le \sum_{n\in\mathbb{Z}^d\setminus\{0\}} M_0 e^{-|n|\sigma_0} c\gamma^{-1}|n|^\tau e^{|n|\sigma_1} = cM_0\gamma^{-1} \sum_{n\in\mathbb{Z}^d\setminus\{0\}} e^{-|n|(\sigma_0-\sigma_1)}|n|^\tau$$

and this series converges for every $\sigma_1 < \sigma_0$; we so proved that $D \in \mathcal{H}(\mathbb{T}^d_{\sigma_1})$. Equivalently it can be shown that $\frac{\partial^{|l|}}{\partial\omega^l} D(x,\omega)$ belongs to $\mathcal{H}(\mathbb{T}^d_{\sigma_1})$ for every $l \in \mathbb{N}$ with $|l|_1 = 1$ since for such l we have

$$\begin{aligned}
\sup_\Omega |d^{(l)}_n(\omega)| &\le \sup_\Omega |f_{n,1}(\omega)\tilde{g}_n(\omega)| + \sup_\Omega |f_n(\omega)\tilde{g}^{(l)}_n(\omega)| \le \\
&\le cM_1 e^{-|n|\sigma_0}\gamma^{-1}|n|^\tau + cM_0 e^{-|n|\sigma_0}\gamma^{-2}|n|^{2\tau+1} \le \\
&\le c\max\left\{M_1\gamma^{-1}, M_0\gamma^{-2}\right\} |n|^{2\tau+1} e^{-|n|\sigma_0}
\end{aligned}$$

and hence for every $x \in \mathbb{T}^d_{\sigma_1}$ we get

$$\sum_{n\in\mathbb{Z}^d\setminus\{0\}} \sup_\Omega |d^{(l)}_n(\omega)||e^{in\cdot x}| \le$$

$$\le c\max\left\{M_1\gamma^{-1}, M_0\gamma^{-2}\right\}\gamma^{-1} \sum_{n\in\mathbb{Z}^d\setminus\{0\}} e^{-|n|(\sigma_0-\sigma_1)}|n|^{2\tau+1} < \infty\ .$$

Using the same estimates we can obtain that the series

$$\sum_{n\in\mathbb{Z}^d\setminus\{0\}} d^{(l_1)}_n(\omega) n^{l_2} e^{in\cdot x}$$

converges uniformly on $\mathbb{T}^d_{\sigma_1} \times \Omega$ for every $l = (l_1, l_2)$ with $|l|_1 = 1$ and therefore $D(x,\omega) \in C^1(\mathbb{T}^d \times \Omega)$. This completes the proof $\square$

Remark B.3.3. *The class of function $f \in \mathcal{C}^1_a(\mathbb{T}^d \times \Omega\,;\,\sigma)$ is closed under the operation of averaging on $\mathbb{T}^d$, inverting never vanishing functions and making products.*

Theorem B.3.2. *Let*

$$\Phi : (y', x'; \omega) \in B \times \mathbb{T}^d \times \Omega_\star \longrightarrow \Phi(y', x'; \omega) = (x, y; \omega) \in B \times \mathbb{T}^d \times \Omega_\star$$

be the symplectic transformation that puts $\hat{H}_\epsilon$ into Kolmogorov's normal form (see (B.2.1)), consistently with all the notations adopted in this section and in section B.2. Consider $f(x',\omega) = \Phi(0,x';\omega)$ defined on $\mathbb{T}^d \times \Omega_\star$; then, f can be extended to $\tilde{f}$ belonging to $\mathcal{C}^1_a(\mathbb{T}^d \times \Omega; \sigma_\infty)$.

Proof First of all, referring to the iterative scheme represented in (A.3.1), proposition A.2.1, lemma A.3.1 and proposition A.3.1 we write the general explicit formula for the symplectic transformation at the j-th step, once the Hamiltonian $H^{(j)} = E^{(j)} + \omega\cdot y^{(j)} + Q^{(j)} + \epsilon^{2^j} P^{(j)}$ is given. In the case of the dependence on additional parameters (in agreement with theorem A.4.1) we have that

$$\Phi^{(j)} : (y^{(j+1)}, x^{(j+1)}; \omega) \in B^d_{r_{j+1}} \times \mathbb{T}^d_{\sigma_{j+1}} \times \Omega_\star \longrightarrow (y^{(j)}, x^{(j)}; \omega) \in B^d_{r_j} \times \mathbb{T}^d_{\sigma_j} \times \Omega_\star$$

so that

$$H^{(j)} \circ \Phi^{(j)} = H^{(j+1)} = E^{(j+1)} + \omega \cdot y^{(j+1)} + Q^{(j+1)} + \epsilon^{2^{j+1}} P^{(j+1)}$$

is generated by

$$F_j\left(y^{(j+1)}, x^{(j)}; \omega\right) = y^{(j+1)} \cdot x^{(j)} + \epsilon^{2^j} g_j(y^{(j+1)}, x^{(j)}; \omega) \tag{B.3.8}$$

with

$$g_j(y^{(j+1)}, x^{(j)}; \omega) = \left[b_j(\omega) \cdot x^{(j)} + s_j(x^{(j)}; \omega) + a_j(x^{(j)}; \omega) \cdot y^{(j+1)}\right] \ .$$

In analogy with formulas (A.2.4), (A.2.6) and (A.2.7) the components of g_j are

$$s_j(x;\omega) \;=\; \mathcal{D}_\omega{}^{-1}(P^{(j)}(0,x;\omega) - \langle P^{(j)}(0,x;\omega)\rangle) \tag{B.3.9}$$

$$\begin{aligned} b_j(\omega) &= -\langle Q^{(j)}_{yy}(0,\cdot;\omega)\rangle^{-1} \langle P^{(j)}_y(0,\cdot;\omega) + \\ &+ Q^{(j)}_{yy}(0,\cdot;\omega) \cdot \partial_x s_j(\cdot;\omega)\rangle \end{aligned} \tag{B.3.10}$$

$$\begin{aligned} a_j(x;\omega) &= -\mathcal{D}_\omega{}^{-1}\left[P^{(j)}_y(0,x;\omega) + \right. \\ &+ \left. Q^{(j)}_{yy}(0,x;\omega) \cdot (b_j(\omega) + \partial_x s_j(x;\omega))\right] \ . \end{aligned} \tag{B.3.11}$$

Denoting with $\tilde{\varphi}_j$ the analytic diffeomorphism of $\mathbb{T}^d$ that inverts $\varphi_j(x;\omega) = x + \epsilon^{2^j} a_j(x;\omega)$ (having then the same dependence of a_j on ω) and taking the derivatives in (B.3.8) we obtain

$$\Phi^{(j)}(y^{(j+1)}, x^{(j+1)}; \omega) = \left\{ \begin{aligned} y^{(j)} &= y^{(j+1)} + \epsilon^{2^j}\left[b_j(\omega) + \tfrac{\partial}{\partial x} s_j(x;\omega)\right. \\ &\quad \left. + \left(\tfrac{\partial}{\partial x} a^T(x;\omega)\right) \cdot y^{(j+1)}\right]\Big|_{x = \tilde{\varphi}_j(x^{(j+1)};\omega)} \\ x^{(j)} &= \tilde{\varphi}_j(x^{(j+1)};\omega) \ . \end{aligned} \right. \tag{B.3.12}$$

Notice that $\Phi^{(j)}$ is a linear application in $y^{(j+1)}$ and therefore analytic for $y^{(j+1)}$ in $B_{r_{j+1}}$.

The components of the new Hamiltonian $H^{(j+1)}(y^{(j+1)}, x^{(j+1)}; \omega)$ are given (in analogy with equations (A.2.8), (A.2.10) and (A.2.9)) by

$$E^{(j+1)}(\omega) = E^{(j)}(\omega) + \epsilon^{2^j}(\omega \cdot b_j(\omega) + \langle P^{(j)}(0, \cdot; \omega)\rangle); \tag{B.3.13}$$

$$\begin{aligned} P^{(j+1)}(y, x; \omega) &= P_y^{(j)}(y, \tilde{\varphi}_j(x; \omega); \omega) \cdot \partial_x g_j(y, \tilde{\varphi}_j(x); \omega) + \\ &\int_0^1 (1-t) H_{yy}^{(j)}(y + t\epsilon^{2^j} \partial_x g_j, \tilde{\varphi}_j(x; \omega); \omega)\, dt \langle \partial_x g_j, \partial_x g_j \rangle \end{aligned} \tag{B.3.14}$$

$$Q^{(j+1)}(y, x; \omega) = Q^{(j)}(y, \tilde{\varphi}_j(x; \omega); \omega) + \epsilon^{2^j} \tilde{Q}^{(j)}(y, \tilde{\varphi}_j(x; \omega); \omega) \tag{B.3.15}$$

where $\tilde{Q}^{(j)} = Q_1^{(j)} + Q_2^{(j)} + Q_3^{(j)}$ with

$$\begin{aligned} Q_1^{(j)}(y, x; \omega) &= Q_y^{(j)}(y, x; \omega) \cdot (\partial_x a_j(x; \omega))^T \cdot y \\ Q_2^{(j)}(y, x; \omega) &= \left(\int_0^1 (1-t) Q_{yyy}^{(j)}(ty, x; \omega)\, dt \right) \langle y, y, b_j(\omega) + \partial_x s_j(x; \omega) \rangle \\ Q_3^{(j)}(y, x; \omega) &= \left(\int_0^1 (1-t) P_{yy}^{(j)}(ty, x; \omega)\, dt \right) \langle y, y \rangle \ . \end{aligned}$$

Now we prove by induction that at each step the symplectic transformation $\Phi^{(j)}\big|_{y=0}$ is an element of $\mathcal{C}_a^1(\mathbb{T}^d \times \Omega; \sigma_\infty)$ as a consequence of the fact that $\Phi^{(j)}\big|_{y=0}$ belongs to $\mathcal{C}_a^1(\mathbb{T}^d \times \Omega; \sigma_{j+1})$ for every $j \in \mathbb{N}$. By hypothesis we have $H^{(0)} := \hat{H}_\epsilon$ analytic for $x \in \mathbb{T}^d_{\sigma_0}$ and $y \in B_r$ and having a C^1 dependence on ω in $\Omega = \omega_0(B)$ (actually the dependence on ω is C^∞ having assumed that the frequency map $\omega_0 = h'$ is such a regular diffeomorphism) and this means that $H^{(0)}\big|_{y=0}$ belongs to $\mathcal{C}_a^1(\mathbb{T}^d \times \Omega; \sigma_0)$. Then by lemma B.3.1 and remark B.3.3 applied in formulas (B.3.9), (B.3.10), (B.3.11) and (B.3.12) for $j = 0$ we see that a_0, b_0, s_0 can be extended to $\mathcal{C}_a^1(\mathbb{T}^d \times \Omega; \sigma_1)$ functions as well as, by consequence, $\Phi^{(0)}$; identifying for simplicity every extension with the original symplectic transformation referring to it with the same name we obtain

$$\Phi^{(0)}\big|_{y=0} \in \mathcal{C}_a^1(\mathbb{T}^d \times \Omega; \sigma_1) \subseteq \mathcal{C}_a^1(\mathbb{T}^d \times \Omega; \sigma_\infty)$$

where σ_j is chosen as in (A.3.2) for all $j \in \mathbb{N}$.

Assume now that $H^{(j)}$ is analytic for $(y, x) \in B_{r_j} \times \mathbb{T}^d_{\sigma_j}$ and

$$\Phi^{(j-1)}\big|_{y=0} \in \mathcal{C}_a^1(\mathbb{T}^d \times \Omega; \sigma_j) \ ;$$

then, always by formulae (B.3.9), (B.3.10), (B.3.11) and (B.3.12) we obtain that

$$\Phi^{(j)}\big|_{y=0} \in \mathcal{C}_a^1(\mathbb{T}^d \times \Omega; \sigma_{j+1}) \subseteq \mathcal{C}_a^1(\mathbb{T}^d \times \Omega; \sigma_\infty) \ .$$

Moreover, observe that from equations (B.3.13), (B.3.14) and (B.3.15), the new Hamiltonian $H^{(j+1)}$ is analytic for $(y^{(j+1)}, x^{(j+1)}) \in B_{r_{j+1}} \times \mathbb{T}^d_{\sigma_{j+1}}$ and restricted to $y^{(j+1)} = 0$ is a $\mathcal{C}_a^1(\mathbb{T}^d \times \Omega; \sigma_{j+1})$ function.

We so proved that

$$f(x,\omega) = \Phi(0,x;\omega) = \lim_{j \longrightarrow \infty} \Phi^{(0)} \circ \Phi^{(1)} \circ \cdots \circ \Phi^{(j)}(0,x;\omega)$$

can be extended to $\tilde{f}$ in $\mathcal{C}_a^1(\mathbb{T}^d \times \Omega; \sigma_\infty)$ □

B.4 Measure of invariant tori

By Kolmogorov's theorem we obtained invariant tori for $H_\epsilon(I,\varphi) = h(I) + \epsilon f(I,\varphi)$ in the form

$$\mathcal{T}_\omega = \{I = \omega_0^{-1}(\omega) + y_0(0,x';\omega), \varphi = x_0(0,x';\omega) \,|\, x' \in \mathbb{T}^d\} \tag{B.4.1}$$

for any $\omega \in \Omega_\star \subset \mathcal{D}_{\gamma,\tau}$; we recall once again that H_ϵ is analytic on $D \times \mathbb{T}^d$ and $\omega_0 = h'$ is a diffeomorphism of class C^∞ on D. We remind also that we defined for $\delta > 0$ the following sets

$$\begin{aligned} B = B(\delta) &= \{I \in D \mid \text{dist}\,(I, \partial D) > \delta\}\,, \\ \Omega &= \omega_0(B), \\ \Omega_\star &= \Omega \cap \mathcal{D}_{\gamma,\tau}, \\ B_\star &= \omega_0^{-1}(\Omega_\star)\,; \end{aligned} \tag{B.4.2}$$

$B_\star$ is then the sets of vectors I such that the image through the "map of frequencies" is diophantine. In the previous section we proved that (for $|\epsilon| < \epsilon_0$) the components of the symplectic transformation Φ_ϵ restricted to $y = 0$, namely $y_0(0,x';\omega)$ and $x_0(0,x';\omega)$, can be extended to $\tilde{y}_0$ and $\tilde{x}_0$ both C^1 functions on $\mathbb{T}^d \times \Omega$.

We now define

$$K_\omega = \bigcup_{\omega \in \Omega_\star} \mathcal{T}_\omega \tag{B.4.3}$$

the union of Kolmogorov's invariant tori for H_ϵ in the parameter space. Adopting the same notation of theorem A.4.1 and writing explicitly the dependence on the parameter ϵ we have

$$\Phi_\epsilon(y',x';\omega) = (y_0(y',x';\omega), x_0(y',x';\omega)) = (y',x') + O(\epsilon) \tag{B.4.4}$$

therefore it results

$$\Phi_\epsilon(0, x'; \omega) = (y_0(0, x'; \omega), x_0(0, x'; \omega)) = (0, x') + O(\epsilon)\,; \tag{B.4.5}$$

more precisely we obtained

$$|y_0(0, x'; \omega)| \leq \epsilon C D \mu r$$

and

$$|x_0(0, x'; \omega) - x'| \leq \epsilon C D \mu r\,.$$

So, if δ is chosen to be greater than $\epsilon C D \mu r$ (C and D as defined in (A.4.1) and (A.4.2)), but sufficiently small to have $B(\delta) \neq \emptyset$ (that is another condition on the size of ϵ) we obtain

$$\omega_0^{-1}(\omega) + y_0(0, x'; \omega) \in \bigcup_{I \in B(\delta)} \{y \,:\, |y - I| \leq \epsilon C D \mu r\} \subseteq \bigcup_{I \in B(\delta)} \{|y - I| \leq \delta\} \subseteq D$$

which implies $\mathcal{T}_\omega \subset D$ and as simple consequence

$$K_\omega \subset D \times \mathbb{T}^d\ .$$

Remark B.4.1. *We can synthesize the condition to impose on δ and by consequence on ϵ, with the inequalities*

$$\epsilon C D \mu r \leq \delta < \sup_{y \in D}\ \sup\{r \,:\, B(y, r) \subseteq D\}\ \ ; \tag{B.4.6}$$

we will be from now on consistent with this choice.

At this point, before calculating the measure of K_ω, we need some lemmata.

Lemma B.4.1. *Let $B(0, R) = \{x \in \mathbb{R}^d \,:\, |x| < R\}$ be the open ball in $\mathbb{R}^d$ with center in the origin and radius $R > 0$; if $B_\star(0, R) = B(0, R) \cap \mathcal{D}^d_{\gamma,\tau}$ for fixed γ and $\tau > d - 1$ then*

$$\textit{meas}\ (B(0, R) \smallsetminus B_\star(0, R)) \leq c_{\tau,d} R^{d-1} \gamma \tag{B.4.7}$$

with

$$c_{\tau,d} = 2^d \sum_{n \in \mathbb{Z}^d \smallsetminus \{0\}} |n|^{-(\tau+1)} \tag{B.4.8}$$

Proof By the definition of $\mathcal{D}^d_{\gamma,\tau}$

$$B_\star(0, R) = \left\{ |\omega| < R \,:\, |\omega \cdot n| \geq \frac{\gamma}{|n|^\tau}\ \ \forall\ n \in \mathbb{Z}^d \smallsetminus \{0\} \right\}$$

and then putting $S = B(0,R) \setminus B_\star(0,R)$ we can write

$$\begin{aligned} S &= \left\{ |\omega| < R : \ \exists\, n \in \mathbb{Z}^d \setminus \{0\} \ : |\omega \cdot n| < \frac{\gamma}{|n|^\tau} \right\} \\ &= \left\{ |\omega| < R : \ \exists\, n \in \mathbb{Z}^d \setminus \{0\} \ : |\omega \cdot \frac{n}{|n|}| < \frac{\gamma}{|n|^{\tau+1}} \right\} \\ &= \bigcup_{n \in \mathbb{Z}^d \setminus \{0\}} \left\{ |\omega| < R : |\omega \cdot \frac{n}{|n|}| < \frac{\gamma}{|n|^{\tau+1}} \right\} \ . \end{aligned}$$

Observe now that $v = \frac{n}{|n|}$ is a unit vector in $\mathbb{R}^d$ and therefore there exists a rotation that maps v into $e_1 = (1, 0, \ldots, 0) \in \mathbb{R}^d$; name U such rigid movement of $\mathbb{R}^d$ with the property

$$Uv = U\frac{n}{|n|} = e_1 \ .$$

Since Lebesgue measure as well as the inner product in $\mathbb{R}^d$ are invariant under rotations, we can write

$$\begin{aligned} \operatorname{meas} S &= \operatorname{meas} U(S) = \\ &= \operatorname{meas} \bigcup_{n \in \mathbb{Z}^d \setminus \{0\}} \left\{ \omega \in \mathbb{R}^d, |U^{-1}\omega| < R : |U^{-1}\omega \cdot v| < \frac{\gamma}{|n|^{\tau+1}} \right\} = \\ &= \operatorname{meas} \bigcup_{n \in \mathbb{Z}^d \setminus \{0\}} \left\{ \omega \in \mathbb{R}^d, |U^{-1}\omega| < R \ : \ |\omega \cdot Uv| < \frac{\gamma}{|n|^{\tau+1}} \right\} = \\ &= \operatorname{meas} \bigcup_{n \in \mathbb{Z}^d \setminus \{0\}} \left\{ \omega \in \mathbb{R}^d, |\omega| < R \ : \ |\omega \cdot e_1| < \frac{\gamma}{|n|^{\tau+1}} \right\} \ ; \end{aligned}$$

(we obviously used that the set $\{\omega \in \mathbb{R}^d \ : \ |x| < R\}$ is invariant under rotations); carrying on this equalities we obtain

$$\begin{aligned} \operatorname{meas} S &= \operatorname{meas} \bigcup_{n \in \mathbb{Z}^d \setminus \{0\}} \left\{ \omega \in \mathbb{R}^d, |\omega| < R \ : \ |\omega \cdot e_1| < \frac{\gamma}{|n|^{\tau+1}} \right\} \leq \\ &\leq \sum_{n \in \mathbb{Z}^d \setminus \{0\}} \left\{ \omega \in \mathbb{R}^d, |\omega| < R \ : \ |\omega_1| < \frac{\gamma}{|n|^{\tau+1}} \right\} \leq \\ &\leq \sum_{n \in \mathbb{Z}^d \setminus \{0\}} 2^d R^{d-1} \frac{\gamma}{|n|^{\tau+1}} = c_{\tau,d} R^{d-1} \gamma \ . \end{aligned}$$

The last thing to be observed to complete the proof is that the sum that defines $c_{\tau,d}$ converges:

$$\sum_{n\in\mathbb{Z}^d\setminus\{0\}} \frac{1}{|n|^{\tau+1}} \leq c\int_{|x|\geq 1}\frac{dx}{|x|^{\tau+1}} \leq c'\int_{\rho\geq 1}\frac{1}{\rho^{\tau+1}}\rho^{d-1}d\rho = c'\int_{\rho\geq 1}\frac{d\rho}{\rho^{\tau-d+2}} < +\infty$$

since $\tau - d + 2 > (d-1) - d + 2 = 1$ $\square$

Corollary B.4.1. *Let $A \subseteq \mathbb{R}^d$ bounded, define $A_\star = A \cap \mathcal{D}^d_{\gamma,\tau}$, for fixed γ and $\tau > d-1$, and*

$$c_A = c_{\tau,d}(\textit{ meas } A)^{-1}\sup_{x\in A}|x|^{d-1} \tag{B.4.9}$$

where $c_{\tau,d}$ is defined in (B.4.8); then

$$\textit{meas } (A \setminus A_\star) \leq c_A(\textit{ meas } A)\gamma\ . \tag{B.4.10}$$

Proof Define

$$R = \sup_{x\in A}|x|$$

by the hypothesis of boundness on A, R is finite and we have $A \subseteq B(0,R)$. Then simply observe

$$A\setminus A_\star = \bigcup_{n\in\mathbb{Z}^d\setminus\{0\}}\left\{\omega\in A : |\omega\cdot n| < \frac{\gamma}{|n|^{\tau+1}}\right\} \subseteq$$

$$\subseteq \bigcup_{n\in\mathbb{Z}^d\setminus\{0\}}\left\{\omega\in B(0,R) : |\omega\cdot n| < \frac{\gamma}{|n|^{\tau+1}}\right\} = B(0,R)\setminus B_\star(0,R)\,.$$

Therefore by lemma B.4.1, taking c_A as in (B.4.9), it results

$$\begin{aligned}\text{meas } (A\setminus A_\star) &\leq \text{meas } (B(0,R)\setminus B_\star(0,R)) \leq \\ &\leq c_{\tau,d}R^{d-1}\gamma = c_A(\text{ meas } A)\gamma \qquad \square\end{aligned}$$

We are now ready to state and prove the main theorem of this chapter concerning the measure of the union K_ω of maximal invariant tori for H_ϵ carrying quasi-periodic motions:

Theorem B.4.1. *Let $H_\epsilon(I,\varphi) = h(I) + \epsilon f(I,\varphi)$ be a real-analytic Hamiltonian on $D\times\mathbb{T}^d$ (where D is some open set in $\mathbb{R}^d$). Let the usual condition $\epsilon CD\mu < 1$ be satisfied for C and D as in (A.4.1) and (A.4.2) respectively. Then,*

$$\textit{meas}_{\,2d}K_\omega \geq (1-|O(\epsilon)|)\,(1-c_\Omega\gamma)\inf_{\omega\in\Omega}\left|\det\frac{\partial}{\partial\omega}\omega_0^{-1}(\omega)\right|(2\pi)^d(\textit{ meas}_{\,d}\Omega) \tag{B.4.11}$$

where ω_0 is defined in (B.1.2), c_Ω in (B.4.9), $\Omega(\gamma,\tau)$ is defined by (B.1.4) and (B.1.3) and K_ω by (B.4.3) and (B.4.1).

Proof Recall first that

$$K_\omega = \bigcup_{\omega\in\Omega_\star} \mathcal{T}_\omega$$

where $\mathcal{T}_\omega$ is an invariant torus for H_ϵ, for any $\omega \in \Omega_\star = \omega_0(B) \cap \mathcal{D}^d_{\gamma,\tau}$. We can parametrize $\mathcal{T}_\omega$ by equation (B.4.1) as follows:

$$\mathcal{T}_\omega = \left\{(I,\varphi) \in D \times \mathbb{T}^d \,:\, I = \eta(x,\omega), \varphi = \xi(x,\omega)\,,\, x \in \mathbb{T}^d\right\}$$

where

$$\eta(x,\omega) \;=\; \omega_0^{-1}(\omega) + \tilde{y}_0(0,x;\omega) \tag{B.4.12}$$

$$\xi(x,\omega) \;=\; \tilde{x}_0(0,x;\omega)\;. \tag{B.4.13}$$

Remind that we defined $\tilde{x}_0$ and $\tilde{y}_0$ as C^1 extensions of the components of the symplectic transformation $\Phi_\epsilon|_{y=0}$; moreover, by hypothesis on H_ϵ, it results that ω_0 (the frequency map) is a C^∞ diffeomorphism and therefore we have that η and ξ are both C^1 functions defined on the whole space $\mathbb{T}^d \times \Omega$.

To estimate the measure of K_ω in the phase space $D \times \mathbb{T}^d$, we use the change of variables theorem for integrals in $\mathbb{R}^n$ obtaining

$$\text{meas}\, K_\omega = \text{meas} \bigcup_{\omega\in\Omega_\star} \mathcal{T}_\omega = \int_{\bigcup_{\omega\in\Omega_\star}\mathcal{T}_\omega} dI d\varphi = \iint_{\mathbb{T}^d\times\Omega_\star} \left|\det \frac{\partial(\eta,\xi)}{\partial(x,\omega)}\right| \, dx d\omega. \tag{B.4.14}$$

Our task is now to estimate

$$\sup_{\mathbb{T}^d\times\Omega_\star} |d(x,\omega)| \qquad \text{with} \qquad d(x,\omega) := \det \frac{\partial(\eta,\xi)}{\partial(x,\omega)}\;.$$

More explicitly, from equations (B.4.12) and (B.4.13), we can write

$$d(x,\omega) = \det \begin{pmatrix} \frac{\partial}{\partial x}\tilde{y}_0(0,x;\omega) & \frac{\partial}{\partial\omega}\omega_0^{-1}(\omega) + \frac{\partial}{\partial\omega}\tilde{y}_0(0,x;\omega) \\ \frac{\partial}{\partial x}\tilde{x}_0(0,x;\omega) & \frac{\partial}{\partial\omega}\tilde{x}_0(0,x;\omega) \end{pmatrix}\;.$$

Thus, from equation (B.4.5) we get

$$\frac{\partial}{\partial x}\tilde{y}_0(0,x;\omega) = \frac{\partial}{\partial\omega}\tilde{x}_0(0,x;\omega) = \frac{\partial}{\partial\omega}\tilde{y}_0(0,x;\omega) = O_{d\times d}(\epsilon)$$

$$\frac{\partial}{\partial x}\tilde{x}_0(0,x;\omega) = \mathbb{I}_d + O_{d\times d}(\epsilon)$$

where $\mathbb{I}_d$ is the unit matrix $d \times d$ and $O_{d\times d}(\epsilon)$ denotes some $d \times d$ matrix such that each element is $O(\epsilon)$. Since $\frac{\partial}{\partial\omega}\omega_0^{-1}(\omega) = h''(I)^{-1}\big|_{I=\omega_0^{-1}(\omega)}$, we obtain

$$d(x,\omega) = \det \begin{pmatrix} O_{d\times d}(\epsilon) & h''(I)^{-1}\big|_{I=\omega_0^{-1}(\omega)} + O_{d\times d}(\epsilon) \\ \\ \mathbb{I}_d + O_{d\times d}(\epsilon) & O_{d\times d}(\epsilon) \end{pmatrix} . \tag{B.4.15}$$

By indicating with $\mathbb{O}_{d\times d}$ the null matrix $d \times d$ we define

$$A = \begin{pmatrix} \mathbb{O}_{d\times d} & h''(I)^{-1}\big|_{I=\omega_0^{-1}(\omega)} \\ \\ \mathbb{I}_{d\times d} & \mathbb{O}_{d\times d} \end{pmatrix} \tag{B.4.16}$$

so that, from equation (B.4.15), it results

$$\begin{aligned} d(x,\omega) &= \det\left(A + O_{2d\times 2d}(\epsilon)\right) = (\det A)\det\left(\mathbb{I}_{2d\times 2d} + A^{-1}O_{2d\times 2d}(\epsilon)\right) = \\ &= (\det A)\det\left(\mathbb{I}_{2d\times 2d} + O_{2d\times 2d}(\epsilon)\right) = (\det A)\left(1 + O(\epsilon)\right) \end{aligned}$$

(observe that A^{-1} obviously exists). Finally, by definition of A in (B.4.16), we have

$$\begin{aligned} d(x,\omega) &= (\det A)\left(1 + O(\epsilon)\right) = \left(\det h''(\omega_0^{-1}(\omega))\right)^{-1}\left(1 + O(\epsilon)\right) = \\ &= \left(\det \frac{\partial}{\partial\omega}\omega_0^{-1}(\omega)\right)\left(1 + O(\epsilon)\right) . \end{aligned}$$

We now come back to equation (B.4.14) to obtain the wanted estimate in (B.4.11) using (B.4.10) and definition (B.4.9):

$$\begin{aligned} \operatorname{meas} K_\omega &= \iint_{\mathbb{T}^d\times\Omega_\star} |d(x,\omega)| \; dx d\omega = \\ &= (1 + O(\epsilon)) \iint_{\mathbb{T}^d\times\Omega_\star} \left|\det \frac{\partial}{\partial\omega}\omega_0^{-1}(\omega)\right| dx d\omega \geq \\ &\geq (1 - |O(\epsilon)|) \inf_{\omega\in\Omega} \left|\det \frac{\partial}{\partial\omega}\omega_0^{-1}(\omega)\right| \operatorname{meas}\left(\mathbb{T}^d \times \Omega_\star\right) \geq \\ &\geq (1 - |O(\epsilon)|) \inf_{\omega\in\Omega} \left|\det \frac{\partial}{\partial\omega}\omega_0^{-1}(\omega)\right| (2\pi)^d (\operatorname{meas}\Omega)(1 - c_\Omega\gamma) \quad \square \end{aligned}$$

Appendix C

Rüßmann's theory for lower dimensional elliptic tori

In this Appendix we review Rüßmann's theorem on the existence of lower dimensional elliptic tori for nearly-integrable and analytic Hamiltonian systems. All results discussed can be found in [Rüßm01] to which we always refer for fully detailed proofs.

C.1 Preliminaries to Rüßmann's theorem

We first introduce some notations

- Consider vectorial spaces $V_j = \mathbb{C}^{n_j}$ for $j = 1 \dots m$; then we denote for $s \in \mathbb{R}^m_+$

$$\mathcal{D}(s) = \{v = (v_1, \dots, v_m),\ v_j \in V_j \ :\ |v_j| < s_j \ \ \forall\ j = 1 \dots m\}\ . \quad \text{(C.1.1)}$$

- Let $X, M \in \ \text{mat}_{\mathbb{C}}(m \times m)$ with $X = X^t$, then we denote with $[[M]]$ the matrix of the linear mapping

$$X \longmapsto MX + XM^t = B \quad \text{(C.1.2)}$$

  (this implies obviously $B = B^t$) provided that only the elements X_{jl} with $j \leq l$ are considered.

  More precisely if we associate to any matrix $A = (A_{jl}) \in \ \text{mat}_{\mathbb{C}}(m \times m)$ the vector

$$\vec{A} = (A_{11}, A_{12}, \dots, A_{1m}, A_{22}, \dots, A_{2m}, \dots, A_{m-1m-1}, A_{m-1m}, A_{mm})$$

belonging to $\mathbb{C}^{\frac{m(m+1)}{2}}$, then we can represent equation (C.1.2) in the form

$$[[M]]\,\vec{X} = \vec{B}$$

where $[[M]] = \left([[M]]_{jk,ln}\right) \in \ \mathrm{mat}_{\mathbb{C}}\left(\frac{m(m+1)}{2} \times \frac{m(m+1)}{2}\right)$ is defined by the expression

$$\left([[M]]_{jk,ln}\right) = \left((1 - \frac{1}{2}\delta_{jk})(M_{lj}\delta_{nk} + M_{lk}\delta_{nj} + M_{nk}\delta_{lj} + M_{nl}\delta_{kj})\right)$$

where δ_{ij} are the classical Kronecker symbols.

Now we recall the definition of approximation function already given in 2.2.1 for easier further references:

Definition C.1.1 (Approximation function). *A continuous function* Φ: $[0, \infty) \to \mathbb{R}$ *is called an approximation function if:*

1. $1 = \Phi(0) \geq \Phi(s) \geq \Phi(t) > 0$ *for* $0 \leq s < t < \infty$*;*
2. $\Phi(1) = 1$ *so that* $\Phi(s) = 1$ *for any* $0 \leq s \leq 1$*;*
3. $s^{\lambda}\Phi(s) \overset{s\to\infty}{\longrightarrow} 0$ *for any* $\lambda > 0$ *so that* $\sup_{s\geq 1} s^{\lambda}\Phi(s)^{\frac{1}{\mu}} < \infty$ *for all* $\mu > 0$ *and* $\lambda \geq 0$*;*
4. $\int_1^{\infty} \log \frac{1}{\Phi(s)} \frac{ds}{s^2} < \infty$.

C.1.1 Consequences of weakly non-degeneration

In this paragraph we consider a real-analytic function

$$\chi = (\omega, A) : B \longrightarrow \mathbb{R}^d \times \mathbb{R}^{2p\times 2p} \tag{C.1.3}$$

defined on a domain B, where A is a $2p \times 2p$ symmetric matrix satisfying the following conditions:

1. there exist p real-analytic functions on B, $\Omega_1(y), \ldots, \Omega_p(y)$, such that

   $$i\Omega_1(y), \ldots, i\Omega_p(y), -i\Omega_1(y), \ldots, -i\Omega_p(y)$$

   are the $2p$ eigenvalues of the matrix AJ_{2p}.
2. the vector

   $$(\omega, \Omega_1, \ldots, \Omega_p) : B \longrightarrow \mathbb{R}^d \times \mathbb{R}^p$$

   is *weakly non-degenerate*.

Lemma C.1.1. *Let* (ω, Ω) *a real-analytic and weakly non-degenerate function; then, in the notation of definition 2.1.2, we have*

$$\mathbb{Z}^p_{\omega\Omega} = U \cup V \ \ \textit{with } U \cap V = \emptyset$$

where

$$\begin{array}{rcl} U & = & \{l \in \mathbb{Z}^p_{\omega\Omega} \mid (\omega, \langle l, \Omega\rangle) \ \textit{ is non-degenerate}\} \\ V & = & \{l \in \mathbb{Z}^p_{\omega\Omega} \mid \ \exists \ \tau_l \in \mathbb{R}^d \setminus \mathbb{Z}^d : \langle \tau_l, \omega\rangle = \langle l, \Omega\rangle\}, \end{array}$$

with the numbers τ_l *uniquely determined. Besides almost all points* $b \in B$ *satisfy* $\langle k, \omega(b)\rangle \neq 0$ *for all* $k \in \mathbb{Z}^d$ *and* $\langle k, \omega(b)\rangle + \langle l, \Omega(b)\rangle \neq 0$ *for all* $(k, l) \in \mathbb{Z}^d \times \mathbb{Z}^p_{\omega\Omega}$*.*

In view of this lemma we can define the first of the important numbers needed to "quantize" the non-degeneracy of χ:

Definition C.1.2 (Amount of degeneracy of χ). *The amount of degeneracy of a weakly non-degenerate function is defined by*

$$\alpha = \alpha(\chi) = \begin{cases} \min_{k \in \mathbb{Z}^d, l \in V} |k + \tau_l|_2 |(k, 1)|_2{}^{-1} & V \neq \emptyset \\ \\ 1 & V = \emptyset \end{cases} \tag{C.1.4}$$

Lemma C.1.2. *Let* $f_j : B \longrightarrow \mathbb{R}^{m_j}$ *be real-analytic and non-degenerate functions defined on a domain* $B \subseteq \mathbb{R}^d$*, for each* $j = 1 \ldots N$*. Consider*

$$\mathcal{C} = \mathbb{R}^{m_1} \times \cdots \times \mathbb{R}^{m_N}$$

and let $c = (c_1, \ldots, c_N) \in \mathcal{C}$ *be some parameters. If we define* $f : \mathcal{C} \times B \longrightarrow \mathbb{R}$ *as the real-analytic function (with respect to the* y *variables)*

$$f(c, y) = \prod_{j=1}^{N} \langle c_j, f_j(y)\rangle \tag{C.1.5}$$

and $\mathcal{S}$ *as the following subset of* $\mathcal{C}$

$$\mathcal{S} = \{c = (c_1, \ldots, c_N) \in \mathcal{C} \mid |c_j|_2 = 1 \ \ \forall \ j = 1 \ldots N\},$$

then for any non-void compact set $\mathcal{K} \subset B$ *there exist numbers* $\mu_0 = \mu_0(f, \mathcal{K}) \in \mathbb{N}$ *and* $\beta = \beta(f, \mathcal{K}) > 0$ *such that*

$$\max_{0 \leq \mu \leq \mu_0} |D^\mu f(c, y)| \geq \beta \quad \forall \ c \in \mathcal{S}, \ \ \forall \ y \in \mathcal{K} \tag{C.1.6}$$

(here and in the sequel, D refers to the y *variables whereas* c *is considered a parameter).*

Observe that if we consider a function f as in (C.1.5), then the function

$$(c, y) \in \mathcal{C} \times B \longrightarrow \max_{0 \leq \nu \leq \mu} |D^{\nu} f(c, y)|$$

is continuous in $\mathcal{C} \times B$ for every $\mu \in \mathbb{N}$. Therefore the number

$$\beta(f, \mu, \mathcal{K}) := \min_{y \in \mathcal{K}, |c|_2 = 1} \max_{0 \leq \nu \leq \mu} |D^{\nu} f(c, y)|$$

is well defined for any compact set $\mathcal{K} \subset B$ and verifies $\beta(f, \mu_1, \mathcal{K}) \leq \beta(f, \mu_2, \mathcal{K})$ for every $0 \leq \mu_1 \leq \mu_2$. Then, by Lemma C.1.2 we can well define the numbers $\mu_0(f, \mathcal{K})$ and $\beta(f, \mu_0, \mathcal{K})$ as follows

Definition C.1.3. *We call* **index of non-degeneracy** *of f with respect to $\mathcal{K}$ the first integer μ_0 such that $\beta(f, \mu_0, \mathcal{K}) > 0$ (while $\beta(f, \mu, \mathcal{K}) = 0$ for every $\mu < \mu_0$); we call the number $\beta(f, \mu_0(f, \mathcal{K}), \mathcal{K})$* **amount of non-degeneracy** *of f with respect to $\mathcal{K}$.*

What we want to do now is to define the index of non-degeneracy and the amount of non-degeneracy of the real-analytic function $\chi = (\omega, A)$. We define the following three functions:

$$[\chi]_k^{(1)}(y) := |\langle k, \omega \rangle|^2 |k|_2{}^{-2} \tag{C.1.7}$$

$$[\chi]_k^{(2)}(y) := C_k^{(2)} |\det(i\langle k, \omega \rangle \mathbb{I}_{2p} + AJ_{2p})|^2 \tag{C.1.8}$$

$$[\chi]_k^{(3)}(y) := C_k^{(3)} |k|_2{}^{-2p-2r} |\det\left(i\langle k, \omega \rangle \mathbb{I}_{2p^2+p} + [[AJ_{2p}]]\right)|^2 \tag{C.1.9}$$

where

$$r = \#\{l \in \mathbb{Z}^p \,|\, |l|_1 = 2\,,\ \langle l, \Omega \rangle = 0\} \tag{C.1.10}$$

and

$$C_k^{(2)} = |(k, 1)|_2{}^{-2\#U_1} \prod_{l \in V_1} |k + \tau_l|_2{}^{-1} |-k + \tau_l|_2{}^{-1} \tag{C.1.11}$$

$$C_k^{(3)} = |(k, 1)|_2{}^{-2\#U_2} \prod_{l \in V_2} |k + \tau_l|_2{}^{-1} |-k + \tau_l|_2{}^{-1} \tag{C.1.12}$$

with

$$U_j = \{l \in \mathbb{Z}^p \,|\, |l|_1 = j \;:\; (\omega, \langle l, \Omega \rangle) \text{ is non-degenerate}\} = U \cap \mathbb{Z}_j^p \tag{C.1.13}$$

$$V_j = \{l \in \mathbb{Z}^p \,|\, |l|_1 = j,\ \exists\, \tau_l \in \mathbb{R}^d \smallsetminus \mathbb{Z}^d : \langle \tau_l, \omega \rangle = \langle l, \Omega \rangle\} = V \cap \mathbb{Z}_j^p \tag{C.1.14}$$

for $j = 1, 2$.

Proposition C.1.1. *Let $[\chi]_k^{(j)}(y)$ be as defined in (C.1.7), (C.1.8), and (C.1.9); then, under the hypotheses made on $\chi = (\omega, A)$, they can be represented (for every $k \geq |j-2|$) in the form (C.1.5) with the properties described in lemma C.1.2.*

Before proving this proposition we need two preliminary lemmata; we refer to [Rüßm01, page 141-142] for the proof of the first lemma while the second can be obtained as a corollary with the help of some easy calculations.

Lemma C.1.3. *Let $M \in \mathit{mat}_{\mathbb{C}}(m \times m)$ with eigenvalues $\lambda_1, \ldots, \lambda_m$. Then the eigenvalues of the matrix $[[M]] \in \mathit{mat}_{\mathbb{C}}\left(\frac{1}{2}m(m+1) \times \frac{1}{2}m(m+1)\right)$ (see section C.1 for the definition of $[[M]]$) are the $\frac{1}{2}m(m+1)$ functions $\lambda_i + \lambda_j$ for $1 \leq i \leq j \leq m$.*

Lemma C.1.4. *Let $\lambda \in \mathbb{R}$ then the following two equalities hold:*

$$\begin{aligned} |\det(i\lambda \mathbb{I}_{2p} + AJ_{2p})|^2 &= \prod_{l \in U_1 \cup V_1} (\lambda + \langle l, \Omega \rangle)(-\lambda + \langle l, \Omega \rangle) \\ |\det \lambda \mathbb{I}_{2p^2+p} + [[AJ_{2p}]]|^2 &= \lambda^{2p+2r} \prod_{l \in U_2 \cup V_2} (\lambda + \langle l, \Omega \rangle)(-\lambda + \langle l, \Omega \rangle) \end{aligned}$$

for U_j, V_j defined in (C.1.13) and (C.1.14) and $r = \#\{l \in \mathbb{Z}^p \,|\, |l|_1 = 2\,,\, \langle l, \Omega \rangle = 0\}$.

Proof [proposition C.1.1] As it can be easily seen $[\chi]_k^{(1)}$ is already put in the form (C.1.5), in fact:

$$[\chi]_k^{(1)} = \prod_{j=1}^{N} \langle c_j, f_j(y) \rangle \text{ with } N = 2, f_j(y) = \omega(y), c_j = k|k|^{-1} \text{ for } j = 1, 2.$$

For what concerns $[\chi]_k^{(2)}$, recall the definitions given in (C.1.8) and (C.1.11) and apply lemma C.1.4 with $\lambda = \langle \omega, k \rangle$ using the definition of V_1 (and consequently of τ_l) in (C.1.14); thus, we may write the following equalities:

$$
\begin{aligned}
[\chi]_k^{(2)} &= C_k^{(2)} |\det(i\langle k, \omega\rangle \mathbb{I}_{2p} + AJ_{2p})|^2 = \\
&= C_k^{(2)} \prod_{l\in U_1} (\langle k, \omega(y)\rangle + \langle l, \Omega(y)\rangle)(\langle -k, \omega(y)\rangle + \langle l, \Omega(y)\rangle) \\
&* \prod_{l\in V_1} (\langle k, \omega(y)\rangle + \langle l, \Omega(y)\rangle)(\langle -k, \omega(y)\rangle + \langle l, \Omega(y)\rangle) = \\
&= C_k^{(2)} \prod_{l\in U_1} (\langle k, \omega(y)\rangle + \langle l, \Omega(y)\rangle)(\langle -k, \omega(y)\rangle + \langle l, \Omega(y)\rangle) \\
&* \prod_{l\in V_1} \langle k + \tau_l, \omega(y)\rangle \langle -k + \tau_l, \omega(y)\rangle = \\
&= \prod_{l\in U_1} \frac{\langle k, \omega(y)\rangle + \langle l, \Omega(y)\rangle}{|(k,1)|_2} \prod_{l\in U_1} \frac{\langle -k, \omega(y)\rangle + \langle l, \Omega(y)\rangle}{|(k,1)|_2} \\
&* \prod_{l\in V_1} \frac{\langle k + \tau_l, \omega(y)\rangle}{|k + \tau_l|_2} \prod_{l\in V_1} \frac{\langle -k + \tau_l, \omega(y)\rangle}{|-k + \tau_l|_2};
\end{aligned}
$$

observing the non-degeneracy of $(\omega, \langle l, \Omega\rangle)$ for $l \in U_1$ and the non-degeneracy of ω we conclude that effectively $[\chi]_k^{(2)}$ is in the form (C.1.5) and satisfies all the hypotheses in lemma C.1.2.

A completely analogous proceeding can be adopted for $[\chi]_k^{(3)}$; in fact the following equalities are verified in view of definition (C.1.9) (together with (C.1.12)), lemma C.1.4 with $\lambda = \langle \omega, k\rangle$, the definition of V_2 (and consequently of τ_l) in

(C.1.14) and the definition of r in (C.1.10):

$$\begin{aligned}
[\chi]_k^{(3)} &= C_k^{(3)} |\det\left(i\langle k,\omega\rangle \mathbb{I}_{2p^2+p} + [[AJ_{2p}]]\right)|^2 = \\
&= C_k^{(3)} \prod_{j=1}^{2p+2r} \langle k,\omega\rangle \prod_{l\in U_2} \left(\langle k,\omega\rangle + \langle l,\Omega\rangle\right)\left(\langle -k,\omega\rangle + \langle l,\Omega\rangle\right) \\
&* \prod_{l\in V_2} \left(\langle k,\omega\rangle + \langle l,\Omega\rangle\right)\left(\langle -k,\omega\rangle + \langle l,\Omega\rangle\right) = \\
&= C_k^{(3)} \prod_{j=1}^{2p+2r} \langle k,\omega\rangle \prod_{l\in U_2} \left(\langle k,\omega\rangle + \langle l,\Omega\rangle\right)\left(\langle -k,\omega\rangle + \langle l,\Omega\rangle\right) \\
&* \prod_{l\in V_2} \langle k+\tau_l,\omega\rangle \langle -k+\tau_l,\omega\rangle = \\
&= \prod_{j=1}^{2p+2r} \frac{\langle k,\omega\rangle}{|k|_2} \prod_{l\in U_2} \frac{\langle k,\omega\rangle + \langle l,\Omega\rangle}{|(k,1)|_2} \frac{\langle -k,\omega\rangle + \langle l,\Omega\rangle}{|(k,1)|_2} \\
&* \prod_{l\in V_2} \frac{\langle k+\tau_l,\omega\rangle}{|k+\tau_l|_2} \frac{\langle -k+\tau_l,\omega\rangle}{|-k+\tau_l|_2};
\end{aligned}$$

recalling the non-degeneracy of ω and the non-degeneracy of $(\omega, \langle l,\Omega\rangle)$ for $l \in U_2$ we conclude the proof $\square$

We can finally state

Proposition C.1.2 (Index and amount of non-degeneracy). *Let $[\chi]_k^{(j)}$ as defined in (C.1.7), (C.1.8) and (C.1.9), then, for every $k \in \mathbb{Z}^d$ such that $|k|_2 \geq |j-2|$ for $j = 1,2,3$, there exist integers $\mu_0^{(j)} = \mu_0([\chi]_k^{(j)}, \mathcal{K}) \in \mathbb{N}$ and numbers $\beta^{(j)} = \beta([\chi]_k^{(j)}, \mathcal{K}) > 0$ characterized by*

$$\min_{y\in\mathcal{K}} \max_{0\leq\nu\leq\mu_0^{(j)}} \left| D^\nu [\chi]_k^{(j)}(y) \right| \geq \beta^{(j)}, \ \forall\ k \in \mathbb{Z}^d,$$

$$\min_{y\in\mathcal{K}} \left| D^\mu [\chi]_k^{(j)}(y) \right| = 0, \ \forall\ \mu < \mu_0^{(j)}.$$

We call **index of non-degeneracy** *of $\chi(y) = (\omega(y), A(y))$ with respect to $\mathcal{K}$ the number*

$$\mu_0 = \max\{\mu_0^{(1)}, \mu_0^{(2)}, \mu_0^{(3)}\}$$

and we call **amount of non-degeneracy** *of $\chi(y) = (\omega(y), A(y))$ with respect to $\mathcal{K}$ the greatest number characterized by*

$$\min_{y\in\mathcal{K}} \max_{0\leq\nu\leq\mu_0} \left| D^\nu [\chi]_k^{(j)}(y) \right| \geq \beta > 0, \ \forall\ k \in \mathbb{Z}^d, \ j = 1,2,3.$$

C.1.2 Estimate for $\exp(sAJ_{2p})$

In this subsection we display an estimate for the exponential $\exp(sA(y)J_{2p})$ where $s \in \mathbb{R}$ and A is a $2p \times 2p$ symmetric matrix which is real-analytic on a domain B in $\mathbb{R}^d$. Let $\mathcal{K}$ be any compact subset of B and choose $\vartheta > 0$ such that

$$\mathcal{K} + 4\vartheta \subseteq B.$$

We start stating two preliminary lemmata:

Lemma C.1.5. *Let $M \in \mathcal{H}(\mathcal{K}+2\vartheta)$ an $m \times m$ matrix function having eigenvalues $\lambda_1(y), \dots, \lambda_m(y)$. Suppose to have satisfied for a number $a \in \mathbb{R}$ the inequality Re $\lambda_j(y) < a$ for every $y \in \mathcal{K}$ and for all $j = 1, \dots, m$; then there exist numbers $a_\star < a$ and $t_\star \in (0, \vartheta)$ such that for $j = 1, \dots, m$*

$$\text{Re } \lambda_j(y) \leq a_\star \tag{C.1.15}$$

for all $y \in \mathcal{K} + t_\star$.

Proof The proof of this lemma can be found in [Rüßm01, page 166] $\square$

Lemma C.1.6. *Consider $M \in \mathcal{H}(\mathcal{K} + 2\vartheta)$ as in lemma C.1.5 and let $a_\star \neq 0$ (it is obviously always possible to suppose that) and $t_\star \in (0, \vartheta)$ such that Re $\lambda_j(y) \leq a_\star$ for all $y \in \mathcal{K} + t_\star$ and $j = 1, \dots, m$; then we have the following estimate*

$$|\exp(sM)|_{\mathcal{K}+t} \leq e^{sa_\star + s\beta|a_\star|} \left\{ 1 + \sum_{j=1}^{m-1} \left(\frac{2|M|_{\mathcal{K}+t}}{\beta|a_\star|} \right)^j (2\pi j)^{-\frac{1}{2}} \right\} \tag{C.1.16}$$

for every $s \geq 0$, $\beta > 0$ and $t \in (0, t_\star]$.

Proof The proof of this lemma can be found in [Rüßm01, pages 166-167] $\square$

Consider now the specific case of a $2p \times 2p$ matrix function $A \in \mathcal{H}(\mathcal{K} + 2\vartheta)$ (in the case we are going to analyze, A will be the matrix of the coefficients of the elliptic variables in the integrable part of the Hamiltonian considered) such that AJ_{2p} has all purely imaginary eigenvalues

$$\begin{aligned} &\lambda_1(y) = i\Omega_1(y), \dots, \lambda_p(y) = i\Omega_p(y), \\ &\lambda_{p+1}(y) = -i\Omega_1(y), \dots, \lambda_{2p}(y) = -i\Omega_p(y) \end{aligned}$$

where Ω_j are real-analytic functions for $j = 1, \dots, p$.

Proposition C.1.3. *Under the hypotheses just done on the matrix function A, there exist constants $a_\star = a_\star(A,\mathcal{K}) \in (0,1)$ and $t_\star = t_\star(A,\mathcal{K}) \in (0.\vartheta)$ so that for all $s \in \mathbb{R}$ we have*

$$|\exp(sAJ_{2p})|_{\mathcal{K}+t} \le e^{3a_\star|s|}\left\{1+\sum_{j=1}^{2p-1}\left(\frac{|A|_{\mathcal{K}+t}}{|a_\star|}\right)^j(2\pi j)^{-\frac{1}{2}}\right\} \qquad \text{(C.1.17)}$$

for every $t \in (0,t_\star]$.

Proof By the hypotheses made on A we can choose any positive number as an upper bound on $\mathcal{K}$ for the real parts of the $2p$ eigenvalues. For instance we can take $a = 1$ such that

$$\sup_{y\in\mathcal{K}} \text{ Re } \lambda_j(y) < a = 1\,, \quad \forall\, j = 1,\ldots,p.$$

Applying lemma C.1.5 we obtain the existence of $t_\star \in (0,\vartheta)$ and $a_\star$ such that

$$\sup_{y\in\mathcal{K}+t_\star} \text{ Re } \lambda_j(y) \le a_\star < 1\,, \quad \forall\, j = 1,\ldots,p. \qquad \text{(C.1.18)}$$

It is sufficient now to apply lemma C.1.16 with $M = \pm AJ_{2p}$, $m = 2p$, $\beta = 2$ and to observe that $a_\star > 0$ and $|AJ_{2p}|_{\mathcal{K}+t} \le |A|_{\mathcal{K}+t} = |A|_{\mathcal{K}+t}$ for any $t \le t_\star$ $\square$

C.2 The main theorem

Theorem C.2.1. *Let*

$$H(x,y,z) = h(y) + \frac{1}{2}\langle z, zA(y)\rangle + f(x,y,z)\otimes z^3 + P(x,y,z) \qquad \text{(C.2.1)}$$

be a real-analytic Hamiltonian defined for

$$(x,y,z) \in \mathbb{T}^d \times B \times V \subseteq \mathbb{T}^d \times \mathbb{R}^d \times \mathbb{R}^{2p}$$

where B is an open connected set of $\mathbb{R}^d$, V is in an open neighborhood of the origin in $\mathbb{R}^{2p}$ and A is a $2p \times 2p$ symmetric matrix. Let D a complex domain on which H can be holomorphically extended; let $\mathcal{K}$ be any non-empty compact subset of B with positive d-dimensional Lebesgue measure $\text{meas}_{\,d}\mathcal{K} > 0$ and let $\mathcal{A} \subseteq D$ be an open set such that

$$\mathbb{T}^d \times \mathcal{K} \times \{0\} \subseteq \mathcal{A}\ . \qquad \text{(C.2.2)}$$

Choose $\vartheta \in (0,1)$ such that

$$\mathbb{T}^d(\vartheta) \times (\mathcal{K}+4\vartheta) \times (\{0\}+\vartheta) \subseteq \mathcal{A} \qquad \text{(C.2.3)}$$

and

$$(\mathcal{K} +_{\mathbb{R}} 2\vartheta) \subseteq B \ . \tag{C.2.4}$$

Define $\omega(y) = \frac{\partial}{\partial y} h(y)$ *and take*

$$\begin{aligned} C_1 &= |\omega|_{\mathcal{K}+3\vartheta} && \text{(C.2.5)} \\ C_2 &= |A|_{\mathcal{K}+3\vartheta} && \text{(C.2.6)} \\ C_3 &= |f|_{\mathcal{A}} \ . && \text{(C.2.7)} \end{aligned}$$

Make the following two hypotheses:

1. *the symmetric matrix* $A(y)$ *is such that* $\det A \neq 0$ *and the eigenvalues of* $A(y)J_{2p}$ *are* $i\Omega_1(y) \ldots i\Omega_p(y), -i\Omega_1(y) \cdots - i\Omega_p(y)$ *with* $\Omega_1(y) \ldots \Omega_p(y)$ *real-analytic functions on* B *(this implies in particular that* $z = 0$ *is an elliptic equilibrium point for* Φ_H^t*).*

2. *the function*

$$(\omega, \Omega_1 \ldots \Omega_p) : B \subset \mathbb{R}^d \longrightarrow \mathbb{R}^d \times \mathbb{R}^p$$

   is weakly non-degenerate and extreme ((ω, Ω) *is usually called the frequency vector and his components are called respectively the tangential and normal frequencies).*

Then for any $\epsilon^\star$ *with* $0 < \epsilon^\star <$ *meas* ${}_d\mathcal{K}$ *there exist positive numbers* ϵ_0 *and* γ *(see subsection C.2.1 for details) depending on* $\omega, A, \mathcal{K}, \epsilon^\star, \mathcal{A}, \Phi$*, such that assuming*

$$|P|_{\mathcal{A}} \leq \frac{1}{2}\epsilon_0 \tag{C.2.8}$$

and taking real numbers σ_0, r_0, u_0 *verifying*

$$\begin{aligned} \sigma_0 &= \vartheta && \text{(C.2.9)} \\ r_0 &= \left(\frac{\epsilon_0\vartheta}{(C_1 + C_2 + C_3)}\right)^{\frac{1}{2}} && \text{(C.2.10)} \\ u_0 &= \min\left\{\left(\frac{\epsilon_0\vartheta}{C_1 + C_2 + C_3}\right)^{\frac{1}{4}}, \left(\frac{\epsilon_0}{2(C_1 + C_2 + C_3)}\right)^{\frac{1}{3}}\right\} && \text{(C.2.11)} \end{aligned}$$

there exists a compact subset $\mathcal{H} \subseteq \mathcal{K}$ *with meas* ${}_d\mathcal{H} >$ *meas* ${}_d\mathcal{K} - \epsilon^\star$ *and a continuous mapping*

$$X : (b, \xi, (\eta, \zeta)) \in \mathcal{H} \times \mathbb{T}^d \times \mathcal{U} \longrightarrow D$$

where $\mathcal{U}$ *is an open neighborhood of the origin in* $\mathbb{R}^d \times \mathbb{R}^{2p}$ *such that*

- *the mapping*

$$(\xi, \eta, \zeta) \longmapsto (x, y, z) = X(b, \xi, \eta, \zeta) \tag{C.2.12}$$

defines (for every $b \in \mathcal{H}$) an holomorphic canonical transformation on $\mathbb{T}^d\left(\frac{\sigma_0}{5}\right) \times \mathcal{U}_\rho$ and

$$X\left(\mathcal{H} \times \mathbb{T}^d\left(\frac{\sigma_0}{5}\right) \times \mathcal{U}_\rho\right) \subseteq \mathbb{T}^d_{\sigma_0} \times B_{r_0} \times V_{u_0}$$

for sufficiently small $\rho > 0$;

- *the transformed Hamiltonian is in the form:*

$$H\left(X(b,\xi,\eta,\zeta)\right) = h^\star(b)+\langle\omega^\star(b),\eta\rangle+\frac{1}{2}\langle\zeta,\zeta A^\star(b)\rangle+O(|\eta|^2+|\eta||\zeta|+|\zeta|^3) \tag{C.2.13}$$

  for every $b \in \mathcal{H}$ and $(\xi,\eta,\zeta) \in \mathbb{T}^d\left(\frac{\sigma_0}{5}\right) \times \mathcal{U}_\rho$;

- *the new frequency vector $\omega^\star$ and the new symmetric matrix $A^\star$ satisfy for all b in $\mathcal{H}$ the diophantine inequalities*

$$\begin{aligned}
|\langle k,\omega^\star(b)|\rangle &\geq \gamma\Phi(|k|),\ \ \forall\ k \in \mathbb{Z}^d \smallsetminus \{0\} \\
|\det\left(i\langle k,\omega^\star(b)\rangle I_{2p} + A^\star(b)J_{2p}\right)| &\geq \gamma\Phi(|k|),\ \ \forall\ k \in \mathbb{Z}^d \\
|\langle\det\left(i\langle k,\omega^\star(b)\rangle I_{2p^2+p} + [[A^\star(b)J_{2p}]]\right)| &\geq \gamma\Phi(|k|),\ \ \forall\ k \in \mathbb{Z}^d \smallsetminus \{0\}\ .
\end{aligned}$$

We conclude remarking that the transformed Hamiltonian system possesses the solutions

$$\begin{aligned}
\xi &= \omega^\star(b)t + c \\
\eta &= 0 = \zeta
\end{aligned} \tag{C.2.14}$$

so that the system described by H in (C.2.1) possesses the invariant torus

$$(x,y,z) = X(b,\xi,0,0)$$

for ξ in $\mathbb{T}^d$, with quasi-periodic flow (C.2.14) for all b in $\mathcal{H}$.

The statement of this theorem is greatly inspired by Rüßmann's theorem in [Rüßm01, page 126] but puts together different results (especially for what concerns the quantitative claims) contained in his work; we cite here, for references and clearness, the major results contained in Rüßmann's work on which theorem C.2.1 is based: theorem 1.7 on page 126, lemma 13.4 on page 158, theorem 15.5 on page 175 and theorem 16.7 on page 179.

The differences between Rüßmann's theorem and theorem C.2.1 (that consist mainly in the addition of a cubic term in the elliptic variables) will be discussed further on in this section. The strategy to prove theorem C.2.1 using Rüßmann's scheme is explained in subsection C.2.2.

C.2.1 Explicit estimate for ϵ_0 in Rüßmann's theorem

Now referring to the estimate Rüßmann gives in [Rüßm01, page 171], we are going to formulate an explicit estimate on the value of ϵ_0 (that is an estimate on the admissible size of the perturbation P in Hamiltonian (C.2.1)) under the hypotheses in theorem C.2.1. We first resume briefly, to be as much clearer as possible, all the quantities intervening in this estimate not mentioned in the statement theorem:

- $a_\star \in (0,1)$ and $t_\star \in (0,\vartheta)$ are two real numbers, whose existence is guaranteed by proposition C.1.3, such that for all $s \in \mathbb{R}$ we have

$$|\exp(sAJ_{2p})|_{\mathcal{K}+t} \leq e^{3a_\star|s|}\left\{1+\sum_{j=1}^{2p-1}\left(\frac{|A|_{\mathcal{K}+t}}{|a_\star|}\right)^j (2\pi j)^{-\frac{1}{2}}\right\} \tag{C.2.15}$$

  for every $t \in (0, t_\star]$;

- Let $\chi(y) = (\omega(y), A(y))$ (where $\omega = h'$, h and A are the real-analytic functions appearing in Hamiltonian (C.2.1)) then $\alpha(\chi,\mathcal{K})$ is the *amount of degeneracy* of χ with respect to $\mathcal{K}$ as in definition C.1.2);

- $\mu_0(\chi,\mathcal{K})$ and $\beta(\chi,\mathcal{K})$ are respectively the *index of non-degeneracy* and the *amount of non-degeneracy* of χ with respect to $\mathcal{K}$ defined in proposition C.1.2 (and whose determination can be found along the procedure in section C.1.1);

- We recall that $\epsilon^\star$ is an arbitrarily chosen number such that $0 < \epsilon^\star < \operatorname{meas}_d\mathcal{K}$, $\vartheta < 1$ is a positive number small enough to verify inclusion (C.2.3) while $C_1 = |\omega|_{\mathcal{K}+3\vartheta}, C_2 = |\Omega|_{\mathcal{K}+3\vartheta}$ and $C_3 = |f|_{\mathcal{A}}$. In addiction to that we define d_0 the diameter of $\mathcal{K}$:

$$d_0 = \sup_{x,y\in\mathcal{K}} |x-y| > 0.$$

We now define some numerical constants obtained by simplifying the values indicated by Rüßmann in the cited estimate (values that we consider as given upper bounds) and some other quantities needed for the wanted estimate:

- Let Φ be the chosen approximation function in agreement with definition C.1.1, take $T_0 \geq 1$ such that

$$\Phi(T_0) \leq e^{-r} \quad \text{for} \quad r := \max\{d+1, 2p^2+p\} \tag{C.2.16}$$

  and the following inequality is verified

$$\int_{T_0}^{\infty} \log\frac{1}{\Phi(T)}\frac{dT}{T^2} \leq \frac{\vartheta\log 2}{12\mu_0 d\log(3^{25}d)}; \tag{C.2.17}$$

- Let Φ be the chosen approximation function we put

$$\Phi_{d\mu_0} = \sup_{s \geq 1} s^d \Phi(s)^{\frac{1}{\mu_0}} < \infty;$$

- Define

$$M^\star = (2p^2+p)^{\mu_0+1} 2^{(2p^2+p)(\mu_0+5)} \left[\mu_0! \vartheta^{-\mu_0} (C_1 + C_2 + 1)\right]^{4p^2+2p-1}$$

and

$$C^\star = (2p^2+p)^{\mu_0+1} 2^{(2p^2+p)(\mu_0+4)} \frac{(\mu_0+1)^{\mu_0+2}}{(\mu_0+1)!} (\Phi_{d\mu_0})^2$$
$$* \left[\frac{(\mu_0+1)!}{\vartheta^{\mu_0+1}} (C_1 + C_2 + 1)\right]^{4p^2+2p}$$

- We can now define the value of γ appearing in theorem C.2.1, which obviously gives contribution to the estimate for ϵ_0: let

$$\gamma_1 = \left[3^{d+3} (2\pi e)^{\frac{d}{2}} d_0^{d-1} \left(d^{-\frac{1}{2}} + 2d_0 + \vartheta^{-1} d_0\right) C^\star\right]^{-\frac{\mu_0}{2}} \alpha^{2p^2(\mu_0+1)} \beta^{\frac{\mu_0+1}{2}} \epsilon^{\star\,\frac{\mu_0}{2}}$$

then

$$\gamma = \min\{2a_\star, \gamma_1\}$$

- Define $t_0 > 0$ by

$$t_0 = \min\left\{t_\star\,,\ \frac{\gamma T_0^{-r} \Phi(T_0) \vartheta}{8(2p^2+p)(3(C_1+C_2+1))^{2p^2+p}}\right\}; \tag{C.2.18}$$

- Let $\bar{g} : \mathbb{R}^d \longrightarrow \mathbb{R}$ such that

$$\bar{g}(x) = \begin{cases} e^{-\frac{1}{1-|x|^2}} & |x| < 1 \\ 0 & |x| \geq 1 \end{cases}$$

and consider the function

$$g(s) = \bar{g}(s) \left(\int_{\mathbb{R}^d} \bar{g}(t)\, dt\right)^{-1},$$

then define

$$C(d,n) = 3^n n! \sup_{a \in \mathbb{R}^d,\, |a|_2 = 1} \int_{\mathbb{R}^d} \sum_{j=0}^{n} \frac{1}{j!} |D^j g(s)(a^j)|\, ds$$

taking $C(d,0) = 1$;

- Define at last

$$a_1 = 1 + \sum_{j=1}^{2p-1} \left(a_\star^{-1} C_2\right)^j (2\pi j)^{-\frac{1}{2}} \tag{C.2.19}$$

$$a_2 = 3^d (2p+1)^2 (7(C_1 + C_2 + 1))^{2p^2+p-1}; \tag{C.2.20}$$

then the estimate on ϵ_0 is given by

$$\epsilon_0 = \frac{\vartheta}{C_1 + C_2 + C_3} \left(\min\{E_1, E_2, E_3\}\right)^2 \tag{C.2.21}$$

where

$$\begin{aligned} E_1 &= \frac{\gamma \Phi(T_0) T_0^{-r}}{3^{27} 2(a_1 + a_2)} \min\left\{\vartheta\,,\ \frac{1}{3^{17}}\right\} \\ E_2 &= \frac{\gamma \Phi(T_0) T_0^{-r}}{3^{22}(2p^2+p)\,(3(C_1+C_2+1))^{2p^2+p-1}} \left(1 - \frac{1}{2^{\mu_0}}\right)\left(1 - \frac{1}{2^{5-\frac{2}{\mu_0}}}\right) \\ E_3 &= \frac{1}{3^{19} T_0} \min\left\{\frac{a_\star}{4a_1}\,,\ \frac{\alpha^{4p^2} \beta {t_0}^{\mu_0}}{2M^\star C(d,\mu_0)}\right\} \end{aligned}$$

C.2.2 Brief explanation of the strategy adopted by Rüßmann

We are going to dedicate the remaining part of this chapter to the proof of theorem C.2.1. Actually, we will entirely use Rüßmann's proof contained in [Rüßm01] and perform just a few preliminary steps to prove the slighlty different version his theorem given in C.2.1. The first step consists in listing the conditions indicated by Rüßmann in [Rüßm01, page 157], under which the n-th step of the iteration process, adopted to prove his main theorem of, can be carried out.

We observe at first that the scheme adopted by Rüßmann conjugates a real-analytic Hamiltonian

$$H(x,y,z) = h(y) + \frac{1}{2}\langle z, zA(y)\rangle + P(x,y,z) \tag{C.2.22}$$

to an Hamiltonian in the form (C.2.13), that is

$$H\left(b,\xi,\eta,\zeta\right) = h^\star(b) + \langle \omega^\star(b), \eta\rangle + \frac{1}{2}\langle \zeta, \zeta A^\star(b)\rangle + O(|\eta|^2 + |\eta||\zeta| + |\zeta|^3)$$

(refer to the statement in [Rüßm01, page 126] without considering the initial and final hyperbolic variables respectively w and θ). It can be easily observed that the

Hamiltonian that entries Rüßmann's scheme is slightly different from the Hamiltonian function (C.2.1) considered in the main theorem of this section. In fact, this last possesses in addition, as already remarked, a cubic part in the elliptic variables

$$f(x, y, z) \otimes z^3 .$$

The second step of our proof is therefore dedicated to manipulate Hamiltonian (C.2.1) in order to put it in the form

$$H_0(b, x, y, z) = N_0(b, y, z) + P_0(b, x, y, z) \tag{C.2.23}$$

with

$$N_0(b, y, z) = h_0(b) + \langle \omega_0(b), y \rangle + \frac{1}{2}\langle z, zA_0(y) \rangle$$

which is the form required by Rüßmann to get into the initial step of the iteration process. In fact, in a general n-th step ($n \geq 0$) we should have an Hamiltonian

$$H_n(b, x, y, z) = N_n(b, y, z) + P_n(b, x, y, z) ,$$

where the normal part is given by

$$N_n(b, y, z) = h_n(b) + \langle \omega_n(b), y \rangle + \frac{1}{2}\langle z, zA_n(y) \rangle ,$$

depending analytically on x, y and z and, additionally, on some parameters b varying in an open and bounded set $\mathcal{B}_n \subset \mathbb{R}^d$ (as it will be later described, thay are substantially the action variables whose frequencies are non-resonant and can be controlled by means of the chosen approximation function up to a certain order).

The last step consists in applying Rüßmann's scheme to the Hamiltonian function (C.2.23) verifying that all the conditions listed during our first step hold for $n = 0$. The proof of theorem C.2.1 will follow by consequence of Rüßmann's theorem.

C.2.3 Conditions to carry out the n-th step of the iteration process

We want now to list, referring to [Rüßm01, page 157], the conditions under which the n-th step of the iteration process can work out. We first need to set some definitions:

- Choose and fix

$$\begin{aligned} &0 < \tau \leq \tau_0 = \tfrac{1}{9}, \quad \kappa = \tfrac{24}{25}, \quad \delta = \left(\tfrac{\kappa}{2}\right)^{\frac{1}{1-\kappa}} = \left(\tfrac{12}{25}\right)^{25}, \\ &\bar{\sigma} = \tfrac{2}{3}, \qquad\qquad \sigma = \tfrac{3}{4}, \quad \varphi = \tfrac{3}{8} \end{aligned} \tag{C.2.24}$$

- Let Φ be the chosen approximation function and T_0 the real number verifying conditions (C.2.16) and (C.2.17), we put

$$\Psi(T) := T^{-r}\Phi(T) \ \text{ with } \ r = \max\{d+1, 2p^2+p\}$$

and

$$T_n := \Psi^{-1}\left(\Psi(T_0)\delta^{\tau n}\right) \tag{C.2.25}$$

for τ and δ as previously defined.

- $\mathcal{K}_n$ is defined recursively as follows:

$$\mathcal{K}_0 := \mathcal{K}$$

is an arbitrarily chosen non-empty compact subset of B (the domain of the action variables) as it appears in the statement of theorem C.2.1 and

$$\mathcal{K}_{n+1} := \mathcal{K}^{(1)}_{n+1} \cap \mathcal{K}^{(2)}_{n+1} \cap \mathcal{K}^{(3)}_{n+1}$$

with

$$\begin{aligned}
\mathcal{K}^{(1)}_{n+1} \ &:= \ \{b \in \mathcal{K}_n \ : \ |\langle k, \omega_n(b)\rangle| \geq 2^{-1}\Phi(T_n),\ k \in \mathbb{Z}^d,\ 0 < |k|_2 \leq T_n\}\,, \\
\mathcal{K}^{(2)}_{n+1} \ &:= \ \{b \in \mathcal{K}_n \ : \ |\det(i\langle k, \omega_n(b)\rangle \mathbb{I}_{2p} + A_n(b)J_{2p})| \geq 2^{-1}\Phi(T_n)\,, \\
&\qquad k \in \mathbb{Z}^d,\ |k|_2 \leq T_n\}, \\
\mathcal{K}^{(3)}_{n+1} \ &:= \ \{b \in \mathcal{K}_n \ : \ |\det(i\langle k, \omega_n(b)\rangle \mathbb{I}_{2p^2+p} + [[A_n(b)J_{2p}]])| \geq 2^{-1}\Phi(T_n), \\
&\qquad k \in \mathbb{Z}^d,\ 0 < |k|_2 \leq T_n\}\,.
\end{aligned}$$

- We put

$$\epsilon_n := \epsilon_0 \delta^{\tau n} \tag{C.2.26}$$

for ϵ_0 verifying (C.2.8) (and being estimated by (C.2.21)) and

$$t_n := t_0 \delta^{\tau n} \tag{C.2.27}$$

for t_0 defined by equation (C.2.18).

- Let $\mathcal{B} \subseteq \mathbb{C}^d$ be an open set and let $T > 0$. We define $P(\mathcal{B}, T)$ as the set of all trigonometric-algebraic polynomials F, depending on $(x, y, z) \in \mathbb{C}^d \times \mathbb{C}^d \times \mathbb{C}^{2p}$ and on parameters $b \in \mathcal{B}$, having the form

$$F(b, x, y, z) = \sum_{(k,j,l) \in I} F_{kjl}(b) e^{i\langle k, x\rangle} y^j z^l \tag{C.2.28}$$

  with

$$I = \left\{(k, j, l) \in \mathbb{Z}^d \times \mathbb{N}^d \times \mathbb{N}^{2p} \,:\, |k|_2 \leq T,\, 2|j|_1 + |l|_1 \leq 2\right\};$$

  and where the coefficients $F_{jkl} : \mathcal{B} \to \mathbb{E}$, for $\mathbb{E} = \mathbb{C}^d$ or $\mathbb{C}^{2p\times 2p}$ are holomorphic (so that $P(\mathcal{B}, T) \subset \mathcal{H}(\mathcal{B} \times \mathbb{C}^d \times \mathbb{C}^d \times \mathbb{C}^{2p})$).

The conditions to carry out the n-step of the iteration process are:

1. **[Condition on $\mathcal{K}_n$]** The compact set $\mathcal{K}_n \subset \mathbb{R}^d$ is such that $\mathcal{K}_n \neq \emptyset$ (recall that $\mathcal{K}_0 = \mathcal{K}$ by definition).

2. **[Conditions on N_n]** The normal part N_n of the Hamiltonian at the n-th step belongs to $P(\mathcal{B}_n, 0)$ with $\mathcal{B}_n = \mathcal{K}_n + t_n \subseteq \mathbb{C}^d$ (see definitions given above for $P(\mathcal{B}, T)$ and t_n) and has the form

$$N_n(b, y, z) = h_n(b) + \langle \omega_n(b), y\rangle + \frac{1}{2}\langle z, zA_n(y)\rangle \tag{C.2.29}$$

   with $A_n = A_n^t$; we remark that by definition of $P(\mathcal{B}, T)$ given above we have $h_n, \omega_n, A_n \in \mathcal{H}(\mathcal{B})$.

3. **[Condition on A_0]** There exist constants $c \in (0, 1)$ and $c_1 \geq 1$ such that the $2p \times 2p$ symmetric matrix function $A_0 \in \mathcal{H}(\mathcal{B}_0)$ satisfies

$$\left|e^{\lambda A_0 J_{2p}}\right|_{\mathcal{B}_0} \leq c_1 e^{3c|\lambda|} \tag{C.2.30}$$

   for all $\lambda \in \mathbb{R}$.

4. **[Conditions on $|\omega_n - \omega_0|$ and $|A_n - A_0|$]** We require the following two inequalities to be satisfied:

$$|\omega_n - \omega_0|_{\mathcal{B}_n} \leq \frac{3\epsilon_0}{2r_0}\delta^{-\bar{\sigma}}\frac{1 - \delta^{(\kappa-\sigma)n}}{1 - \delta^{\kappa-\sigma}} \tag{C.2.31}$$

$$|A_n - A_0|_{\mathcal{B}_n} \leq 3T_0\frac{\epsilon_0}{r_0}\delta^{-\bar{\sigma}}\frac{1 - \delta^{(\kappa-\sigma)n}}{1 - \delta^{\kappa-\sigma}} \tag{C.2.32}$$

   where ϵ_0 and r_0 appear in theorem C.2.1 and are respectively defined by (C.2.8) and (C.2.10), δ, κ, σ and $\bar{\sigma}$ are fixed according to (C.2.24) and T_0 verifies conditions (C.2.16) and (C.2.17).

5. **[Condition on the derivatives of ω_n and A_n]** Let C_1 and C_2 be the constants in (C.2.5) and (C.2.6), let $\vartheta \in (0,1)$ be small enough to verify inclusion (C.2.3) and let t_0 be defined by (C.2.18), then the following inequality must be verified:

$$T_0|D\omega_n|_{\mathcal{B}_n} + |DA_n|_{\mathcal{B}_n} \leq T_0 \left(\frac{C_1 + C_2}{\vartheta} + \frac{4\epsilon_0}{r_0 t_0} \delta^{-\bar{\sigma}} \frac{1 - \delta^{(\kappa-\sigma-\tau)n}}{(1-\delta^\tau)(1-\delta^{\kappa-\sigma-\tau})} \right)$$

6. **[Condition on P_n]** The perturbative part P_n of the Hamiltonian resulting after n steps of the iteration process is an holomorphic function for $(b,x,y,z) \in \mathcal{B}_n \times \mathcal{D}_n$, where $\mathcal{D}_n = \mathcal{D}(\sigma_n, r_n, u_n)$ (see (C.1.1) for the definition of $\mathcal{D}(s)$) with

$$\begin{aligned} \sigma_n &:= \sigma_0 - 2d \log \frac{4d}{\delta} \sum_{k=0}^{n-1} \frac{1}{T_k} \\ r_n &:= r_0 \delta^{\sigma n} \\ u_n &:= u_0 \delta^{\varphi n} \end{aligned}$$

(σ_0, r_0 and u_0 are defined in (C.2.9), (C.2.10) and (C.2.11), δ, σ and φ in (C.2.24) and T_n in (C.2.25)). Besides P_n has to be 2π-periodic with respect to the angle-variables that is

$$P_n(b,x,y,z) = P_n(b, x + 2k\pi, y, z), \;\; \forall \; (b,x,y,z) \in \mathcal{B}_n \times \mathcal{D}_n, \;\; \forall \; k \in \mathbb{Z}^d$$

and has to verify

$$|P_n|_{\mathcal{B}_n \times \mathcal{D}_n} \leq \epsilon_n \tag{C.2.33}$$

for ϵ_n defined in (C.2.26).

7. **[Conditions on H_n]** The Hamiltonian function at the n-th step of the iteration has the form

$$H_n(b,x,y,z) = N_n(b,y,z) + P_n(b,x,y,z). \tag{C.2.34}$$

with N_n and P_n verifying all conditions previously listed.

C.2.4 Survey of the conditions to carry out the n-th step of the iteration process for $n = 0$

In this subsection we show that all conditions from C.2.3.1 to C.2.3.7 can be verified for $n = 0$ under the hypotheses of theorem C.2.1. Recall that the considered

Hamiltonian

$$H(x,y,z) = h(y) + \frac{1}{2}\langle z, zA(y)\rangle + f(x,y,z)\otimes z^3 + P(x,y,z) \qquad \text{(C.2.35)}$$

is real-analytic for $(x,y,z) \in \mathbb{T}^d \times B \times U$ (where B is a domain in $\mathbb{R}^d$ and U is ball around the origin in $\mathbb{R}^{2p}$), $\mathcal{K}$ is a non-empty compact subset of B, arbitrarily chosen, with positive d-dimensional Lebesgue measure and $\mathcal{A}$ is an open subset of D such that

$$\mathbb{T}^d(\vartheta) \times (\mathcal{K} + 4\vartheta) \times (\{0\} + \vartheta) \subseteq \mathcal{A}$$

for sufficiently small $\vartheta \in (0,1)$.

Since it is more comfortable to work with complex neighborhoods of $\mathbb{R}^d$ instead that with complex domains containing the torus $\mathbb{T}^d$, we shall consider from now on H as a function defined on $\mathbb{R}^d \times B \times U$ and being 2π-periodic in the x variables. More precisely this means that f and P shall be considered as real-analytic functions on $\mathbb{R}^d \times B \times U$ satisfying

$$f(x,y,z) = f(x+2k\pi, y, z)\,,\ P(x,y,z) = P(x+2k\pi,y,z)\ \ \forall\ k \in \mathbb{Z}^d.$$

Furthermore, D has to be considered as a complex domain in $\mathbb{C}^{2d+2p}$ containing $\mathbb{R}^d \times B \times U$ and $\mathcal{A}$ as an open set such that

$$(\mathbb{R}^d + \vartheta) \times (\mathcal{K} + 4\vartheta) \times (\{0\} + \vartheta) \subseteq \mathcal{A} \subseteq D \qquad \text{(C.2.36)}$$

and

$$(x,y,z) \in \mathcal{A} \Leftrightarrow (x+2k\pi, y, z) \in \mathcal{A}\ \ \forall\ k \in \mathbb{Z}^d$$

After this simple agreement we are ready to verify the above mentioned conditions. We start observing that condition C.2.3.1 is verified for $n = 0$ by mere definition of

$$\mathcal{K}_0 = \mathcal{K} \neq \emptyset$$

and condition C.2.3.4 trivially holds for $n = 0$.

Now, let $b \in \mathcal{K} + 3\vartheta$, then for every $(x,y,z) \in \mathcal{D}(\vartheta,\vartheta,\vartheta)$ (see definition (C.1.1)) we may write

$$\begin{aligned}
& H_0(b,x,y,z) = H(x,b+y,z) = \\
= \ & h(b+y) + \frac{1}{2}\langle z, zA(b+y)\rangle + f(x,b+y,z)\otimes z^3 + P(x,b+y,z) = \\
= \ & h(b) + \langle \frac{\partial h(b)}{\partial y}, y\rangle + \frac{1}{2}\langle z, zA(b)\rangle + h(b+y) - h(b) - \langle \frac{\partial h(b)}{\partial y}, y\rangle \\
+ \ & \frac{1}{2}\langle z, z(A(b+y) - A(b))\rangle + f(x,b+y,z)\otimes z^3 + P(x,b+y,z) = \\
= \ & N_0(b,y,z) + P_0(b,x,y,z)
\end{aligned}$$

according to (C.2.34) and (C.2.29) with $n = 0$, having established the correspondences

$$h = h_0\,,\ \ \frac{\partial h}{\partial y} = \omega_0\,,\ \ A = A_0$$

and putting

$$\begin{aligned} P_0(b,x,y,z) &= h(b+y) - h(b) - \langle \frac{\partial h(b)}{\partial y}, y\rangle + \frac{1}{2}\langle z, z(A(b+y) - A(b))\rangle \\ &+ f(x, b+y, z) \otimes z^3 + P(x, b+y, z). \end{aligned} \tag{C.2.37}$$

As it can be immediately seen $N_0 \in P(\mathcal{B}_0, 0)$ since h and A belong to $\mathcal{H}(\mathcal{K} + 4\vartheta)$ (as inclusion (C.2.36) shows) and

$$\mathcal{B}_0 = \mathcal{K}_0 + t_0 = \mathcal{K} + t_0 \subseteq \mathcal{K} + t_\star \subseteq \mathcal{K} + \vartheta$$

by definition of t_0 in (C.2.18) and $t_\star$ in proposition (C.1.3). This proves condition C.2.3.2 for $n = 0$.

Now $A_0 = A = A_0^t$ verifies the hypothesis in theorem C.2.1 concerning its eigenvalues; therefore we may apply proposition C.1.3, and in particular inequality (C.1.17), in order to have condition C.2.3.3 satisfied with

$$\begin{aligned} c &= a_\star \\ c_1 &= 1 + \sum_{j=1}^{2p-1} \left(\frac{|A|_{\mathcal{B}_0}}{|a_\star|} \right)^j (2\pi j)^{-\frac{1}{2}} \end{aligned}$$

(where we used once again $t_0 \le t_\star$).

To prove the effectiveness of condition C.2.3.5 in the case $n = 0$ we claim that

$$\begin{aligned} & T_0|D\omega_0|_{\mathcal{B}_0} + |DA_0|_{\mathcal{B}_0} \le T_0\left(|D\omega_0|_{\mathcal{K}_0+t_0} + |DA_0|_{\mathcal{K}_0+t_0}\right) \le \\ \le\ & T_0\left(|D\omega_0|_{\mathcal{K}+2\vartheta} + |DA_0|_{\mathcal{K}+2\vartheta}\right) \le \frac{T_0}{\vartheta}\left(|\omega_0|_{\mathcal{K}+3\vartheta} + |A_0|_{\mathcal{K}+3\vartheta}\right) \le \\ \le\ & T_0\left(\frac{C_1 + C_2}{\vartheta}\right) \end{aligned}$$

having used $T_0 \ge 1$, $t_0 \le t_\star \le \vartheta$, Cauchy's estimate (with a loss of analycity ϑ), and definitions of C_1 and C_2 in (C.2.5) and (C.2.6).

To conclude, we need to prove condition C.2.3.6 (since C.2.3.7 runs automatically by consequence of all other conditions). Before estimating $|P_0|_{\mathcal{B}_0\times\mathcal{D}_0}$

(accordingly with notations in C.2.3.6 in the case $n = 0$) we shall prove the following

Lemma C.2.1. *Given ϵ_0 as determined by (C.2.21), C_1,C_2 and C_3 as defined in (C.2.5), (C.2.6) and (C.2.7) and $\vartheta \in (0,1)$, the following relation holds:*

$$\epsilon_0 \leq (C_1 + C_2 + C_3)\vartheta^3. \tag{C.2.38}$$

Proof From the estimate on ϵ_0 in (C.2.21) we have

$$\epsilon_0 \leq \frac{\vartheta}{C_1 + C_2 + C_3} E_1^2 \tag{C.2.39}$$

where we recall that E_1 is given by (see (C.2.40) and the whole section C.2.1 for notations)

$$E_1 = \frac{\gamma \Phi(T_0) T_0^{-r}}{3^{27} 2(a_1 + a_2)} \min \left\{ \vartheta \, , \, \frac{1}{3^{17}} \right\}. \tag{C.2.40}$$

Now, theorem 2.5.1 assures the existence of some $b \in \mathcal{K}_0 \subset \mathcal{B}_0$ that verifies

$$|\langle k, \omega_0(b) \rangle| \geq \frac{\gamma}{2} \Phi(T_0)$$

for every $k \in \mathbb{Z}^d$ such that $|k|_2 \leq T_0$ (that is equivalent to infer $\mathcal{K}_1 \neq \emptyset$); so we have

$$\frac{\gamma}{2} \Phi(T_0) \leq |\langle k, \omega_0(b) \rangle| \leq T_0 |\omega|_{\mathcal{B}_0} \leq T_0 (C_1 + C_2 + C_3).$$

Substituting this last inequality in (C.2.40) we obtain

$$E_1 \leq \frac{T_0^{-r+1}}{3^{27}(a_1 + a_2)} (C_1 + C_2 + C_3) \vartheta;$$

since $a_1, a_2 \geq 1$ (see definitions (C.2.19) and (C.2.20)), $T_0 \geq 1$ and $r = \max\{d + 1, 2p^2 + p\} \geq 1$, we have

$$E_1 \leq (C_1 + C_2 + C_3) \vartheta.$$

With inequality (C.2.39) we get the statement in (C.2.38) $\square$

With the result in (C.2.38) and the definitions of σ_0, r_0 and u_0 in (C.2.9), (C.2.10) and (C.2.11), we obtain

$$\sigma_0, r_0, u_0 \leq \vartheta$$

and hence the following inclusion

$$\mathcal{D}_0 := \mathcal{D}(\sigma_0, r_0, u_0) \subseteq \mathcal{D}(\vartheta, \vartheta, \vartheta). \tag{C.2.41}$$

Now, coming back to the definition of P_0 in (C.2.37) we use Taylor's formula for the integral remaining of an analytic series to obtain (denoting as usual $\omega = \frac{\partial h}{\partial y}$)

$$\begin{aligned} P_0(b,x,y,z) &= \int_0^1 (1-t)\langle D\omega(b+ty)(y), y\rangle + \frac{1}{2}\langle z, zA(b+ty)(y)\rangle \; dt + \\ &+ \; f(x, b+y, z) \otimes z^3 + P(x, b+y, z). \end{aligned}$$

In view of this equality, inclusions (C.2.41) and (C.2.36) and recalling that $\mathcal{B}_0 = \mathcal{K} + t_0 \subseteq \mathcal{K} + \vartheta$, we have the following estimate:

$$\begin{aligned} |P_0|_{\mathcal{B}_0 \times \mathcal{D}_0} &\leq \sup_{\mathcal{D}_0} \left(\frac{1}{2}|D\omega|_{\mathcal{K}+2\vartheta}|y|^2 + \frac{1}{2}|A|_{\mathcal{K}+2\vartheta}|z|^2|y| + |f|_{\mathcal{A}}|z|^3 \right) + |P|_{\mathcal{A}} \leq \\ &\leq \frac{1}{2}|\omega|_{\mathcal{K}+3\vartheta} r_0^2 \vartheta^{-1} + \frac{1}{2}|A|_{\mathcal{K}+3\vartheta} u_0^2 r_0 \vartheta^{-1} + |f|_{\mathcal{A}} u_0^3 + |P|_{\mathcal{A}} \; . \end{aligned}$$

Then, in view of the definitions of σ_0, r_0 and u_0 given in theorem C.2.1 (namely (C.2.9), (C.2.10) and (C.2.11)), inequality (C.2.8) and definitions (C.2.5), (C.2.6) and (C.2.7), we get

$$|P_0|_{\mathcal{B}_0 \times \mathcal{D}_0} \leq \frac{1}{2} C_1 r_0^2 \vartheta^{-1} + \frac{1}{2} C_2 u_0^2 r_0 \vartheta^{-1} + C_3 u_0^3 + |P|_{\mathcal{A}} \leq \frac{\epsilon_0}{2} + |P|_{\mathcal{A}} \leq \epsilon_0$$

that is condition C.2.3.6 for $n = 0$.

With this last result we have proved that all conditions listed in subsection C.2.3, which are needed for the iteration process to work, hold for $n = 0$, under hypotheses in theorem C.2.1. This allows to enter Rüßmann's iterative scheme (described in particular in [Rüßm01, pages 156-164]) with the Hamiltonian function $H_0(b,x,y,z) = N_0(b,x,y,z) + P_0(b,x,y,z) = H(x, b+y, z)$, obtaining the proof of theorem C.2.1 as a consequence.

Bibliography

[Arn63a] V. I. Arnold. Proof of a theorem by A. N. Kolmogorov on the invariance of quasi-periodic motions under small perturbations of the Hamiltonian. *Russian Math. Survey*, 18:13–40, 1963.

[Arn63b] V.I. Arnold. Small Denominators and Problems of Stability of Motion in Classical and Celestial Mechanics. *Usephi Mat. Nauk*, 18(6 (114)):91–192, 1963.

[Chi97] L. Chierchia. *Lezioni di Analisi Matematica 2*. Aracne, 1997.

[Chi05a] A. Celletti, L. Chierchia. KAM Tori for N-Body Problems. *Kluwer Academic Publisher*, 2005.

[Chi05b] L. Chierchia. Il Problema dei 3 Corpi Ristretto Circolare Piano, 2004-2005. Lecture notes from the course AM8. Università degli Studi Roma Tre.

[Chi06] L. Chierchia. KAM theory and Celestial Mechanics (2005). *Encyclopedia of Mathematical Physics*, 3, 2006. J.P. Françoise, G.L. Naber and Tsou S.T. Oxford.

[Féj04] J. Féjoz. Démonstration du "Théorèm d'Arnold" sur la Stabilité du Système Planetaire (d'après M. Herman). *Ergod. Th. & Dynam. Sys*, 24:1–62, 2004.

[Ham74] R.S. Hamilton. The inverse function theorem of Nash and Moser, 1974. Manuscript, Cornell University.

[Her98] M.R. Herman. Démonstration d'un théorème de V.I. Arnold. *Séminaire de Systèmes Dynamiques et manuscripts*, 1998.

[Kol54] A.N. Kolmogorov. On the Conservation of Conditionally Periodic Motions under Small Perturbation of the Hamiltonian. *Dokl. Akad. Nauk SSR*, 98:527–530, 1954.

[Las91] J. Laskar. Analytical framework in Poincaré variables for the motion of the solar system. *Predictability, Stability and Chaos in n-Body Dynamical Systems,* pages 93–114, 1991.

[Poi07] H. Poincaré. Leçons de Mécanique Céleste, 1905-1907. Gauthier-Villars, Paris.

[Pös82] J. Pöschel. Integrability of Hamiltonian Systems on Cantor Sets. *Communication on Pure and Applied Mathematics*, 35:653–695, 1982. John Wiley and Sons Inc.

[Pya69] A.S. Pyartli. Diophantine approximations on Euclidean submanifolds. *Functional. Anal. Appl.*, 3:303–306, 1969. (in Russian) Funckcional. Anal. i Prilozen, volume 3, pages 59-62.

[Rob95] J. Laskar, P. Robuteil. Stability of the Planetary Three-body Problem. *Celestial Mech. Dynam. Astronom.*, 62:193–217, 1995. (I. Expansion of the planetary Hamiltonian).

[Rüßm01] H. Rüßmann. Invariant Tori in Non-Degenerate Nearly Integrable Hamiltonian Systems. *R. & C. Dynamics*, 2(6):119–203, March 2001.

[Val03] L. Biasco, L. Chierchia & E. Valdinoci. Elliptic Two-Dimensional Invariant Tori for the Planetary Three-Body Problem. *Arch. Rational Mech. Anal.*, 170:91–135, March 2003.

[Whi34] H. Whitney. Analytic Extensions of Differentiable Functions Defined in Closed Sets. *Transaction A.M.S.,* 36:63–89, 1934.

[Zeh94] H. Hofer, E. Zehnder. *Symplectic Invariants and Hamiltonian Dynamics*. Birkhäuser Verlag, 1994.