

## A KAM theorem without action-angle variables for elliptic lower dimensional tori

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

2011 Nonlinearity 24 1033

(<http://iopscience.iop.org/0951-7715/24/4/003>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 147.83.133.63

The article was downloaded on 07/05/2013 at 12:16

Please note that [terms and conditions apply](#).

# A KAM theorem without action-angle variables for elliptic lower dimensional tori

Alejandro Luque and Jordi Villanueva

Departament de Matemàtica Aplicada I, Universitat Politècnica de Catalunya, Diagonal 647,  
08028 Barcelona, Spain

E-mail: [alejandro.luque@upc.edu](mailto:alejandro.luque@upc.edu) and [jordi.villanueva@upc.edu](mailto:jordi.villanueva@upc.edu)

Received 15 February 2010, in final form 25 January 2011

Published 17 February 2011

Online at [stacks.iop.org/Non/24/1033](http://stacks.iop.org/Non/24/1033)

Recommended by D V Treschev

## Abstract

We study elliptic lower dimensional invariant tori of Hamiltonian systems via parametrizations. The method is based on solving iteratively the functional equations that stand for invariance and reducibility. In contrast with classical methods, we do not assume that the system is close to an integrable one nor that it is written in action-angle variables. We only require an approximation of an invariant torus with a fixed vector of basic frequencies and a basis along the torus that approximately reduces the normal variational equations to constant coefficients. We want to highlight that this approach presents many advantages compared with methods which are built in terms of canonical transformations, e.g., it produces simpler and more constructive proofs that lead to more efficient numerical algorithms for the computation of these objects. Such numerical algorithms are suitable to be adapted in order to perform computer assisted proofs.

Mathematics Subject Classification: 37J40

## 1. Introduction

Persistence of quasi-periodic solutions has been a subject of remarkable importance in dynamical systems for a long time. Roughly speaking, KAM theory—named after Kolmogorov [42], Arnold [1] and Moser [50]—deals with the effect of small perturbations on dynamical systems (typically Hamiltonian) which admit invariant tori carrying quasi-periodic motion. Nowadays, KAM theory is a vast area of research that involves a large collection of methods and applications to a wide set of contexts: Hamiltonian systems, reversible systems, volume-preserving systems, symplectic maps, PDEs and lattices, just to mention a few. We refer to [2, 4, 10, 16, 60] for different surveys or tutorials that collect many aspects of the theory and cover a large amount of bibliography.

In this work we are concerned with lower dimensional (isotropic) tori of Hamiltonian systems. Let us consider a real analytic Hamiltonian system with  $n$  degrees of freedom having an invariant torus of dimension  $r < n$ , carrying quasi-periodic dynamics with a vector of basic frequencies  $\omega \in \mathbb{R}^r$ . The variational equations around such a torus correspond to a  $2n$ -dimensional linear quasi-periodic system with a vector of frequencies  $\omega$ . For this linear system we have  $2r$  trivial directions (i.e. zero eigenvalues of the reduced matrix of the system restricted to these directions) associated with the tangent directions of the torus and the symplectic conjugate ones (these trivial directions are usually referred to as the central directions of the torus). If the remaining  $2(n - r)$  directions (normal directions of the torus) are hyperbolic, we say that the torus is hyperbolic<sup>1</sup> or whiskered. Hyperbolic tori are very robust under perturbations [19, 24, 27, 32, 45]. For example, if we consider a perturbation of the system depending analytically on external parameters, one can establish, under suitable conditions, the existence of an analytic (with respect to these parameters) family of hyperbolic tori having the same basic frequencies. In the above setting, if the torus possesses some elliptic (oscillatory) normal directions we say that it is elliptic or partially elliptic. In this case the situation is completely different, since we have to take into account combinations between basic and normal frequencies in the small divisors that appear in the construction of these tori (the corresponding non-resonance conditions are usually referred to as Melnikov conditions [47, 48]). As a consequence, families of elliptic or partially elliptic invariant tori with fixed basic frequencies  $\omega$  cannot be continuous in general, but they turn out to be Cantorian with respect to parameters. First rigorous proofs of existence of elliptic tori were given in [54] for  $r = n - 1$  and in [21, 44] for  $r < n$ . We refer also to [9–11, 28, 35, 39, 40, 57, 59, 69, 71] as interesting contributions covering different points of view.

The main source of difficulty in the presence of elliptic normal directions is the so-called lack of parameters problem [10, 54, 68]. Basically, since we have only as many internal parameters ('actions') as the number of basic frequencies of the torus, we cannot simultaneously control the normal ones, so we cannot prevent them from 'falling into resonance'. This is equivalent to say that, for a given Hamiltonian system, we cannot construct a torus with a fixed set of basic and normal frequencies because there are not enough parameters. The previous fact leads to the exclusion of a small set of these internal parameters in order to avoid resonances involving normal frequencies. To control the measure of the set of excluded parameters, typically one assumes that the normal frequencies 'move' as a function of the internal parameters. Another possibility to overcome the lack of parameters problem is to apply the so-called Broer–Huiteima–Takens theory (see [11]). This consists in adding as many (external) parameters as needed to control simultaneously the values of both basic and normal frequencies (this process is referred to as unfolding). With this setting, we can prove that—under small perturbations—there exist invariant tori for a nearly full-measure Cantor set of parameters. The  $C^\infty$ -Whitney smoothness of this construction is also established. Finally, in order to ensure the existence of invariant tori for the original system (free of parameters), one can apply the so-called Herman's method. Indeed, external parameters can be eliminated—under very weak non-degeneracy conditions—by means of an appropriate technical result concerning the Diophantine approximation on submanifolds (see [10, 64, 66–69]).

Another issue linked to the persistence of lower dimensional invariant tori is the reducibility of the normal variational equations (at least in the elliptic directions). This is usually asked to simplify the study of the linearized equations involved. To control the small divisors of the cohomological equations that appear in the construction of the reduced matrix, it is typical

<sup>1</sup> As it is customary in the Hamiltonian context, the term 'hyperbolic torus' is not used in the sense of the general theory of normally hyperbolic invariant manifolds. We refer to [3] for a detailed discussion of this definition in the Hamiltonian setting.

to consider second Melnikov conditions [47, 48]. Other approaches for studying persistence of invariant tori in the elliptic context, without second Melnikov conditions, are discussed in remark 3.10.

Classical methods for studying the persistence of lower dimensional tori are based on canonical transformations performed on the Hamiltonian function. These methods typically deal with a perturbative setting, i.e. the problem is written as a perturbation of an ‘integrable’ Hamiltonian (in the sense that it has a continuous family of reducible invariant tori). They also take advantage of the existence of action-angle-like coordinates for the unperturbed Hamiltonian system. These coordinates play an important role in solving the cohomological equations involved in the iterative KAM process. They also allow us to guarantee (at every step) the isotropic character of the approximately invariant tori, thus simplifying a lot of details. However, classical approaches present some shortcomings, mainly due to the fact that they only allow us to face perturbative problems. For example:

- In many practical applications (design of space missions [25, 26], study of models in Celestial Mechanics [15], Molecular Dynamics [58, 70] or Plasma-Beam Physics [49], just to mention a few) we have to consider non-perturbative systems. We can obtain approximately invariant tori for these systems by means of numerical computations or asymptotic expansions, but in general we cannot apply classical results to prove the existence of these objects. In some cases it is possible to identify an integrable approximation of the system but the remaining part cannot be considered as an arbitrarily small perturbation.
- Even if we are studying a concrete perturbative problem, sometimes it is very complicated to establish action-angle variables for the unperturbed Hamiltonian. In some cases action-angle variables are not explicit, become singular or introduce problems of regularity (for example, when we approach a separatrix). Although in many contexts this shortcoming has been solved by means of several techniques (see for example [20, 31, 56] for a construction in the case of an integrable Hamiltonian and [12, 41] for a construction around a particular object), it introduces more technical difficulties in the problem.
- From the computational viewpoint, methods based on transformations are sometimes inefficient and quite expensive. This is a serious difficulty when implementing numerical methods or computer assisted proofs based on them.

An alternative to the classical approach are the so-called parametrization methods. They consist in performing an iterative scheme, without using canonical transformations, to solve the invariance equation of the torus. This scheme is carried out by adding a small function to the previous approximation of the torus. This function is obtained by solving (approximately) the linearized invariance equation around the approximated torus (the Newton method). We point out that the geometry of the problem plays an important role in the study of these equations. This approach is suitable for obtaining invariant tori of Hamiltonian systems using neither action-angle variables nor a perturbative setting. Such geometric approach—also referred to as KAM theory without action-angle variables—was introduced in [17] for Lagrangian tori and extended in [24] to hyperbolic lower dimensional tori, following long-time developed ideas (relevant work can be found in [15, 35, 52, 53, 63, 65, 72]). Roughly speaking, the insight of these methods is summarized in the following quote from [35]:

“...near approximate solutions of certain equations satisfying certain non-degeneracy assumptions, we can find true solutions defined on a large set.”

The aim of this paper is to adapt parametrization methods to study normally elliptic tori without using action-angle variables and in a non-perturbative setting. Concretely, we

assume that we have a 1-parameter family of Hamiltonian systems for which we know a 1-parameter family of approximately invariant lower dimensional elliptic tori—all of them with the same vector of basic frequencies—and also approximations of the vectors of normal frequencies and the normal directions associated with these frequencies (i.e. a basis of the normal directions along each torus that approximately reduces the normal variational equations to constant coefficients). Under suitable hypotheses of non-resonance and non-degeneracy, we show that for a Cantorian subset of parameters there exists a true elliptic torus close to the approximate one, having the same basic frequencies and slightly modified normal frequencies. This Cantorian set has a large relative Lebesgue measure. The scheme to deal with reducibility of the normal directions of these tori is the main contribution of this paper, and it consists in performing suitable (small) corrections in the normal directions at each step of the iterative procedure. We emphasize that both the preceding works [17, 24] devoted to parametrization methods considered invariant tori for individual systems (rather than parametric families of systems). In this paper we introduce an external parameter to enjoy at least a partial control over the normal frequencies, thus avoiding normal-internal resonances.

The setting of the paper has been selected in order to simplify some technical aspects of the result—both in the assumptions and in the proof—thus highlighting the geometric construction of the paper. We point out that all the basic ideas linked to parametrization methods, without using action-angle variables, for reducible lower dimensional tori are present in our approach. In section 3 we discuss several extensions and generalizations that can be tackled with the method presented in this paper.

Parametrization methods are computationally oriented in the sense that they can be implemented numerically, thus obtaining very efficient algorithms for the computation of invariant tori. For example, if we approximate a torus using  $N$  Fourier modes, such methods allow us to compute the object with a cost of order  $\mathcal{O}(N \log N)$  in time and  $\mathcal{O}(N)$  in memory (see remark 3.11). This is another advantage of our approach in contrast with classical methods based on transformation theory. The reader interested in such algorithms is referred to [18] for the implementation of the ideas in [17, 24] for Lagrangian and whiskered tori (see also [13] for the case of lattices and twist maps) and to [36] for the implementation of the ideas of [38] for reducible elliptic and hyperbolic tori for quasi-periodic skew-product maps (in this case, which corresponds to quasi-periodic perturbations of equilibrium points for flows, the geometric part discussed in this paper is not required).

Finally, we observe that in the presence of hyperbolic directions one can approach the problem by combining techniques in [24] (for studying hyperbolic directions) together with those introduced here (for studying elliptic directions). Indeed, the methodology presented in this work can be adapted to deal with invariant tori with reducible hyperbolic directions, but this assumption is quite restrictive (see [29]) in the hyperbolic context (reducibility is not required in [24]).

The paper is organized as follows. In section 2 we provide some notations, definitions and background of the problem. In section 3 we state the main result of this paper and we discuss several extensions and generalizations of the method presented. A motivating sketch of the construction performed in the proof of this result is given in section 4, together with a detailed description of some geometric properties of elliptic lower dimensional invariant tori of Hamiltonian systems. Next, in section 5 we perform one step of the iterative method to correct both an approximation of an elliptic invariant torus and a basis along this torus that approximately reduces the normal variational equations to constant coefficients. The new errors in invariance and reducibility are quadratic in terms of the previous ones. Finally, the main result is proved in section 6.

## 2. General background

In this section we introduce some notation and, in order to help the reader, we recall the basic terminology and concepts related to the problem. Thus, after setting the notation used in the paper in section 2.1, we provide the basic definitions regarding lower dimensional invariant tori of Hamiltonian systems (section 2.2) and their normal behaviour (section 2.3).

### 2.1. Basic notations

Given a real or complex scalar function  $f$  of several variables, we denote by  $Df$  its Jacobian matrix. If  $f$  is scalar, we also use the notation  $\text{grad } f = Df^\top$  for the gradient vector and  $\text{hess } f = D^2 f$  the Hessian matrix.

For any complex number  $z \in \mathbb{C}$  we denote  $z^* \in \mathbb{C}$  its complex conjugate number and  $\text{Re}(z)$ ,  $\text{Im}(z)$  the real and imaginary parts of  $z$ , respectively. We extend these notations to complex vectors and matrices. Note that if  $N$  is a matrix with complex entries then  $N^*$  is the matrix with complex conjugate entries without transposing.

Given a complex vector  $v \in \mathbb{C}^l$  we denote by  $\text{diag}(v) \in M_{l \times l}(\mathbb{C})$  the diagonal matrix having the components of  $v$  in the diagonal. Moreover, given  $Z \in M_{l \times l}(\mathbb{C})$ , we denote by  $\text{diag}(Z) \in M_{l \times l}(\mathbb{C})$  the diagonal matrix having the same diagonal entries as  $Z$ .

For any  $k \in \mathbb{Z}^r$ , we denote  $|k|_1 = |k_1| + \dots + |k_r|$ . Given a vector  $x \in \mathbb{C}^l$ , we set  $|x| = \sup_{j=1, \dots, l} |x_j|$  for the supremum norm and we extend the notation to the induced norm for complex matrices. Furthermore, given an analytic function  $f$ , with bounded derivatives in a complex domain  $\mathcal{U} \subset \mathbb{C}^l$ , and  $m \in \mathbb{N}$  we introduce the  $\mathcal{C}^m$ -norm for  $f$  as

$$\|f\|_{\mathcal{C}^m, \mathcal{U}} = \sup_{\substack{k \in (\mathbb{N} \cup \{0\})^l \\ 0 \leq |k|_1 \leq m}} \sup_{z \in \mathcal{U}} |D^k f(z)|.$$

We denote by  $\mathbb{T}^r = \mathbb{R}^r / (2\pi\mathbb{Z})^r$  the real  $r$ -dimensional torus, with  $r \geq 1$ . We use the  $|\cdot|$ -norm introduced above to define the complex strip around  $\mathbb{T}^r$  of width  $\rho > 0$  as

$$\Delta(\rho) = \{\theta \in \mathbb{C}^r / (2\pi\mathbb{Z})^r : |\text{Im}(\theta)| \leq \rho\}.$$

Accordingly, we will consider the Banach space of analytic functions  $f : \Delta(\rho) \rightarrow \mathbb{C}$  equipped with the norm

$$\|f\|_\rho = \sup_{\theta \in \Delta(\rho)} |f(\theta)|.$$

Similarly, if  $f$  takes values in  $\mathbb{C}^l$ , we set  $\|f\|_\rho = (|f_1|_\rho, \dots, |f_l|_\rho)$ . If  $f$  is a matrix-valued function, we extend  $\|f\|_\rho$  by computing the  $|\cdot|$ -norm of the constant matrix defined by the  $\|\cdot\|_\rho$ -norms of the entries of  $f$ . We observe that if the matrix product is defined then this space is a Banach algebra and we have  $\|f_1 f_2\|_\rho \leq \|f_1\|_\rho \|f_2\|_\rho$ . In addition, we can use Cauchy estimates

$$\left\| \frac{\partial f}{\partial \theta_j} \right\|_{\rho-\delta} \leq \frac{\|f\|_\rho}{\delta}, \quad j = 1, \dots, r.$$

For any function  $f$  analytic on  $\mathbb{T}^r$  and taking values in  $\mathbb{C}$ ,  $\mathbb{C}^l$  or in a space of complex matrices, we denote its Fourier series as

$$f(\theta) = \sum_{k \in \mathbb{Z}^r} \hat{f}_k e^{i\langle k, \theta \rangle}, \quad \hat{f}_k = \frac{1}{(2\pi)^r} \int_{\mathbb{T}^r} f(\theta) e^{-i\langle k, \theta \rangle} d\theta$$

and its average as  $[f]_{\mathbb{T}^r} = \hat{f}_0$ . We also set  $\tilde{f}(\theta) = f(\theta) - [f]_{\mathbb{T}^r}$ . Moreover, we have the following bounds:

$$|[f]_{\mathbb{T}^r}| \leq \|f\|_\rho, \quad \|\tilde{f}\|_\rho \leq 2\|f\|_\rho, \quad |\hat{f}_k| \leq \|f\|_\rho e^{-\rho|k|_1}.$$

Now, we introduce some notation regarding Lipschitz regularity. Assume that  $f(\mu)$  is a function defined for  $\mu \in I \subset \mathbb{R}$ —the subset  $I$  may not be an interval—taking values in  $\mathbb{C}$ ,  $\mathbb{C}^l$  or  $M_{l_1 \times l_2}(\mathbb{C})$ . We say that  $f$  is Lipschitz with respect to  $\mu$  on the set  $I$  if

$$\text{Lip}_I(f) = \sup_{\substack{\mu_1, \mu_2 \in I \\ \mu_1 \neq \mu_2}} \frac{|f(\mu_2) - f(\mu_1)|}{|\mu_2 - \mu_1|} < \infty.$$

The value  $\text{Lip}_I(f)$  is called the Lipschitz constant of  $f$  on  $I$ . For these functions we define  $\|f\|_I = \sup_{\mu \in I} |f(\mu)|$ . Similarly, if we have a family  $\mu \in I \subset \mathbb{R} \mapsto f_\mu$ , where  $f_\mu$  is a function on  $\mathbb{T}^r$  taking values in  $\mathbb{C}$ ,  $\mathbb{C}^l$  or  $M_{l_1 \times l_2}(\mathbb{C})$ , we extend the previous notations as

$$\text{Lip}_{I,\rho}(f) = \sup_{\substack{\mu_1, \mu_2 \in I \\ \mu_1 \neq \mu_2}} \frac{\|f_{\mu_2} - f_{\mu_1}\|_\rho}{|\mu_2 - \mu_1|}, \quad \|f\|_{I,\rho} = \sup_{\mu \in I} \|f_\mu\|_\rho.$$

Analogously, given a family  $\mu \in I \subset \mathbb{R} \mapsto f_\mu$ , where  $f_\mu$  is an analytic function with bounded derivatives in a complex domain  $\mathcal{U} \subset \mathbb{C}^l$ , we introduce for  $m \in \mathbb{N}$

$$\text{Lip}_{I,\mathcal{C}^m,\mathcal{U}}(f) = \sup_{\substack{\mu_1, \mu_2 \in I \\ \mu_1 \neq \mu_2}} \frac{\|f_{\mu_1} - f_{\mu_2}\|_{\mathcal{C}^m,\mathcal{U}}}{|\mu_1 - \mu_2|}, \quad \|f\|_{I,\mathcal{C}^m,\mathcal{U}} = \sup_{\mu \in I} \|f_\mu\|_{\mathcal{C}^m,\mathcal{U}}.$$

Finally, we say that  $f$  is Lipschitz from below with respect to  $\mu$  on the set  $I$  if

$$\text{lip}_I(f) = \inf_{\substack{\mu_1, \mu_2 \in I \\ \mu_1 \neq \mu_2}} \frac{|f(\mu_2) - f(\mu_1)|}{|\mu_2 - \mu_1|} > 0.$$

In this work we are concerned with Hamiltonian systems in  $\mathbb{R}^{2n}$  with respect to the standard symplectic form  $\Omega^0$ , given by  $\Omega^0(\xi, \eta) = \xi^\top J_n \eta$  where

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

is the canonical skew-symmetric matrix. We extend the notation above to write  $J_j$  for any  $1 \leq j \leq n$ , and  $\text{Id}_j$  for any  $1 \leq j \leq 2n$ . For the sake of simplicity, we denote  $J = J_n$  and  $\text{Id} = \text{Id}_{2n}$ .

Finally, given matrix-valued functions  $A : \mathbb{T}^r \rightarrow M_{2n \times l_a}(\mathbb{C})$  and  $B : \mathbb{T}^r \rightarrow M_{2n \times l_b}(\mathbb{C})$ , we set the notations  $G_{A,B}(\theta) = A(\theta)^\top B(\theta)$ ,  $\Omega_{A,B}(\theta) = A(\theta)^\top J B(\theta)$ ,  $G_A(\theta) = G_{A,A}(\theta)$  and  $\Omega_A(\theta) = \Omega_{A,A}(\theta)$ .

## 2.2. Invariant and approximately invariant tori

Given a Hamiltonian function  $h : U \subset \mathbb{R}^{2n} \rightarrow \mathbb{R}$ , we study the existence of lower dimensional quasi-periodic invariant tori for the Hamiltonian vector field  $X_h(x) = J \text{grad } h(x)$ .

**Definition 2.1.** For any integer  $1 \leq r \leq n$ ,  $\mathcal{T} \subset U$  is an  $r$ -dimensional invariant torus carrying conditionally periodic motion with basic frequencies  $\omega \in \mathbb{R}^r$  for  $X_h$ , if  $\mathcal{T}$  is invariant under the flow of  $X_h$  and there exists a parametrization given by an embedding  $\tau : \mathbb{T}^r \rightarrow U$  such that  $\mathcal{T} = \tau(\mathbb{T}^r)$ , making the following diagram commute

$$\begin{array}{ccc} \mathbb{T}^r & \xrightarrow{T_{I,\omega}} & \mathbb{T}^r \\ \tau \downarrow & & \downarrow \tau \\ \mathcal{T} & \xrightarrow{\phi_t|_{\mathcal{T}}} & \mathcal{T} \end{array} \quad (1)$$

where  $T_{t,\omega}(\theta) = \theta + \omega t$  is the (parallel) flow of the constant vector field

$$L_\omega = \omega_1 \frac{\partial}{\partial \theta_1} + \cdots + \omega_r \frac{\partial}{\partial \theta_r}$$

and  $\phi_t$  is the flow of  $X_h$ . In addition, if  $\omega$  is non-resonant, i.e.

$$\langle k, \omega \rangle \neq 0, \quad \forall k \in \mathbb{Z}^r \setminus \{0\}, \quad (2)$$

then we say that  $\mathcal{T}$  is an  $r$ -dimensional quasi-periodic torus.

The fact that  $\mathcal{T} = \tau(\mathbb{T}^r)$  and diagram (1) commutes is equivalent to that the embedding  $\tau$  satisfies

$$L_\omega \tau(\theta) = X_h(\tau(\theta)). \quad (3)$$

If  $\omega \in \mathbb{R}^r$  is non-resonant, then the quasi-periodic function  $z(t) = \tau(\omega t + \theta_0)$  is an integral curve of  $X_h$  for any  $\theta_0 \in \mathbb{T}^r$  that fills densely  $\mathcal{T}$ .

By means of  $\tau$  we can pull-back to  $\mathbb{T}^r$  both the restrictions to  $\mathcal{T}$  of the standard metric and the symplectic structure, obtaining the following matrix representations:

$$G_{D\tau}(\theta) = D\tau(\theta)^\top D\tau(\theta), \quad \Omega_{D\tau}(\theta) = D\tau(\theta)^\top J D\tau(\theta), \quad \theta \in \mathbb{T}^r.$$

**Remark 2.2.** We note that as  $\tau$  is embedding we have  $\text{rank}(D\tau(\theta)) = r$  for every  $\theta \in \mathbb{T}^r$ , so it turns out that  $\det G_{D\tau}(\theta) \neq 0$  for every  $\theta \in \mathbb{T}^r$ . Moreover, the average  $[\Omega_{D\tau}]_{\mathbb{T}^r}$  is zero since the symplectic form is exact. Indeed, if we write  $\tau(\theta) = (x(\theta), y(\theta))$  then we have  $\Omega_{D\tau}(\theta) = D\alpha(\theta) - D\alpha(\theta)^\top$ , where  $\alpha(\theta) = Dx(\theta)^\top y(\theta)$  and, by definition,  $[D\alpha]_{\mathbb{T}^r} = 0$ .

Next we state a version of a classical lemma—see for example [10, 51]—which says that any quasi-periodic torus of a locally Hamiltonian vector field is isotropic provided that the symplectic form is exact.

**Lemma 2.3.** *Let  $h : U \subset \mathbb{R}^{2n} \rightarrow \mathbb{R}$  be a Hamiltonian function and  $\mathcal{T}$  an  $r$ -dimensional invariant torus for  $X_h$  of non-resonant frequencies  $\omega$ . Then the submanifold  $\mathcal{T}$  is isotropic, i.e.  $\Omega_{D\tau}(\theta) = 0$  for every  $\theta \in \mathbb{T}^r$ . In particular, if  $r = n$  then  $\mathcal{T}$  is Lagrangian.*

**Remark 2.4.** In the text there appear many functions depending on  $\theta \in \mathbb{T}^r$ . In order to simplify the notation sometimes we omit the dependence on  $\theta$ —eventually we even omit the fact that some functions are evaluated at  $\tau(\theta)$  if there is no source of confusion.

**Proof of lemma 2.3.** Here we adapt a proof given in [17] that will be useful later on. First, we compute

$$\begin{aligned} L_\omega(\Omega_{D\tau}) &= L_\omega(D\tau^\top J D\tau) = [D(L_\omega \tau)]^\top J D\tau + D\tau^\top J D(L_\omega \tau) \\ &= [J \text{hess } h(\tau) D\tau]^\top J D\tau + D\tau^\top J J \text{hess } h(\tau) D\tau = 0, \end{aligned}$$

where we used that  $D \circ L_\omega = L_\omega \circ D$ , the hypothesis  $L_\omega \tau(\theta) = X_h(\tau(\theta))$  and the properties  $J^\top = -J$  and  $J^2 = -\text{Id}$ . Then, since  $\omega$  is non-resonant, the fact that the derivative  $L_\omega$  vanishes implies that  $\Omega_{D\tau} = [\Omega_{D\tau}]_{\mathbb{T}^r}$ . From remark 2.2 we conclude that  $\Omega_{D\tau} = 0$ .  $\square$

Finally, we set the idea of parametrization of an approximately invariant torus. Essentially, we measure how far to commute in diagram (1).

**Definition 2.5.** *Given a Hamiltonian  $h : U \subset \mathbb{R}^{2n} \rightarrow \mathbb{R}$  and an integer  $1 \leq r \leq n$ , we say that  $\mathcal{T} \subset U$  is an  $r$ -dimensional approximately quasi-periodic invariant torus with non-resonant basic frequencies  $\omega \in \mathbb{R}^r$  for  $X_h$  provided that there exists an embedding  $\tau : \mathbb{T}^r \rightarrow U$ , such that  $\mathcal{T} = \tau(\mathbb{T}^r)$ , satisfying*

$$L_\omega \tau(\theta) = J \text{grad } h(\tau(\theta)) + e(\theta),$$

where  $e : \mathbb{T}^r \rightarrow \mathbb{R}^{2n}$  is ‘small’ in a suitable norm.



Among the conditions needed to find a true invariant torus around an approximately invariant one, we are concerned with Diophantine conditions on the vector of basic frequencies.

**Definition 2.6.** We say that  $\omega \in \mathbb{R}^r$  satisfies Diophantine conditions of  $(\gamma, \nu)$ -type, for  $\gamma > 0$  and  $\nu > r - 1$ , if

$$|\langle k, \omega \rangle| \geq \frac{\gamma}{|k|_1^\nu}, \quad \forall k \in \mathbb{Z}^r \setminus \{0\}. \quad (4)$$

It is well known that if we consider a fixed  $\nu$  then, for almost every  $\omega \in \mathbb{R}^r$ , there is  $\gamma > 0$  for which (4) is fulfilled (see [46]).

### 2.3. Linear normal behaviour of invariant tori

In order to study the behaviour of the solutions in a neighbourhood of an  $r$ -dimensional quasi-periodic invariant torus of basic frequencies  $\omega$ —parametrized by  $\tau$ —it is usual to consider the variational equations around the torus,

$$L_\omega \xi(\theta) = J \text{hess } h(\tau(\theta)) \xi(\theta). \quad (5)$$

If  $r = 1$  system (5) is  $2\pi/\omega$ -periodic. Then, following Floquet's theorem, there exists a linear periodic change of variables that reduces the system to constant coefficients. If  $r > 1$ , then we consider reducibility to constant coefficients (in the sense of Lyapunov–Perron) as follows.

**Definition 2.7.** We say that the invariant torus  $\mathcal{T}$  in definition 2.1 is reducible if there exists a linear change of coordinates  $\xi = M(\theta)\eta$ , defined for  $\theta \in \mathbb{T}^r$ , such that the variational equations (5) turn out to be  $L_\omega \eta(\theta) = B\eta(\theta)$ , where  $B \in M_{2n \times 2n}(\mathbb{C})$ .

This property is equivalent to the fact that  $M$  satisfies the differential equation

$$L_\omega M(\theta) = J \text{hess } h(\tau(\theta)) M(\theta) - M(\theta) B. \quad (6)$$

In the Lagrangian case  $r = n$  such transformation exists, under regularity assumptions, provided  $\omega$  satisfies (4)—see [17]. Indeed, we can take derivatives at both sides of the invariance equation (3), thus obtaining

$$L_\omega D\tau(\theta) = J \text{hess } h(\tau(\theta)) D\tau(\theta).$$

Then, the matrix  $M(\theta)$  can be given by the columns of  $D\tau(\theta)$  and  $JD\tau(\theta)G_{D\tau}^{-1}(\theta) + D\tau(\theta)C(\theta)$ , where  $C(\theta)$  is a suitable  $n \times n$  matrix—obtained by solving certain cohomological equation. The reduced matrix turns out to be of the form

$$B = \begin{pmatrix} 0 & B^C \\ 0 & 0 \end{pmatrix},$$

where  $B^C \in M_{n \times n}(\mathbb{R})$  is symmetric. The  $2n$  zero eigenvalues correspond to the tangent directions to the torus together with their symplectic conjugate ones, meanwhile the matrix  $B^C$  controls the variation of the frequencies of the torus—the twist condition reads  $\det B^C \neq 0$ —when moving the ‘actions’ of the system.

In the lower dimensional case  $1 < r < n$  we cannot guarantee, in general, reducibility to constant coefficients (we refer to [33, 34, 61]). Nevertheless, if we consider a family of quasi-periodic linear perturbations of a linear system with constant coefficients then, under some generic hypotheses of non-resonance and non-degeneracy, we can state the reducibility of a large subfamily. On the one hand, if we restrict  $M(\theta)$  to the space of close-to-the-identity matrices, then we can prove that the reducible subfamily is Cantorian and has a large Lebesgue measure (we refer to [37, 38]). On the other hand, considering a more general

class of matrices (see ideas introduced in [22, 23, 43, 55]) reducibility can be extended to a full-measure subfamily (this was conjectured in [23] and proved in [30]).

If system (5) is reducible, then the geometry of the problem allows us to choose the matrix  $B$  with the following block structure:

$$B = \left( \begin{array}{cc|c} 0 & B^C & 0 \\ 0 & 0 & 0 \\ 0 & 0 & B^N \end{array} \right),$$

where  $B^C \in M_{r \times r}(\mathbb{R})$  is symmetric (it plays the same ‘twist’ role as in the Lagrangian case), and  $B^N \in M_{2(n-r) \times 2(n-r)}(\mathbb{C})$  can be written as  $B^N = J_{n-r} S$ , where  $S$  is also symmetric. The matrix  $B^N$  gives the normal linear behaviour of the torus. The real parts of the eigenvalues of  $B^N$  correspond to Lyapunov exponents and their imaginary parts to normal frequencies. As discussed in the introduction, in this work we are interested in the normally elliptic case, in which all the eigenvalues of  $B^N$  have a vanishing real part, i.e.

$$\text{spec}(B^N) = \{i\lambda_1, \dots, i\lambda_{n-r}, -i\lambda_1, \dots, -i\lambda_{n-r}\},$$

where  $\lambda_j \in \mathbb{R} \setminus \{0\}$  are the normal frequencies. Throughout the paper we assume that they have different moduli.

In order to simplify the resolution of the obtained cohomological equations, it is convenient to put the matrix  $B^N$  in diagonal form. In the classical KAM approach—using symplectic transformations and action-angle variables adapted to the torus—this is possible with a complex canonical change of coordinates. This transformation casts the initial real Hamiltonian into a complex one, having some symmetries. As these symmetries are preserved by the canonical transformations performed along these classical proofs, the final Hamiltonian can be made real and thus the obtained tori are real. In this paper we perform this complexification by selecting a complex matrix function  $N : \mathbb{T}^r \rightarrow M_{2n \times (n-r)}(\mathbb{C})$  associated with the eigenfunctions of eigenvalues  $i\lambda_1, \dots, i\lambda_{n-r}$ . It is clear that the real and imaginary parts of these vectors span the associated real normal subspace at any point of the torus. Indeed, from equation (6), the matrix function  $N$  satisfies

$$L_\omega N(\theta) = J \text{hess } h(\tau(\theta)) N(\theta) - N(\theta) \Lambda,$$

where  $\Lambda = \text{diag}(i\lambda) = \text{diag}(i\lambda_1, \dots, i\lambda_{n-r})$ . Then, together with these vectors, we resort to the use of the complex conjugate ones, which clearly satisfy

$$L_\omega N^*(\theta) = J \text{hess } h(\tau(\theta)) N^*(\theta) + N^*(\theta) \Lambda,$$

to span a basis of the complexified normal space along the torus (this is guaranteed by the conditions  $\det G_{N, N^*} \neq 0$  on  $\mathbb{T}^r$ ).

As we have pointed out in the introduction, to face the resolution of the cohomological equations standing for invariance and reducibility, we assume additional non-resonance conditions apart from (2).

**Definition 2.8.** We say that the normal frequencies  $\lambda \in \mathbb{R}^{n-r}$  are non-resonant with respect to  $\omega \in \mathbb{R}^r$  if

$$\langle k, \omega \rangle + \lambda_i \neq 0, \quad \forall k \in \mathbb{Z}^r, \quad i = 1, \dots, n-r, \quad (7)$$

and

$$\langle k, \omega \rangle + \lambda_i \pm \lambda_j \neq 0, \quad \forall k \in \mathbb{Z}^r \setminus \{0\}, \quad i, j = 1, \dots, n-r. \quad (8)$$

Conditions (7) and (8) are referred to as first and second Melnikov conditions, respectively (see [47, 48]).

In the spirit of definition 2.5, we introduce the idea of approximate reducibility as follows.

**Definition 2.9.** We say that the approximately invariant torus  $\mathcal{T}$  in definition 2.5 is approximately elliptic if there exists a map  $N : \mathbb{T}^r \rightarrow M_{2n \times (n-r)}(\mathbb{C})$  and normal frequencies  $\lambda \in \mathbb{R}^{n-r}$ , which are non-resonant with respect to  $\omega$ , satisfying

$$L_\omega N(\theta) = J \text{hess } h(\tau(\theta)) N(\theta) - N(\theta) \Lambda + R(\theta),$$

where  $\Lambda = \text{diag}(i\lambda)$ ,  $\det G_{N, N^*} \neq 0$  on  $\mathbb{T}^r$  and  $R : \mathbb{T}^r \rightarrow M_{2n \times (n-r)}(\mathbb{C})$  is ‘small’ in a suitable norm.

In order to avoid the effect of the small divisors associated with (7) and (8), we assume additional Diophantine conditions.

**Definition 2.10.** Let us consider non-resonant basic and normal frequencies  $(\omega, \lambda) \in \mathbb{R}^r \times \mathbb{R}^{n-r}$  and constants  $\gamma > 0$  and  $\nu > r - 1$ . We say that  $\lambda$  satisfies Diophantine conditions of  $(\gamma, \nu)$ -type with respect to  $\omega$  if

$$|\langle k, \omega \rangle + \lambda_i| \geq \frac{\gamma}{|k|_1^\nu}, \quad |\langle k, \omega \rangle + \lambda_i \pm \lambda_j| \geq \frac{\gamma}{|k|_1^\nu}, \quad (9)$$

$\forall k \in \mathbb{Z}^r \setminus \{0\}$  and  $i, j = 1, \dots, n - r$ .

### 3. Statement of the main result

In this section we state the main result of the paper. We consider a 1-parameter family of Hamiltonian systems for which we know a family of parametrizations of approximately (with small error) elliptic lower dimensional invariant tori, all with the same basic frequencies. Then, under certain non-degeneracy conditions, we use the parameter to control the normal frequencies in order to prove that there exists a large set of parameters for which we have a true elliptic invariant torus close to the approximate one. We emphasize that we do not assume that the system is given in action-angle-like coordinates nor that the Hamiltonians are close-to-integrable ones.

**Theorem 3.1.** Let us consider a family of Hamiltonians  $\mu \in I \subset \mathbb{R} \mapsto h_\mu$  with  $h_\mu : U \subset \mathbb{R}^{2n} \rightarrow \mathbb{R}$ , where  $I$  is a finite interval and  $U$  is an open set. Let  $\omega \in \mathbb{R}^r$  be a vector of basic frequencies satisfying Diophantine conditions (4) of  $(\hat{\gamma}, \hat{\nu})$ -type, with  $\hat{\gamma} > 0$  and  $\hat{\nu} > r - 1$ . Assume that the following hypotheses hold:

- H<sub>1</sub> The functions  $h_\mu$  are real analytic and can be holomorphically extended to some complex neighbourhood  $\mathcal{U}$  of  $U$ . Moreover, we assume that  $\|h\|_{I, \mathcal{C}^4, \mathcal{U}} \leq \sigma_0$ .
- H<sub>2</sub> There exists a family of approximately invariant and elliptic tori of  $h_\mu$  in the sense of definitions 2.5 and 2.9, i.e. we have families of embeddings  $\mu \in I \subset \mathbb{R} \mapsto \tau_\mu$ , matrix functions  $\mu \in I \subset \mathbb{R} \mapsto N_\mu$  and approximated normal eigenvalues  $\Lambda_\mu = \text{diag}(i\lambda_\mu)$ , with  $\lambda_\mu \in \mathbb{R}^{n-r}$ , satisfying

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

for certain error functions  $e_\mu$  and  $R_\mu$ , where  $\tau_\mu$  and  $N_\mu$  are analytic and can be holomorphically extended to  $\Delta(\rho)$  for certain  $0 < \rho < 1$ , satisfying  $\tau_\mu(\Delta(\rho)) \subset \mathcal{U}$ . Assume also that we have constants  $\sigma_1, \sigma_2$  such that

$$\|D\tau\|_{I, \rho}, \|N\|_{I, \rho}, \|G_{D\tau}^{-1}\|_{I, \rho}, \|G_{N, N^*}^{-1}\|_{I, \rho} < \sigma_1, \quad \text{dist}(\tau_\mu(\Delta(\rho)), \partial\mathcal{U}) > \sigma_2 > 0,$$

for every  $\mu \in I$ , where  $\partial\mathcal{U}$  stands for the boundary of  $\mathcal{U}$ .

H<sub>3</sub> We have  $\text{diag}[\Omega_{N_\mu, N_\mu^*}]_{\mathbb{T}^r} = i\text{Id}_{n-r}$  for every  $\mu \in I$ .

H<sub>4</sub> The family of matrix functions

$$A_{1,\mu}(\theta) = G_{D\tau_\mu}^{-1}(\theta) D\tau_\mu(\theta)^\top [T_{1,\mu}(\theta) + T_{2,\mu}(\theta) + T_{2,\mu}(\theta)^\top] D\tau_\mu(\theta) G_{D\tau_\mu}^{-1}(\theta),$$

where

$$T_{1,\mu}(\theta) = J^\top \text{hess } h_\mu(\tau_\mu(\theta)) J - \text{hess } h_\mu(\tau_\mu(\theta)),$$

$$T_{2,\mu}(\theta) = T_{1,\mu}(\theta) J [D\tau_\mu(\theta) G_{D\tau_\mu}(\theta)^{-1} D\tau_\mu(\theta)^\top - \text{Id}] \text{Re}(iN_\mu(\theta) N_\mu^*(\theta)^\top)$$

satisfies the non-degeneracy (twist) condition  $\|[A_1]_{\mathbb{T}^r}^{-1}\|_I < \sigma_1$ .

H<sub>5</sub> There exist constants  $\sigma_3, \sigma_4$  such that for every  $\mu \in I$  the approximated normal frequencies

$\lambda_\mu = (\lambda_{1,\mu}, \dots, \lambda_{n-r,\mu})$  satisfy

$$0 < \frac{\sigma_3}{2} < |\lambda_{i,\mu}| < \frac{\sigma_4}{2}, \quad 0 < \sigma_3 < |\lambda_{i,\mu} \pm \lambda_{j,\mu}|,$$

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

H<sub>6</sub> The objects  $h_\mu, \tau_\mu, N_\mu$  and  $\lambda_\mu$  are at least  $\mathcal{C}^1$  with respect to  $\mu$ , and we have

$$\left\| \frac{dh}{d\mu} \right\|_{I, \mathcal{C}^3, \mathcal{U}}, \left\| \frac{d\tau}{d\mu} \right\|_{I, \rho}, \left\| \frac{dD\tau}{d\mu} \right\|_{I, \rho}, \left\| \frac{dN}{d\mu} \right\|_{I, \rho}, \left\| \frac{d\lambda_i}{d\mu} \right\|_I < \sigma_5,$$

for  $i = 1, \dots, n-r$ . Moreover, we have the following separation conditions:

$$0 < \frac{\sigma_6}{2} < \left| \frac{d}{d\mu} \lambda_{i,\mu} \right|, \quad 0 < \sigma_6 < \left| \frac{d}{d\mu} \lambda_{i,\mu} \pm \frac{d}{d\mu} \lambda_{j,\mu} \right|,$$

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

Under these assumptions there is a constant  $C_1$  such that, if

$$\varepsilon_* = \|e\|_{I, \rho} + \|R\|_{I, \rho} + \left\| \frac{de}{d\mu} \right\|_{I, \rho} + \left\| \frac{dR}{d\mu} \right\|_{I, \rho}$$

satisfies  $\varepsilon_* \leq C_1 \gamma_0^8$  for a given  $\gamma_0 \leq \frac{1}{2} \min\{1, \hat{\gamma}\}$ , then there exists a Cantorian subset  $I_{(\infty)} \subset I$  such that for every  $\mu \in I_{(\infty)}$  the Hamiltonian  $h_\mu$  has an  $r$ -dimensional elliptic invariant torus  $\mathcal{T}_{\mu, (\infty)}$  with basic frequencies  $\omega$ . The normal frequencies  $\lambda_{\mu, (\infty)}$  satisfy Diophantine conditions of the form

$$|\langle k, \omega \rangle + \lambda_{i, \mu, (\infty)}| \geq \frac{\gamma_0}{|k|_1^v}, \quad |\langle k, \omega \rangle + \lambda_{i, \mu, (\infty)} \pm \lambda_{j, \mu, (\infty)}| \geq \frac{\gamma_0}{|k|_1^v}, \quad \forall k \in \mathbb{Z}^r \setminus \{0\},$$

for  $v > \hat{v}$  and

$$\|\lambda_{i, (\infty)} - \lambda_i\|_{I_{(\infty)}} \leq \frac{C_2 \varepsilon_*}{\gamma_0^2}, \quad (10)$$

for  $i, j = 1, \dots, n-r$ . The parametrizations  $\tau_{\mu, (\infty)}$  and normal directions  $N_{\mu, (\infty)}$  satisfy

$$\|\tau_{(\infty)} - \tau\|_{I_{(\infty), \rho/2}} \leq \frac{C_2 \varepsilon_*}{\gamma_0^2}, \quad \|N_{(\infty)} - N\|_{I_{(\infty), \rho/2}} \leq \frac{C_2 \varepsilon_*}{\gamma_0^4}. \quad (11)$$

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

$$\text{meas}_{\mathbb{R}}(I \setminus I_{(\infty)}) \leq C_3 \gamma_0. \quad (12)$$

The constants  $C_1, C_2$  and  $C_3$  depend on  $|\omega|, v, r, n, \sigma_0, \sigma_1, \sigma_2, \sigma_3, \sigma_4, \sigma_5$  and  $\sigma_6$ . The constant  $C_3$  also depends on  $\hat{\gamma}, \hat{v}$ .

**Remark 3.2.** We will see that, if  $\|e\|_{I,\rho}$  and  $\|R\|_{I,\rho}$  are small enough, then hypotheses  $H_2$  and  $H_5$ , together with suitable Diophantine conditions on  $\omega$  and  $\lambda$ , imply that the matrix  $\Omega_{N,N^*}$  is pure imaginary, approximately constant and close to diagonal (see propositions 4.1 and 5.3 for details). In order to follow our approach for constructing an approximately symplectic basis along the torus, we assume that the average of this matrix is non-singular. According to this, it is clear that we can assume (after a suitable choice of the sign of the components of  $\lambda$  and scaling of the columns of  $N$ ) that  $\text{diag}[\Omega_{N,N^*}]_{\mathbb{T}^r} = i\text{Id}_{n-r}$ , as it is done in hypothesis  $H_3$  of theorem 3.1.

**Remark 3.3.** As it is customary in parametrization methods—we encourage the reader to compare this result with those in [17, 24, 35]—the conditions of theorem 3.1 can be verified using information provided by the initial approximations. This fact is useful in the validation of numerical computations that consist in looking for trigonometric functions that satisfy invariance and reducibility equations approximately. Concretely, let us assume that for a given parameter  $\mu_0 \in I$  we have computed approximations  $\tau_{\mu_0}$ ,  $N_{\mu_0}$  and  $\Lambda_{\mu_0}$  satisfying the explicit conditions of theorem 3.1 for certain  $\omega \in \mathbb{R}^r$ . Then, for most of the values of  $\mu$  close to  $\mu_0$ , there exists an elliptic quasi-periodic invariant torus nearby, whose normal frequencies are just slightly changed.

**Remark 3.4.** Hypothesis  $H_4$  is called twist condition because when applying this result in a perturbative setting it stands for the Kolmogorov non-degeneracy condition (see the computations performed for Hamiltonian (13)). Observe that in the Lagrangian case the function  $A_{1,\mu}$  reads as  $A_{1,\mu} = G_{D\tau_\mu}^{-1} D\tau_\mu^\top T_{1,\mu} D\tau G_{D\tau_\mu}^{-1}$  for the same matrix  $T_{1,\mu}$ , thus recovering the condition in [17].

**Remark 3.5.** Let us assume that  $0 \in I$  and that for  $\mu = 0$  we have a true elliptic quasi-periodic invariant torus satisfying the Diophantine and non-degeneracy conditions of theorem 3.1. In this case, it is expected that the measure of invariant tori nearby is larger than the one predicted by theorem 3.1. Actually, it is known that the complementary set  $[-\mu_0, \mu_0] \setminus I_{(\infty)}$  has measure exponentially small when  $\mu_0 \rightarrow 0$  (see [39, 40]). To obtain such estimates we would need to slightly modify some details of the proof performed here—but not the scheme—asking for Diophantine conditions as those used in [38, 40] (which turn out to be exponentially small in  $|k|_1$ ).

**Remark 3.6.** Theorem 3.1 can be extended to exact symplectic maps. Actually, the parametrization approach in the context of maps is the main setting in [17, 24, 35]. To this end, we should ‘translate’ the computations performed along the paper to the context of maps, following the ‘dictionary’ of these references. Attention should be taken in order to adapt the geometric conditions that we highlight in remarks 4.5 and 4.7, which are not true for maps, but satisfied up to quadratic terms (this is enough for the convergence of the scheme).

**Remark 3.7.** It would also be interesting to extend the result in order to deal with symplectic vector fields or symplectic maps. Let us recall that a vector field  $X$  on a symplectic manifold with 2-form  $\Omega$  is said to be symplectic (or locally Hamiltonian) if  $L_X \Omega = 0$ , i.e. if the 2-form is preserved along the flow of  $X$  (symplectic vector fields that are not Hamiltonian can be found for example in the context of magnetic fields). In this situation, the method of ‘translated torus’ should be adapted as it is done in [24] for the hyperbolic case. To this end, it must be taken into account that the cohomology of the torus must be compatible with the cohomology class of the contraction  $\Omega(\cdot, X)$ .

**Remark 3.8.** The scheme of the proof of theorem 3.1 can also be used for proving the existence of reducible tori having some hyperbolic directions, under the assumption of first and second

Melnikov conditions. In this case, we need to adapt the geometrical ideas of the paper in order to deal simultaneously with elliptic and hyperbolic directions. However, as hyperbolic tori are known to exist beyond the breakdown of reducibility (see [29]), it is interesting to approach the problem of partially elliptic tori by combining techniques in [24] (for studying hyperbolic directions) together with those presented here (for studying elliptic directions).

**Remark 3.9.** The scheme can be also adapted to deal with the classical Broer–Huitema–Takens approach (see [11]) explained in the introduction. On the one hand, this allows obtaining  $C^\infty$ -Whitney regularity for the constructed tori, and on the other hand this permits to deal with degenerate cases where the Kolmogorov condition does not hold. For example, if we have other higher-order non-degeneracy conditions such as the so-called Rüssmann’s non-degeneracy condition (see [67]). These ideas could also be used to deal with resonant situations as those considered in [7, 8], where double frequencies and in some cases also vanishing frequencies are allowed.

**Remark 3.10.** After the work in [5, 6, 23, 30, 71] it is known that second Melnikov conditions are not necessary for proving existence of lower dimensional tori in the elliptic context. For example, Bourgain approached the problem without using reducibility, thus avoiding to ask for these non-resonance conditions. However, cumbersome multiscale analysis is required to approximate the solution of truncated cohomological equations, thus leading to a process which is not suitable for numerical implementations—at each step, one has to invert a large matrix which has a huge computational cost. Nevertheless, asking for reducibility we end up inverting a diagonal matrix in Fourier space (see remark 3.11). Another approach to avoid second Melnikov conditions was proposed by Eliasson in [23] and consists in performing a far-from-identity transformation when we have to deal with such resonant frequencies. Concretely, if  $\lambda \in \mathbb{R}^{n-r}$  does not satisfy second Melnikov conditions, then we can introduce new normal frequencies  $\tilde{\lambda}_j = \lambda_j - \langle m_j, \omega/2 \rangle$ , and we can carefully choose the vectors  $m_j \in \mathbb{Z}^r$  in such a way that second Melnikov conditions are satisfied (it is also necessary to work in the covering  $2\mathbb{T}^r = \mathbb{R}^r / (4\pi\mathbb{Z})^r$  of the torus). In this paper we study reducible tori without using Eliasson’s method (thus emphasizing the geometric ideas linked with parametrization methods). We ask for second Melnikov conditions paying the price of excluding a small set of invariant tori. Nevertheless, when numerically implementing this method, the use of Eliasson’s transformation is very useful (this was used in [29] to continue elliptic tori beyond their bifurcation to hyperbolic tori).

**Remark 3.11.** All the computations performed in the proof of theorem 3.1 can be implemented very efficiently in a computer. For example, the solution of cohomological equations with constant coefficients and the computation of derivatives like  $D\tau$  or  $L_\omega\tau$  correspond to diagonal operators in Fourier space. Other algebraic manipulations can be performed efficiently in real space and there are very fast and robust fast Fourier transform algorithms that allow passing from real (or complex) space to Fourier space (and ‘vice versa’). Accordingly, if we approximate a torus using  $N$  Fourier modes, we can implement an algorithm to compute the object with a cost of order  $\mathcal{O}(N \log N)$  in time and  $\mathcal{O}(N)$  in memory. We refer to the works [13, 18, 36] to analogous algorithms in several contexts. Therefore, this approach presents significant advantages in contrast with methods which require to deal with full matrices, since they represent a cost of  $\mathcal{O}(N^2)$  in memory and  $\mathcal{O}(N^3)$  in time (we refer for example to [14]).

Although one of the main features of both the formulation and the proof of theorem 3.1 is that we do not require to write the problem in action-angle coordinates, we think that it can be illustrative to express this result for a close-to-integrable system, in order to clarify the

meaning of hypotheses  $H_3$  and  $H_4$  in this context. Indeed, let us consider the following family of Hamiltonian systems written in action-angle-like coordinates  $(\varphi, y, z) \in \mathbb{T}^r \times \mathbb{R}^r \times \mathbb{R}^{2(n-r)}$ :

$$h_\mu(\varphi, y, z) = h_0(y, z) + \mu f(\varphi, y, z) \quad (13)$$

such that for  $y = 0$ , we have that  $z = 0$  is an elliptic non-degenerate equilibrium for the system  $h_0(y, z)$ . In particular,  $\tau_0(\theta) = (\theta, 0, 0)$  gives a parametrization of an invariant torus of  $h_0$  with basic frequencies  $\omega = \text{grad}_y h_0(0, 0) \in \mathbb{R}^r$ . By performing a suitable canonical change of variables in order to eliminate crossed quadratic terms in  $(y, z)$ , we can assume that

$$h_0(y, z) = \langle \omega, y \rangle + \frac{1}{2} \langle y, Ay \rangle + \frac{1}{2} \langle z, Bz \rangle + \mathcal{O}_3(y, z)$$

close to  $(y, z) = (0, 0)$ , where  $A$  and  $B$  are symmetric matrices, such that

$$\text{spec}(J_{n-r}B) = \{i\lambda_1, \dots, i\lambda_{n-r}, -i\lambda_1, \dots, -i\lambda_{n-r}\}$$

are the normal eigenvalues of the torus  $\mathbb{T}^r \times \{0\} \times \{0\}$ . The associated normal directions are given by the real and imaginary parts of the matrix of eigenvectors satisfying  $J_{n-r}B\hat{N} = \hat{N}\Lambda$ , where  $\Lambda = \text{diag}(i\lambda) = \text{diag}(i\lambda_1, \dots, i\lambda_{n-r})$ . Using symplectic properties, we can select the signs of the components of  $\lambda$  and the complex matrix  $\hat{N}$  in such a way that it satisfies  $\hat{N}^\top J_{n-r} \hat{N}^* = i\text{Id}_{n-r}$ .

Then, to apply theorem 3.1 to the family of Hamiltonians  $h_\mu$  given by (13), for small  $|\mu|$ , we consider the family of approximately elliptic and invariant tori  $\tau_\mu(\theta) = \tau_0(\theta) + \mathcal{O}(\mu)$  with normal frequencies  $\lambda_\mu = \lambda + \mathcal{O}(\mu)$  and normal vectors  $N_\mu(\theta)^\top = (0 \ 0 \ \hat{N}^\top) + \mathcal{O}(\mu)$ , where the terms  $\mathcal{O}(\mu)$  stand for the first order corrections in  $\mu$ —they can be computed by means of Lindstedt series or normal forms with respect to  $\mu$ —that are needed in order to check that the normal frequencies ‘move’ as a function of  $\mu$ . This family satisfies

$$\begin{aligned} L_\omega \tau_\mu(\theta) &= J \text{grad } h_\mu(\tau_\mu(\theta)) + \mathcal{O}_2(\mu), \\ L_\omega N_\mu(\theta) &= J \text{hess } h_\mu(\tau_\mu(\theta)) N_\mu(\theta) - N_\mu(\theta) \Lambda_\mu + \mathcal{O}_2(\mu), \end{aligned}$$

and, for  $\mu = 0$ , we have

$$D\tau_0(\theta) = \begin{pmatrix} \text{Id}_r \\ 0 \\ 0 \end{pmatrix}, \quad N_0(\theta) = \begin{pmatrix} 0 \\ 0 \\ \hat{N} \end{pmatrix}, \quad G_{D\tau_0}^{-1}(\theta) = \text{Id}_r, \quad \Omega_{N_0, N_0^*}(\theta) = i\text{Id}_{n-r}.$$

Moreover, it is not difficult to check that the matrix  $A_{1,\mu}(\theta)$  in  $H_4$  at  $\mu = 0$  reads as  $A_{1,0}(\theta) = -A$ , which implies that  $H_4$  is equivalent to the standard (Kolmogorov) non-degeneracy condition for the unperturbed system.

#### 4. Overview and heuristics of the method

In this section we outline the main ideas of the presented approach. Our aim is to emphasize the geometric interpretation of our construction, thus highlighting the additional difficulties with respect to the Lagrangian and normally hyperbolic cases. First, in section 4.1, we sketch briefly the proof of theorem 3.1. Although some parts of the proof involve quite cumbersome computations, the construction of the iterative procedure is fairly natural. In section 4.2 we focus on the geometric properties of the invariant and elliptic case. These properties allow us to obtain approximate solutions for the equations derived in section 4.1 associated with approximately invariant and elliptic tori.

#### 4.1. Sketch of the proof

Let  $h : U \subset \mathbb{R}^{2n} \rightarrow \mathbb{R}$  be a Hamiltonian function and let us suppose that  $\mathcal{T}$  is an approximately invariant and elliptic torus of basic frequencies  $\omega \in \mathbb{R}^r$  and normal ones  $\lambda \in \mathbb{R}^{n-r}$ , satisfying strong non-resonance conditions (4) and (9). The translation of definitions 2.5 and 2.9 into a functional setting is

$$\mathcal{F}(\tau) = e, \quad \mathcal{G}(\tau, N, \Lambda) = R,$$

with  $\Lambda = \text{diag}(i\lambda)$ , where we have introduced the following operators:

$$\begin{aligned} \mathcal{F}(\tau) &= L_\omega \tau - J \text{grad } h(\tau), \\ \mathcal{G}(\tau, N, \Lambda) &= L_\omega N - J \text{hess } h(\tau) N + N \Lambda. \end{aligned}$$

Then, we look for an embedding  $\bar{\tau} : \mathbb{T}^r \rightarrow U$ , normal directions  $\bar{N} : \mathbb{T}^r \rightarrow M_{2n \times (n-r)}(\mathbb{C})$  and normal frequencies  $\bar{\lambda}$ , satisfying

$$\mathcal{F}(\bar{\tau}) = 0, \quad \mathcal{G}(\bar{\tau}, \bar{N}, \bar{\Lambda}) = 0,$$

with  $\bar{\Lambda} = \text{diag}(i\bar{\lambda})$ . Since these equations have a triangular structure, we approach first the correction of the parametrization of the torus, i.e. we look for  $\bar{\tau} = \tau + \Delta_\tau$  satisfying the above expressions. We write the first equation as

$$\mathcal{F}(\tau + \Delta_\tau) = e + L_\omega \Delta_\tau - J \text{hess } h(\tau) \Delta_\tau + O_2(\Delta_\tau) = 0.$$

If we neglect terms  $O_2(\Delta_\tau)$  we obtain the following linearized equation (the Newton method):

$$L_\omega \Delta_\tau - J \text{hess } h(\tau) \Delta_\tau = -e, \quad (14)$$

which allows us to correct the invariance of the torus up to terms of second order in  $e$ . In a similar way, we look for  $\bar{N} = N + \Delta_N$  and  $\bar{\Lambda} = \Lambda + \Delta_\Lambda$  such that

$$\mathcal{G}(\bar{\tau}, \bar{N}, \bar{\Lambda}) = \hat{R} + L_\omega \Delta_N - J \text{hess } h(\tau) \Delta_N + N \Delta_\Lambda + \Delta_N \Lambda + O_2(\Delta_N, \Delta_\Lambda) = 0,$$

where

$$\hat{R} = R + J \text{hess } h(\tau) N - J \text{hess } h(\bar{\tau}) N \quad (15)$$

includes both the error in reducibility and the one introduced when correcting the torus (which is expected to be of order of the size of  $e$ ). Hence, in order to apply one step of the Newton method to correct reducibility, we have to solve the following linearized equation for  $\Delta_N$  and  $\Delta_\Lambda$ :

$$L_\omega \Delta_N - J \text{hess } h(\tau) \Delta_N + N \Delta_\Lambda + \Delta_N \Lambda = -\hat{R}. \quad (16)$$

For convenience, once we fix  $\tau$ ,  $N$  and  $\Lambda$ , we define the following differential operators (acting on vectors or matrices of  $2n$  rows):

$$\mathcal{R}(\xi) = L_\omega \xi - J \text{hess } h(\tau) \xi, \quad (17)$$

$$\mathcal{S}(\xi, \eta) = \mathcal{R}(\xi) + N \eta + \xi \Lambda, \quad (18)$$

so equations (14) and (16) are equivalent to invert  $\mathcal{R}$  and  $\mathcal{S}$

$$\mathcal{R}(\Delta_\tau) = -e, \quad \mathcal{S}(\Delta_N, \Delta_\Lambda) = -\hat{R}. \quad (19)$$

As it was done in [17], the main idea is to transform the linearized equation (14) into a simpler linear equation. This is done using a suitable basis along the approximate torus. The obtained equation can be approximately solved by means of the Fourier series. Indeed, an approximate solution with an error of quadratic size in  $e$  and  $R$  is enough for the convergence of the scheme—the Newton method still converges quadratically if we have a good enough



approximation of the Jacobian matrix. Under suitable conditions of non-resonance and non-degeneracy, iteration of this process leads to a quadratic scheme that allows us to overcome the effect of the small divisors of the problem. The main contribution of this paper is to adapt this construction (that we describe next in a more precise way) to deal with equations (14) and (16) simultaneously.

Let us discuss the construction of the basis mentioned above. In the Lagrangian case we only have to deal with equation (14) and the columns of the matrices  $D\tau$  and  $JD\tau G_{D\tau}^{-1}$  give us an approximately symplectic basis of  $\mathbb{R}^{2n}$  at any point of the torus. Moreover, it turns out that  $\mathcal{R}(D\tau) = 0 + \mathcal{O}(e)$  and  $\mathcal{R}(JD\tau G_{D\tau}^{-1}) = D\tau A_1 + \mathcal{O}(e)$ , where  $A_1 : \mathbb{T}^n \rightarrow M_{n \times n}(\mathbb{R})$  is a symmetric matrix. Using this basis we can write the linearized equation (14) in ‘triangular form’ with respect to the projections of  $\Delta_\tau$  over  $D\tau$  and  $JD\tau G_{D\tau}^{-1}$ , in such a way that the problem is reduced to solve two cohomological equations with constant coefficients. However, in the lower dimensional case the previous construction is not enough since we also have to take into account the normal directions of the torus. As mentioned in the introduction, this scheme has been recently adapted in [24] for the normally hyperbolic case, without requiring reducibility of the normal variational equations. The main ingredient is that there exists a splitting between the centre and the hyperbolic directions of the torus and we can reduce the study of equation (14) to the projections according to this splitting. The dynamics on the hyperbolic directions is characterized by asymptotic (geometric) growth conditions<sup>2</sup>—both in the future and in the past—and the linearized equation (14) restricted to the centre subspace follows as in the Lagrangian case (now the ambient space is  $\mathbb{R}^{2r}$ ).

In the normally elliptic context, we ask for reducibility in order to express equation (14) in a simple form. Hence, we solve simultaneously equation (16), thus obtaining a basis that reduces the normal variational equations of the torus to constant coefficients up to a quadratic error. In this case, the approximately (with an error of the order of the size of  $e$  and  $R$ ) symplectic basis is obtained by completing the columns of  $D\tau$ ,  $N$  and  $iN^*$  with the columns of a suitably constructed matrix  $V : \mathbb{T}^r \rightarrow M_{2n \times r}(\mathbb{R})$ . Basically, we take advantage of the fact that  $V$  satisfies  $\mathcal{R}(V) = D\tau A_1$  modulo terms of order  $e$  and  $R$ , where  $A_1 : \mathbb{T}^r \rightarrow M_{r \times r}(\mathbb{R})$  will be specified later on. Hence, we find approximately solutions for equations (19) in terms of the constructed basis as follows:

$$\begin{aligned}\Delta_\tau &= D\tau \Delta_1 + V \Delta_2 + N \Delta_3 + iN^* \Delta_4, \\ \Delta_N &= D\tau P_1 + V P_2 + N P_3 + iN^* P_4,\end{aligned}$$

where  $\{\Delta_i\}, \{P_i\}$ , with  $i = 1, \dots, 4$ , are the solutions of cohomological equations (37)–(40) and (42)–(45), respectively. The correction  $\Delta_\Lambda$  in the normal eigenvalues is determined from the compatibility condition of these last equations.

Let us observe that in order to correct the reducibility of the torus we have to change slightly the normal directions and the normal frequencies. Since the normal frequencies  $\lambda$  are modified at each step of the process, we do not know in advance if they will satisfy the required Diophantine conditions for all steps—unless we have enough parameters to control the value of all of them simultaneously. To deal with this problem we require some control on the change of these frequencies, in such a way that we can remove parameters that give rise to resonant frequencies. Since at every step of the inductive process we are removing a dense set of parameters, we have to work with sets of empty interior. This does not allow us to keep any kind of smooth dependence (in the usual sense) with respect to the parameter.

There are several methods in the literature to deal with this problem. The first approach was due to Arnold (see [1]) and it consists in working, at every step of the inductive procedure,

<sup>2</sup> Concretely, the solutions for the equations projected into the hyperbolic directions are obtained by means of an absolutely convergent power series. See details in [24].

with a finite number of terms in the Fourier expansions ('ultraviolet cut-off'). Then, since we only need to deal with a finite number of resonances at every step, we can work on open sets of parameters and keep the smooth dependence on these sets. Another possibility is to consider Lipschitz parametric dependence and to check that this dependence is preserved along the iterative procedure (this is the method used in [37–40]). Lipschitz regularity suffices to control the measure of the resonant sets. In this paper we follow the Lipschitz approach because it allows us to present the geometric construction in a clear way, assuming that we have Diophantine conditions up to any order. Then, the technical details corresponding to the Lipschitz regularity are performed separately in section 6.2.

#### 4.2. Characterization of the invariant and reducible case

Our goal now is to formally 'invert' the linear operators  $\mathcal{R}$  given by (17) and  $\mathcal{S}$  given by (18)—see propositions 4.4 and 4.6, respectively—when the corresponding torus  $\mathcal{T}$  is invariant and normally elliptic. In order to do this, first we characterize at a formal level some geometric properties of lower dimensional elliptic invariant tori. Later on, the same construction provided in this section will be used to study approximately invariant tori in order to solve equations in (19) with a small error (controlled by the errors of invariance and reducibility).

All along this section we consider an  $r$ -dimensional normally elliptic quasi-periodic invariant torus  $\mathcal{T}$  for a Hamiltonian  $h$  of basic frequencies  $\omega \in \mathbb{R}^r$  and normal frequencies  $\lambda \in \mathbb{R}^{n-r}$  satisfying non-resonance conditions (2), (7) and (8), i.e. we have

$$L_\omega \tau(\theta) = J \operatorname{grad} h(\tau(\theta)), \quad (20)$$

$$L_\omega N(\theta) = J \operatorname{hess} h(\tau(\theta)) N(\theta) - N(\theta) \Lambda, \quad (21)$$

with  $\Lambda = \operatorname{diag}(i\lambda)$ . We assume also that the matrices  $G_{D\tau}(\theta)$  and  $G_{N,N^*}(\theta)$  are invertible for every  $\theta \in \mathbb{T}^r$ . Then, we claim (see the proof of proposition 4.1) that under these conditions  $\Omega_{N,N^*}$  is constant, pure imaginary and diagonal. If we assume that this matrix is non-singular, then we can suppose that (see remark 3.2)

$$\Omega_{N,N^*}(\theta) = i \operatorname{Id}_{n-r}. \quad (22)$$

**Proposition 4.1.** *Given  $\mathcal{T}$  an invariant and elliptic torus as above, we define the matrix functions*

$$N_1(\theta) = N(\theta), \quad N_2(\theta) = iN^*(\theta),$$

*and the real matrix*

$$V(\theta) = J D\tau(\theta) G_{D\tau}^{-1}(\theta) + N_1(\theta) B_1(\theta) + N_2(\theta) B_2(\theta) + D\tau(\theta) B_3(\theta), \quad (23)$$

*where*

$$B_1(\theta) = G_{N_2, D\tau}(\theta) G_{D\tau}^{-1}(\theta), \quad (24)$$

$$B_2(\theta) = -G_{N_1, D\tau}(\theta) G_{D\tau}^{-1}(\theta), \quad (25)$$

$$B_3(\theta) = \operatorname{Re}(G_{B_2, B_1}(\theta)). \quad (26)$$

*Then, the columns of the matrices  $D\tau(\theta)$ ,  $V(\theta)$ ,  $N_1(\theta)$  and  $N_2(\theta)$  form a symplectic basis for any  $\theta \in \mathbb{T}^r$ , in the sense that the matrices  $\Omega_{D\tau}(\theta)$ ,  $\Omega_V(\theta)$ ,  $\Omega_{N_i}(\theta)$ ,  $\Omega_{D\tau, N_i}(\theta)$  and  $\Omega_{N_i, V}(\theta)$  vanish, for  $i = 1, 2$ , and*

$$\Omega_{N_2, N_1}(\theta) = \operatorname{Id}_{n-r}, \quad \Omega_{V, D\tau}(\theta) = \operatorname{Id}_r.$$

**Proof.** To obtain the geometric properties associated with the matrices  $D\tau$ ,  $N_1$  and  $N_2$  we proceed as in the proof of lemma 2.3, where we proved that  $\Omega_{D\tau} = 0$ . Let us start studying the matrix  $\Omega_{N_2, N_1}$  by computing

$$\begin{aligned} L_\omega \Omega_{N_2, N_1} &= L_\omega (N_2^\top J N_1) = (L_\omega N_2)^\top J N_1 + N_2^\top J L_\omega N_1 \\ &= (J \text{hess } h(\tau) N_2 + N_2 \Lambda)^\top J N_1 + N_2^\top J (J \text{hess } h(\tau) N_1 - N_1 \Lambda) \\ &= \Lambda N_2^\top J N_1 - N_2^\top J N_1 \Lambda = \Lambda \Omega_{N_2, N_1} - \Omega_{N_2, N_1} \Lambda. \end{aligned}$$

Then, if we expand  $\Omega_{N_2, N_1}$  in Fourier series we obtain

$$(\langle k, \omega \rangle - \lambda_i + \lambda_j) (\widehat{\Omega}_{N_2, N_1})_k^{(i, j)} = 0,$$

where  $(\Omega_{N_2, N_1})^{(i, j)}$  denotes the  $(i, j)$ th entry of  $\Omega_{N_2, N_1}$ . Recalling the non-resonance hypotheses  $\langle k, \omega \rangle - \lambda_i + \lambda_j \neq 0$  (if  $i \neq j$  or  $k \neq 0$ ) we obtain that  $(\widehat{\Omega}_{N_2, N_1})_k^{(i, j)} = 0$ , for all  $k \in \mathbb{Z}^r \setminus \{0\}$ , and  $(\widehat{\Omega}_{N_2, N_1})_0^{(i, j)} = 0$  if  $i \neq j$ , so this matrix is constant and diagonal. Moreover,  $\Omega_{N_2, N_1}^\top = \Omega_{N_2, N_1}^*$  so its entries are real. Finally, using hypothesis (22) we write  $\Omega_{N_2, N_1} = i\Omega_{N^*, N} = -i\Omega_{N, N^*}^\top = \text{Id}_{n-r}$ .

To prove that  $\Omega_{N_1}$  vanishes we compute

$$L_\omega \Omega_{N_1} = -\Lambda \Omega_{N_1} - \Omega_{N_1} \Lambda.$$

Now, the Fourier coefficients of  $\Omega_{N_1}$  satisfy

$$(\langle k, \omega \rangle + \lambda_i + \lambda_j) (\widehat{\Omega}_{N_1})_k^{(i, j)} = 0,$$

so it turns out that all of them vanish (using the non-resonance conditions). Moreover, taking derivatives at equation (20) we obtain

$$L_\omega D\tau = J \text{hess } h D\tau,$$

that together with equation (21), leads to

$$L_\omega \Omega_{D\tau, N_1} = -\Omega_{D\tau, N_1} \Lambda.$$

This implies that  $\Omega_{D\tau, N_1} = 0$ , since the Fourier coefficients of the component functions satisfy

$$(\langle k, \omega \rangle + \lambda_i) (\widehat{\Omega}_{D\tau, N_1})_k^{(i, j)} = 0.$$

Finally, it is easy to see that  $\Omega_{N_2} = -\Omega_{N_1}^*$  and  $\Omega_{D\tau, N_2} = i\Omega_{D\tau, N_1}^*$ , so these matrices also vanish.

Next, we see that the columns of the (real) matrices  $D\tau$ ,  $J D\tau G_{D\tau}^{-1}$ ,  $\text{Re}(N)$  and  $\text{Im}(N)$  form an  $\mathbb{R}$ -basis of  $\mathbb{R}^{2n}$ . To this end, it suffices to check that the columns of  $D\tau$ ,  $J D\tau G_{D\tau}^{-1}$ ,  $N_1$  and  $N_2$  are  $\mathbb{C}$ -independent on  $\mathbb{C}^{2n}$ . Thus, let us consider a linear combination

$$D\tau a + J D\tau G_{D\tau}^{-1} b + N_1 c + N_2 d = 0,$$

for vector functions  $a, b : \mathbb{T}^r \rightarrow \mathbb{C}^r$  and  $c, d : \mathbb{T}^r \rightarrow \mathbb{C}^{n-r}$ . Multiplying by  $D\tau^\top$ ,  $D\tau^\top J$ ,  $N_2^\top J$  and  $N_1^\top J$  and using the geometric properties proved above, we obtain the following system of equations:

$$\underbrace{\begin{pmatrix} G_{D\tau} & 0 & G_{D\tau, N_1} & G_{D\tau, N_2} \\ 0 & -\text{Id}_r & 0 & 0 \\ 0 & -G_{N_2, D\tau} G_{D\tau}^{-1} & \text{Id}_{n-r} & 0 \\ 0 & -G_{N_1, D\tau} G_{D\tau}^{-1} & 0 & -\text{Id}_{n-r} \end{pmatrix}}_{M_1} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad (27)$$

where  $\det M_1 = \det G_{D\tau} \neq 0$ , so we conclude that  $a = b = 0$  and  $c = d = 0$ .

To check that the matrix  $V$  is real, we use the expressions  $N_1^* = -iN_2$  and  $B_1^* = iB_2$ . Then, we compute  $N_1^* B_1^* = -i^2 N_2 B_2 = N_2 B_2$ , thus concluding that  $V^* = V$ .

Finally, the following computations are straightforward:

$$\begin{aligned}\Omega_{D\tau, V} &= -\text{Id}_r + \Omega_{D\tau, N_1} B_1 + \Omega_{D\tau, N_2} B_2 + \Omega_{D\tau} B_3 = -\text{Id}_r, \\ \Omega_{N_1, V} &= -G_{N_1, D\tau} G_{D\tau}^{-1} + \Omega_{N_1} B_1 + \Omega_{N_1, N_2} B_2 + \Omega_{N_1, D\tau} B_3 \\ &= -G_{N_1, D\tau} G_{D\tau}^{-1} - B_2 = 0, \\ \Omega_{N_2, V} &= -G_{N_2, D\tau} G_{D\tau}^{-1} + B_1 = 0, \\ \Omega_V &= (G_{D\tau}^{-1} D\tau^\top J^\top + B_1^\top N_1^\top + B_2^\top N_2^\top + B_3^\top D\tau^\top) J V \\ &= G_{D\tau}^{-1} G_{D\tau, N_1} B_1 + G_{D\tau}^{-1} G_{N_2, D\tau} B_2 + B_3 - B_3^\top \\ &= -G_{B_2, B_1} + G_{B_1, B_2} + B_3 - B_3^\top \\ &= i\text{Im}(G_{B_1, B_2} - G_{B_2, B_1}) = 0.\end{aligned}$$

In the last computation we used that  $V$  is real.  $\square$

**Remark 4.2.** Note that the matrix  $B_3$  can be taken modulo the addition of a symmetric real matrix. This freedom can be used to ask for reducibility also in the ‘central directions’ of the torus. Hence, instead of the matrix  $A_1$  that appears in lemma 4.3 we would obtain its average  $[A_1]_{\mathbb{T}^r}$ . Since this does not give us any significant advantage, we do not resort to this fact.

In the invariant and reducible case, we characterize the action of  $\mathcal{R}$  on  $D\tau(\theta)$ ,  $N_1(\theta)$  and  $N_2(\theta)$  in a very simple way

$$\mathcal{R}(D\tau(\theta)) = 0, \quad \mathcal{R}(N_1(\theta)) = -N_1(\theta)\Lambda, \quad \mathcal{R}(N_2(\theta)) = N_2(\theta)\Lambda. \quad (28)$$

The first expression follows immediately from equation (20)—invariance—and the other ones from equation (21)—reducibility. Moreover, we have the following result for  $V(\theta)$ .

**Lemma 4.3.** *Under the setting of proposition 4.1, we have that*

$$\mathcal{R}(V(\theta)) = D\tau(\theta)A_1(\theta),$$

where  $A_1 : \mathbb{T}^r \rightarrow M_{r \times r}(\mathbb{R})$  is given by the real symmetric matrix

$$A_1(\theta) = G_{D\tau}^{-1}(\theta) D\tau(\theta)^\top [T_1(\theta) + T_2(\theta) + T_2(\theta)^\top] D\tau(\theta) G_{D\tau}^{-1}(\theta), \quad (29)$$

where

$$T_1(\theta) = J^\top \text{hess } h(\tau(\theta)) J - \text{hess } h(\tau(\theta)), \quad (30)$$

$$T_2(\theta) = T_1 J [D\tau(\theta) G_{D\tau}(\theta)^{-1} D\tau(\theta)^\top - \text{Id}] \text{Re}(N_1(\theta) N_2(\theta)^\top).$$

**Proof.** We only have to write the expression for  $\mathcal{R}(V)$  in terms of the previously constructed symplectic basis

$$\mathcal{R}(V) = D\tau A_1 + V A_2 + N_1 A_3 + N_2 A_4, \quad (31)$$

and then to show that  $A_1$  is given by (29) and  $A_2 = A_3 = A_4 = 0$ . First, we use (23) and (28) to express  $\mathcal{R}(V)$  as

$$\mathcal{R}(V) = \mathcal{R}(J D\tau G_{D\tau}^{-1}) + N_1(L_\omega B_1 - \Lambda B_1) + N_2(L_\omega B_2 + \Lambda B_2) + D\tau L_\omega B_3.$$

Then, multiplying both sides of equation (31) by  $V^\top J$ ,  $D\tau^\top J$ ,  $N_2^\top J$ ,  $N_1^\top J$  and using the symplectic properties of the basis we obtain the following expressions:

$$A_1 = L_\omega B_3 + V^\top J \mathcal{R}(J D\tau G_{D\tau}^{-1}), \quad (32)$$

$$A_2 = -D\tau^\top J \mathcal{R}(J D\tau G_{D\tau}^{-1}), \quad (33)$$

$$A_3 = L_\omega B_1 - \Lambda B_1 + N_2^\top J \mathcal{R}(J D\tau G_{D\tau}^{-1}), \quad (34)$$

$$A_4 = L_\omega B_2 + \Lambda B_2 - N_1^\top J \mathcal{R}(J D\tau G_{D\tau}^{-1}). \quad (35)$$

First, introducing  $B_1 = G_{N_2, D\tau} G_{D\tau}^{-1}$  into equation (34), we obtain

$$\begin{aligned} A_3 &= L_\omega(N_2^\top D\tau G_{D\tau}^{-1}) - \Lambda N_2^\top D\tau G_{D\tau}^{-1} + N_2^\top J\mathcal{R}(JD\tau G_{D\tau}^{-1}) \\ &= L_\omega N_2^\top D\tau G_{D\tau}^{-1} + N_2^\top L_\omega(D\tau G_{D\tau}^{-1}) - \Lambda N_2^\top D\tau G_{D\tau}^{-1} \\ &\quad + N_2^\top J L_\omega(JD\tau G_{D\tau}^{-1}) + N_2^\top \text{hess } h JD\tau G_{D\tau}^{-1} \\ &= \underbrace{(L_\omega N_2 - J\text{hess } h N_2 - N_2 \Lambda)^\top}_{\mathcal{R}(N_2)} D\tau G_{D\tau}^{-1} = 0, \end{aligned}$$

where we used the property (28) for  $N_2$ . Recalling that  $N_2 = iN_1^*$  we observe that  $A_3^* = iA_4$  so we also have  $A_4 = 0$ .

Now, we expand the expression for  $\mathcal{R}(JD\tau G_{D\tau}^{-1})$ , obtaining

$$\begin{aligned} \mathcal{R}(JD\tau G_{D\tau}^{-1}) &= \mathcal{R}(JD\tau)G_{D\tau}^{-1} + JD\tau L_\omega(G_{D\tau}^{-1}) \\ &= -\text{hess } h D\tau G_{D\tau}^{-1} - JD\tau G_{D\tau}^{-1}(D\tau^\top J\text{hess } h - D\tau^\top \text{hess } h J)D\tau G_{D\tau}^{-1} \\ &\quad - J\text{hess } h JD\tau G_{D\tau}^{-1} = (\text{Id}_r + JD\tau G_{D\tau}^{-1} D\tau^\top J)T_1 D\tau G_{D\tau}^{-1}, \end{aligned}$$

where we used expression (30) for  $T_1$ . Then, on the one hand we have  $D\tau^\top J\mathcal{R}(JD\tau G_{D\tau}^{-1}) = 0$ —in combination with (33) this implies that  $A_2 = 0$ —and on the other hand we have

$$\begin{aligned} V^\top J\mathcal{R}(JD\tau G_{D\tau}^{-1}) &= (B_3^\top D\tau^\top + B_2^\top N_2^\top + B_1^\top N_1^\top + G_{D\tau}^{-1} D\tau^\top J^\top)J\mathcal{R}(JD\tau G_{D\tau}^{-1}) \\ &= -B_2^\top (L_\omega B_1 - \Lambda B_1) + B_1^\top (L_\omega B_2 + \Lambda B_2) + G_{D\tau}^{-1} D\tau^\top T_1 D\tau G_{D\tau}^{-1}, \end{aligned}$$

where we have used equations (34) and (35) taking into account that  $A_3 = A_4 = 0$ . Finally, we introduce this last expression into (32) and recall that  $B_3 = \text{Re}(G_{B_2, B_1})$  in order to obtain

$$\begin{aligned} A_1 &= \text{Re}(L_\omega(B_2^\top B_1)) - B_2^\top (L_\omega B_1 - \Lambda B_1) + B_1^\top (L_\omega B_2 + \Lambda B_2) + G_{D\tau}^{-1} D\tau^\top T_1 D\tau G_{D\tau}^{-1} \\ &= \text{Re}(L_\omega B_2^\top B_1 - B_2^\top L_\omega B_1 + 2B_2^\top \Lambda B_1) + G_{D\tau}^{-1} D\tau^\top T_1 D\tau G_{D\tau}^{-1}, \end{aligned}$$

where we used that  $(B_1^\top (L_\omega B_2 + \Lambda B_2))^* = -B_2^\top (L_\omega B_1 - \Lambda B_1)$ . Now we replace  $B_1$  and  $B_2$  by equations (24) and (25) respectively, and we expand the expression for  $L_\omega B_1$  and  $L_\omega B_2$  as follows (we also use that  $B_1 = -iB_2^*$ ):

$$\begin{aligned} L_\omega B_2 &= -L_\omega N_1^\top D\tau G_{D\tau}^{-1} - N_1^\top L_\omega(D\tau G_{D\tau}^{-1}) \\ &= \Lambda N_1^\top D\tau G_{D\tau}^{-1} - N_1^\top \text{hess } h J^\top D\tau G_{D\tau}^{-1} - N_1^\top L_\omega(D\tau G_{D\tau}^{-1}) \\ &= -\Lambda B_2 + N_1^\top J T_1 D\tau G_{D\tau}^{-1} - N_1^\top D\tau L_\omega(G_{D\tau}^{-1}) \\ &= -\Lambda B_2 + N_1^\top J T_1 D\tau G_{D\tau}^{-1} - N_1^\top D\tau G_{D\tau}^{-1} D\tau^\top J T_1 D\tau G_{D\tau}^{-1}, \\ L_\omega B_1 &= \Lambda B_1 - N_2^\top J T_1 D\tau G_{D\tau}^{-1} + N_2^\top D\tau G_{D\tau}^{-1} D\tau^\top J T_1 D\tau G_{D\tau}^{-1}. \end{aligned}$$

From these expressions we observe that term  $2B_2^\top \Lambda B_1$  in  $A_1$  is cancelled. Finally, since  $(N_1 N_2^\top)^* = -N_2 N_1^\top = (-N_1 N_2^\top)^\top$ , it turns out that  $\text{Re}(N_1 N_2^\top) = \text{Re}((-N_1 N_2^\top)^\top)$  so we obtain expression (29) for  $A_1$ .  $\square$

Now we have all the ingredients to formally invert the operator  $\mathcal{R}$ .

**Proposition 4.4.** *Under the setting of proposition 4.1, we assume that the matrix  $A_1$  given in (29) satisfies the twist condition  $\det[A_1]_{\mathbb{T}^r} \neq 0$ . Then, given a function  $e : \mathbb{T}^r \rightarrow \mathbb{R}^{2n}$  satisfying  $[D\tau^\top J e]_{\mathbb{T}^r} = 0$ , we obtain a formal solution for the equation*

$$\mathcal{R}(\Delta_\tau(\theta)) = L_\omega \Delta_\tau - J\text{hess } h(\tau) \Delta_\tau = -e(\theta),$$

which is unique up to terms in  $\ker(\mathcal{R}) = \{D\tau A : A \in M_{r \times r}(\mathbb{R})\}$ .

**Proof.** We express the unknown  $\Delta_\tau(\theta)$  in terms of the constructed symplectic basis

$$\Delta_\tau = D\tau\Delta_1 + V\Delta_2 + N_1\Delta_3 + N_2\Delta_4, \quad (36)$$

expand  $\mathcal{R}(\Delta_\tau)$  and project to compute the functions  $\{\Delta_i\}_{i=1,\dots,4}$ . Concretely, we have

$$\begin{aligned} \mathcal{R}(\Delta_\tau) &= \mathcal{R}(D\tau)\Delta_1 + \mathcal{R}(V)\Delta_2 + \mathcal{R}(N_1)\Delta_3 + \mathcal{R}(N_2)\Delta_4 \\ &\quad + D\tau L_\omega\Delta_1 + V L_\omega\Delta_2 + N_1 L_\omega\Delta_3 + N_2 L_\omega\Delta_4 \\ &= D\tau(L_\omega\Delta_1 + A_1\Delta_2) + V L_\omega\Delta_2 + N_1(L_\omega\Delta_3 - \Lambda\Delta_3) + N_2(L_\omega\Delta_4 + \Lambda\Delta_4). \end{aligned}$$

Multiplying both sides of this expression by  $V^\top J$ ,  $D\tau^\top J$ ,  $N_2^\top J$  and  $N_1^\top J$ , we obtain the following four cohomological equations:

$$L_\omega\Delta_1 + A_1\Delta_2 = -V^\top J e, \quad (37)$$

$$L_\omega\Delta_2 = D\tau^\top J e, \quad (38)$$

$$L_\omega\Delta_3 - \Lambda\Delta_3 = -N_2^\top J e, \quad (39)$$

$$L_\omega\Delta_4 + \Lambda\Delta_4 = N_1^\top J e. \quad (40)$$

As  $[D\tau^\top J e]_{\mathbb{T}^r} = 0$ , the solution of equation (38) is unique, up to an arbitrary average  $[\Delta_2]_{\mathbb{T}^r}$ , provided that the non-resonance condition (2) holds. Then, using the non-degeneracy condition  $\det[A_1]_{\mathbb{T}^r} \neq 0$ , we choose

$$[\Delta_2]_{\mathbb{T}^r} = [A_1]_{\mathbb{T}^r}^{-1}([-V^\top J e]_{\mathbb{T}^r} - [\tilde{A}_1 \tilde{\Delta}_2]_{\mathbb{T}^r}) \quad (41)$$

in such a way that  $[A_1\Delta_2 + V^\top J e]_{\mathbb{T}^r} = 0$  so we have a unique solution for  $\Delta_1$  up to the freedom of fixing  $[\Delta_1]_{\mathbb{T}^r}$ . Actually, it is easy to check that (39) and (40) have a unique solution for  $\Delta_3$  and  $\Delta_4$  provided that the non-resonance condition (9) is fulfilled. Moreover, since  $e$  is a real function, we conclude that  $\Delta_3^* = i\Delta_4$  and this allows us to guarantee that expression (36) is also real.  $\square$

**Remark 4.5.** Later on—see computations in (93)—we show that if  $\tau$  also parametrizes an approximately invariant torus and  $e$  is the error of invariance, then the compatibility condition  $[D\tau^\top J e]_{\mathbb{T}^r} = 0$  is automatically fulfilled.

**Proposition 4.6.** *Under the setting of proposition 4.1, given a function  $\hat{R} : \mathbb{T}^r \mapsto M_{2n \times (n-r)}(\mathbb{C})$ , we obtain a solution for the equation*

$$S(\Delta_N, \Delta_\Lambda) = \mathcal{R}(\Delta_N) + N\Delta_\Lambda + \Delta_N\Lambda = -\hat{R},$$

which is unique for  $\Delta_\Lambda$  and for  $\Delta_N$  up to terms in  $\ker(S) = \{ND : D = \text{diag}(d), d \in \mathbb{C}^{n-r}\}$ .

**Proof.** As before, we write the solution  $\Delta_N$  of this equation in terms of the symplectic basis as

$$\Delta_N = D\tau P_1 + V P_2 + N_1 P_3 + N_2 P_4.$$

Then, we compute the action of  $S$  on the pair  $(\Delta_N, \Delta_\Lambda)$ , thus obtaining

$$\begin{aligned} S(\Delta_N, \Delta_\Lambda) &= \mathcal{R}(\Delta_N) + N_1\Delta_\Lambda + \Delta_N\Lambda \\ &= \mathcal{R}(D\tau)P_1 + \mathcal{R}(V)P_2 + \mathcal{R}(N_1)P_3 + \mathcal{R}(N_2)P_4 \\ &\quad + D\tau L_\omega P_1 + V L_\omega P_2 + N_1 L_\omega P_3 + N_2 L_\omega P_4 + N_1\Delta_\Lambda \\ &\quad + D\tau P_1\Lambda + V P_2\Lambda + N_1 P_3\Lambda + N_2 P_4\Lambda \\ &= D\tau(L_\omega P_1 + P_1\Lambda + A_1 P_2) + V(L_\omega P_2 + P_2\Lambda) \\ &\quad + N_1(L_\omega P_3 + P_3\Lambda - \Lambda P_3 + \Delta_\Lambda) + N_2(L_\omega P_4 + P_4\Lambda + \Lambda P_4) = \hat{R}. \end{aligned}$$

If we multiply this expression by  $V^\top J$ ,  $D\tau^\top J$ ,  $N_2^\top J$  and  $N_1^\top J$ , we end up with the following four cohomological equations:

$$L_\omega P_1 + P_1 \Lambda + A_1 P_2 = -V^\top J \hat{R}, \quad (42)$$

$$L_\omega P_2 + P_2 \Lambda = D\tau^\top J \hat{R}, \quad (43)$$

$$L_\omega P_3 + P_3 \Lambda - \Lambda P_3 = -N_2^\top J \hat{R} - \Delta_\Lambda, \quad (44)$$

$$L_\omega P_4 + P_4 \Lambda + \Lambda P_4 = N_1^\top J \hat{R}. \quad (45)$$

Let us observe that, under the assumed non-resonance conditions (7) and (8), the only unavoidable resonances are those in the diagonal of the average of equation (44), so we require that the diagonal of the average of the right-hand side of this equation vanishes. This is attained by fixing the correction of the normal eigenvalues  $\Delta_\Lambda = -\text{diag}[N_2^\top J \hat{R}]_{\mathbb{T}^r}$ . Therefore, we obtain unique solutions for  $P_1$ ,  $P_2$ ,  $P_3$ ,  $P_4$  and  $\Delta_\Lambda$ —modulo terms in  $\text{diag}[P_3]_{\mathbb{T}^r}$ —of this system of equations.  $\square$

**Remark 4.7.** We will see that if  $\hat{R}$  corresponds to the error in reducibility, as defined in equation (15), then the geometry imposes that the correction  $\Delta_\Lambda$  is a pure imaginary diagonal matrix, thus preserving the elliptic normal behaviour—see computations in (96).

## 5. One step of the Newton method

In this section we perform one step of the Newton method to correct an approximately invariant and elliptic torus. To this end, we follow the scheme presented in section 4.2 for the case of a true elliptic invariant torus. The main difficulty is that we have to handle with ‘noise’ introduced by the approximately invariant and reducible character.

**Proposition 5.1.** *Let us consider a Hamiltonian  $h : U \subset \mathbb{R}^{2n} \rightarrow \mathbb{R}$ , where  $U$  is an open set, and a vector of basic frequencies  $\omega \in \mathbb{R}^r$ . Let us assume that the following hypotheses hold:*

- $H_1$  *The Hamiltonian  $h$  is real analytic and can be holomorphically extended to some complex neighbourhood  $\mathcal{U}$  of  $U$ . Moreover, we assume that  $\|h\|_{C^3, \mathcal{U}} \leq \sigma_0$ .*
- $H_2$  *There exists an approximately invariant and elliptic torus in the sense of definitions 2.5 and 2.9, i.e. we have an embedding  $\tau$ , a matrix function  $N$  and approximated normal eigenvalues  $\Lambda = \text{diag}(i\lambda)$ , with  $\lambda \in \mathbb{R}^{n-r}$ , satisfying*

$$L_\omega \tau(\theta) = J \text{grad } h(\tau(\theta)) + e(\theta), \quad (46)$$

$$L_\omega N(\theta) = J \text{hess } h(\tau(\theta)) N(\theta) - N(\theta) \Lambda + R(\theta), \quad (47)$$

*for certain error functions  $e$  and  $R$ , where the functions  $\tau$  and  $N$  are analytic and can be holomorphically extended to  $\Delta(\rho)$  for certain  $0 < \rho < 1$ , satisfying  $\tau(\Delta(\rho)) \subset \mathcal{U}$ . Assume also that we have constants  $\sigma_1, \sigma_2$  such that*

$$\|D\tau\|_\rho, \|N\|_\rho, \|G_{D\tau}^{-1}\|_\rho, \|G_{N, N^*}^{-1}\|_\rho < \sigma_1, \quad \text{dist}(\tau(\Delta(\rho)), \partial\mathcal{U}) > \sigma_2 > 0.$$

- $H_3$  *We have  $\text{diag}[\Omega_{N, N^*}]_{\mathbb{T}^r} = i\text{Id}_{n-r}$ .*

- $H_4$  *The real symmetric matrix  $A_1$  given by*

$$A_1(\theta) = G_{D\tau}^{-1}(\theta) D\tau(\theta)^\top [T_1(\theta) + T_2(\theta) + T_2(\theta)^\top] D\tau(\theta) G_{D\tau}^{-1}(\theta), \quad (48)$$

*where*

$$T_1(\theta) = J^\top \text{hess } h(\tau(\theta)) J - \text{hess } h(\tau(\theta)), \quad (49)$$

$$T_2(\theta) = T_1(\theta) J [D\tau(\theta) G_{D\tau}(\theta)^{-1} D\tau(\theta)^\top - \text{Id}] \text{Re}(iN(\theta) N^*(\theta)^\top), \quad (50)$$

*satisfies the non-degeneracy (twist) condition  $|[A_1]_{\mathbb{T}^r}^{-1}| < \sigma_1$ .*

H<sub>5</sub> There exist constants  $\sigma_3, \sigma_4$  such that the approximated normal frequencies satisfy

$$0 < \frac{\sigma_3}{2} < |\lambda_i| < \frac{\sigma_4}{2}, \quad 0 < \sigma_3 < |\lambda_i \pm \lambda_j|,$$

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

H<sub>6</sub> The basic frequencies  $\omega \in \mathbb{R}^r$  and the normal frequencies  $\lambda \in \mathbb{R}^{n-r}$  satisfy Diophantine conditions (4) and (9) of  $(\gamma, \nu)$ -type, for certain  $0 < \gamma < 1$  and  $\nu > r-1$ .

Then, there exists a constant  $\bar{\alpha} > 1$  depending on  $\nu, r, n, |\omega|, \sigma_0, \sigma_1, \sigma_2, \sigma_3$  and  $\sigma_4$  such that if the following bounds are satisfied:

$$\frac{\bar{\alpha}}{\gamma^4 \delta^{4\nu-1}} \left( \frac{\|e\|_\rho}{\delta} + \|R\|_\rho \right) < \min\{1, \sigma_1 - \sigma^*\}, \quad (51)$$

$$\text{dist}(\tau(\Delta(\rho)), \partial\mathcal{U}) - \frac{\bar{\alpha}}{\gamma^2 \delta^{2\nu}} \|e\|_\rho > \sigma_2, \quad (52)$$

$$\min_{i \neq j} |\lambda_i \pm \lambda_j| - \frac{\bar{\alpha}}{\gamma^2 \delta^{2\nu-1}} \left( \frac{\|e\|_\rho}{\delta} + \|R\|_\rho \right) > \sigma_3, \quad (53)$$

$$\min_i |\lambda_i| - \frac{\bar{\alpha}}{\gamma^2 \delta^{2\nu-1}} \left( \frac{\|e\|_\rho}{\delta} + \|R\|_\rho \right) > \frac{\sigma_3}{2}, \quad (54)$$

$$\max_i |\lambda_i| + \frac{\bar{\alpha}}{\gamma^2 \delta^{2\nu-1}} \left( \frac{\|e\|_\rho}{\delta} + \|R\|_\rho \right) < \frac{\sigma_4}{2}, \quad (55)$$

where

$$\sigma^* = \max \left\{ \|D\tau\|_\rho, \|N\|_\rho, \|G_{D\tau}^{-1}\|_\rho, \|G_{N,N^*}^{-1}\|_\rho, |[A_1]_{\mathbb{T}^r}^{-1}| \right\},$$

for some  $0 < \delta < \rho/4$ , then we have an approximately invariant and elliptic torus  $\tilde{\mathcal{T}}$  for  $X_h$  of the same basic frequencies  $\omega$ . This means that we have an embedding  $\bar{\tau} = \tau + \Delta_\tau$ , with  $\bar{\tau}(\mathbb{T}^r) = \tilde{\mathcal{T}}$ , a matrix function  $\bar{N} = N + \Delta_N$ , which are analytic in  $\Delta(\rho - 2\delta)$  and  $\Delta(\rho - 4\delta)$ , respectively, and approximated normal eigenvalues  $\bar{\Lambda} = \text{diag}(i\bar{\lambda}) = \Lambda + \Delta_\Lambda$ , with  $\bar{\lambda} \in \mathbb{R}^{n-r}$ , such that

$$L_\omega \bar{\tau}(\theta) = J \text{grad } h(\bar{\tau}(\theta)) + \bar{e}(\theta),$$

$$L_\omega \bar{N}(\theta) = J \text{hess } h(\bar{\tau}(\theta)) \bar{N}(\theta) - \bar{N}(\theta) \bar{\Lambda} + \bar{R}(\theta).$$

In addition, the following estimates hold:

$$\|\Delta_\tau\|_{\rho-2\delta} \leq \frac{\bar{\alpha}}{\gamma^2 \delta^{2\nu}} \|e\|_\rho, \quad (56)$$

$$\|\bar{e}\|_{\rho-3\delta} \leq \frac{\bar{\alpha}}{\gamma^4 \delta^{4\nu-1}} \left( \frac{\|e\|_\rho}{\delta} + \|R\|_\rho \right) \|e\|_\rho, \quad (57)$$

$$|\Delta_\Lambda| \leq \frac{\bar{\alpha}}{\gamma^2 \delta^{2\nu-1}} \left( \frac{\|e\|_\rho}{\delta} + \|R\|_\rho \right), \quad (58)$$

$$\|\Delta_N\|_{\rho-4\delta} \leq \frac{\bar{\alpha}}{\gamma^4 \delta^{4\nu-1}} \left( \frac{\|e\|_\rho}{\delta} + \|R\|_\rho \right), \quad (59)$$

$$\|\bar{R}\|_{\rho-4\delta} \leq \frac{\bar{\alpha}}{\gamma^8 \delta^{8\nu-2}} \left( \frac{\|e\|_\rho}{\delta} + \|R\|_\rho \right)^2, \quad (60)$$

$$\|G_{D\tau}^{-1} - G_{D\tau}^{-1}\|_{\rho-3\delta} \leq \frac{\bar{\alpha}}{\gamma^2 \delta^{2\nu+1}} \|e\|_\rho, \quad (61)$$

$$\|G_{\bar{N}, \bar{N}^*}^{-1} - G_{N, N^*}^{-1}\|_{\rho-4\delta} \leq \frac{\bar{\alpha}}{\gamma^4 \delta^{4\nu-1}} \left( \frac{\|e\|_\rho}{\delta} + \|R\|_\rho \right), \quad (62)$$

$$|[\bar{A}_1]_{\mathbb{T}^r}^{-1} - [A_1]_{\mathbb{T}^r}^{-1}| \leq \frac{\bar{\alpha}}{\gamma^4 \delta^{4\nu-1}} \left( \frac{\|e\|_\rho}{\delta} + \|R\|_\rho \right). \quad (63)$$



Furthermore, the new objects satisfy the following conditions:

$$\text{dist}(\bar{\tau}(\Delta(\rho - 2\delta)), \partial\mathcal{U}) > \sigma_2, \quad \frac{\sigma_3}{2} < |\bar{\lambda}_j| < \frac{\sigma_4}{2}, \quad \sigma_3 < |\bar{\lambda}_i \pm \bar{\lambda}_j|, \quad (64)$$

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

$$\max \left\{ \|D\bar{\tau}\|_{\rho-3\delta}, \|\bar{N}\|_{\rho-4\delta}, \|G_{D\bar{\tau}}^{-1}\|_{\rho-3\delta}, \|G_{\bar{N}, \bar{N}^*}^{-1}\|_{\rho-4\delta}, \|[\bar{A}_1]_{\mathbb{T}^r}^{-1}\| \right\} < \sigma_1, \quad (65)$$

where  $\bar{A}_1$  corresponds to formulae (48), (49) and (50) for  $\bar{\tau}$  and  $\bar{N}$ . Moreover, the columns of  $\bar{N}$  are normalized in such a way that  $\text{diag}[\Omega_{\bar{N}, \bar{N}^*}]_{\mathbb{T}^r} = i\text{Id}_{n-r}$ .

To prove this result, we first construct an approximately symplectic basis along the torus following the ideas of section 4.2. This is done in proposition 5.3. The geometric properties of this basis allow us to approximately invert the operators  $\mathcal{R}$  and  $\mathcal{S}$ —given by (17) and (18), respectively—as it is required to obtain the iterative result of proposition 5.1. Basically, it turns out that the solutions of the cohomological equations derived in section 4.2 are enough to get the desired result. Before that, we state the following standard result that allows us to control the small divisors.

**Lemma 5.2 (Rüssmann estimates).** *Let  $g : \mathbb{T}^r \rightarrow \mathbb{C}$  be an analytic function on  $\Delta(\rho)$  and bounded in the closure. Given  $\omega \in \mathbb{R}^r \setminus \{0\}$  and  $d \in \mathbb{R} \setminus \{0\}$  we consider the sets of complex numbers  $\{d_k^0\}_{k \in \mathbb{Z}^r \setminus \{0\}}$ ,  $\{d_k^1\}_{k \in \mathbb{Z}^r}$  given by  $d_k^0 = \langle k, \omega \rangle$ ,  $d_k^1 = \langle k, \omega \rangle + d$ , satisfying*

$$|d_k^0|, |d_k^1| \geq \gamma/|k|_1^v, \quad \forall k \in \mathbb{Z}^r \setminus \{0\}$$

for certain  $\gamma > 0$  and  $v > r - 1$ . Then, the functions  $f^0$  and  $f^1$  whose Fourier coefficients are given by

$$\begin{aligned} \hat{f}_k^0 &= \hat{g}_k/d_k^0, & k \in \mathbb{Z}^r \setminus \{0\}, & \quad \hat{f}_0^0 = 0, \\ \hat{f}_k^1 &= \hat{g}_k/d_k^1, & k \in \mathbb{Z}^r, \end{aligned}$$

satisfy

$$\|f^0\|_{\rho-\delta} \leq \frac{\alpha_0}{\gamma\delta^v} \|g\|_{\rho}, \quad \|f^1\|_{\rho-\delta} \leq \left( \frac{1}{|d|} + \frac{\alpha_0}{\gamma\delta^v} \right) \|g\|_{\rho},$$

for any  $\delta \in (0, \min\{1, \rho\})$ , where  $\alpha_0 \geq 1$  is a constant depending on  $r$  and  $v$ .

**Proof.** We can control the functions  $\tilde{f}^i(\theta) = f^i(\theta) - [f^i]_{\mathbb{T}^r}$  as

$$\|\tilde{f}^i\|_{\rho-\delta} \leq \sum_{k \in \mathbb{Z}^r \setminus \{0\}} \frac{|\hat{g}_k|}{|d_k^i|} e^{|k|_1(\rho-\delta)} \leq \left( \sum_{k \in \mathbb{Z}^r \setminus \{0\}} |\hat{g}_k|^2 e^{2|k|_1\rho} \right)^{1/2} \left( \sum_{k \in \mathbb{Z}^r \setminus \{0\}} \frac{1}{|d_k^i|^2} e^{-2|k|_1\delta} \right)^{1/2},$$

for  $i = 0, 1$ , where we used Cauchy–Schwarz inequality. On the one hand, it is not difficult to see—using Bessel’s inequality, see details in [62]—that the first term can be bounded by

$$\sum_{k \in \mathbb{Z}^r \setminus \{0\}} |\hat{g}_k|^2 e^{2|k|_1\rho} \leq 2^r \|\tilde{g}\|_{\rho}^2,$$

and on the other hand, the second term is controlled by estimating the sum

$$\sum_{k \in \mathbb{Z}^r \setminus \{0\}} \frac{1}{|d_k^i|^2} e^{-2|k|_1\delta} = \sum_{l=1}^{\infty} \left( \sum_{\substack{k \in \mathbb{Z}^r \setminus \{0\} \\ |k|_1 \leq l}} \frac{1}{|d_k^i|^2} \right) (e^{-2l\delta} - e^{-2(l+1)\delta}). \quad (66)$$

Now, we study in detail the case of  $d_k^1$  (the case of  $d_k^0$  is analogous). First, we observe that the divisors  $d_k^1 = \langle k, \omega \rangle + d$  satisfy  $d_{k_1} \neq d_{k_2}$  if  $k_1 \neq k_2$ . Then, given  $l \in \mathbb{N}$ , we define

$$\mathcal{D}_l = \{k \in \mathbb{Z}^r \setminus \{0\} : |k|_1 \leq l \text{ and } d_k^1 > 0\}$$

and we sort the divisors according to  $0 < d_{k_1} < \dots < d_{k_{\#\mathcal{D}_l}}$  with  $k_j \in \mathcal{D}_l$ , for  $j = 1, \dots, \#\mathcal{D}_l$ . Then, we observe that (since  $|k_j - k_{j-1}| \leq 2l$ )

$$d_{k_j}^1 - d_{k_{j-1}}^1 = |\langle k_j - k_{j-1}, \omega \rangle| \geq d_{2l, \min}^0, \quad (67)$$

where we have introduced the notation

$$d_{l, \min}^i = \min_{\substack{k \in \mathbb{Z}^r \setminus \{0\} \\ |k|_1 \leq l}} |d_k^i|.$$

From expression (67) we obtain recursively

$$d_{k_j}^1 = d_{k_{j-1}}^1 + d_{k_j}^1 - d_{k_{j-1}}^1 \geq d_{k_{j-1}}^1 + d_{2l, \min}^0 \geq d_{l, \min}^1 + (j-1)d_{2l, \min}^0.$$

Then, using that  $d_{2l, \min}^0 \geq \gamma/(2l)^\nu$  and  $d_{l, \min}^1 \geq \gamma/l^\nu$ , we have

$$\sum_{j=1}^{\#\mathcal{D}_l} \frac{1}{(d_{k_j}^1)^2} \leq \sum_{j=1}^{\#\mathcal{D}_l} \frac{1}{(d_{l, \min}^1 + (j-1)d_{2l, \min}^0)^2} \leq \sum_{j=1}^{\infty} \frac{l^{2\nu}}{\gamma^2(1+(j-1)2^{-\nu})^2} \leq \frac{\alpha(\nu)}{\gamma^2} l^{2\nu},$$

and using a similar argument for  $d_{k_j}^1 < 0$ , we obtain

$$\sum_{\substack{k \in \mathbb{Z}^r \setminus \{0\} \\ |k|_1 \leq l}} \frac{1}{|d_k^1|^2} \leq \frac{2\alpha(\nu)}{\gamma^2} l^{2\nu},$$

so we can control the sum (66) as follows:

$$\sum_{k \in \mathbb{Z}^r \setminus \{0\}} \frac{1}{|d_k^1|^2} e^{-2|k|_1 \delta} \leq \sum_{l=1}^{\infty} \frac{2\delta\alpha(\nu)}{\gamma^2} \int_l^{l+1} x^{2\nu} e^{-2\delta x} dx \leq \frac{\alpha(\nu)}{\gamma^2(2\delta)^{2\nu}} \Gamma(2\nu+1).$$

Combining the obtained expressions—and using that  $|\llbracket f^1 \rrbracket_{\mathbb{T}^r}| = |\hat{g}_0|/|d|$ —we end up with the stated estimates.  $\square$

**Proposition 5.3.** *Under the notations and assumptions of proposition 5.1, we define the matrix functions*

$$N_1(\theta) = N(\theta), \quad N_2(\theta) = iN^*(\theta),$$

and the real analytic matrix  $V(\theta)$  given by (23)–(26). Then, for any  $0 < \delta < \rho/2$  the following estimates hold:

$$\|\Omega_{D\tau}\|_{\rho-2\delta} \leq \frac{\hat{\alpha}}{\gamma\delta^{\nu+1}} \|e\|_{\rho}, \quad (68)$$

$$\|\Omega_{N_i}\|_{\rho-\delta} \leq \frac{\hat{\alpha}}{\gamma\delta^{\nu}} \|R\|_{\rho}, \quad (69)$$

$$\|\Omega_{D\tau, N_i}\|_{\rho-2\delta} \leq \frac{\hat{\alpha}}{\gamma\delta^{\nu}} \left( \frac{\|e\|_{\rho}}{\delta} + \|R\|_{\rho} \right), \quad (70)$$

$$\|\Omega_{N_2, N_1} - \text{Id}_{n-r}\|_{\rho-\delta} \leq \frac{\hat{\alpha}}{\gamma\delta^{\nu}} \|R\|_{\rho}, \quad (71)$$

$$\|\Omega_{V, D\tau} - \text{Id}_r\|_{\rho-2\delta} \leq \frac{\hat{\alpha}}{\gamma\delta^{\nu}} \left( \frac{\|e\|_{\rho}}{\delta} + \|R\|_{\rho} \right), \quad (72)$$

$$\|\Omega_{V, N_i}\|_{\rho-2\delta} \leq \frac{\hat{\alpha}}{\gamma\delta^{\nu}} \left( \frac{\|e\|_{\rho}}{\delta} + \|R\|_{\rho} \right), \quad (73)$$

$$\|\Omega_V\|_{\rho-2\delta} \leq \frac{\hat{\alpha}}{\gamma\delta^{\nu}} \left( \frac{\|e\|_{\rho}}{\delta} + \|R\|_{\rho} \right), \quad (74)$$

for  $i = 1, 2$ , where  $\hat{\alpha} > 1$  is a constant depending on  $v, r, n, |\omega|, \sigma_0, \sigma_1, \sigma_3$  and  $\sigma_4$ . Furthermore, if the errors  $\|e\|_\rho$  and  $\|R\|_\rho$  satisfy

$$\frac{\hat{\alpha}}{\gamma\delta^v} \left( \frac{\|e\|_\rho}{\delta} + \|R\|_\rho \right) \leq \frac{1}{2}, \quad (75)$$

then the columns of  $D\tau(\theta), V(\theta), N_1(\theta), N_2(\theta)$  form an approximately symplectic basis for every  $\theta \in \mathbb{T}^r$ . In addition, it turns out that the action of the operator  $\mathcal{R}$  given in (17) on  $V$  is expressed in terms of this basis as

$$\mathcal{R}(V(\theta)) = D\tau(\theta)(A_1(\theta) + A_1^+(\theta)) + V(\theta)A_2^+(\theta) + N_1(\theta)A_3^+(\theta) + N_2(\theta)A_4^+(\theta), \quad (76)$$

where  $A_1$  is the matrix (48) and  $A_1^+, A_2^+, A_3^+$  and  $A_4^+$  satisfy the estimate

$$\|A_i^+\|_{\rho-2\delta} \leq \frac{\hat{\alpha}}{\gamma\delta^{v+1}} \left( \frac{\|e\|_\rho}{\delta} + \|R\|_\rho \right), \quad (77)$$

for  $i = 1, \dots, 4$ .

**Proof.** For the sake of simplicity, we redefine (enlarge) the constant  $\hat{\alpha}$  along the proof to meet the different conditions given in the statement. For example, we observe that there exists a constant  $\hat{\alpha} > 0$ , depending on  $r, n, |\omega|, \sigma_0$  and  $\sigma_1$ , such that

$$\|B_i\|_\rho, \|T_1\|_\rho, \|T_2\|_\rho, \|A_1\|_\rho, \|V\|_\rho \leq \hat{\alpha}, \quad \|L_\omega B_i\|_{\rho-\delta} \leq \frac{\hat{\alpha}}{\delta}, \quad (78)$$

for  $i = 1, 2, 3$ —we recall that  $T_1$  and  $T_2$  are given in (49) and (50), respectively. Now we take derivatives at both sides of the approximated invariance equation in (46) and we read the reducibility equations in (47) for  $N_1$  and  $N_2$

$$\begin{aligned} L_\omega D\tau &= J \text{hess } h(\tau) D\tau + De, \\ L_\omega N_1 &= J \text{hess } h(\tau) N_1 - N_1 \Lambda + R, \\ L_\omega N_2 &= J \text{hess } h(\tau) N_2 + N_2 \Lambda + iR^*. \end{aligned} \quad (79)$$

Using the previous expressions, we compute the derivative  $L_\omega$  of the matrices  $\Omega_{D\tau}, \Omega_{N_1}, \Omega_{D\tau, N_1}$  and  $\Omega_{N_2, N_1}$  thus obtaining

$$L_\omega(\Omega_{D\tau}) = \Omega_{De, D\tau} + \Omega_{D\tau, De}, \quad (80)$$

$$L_\omega(\Omega_{N_1}) = -\Lambda \Omega_{N_1} - \Omega_{N_1} \Lambda + \Omega_{R, N_1} + \Omega_{N_1 R}, \quad (81)$$

$$L_\omega(\Omega_{D\tau, N_1}) = -\Omega_{D\tau, N_1} \Lambda + \Omega_{De, N_1} + \Omega_{D\tau, R}, \quad (82)$$

$$L_\omega(\Omega_{N_2, N_1}) = \Lambda \Omega_{N_2, N_1} - \Omega_{N_2, N_1} \Lambda + i\Omega_{R^*, N_1} + \Omega_{N_2, R}. \quad (83)$$

First, we get estimate (68) for  $\Omega_{D\tau}$  by applying lemma 5.2 to the  $(i, j)$ -component of  $\Omega_{D\tau}$  obtained from (80), i.e. taking  $d_k^0 = \langle \omega, k \rangle$  and  $g = -i(\Omega_{De, D\tau} + \Omega_{D\tau, De})^{(i, j)}$  that (using Cauchy estimates) is analytic in  $\Delta(\rho - \delta)$ . Moreover, since  $[\Omega_{D\tau}]_{\mathbb{T}^r} = 0$  (see remark 2.2), we obtain

$$\|\Omega_{D\tau}\|_{\rho-2\delta} \leq \frac{\alpha_0}{\gamma\delta^v} \|g\|_{\rho-\delta} \leq \frac{\hat{\alpha}}{\gamma\delta^{v+1}} \|e\|_\rho.$$

Then, we proceed in a similar way to get (69) for  $N_1$ , by applying lemma 5.2 to the  $(i, j)$ -component of  $\Omega_{N_1}$  obtained from (81), i.e. taking  $d_k^1 = \langle \omega, k \rangle + \lambda_i + \lambda_j$  and  $g = -i(\Omega_{R, N_1} + \Omega_{N_1 R})^{(i, j)}$ , analytic in  $\Delta(\rho)$ . To bound the average of  $\Omega_{N_1}$ , we use hypothesis  $H_5$  of proposition 5.1,

$$\|\Omega_{N_1}\|_{\rho-\delta} \leq \left( \frac{1}{\min_{i,j} |\lambda_i + \lambda_j|} + \frac{\alpha_0}{\gamma\delta^v} \right) \|g\|_\rho \leq \left( \frac{1}{\sigma_3} + \frac{\alpha_0}{\gamma\delta^v} \right) \leq \frac{\hat{\alpha}}{\gamma\delta^v} \|R\|_\rho.$$

Analogous computations from equations (82) and (83) allow us to obtain estimate (70) for  $N_1$  and (71). Of course, to obtain (71) we resort to the hypothesis  $\text{diag}[\Omega_{N,N^*}]_{\mathbb{T}^r} = i\text{Id}_{n-r}$  in  $H_3$  of proposition 5.1. Estimates (69) and (70) corresponding to  $N_2$  are straightforward using that  $\Omega_{N_2} = -\Omega_{N_1}^*$  and  $\Omega_{D\tau, N_2} = i\Omega_{D\tau, N_1}^*$ .

Next we show that the columns of  $D\tau$ ,  $JD\tau G_{D\tau}^{-1}$ ,  $\text{Re}(N_1)$  and  $\text{Im}(N_1)$  form an  $\mathbb{R}$ -basis of  $\mathbb{R}^{2n}$ . As in the proof of proposition 4.1, we consider a linear combination

$$D\tau a + JD\tau G_{D\tau}^{-1}b + N_1c + N_2d = 0,$$

for functions  $a, b : \mathbb{T}^r \rightarrow \mathbb{C}^r$  and  $c, d : \mathbb{T}^r \rightarrow \mathbb{C}^{n-r}$ . We project this equation multiplying by  $D\tau^\top$ ,  $D\tau^\top J$ ,  $N_2^\top J$  and  $N_1^\top J$ , thus obtaining

$$\left( M_1 + \underbrace{\begin{pmatrix} 0 & \Omega_{D\tau} G_{D\tau}^{-1} & 0 & 0 \\ \Omega_{D\tau} & 0 & \Omega_{D\tau, N_1} & \Omega_{D\tau, N_2} \\ \Omega_{N_2, D\tau} & 0 & \Omega_{N_2, N_1} - \text{Id}_{n-r} & \Omega_{N_2} \\ \Omega_{N_1, D\tau} & 0 & \Omega_{N_1} & \Omega_{N_1, N_2} + \text{Id}_{n-r} \end{pmatrix}}_{M_2} \right) \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix},$$

where  $M_1$  is the same matrix that appears in equation (27). Now, we have to invert the matrix  $M_1 + M_2 = M_1(\text{Id} + M_1^{-1}M_2)$ , where

$$M_1^{-1} = \begin{pmatrix} G_{D\tau}^{-1} & M_{1,2} & -G_{D\tau}^{-1}G_{D\tau, N_1} & G_{D\tau}^{-1}G_{D\tau, N_2} \\ 0 & -\text{Id}_r & 0 & 0 \\ 0 & -G_{N_2, D\tau}G_{D\tau}^{-1} & \text{Id}_{n-r} & 0 \\ 0 & G_{N_1, D\tau}G_{D\tau}^{-1} & 0 & -\text{Id}_{n-r} \end{pmatrix},$$

with  $M_{1,2} = G_{D\tau}^{-1}(G_{D\tau, N_1}G_{N_2, D\tau} - G_{D\tau, N_2}G_{N_1, D\tau})G_{D\tau}^{-1}$ , so it is clear that  $\|M_1^{-1}\|_\rho \leq \hat{\alpha}$ . By means of Neumann series we obtain

$$\|(\text{Id} + M_1^{-1}M_2)^{-1}\|_{\rho-2\delta} \leq \frac{1}{1 - \|M_1^{-1}M_2\|_{\rho-2\delta}},$$

which is well posed since (using bounds (68)–(71) and applying hypothesis (75))

$$\|M_1^{-1}M_2\|_{\rho-2\delta} \leq \frac{\hat{\alpha}}{\gamma\delta^v} \left( \frac{\|e\|_\rho}{\delta} + \|R\|_\rho \right) \leq \frac{1}{2}.$$

Then, it must be  $a = b = 0$  and  $c = d = 0$  along  $\mathbb{T}^r$ .

Now, we consider the basis defined by the columns of  $D\tau$ ,  $V$ ,  $N_1$  and  $N_2$ , where  $V$  is given by (23)–(26), and we characterize the fact that the new basis is approximately symplectic. It is straightforward to compute

$$\begin{aligned} \Omega_{D\tau, V} &= -\text{Id}_r + \Omega_{D\tau, N_1}B_1 + \Omega_{D\tau, N_2}B_2 + \Omega_{D\tau}B_3, \\ \Omega_{N_1, V} &= -G_{N_1, D\tau}G_{D\tau}^{-1} + \Omega_{N_1}B_1 + \Omega_{N_1, N_2}B_2 + \Omega_{N_1, D\tau}B_3 \\ &= \Omega_{N_1}B_1 + (\Omega_{N_1, N_2} + \text{Id}_{n-r})B_2 + \Omega_{N_1, D\tau}B_3, \\ \Omega_V &= B_3^\top(\Omega_{D\tau, V} + \text{Id}_r) + B_1^\top\Omega_{N_1, V} + B_2^\top\Omega_{N_2, V} + G_{D\tau}^{-1}\Omega_{D\tau}G_{D\tau}^{-1}, \end{aligned}$$

and  $\Omega_{N_2, V} = i\Omega_{N_1, V}^*$ . Then, estimates (72)–(74) follow from (68)–(71) and (78).

Let us characterize the action of the linear operator  $\mathcal{R}$  on the elements of this basis. By hypothesis, we immediately have that

$$\mathcal{R}(D\tau) = De, \quad \mathcal{R}(N_1) = -N_1\Lambda + R, \quad \mathcal{R}(N_2) = N_2\Lambda + iR^*, \quad (84)$$

and we have to see that if we write

$$\mathcal{R}(V) = D\tau(A_1 + A_1^+) + VA_2^+ + N_1A_3^+ + N_2A_4^+,$$

where  $A_1$  is the matrix (29), then the functions  $A_1^+$ ,  $A_2^+$ ,  $A_3^+$  and  $A_4^+$  are small—i.e. they satisfy (77). To this end, expanding  $\mathcal{R}(V)$  in the previous expression as

$$\begin{aligned}\mathcal{R}(V) &= \mathcal{R}(JD\tau G_{D\tau}^{-1}) + N_1(L_\omega B_1 - \Lambda B_1) + N_2(L_\omega B_2 + \Lambda B_2) + D\tau L_\omega B_3 \\ &\quad + RB_1 + iR^*B_2 + DeB_3,\end{aligned}$$

and multiplying both sides of this equation by  $V^\top J$ ,  $D\tau^\top J$ ,  $N_2^\top J$  and  $N_1^\top J$ , we obtain the linear system

$$(\text{Id} + M_3) \begin{pmatrix} A_1^+ \\ A_2^+ \\ A_3^+ \\ A_4^+ \end{pmatrix} = \begin{pmatrix} C_1 \\ C_2 \\ C_3 \\ C_4 \end{pmatrix}, \quad (85)$$

where

$$M_3 = \begin{pmatrix} \Omega_{V,D\tau} - \text{Id}_r & \Omega_V & \Omega_{V,N_1} & \Omega_{V,N_2} \\ -\Omega_{D\tau} & \Omega_{V,D\tau} - \text{Id}_r & -\Omega_{D\tau,N_1} & -\Omega_{D\tau,N_2} \\ \Omega_{N_2,D\tau} & \Omega_{N_2,V} & \Omega_{N_2,N_1} - \text{Id}_{n-r} & \Omega_{N_2} \\ -\Omega_{N_1,D\tau} & -\Omega_{N_1,V} & -\Omega_{N_1} & \Omega_{N_2,N_1} - \text{Id}_{n-r} \end{pmatrix} \quad (86)$$

and the functions  $C_1$ ,  $C_2$ ,  $C_3$  and  $C_4$  have the following form

$$\begin{aligned}C_1 &= \overbrace{V^\top J\mathcal{R}(JD\tau G_{D\tau}^{-1}) + L_\omega B_3 - A_1 + \Omega_{V,N_1}(L_\omega B_1 - \Lambda B_1) + \Omega_{V,N_2}(L_\omega B_2 + \Lambda B_2)}^{C_1^+} \\ &\quad + V^\top J(RB_1 + iR^*B_2 + DeB_3) + (\Omega_{V,D\tau} - \text{Id}_r)(L_\omega B_3 - A_1), \\ C_2 &= \overbrace{-D\tau^\top J\mathcal{R}(JD\tau G_{D\tau}^{-1}) - \Omega_{D\tau,N_1}(L_\omega B_1 - \Lambda B_1) - \Omega_{D\tau,N_2}(L_\omega B_2 + \Lambda B_2)}^{C_2^+} \\ &\quad - D\tau^\top J(RB_1 + iR^*B_2 + DeB_3) + \Omega_{D\tau}(A_1 - L_\omega B_3), \\ C_3 &= \overbrace{N_2^\top J\mathcal{R}(JD\tau G_{D\tau}^{-1}) + L_\omega B_1 - \Lambda B_1 + \Omega_{N_2}(L_\omega B_2 + \Lambda B_2) + \Omega_{N_2,D\tau}(L_\omega B_3 - A_1)}^{C_3^+} \\ &\quad + (\Omega_{N_2,N_1} - \text{Id}_{n-r})(L_\omega B_1 - \Lambda B_1) + N_2^\top J(RB_1 + iR^*B_2 + DeB_3), \\ C_4 &= \overbrace{-N_1^\top J\mathcal{R}(JD\tau G_{D\tau}^{-1}) + L_\omega B_2 + \Lambda B_2 - \Omega_{N_1}(L_\omega B_1 - \Lambda B_1) + \Omega_{N_1,D\tau}(A_1 - L_\omega B_3)}^{C_4^+} \\ &\quad - (\Omega_{N_1,N_2} + \text{Id}_{n-r})(L_\omega B_2 + \Lambda B_2) - N_1^\top J(RB_1 + iR^*B_2 + DeB_3),\end{aligned}$$

and we observe that  $C_4 = -iC_3^*$  and  $C_4^+ = -i(C_3^+)^*$ . Apart from  $C_1^+$ ,  $C_2^+$ ,  $C_3^+$  and  $C_4^+$ , the size of the other terms that appear in the above expressions are easily controlled in terms of  $\|e\|_\rho$  and  $\|R\|_\rho$ —using approximately symplectic properties in (68)–(74). We see next that  $C_j^+$ , for  $j = 1, \dots, 4$ , are also controlled in a similar way, since they are given by equations which are close to (32)–(35) for the invariant and reducible case. For example, using equation (24) for  $B_1$  in the expressions of  $C_3^+$  we obtain

$$C_3^+ = (L_\omega N_2 - J\text{hess } hN_2 - N_2\Lambda)^\top D\tau G_{D\tau}^{-1} = iG_{R^*,D\tau} G_{D\tau}^{-1}, \quad (87)$$

where we used equation (79). To control  $C_1^+$  and  $C_2^+$  we have to compute the action of  $\mathcal{R}$  on the matrix  $J D\tau G_{D\tau}^{-1}$

$$\begin{aligned}\mathcal{R}(J D\tau G_{D\tau}^{-1}) &= \mathcal{R}(J D\tau) G_{D\tau}^{-1} + J D\tau L_\omega(G_{D\tau}^{-1}) = (\text{Id} + J D\tau G_{D\tau}^{-1} D\tau^\top J) T_1 D\tau G_{D\tau}^{-1} \\ &\quad + J D\tau G_{D\tau}^{-1} - J D\tau G_{D\tau}^{-1} [G_{De, D\tau} + G_{D\tau, De}] G_{D\tau}^{-1},\end{aligned}$$

where  $T_1$  is given by (49). Then, if we multiply this expression by  $D\tau^\top J$  we get

$$C_2^+ = -G_{De, D\tau} G_{D\tau}^{-1} \quad (88)$$

and if we multiply by  $V^\top J$  and use the definitions of  $C_2^+$ ,  $C_3^+$  and  $C_4^+$ , we obtain

$$\begin{aligned}C_1^+ &= (B_3^\top D\tau^\top + B_2^\top N_2^\top + B_1^\top N_1^\top + G_{D\tau}^{-1} D\tau^\top J^\top) J \mathcal{R}(J D\tau G_{D\tau}^{-1}) + L_\omega B_3 - A_1 \\ &= -B_3^\top C_2^+ + B_2^\top (C_3^+ - L_\omega B_1 + \Lambda B_1) + B_1^\top (-C_4^+ + L_\omega B_2 + \Lambda B_2) \\ &\quad + G_{D\tau}^{-1} D\tau^\top \mathcal{R}(J D\tau G_{D\tau}^{-1}) + L_\omega B_3 - A_1 \\ &= -B_3^\top C_2^+ + B_2^\top C_3^+ - B_1^\top C_4^+ + G_{D\tau}^{-1} \Omega_{D\tau} G_{D\tau}^{-1} D\tau^\top J T_1 D\tau G_{D\tau}^{-1} \\ &\quad + G_{D\tau}^{-1} \Omega_{D\tau, De} G_{D\tau}^{-1} - G_{D\tau}^{-1} \Omega_{D\tau} G_{D\tau}^{-1} (G_{De, D\tau} + G_{D\tau, De}) G_{D\tau}^{-1} + C_1^{++},\end{aligned} \quad (89)$$

where  $C_1^{++}$  is given as

$$\begin{aligned}C_1^{++} &= G_{D\tau}^{-1} D\tau^\top T_1 D\tau G_{D\tau}^{-1} + B_1^\top (L_\omega B_2 + \Lambda B_2) - B_2^\top (L_\omega B_1 - \Lambda B_1) + L_\omega B_3 - A_1 \\ &= G_{D\tau}^{-1} D\tau^\top T_1 D\tau G_{D\tau}^{-1} + \text{Re}(L_\omega B_2^\top B_1 - B_2^\top L_\omega B_1 + 2B_2^\top \Lambda B_1) - A_1,\end{aligned}$$

where we used that  $B_3 = \text{Re}(G_{B_2, B_1})$  and  $(B_1^\top (L_\omega B_2 + \Lambda B_2))^* = -B_2^\top (L_\omega B_1 - \Lambda B_1)$ . By introducing expression (48) for  $A_1$ , expanding  $L_\omega B_1$  and  $L_\omega B_2$  as in lemma 4.3

$$\begin{aligned}L_\omega B_2 &= -\Lambda B_2 + N_1^\top J T_1 D\tau G_{D\tau}^{-1} - N_1^\top D\tau G_{D\tau}^{-1} D\tau^\top J T_1 D\tau G_{D\tau}^{-1} \\ &\quad - G_{R, D\tau} G_{D\tau}^{-1} - G_{N_1, De} G_{D\tau}^{-1} + G_{N_1, D\tau} G_{D\tau}^{-1} (G_{De, D\tau} + G_{D\tau, De}) G_{D\tau}^{-1}, \\ L_\omega B_1 &= \Lambda B_1 - N_2^\top J T_1 D\tau G_{D\tau}^{-1} + N_2^\top D\tau G_{D\tau}^{-1} D\tau^\top J T_1 D\tau G_{D\tau}^{-1} \\ &\quad + iG_{R^*, D\tau} G_{D\tau}^{-1} + G_{N_2, De} G_{D\tau}^{-1} - G_{N_2, D\tau} G_{D\tau}^{-1} (G_{De, D\tau} + G_{D\tau, De}) G_{D\tau}^{-1},\end{aligned}$$

and using that  $\text{Re}(N_1 N_2^\top) = -\text{Re}((N_1 N_2^\top)^\top)$  and  $\text{Re}(G_{D\tau, R} G_{N_2, D\tau})^\top = -\text{Re}(iG_{D\tau, N_1} G_{R^*, D\tau})$ , we obtain (after some cancellations)

$$C_1^{++} = \text{Re}(T_3 + T_3^\top), \quad (90)$$

where

$$T_3 = -G_{D\tau}^{-1} (G_{D\tau, R} + G_{De, N_1} - (G_{De, D\tau} + G_{D\tau, De}) G_{D\tau}^{-1} G_{D\tau, N_1}) G_{N_2, D\tau} G_{D\tau}^{-1}. \quad (91)$$

Now, we control expressions (88), (87), (91) and (90) as

$$\|C_2^+\|_{\rho-\delta} \leq \frac{\hat{\alpha}}{\delta} \|e\|_\rho, \quad \|C_3^+\|_{\rho-\delta}, \|C_4^+\|_{\rho-\delta} \leq \hat{\alpha} \|R\|_\rho,$$

$$\|T_3\|_{\rho-\delta}, \|C_1^{++}\|_{\rho-\delta} \hat{\alpha} \left( \frac{\|e\|_\rho}{\delta} + \|R\|_\rho \right).$$

We use these bounds to control expression (89) as follows:

$$\|C_1^+\|_{\rho-2\delta} \leq \hat{\alpha} \left( \frac{\|e\|_\rho}{\gamma \delta^{v+1}} + \frac{\|e\|_\rho^2}{\gamma \delta^{v+2}} + \|R\|_\rho \right).$$

Then we use hypothesis (75) to get rid of the quadratic terms, thus obtaining

$$\|C_1^+\|_{\rho-2\delta} \leq \frac{\hat{\alpha}}{\gamma \delta^v} \left( \frac{\|e\|_\rho}{\delta} + \|R\|_\rho \right).$$

Therefore, we have

$$\|C_i\|_{\rho-2\delta} \leq \frac{\hat{\alpha}}{\gamma \delta^{v+1}} \left( \frac{\|e\|_\rho}{\delta} + \|R\|_\rho \right),$$

for  $i = 1, \dots, 4$ . Finally, we obtain estimates for the inverse of the matrix  $\text{Id} + M_3$  that appears in system (85), given by

$$\|(\text{Id} + M_3)^{-1}\|_{\rho-2\delta} \leq \frac{1}{1 - \|M_3\|_{\rho-2\delta}}, \quad (92)$$

that, using hypothesis (75) again, is well posed since

$$\|M_3\|_{\rho-2\delta} \leq \frac{\hat{\alpha}}{\gamma\delta^v} \left( \frac{\|e\|_\rho}{\delta} + \|R\|_\rho \right) \leq \frac{1}{2}.$$

Therefore, we obtain (77) for the functions  $\{A_i^+\}_{i=1,\dots,4}$ .  $\square$

**Proof of proposition 5.1.** We organize the proof of this iterative procedure in three parts. In part **I**), we correct the invariance of the torus by approximately solving the linearized equation  $\mathcal{R}(\Delta_\tau) = -e$ , given by (17), as it was explained in proposition 4.4. Next, in part **II**) we correct the reducibility of the torus by approximately solving the linearized equation  $\mathcal{S}(\Delta_N, \Delta_\Lambda) = -\hat{R}$ , given by (15) and (18), as it was explained in proposition 4.6. Finally, in part **III**) we compute some additional estimates regarding the non-degeneracy conditions for the new torus.

Firstly, let us observe that condition (51) implies condition (75) in proposition 5.3 by taking a constant  $\bar{\alpha}$  larger than  $\hat{\alpha}$ . Then, we use proposition 5.3 to construct an approximately symplectic basis at every point of the torus. As before, we redefine (enlarge) the constant  $\bar{\alpha}$  along the proof to meet the different conditions given in the statement.

**(I) Correction of the torus.** The idea is that the solution of the equation  $\mathcal{R}(\Delta_\tau) = -e$  obtained in the invariant and reducible case—as discussed in proposition 4.4—provides an approximate solution in the approximately invariant case. To this end, we consider the function

$$\Delta_\tau = D\tau\Delta_1 + V\Delta_2 + N_1\Delta_3 + N_2\Delta_4,$$

where  $\Delta_i$ , for  $i = 1, \dots, 4$ , are solutions of the cohomological equations (37)–(40), taking  $[\Delta_1]_{\mathbb{T}^v} = 0$  and  $[\Delta_2]_{\mathbb{T}^v}$  given by (41). Then we claim that the new embedding  $\bar{\tau} = \tau + \Delta_\tau$  parametrizes an approximately reducible and invariant torus  $\bar{\mathcal{T}}$  with an error which is quadratic in  $\|e\|_\rho$  and  $\|R\|_\rho$ . Of course, first we have to check the compatibility condition  $[D\tau^\top J e]_{\mathbb{T}^v} = 0$ . This follows from

$$D\tau^\top J e = D\tau^\top J (L_\omega \tau - J \text{grad } h(\tau)) = \Omega_{D\tau} \omega + \text{grad}_\theta(h(\tau)), \quad (93)$$

by observing that both terms at the right-hand side have zero average (see remark 2.2). It is important to observe that  $\Delta_3^* = i\Delta_4$  so the correction  $\Delta_\tau$  is real analytic.

As far as the estimates are concerned, we have (using lemma 5.2 to control the solution of the cohomological equations)

$$\|\Delta_1\|_{\rho-2\delta} \leq \frac{\bar{\alpha}}{\gamma^2\delta^{2v}} \|e\|_\rho, \quad \|\Delta_i\|_{\rho-\delta} \leq \frac{\bar{\alpha}}{\gamma\delta^v} \|e\|_\rho,$$

for  $i = 2, 3, 4$ , so we can control the correction  $\Delta_\tau$  in the parametrization as follows:

$$\|\Delta_\tau\|_{\rho-2\delta} \leq \frac{\bar{\alpha}}{\gamma^2\delta^{2v}} \|e\|_\rho,$$

thus obtaining estimate (56). Moreover, we observe that the derivative of the new parametrization can be controlled easily as follows:

$$\|D\bar{\tau}\|_{\rho-3\delta} \leq \|D\tau\|_\rho + \|D\Delta_\tau\|_{\rho-3\delta} \leq \|D\tau\|_\rho + \frac{\bar{\alpha}}{\gamma^2\delta^{2v+1}} \|e\|_\rho < \sigma_1,$$

where we used hypothesis (51), and also the distance of  $\bar{\tau}(\Delta(\rho - 2\delta))$  to the boundary of  $\mathcal{U}$

$$\text{dist}(\bar{\tau}(\Delta(\rho - 2\delta)), \partial\mathcal{U}) \geq \text{dist}(\tau(\Delta(\rho)), \partial\mathcal{U}) - \|\Delta_\tau\|_{\rho-2\delta}$$

$$\geq \text{dist}(\tau(\Delta(\rho)), \partial\mathcal{U}) - \frac{\tilde{\alpha}}{\gamma^2 \delta^{2\nu}} \|e\|_\rho > \sigma_2,$$

where we used hypothesis (52). Note that we have achieved part of (64) and (65).

Next we control the new error in the invariance. To this end, we first introduce  $\Delta_\tau$  into  $\mathcal{R}(\Delta_\tau) + e$  and we use properties (76) and (84) of the operator  $\mathcal{R}$  and also the cohomological equations (37)–(40), thus obtaining

$$\begin{aligned} \mathcal{R}(\Delta_\tau) + e &= \mathcal{R}(D\tau)\Delta_1 + \mathcal{R}(V)\Delta_2 + \mathcal{R}(N_1)\Delta_3 + \mathcal{R}(N_2)\Delta_4 \\ &\quad + D\tau L_\omega \Delta_1 + V L_\omega \Delta_2 + N_1 L_\omega \Delta_3 + N_2 L_\omega \Delta_4 + e \\ &= De\Delta_1 + (D\tau A_1^\top + V A_2^\top + N_1 A_3^\top + N_2 A_4^\top)\Delta_2 + R\Delta_3 + iR^* \Delta_4 \\ &\quad - \underbrace{D\tau V^\top J e + V D\tau^\top J e - N_1 N_2^\top J e + N_2 N_1^\top J e + e}_{e^+} \end{aligned} \quad (94)$$

We note that the terms not included in  $e^+$  are clearly quadratic in  $e$  and  $R$ , since the functions  $\{A_i^\top\}_{i=1,\dots,4}$  and  $\{\Delta_i\}_{i=1,\dots,4}$  are controlled by  $\|e\|_\rho$  and  $\|R\|_\rho$ . Then, it suffices to study the remaining part  $e^+$ . To this end, we write  $e^+$  in terms of the constructed basis

$$e^+ = D\tau e_1^+ + V e_2^+ + N_1 e_3^+ + N_2 e_4^+,$$

and obtain  $\{e_i^+\}_{i=1,\dots,4}$  by multiplying both sides by  $V^\top J$ ,  $D\tau^\top J$ ,  $N_2^\top J$  and  $N_1^\top J$ . This leads to study the linear system

$$(\text{Id} + M_3) \begin{pmatrix} e_1^+ \\ e_2^+ \\ e_3^+ \\ e_4^+ \end{pmatrix} = \begin{pmatrix} D_1 \\ D_2 \\ D_3 \\ D_4 \end{pmatrix}, \quad (95)$$

where  $M_3$  is given in (86) and the matrices in the right-hand side are

$$\begin{aligned} D_1 &= -(\Omega_{V,D\tau} - \text{Id}_r) V^\top J e + \Omega_V D\tau^\top J e - \Omega_{V,N_1} N_2^\top J e + \Omega_{V,N_2} N_1^\top J e, \\ D_2 &= \Omega_{D\tau} V^\top J e - (\Omega_{D\tau,V} + \text{Id}_r) D\tau^\top J e + \Omega_{D\tau,N_1} N_2^\top J e - \Omega_{D\tau,N_2} N_1^\top J e, \\ D_3 &= -\Omega_{N_2,D\tau} V^\top J e + \Omega_{N_2,V} D\tau^\top J e - (\Omega_{N_2,N_1} - \text{Id}_{n-r}) N_2^\top J e + \Omega_{N_2} N_1^\top J e, \\ D_4 &= -iD_3^*. \end{aligned}$$

Now we control these functions using estimates (68)–(74) in proposition 5.3

$$\|D_i\|_{\rho-2\delta} \leq \frac{\tilde{\alpha}}{\gamma \delta^\nu} \left( \frac{\|e\|_\rho}{\delta} + \|R\|_\rho \right) \|e\|_\rho,$$

for  $i = 1, \dots, 4$ . We have shown in the proof of proposition 5.3 that the matrix  $\text{Id} + M_3$  is invertible and that  $\|(\text{Id} + M_3)^{-1}\|_{\rho-2\delta} \leq 2$  (see (92)) so we conclude that

$$\|e^+\|_{\rho-2\delta} \leq \frac{\tilde{\alpha}}{\gamma \delta^\nu} \left( \frac{\|e\|_\rho}{\delta} + \|R\|_\rho \right) \|e\|_\rho.$$

Going back to equation (94) we get

$$\|\mathcal{R}(\Delta_\tau) + e\|_{\rho-3\delta} \leq \frac{\tilde{\alpha}}{\gamma^2 \delta^{2\nu+1}} \left( \frac{\|e\|_\rho}{\delta} + \|R\|_\rho \right) \|e\|_\rho,$$

and therefore, we conclude that  $\mathcal{R}(\Delta_\tau) = -e$  is solved modulo quadratic terms in the errors. Then, we observe that

$$\begin{aligned} \bar{e} &= L_\omega \bar{\tau} - J \text{grad } h(\bar{\tau}) \\ &= \mathcal{R}(\Delta_\tau) + e + J(\text{grad } h(\tau) + \text{hess } h(\tau) \Delta_\tau - \text{grad } h(\tau + \Delta_\tau)). \end{aligned}$$



We control the last terms by estimating the remainder of the Taylor expansion of  $h$  up to second order, thus obtaining

$$\|\text{grad } h(\tau) + \text{hess } h(\tau) \Delta_\tau - \text{grad } h(\tau + \Delta_\tau)\|_{\rho-2\delta} \leq \frac{\bar{\alpha}}{\gamma^4 \delta^{4\nu}} \|e\|_\rho^2.$$

Hence, we end up with

$$\|\bar{e}\|_{\rho-3\delta} \leq \frac{\bar{\alpha}}{\gamma^4 \delta^{4\nu-1}} \left( \frac{\|e\|_\rho}{\delta} + \|R\|_\rho \right) \|e\|_\rho,$$

where we used that  $\nu > r - 1 \geq 1$ , finally obtaining estimate (57).

(II) *Correction of the reducibility.* To square the error in reducibility of the new torus  $\tilde{\mathcal{T}}$  we have to deal with the equation  $\mathcal{S}(\Delta_N, \Delta_\Lambda) = -\hat{R}$ , given by (15) and (18). As before, we solve approximately this equation by taking (the reason of writing  $\hat{\Delta}_N$  rather than  $\Delta_N$  will be clear later on)

$$\hat{\Delta}_N = D\tau P_1 + V P_2 + N_1 P_3 + N_2 P_4,$$

$\{P_i\}_{i=1,\dots,4}$  and  $\Delta_\Lambda$  being the solutions of the cohomological equations (42)–(45) for

$$\hat{R} = R + J \text{hess } h(\tau) N - J \text{hess } h(\bar{\tau}) N,$$

and fixing  $\text{diag}[P_3]_{\mathbb{T}^r} = 0$ . The formal solution of these equations has been discussed in proposition 4.6 so we know that we must take  $\Delta_\Lambda = -\text{diag}[N_2^\top J \hat{R}]_{\mathbb{T}^r}$ .

Firstly, we claim that the geometry of the problem implies that the selected  $\Delta_\Lambda$  is pure imaginary, so our procedure automatically preserves the approximately elliptic character of the torus. To see that, we observe that transposing equation (83) leads to

$$L_\omega \Omega_{N_1, N_2} = -\Lambda \Omega_{N_1, N_2} + \Omega_{N_1, N_2} \Lambda + \Omega_{R, N_2} + i \Omega_{N_1, R^*}.$$

Since the left-hand side of this expression has vanishing average and  $\text{diag}[\Omega_{N_1, N_2}]_{\mathbb{T}^r} = -\text{Id}_{n-r}$ , it turns out that

$$\text{diag}[\Omega_{R, N_2} + i \Omega_{N_1, R^*}]_{\mathbb{T}^r} = 0,$$

and so  $\text{diag}[i \Omega_{N_1, R^*}]_{\mathbb{T}^r} = \text{diag}[\Omega_{N_2, R}]_{\mathbb{T}^r}^\top$ . Then, it is straightforward to compute

$$\begin{aligned} \Delta_\Lambda^* &= -\text{diag}[N_2^\top J \hat{R}]_{\mathbb{T}^r}^* = -\text{diag}[-i \Omega_{N_1, R^*} + N_1^\top (\text{hess } h(\tau) - \text{hess } h(\bar{\tau})) N_2]_{\mathbb{T}^r} \\ &= \text{diag}[\Omega_{N_2, R}]_{\mathbb{T}^r}^\top - \text{diag}[N_2^\top (\text{hess } h(\tau) - \text{hess } h(\bar{\tau})) N_1]_{\mathbb{T}^r}^\top \\ &= \text{diag}[\Omega_{N_2, R} - N_2^\top (\text{hess } h(\tau) - \text{hess } h(\bar{\tau})) N_1]_{\mathbb{T}^r}^\top = -\Delta_\Lambda^\top = -\Delta_\Lambda, \end{aligned} \quad (96)$$

so  $\Delta_\Lambda$  is pure imaginary.

Now obtaining estimates for the solution of the cohomological equations is straightforward after controlling

$$\|\hat{R}\|_{\rho-2\delta} \leq \frac{\bar{\alpha}}{\gamma^2 \delta^{2\nu-1}} \left( \frac{\|e\|_\rho}{\delta} + \|R\|_\rho \right), \quad |\Delta_\Lambda| \leq \frac{\bar{\alpha}}{\gamma^2 \delta^{2\nu-1}} \left( \frac{\|e\|_\rho}{\delta} + \|R\|_\rho \right),$$

and applying lemma 5.2

$$\|P_1\|_{\rho-4\delta} \leq \frac{\bar{\alpha}}{\gamma^4 \delta^{4\nu-1}} \left( \frac{\|e\|_\rho}{\delta} + \|R\|_\rho \right), \quad \|P_i\|_{\rho-3\delta} \leq \frac{\bar{\alpha}}{\gamma^3 \delta^{3\nu-1}} \left( \frac{\|e\|_\rho}{\delta} + \|R\|_\rho \right),$$

for  $i = 2, 3, 4$ . With these estimates we check condition (64) for the new approximate normal frequencies  $\bar{\lambda}$ . For example,

$$|\bar{\lambda}_i \pm \bar{\lambda}_j| \geq |\lambda_i \pm \lambda_j| - 2|\Delta_\Lambda| \geq \min_{i \neq j} |\lambda_i \pm \lambda_j| - 2|\Delta_\Lambda| > \sigma_3,$$

where we used (53). Similar computations allow us to see that  $(\sigma_3/2) < |\bar{\lambda}_j| < (\sigma_4/2)$ , using (54) and (55), respectively.

We also have

$$\|\hat{\Delta}_N\|_{\rho-4\delta} \leq \frac{\bar{\alpha}}{\gamma^4 \delta^{4v-1}} \left( \frac{\|e\|_\rho}{\delta} + \|R\|_\rho \right)$$

and we observe that, if we introduce  $\hat{N} = N + \hat{\Delta}_N$ , then using (51) we obtain

$$\|\hat{N}\|_{\rho-4\delta} \leq \|N\|_\rho + \|\hat{\Delta}_N\|_{\rho-4\delta} \leq \|N\|_\rho + \frac{\bar{\alpha}}{\gamma^4 \delta^{4v-1}} \left( \frac{\|e\|_\rho}{\delta} + \|R\|_\rho \right) < \sigma_1. \quad (97)$$

Now the matrix  $\text{diag}[\Omega_{\hat{N}, \hat{N}^*}]_{\mathbb{T}^r}$  is constant, diagonal and pure imaginary, but it is not  $i\text{Id}_{n-r}$  as we want. Nevertheless, from the following expression:

$$\Omega_{\hat{N}, \hat{N}^*} - \Omega_{N, N^*} = \Omega_{N, \hat{\Delta}_N^*} + \Omega_{\hat{\Delta}_N, N^*}$$

and using hypothesis (51) we obtain

$$\|\Omega_{\hat{N}, \hat{N}^*} - \Omega_{N, N^*}\|_{\rho-4\delta} \leq \frac{\bar{\alpha}}{\gamma^4 \delta^{4v-1}} \left( \frac{\|e\|_\rho}{\delta} + \|R\|_\rho \right),$$

where we recall that  $\text{diag}[\Omega_{N, N^*}]_{\mathbb{T}^r} = i\text{Id}_{n-r}$ . Hence, we have that the elements of  $\text{diag}[\Omega_{\hat{N}, \hat{N}^*}]_{\mathbb{T}^r}$  are of the form  $i(1 + d_i)$  with

$$|d_i| \leq \frac{\bar{\alpha}}{\gamma^4 \delta^{4v-1}} \left( \frac{\|e\|_\rho}{\delta} + \|R\|_\rho \right).$$

Hence, using again hypothesis (51), we have that  $|d_i| \leq 1/2$  for  $i = 1, \dots, n-r$ , so we can normalize  $\hat{N}$  in order to preserve hypothesis  $H_3$ . To this end, we define the real matrix

$$B = \text{diag}(b_1, \dots, b_{n-r}), \quad \text{with} \quad b_i = \sqrt{\frac{1}{1 + d_i}},$$

and it turns out that the matrix  $\bar{N} = \hat{N}B$  satisfies  $\text{diag}[\Omega_{\bar{N}, \bar{N}^*}]_{\mathbb{T}^r} = i\text{Id}_{n-r}$ . We observe that the performed correction is small, since if we take  $\bar{N} = N + \Delta_N$  we have that

$$\Delta_N = N(B - \text{Id}_{n-r}) + \hat{\Delta}_N B,$$

and so

$$\|\Delta_N\|_{\rho-4\delta} \leq \frac{\bar{\alpha}}{\gamma^4 \delta^{4v-1}} \left( \frac{\|e\|_\rho}{\delta} + \|R\|_\rho \right),$$

which corresponds to estimate (59). We see that  $\|\bar{N}\|_{\rho-4\delta} < \sigma_1$  by similar computations as in (97), thus obtaining the corresponding condition in (65).

The rest of this part is devoted to check that the new approximately invariant torus  $\bar{\mathcal{T}}$  is approximately elliptic up to a quadratic error. To this end, we compute

$$\begin{aligned} \bar{R} &= L_\omega \bar{N} - J \text{hess } h(\bar{\tau}) \bar{N} + \bar{N} \bar{\Lambda} \\ &= \mathcal{S}(\Delta_N, \Delta_\Lambda) + \hat{R} + J(\text{hess } h(\tau) - \text{hess } h(\bar{\tau})) \Delta_N + \Delta_N \Delta_\Lambda. \end{aligned} \quad (98)$$

Now we write the action of  $\mathcal{S}$  on  $\Delta_N$  in terms of the action on  $\hat{\Delta}_N$  as follows:

$$\begin{aligned} \mathcal{S}(\Delta_N, \Delta_\Lambda) + \hat{R} &= \mathcal{S}(N(B - \text{Id}_{n-r}) + \hat{\Delta}_N B, \Delta_\Lambda) + \hat{R} \\ &= \mathcal{R}(N(B - \text{Id}_{n-r})) + \mathcal{R}(\hat{\Delta}_N B) + N \Delta_\Lambda + N(B - \text{Id}_{n-r}) \Lambda + \hat{\Delta}_N B \Lambda + \hat{R} \\ &= R(B - \text{Id}_{n-r}) + \mathcal{R}(\hat{\Delta}_N) B + N \Delta_\Lambda + \hat{\Delta}_N B \Delta + \hat{R} \\ &= (\mathcal{S}(\hat{\Delta}_N, \Delta_\Lambda) + \hat{R}) B + (R - \hat{R} - N \Delta_\Lambda + \hat{\Delta}_N \Lambda)(B - \text{Id}_{n-r}), \end{aligned}$$

where we used that  $\mathcal{R}(N) = -N\Lambda + R$  and  $B\Lambda = \Lambda B$ .

Then, we introduce  $\hat{\Delta}_N$  and  $\Delta_\Lambda$  in  $\mathcal{S}(\hat{\Delta}_N, \Delta_\Lambda) + \hat{R}$  and we use the properties (76) and (84) of the operator  $\mathcal{R}$  and also the cohomological equations (42)–(45), thus obtaining

$$\begin{aligned} \mathcal{S}(\hat{\Delta}_N, \Delta_\Lambda) + \hat{R} &= \mathcal{R}(\hat{\Delta}_N) + N_1 \Delta_\Lambda + \hat{\Delta}_N \Lambda + \hat{R} \\ &= \mathcal{R}(D\tau)P_1 + \mathcal{R}(V)P_2 + \mathcal{R}(N_1)P_3 + \mathcal{R}(N_2)P_4 \\ &\quad + D\tau L_\omega P_1 + V L_\omega P_2 + N_1 L_\omega P_3 + N_2 L_\omega P_4 + N_1 \Delta_\Lambda \\ &\quad + D\tau P_1 \Lambda + V P_2 \Lambda + N_1 P_3 \Lambda + N_2 P_4 \Lambda + \hat{R} \\ &= DeP_1 + (D\tau A_1^+ + V A_2^+ + N_1 A_3^+ + N_2 A_4^+)P_2 + R P_3 + iR^* P_4 \\ &\quad - \underbrace{D\tau V^\top J \hat{R} + V D\tau^\top J \hat{R} - N_1 N_2^\top J \hat{R} + N_2 N_1^\top J \hat{R} + \hat{R}}_{R^+}. \end{aligned}$$

As in the case of equation (94), the terms not included in  $R^+$  are clearly quadratic in  $e$  and  $R$ . Then, we express  $R^+$  in terms of the basis

$$R^+ = D\tau R_1^+ + V R_2^+ + N_1 R_3^+ + N_2 R_4^+,$$

and for  $R_j^+$  we get a system like (95) for  $e_j^+$ , simply by replacing  $e$  with  $\hat{R}$  in the definition of  $D_j$ . Hence,

$$\|R^+\|_{\rho-2\delta} \leq \frac{\bar{\alpha}}{\gamma^3 \delta^{3\nu-1}} \left( \frac{\|e\|_\rho}{\delta} + \|R\|_\rho \right)^2.$$

Therefore, we can compute a bound for the error in the solution of the linear equation that corrects reducibility

$$\|\mathcal{S}(\hat{\Delta}_N, \Delta_\Lambda) + \hat{R}\|_{\rho-4\delta} \leq \frac{\bar{\alpha}}{\gamma^4 \delta^{4\nu}} \left( \frac{\|e\|_\rho}{\delta} + \|R\|_\rho \right)^2,$$

so we obtain

$$\|\mathcal{S}(\Delta_N, \Delta_\Lambda) + \hat{R}\|_{\rho-4\delta} \leq \frac{\bar{\alpha}}{\gamma^8 \delta^{8\nu-2}} \left( \frac{\|e\|_\rho}{\delta} + \|R\|_\rho \right)^2.$$

Therefore, recalling (98), we easily show that the new error (60) in reducibility is quadratic

$$\|\bar{R}\|_{\rho-4\delta} \leq \frac{\bar{\alpha}}{\gamma^8 \delta^{8\nu-2}} \left( \frac{\|e\|_\rho}{\delta} + \|R\|_\rho \right)^2.$$

(III) *Additional estimates.* Finally, we have to check estimates that allow us to control the non-degeneracy of the basis and the twist condition. Using that

$$G_{D\bar{\tau}} - G_{D\tau} = G_{D\tau, D\Delta_\tau} + G_{D\Delta_\tau, D\bar{\tau}}$$

and recalling (51) and (56), we get

$$\|G_{D\bar{\tau}} - G_{D\tau}\|_{\rho-3\delta} \leq \frac{\bar{\alpha}}{\gamma^2 \delta^{2\nu+1}} \|e\|_\rho.$$

Now, we observe that  $G_{D\bar{\tau}}^{-1} = (\text{Id}_r + G_{D\tau}^{-1}(G_{D\bar{\tau}} - G_{D\tau}))^{-1} G_{D\tau}^{-1}$  so we can compute the following (again, we make use of (51)):

$$\begin{aligned} \|G_{D\bar{\tau}}^{-1} - G_{D\tau}^{-1}\|_{\rho-3\delta} &\leq \|G_{D\tau}^{-1}\|_\rho \|(\text{Id}_r + G_{D\tau}^{-1}(G_{D\bar{\tau}} - G_{D\tau}))^{-1} - \text{Id}_r\|_{\rho-3\delta} \quad (99) \\ &\leq \frac{\|G_{D\tau}^{-1}\|_\rho^2 \|G_{D\bar{\tau}} - G_{D\tau}\|_{\rho-3\delta}}{1 - \|G_{D\tau}^{-1}\|_\rho \|G_{D\bar{\tau}} - G_{D\tau}\|_{\rho-3\delta}} \leq \frac{\bar{\alpha}}{\gamma^2 \delta^{2\nu+1}} \|e\|_\rho, \end{aligned}$$

thus obtaining (61) and the term in (65) that corresponds to  $G_{D\bar{\tau}}^{-1}$ . Similar computations allow us to control the non-degeneracy of the set of normal vectors, thus getting (62) and (65) for  $G_{\bar{N}, \bar{N}^*}^{-1}$ . Now, we are able to estimate the new twist condition for

$$\bar{A}_1(\theta) = G_{D\bar{\tau}}^{-1}(\theta) D\bar{\tau}(\theta)^\top (\bar{T}_1(\theta) + \bar{T}_2(\theta) + \bar{T}_2(\theta)^\top) D\bar{\tau}(\theta) G_{D\bar{\tau}}^{-1}(\theta),$$

where

$$\begin{aligned}\bar{T}_1(\theta) &= J^\top \text{hess } h(\bar{\tau}(\theta)) J - \text{hess } h(\bar{\tau}(\theta)), \\ \bar{T}_2(\theta) &= \bar{T}_1 J [D\bar{\tau}(\theta) G_{D\bar{\tau}}(\theta)^{-1} D\bar{\tau}(\theta)^\top - \text{Id}] \text{Re}(i\bar{N}(\theta) \bar{N}^*(\theta)^\top).\end{aligned}$$

As before, we first bound

$$|[\bar{A}_1]_{\mathbb{T}^r} - [A_1]_{\mathbb{T}^r}| \leq \frac{\bar{\alpha}}{\gamma^4 \delta^{4\nu-1}} \left( \frac{\|e\|_\rho}{\delta} + \|R\|_\rho \right).$$

Now we estimate the inverse of  $[\bar{A}_1]_{\mathbb{T}^r}$  using the fact that  $\bar{A}_1 = A_1 + \bar{A}_1 - A_1$ . Then, we repeat the same argument used in (99), thus obtaining bounds (63) and (65) for  $[\bar{A}_1]_{\mathbb{T}^r}^{-1}$ .  $\square$

## 6. Proof of the main result

In this section we prove theorem 3.1 by applying inductively proposition 5.1. First, in section 6.1 we study the convergence of the obtained iterative scheme, without worrying about the exclusion of parameters that lead to resonances. As usual, the quadratic convergence overcomes the effect of small divisors. Then, in section 6.2 we prove that Lipschitz regularity is preserved along the iterative procedure. Finally, in section 6.3, we estimate the measure of the set of excluded parameters.

### 6.1. Convergence of the Newton scheme

Given a parameter  $\mu \in I$ , we denote the objects that characterize the corresponding approximately elliptic and invariant torus as (from now on we omit the dependence on the parameter)

$$\tau_{(0)} = \tau_\mu, \quad N_{(0)} = N_\mu, \quad \Lambda_{(0)} = \Lambda_\mu,$$

and we introduce also

$$e_{(0)} = e_\mu, \quad R_{(0)} = R_\mu, \quad A_{1,(0)} = A_{1,\mu}, \quad \lambda_{(0)} = (\lambda_1^{(0)}, \dots, \lambda_{n-r}^{(0)}) = \lambda_\mu.$$

where we recall that  $\Lambda_{(0)} = \text{diag}(i\lambda_{(0)})$ . Moreover, given  $\gamma_0 > 0$  such that  $\gamma_0 \leq \frac{1}{2} \min\{1, \hat{\gamma}\}$ , we define the following quantities (recall that  $0 < \rho < 1$ ):

$$\begin{aligned}\rho_{(0)} &= \rho, \quad \delta_{(0)} = \frac{\rho_{(0)}}{16}, \quad \rho_{(s)} = \rho_{(s-1)} - 4\delta_{(s-1)}, \quad \delta_{(s)} = \frac{\delta_{(0)}}{2^s}, \\ \gamma_{(s)} &= (1 + 2^{-s})\gamma_0,\end{aligned}$$

for any  $s \geq 1$ , and consider the normalized error

$$\varepsilon_{(0)} = \frac{\|e_{(0)}\|_{\rho_{(0)}}}{\delta_{(0)}} + \|R_{(0)}\|_{\rho_{(0)}}. \quad (100)$$

We are going to show that, if the normalized error  $\varepsilon_{(0)}$  is sufficiently small so that

$$\frac{2^{8\nu-1} \bar{\alpha} \varepsilon_{(0)}}{\gamma_0^8 \delta_{(0)}^{8\nu-2}} < \frac{1}{2} \min \left\{ 1, \sigma_1 - \sigma^*, \text{dist}(\tau_{(0)}(\Delta(\rho)), \partial\mathcal{U}) - \sigma_2, \sigma^{**} - \sigma_3, \sigma_4 - 2 \max_j |\lambda_j^{(0)}| \right\}, \quad (101)$$

where

$$\sigma^* = \max \left\{ \|D\tau_{(0)}\|_{\rho_{(0)}}, \|N_{(0)}\|_{\rho_{(0)}}, \|G_{D\tau_{(0)}}^{-1}\|_{\rho_{(0)}}, \|G_{N_{(0)}, N_{(0)}^*}^{-1}\|_{\rho_{(0)}}, |[A_{1,(0)}]_{\mathbb{T}^r}^{-1}| \right\}, \quad (102)$$

$$\sigma^{**} = \min \left\{ \min_{i \neq j} |\lambda_i^{(0)} \pm \lambda_j^{(0)}|, 2 \min_j |\lambda_j^{(0)}| \right\} \quad (103)$$

and  $\bar{\alpha}$  is the constant provided by proposition 5.1, then we can apply recursively proposition 5.1 to the initial approximation. Thus, we obtain a sequence

$$\begin{aligned} \tau_{(s)} &= \bar{\tau}_{(s-1)} = \tau_{(s-1)} + \Delta_{\tau_{(s-1)}}, & e_{(s)} &= \bar{e}_{(s-1)}, \\ N_{(s)} &= \bar{N}_{(s-1)} = N_{(s-1)} + \Delta_{N_{(s-1)}}, & R_{(s)} &= \bar{R}_{(s-1)}, \\ \Lambda_{(s)} &= \bar{\Lambda}_{(s-1)} = \Lambda_{(s-1)} + \Delta_{\Lambda_{(s-1)}}, & A_{1,(s)} &= \bar{A}_{1,(s-1)}, \end{aligned}$$

all these objects being analytic in  $\Delta(\rho_{(s)})$ . Note that, in order to apply  $s$  times proposition 5.1, we restrict the parameter  $\mu$  to the set  $I_{(s-1)}$  defined iteratively by  $I_{(-1)} = I$  and

$$\begin{aligned} I_{(s)} &= \{\mu \in I_{(s-1)} : \lambda_{(s)} \text{ satisfies Diophantine conditions (9)} \\ &\quad \text{of } (\gamma_{(s)}, \nu)\text{-type with respect to } \omega\}. \end{aligned} \quad (104)$$

Let us observe that the basic frequencies  $\omega$  automatically satisfy Diophantine conditions (4) of  $(\gamma_{(s)}, \nu)$ -type, for every  $s \geq 0$ , since they are fixed along the procedure and we have  $\gamma_{(s)} \leq 2\gamma_0 \leq \hat{\gamma}$  and  $\nu > \hat{\nu}$ .

Now we proceed by induction. We suppose that we have applied  $s$  times proposition 5.1, for certain  $s \geq 0$ , and we verify that we can apply it again. To this end, we define  $\varepsilon_{(s)}$ ,  $\sigma_{(s)}^*$  and  $\sigma_{(s)}^{**}$  as in (100), (102) and (103), just by replacing the (0)-objects with (s)-ones. First, we observe that we have

$$\sigma_{(s)}^* < \sigma_1, \quad \text{dist}(\tau_{(s)}(\Delta(\rho_{(s)})), \partial\mathcal{U}) > \sigma_2, \quad \sigma_{(s)}^{**} > \sigma_3, \quad \max_j |\lambda_j^{(s)}| < \frac{\sigma_4}{2},$$

so the construction of the constant  $\bar{\alpha}$  of proposition 5.1 is uniform for all iterative steps—it depends on the constants  $\sigma_0, \sigma_1, \sigma_2, \sigma_3$  and  $\sigma_4$  that remain unchanged along the procedure. Hence, conditions (51)–(55) are fulfilled provided that the normalized error  $\varepsilon_{(s)}$  satisfies

$$\frac{\bar{\alpha}\varepsilon_{(s)}}{\gamma_{(s)}^4 \delta_{(s)}^{4\nu-1}} < \frac{1}{2} \min \left\{ 1, \sigma_1 - \sigma_{(s)}^*, \text{dist}(\tau_{(s)}(\Delta(\rho_{(s)})), \partial\mathcal{U}) - \sigma_2, \sigma_{(s)}^{**} - \sigma_3, \sigma_4 - 2 \max_j |\lambda_j^{(s)}| \right\}. \quad (105)$$

In order to verify this inequality, we start by computing the normalized error at the  $s$ th step (recall that  $\gamma_{(s)} < 1$  and  $\delta_{(s)} < 1$ )

$$\varepsilon_{(s)} = \frac{\|e_{(s)}\|_{\rho_{(s)}}}{\delta_{(s)}} + \|R_{(s)}\|_{\rho_{(s)}} \leq \frac{2\bar{\alpha}}{\gamma_{(s-1)}^8 \delta_{(s-1)}^{8\nu-2}} \varepsilon_{(s-1)}^2 \leq \frac{2^{(s-1)(8\nu-2)+1} \bar{\alpha}}{\gamma_0^8 \delta_{(0)}^{8\nu-2}} \varepsilon_{(s-1)}^2, \quad (106)$$

where we used (57), (60) and the fact that  $\gamma_{(s-1)} \geq \gamma_0$ . Then, by iterating this sequence backwards, we obtain that

$$\varepsilon_{(s)} \leq \frac{\gamma_0^8 \delta_{(0)}^{8\nu-2}}{2\bar{\alpha}} 2^{-(s+1)(8\nu-2)} \left( \frac{2^{8\nu-1} \bar{\alpha} \varepsilon_{(0)}}{\gamma_0^8 \delta_{(0)}^{8\nu-2}} \right)^{2^s}. \quad (107)$$

Using this expression of the error, we verify condition (105) in order to perform the step  $s+1$ . For example, the first term in this condition is straightforward

$$\frac{\bar{\alpha}\varepsilon_{(s)}}{\gamma_{(s)}^4 \delta_{(s)}^{4\nu-1}} \leq \frac{1}{2} \gamma_0^4 \delta_{(0)}^{4\nu-1} 2^{-(4\nu-1)s-8\nu+2} \left( \frac{2^{8\nu-1} \bar{\alpha} \varepsilon_{(0)}}{\gamma_0^8 \delta_{(0)}^{8\nu-2}} \right)^{2^s} < \frac{1}{2},$$

recalling that  $\nu > r - 1 \geq 1$  and (101). In order to verify the remaining conditions in (105), we have to control also the objects  $\|D\tau_{(s)}\|_{\rho_{(s)}}$ ,  $\|N_{(s)}\|_{\rho_{(s)}}$ ,  $|\Lambda_{(s)}|$ , etc. For example, we obtain

$$\|D\tau_{(s)}\|_{\rho_{(s)}} + \frac{\bar{\alpha} \varepsilon_{(s)}}{\gamma_{(s)}^4 \delta_{(s)}^{4\nu-1}} < \sigma_1$$

recursively as follows:

$$\begin{aligned} \|D\tau_{(s)}\|_{\rho_{(s)}} + \frac{\bar{\alpha} \varepsilon_{(s)}}{\gamma_{(s)}^4 \delta_{(s)}^{4\nu-1}} &\leq \|D\tau_{(s-1)}\|_{\rho_{(s-1)}} + \|D\Delta_{\tau_{(s-1)}}\|_{\rho_{(s)}} + \frac{\bar{\alpha} \varepsilon_{(s)}}{\gamma_{(s)}^4 \delta_{(s)}^{4\nu-1}} \\ &\leq \|D\tau_{(0)}\|_{\rho_{(0)}} + \sum_{j=0}^{s-1} \|D\Delta_{\tau_{(j)}}\|_{\rho_{(j+1)}} + \frac{\bar{\alpha} \varepsilon_{(s)}}{\gamma_{(s)}^4 \delta_{(s)}^{4\nu-1}} \\ &\leq \|D\tau_{(0)}\|_{\rho_{(0)}} + \sum_{j=0}^{s-1} \frac{\bar{\alpha} \varepsilon_{(j)}}{\gamma_{(j)}^2 \delta_{(j)}^{2\nu}} + \frac{\bar{\alpha} \varepsilon_{(s)}}{\gamma_{(s)}^4 \delta_{(s)}^{4\nu-1}} \\ &\leq \|D\tau_{(0)}\|_{\rho_{(0)}} + \sum_{j=0}^s \frac{\bar{\alpha} \varepsilon_{(j)}}{\gamma_{(j)}^4 \delta_{(j)}^{4\nu-1}}. \end{aligned} \quad (108)$$

Note that in the above computations we used estimate (56) in proposition 5.1 and the fact that  $\gamma_{(s)}, \delta_{(s)} < 1$ . Then, we introduce the expression for the errors  $\varepsilon_{(j)}$  previously computed and use that  $j+1 \leq 2^j$  to obtain

$$\begin{aligned} \|D\tau_{(s)}\|_{\rho_{(s)}} + \frac{\bar{\alpha} \varepsilon_{(s)}}{\gamma_{(s)}^4 \delta_{(s)}^{4\nu-1}} &\leq \|D\tau_{(0)}\|_{\rho_{(0)}} + \gamma_0^4 \delta_{(0)}^{4\nu-1} 2^{-8\nu+1} \sum_{j=0}^s 2^{-(4\nu-1)j} \left( \frac{2^{8\nu-1} \bar{\alpha} \varepsilon_{(0)}}{\gamma_0^8 \delta_{(0)}^{8\nu-2}} \right)^{2^j} \\ &\leq \|D\tau_{(0)}\|_{\rho_{(0)}} + \gamma_0^4 \delta_{(0)}^{4\nu-1} 2^{-8\nu+1} \sum_{j=0}^{\infty} \left( \frac{2^{8\nu-1} \bar{\alpha} \varepsilon_{(0)}}{\gamma_0^8 \delta_{(0)}^{8\nu-2}} \right)^{j+1} \\ &\leq \|D\tau_{(0)}\|_{\rho_{(0)}} + \frac{2\bar{\alpha} \varepsilon_{(0)}}{\gamma_0^4 \delta_{(0)}^{4\nu-1}} < \sigma_1, \end{aligned}$$

where in the last two inequalities we have used hypothesis (101) to bound the expression by the sum of a geometric progression of ratio 1/2. Analogous computations show that (we use estimates (59), (61), (62) and (63), respectively)

$$\begin{aligned} \|N_{(s)}\|_{\rho_{(s)}} + \frac{\bar{\alpha} \varepsilon_{(s)}}{\gamma_{(s)}^4 \delta_{(s)}^{4\nu-1}} &< \sigma_1, \quad \|G_{D\tau_{(s)}}^{-1}\|_{\rho_{(s)}} + \frac{\bar{\alpha} \varepsilon_{(s)}}{\gamma_{(s)}^4 \delta_{(s)}^{4\nu-1}} < \sigma_1, \\ \|G_{N_{(s)}, N_{(s)}^*}^{-1}\|_{\rho_{(s)}} + \frac{\bar{\alpha} \varepsilon_{(s)}}{\gamma_{(s)}^4 \delta_{(s)}^{4\nu-1}} &< \sigma_1, \quad |[A_{(s)}]_{\mathbb{T}^r}^{-1}| + \frac{\bar{\alpha} \varepsilon_{(s)}}{\gamma_{(s)}^4 \delta_{(s)}^{4\nu-1}} < \sigma_1, \end{aligned}$$

thus obtaining the second condition in (105). Next, to verify the inequality which corresponds to the third term in (105) we observe that

$$\text{dist}(\tau_{(s)}(\Delta(\rho_{(s)})), \partial\mathcal{U}) \geq \text{dist}(\tau_{(s-1)}(\Delta(\rho_{(s-1)})), \partial\mathcal{U}) - \|\Delta_{\tau_{(s-1)}}\|_{\rho_{(s)}},$$

and we use again (56) and (101), thus concluding (computations are analogous to those performed for  $D\tau_{(s)}$  above)

$$\text{dist}(\tau_{(s)}(\Delta(\rho_{(s)})), \partial\mathcal{U}) - \frac{\bar{\alpha} \varepsilon_{(s)}}{\gamma_{(s)}^4 \delta_{(s)}^{4\nu-1}} \geq \text{dist}(\tau_{(0)}(\Delta(\rho_{(0)})), \partial\mathcal{U}) - \frac{2\bar{\alpha} \varepsilon_{(0)}}{\gamma_0^4 \delta_{(0)}^{4\nu-1}} > \sigma_2.$$

Checking fourth and fifth conditions in (105)—which involves estimates (58) for the normal frequencies—is left to the reader, since it can be performed in the same way.

We now observe that hypotheses  $H_1, H_2, H_3, H_4$  and  $H_5$  are automatically satisfied for the  $s$ -objects and Diophantine conditions in  $H_6$  are guaranteed after defining the sets  $I_{(s)}$  of ‘good parameters’. Then, we can apply proposition 5.1 again.

Therefore, we can apply inductively this scheme. Since the sequence of normalized errors satisfies  $\varepsilon_{(s)} \rightarrow 0$  as  $s \rightarrow \infty$  (due to hypothesis (101)) we converge to a true quasi-periodic invariant torus for every  $\mu$  in the set

$$I_{(\infty)} = \bigcap_{s \geq 0} I_{(s)}. \quad (109)$$

Note also that

$$\rho_{(\infty)} = \lim_{s \rightarrow \infty} \rho_{(s)} = \rho_{(0)} - 4 \sum_{s=0}^{\infty} \delta_{(s)} = \rho_{(0)} - 8\delta_{(0)} = \frac{\rho_{(0)}}{2},$$

and that the limit objects are close to the initial (approximate) ones:

$$\begin{aligned} \|\tau_{(\infty)} - \tau_{(0)}\|_{I_{(\infty)}, \rho_{(0)}/2} &\leq \frac{2\bar{\alpha}\varepsilon_{(0)}}{\gamma_0^2 \delta_{(0)}^{2\nu}}, & \|N_{(\infty)} - N_{(0)}\|_{I_{(\infty)}, \rho_{(0)}/2} &\leq \frac{2\bar{\alpha}\varepsilon_{(0)}}{\gamma_0^4 \delta_{(0)}^{4\nu-1}}, \\ |\lambda_{i,(\infty)} - \lambda_{i,(0)}|_{I_{(\infty)}} &\leq \frac{2\bar{\alpha}\varepsilon_{(0)}}{\gamma_0^2 \delta_{(0)}^{2\nu-1}}, \end{aligned}$$

for  $i = 1, \dots, n - r$ . Then, from these expressions we obtain bounds (11) and (10) in the statement of the theorem, just observing that  $\varepsilon_{(0)} \leq \varepsilon_*/\delta_{(0)}$ .

## 6.2. Lipschitz regularity

As we pointed out in section 4.1, to control the measure of the set of removed parameters we cannot use smooth dependence (in the usual sense) with respect to  $\mu$ , because the sets  $I_{(s)}$  have empty interior. Then, following closely [37–40], to control this measure we use a Lipschitz condition from below with respect to  $\mu$  on the eigenvalues of the matrix  $\Lambda_{(s)}$ , for  $s \geq 0$ . In order to guarantee this condition we see that  $\Lambda_{(s)}$  is Lipschitz and then, using that  $\Lambda_{(s)}$  is close to  $\Lambda_{(0)}$ , we prove that  $\Lambda_{(s)}$  is Lipschitz from below. For the sake of completeness, we provide some basic results related to Lipschitz dependence.

**Lemma 6.1.** *Given Lipschitz functions  $f, g : I \subset \mathbb{R} \rightarrow \mathbb{C}$ , we have*

- (i)  $\text{Lip}_I(f + g) \leq \text{Lip}_I(f) + \text{Lip}_I(g)$ .
- (ii)  $\text{Lip}_I(fg) \leq \text{Lip}_I(f)\|g\|_I + \|f\|_I \text{Lip}_I(g)$ .
- (iii)  $\text{Lip}_I(1/f) \leq \|1/f\|_I^2 \text{Lip}_I(f)$ , if  $f$  does not vanish in  $I$ .

Moreover, an equivalent result holds if  $f$  and  $g$  take values in spaces of complex matrices ( $f$  must be invertible in the third item) and also for families  $\mu \mapsto f_\mu$  of functions on  $\mathbb{T}^r$ , using  $\text{Lip}_{I,\rho}(f)$  and  $\|f\|_{I,\rho}$ .

**Proof.** The result is straightforward. □

**Lemma 6.2.** *Given a family  $\mu \in I \subset \mathbb{R} \mapsto f_\mu$ , where  $f_\mu : \mathcal{U} \subset \mathbb{C}^l \rightarrow \mathbb{C}$  is an analytic function with bounded derivatives (that we denote  $D^m f_\mu$ ) in  $\mathcal{U}$ , and given families  $\mu \in I \subset \mathbb{R} \mapsto g_\mu, h_\mu$ , where  $g_\mu, h_\mu : \mathbb{T}^r \rightarrow \mathcal{U}$  are analytic in  $\Delta(\rho)$ , we have*

- (i)  $\text{Lip}_{I,\rho}(f \circ g) \leq \text{Lip}_{I,\mathcal{U}}(f) + \|f\|_{I,\mathbb{C}^1,\mathcal{U}} \text{Lip}_{I,\rho}(g)$ .
- (ii)  $\text{Lip}_{I,\rho}(f \circ g - f \circ h) \leq \beta^{(1)}(\text{Lip}_{I,\rho}(g - h) + \|g - h\|_{I,\rho})$ .

(iii)  $\text{Lip}_{I,\rho}(f \circ g - f \circ h - Df \circ h[g - h]) \leq \beta^{(2)} \|g - h\|_{I,\rho} (\|g - h\|_{I,\rho} + \text{Lip}_{I,\rho}(g - h)).$

The constant  $\beta^{(1)}$  depends on  $\text{Lip}_{I,C^1,\mathcal{U}}(f)$ ,  $\|f\|_{I,C^2,\mathcal{U}}$  and  $\sup_{s \in [0,1]} \text{Lip}_{I,\rho}(h + s(g - h))$ . The constant  $\beta^{(2)}$  depends on  $\text{Lip}_{I,C^2,\mathcal{U}}(f)$ ,  $\|f\|_{I,C^3,\mathcal{U}}$  and  $\sup_{s \in [0,1]} \text{Lip}_{I,\rho}(h + s(g - h))$ .

**Proof.** Item (i) is straightforward. Then, items (ii) and (iii) are obtained using the expressions

$$f_\mu \circ g_\mu - f_\mu \circ h_\mu = \int_0^1 Df_\mu \circ (h_\mu + s(g_\mu - h_\mu)) [g_\mu - h_\mu] ds$$

and

$$f_\mu \circ g_\mu - f_\mu \circ h_\mu - Df_\mu \circ h_\mu [g_\mu - h_\mu] = \int_0^1 D^2 f_\mu \circ (h_\mu + s(g_\mu - h_\mu)) [g_\mu - h_\mu]^{\otimes 2} ds,$$

respectively, and then applying item (i).  $\square$

**Lemma 6.3.** Let  $\mu \in I \subset \mathbb{R} \mapsto g_\mu$  be a family of functions  $g_\mu : \mathbb{T}^r \rightarrow \mathbb{C}$  that are analytic in  $\Delta(\rho)$  and satisfying  $\text{Lip}_{I,\rho}(g) < \infty$ . If we expand  $g$  in Fourier series

$$g_\mu(\theta) = \sum_{k \in \mathbb{Z}^r} \hat{g}_k(\mu) e^{i\langle k, \theta \rangle},$$

then we have

- (i)  $\text{Lip}_I(\hat{g}_k) \leq \text{Lip}_{I,\rho}(g) e^{-|k|_1 \rho}.$
- (ii)  $\text{Lip}_{I,\rho-\delta} \left( \frac{\partial g}{\partial \theta_j} \right) \leq \frac{1}{\delta} \text{Lip}_{I,\rho}(g),$  for  $j = 1, \dots, r.$
- (iii) Given  $\omega \in \mathbb{R}^r \setminus \{0\}$  and a Lipschitz function  $d : I \subset \mathbb{R} \rightarrow \mathbb{C}$ , we consider the sets  $\{d_k^0\}_{k \in \mathbb{Z}^r \setminus \{0\}}, \{d_k^1\}_{k \in \mathbb{Z}^r}$  of complex functions of  $\mu$  given by  $d_k^0 = \langle k, \omega \rangle, d_k^1 = \langle k, \omega \rangle + d(\mu)$ , satisfying  $|d_k^0|, |d_k^1| \geq \gamma/|k|_1^v$ , if  $|k|_1 \neq 0$ , for certain  $1 > \gamma > 0$  and  $v > r - 1$ . Then, the functions  $f^0$  and  $f^1$  whose Fourier coefficients are given by

$$\begin{aligned} \hat{f}_k^0 &= \hat{g}_k / d_k^0, & k \in \mathbb{Z}^r \setminus \{0\}, & \hat{f}_0^0 = 0, \\ \hat{f}_k^1 &= \hat{g}_k / d_k^1, & k \in \mathbb{Z}^r, & \end{aligned}$$

satisfy

$$\begin{aligned} \text{Lip}_{I,\rho-\delta}(f^0) &\leq \frac{\alpha_0}{\gamma \delta^v} \text{Lip}_{I,\rho-\delta}(g), \\ \text{Lip}_{I,\rho-\delta}(f^1) &\leq \beta_0 \left( \frac{\text{Lip}_{I,\rho}(g)}{\gamma \delta^v} + \|g\|_{I,\rho} \frac{\text{Lip}_I(d)}{\gamma^2 \delta^{2v}} \right) + \text{Lip}_{I,\rho}(g) \left\| \frac{1}{d} \right\|_I \\ &\quad + \left\| \frac{1}{d} \right\|_I^2 \text{Lip}_I(d) \|g\|_{I,\rho}, \end{aligned}$$

for any  $\delta \in (0, \min\{1, \rho\})$ , where  $\alpha_0 \geq 1$  is the constant that appears in lemma 5.2, and  $\beta_0 \geq \alpha_0$  is a constant depending on  $r, v$  and  $\alpha_0$ .

**Proof.** Items (i) and (ii) are straightforward (see [38]). Item (iii) follows from the same arguments used in lemma 5.2 and applying the properties in lemma 6.1.  $\square$

Now, we use these elementary results to control recursively the Lipschitz dependence of the constructed objects. To this end, we obtain an ‘extended’ version of proposition 5.1.



**Lemma 6.4 (Addenda to proposition 5.1).** *Let us consider a Lipschitz family of Hamiltonian systems  $\mu \in I \subset \mathbb{R} \mapsto h_\mu$ , where  $I$  is an arbitrary set, with  $h_\mu : U \subset \mathbb{R}^{2n} \rightarrow \mathbb{R}$ , and a vector of basic frequencies  $\omega \in \mathbb{R}^r$ . Assume that there exist families  $\mu \in I \mapsto \tau_\mu, N_\mu, \Lambda_\mu$  satisfying all the hypotheses of proposition 5.1 for every  $\mu \in I$  and also that*

$$\|h\|_{I, C^4, U} \leq \sigma_0, \quad \text{Lip}_{I, C^3, U}(h), \text{Lip}_{I, \rho}(\tau), \text{Lip}_{I, \rho}(D\tau), \text{Lip}_{I, \rho}(N), \text{Lip}_I(\Lambda) < \sigma_5.$$

*Then, there exists a constant  $\tilde{\beta} \geq \tilde{\alpha}$ —where  $\tilde{\alpha}$  is introduced in proposition 5.1—depending on  $r, n, \nu, |\omega|, \sigma_0, \sigma_1, \sigma_2, \sigma_3, \sigma_4$  and  $\sigma_5$ , such that if the condition*

$$\max\{\text{Lip}_{I, \rho}(\tau), \text{Lip}_{I, \rho}(D\tau), \text{Lip}_{I, \rho}(N), \text{Lip}_I(\Lambda)\} + \frac{\tilde{\beta}\hat{\varepsilon}}{\gamma^4\delta^{4\nu-1}} < \sigma_5 \quad (110)$$

*holds, where*

$$\hat{\varepsilon} = \frac{\text{Lip}_{I, \rho}(e)}{\delta} + \text{Lip}_{I, \rho}(R) + \frac{1}{\gamma\delta^\nu} \left( \frac{\|e\|_{I, \rho}}{\delta} + \|R\|_{I, \rho} \right), \quad (111)$$

*then we have that the families  $\mu \mapsto \bar{\tau}_\mu, D\bar{\tau}_\mu, \bar{N}_\mu, \bar{\Lambda}_\mu, \Delta_\tau, \Delta_N, \Delta_\Lambda$  obtained in proposition 5.1 satisfy*

$$\text{Lip}_{I, \rho-2\delta}(\bar{\tau}), \text{Lip}_{I, \rho-3\delta}(D\bar{\tau}), \text{Lip}_{I, \rho-4\delta}(\bar{N}), \text{Lip}_I(\bar{\Lambda}) < \sigma_5, \quad (112)$$

$$\text{Lip}_{I, \rho-2\delta}(\Delta_\tau) \leq \frac{\tilde{\beta}\hat{\varepsilon}}{\gamma^2\delta^{2\nu-1}}, \quad \text{Lip}_{I, \rho-4\delta}(\Delta_N) \leq \frac{\tilde{\beta}\hat{\varepsilon}}{\gamma^4\delta^{4\nu-1}}, \quad \text{Lip}_I(\Delta_\Lambda) \leq \frac{\tilde{\beta}\hat{\varepsilon}}{\gamma^2\delta^{2\nu-1}} \quad (113)$$

*and*

$$\text{Lip}_{I, \rho-3\delta}(\bar{e}) \leq \frac{\tilde{\beta}\hat{\varepsilon}^2}{\gamma^3\delta^{3\nu-2}}, \quad \text{Lip}_{I, \rho-4\delta}(\bar{R}) \leq \frac{\tilde{\beta}\hat{\varepsilon}^2}{\gamma^7\delta^{7\nu-2}}. \quad (114)$$

**Proof.** Basically, it consists in using the properties in lemmata 6.1, 6.2 and 6.3 to control the different functions that appear along the proof of propositions 5.3 (construction of the approximately symplectic basis) and proposition 5.1 (iterative procedure). Since the computations are similar to those detailed in section 5, we will omit some intermediate steps.

First, let us study the objects in proposition 5.3. To this end, we observe that there exists a constant  $\hat{\beta}$  (which is enlarged along the proof in order to include dependence on  $r, n, \nu, |\omega|, \sigma_0, \sigma_1, \sigma_2, \sigma_3, \sigma_4$  and  $\sigma_5$ ) such that

$$\text{Lip}_{I, \rho}(G_{D\tau}^{-1}), \text{Lip}_{I, \rho}(G_{N, N^*}^{-1}), \text{Lip}_{I, \rho}(T_i), \text{Lip}_I([A_1]_{\mathbb{T}^r}^{-1}), \text{Lip}_{I, \rho}(V), \text{Lip}_{I, \rho}(B_j) \leq \hat{\beta},$$

$$\text{Lip}_{I, \rho-\delta}(L_\omega B_i) \leq \frac{\hat{\beta}}{\delta}$$

for  $i = 1, 2$  and  $j = 1, 2, 3$ . For example, we have that

$$\begin{aligned} \text{Lip}_{I, \rho}(G_{D\tau}^{-1}) &\leq \|G_{D\tau}^{-1}\|_{I, \rho}^2 \text{Lip}_{I, \rho}(G_{D\tau}) \\ &\leq \|G_{D\tau}^{-1}\|_{I, \rho}^2 \left( \text{Lip}_{I, \rho}(D\tau^\top) \|D\tau\|_{I, \rho} + \|D\tau^\top\|_{I, \rho} \text{Lip}_{I, \rho}(D\tau) \right) \\ &\leq 4n \|G_{D\tau}^{-1}\|_{I, \rho}^2 \|D\tau\|_{I, \rho} \text{Lip}_{I, \rho}(D\tau) \leq \hat{\beta}. \end{aligned}$$

Then, we estimate Lipschitz constants for the matrices  $\Omega_{D\tau}, \Omega_{N_1}, \dots, \Omega_V$  that characterize the approximately symplectic character of the basis in propositions 5.3. For example, we get  $\text{Lip}_{I, \rho-2\delta}(\Omega_{D\tau})$  by applying item (iii) of lemma 6.3 to the  $(i, j)$ -component of  $\Omega_{D\tau}$  obtained from equation (80). Concretely, taking  $d_k^0 = \langle \omega, k \rangle$  and  $g = -i(\Omega_{De, D\tau} + \Omega_{D\tau, De})^{(i, j)}$ , we obtain

$$\text{Lip}_{I, \rho-2\delta}(\Omega_{D\tau}) \leq \frac{\hat{\beta}}{\gamma\delta^{\nu+1}} (\text{Lip}_{I, \rho}(e) + \|e\|_{I, \rho}).$$

Similarly, to bound  $\text{Lip}_{I,\rho-\delta}(\Omega_{N_1})$  we proceed in the same way taking  $d_k^1 = \langle k, \omega \rangle + \lambda_i + \lambda_j$  and  $g = -i(\Omega_{R,N_1} + \Omega_{N_1,R})^{(i,j)}$ . We obtain

$$\text{Lip}_{I,\rho-\delta}(\Omega_{N_1}) \leq \frac{\hat{\beta}}{\gamma\delta^v} \left( \text{Lip}_{I,\rho}(R) + \frac{\|R\|_{I,\rho}}{\gamma\delta^v} \right).$$

In this way, we have the following bounds in terms of the error  $\hat{\varepsilon}$  defined in (111):

$$\text{Lip}_{I,\rho-2\delta}(\Omega_{D\tau}), \text{Lip}_{I,\rho-\delta}(\Omega_{N_i}), \text{Lip}_{I,\rho-\delta}(\Omega_{N_2,N_1}) \leq \frac{\hat{\beta}\hat{\varepsilon}}{\gamma\delta^v},$$

$$\text{Lip}_{I,\rho-2\delta}(\Omega_{D\tau,N_i}), \text{Lip}_{I,\rho-2\delta}(\Omega_{V,D\tau}), \text{Lip}_{I,\rho-2\delta}(\Omega_{V,N_i}), \text{Lip}_{I,\rho-2\delta}(\Omega_V) \leq \frac{\hat{\beta}\hat{\varepsilon}}{\gamma\delta^v},$$

for  $i = 1, 2$ . Furthermore, by performing similar computations to estimate the Lipschitz constants of  $M_3$  in (86),  $C_1^+$  in (89),  $C_2^+$  in (88),  $C_3^+$  in (87) and  $(\text{Id} + M_3)^{-1}$  in (92) we obtain that the functions  $A_i^+$ , for  $i = 1, \dots, 4$ , in the statement of proposition 5.3 are controlled by

$$\text{Lip}_{I,\rho-2\delta}(A_i^+) \leq \frac{\hat{\beta}\hat{\varepsilon}}{\gamma\delta^{v+1}},$$

provided  $\hat{\varepsilon}$  is small enough—indeed, under condition (51) in proposition 5.1.

Now we can estimate the Lipschitz constants of  $\Delta_i$ ,  $i = 1, \dots, 4$ , defined as the solutions of cohomological equations (37)–(40). In analogy with the notation in proposition 5.1, we introduce a constant  $\bar{\beta} \geq \hat{\beta}$  depending on the same variables as  $\hat{\beta}$ . We have

$$\text{Lip}_{I,\rho-2\delta}(\Delta_1) \leq \frac{\bar{\beta}\hat{\varepsilon}}{\gamma^2\delta^{2v-1}}, \quad \text{Lip}_{I,\rho-\delta}(\Delta_i) \leq \frac{\bar{\beta}\hat{\varepsilon}}{\gamma\delta^{v-1}}, \quad \text{Lip}_{I,\rho-2\delta}(\Delta_\tau) \leq \frac{\bar{\beta}\hat{\varepsilon}}{\gamma^2\delta^{2v-1}},$$

for  $i = 2, 3, 4$ . In particular, we observe that condition (110) guarantees that

$$\text{Lip}_{I,\rho-2\delta}(\bar{\tau}), \text{Lip}_{I,\rho-3\delta}(D\bar{\tau}) < \sigma_5.$$

Now, to control the Lipschitz constant of expression (94) we compute

$$\text{Lip}_{I,\rho-2\delta}(D_i), \text{Lip}_{I,\rho-2\delta}(e^+) \leq \bar{\beta}\delta\hat{\varepsilon}^2, \quad \text{Lip}_{I,\rho-3\delta}(\mathcal{R}(\Delta_\tau) + e) \leq \frac{\bar{\beta}\hat{\varepsilon}^2}{\gamma\delta^v},$$

for  $i = 1, 2, 3, 4$ . Finally, an estimate for the Lipschitz constant of  $\bar{e}$  follows by applying item (iii) in lemma 6.2 that allows controlling the Taylor remainder, thus obtaining

$$\text{Lip}_{I,\rho-3\delta}(\bar{e}) \leq \frac{\bar{\beta}\hat{\varepsilon}^2}{\gamma^3\delta^{3v-2}}.$$

Similarly we control the Lipschitz constant of the new normal eigenvalues of the new reducibility error. As in section 5, we start by controlling

$$\text{Lip}_{I,\rho-2\delta}(\hat{R}) \leq \frac{\bar{\beta}\hat{\varepsilon}}{\gamma^2\delta^{2v-1}}, \quad \text{Lip}_I(\Delta_\Lambda) \leq \frac{\bar{\beta}\hat{\varepsilon}}{\gamma^2\delta^{2v-1}},$$

where we used item (ii) in lemma 6.2. Then we apply lemma 6.3 in order to obtain

$$\text{Lip}_{I,\rho-3\delta}(P_i) \leq \frac{\bar{\beta}\hat{\varepsilon}}{\gamma^3\delta^{3\delta-1}}, \quad \text{Lip}_{I,\rho-4\delta}(P_1) \leq \frac{\bar{\beta}\hat{\varepsilon}}{\gamma^4\delta^{4\delta-1}},$$

for  $i = 2, 3, 4$ . From these estimates it follows that

$$\text{Lip}_{I,\rho-4\delta}(\hat{\Delta}_N), \text{Lip}_I(d_i), \text{Lip}_I(b_i), \text{Lip}_{I,\rho-4\delta}(\Delta_N) \leq \frac{\bar{\beta}\hat{\varepsilon}}{\gamma^4\delta^{4v-1}},$$

and we observe that condition (110) guarantees that

$$\text{Lip}_{I,\rho-4\delta}(\bar{N}), \text{Lip}_I(\bar{\Lambda}) < \sigma_5.$$

In order to control the Lipschitz constant of the new error in reducibility  $\bar{R}$  we have to compute

$$\text{Lip}_{I, \rho-2\delta}(R^+) \leq \frac{\bar{\beta}\hat{\varepsilon}^2}{\gamma^2\delta^{2v-1}}, \quad \text{Lip}_{I, \rho-4\delta}(S(\hat{\Delta}_N, \Delta_\Lambda) + \hat{R}) \leq \frac{\bar{\beta}\hat{\varepsilon}^2}{\gamma^3\delta^{3v}}$$

and

$$\text{Lip}_{I, \rho-4\delta}(S(\Delta_N, \Delta_\Lambda) + \hat{R}) \leq \frac{\bar{\beta}\hat{\varepsilon}^2}{\gamma^7\delta^{7v-2}}.$$

Finally, estimate (114) for the Lipschitz constant of  $\bar{R}$  follows by applying lemma 6.2 that allows us to control the Taylor remainder in (98).  $\square$

In order to prove that the Lipschitz dependence is preserved along the iterative scheme, we only have to check—together with conditions for the convergence of the quadratic method—that condition (110) is satisfied at every step of the procedure. As in section 6.1, assuming that we have applied  $s$  times proposition 5.1 and lemma 6.4, we have to ensure that

$$\max\{\text{Lip}_{I_{(s-1)}, \rho_{(s)}}(\tau_{(s)}), \text{Lip}_{I_{(s-1)}, \rho_{(s)}}(D\tau_{(s)}), \text{Lip}_{I_{(s-1)}, \rho_{(s)}}(N_{(s)}), \text{Lip}_{I_{(s-1)}}(\Lambda_{(s)})\} + \frac{\bar{\beta}\hat{\varepsilon}_{(s)}}{\gamma_{(s)}^4\delta_{(s)}^{4v-1}} < \sigma_5$$

To this end, we compute the normalized error (111) at the  $s$ th step

$$\hat{\varepsilon}_{(s)} = \frac{\text{Lip}_{I_{(s-1)}, \rho_{(s)}}(e_{(s)})}{\delta_{(s)}} + \text{Lip}_{I_{(s-1)}, \rho_{(s)}}(R_{(s)}) + \frac{1}{\gamma_{(s)}\delta_{(s)}^v} \left( \frac{\|e_{(s)}\|_{I_{(s-1)}, \rho_{(s)}}}{\delta_{(s)}} + \|R_{(s)}\|_{I_{(s-1)}, \rho_{(s)}} \right) \quad (115)$$

in terms of  $\hat{\varepsilon}_{(0)}$ . Analogous computations as those performed in (107) show that

$$\hat{\varepsilon}_{(s)} \leq \frac{2^{(s-1)(7v-2)+v+2}\bar{\beta}}{\gamma_0^7\delta_{(0)}^{7v-2}}\hat{\varepsilon}_{(s-1)}^2$$

where we used (57), (60) and (114). Then, by iterating this sequence backwards, we obtain

$$\hat{\varepsilon}_{(s)} \leq \frac{\gamma_0^7\delta_{(0)}^{7v-2}}{\bar{\beta}} 2^{-(s+1)(7v-2)-v-2} \left( \frac{2^{8v}\bar{\beta}\hat{\varepsilon}_{(0)}}{\gamma_0^7\delta_{(0)}^{7v-2}} \right)^{2^s}.$$

The convergence of the Lipschitz procedure follows from similar computations as those in (108) (but using (113)), asking for the condition:

$$\frac{2^{8v}\bar{\beta}\hat{\varepsilon}_{(0)}}{\gamma_0^7\delta_{(0)}^{7v-2}} \leq \frac{1}{2} \min\{1, \sigma_5 - \sigma^{***}\}, \quad (116)$$

where

$$\sigma^{***} = \max\{\text{Lip}_{I, \rho_{(0)}}(\tau_{(0)}), \text{Lip}_{I, \rho_{(0)}}(D\tau_{(0)}), \text{Lip}_{I, \rho_{(0)}}(N_{(0)}), \text{Lip}_I(\Lambda_{(0)})\}.$$

Next, we show that the Lipschitz constants from below of the functions

$$\mu \in I_{(s-1)} \mapsto \lambda_{i,(s)}(\mu), \quad \mu \in I_{(s-1)} \mapsto \lambda_{i,(s)}(\mu) \pm \lambda_{j,(s)}(\mu),$$

for  $i \neq j = 1, \dots, n-r$ , have a lower bound that does not depend on the step  $s$ —note that Lipschitz (from above) constants are controlled for every  $s$  as (112). Indeed, we have

$$\text{Lip}_{I_{(s-1)}}(\Lambda_{(s)} - \Lambda_{(0)}) \leq \sum_{j=0}^{s-1} \text{Lip}_{I_{(j)}}(\Delta_{\Lambda_{(j)}}) \leq \frac{2\bar{\beta}\hat{\varepsilon}_{(0)}}{\gamma_0^2\delta_{(0)}^{2v-1}},$$

where we used condition (116).

Finally, using that

$$\text{lip}_{I_{(s-1)}}(\lambda_{i,(s)}) \geq \text{lip}_I(\lambda_{i,(0)}) - \text{Lip}_{I_{(s-1)}}(\lambda_{i,(s)} - \lambda_{i,(0)})$$

and provided that (we use  $H_6$  in theorem 3.1)

$$\frac{2\bar{\beta}\hat{\varepsilon}_{(0)}}{\gamma_0^2\delta_{(0)}^{2v-1}} < \frac{\sigma_6}{4}, \quad (117)$$

we end up with the bounds

$$\text{lip}_{I_{(s-1)}}(\lambda_{i,(s)}) \geq \frac{\sigma_6}{4}, \quad \text{lip}_{I_{(s-1)}}(\lambda_{i,(s)} \pm \lambda_{j,(s)}) \geq \frac{\sigma_6}{2}, \quad (118)$$

for  $i, j = 1, \dots, n-r$ , with  $i \neq j$ , and  $s \geq 0$ . Analogous computations hold for  $\lambda_{i,(s)}(\mu) \pm \lambda_{j,(s)}(\mu)$ .

Therefore, the quadratic procedure to obtain invariant tori with Lipschitz dependence converges for  $\mu \in I_{(\infty)}$ —see (109)—provided conditions (101), (116) and (117) hold. Moreover, note that we can control the errors  $\varepsilon_{(0)}$  and  $\hat{\varepsilon}_{(0)}$  (given by (100) and (115), respectively) in terms of the error  $\varepsilon_*$  that appears in the statement of theorem 3.1 as follows:

$$\varepsilon_{(0)} \leq \frac{\varepsilon_*}{\delta_{(0)}}, \quad \hat{\varepsilon}_{(0)} \leq \frac{\varepsilon_*}{\gamma_0\delta_{(0)}^{v+1}}. \quad (119)$$

Then we have convergence provided that  $\varepsilon_* \leq C_1\gamma_0^8$ , where  $C_1$  is taken in order to meet all the required conditions.

### 6.3. Measure of the set of excluded parameters

It remains to control the measure of the set  $I_{(\infty)}$  given by (104) and (109), for which all steps performed along the iterative procedure of section 6.1 are well posed. Let us recall that  $I_{(\infty)}$  is constructed by taking out, in recursive form, the set of parameters  $\mu$  for which (4) and (9) do not hold at any step of the KAM process. Concretely, we bound the measure of the complementary set  $I \setminus I_{(\infty)}$ , which we write as

$$I \setminus I_{(\infty)} = (I \setminus I_{(0)}) \cup (I_{(0)} \setminus I_{(\infty)}) = (I \setminus I_{(0)}) \cup \bigcup_{s \geq 1} I_{(s-1)} \setminus I_{(s)}.$$

We start by controlling the measure of  $I_{(s-1)} \setminus I_{(s)}$ , for  $s \geq 1$ . To simplify the notation, in the following discussion we consider a generic divisor of the form  $\langle \omega, k \rangle - d_{(s)}(\mu)$ , where  $d_{(s)}(\mu)$  is either  $\lambda_{j,(s)}(\mu)$  or  $\lambda_{j,(s)}(\mu) \pm \lambda_{i,(s)}(\mu)$ . For this purpose, we introduce the  $k$ th resonant set for the divisor  $d_{(s)}$  as

$$\text{Res}_k^{(s)} = \left\{ \mu \in I_{(s-1)} : |\langle \omega, k \rangle - d_{(s)}(\mu)| < \frac{\gamma_{(s)}}{|k|_1^v} \right\},$$

and we control

$$\text{meas}_{\mathbb{R}} \left( \bigcup_{k \in \mathbb{Z}^r \setminus \{0\}} \text{Res}_k^{(s)} \right) \quad (120)$$

using the following two elementary results.

**Lemma 6.5.** *In the above setting, let us assume that  $\mu \in I_{(s-1)}$ , with  $s \geq 1$ . Then, there exists  $K^*(s) \in \mathbb{N}$  such that  $\mu \notin \text{Res}_k^{(s)}$  provided  $|k|_1 \leq K^*(s)$ . Concretely,*

$$K^*(s) = \lfloor \tilde{\varepsilon}^{-\frac{2s-1}{v}} \rfloor, \quad \text{where} \quad \tilde{\varepsilon} = \frac{2^{8v-1}\bar{\alpha}\varepsilon_{(0)}}{\gamma_0^8\delta_{(0)}^{8v-2}}. \quad (121)$$

**Proof.** To prove this result, we observe that the correction of the normal frequencies—and hence of  $d_{(s)}(\mu)$ —is smaller at each step of the iterative procedure. Indeed, using (58) at the  $s$ th step, (106), (107) and (119), we have that

$$\|d_{(s)} - d_{(s-1)}\|_{I_{(s-1)}} \leq 2\|\Delta_{\Lambda_{(s-1)}}\|_{I_{(s-1)}} \leq \gamma_0^6 \delta_{(0)}^{6\nu-1} 2^{-(6\nu-1)s-2\nu+1} \tilde{\varepsilon}^{2^{s-1}}.$$

Then, given  $\mu \in I_{(s-1)}$ , it turns out that  $\mu \notin \text{Res}_k^{(s)}$  provided the quantity

$$\begin{aligned} |\langle \omega, k \rangle - d_{(s)}(\mu)| &\geq |\langle k, \omega \rangle - d_{(s-1)}(\mu)| - \|d_{(s)} - d_{(s-1)}\|_{I_{(s-1)}} \\ &\geq \gamma_{(s-1)} |k|_1^{-\nu} - \gamma_0^6 \delta_{(0)}^{6\nu-1} 2^{-(6\nu-1)s-2\nu+1} \tilde{\varepsilon}^{2^{s-1}} \end{aligned}$$

is larger than  $\gamma_{(s)} |k|_1^{-\nu}$ . This is equivalent to ask for (recall that  $\gamma_{(s)} = (1 + 2^{-s})\gamma_0$ )

$$\gamma_0^6 \delta_{(0)}^{6\nu-1} 2^{-(6\nu-1)s-2\nu+1} \tilde{\varepsilon}^{2^{s-1}} \leq 2^{-s} \gamma_0 |k|_1^{-\nu},$$

which is satisfied for every  $|k|_1 \leq K^*(s)$ , where  $K^*(s)$  is given in (121).  $\square$

**Lemma 6.6.** *Let us consider a vector  $\omega \in \mathbb{R}^r$  satisfying Diophantine conditions (4) of  $(\hat{\gamma}, \hat{\nu})$ -type, with  $\hat{\gamma} > 0$  and  $\hat{\nu} > r - 1$ . Then, given  $J \subset [\alpha, \beta] \subset \mathbb{R}$ , with  $\alpha > 0$ ,  $\gamma > 0$ ,  $\nu > \hat{\nu}$  and  $K \in \mathbb{N}$ , we have that the measure of the set*

$$\Delta_K = \left\{ d \in J : |\langle k, \omega \rangle - d| < \frac{\gamma}{|k|_1^\nu}, \text{ for some } k \in \mathbb{Z}^r, \text{ with } |k|_1 > K \right\} \quad (122)$$

is controlled as

$$\text{meas}_{\mathbb{R}}(\Delta_K) \leq 2^{\hat{\nu}+1} \hat{\nu}(\beta - \alpha) \frac{\gamma}{\hat{\gamma}} \sum_{j>K} \frac{1}{j^{\nu-\hat{\nu}+1}}. \quad (123)$$

**Proof.** Let us introduce the following notation:

$$\text{Res}_k = \left\{ d \in J : |\langle k, \omega \rangle - d| < \frac{\gamma}{|k|_1^\nu} \right\}$$

if  $|k|_1 > K$ ,  $\text{Res}_k = \emptyset$  if  $|k|_1 \leq K$ , and also  $\widetilde{\text{Res}}_j = \bigcup_{|k|_1=j} \text{Res}_k$ . Then, we have that

$$\text{meas}_{\mathbb{R}}(\Delta_K) = \text{meas}_{\mathbb{R}}\left(\bigcup_{j>K} \widetilde{\text{Res}}_j\right) \leq \sum_{j>K} \text{meas}_{\mathbb{R}}(\widetilde{\text{Res}}_j),$$

thus reducing the problem to study the sets  $\widetilde{\text{Res}}_j$ , which only contain resonances of order  $j$ . Now, let us observe that the width of one resonant set  $\text{Res}_k$  of order  $j$  is controlled by  $\text{meas}_{\mathbb{R}}(\text{Res}_k) \leq 2\gamma j^{-\nu}$ . Hence, it remains to estimate the number of resonant sets  $\text{Res}_k$ , with  $|k|_1 = j$ , that intersect  $J$ . This follows using similar arguments as in lemma 5.2.

For this purpose we introduce some notation. We define  $d_k = \langle k, \omega \rangle$ , which corresponds to the exact resonant value  $d \in \text{Res}_k$ , and

$$\tilde{D}_j = [\alpha, \beta] \cap \left( \bigcup_{|k|_1=j} \{d_k\} \right), \quad \hat{D}_j = \bigcup_{l \leq j} \tilde{D}_l.$$

With these ingredients, we have

$$\text{meas}_{\mathbb{R}}(\Delta_K) \leq \sum_{j>K} \#(\tilde{D}_j) \frac{2\gamma}{j^\nu} = 2\gamma \sum_{j>K} \#\hat{D}_j \left( \frac{1}{j^\nu} - \frac{1}{(j+1)^\nu} \right) \leq 2\gamma \nu \sum_{j>K} \frac{\#\hat{D}_j}{j^{\nu+1}}, \quad (124)$$

where we used the convention  $\text{Res}_k = \emptyset$  if  $|k|_1 \leq K$ .

In order to estimate  $\#\hat{D}_j$ , we sort the resonances  $d_k$  for  $|k|_1 \leq j$  according to

$$\dots < d_{k_{-2}} < d_{k_{-1}} < 0 < d_{k_1} < d_{k_2} < \dots$$

and we observe that  $d_{k_l} \neq d_{k_j}$  if  $k_l \neq k_j$ . Then, using that  $\omega$  is Diophantine of  $(\hat{\gamma}, \hat{\nu})$ -type, we have  $|d_{k_{j+1}} - d_{k_j}| \geq \hat{\gamma}/(2j)^{\hat{\nu}}$ . Hence

$$\#(\hat{D}_j) \leq \frac{|\beta - \alpha| 2^{\hat{\nu}} j^{\hat{\nu}}}{\hat{\gamma}}.$$

By introducing this expression into (124) we end up with (123).  $\square$

Now let us control (120). First we use lemma 6.5 in order to restrict the indexes in (120) to  $k \in \mathbb{Z}^r \setminus \{0\}$  such that  $|k|_1 > K^*(s)$ —see (121). Then, we use lemma 6.6 to control the corresponding measure. To do that, we observe that  $\text{lip}_{I_{(s-1)}}(d_{(s)}) \geq \sigma_6/4$  by (118) and that  $d_{(s)}(I_{(s-1)}) \subset [\sigma_3/2, \sigma_4]$ —this follows from the fact that (64) is preserved along the iterative procedure. Then, the Lipschitz constant from below of  $d_{(s)}$  allows moving the measure of the ‘resonant’ sets (122), controlled in terms of  $d = d_{(s)}(\mu)$ , to the corresponding measure in terms of  $\mu$ . Hence, we get

$$\text{meas}_{\mathbb{R}}\left(\bigcup_{k \in \mathbb{Z}^r \setminus \{0\}} \text{Res}_k^{(s)}\right) = \text{meas}_{\mathbb{R}}\left(\bigcup_{|k|_1 > K^*(s)} \text{Res}_k^{(s)}\right) \leq 2^{\hat{\nu}+4} \hat{\nu} \left(\frac{\sigma_4 - \frac{\sigma_3}{2}}{\sigma_6}\right) \frac{\gamma_0}{\hat{\gamma}} \sum_{j > K^*(s)} \frac{1}{j^{v-\hat{\nu}+1}},$$

where we used that  $\gamma_{(s)} \leq 2\gamma_0$ . Note that this estimate does not depend on the selected  $d_{(s)}(\mu)$ . Then, we can control the measure of the set  $I_{(0)} \setminus I_{(\infty)}$  multiplying the obtained bound by  $2(n-r)^2$  to take into account all possible combinations of normal frequencies. Concretely, we have

$$\text{meas}_{\mathbb{R}}(I_{(0)} \setminus I_{(\infty)}) \leq \frac{\overbrace{2^{\hat{\nu}+5} \hat{\nu} (n-r)^2}^C}{\hat{\gamma}} \left(\frac{\sigma_4 - \frac{\sigma_3}{2}}{\sigma_6}\right) \gamma_0 \sum_{s=1}^{\infty} \sum_{j > K^*(s)} \frac{1}{j^{v-\hat{\nu}+1}} \leq \frac{C\gamma_0}{v-\hat{\nu}} \sum_{s=1}^{\infty} \frac{1}{K^*(s)^{v-\hat{\nu}}},$$

where we used  $v > \hat{\nu}$ . Then, we introduce expression (121) for  $K^*(s)$  and using that  $\tilde{\varepsilon} \leq 1/2$  (see (101)) and that  $s \leq 2^{s-1}$ , we have

$$\text{meas}_{\mathbb{R}}(I_{(0)} \setminus I_{(\infty)}) \leq \frac{C\gamma_0}{v-\hat{\nu}} \sum_{s=1}^{\infty} (\tilde{\varepsilon}^{\frac{v-\hat{\nu}}{v}})^{2^{s-1}} \leq \frac{C\gamma_0}{v-\hat{\nu}} \sum_{s=1}^{\infty} 2^{-s} \leq \frac{C}{v-\hat{\nu}} \gamma_0.$$

Finally, we estimate the measure of the set  $I \setminus I_{(0)} = I_{(-1)} \setminus I_{(0)}$  using the same arguments, and it turns out that it is also proportional to  $\gamma_0$ . Hence, we obtain (12) as follows:

$$\text{meas}_{\mathbb{R}}(I \setminus I_{(\infty)}) \leq \text{meas}_{\mathbb{R}}(I \setminus I_{(0)}) + \text{meas}_{\mathbb{R}}(I_{(0)} \setminus I_{(\infty)}) \leq C_3 \gamma_0.$$

## Acknowledgments

The authors wish to thank Rafael de la Llave for valuable comments and suggestions, and also Yannick Sire for encouragement and for interesting discussions on [24] before it appeared. In addition, they thank the anonymous referees for their careful reading of this paper and for their valuable suggestions. The authors have been partially supported by MCyT/FEDER grant MTM2006-00478, MICINN-FEDER grant MTM2009-06973 and CUR-DIUE grant 2009SGR859. Moreover, the research of AL has been supported by the fellowship grant FPU AP2005-2950 and PTA2008-1693-P.

## References

- [1] Arnold V I 1963 Proof of a theorem of A. N. Kolmogorov on the invariance of quasi-periodic motions under small perturbations *Russ. Math. Surv.* **18** 9–36
- [2] Arnold V I, Kozlov V V and Neishtadt A I 2006 *Mathematical Aspects of Classical and Celestial Mechanics (Encyclopaedia of Mathematical Sciences vol 3)* 3rd edn (Berlin: Springer)
- [3] Bolotin S V and Treschev D V 2000 Remarks on the definition of hyperbolic tori of Hamiltonian systems *Regular Chaotic Dyn.* **5** 401–12
- [4] Bost J B 1986 Tores invariants des systèmes dynamiques hamiltoniens (d'après Kolmogorov, Arnold, Moser, Rüssmann, Zehnder, Herman, Pöschel,...) *Astérisque* **133-134** 113–157 (Seminar Bourbaki, vol 1984/85)
- [5] Bourgain J 1994 Construction of quasi-periodic solutions for Hamiltonian perturbations of linear equations and applications to nonlinear PDE *Int. Math. Res. Not.* **1994** 475–97
- [6] Bourgain J 1997 On Melnikov's persistency problem *Math. Res. Lett.* **4** 445–58
- [7] Broer H W, Ciocci M C, Hanßmann H and Vanderbauwhede A 2009 Quasi-periodic stability of normally resonant tori *Physica D* **238** 309–318
- [8] Broer H W, Hoo J and Naudot V 2007 Normal linear stability of quasi-periodic tori *J. Diff. Eqns* **232** 355–418
- [9] Broer H W, Hanßmann H, Jorba À, Villanueva J and Wagener F 2003 Normal-internal resonances in quasi-periodically forced oscillators: a conservative approach *Nonlinearity* **16** 1751–91
- [10] Broer H W, Huitema G B and Sevryuk M B 1996 *Quasi-periodic motions in families of dynamical systems. Order amidst chaos (Lecture Notes in Mathematics vol 1645)* (Berlin: Springer)
- [11] Broer H W, Huitema G B and Takens F 1990 Unfoldings of quasi-periodic tori *Mem. Am. Math. Soc.* **83** 1–81, 171–5
- [12] Bruno A D 1989 Normalization of a Hamiltonian system near an invariant cycle or torus *Russ. Math. Surv.* **44** 53–89
- [13] Calleja R and de la Llave R 2009 Fast numerical computation of quasi-periodic equilibrium states in 1-D statistical mechanics, including twist maps *Nonlinearity* **22** 1311–36
- [14] Castellà E and Jorba À 2000 On the vertical families of two-dimensional tori near the triangular points of the bicircular problem *Celestial Mech. Dyn. Astronom.* **76** 35–54
- [15] Celletti A and Chierchia L 1997 On the stability of realistic three-body problems *Commun. Math. Phys.* **186** 413–49
- [16] de la Llave R 2001 A tutorial on KAM theory *Smooth Ergodic Theory and its Applications (Seattle, WA, 1999) (Proceedings of Symposia in Pure Mathematics vol 69)* (Providence, RI: American Mathematical Society) pp 175–292
- [17] de la Llave R, González A, Jorba À and Villanueva J 2005 KAM theory without action-angle variables *Nonlinearity* **18** 855–95
- [18] de la Llave R, Huguet G and Sire Y 2009 Fast numerical algorithms for the computation of invariant tori in Hamiltonian Systems, Preprint available electronically at [www.ma.utexas.edu/mp\\_arc-bin/mpa?yn=09-2](http://www.ma.utexas.edu/mp_arc-bin/mpa?yn=09-2)
- [19] de la Llave R and Wayne C E 2004 Whiskered and low dimensional tori in nearly integrable Hamiltonian systems *Math. Phys. Electron. J.* **10** Paper 5, 45pp (electronic)
- [20] Delshams A and Gutiérrez P 1996 Estimates on invariant tori near an elliptic equilibrium point of a Hamiltonian system *J. Diff. Eqns* **131** 277–303
- [21] Eliasson L H 1988 Perturbations of stable invariant tori for Hamiltonian systems *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)* **15** 115–47
- [22] Eliasson L H 1992 Floquet solutions for the 1-dimensional quasi-periodic Schrödinger equation *Commun. Math. Phys.* **146** 447–82
- [23] Eliasson L H 2001 Almost reducibility of linear quasi-periodic systems *Smooth Ergodic Theory and its Applications (Seattle, WA, 1999) (Proceedings of Symposia in Pure Mathematics vol 69)* (Providence, RI: American Mathematics Society) pp 679–705
- [24] Fontich E, de la Llave R and Sire Y 2009 Construction of invariant whiskered tori by a parameterization method: I. Maps and flows in finite dimensions *J. Diff. Eqns* **246** 3136–213
- [25] Gómez G, Jorba À, Simó C and Masdemont J 2001 *Dynamics and Mission Design Near Libration Points, vol III Advanced Methods for Collinear Points (World Scientific Monograph Series in Mathematics vol 4)* (River Edge, NJ: World Scientific)
- [26] Gómez G, Jorba À, Simó C and Masdemont J 2001 *Dynamics and Mission Design Near Libration Points, vol IV, Advanced Methods for Triangular Points (World Scientific Monograph Series in Mathematics vol 5)* (River Edge, NJ: World Scientific)
- [27] Graff S M 1974 On the conservation of hyperbolic invariant tori for Hamiltonian systems *J. Diff. Eqns* **15** 1–69
- [28] Han Y, Li Y and Yi Y 2006 Degenerate lower-dimensional tori in Hamiltonian systems *J. Diff. Eqns* **227** 670–91

- [29] Haro A and de la Llave R 2007 A parameterization method for the computation of invariant tori and their whiskers in quasi-periodic maps: explorations and mechanisms for the breakdown of hyperbolicity *SIAM J. Appl. Dyn. Syst.* **6** 142–207
- [30] Her H-L and You J 2008 Full measure reducibility for generic one-parameter family of quasi-periodic linear systems *J. Dyn. Diff. Eqns* **20** 831–66
- [31] Herman M R 1983 *Sur les Courbes Invariantes par les Difféomorphismes de l'anneau* vol 1 (*Astérisque* vol 103) (Paris: Société Mathématique de France) (With an appendix by Albert Fathi, and an English summary)
- [32] Huang D and Liu Z 2000 On the persistence of lower-dimensional invariant hyperbolic tori for smooth Hamiltonian systems *Nonlinearity* **13** 189–202
- [33] Johnson R A and Moser J 1982 The rotation number for almost periodic potentials *Commun. Math. Phys.* **84** 403–38
- [34] Johnson R A and Sell G R 1981 Smoothness of spectral subbundles and reducibility of quasiperiodic linear differential systems *J. Diff. Eqns* **41** 262–88
- [35] Jorba À, de la Llave R and Zou M 1999 Lindstedt series for lower-dimensional tori *Hamiltonian Systems with Three or More Degrees of Freedom* (S'Agaró, 1995) *NATO Advanced Science Institute Series C: Mathematics and Physical Sciences* vol 533) (Dordrecht: Kluwer) pp 151–67
- [36] Jorba À and Olmedo E 2009 On the computation of reducible invariant Tori on a parallel computer *SIAM J. Appl. Dyn. Syst.* **8** 1382–404
- [37] Jorba À and Simó C 1992 On the reducibility of linear differential equations with quasiperiodic coefficients *J. Diff. Eqns* **98** 111–24
- [38] Jorba À and Simó C 1996 On quasi-periodic perturbations of elliptic equilibrium points *SIAM J. Math. Anal.* **27** 1704–37
- [39] Jorba À and Villanueva J 1997 On the normal behaviour of partially elliptic lower-dimensional tori of Hamiltonian systems *Nonlinearity* **10** 783–822
- [40] Jorba À and Villanueva J 1997 On the persistence of lower-dimensional invariant tori under quasi-periodic perturbations *J. Nonlinear Sci.* **7** 427–73
- [41] Jorba À and Villanueva J 1998 Numerical computation of normal forms around some periodic orbits of the restricted three-body problem *Physica D* **114** 197–229
- [42] Kolmogorov A N 1954 On conservation of conditionally periodic motions for a small change in Hamilton's function *Dokl. Akad. Nauk SSSR (N.S.)* **98** 527–30  
Casati G and Ford J (ed) 1979 *Stochastic Behavior in Classical and Quantum Hamiltonian Systems, Como 1977* (*Lecture Notes in Physics* vol 93) (Berlin: Springer) pp 51–6
- [43] Krikorian R 1999 Réductibilité presque partout des flots fibrés quasi-périodiques à valeurs dans des groupes compacts *Ann. Sci. École Norm. Sup. (4)* **32** 187–240
- [44] Kuksin S B 1988 Perturbation of conditionally periodic solutions of infinite-dimensional Hamiltonian systems *Izv. Akad. Nauk SSSR Ser. Mat.* **52** 41–63, 240  
Kuksin S B 1989 *Math. USSR-Izv.* **32** 39–62 (Engl. Transl.)
- [45] Li Y and Yi Y 2005 Persistence of hyperbolic tori in Hamiltonian systems *J. Diff. Eqns* **208** 344–87
- [46] Lochak P and Meunier C 1988 *Multiphase Averaging for Classical Systems (Applied Mathematical Sciences* vol 72) (New York: Springer)
- [47] Melnikov V K 1965 On some cases of conservation of conditionally periodic motions under a small change of the Hamiltonian function *Sov. Math. Dokl.* **6** 1592–6
- [48] Melnikov V K 1968 A family of conditionally periodic solutions of a Hamiltonian systems *Sov. Math. Dokl.* **9** 882–6
- [49] Michelotti L 1995 *Intermediate Classical Dynamics with Applications to Beam Physics (Wiley Series in Beam Physics and Accelerator Technology)* (New York: Wiley) (A Wiley-Interscience Publication)
- [50] Moser J 1962 On invariant curves of area-preserving mappings of an annulus *Nachr. Akad. Wiss. Göttingen Math.-Phys. Kl. II* **1962** 1–20
- [51] Moser J 1966 On the theory of quasiperiodic motions *SIAM Rev.* **8** 145–72
- [52] Moser J 1966 A rapidly convergent iteration method and non-linear partial differential equations: I *Ann. Scuola Norm. Sup. Pisa (3)* **20** 265–315
- [53] Moser J 1966 A rapidly convergent iteration method and non-linear differential equations: II *Ann. Scuola Norm. Sup. Pisa (3)* **20** 499–535
- [54] Moser J 1967 Convergent series expansions for quasi-periodic motions *Math. Ann.* **169** 136–76
- [55] Moser J and Pöschel J 1984 An extension of a result by Dinaburg and Sinai on quasiperiodic potentials *Comment. Math. Helv.* **59** 39–85
- [56] Neishtadt A I 1985 The separation of motions in systems with rapidly rotating phase *J. Appl. Math. Mech.* **48** 133–9



- [57] Ollé M, Pacha J R and Villanueva J 2008 Kolmogorov–Arnold–Moser aspects of the periodic Hamiltonian Hopf bifurcation *Nonlinearity* **21** 1759–811
- [58] Paskauskas R, Chandre C and Uzer T 2008 Dynamical bottlenecks to intramolecular energy flow *Phys. Rev. Lett.* **100** 083001
- [59] Pöschel J 1989 On elliptic lower-dimensional tori in Hamiltonian systems *Math. Z.* **202** 559–608
- [60] Pöschel J 2001 A lecture on the classical KAM theorem *Smooth Ergodic Theory and its Applications (Seattle, WA, 1999) (Proceedings of Symposia in Pure Mathematics vol 69)* (Providence, RI: American Mathematical Society) pp 707–32
- [61] Puig J Reducibility of linear differential equations with quasi-periodic coefficients: a survey. Preprint available electronically at [www.ma1.upc.edu/~jpuig/preprints/qpred.ps](http://www.ma1.upc.edu/~jpuig/preprints/qpred.ps)
- [62] Rüssmann H 1975 On optimal estimates for the solutions of linear partial differential equations of first order with constant coefficients on the torus *Dynamical Systems, Theory and Applications (Rencontres, Battelle Research Institute, Seattle, WA, 1974) (Lecture Notes in Physics vol 38)* (Berlin: Springer) pp 598–624
- [63] Rüssmann H 1976 On a new proof of Moser’s twist mapping theorem *Proc. 5th Conf. on Mathematical Methods in Celestial Mechanics (Oberwolfach, Germany 1975), Part I. Celestial Mech.* **14** 19–31
- [64] Rüssmann H 2001 Invariant tori in non-degenerate nearly integrable Hamiltonian systems *Regular Chaotic Dyn.* **6** 119–204
- [65] Salamon D and Zehnder E 1989 KAM theory in configuration space *Comment. Math. Helv.* **64** 84–132
- [66] Sevryuk M B 1995 KAM-stable Hamiltonians *J. Dyn. Control Syst.* **1** 351–66
- [67] Sevryuk M B 1996 Invariant tori of Hamiltonian systems that are nondegenerate in the sense of Rüssmann *Dokl. Math.* **53** 69–72
- [68] Sevryuk M B 1999 The lack-of-parameters problem in the KAM theory revisited *Hamiltonian Systems with Three or More Degrees of Freedom (S’Agaró, 1995) (NATO Advanced Science Institutes Series C: Mathematical and Physical Sciences vol 533)* (Dordrecht: Kluwer) pp 568–72
- [69] Sevryuk M B 2007 Invariant tori in quasi-periodic non-autonomous dynamical systems via Herman’s method *Discrete Contin. Dyn. Syst.* **18** 569–95
- [70] Tarquis A M, Losada J C, Benito R M and Borondo F 2002 Multifractal analysis of tori destruction in a molecular Hamiltonian system *Phys. Rev. E* **65** 016213
- [71] Xu J and You J 2001 Persistence of lower-dimensional tori under the first Melnikov’s non-resonance condition *J. Math. Pures Appl. (9)* **80** 1045–67
- [72] Zehnder E 1976 Generalized implicit function theorems with applications to some small divisor problems: II *Commun. Pure Appl. Math.* **29** 49–111