On the Normal Behaviour of Partially Elliptic Lower Dimensional Tori of Hamiltonian Systems

Àngel Jorba and Jordi Villanueva

October 16th, 1996

Departament de Matemàtica Aplicada I
Universitat Politècnica de Catalunya
Diagonal 647, 08028 Barcelona, Spain.
E-mails: `angel@tere.upc.es`, `jordi@tere.upc.es`

Abstract

The purpose of this paper is to study the dynamics near a reducible lower dimensional invariant tori of a finite-dimensional autonomous Hamiltonian system with ℓ degrees of freedom. We will focus in the case in which the torus has (some) elliptic directions.

First, let us assume that the torus is totally elliptic. In this case, it is shown that the diffusion time (the time to move away from the torus) is exponentially big with the initial distance to the torus. The result is valid, in particular, when the torus is of maximal dimension and when it is of dimension 0 (elliptic point). In the maximal dimension case, our results coincide with previous ones. In the zero dimension case, our results improve the existing bounds in the literature.

Let us assume now that the torus (of dimension r, $0 \leq r < \ell$) is partially elliptic (let us call m_e to the number of these directions). In this case we show that, given a fixed number of elliptic directions (let us call $m_1 \leq m_e$ to this number), there exist a Cantor family of invariant tori of dimension $r + m_1$, that generalize the linear oscillations corresponding to these elliptic directions. Moreover, the Lebesgue measure of the complementary of this Cantor set (in the frequency space $\mathbb{R}^{r+m_1}$) is proven to be exponentially small with the distance to the initial torus. This is a sort of "Cantorian central manifold" theorem, in which the central manifold is completely filled up by invariant tori and it is uniquely defined.

The proof of these results is based on the construction of suitable normal forms around the initial torus.

Contents

1 Introduction **3**

2 Summary **4**

2.1 Notation and formulation of the problem 4

2.1.1 Reducibility 5

2.1.2 Linear normal behaviour of the torus 5

2.1.3 Seminormal form: formal description 6

2.2 Results and main ideas 7

2.2.1 Seminormal form: bounds on the remainder 8

2.2.2 Elliptic tori are very sticky 8

2.2.3 Cantor families of invariant tori 9

3 Normal form and effective stability **11**

3.1 Notation 11

3.2 Bounding the remainder of the normal form 12

3.3 Effective stability 21

4 Estimates on the families of lower dimensional tori **23**

4.1 Nondegeneracy conditions 24

4.1.1 Nondegeneracy of the intrinsic frequencies 24

4.1.2 Nondegeneracy of the normal frequencies 26

4.2 Main theorems 26

4.3 Proof of Theorem 4 28

4.3.1 Preliminaries 29

4.3.2 The iterative scheme 32

4.3.3 Convergence of the iterative scheme 37

4.3.4 Bounds on the measure 38

5 Basic lemmas **40**

6 Acknowledgements **44**

References **44**

1 Introduction

The study of the solutions close to an invariant object is a classical subject in Dynamical Systems. Here we will address the problem of describing the phase space near an invariant torus of a Hamiltonian system. To fix the notation, let us call H to a real analytic Hamiltonian with ℓ degrees of freedom, and let us assume it has an invariant r-dimensional torus, $0 \leq r \leq \ell$. Note that we are including the two limit cases, that is, when it is an equilibrium point and when it is a maximal dimensional torus.

To start the discussion, let us assume that the torus has some elliptic directions, this is, that the linearized normal flow contains some harmonic oscillators. A natural question is if these oscillations persist when the nonlinear part of the Hamiltonian is added. If the torus is totally elliptic, another natural problem is the (nonlinear) stability around this torus.

There are known answers to these questions in some concrete cases. If $r = 0$ (the torus is an equilibrium point) and it is totally elliptic, KAM theory says that there is plenty of maximal dimension invariant tori around the point (see [6]): the complementary of the set of invariant tori has measure exponentially small with the distance to the point. It is well known that if $\ell = 2$, the maximal dimensional tori split the energy levels $H = h$ in disconnected components. This is the basis to prove the nonlinear stability of the point. Unfortunately, if $\ell > 2$, the invariant tori do not separate the energy levels. In this case it is generally believed that some diffusion can take place in the phase space (see [2]). Nevertheless, it is still possible to give lower bounds on the diffusion time, that are exponentially big with the distance to the point (they follow immediately from [6]).

If $r = \ell$ (the torus has maximal dimension) we can not speak about normal behaviour since there are not "available directions". The nonlinear stability has been studied in [14] and [13] (among others), where it is shown that the diffusion time is also bounded by an exponentially big (with the distance to the initial torus) quantity. In [13] it is also related this exponentially big stability time with the density of invariant maximal dimensional tori around the initial one, by showing that the total measure of the gaps between the invariant tori nearby is not bigger than an exponentially small quantity with respect to the distance to the initial one. In fact, in [13] it is proved that under an extra steepness condition the diffusion time is, at least, superexponential. This condition corresponds to the classical quasi-convexity hypothesis used to obtain "global" and exponentially big stability time of a perturbed integrable Hamiltonian system with respect to the size of the perturbation (see [5] and references therein).

In this work we will consider these problems, without any steepness condition, for a lower dimensional torus. The two limit cases mentioned above are included, and the results obtained can be summarized as follows: for a totally elliptic torus, we have obtained lower bounds for the diffusion time. They agree with the bounds of [13] in the case $r = \ell$ but, for the case $r = 0$, they are better than the ones directly derived from [6]. Moreover, we show the existence of quasiperiodic solutions that generalize the linear oscillations of the normal flow to the complete system. If the torus has normal behaviour of the kind "some centres"$\times$"some saddles" we obtain, for any combination of centres, a Cantor family of invant tori around the initial one, by adding to the initial set of frequencies new ones that come from the nonlinear oscillations associated to the chosen centres. Those invariant tori have the same normal behaviour as the initial one (of course, skipping the centres that give rise to the family). This result is a sort of "Cantorian central manifold" theorem, in which we obtain an invariant manifold parametrized on a Cantor set and completely filled up by invariant tori. We note that we obtain a Cantorian central "submanifold" for each combination of centres, and that it is uniquely defined.

The proofs are based on the construction of suitable normal forms. The estimates on the difussion time are obtained bounding the remainder of this normal form, while the existence of families of lower dimensional tori is proved by applying a KAM scheme to this remainder.

The paper has been organized in the following way: Section 2 summarizes the main ideas and results contained in the work. Section 3 contains the details concerning the normal form and the bounds on the diffusion time. Section 4 is devoted to the existence of families of tori near the initial one and, finally, in Section 5, we have included some basic lemmas used along the paper.

2 Summary

Here we have included a technical description of the problem, the methodology used in the proofs and the results obtained. We have ommitted the technical details of the proofs in order to simplify the reading.

2.1 Notation and formulation of the problem

Let H be a Hamiltonian system of ℓ degrees of freedom defined on $\mathbb{R}^{2\ell}$, having an invariant r-dimensional isotropic torus (that is, the canonical 2-form of $\mathbb{R}^{2\ell}$ restricted to the tangent bundle of the torus vanish), $0 \leq r \leq \ell$, with a quasiperiodic flow given by the vector of basic frequencies $\hat{\omega}^{(0)} \in \mathbb{R}^r$. We assume, from the isotropic character of the torus, that we can introduce (with a canonical change of coordinates) r angular variables $\hat{\theta}$ describing the initial torus. Hence, the Hamiltonian in these coordinates takes the form

$$H(\hat{\theta}, x, \hat{I}, y) = \hat{\omega}^{(0)\top}\hat{I} + \frac{1}{2}z^\top\mathcal{B}(\hat{\theta})z + H_1(\hat{\theta}, x, \hat{I}, y),$$

where $z^\top = (x^\top, y^\top)$. Here, x, y are m-dimensional real vectors, and $\hat{\theta}$, $\hat{I}$ belong to $\mathbb{R}^r$, $r+m = \ell$. Of course, $\hat{\theta}$, x are the positions and $\hat{I}$, y the respective conjugate momenta. As $\hat{\theta}$ is an angular variable, we assume that H depends on it in a 2π-periodic way. Moreover, we will use $u^\top v$ to denote the scalar product of two vectors.

We also suppose that the Hamiltonian H can be extended to a real analytic function defined on the set $\mathcal{D}_{r,m}(\rho_0, R_0)$ given by

$$\mathcal{D}_{r,m}(\rho_0, R_0) = \{(\hat{\theta}, x, \hat{I}, y) \in \mathbb{C}^r \times \mathbb{C}^m \times \mathbb{C}^r \times \mathbb{C}^m : \ |\mathrm{Im}\hat{\theta}| \leq \rho_0, \ |z| \leq R_0, \ |\hat{I}| \leq R_0^2\}, \quad (1)$$

where $|.|$ denotes the infinity norm of a complex vector (we will use the same notation for the matrix norm induced). The different scaling for the variables z and $\hat{I}$ in $\mathcal{D}_{r,m}(\rho_0, R_0)$ is motivated by the definition of degree for a monomial of the Taylor expansion (with respect to z and $\hat{I}$, see (10)) used along the paper:

$$\deg\left(h_{l,s}(\hat{\theta})z^l\hat{I}^s\right) = |l|_1 + 2|s|_1, \quad (2)$$

with $l \in \mathbb{N}^{2m}$, $s \in \mathbb{N}^r$, and where $|k|_1$ is defined as $\sum_j |k_j|$. The reason for counting twice the exponent s will be clear later (it is motivated, basically, by the properties of the Poisson bracket).

We assume the initial invariant torus is given by $z = 0$ and $\hat{I} = 0$. Hence, we can take $\mathcal{B}(\hat{\theta})$ as a symmetric $2m$-dimensional matrix, with real coefficients that depend on $\hat{\theta}$ in analytic and 2π-periodic way. Moreover, the Taylor expansion of H_1 around $z = 0$, $\hat{I} = 0$ begins with terms of degree at least three.

2.1.1 Reducibility

We will assume that the normal variational flow around this torus (given by the matrix $J_m \mathcal{B}(\hat{\omega}^{(0)} t)$, where J_m is the canonical 2-form of $\mathbb{C}^{2m}$) can be reduced to constant coefficients with a real linear change of variables that depends quasiperiodically on $\hat{\theta}$, having $\hat{\omega}^{(0)}$ as a vector of basic frequencies (quasiperiodic Floquet reduction).[1] The hypothesis does not seem to be very restrictive in our context, since all the partially elliptic tori obtained by KAM techniques have reducible normal flow (see, for instance, [7], [15], [8], [11]). This property allows to construct a canonical change of coordinates that transforms the matrix $\mathcal{B}(\hat{\theta})$ to constant coefficients. Hence, we will assume that $\mathcal{B}$ is a real symmetric matrix, independent from $\hat{\theta}$, and that the initial Hamiltonian in those Floquet variables looks like:

$$H(\hat{\theta}, x, \hat{I}, y) = \hat{\omega}^{(0)\top} \hat{I} + \frac{1}{2} z^\top \mathcal{B} z + H_2(\hat{\theta}, x, \hat{I}, y), \tag{3}$$

where H_2 begins with terms of degree at least three.

2.1.2 Linear normal behaviour of the torus

We also assume that the matrix $J_m \mathcal{B}$ has different eigenvalues, given by the complex vector $\lambda \in \mathbb{C}^{2m}$, that takes the form $\lambda^\top = (\lambda_1, \ldots, \lambda_m, -\lambda_1, \ldots, -\lambda_m)$ (this structure comes from the canonical character of the system). We note that in this case, different eigenvalues also means nonzero eigenvalues. We will refer to those eigenvalues as the normal eigenvalues of the torus. We remark that if $\lambda_j = i\beta$ (with $\beta \in \mathbb{R} \setminus \{0\}$ and $i = \sqrt{-1}$) is an eigenvalue, then $\lambda_{j+m} = -i\beta$. The vectors of $\mathbb{R}^{2m}$ that are combination of eigenvectors corresponding to (couples of) eigenvalues of this form are called the elliptic directions of the torus.

The study of the behaviour of the initial torus in those directions is the main issue in this paper. Moreover, there may be other eigenvalues with real part different from zero, that define the hyperbolic directions of the torus. They can be grouped in one of these two following forms:

1. if $\lambda_j = \alpha \in \mathbb{R} \setminus \{0\}$, then $\lambda_{j+m} = -\alpha$,
2. if $\lambda_j = \alpha + i\beta$ (with $\alpha, \beta \in \mathbb{R} \setminus \{0\}$), then, from the real character of the matrix $\mathcal{B}$, we can take $\lambda_{j+1} = \alpha - i\beta$, and hence, $\lambda_{j+m} = -\alpha - i\beta$ and $\lambda_{j+m+1} = -\alpha + i\beta$.

The imaginary parts of the eigenvalues are usually called normal frequencies of the torus.

For reasons that will be clear later, it is very convenient to put the matrix $J_m \mathcal{B}$ in diagonal form. This is possible with a complex canonical change of basis, that transforms the

[1] If the torus is reducible except by an small remainder, it is still possible to derive similar results (by adding a perturbative parameter). See [9] and [11] for the main ideas and related results.

initial real Hamiltonian system into a complex one. Thus, the complexified Hamiltonian has some symmetries because it comes from a real one. As this symmetries are preserved by the transformations used along the proofs, the final Hamiltonian can be realified. In fact, complexification is not necessary, but it simplifies the proofs. Nevertheless, in the proofs we have not written explicitly the preservation of those symmetries. This is because the details are very tedious and cumbersome and, on the other hand, the interested reader should not have problems in writting them (it is a very standard methodology). For further uses, we denote by $Z^\top = (X^\top, Y^\top)$ those complex (canonical) variables, and by $\mathcal{B}^*$ the complex symmetric matrix such that $J_m\mathcal{B}^* = \mathrm{diag}(\lambda)$.

2.1.3 Seminormal form: formal description

Now we take a subbundle of $\mathbb{R}^{2m}$, $\mathcal{G} \subset \mathbb{R}^{2m}$, invariant by the action of the matrix $J_m\mathcal{B}$, and such that it only contains eigenvectors of elliptic type. We put $2m_1 = \dim(\mathcal{G})$ (we recall that this dimension is always even) and we call $\tilde{\omega}^{(0)} \in \mathbb{R}^{m_1}$ to the vector of normal frequencies associated to this subbundle. As $\mathcal{G}$ will be fixed along the paper, we introduce some notation related to it. First, we assume that the first m_1 eigenvalues of λ are the ones associated to $\mathcal{G}$, that is, $\lambda_j = i\tilde{\omega}_j^{(0)}$, $j = 1, \ldots, m_1$. We also denote by $\hat{\lambda} \in \mathbb{C}^{2(m-m_1)}$ the vector obtained skipping from λ the $2m_1$ eigenvalues associated to $\mathcal{G}$. This introduces in a natural way the decomposition $X^\top = (\tilde{X}^\top, \hat{X}^\top)$, $Y^\top = (\tilde{Y}^\top, \hat{Y}^\top)$, obtained taking apart the first m_1 components from the last $m-m_1$. Moreover, we define $\tilde{Z}^\top = (\tilde{X}^\top, \tilde{Y}^\top)$ and $\hat{Z}^\top = (\hat{X}^\top, \hat{Y}^\top)$. A similar notation can be used for any vector $l \in \mathbb{N}^{2m}$, splitting $l^\top = (l_X^\top, l_Y^\top)$, where l_X and l_Y are the exponents of X and Y in the monomial Z^l ($Z^l = X^{l_X}Y^{l_Y}$). Then, we introduce $\omega^{(0)} \in \mathbb{R}^{r+m_1}$ as $\omega^{(0)\top} = (\hat{\omega}^{(0)\top}, \tilde{\omega}^{(0)\top})$, and we ask for a Diophantine condition of the following form,[2]

$$|ik^\top\omega^{(0)} + l^\top\hat{\lambda}| \geq \frac{\mu}{|k|_1^\gamma}, \quad k \in \mathbb{Z}^{r+m_1} \setminus \{0\}, \quad l \in \mathbb{N}^{2(m-m_1)}, \quad 0 \leq |l|_1 \leq 2, \tag{4}$$

being $\mu > 0$ and $\gamma > r + m_1$. This nonresonance condition allows to construct (formally) a seminormal form related to the chosen $\mathcal{G}$. If we express the Hamiltonian in terms of the variables Z, this seminormal form is done by removing from H the monomials of the following form (see (11) for the notations):

$$h_{l,s,k}\exp(ik^\top\hat{\theta})Z^l\hat{I}^s, \quad l \in \mathbb{N}^{2m}, \quad s \in \mathbb{N}^r, \quad k \in \mathbb{Z}^r, \quad |k|_1 + |l_X - l_Y|_1 \neq 0, \quad |\hat{l}|_1 \leq 2, \tag{5}$$

where $\hat{l}$ is the part of l that corresponds to $\hat{\lambda}$. After this normal form process, using the preservation of the symmetries that come from the complexification, we can rewrite this (formal) seminormal form, in terms of suitable real variables, in the following form:

$$H(\hat{\theta}, x, \hat{I}, y) = \omega^{(0)\top}I + \frac{1}{2}\hat{z}^\top\hat{\mathcal{B}}\hat{z} + \mathcal{F}(I) + \frac{1}{2}\hat{z}^\top\mathcal{Q}(I)\hat{z} + O_3(\hat{z}), \tag{6}$$

where, for simplicity, we do not change the name of the Hamiltonian, and where we extend the decomposition introduced above to the variables (x, y). Here, the matrix $\hat{\mathcal{B}}$ is a real

[2]The Diophantine condition can be relaxed when $|l|_1 = 2$ and $l^\top\hat{\lambda}$ only involves hyperbolic eigenvalues. In this case, the results are proved using a combined method based on a fixed point scheme for the hyperbolic directions and a Newton method for the remaining ones. This technique allows to have multiple hyperbolic eigenvalues.

symmetric matrix obtained by projecting $\mathcal{B}$ on the directions given by the eigenvalues corresponding to the eigenvectors $\hat{\lambda}$. I is a compact notation for $I^\top = (\hat{I}^\top, \tilde{I}^\top)$, where the actions $\tilde{I}$ can be taken as $\tilde{I}_j = \frac{1}{2}(x_j^2 + y_j^2)$, $j = 1, \ldots, m_1$, if we choose the real normal form variables (x, y) associated to the considered elliptic directions in adequate (and standard) way (see (44) in the proof of Theorem 2). Of course, $\mathcal{F} = O_2(I)$ and $\mathcal{Q} = O_1(I)$.

Now, we proceed to describe the normal behaviour of the torus derived from this seminormal form. It is not difficult to check that we have the following (formal) quasiperiodic solutions for the canonical equations of (6):

$$
\begin{aligned}
\hat{\theta}(t) &= \left(\hat{\omega}^{(0)} + \frac{\partial \mathcal{F}}{\partial \hat{I}}(I(0))\right) t + \hat{\theta}(0), \\
\hat{I}(t) &= \hat{I}(0), \\
\tilde{x}_j(t) &= \sqrt{2\tilde{I}_j(0)} \sin\left(\left(\tilde{\omega}^{(0)} + \frac{\partial \mathcal{F}}{\partial \tilde{I}_j}(I(0))\right) t + \tilde{\theta}_j(0)\right), \\
\tilde{y}_j(t) &= \sqrt{2\tilde{I}_j(0)} \cos\left(\left(\tilde{\omega}^{(0)} + \frac{\partial \mathcal{F}}{\partial \tilde{I}_j}(I(0))\right) t + \tilde{\theta}_j(0)\right), \\
\hat{z}(t) &= 0.
\end{aligned}
\tag{7}
$$

That is, we obtain a $2(r + m_1)$-dimensional invariant manifold ($\hat{z} = 0$) foliated by a continuous $(r + m_1)$-dimensional family of $(r + m_1)$-dimensional invariant reducible tori, parametrized by $I(0)$. The selection of the parameter $I(0)$ is natural, as $I_1, \ldots, I_{r+m_1}$ are first integrals of the Hamiltonian (6) restricted to the invariant manifold $\hat{z} = 0$. We remark that the tori of the family collapse to lower dimensional ones when any of the $\tilde{I}_j(0)$ become zero. In particular, if we take $I(0) = 0$ we recover the initial r-dimensional one. In fact, for every $0 \le m_2 \le m_1$ we have, for this seminormal form, $\binom{m_1}{m_2}$ different $(r + m_2)$-dimensional families of $(r + m_2)$-dimensional invariant tori. They are associated to every invariant real subbundle contained in $\mathcal{G}$. The skeleton of these families comes from the natural r-dimensional family of r-dimensional tori containing the initial one. This family is associated to the neutral directions of the torus (the neutral directions are conjugated to the tangent ones), an it is obtained taking $\mathcal{G} = 0$ in our notation. Moreover, we also remark that in (7) we only have real tori when all the $\tilde{I}_j \ge 0$. This comes directly from the definition of $\tilde{I}$ as a function of the real normal form variables. To explain this fact let us give the classical example of a 1-dimensional pendulum near the elliptic equilibrium point, $\ddot{x} + \sin(x) = 0$. The linear (normal) frequency at the equilibrium point is 1. Moving the energy level in the real phase space we obtain periodic orbits with frequency smaller than 1. If one wants periodic orbits with frequency bigger than 1, one is forced to extend the phase space from $\mathbb{R}^2$ to $\mathbb{C}^2$, keeping the time in $\mathbb{R}$. This same phenomenon happens when we study the normal elliptic directions of a torus. It is important to note that, for us, a s-dimensional complex torus is a map from $\mathbb{T}^s$ to $\mathbb{C}^{2\ell}$. Hence, we will use the word "dimension" to refer to the real dimension.

2.2 Results and main ideas

A basic result in this paper is the quantitative version of the seminormal form, if we only kill the monomials like (5) up to some finite order. From the estimates on this seminormal form, we deduce (under certain nondegeneracy conditions) that the normal behaviour of the initial torus described in Section 2.1.3 is "correct" in the sense of the classical KAM ideas: the "majority" of these tori really exist (but slightly deformed) in the initial Hamiltonian system. Moreover, we also deduce the long time effective stability of any real trajectory close to a totally elliptic torus. In the following sections we present the explicit description of those results, and we explain the main ideas used in the proofs.

2.2.1 Seminormal form: bounds on the remainder

We start with the Hamiltonian (3), where the normal flow is reduced to constant coefficients. Then, we perform a finite number of (semi)normal form steps, by using suitable canonical transformations that remove the monomials (5) up to a finite degree. This allows to show the convergence of the process on the set $\mathcal{D}_{r,m}(\rho_1, R)$, where ρ_1 is independent from R and R is small enough. By selecting the order up to which the seminormal form is done as a suitable function of R, it is possible to obtain a remainder for the seminormal form which is exponentially small with R. This is contained in Theorem 1.

2.2.2 Elliptic tori are very sticky

Now let us assume that the inital torus has all the normal directions of elliptic type. In this case we can take $\mathcal{G} = \mathbb{R}^{2m}$, the whole set of normal directions.

Then, using the normal form explained above, one can write the initial Hamiltonian as an integrable one plus an exponentially small perturbation. Hence, it is very natural to obtain exponentially big estimates for the diffusion time: the time needed for a real trajectory to go away from the set $\mathcal{D}_{r,m}(\rho, R)$ (for a precise definition of "going away" see Theorem 2) is bigger than

$$T(R) = const.\exp\left(const.\left(\frac{1}{R}\right)^{\frac{2}{\gamma+1}}\right), \tag{8}$$

being the constants on the definition of $T(R)$ independent from R. As usual, we call the exponent $\frac{2}{\gamma+1}$ the stability exponent.

Let us compare this result with previous ones. In the case in which the initial torus is of maximal dimension, note that the normal variables $z^\top = (x^\top, y^\top)$ are missing everywhere. So, the set $\mathcal{D}_{r,m}(\rho, R)$ (see (1)) reads

$$\mathcal{D}_{\ell,0}(\rho, R) = \{(\hat{\theta}, \hat{I}) \in \mathbb{C}^\ell \times \mathbb{C}^\ell :\ |\mathrm{Im}\hat{\theta}| \le \rho,\ ,\ |\hat{I}| \le R^2\}.$$

To compare with [13] we must redefine R^2 as R, in order to have the same units. Then, the stability exponent in (8) coincides with [13].

If the initial torus is an equilibrium point, the variables $\hat{\theta}$ and $\hat{I}$ are the ones that are missing. Hence, $\mathcal{D}_{r,m}(\rho, R)$ becomes

$$\mathcal{D}_{0,\ell}(0, R) = \{(x, y) \in \mathbb{C}^\ell \times \mathbb{C}^\ell :\ |(x, y)| \le R\}.$$

Hence, no rescaling is necessary to compare the diffusion time of (8) with the one derived from [6]: the improvement is that the exponent $\frac{1}{\gamma+1}$ in [6] here becomes $\frac{2}{\gamma+1}$. We note that this improvement is not only on the diffusion time, but also on the measure of the destroyed tori (see Remark 10).

2.2.3 Cantor families of invariant tori

It is clear from Section 2.1.3 that computing the seminormal formal form associated to $\mathcal{G}$ around the initial torus, up to finite order, and skipping the non-integrable remainder, those elliptic directions define a unique $(r+m_1)$-dimensional family of $(r+m_1)$-dimensional tori around the initial r-dimensional one. When we approach the inital torus, the intrinsic frequencies of the tori of the family can be selected such that they tend to $\omega^{(0)}$.

In this case we will show that when we add the remainder of the seminormal form most of these tori still persist in the complete system H, having also reducible normal flow. The normal eigenvalues of these tori are close to the eigenvalues $\hat{\lambda}_j$ (that are the ones not related with $\mathcal{G}$). Of course, due to the different small divisors involved in the problem, we can not prove the persistence of all the invariant tori predicted by the normal form.

The hypotheses needed are usual in KAM methods. The first one is a non-resonance condition involving the frequencies $\hat{\omega}^{(0)}$ and the normal ones λ, that depends on the concrete selection of $\mathcal{G}$ and it is explicitly given in (4). The second hypothesis is a nondegeneracy condition, asking that all the frequencies vary with the actions. Note that, in general, we have more frequencies $(r+m)$ than actions $(r+m_1,\ m_1 \leq m)$. This introduces the classical lack-of-parameters problem when working with lower dimensional torus, that needs a special treatement (for related results, see [4], [16] and [11]). The idea that we have used here is to choose a suitable $(r+m_1)$-dimensional set of parameters, and to ask for the existence of lower dimensional tori associated to some of the values of these parameters. Here, the natural parameter is the vector of intrinsic frequencies $\omega \in \mathbb{R}^{r+m_1}$ of the invariant tori. To use this parametrization we need a typical nondegeneracy condition on the frequency map from I to ω, this is, that this map be a (local) diffeomorphism around $I=0$. This condition can be explicitly formulated computing the normal form of Section 2.1.3 up to degree 4 and it is given in (50). The control of the remaining $m-m_1$ normal frequencies (normal to the $(r+m_1)$-dimensional family of tori) is more difficult, since there are no free parameters to control them. Note that those frequencies are functions of the intrinsic ones. Then, the idea is to eliminate all the frequencies for which the Diophantine conditions needed to construct invariant tori are not satisfied. This will lead us to eliminate values of ω to: a) control the intrinsic frequencies ω and b) control the normal ones as a function of the intrinsic ones. To control the measure of the set of intrisic frequencies for which the associated normal ones are close to resonance, we use the same kind of method of [11]: we ask for a extra set of nondegeneracy conditions for the dependence of these normal frequencies with respect to the intrinsic ones. Those conditions are given in (54). They have already been considered in [12] and [7].

With the formulation given above, the result is that the measure of the complementary of the preserved tori is exponentially small: we introduce

$$\mathcal{U}(A) = \left\{ \omega \in \mathbb{R}^{r+m_1} \,:\, |\omega - \omega^{(0)}| \leq A \right\}, \quad A > 0, \tag{9}$$

and let us define $\mathcal{A}(A)$ as the set of frequencies of $\mathcal{U}(A)$ for which we have reducible invariant tori. Then, if A is small enough, we have

$$\frac{\mathrm{mes}(\mathcal{U}(A) \setminus \mathcal{A}(A))}{\mathrm{mes}(\mathcal{U}(A))} \leq const. \exp\left(-const. \left(\frac{1}{A}\right)^{\frac{1}{\gamma+1}} \right),$$

where mes$(\cdot)$ denotes the Lebesgue measure of $\mathbb{R}^{r+m_1}$, γ is the exponent of the Diophantine condition (4), and the constants that appear in this bound are positive and independent from A. This result is formulated in Theorem 3. Nevertheless, as we have noted in Section 2.1.3, some of the frequencies of $\mathcal{A}(A)$ give rise to complex tori. If one wants to ensure that the obtained tori are real tori, one can look at the formulation of Theorem 4.

Let us describe how those result are proved. For this purpose, we start from the seminormal form provided by Theorem 1, and we assume that the reader is familiar with the standard KAM techniques (see [3] and references therein).

Initially, we have the seminormal form tori parametrized by the vector of "actions" $I \in \mathbb{R}^{r+m_1}$ (see (7)). By using the nondegeneracy condition of (50), we can replace this parameter by the $(r+m_1)$-dimensional vector of frequencies (see Lemma 3 for the details).

The main issue is to kill, for a given frequency, the part of the remainder that obstructs the existence of the corresponding invariant torus in the complete Hamiltonian. This will be done by a standard iterative Newton method. As usual, we need to have some control on the combinations of intrinsic frequencies and normal eigenvalues that appear in the divisors of the series used to keep them satisfying a suitable Diophantine condition (like (4)). This control can be done using the nondegeneracy conditions of (54). As we will start these iterations from an integrable Hamiltonian (at least in the direction $\mathcal{G}$) with an exponentially small perturbation, we can take the μ in (4) of the same order. This produces convergence except for a set of "bad frequencies" with exponentially small measure.

Now, we use Poincaré variables (see (57) and (58)) to introduce extra m_1 angular variables to describe the invariant $(r + m_1)$-dimensional tori of the seminormal form. When we introduce those variables, there is also another source for degeneracy that, essentially, is due to the fact that the family of $(r + m_1)$-dimensional tori comes from an r-dimensional one. It causes the Poincaré variables to become singular when some of the $\tilde{I}_j$ are zero.[3] We remark that this degeneracy corresponds, in the seminormal form (6), to the families of invariant tori of dimension between r and $r+m_1-1$ (assuming $m_1 \geq 1$). If we ask for real invariant tori, this degeneracy also corresponds to the transition manifold from real to complex tori. We remark that we have an exact knowledge of the manifold of degenerate frequencies for the seminormal form, but the exponentially small remainder makes that we only know the set of degenerate frequencies up to an exponentially small error. This is the main reason that forces to refine the seminormal form as we approach to the initial torus. Moreover, the same remarks apply when we look for real tori: we know the boundary of the set of frequencies that are candidate to give a real torus in the complete Hamiltonian, with an exponentially small error. To remove the degeneracy, we will take out a neighbourhood of the frequencies corresponding to the transition manifold. As the terms of the remainder are exponentially small with R, this neighbourhood can be selected with exponentially small measure with respect to R.

Finally, we note that the application of the results mentioned above show that, around the initial torus, there exists (Cantorian) families of tori of dimensions between r and m_e (we recall m_e is the number of elliptic directions of the initial torus), under generic conditions of non-resonance and nondegeneracy.

[3]This is the same problem that appears when we put action-angle variables around an elliptic equilibrium point of a one degree of freedom Hamiltonian system. A neighbourhood of the origin has to be excluded since the change of variables is singular there.

3 Normal form and effective stability

This section contains the technical details of the seminormal form process with rigorous bounds on the remainder, as well as bounds on the diffusion time around an elliptic torus.

3.1 Notation

First, let us introduce some notation. We will consider analytic functions $h(\hat{\theta}, x, \hat{I}, y)$ defined on $\mathcal{D}_{r,m}(\rho, R)$, for some $\rho > 0$ and $R > 0$, and 2π-periodic with respect to $\hat{\theta}$. We denote the Taylor series of h as

$$h = \sum_{(l,s)\in\mathbb{N}^{2m}\times\mathbb{N}^r} h_{l,s}(\hat{\theta}) z^l \hat{I}^s. \tag{10}$$

Moreover, the coefficients $h_{l,s}$ will be expanded in Fourier series,

$$h_{l,s}(\hat{\theta}) = \sum_{k\in\mathbb{Z}^m} h_{l,s,k} \exp{(ik^\top\hat{\theta})}. \tag{11}$$

We will denote by $\bar{h}_{l,s} = h_{l,s,0}$ the average of $h_{l,s}(\hat{\theta})$, and let us define $\tilde{h}_{l,s}(\hat{\theta}) = h_{l,s}(\hat{\theta}) - \bar{h}_{l,s}$. Then, we use the expressions (10) and (11) to introduce the following norms:

$$|h_{l,s}|_\rho = \sum_{k\in\mathbb{Z}^m} |h_{l,s,k}| \exp{(|k|_1\rho)}, \tag{12}$$

$$|h|_{\rho,R} = \sum_{(l,s)\in\mathbb{N}^{2m}\times\mathbb{N}^r} |h_{l,s}|_\rho R^{|l|_1+2|s|_1}. \tag{13}$$

Some basic properties of these norms are given in Section 5. Here we only note that, if those norms are convergent, they are bounds for the supremum norms of $h_{l,s}(\hat{\theta})$ (on the complex strip of width $\rho > 0$) and of h (on $\mathcal{D}_{r,m}(\rho, R)$).

Let us recall the definition of Poisson bracket of two functions depending on $(\hat{\theta}, x, \hat{I}, y)$:

$$\{f, g\} = \frac{\partial f}{\partial\hat{\theta}}\left(\frac{\partial g}{\partial\hat{I}}\right)^\top - \frac{\partial f}{\partial\hat{I}}\left(\frac{\partial g}{\partial\hat{\theta}}\right)^\top + \frac{\partial f}{\partial z} J_m \left(\frac{\partial g}{\partial z}\right)^\top.$$

We use a similar definition when f and g depend on $(\hat{\theta}, X, \hat{I}, Y)$. Note that, with our definition of degree (see (2)), if f and g are homogeneous polynomials and $\{f, g\} \neq 0$, one has

$$\deg{(\{f, g\})} = \deg(f) + \deg(g) - 2. \tag{14}$$

To introduce more notation, let us define $\mathcal{N} = \{(l, s) \in \mathbb{N}^{2m} \times \mathbb{N}^r \,:\, |l|_1 + 2|s|_1 \geq 3\}$, and let $\mathcal{S}$ be a subset of $\mathcal{N}$. We will say that $h \in \mathcal{M}(\mathcal{S})$ if $h_{l,s} = 0$ when $(l, s) \notin \mathcal{S}$. We will also use the following decomposition: given $h \in \mathcal{M}(\mathcal{N})$, we write $h = \mathcal{S}(h) + (\mathcal{N}\setminus\mathcal{S})(h)$, where $\mathcal{S}(h) \in \mathcal{M}(\mathcal{S})$ and $(\mathcal{N}\setminus\mathcal{S})(h) \in \mathcal{M}(\mathcal{N}\setminus\mathcal{S})$.

Let us split $l^\top = (l_x^\top, l_y^\top)$, with $l_x, l_y \in \mathbb{N}^m$. Now, given $h \in \mathcal{M}(\mathcal{S})$, we say that $h \in \overline{\mathcal{M}}(\mathcal{S})$ when $h_{l,s} = 0$ for all $(l, s) \in \mathcal{S}$ such that $l_x \neq l_y$, and $h_{l,s} = \bar{h}_{l,s}$ if $l_x = l_y$. We say that $h \in \widetilde{\mathcal{M}}(\mathcal{S})$ if $\bar{h}_{l,s} = 0$ for all $(l, s) \in \mathcal{S}$ such that $l_x = l_y$. Note that, for any $h \in \mathcal{M}(\mathcal{S})$, we have $h = \overline{\mathcal{S}}(h) + \widetilde{\mathcal{S}}(h)$, with $\overline{\mathcal{S}}(h) \in \overline{\mathcal{M}}(\mathcal{S})$ and $\widetilde{\mathcal{S}}(h) \in \widetilde{\mathcal{M}}(\mathcal{S})$. We remark that the functions in $\overline{\mathcal{M}}(\mathcal{S})$ only depend on $\hat{I}$ and on the products $x_j y_j$, $j = 1, \ldots, m$.

3.2 Bounding the remainder of the normal form

We introduce $\mathcal{S} \subset \mathcal{N}$ in the following form: we recall that the first m_1 components of λ are eigenvalues associated to $\mathcal{G}$, and then, we put $\mathcal{S} = \mathcal{S}_1 \cup \mathcal{S}_2$, with

$$\mathcal{S}_1 = \{(l,s) \in \mathcal{N} : \ |l_{m_1+1}| + \ldots + |l_m| + |l_{m+m_1+1}| + \ldots + |l_{2m}| \le 1\}, \tag{15}$$
$$\mathcal{S}_2 = \{(l,s) \in \mathcal{N} : \ |l_{m_1+1}| + \ldots + |l_m| + |l_{m+m_1+1}| + \ldots + |l_{2m}| = 2\}. \tag{16}$$

This splitting of $\mathcal{S}$ will be used during the proof of Lemma 1, to identify in a precise form the contribution to $\mathcal{M}(\mathcal{S})$ from the Poisson brackets involving monomials of $\mathcal{M}(\mathcal{S})$ (see the bounds (31) and (32)). This is essential to obtain the estimates of Lemma 2.

We take the Hamiltonian $H(\hat{\theta}, x, \hat{I}, y)$ of (3), we write it in the variables Z (introduced at the end of Section 2.1.2) and we decompose it in the following form:

$$H(\hat{\theta}, X, \hat{I}, Y) = \hat{\omega}^{(0)\top}\hat{I} + \frac{1}{2}Z^\top \mathcal{B}^* Z + N(X, \hat{I}, Y) + S(\hat{\theta}, X, \hat{I}, Y) + T(\hat{\theta}, X, \hat{I}, Y), \tag{17}$$

with $N \in \overline{\mathcal{M}}(\mathcal{S})$, $S \in \widetilde{\mathcal{M}}(\mathcal{S})$ and $T \in \mathcal{M}(\mathcal{N} \setminus \mathcal{S})$. We also define $S_1 := \mathcal{S}_1(S)$ and $S_2 := \mathcal{S}_2(S)$. With this formulation, we say that N is in normal form with respect to $\mathcal{S}$, that S contains the terms of H that are not in normal form with respect to $\mathcal{S}$, and that T contains the terms of the Taylor expansion of H not associated to $\mathcal{S}$. We will show that, assuming the Diophantine conditions of (4), we can put H in normal form with respect to $\mathcal{S}$, with a canonical transformation defined around the initial r-dimensional torus, $Z = 0$ and $\hat{I} = 0$, leaving a small remainder of non-resonant terms. This remainder will be exponentially small with respect to R on the set $\mathcal{D}_{r,m}(\rho_1, R)$, provided that R be small enough, and for certain $\rho_1 > 0$ independent from R. This is done with a finite iterative scheme, with general step described in the following lemma:

Lemma 1 *We consider the Hamiltonian H given in* (17)*. We assume that it is defined on $\mathcal{D}_{r,m}(\rho, R)$, with $0 < \rho < 1$ and $0 < R < 1$, and that there exists $\mu_0 > 0$ and $\gamma > r + m_1$ such that*

$$|ik^\top \hat{\omega}^{(0)} + l^\top \lambda| \ge \frac{\mu_0}{(|k|_1 + |l_x - l_y|_1)^\gamma} \quad \forall (l,s) \in \mathcal{S}\ \forall k \in \mathbb{Z}^r \text{ with } |l_x - l_y|_1 + |k|_1 \ne 0,$$

and, given $\delta > 0$, let us introduce $\rho_j = \rho - j\delta$ and $R_j = R\exp(-j\delta)$. Then, we can construct an analytic function $G(\hat{\theta}, X, \hat{I}, Y) \in \widetilde{\mathcal{M}}(\mathcal{S})$, such that for any $0 < \delta \le \rho/8$ we have the following properties :

1. *G is defined on $\mathcal{D}_{r,m}(\rho_1, R_1)$, and if we decompose $G = G_1 + G_2$, being $G_1 = \widetilde{\mathcal{S}}_1(G)$ and $G_2 = \widetilde{\mathcal{S}}_2(G)$, the bounds for $|G|_{\rho_1,R_1}$, $|G_1|_{\rho_1,R_1}$ and $|G_2|_{\rho_1,R_1}$ are given in* (24).

2. *Let us denote by Ψ_t^G the flow at time t of the Hamiltonian system G. Then, if*

$$\Delta \frac{|S|_{\rho,R}}{\delta^{\gamma+2} R^2} \le 1, \tag{18}$$

where Δ depends only on γ and μ_0, we have

$$\Psi_1^G, \Psi_{-1}^G : \mathcal{D}_{r,m}(\rho_4, R_4) \longrightarrow \mathcal{D}_{r,m}(\rho_3, R_3).$$

3. *If we take* $(\hat\theta, X, \hat I, Y) \in \mathcal{D}_{r,m}(\rho_4, R_4)$*, and we put* $(\hat\theta^*, X^*, \hat I^*, Y^*) = \Psi_1^G(\hat\theta, X, \hat I, Y)$*, then we have* $|\hat\theta^* - \hat\theta| \le \delta$*,* $|Z^* - Z| \le R\delta \exp(-1/2)/2$*,* $|\hat I^* - \hat I| \le R^2\delta\exp(-1)$*. The same bounds also hold for* Ψ_{-1}^G*.*

4. Ψ_1^G *transforms*

$$H^{(1)} := H \circ \Psi_1^G = \hat\omega^{(0)\top}\hat I + \frac{1}{2}Z^\top \mathcal{B}^* Z + N^{(1)} + S^{(1)} + T^{(1)}, \tag{19}$$

decomposition analogous to (17)*, with the bounds* (29)*–*(33)*.*

Remark 1 *In the Diophantine condition of the statement of the lemma, we remark that if we write* $\lambda^\top = (\underline{\lambda}^\top, -\underline{\lambda}^\top)$*, then, for any* $l \in \mathbb{N}^{2m}$*, we have* $l^\top\lambda = (l_x - l_y)^\top\underline{\lambda}$*. Hence, this condition is equivalent to the one formulated in* (4)*, and one can take* μ_0 *as the minimum of* μ *and* $\min\{|\hat l^\top\hat\lambda|\}$*, where this last expression is taken on the* $\hat l \in \mathbb{Z}^{2(m-m_1)}$ *with* $0 < |\hat l|_1 \le 2$ *and* $\hat l_x - \hat l_y \ne 0$*. Moreover, we remark that if we take* $\gamma > r + m_1$*, the set of vectors* $\hat\omega^{(0)}$ *and* λ *for which any Diophantine condition of this kind is not satisfied, has zero measure.*

Remark 2 *The canonical transformation generated by* G *has been chosen to remove the term* S *in the decomposition* (17)*, formulating the homological equation in terms of the monomials of degree* 2 *of the Hamiltonian. This is, in fact, a classical (and linearly convergent) normal form scheme.*

Remark 3 *The bounds on* $H^{(1)}$ *given by Lemma* 1 *are not very concrete. This is because we will use this lemma in iterative form, but the estimates used in the first steps will be different from the ones used in a general step of the iterative process. A description of a general step is given in Lemma* 2*.*

Proof: We look for a generating function $G \in \widetilde{\mathcal{M}}(\mathcal{S})$, such that

$$S + \{\hat\omega^{(0)\top}\hat I + \frac{1}{2}Z^\top\mathcal{B}^*Z, G\} = 0.$$

From the definition of the Poisson bracket, we have

$$S + \left(-\frac{\partial G}{\partial\hat\theta}\hat\omega^{(0)} + Z^\top\mathcal{B}^*J_m\left(\frac{\partial G}{\partial Z}\right)^\top\right) = 0.$$

Expanding G and S, we obtain

$$G_{l,s,k} = \frac{S_{l,s,k}}{ik^\top\hat\omega^{(0)} + (l_x - l_y)^\top\underline{\lambda}}, \tag{20}$$

for the admissible scripts (l,s,k) in the expansion of S (otherwise $G_{l,s,k}$ is defined as 0). Then, from the definition of G, we have

$$\begin{aligned}
H^* &:= H\circ\Psi_1^G - (\hat\omega^{(0)\top}\hat I + \frac{1}{2}Z^\top\mathcal{B}^*Z + N + T + \{N+T,G\}) = \\
&= \int_0^1 \frac{d}{dt}\Big(tH + (1-t)(\hat\omega^{(0)\top}\hat I + \frac{1}{2}Z^\top\mathcal{B}^*Z + N + T + \{N+T,G\})\Big)\circ\Psi_t^G dt = \\
&= \int_0^1 \{tH + (1-t)(\hat\omega^{(0)\top}\hat I + \frac{1}{2}Z^\top\mathcal{B}^*Z + N + T + \{N+T,G\}), G\}\circ\Psi_t^G dt + \\
&\quad + \int_0^1 (H - \hat\omega^{(0)\top}\hat I - \frac{1}{2}Z^\top\mathcal{B}^*Z - N - T - \{N+T,G\})\circ\Psi_t^G dt = \\
&= \int_0^1 \{tS + (1-t)\{N+T,G\}, G\}\circ\Psi_t^G dt. \qquad (21)
\end{aligned}$$

Hence, we have $N^{(1)} = N + \overline{\mathcal{S}}(\{N+T,G\}+H^*)$, $S_1^{(1)} = \tilde{\mathcal{S}}_1(\{N,G_1\}+H^*)$, $S_2^{(1)} = \tilde{\mathcal{S}}_2(\{N,G\}+\{T,G_1\}+H^*)$ and $T^{(1)} = T + (\mathcal{N}\setminus\mathcal{S})(\{N+T,G\}+H^*)$. To give the expressions of $S_1^{(1)}$ and $S_2^{(1)}$, we remark that $\{T,G_1\}\in\mathcal{M}(\mathcal{N}\setminus\mathcal{S}_1)$, $\{T,G_2\}\in\mathcal{M}(\mathcal{N}\setminus\mathcal{S})$ and $\{N,G_2\}\in\mathcal{M}(\mathcal{N}\setminus\mathcal{S}_1)$. Those facts are consequence of the definition of $\mathcal{S}_1$, $\mathcal{S}_2$, the structure of $N\in\overline{\mathcal{M}}(\mathcal{S})$, and the properties of the Poisson bracket.

We proceed to describe the effect of the transformation Ψ_1^G and to bound the transformed Hamiltonian $H\circ\Psi_1^G$. For this purpose, we take a fixed value of δ, $0<\delta\le\rho/8$. Then, for any $(l,s)\in\mathcal{S}$, $k\in\mathbb{Z}^r$, with $|l_x-l_y|_1+|k|_1\neq 0$, we have from (20)

$$\begin{aligned}
&\left|\frac{S_{l,s,k}Z^l\exp(ik^\top\hat{\theta})}{ik^\top\hat{\omega}^{(0)}+(l_x-l_y)^\top\underline{\lambda}}\right|_{\rho_1,R_1}\le\\
&\quad\le\frac{(|k|_1+|l_x-l_y|_1)^\gamma}{\mu_0}\exp(-\delta|k|_1-\delta|l|_1)|S_{l,s,k}|R^{|l|_1}\exp(|k|_1\rho)\le\\
&\quad\le\sup_{\alpha\ge1}\{\alpha^\gamma\exp(-\delta\alpha)\}\frac{|S_{l,s,k}|}{\mu_0}R^{|l|_1}\exp(|k|_1\rho).
\end{aligned}\tag{22}$$

Now, using that for any $\gamma>0$ and $\delta>0$,

$$\sup_{\alpha\ge1}\{\alpha^\gamma\exp(-\delta\alpha)\}\le\left(\frac{\gamma}{\delta\exp(1)}\right)^\gamma,\tag{23}$$

we deduce from (22),

$$|G|_{\rho_1,R_1}\le\left(\frac{\gamma}{\delta\exp(1)}\right)^\gamma\frac{|S|_{\rho,R}}{\mu_0}.\tag{24}$$

Moreover, the same bounds hold for $G_1=\mathcal{S}_1(G)$ and $G_2=\mathcal{S}_2(G)$, if one adds the subscripts "1" or "2" to G and S in (24). Hence, using Lemma 5, we have

$$\left|\frac{\partial G}{\partial\hat{I}}\right|_{\rho_2,R_2}\le\frac{|G|_{\rho_1,R_1}}{R^2\exp(-2\delta)(1-\exp(-2\delta))}\le\frac{|G|_{\rho_1,R_1}}{R^2\delta\exp(-2\delta)},\tag{25}$$

$$\left|\frac{\partial G}{\partial Z}\right|_{\rho_2,R_2}\le\frac{|G|_{\rho_1,R_1}}{R\exp(-\delta)(1-\exp(-\delta))}\le\frac{2|G|_{\rho_1,R_1}}{R\delta\exp(-\delta)},\tag{26}$$

$$\left|\frac{\partial G}{\partial\hat{\theta}}\right|_{\rho_2,R_2}\le\frac{|G|_{\rho_1,R_1}}{\delta\exp(1)},$$

where in (25) and (26) we have used that, if $0<\alpha\le1$, then

$$\frac{\alpha}{2}\le1-\exp(-\alpha).\tag{27}$$

Now, to check the bounds

$$\left|\frac{\partial G}{\partial\hat{I}}\right|_{\rho_2,R_2}\le\delta,\quad\left|\frac{\partial G}{\partial Z}\right|_{\rho_2,R_2}\le\frac{R\delta\exp(-1/2)}{2},\quad\left|\frac{\partial G}{\partial\hat{\theta}}\right|_{\rho_2,R_2}\le R^2\delta\exp(-1),$$

we use (18) with the following Δ:

$$\Delta(\gamma,\mu_0)=\left(\frac{\gamma}{\exp(1)}\right)^\gamma\frac{4\exp(1)}{\mu_0}.$$

If we use the notation $\Psi_t^G - Id = (\hat{\Theta}_t^G, \mathcal{X}_t^G, \hat{\mathcal{I}}_t^G, \mathcal{Y}_t^G)$ (see Lemma 9), then we obtain,

$$\begin{array}{rclrcl} |\hat{\mathcal{I}}_t^G|_{\rho_2,R_2} & \leq & R^2\delta\exp(-1), & |\hat{\Theta}_t^G|_{\rho_2,R_2} & \leq & \delta, \\ |\mathcal{Z}_t^G|_{\rho_2,R_2} & \leq & R\delta\exp(-1/2)/2, & & & \end{array} \tag{28}$$

for any $-1 \leq t \leq 1$, being $\mathcal{Z}_t^G = (\mathcal{X}_t^G, \mathcal{Y}_t^G)$. From the bounds of (28), and using the inequality (27), we can also deduce that the transformations Ψ_1^G and Ψ_{-1}^G act as we describe in the statement. Moreover, (28) and Lemma 7 allows to bound (21) as

$$|H^*|_{\rho_4,R_4} \leq |\{S,G\}|_{\rho_2,R_2} + |\{\{T,G\},G\}|_{\rho_3,R_3} + |\{\{N,G\},G\}|_{\rho_3,R_3}. \tag{29}$$

Finally, the same arguments hold to bound the terms of $H^{(1)}$ in (19) by

$$|N^{(1)} - N|_{\rho_4,R_4} \leq |\{N,G\}|_{\rho_2,R_2} + |\{T,G\}|_{\rho_2,R_2} + |H^*|_{\rho_4,R_4}, \tag{30}$$

$$|S_1^{(1)}|_{\rho_4,R_4} \leq |\{N,G_1\}|_{\rho_2,R_2} + |H^*|_{\rho_4,R_4}, \tag{31}$$

$$|S_2^{(1)}|_{\rho_4,R_4} \leq |\{N,G\}|_{\rho_2,R_2} + |\tilde{\mathcal{S}}_2(\{T,G_1\})|_{\rho_2,R_2} + |H^*|_{\rho_4,R_4}, \tag{32}$$

$$|T^{(1)} - T|_{\rho_4,R_4} \leq |\{N,G\}|_{\rho_2,R_2} + |\{T,G\}|_{\rho_2,R_2} + |H^*|_{\rho_4,R_4}. \tag{33}$$

■

Before giving more concrete estimates on the bounds of Lemma 1, we assume that H is in normal form up to certain order p, to be determined later (the reduction of H to this finite normal form will be described in the proof of Theorem 1). Then, taking advantage of this fact, the bounds of Lemma 1 produce better estimates on the different steps of the normal form process (this is done in Lemma 2). This allows to produce a very accurate bound on the final remainder. We want to stress that these bounds are not so good if the initial Hamiltonian is not in normal form up to degree p.

Let us introduce now the following notation: we break the Hamiltonian (17) as

$$N = N_4 + N^*, \quad T = T_3 + T^*, \tag{34}$$

where N_4 contains the monomials of N of degree 4 and T_3 contains the monomials of degree 3 of T. Then, We assume that, for R small enough, we have the bounds

$$\begin{array}{rclrclrcl} |S_1|_{\rho,R} & \leq & \hat{S}R^{p+1}, & |S_2|_{\rho,R} & \leq & \hat{S}R^p, & |S|_{\rho,R} & \leq & \hat{S}R^p, \\ |N_4|_{\rho,R} & \leq & \hat{N}_4R^4, & |N^*|_{\rho,R} & \leq & \hat{N}^*R^6, & & & \\ |T_3|_{\rho,R} & \leq & \hat{T}_3R^3, & |T^*|_{\rho,R} & \leq & \hat{T}^*R^4, & & & \end{array} \tag{35}$$

being $\hat{S}$, $\hat{N}_4$, $\hat{N}^*$, $\hat{T}_3$ and $\hat{T}^*$ positive constants. Here, $p \in \mathbb{N}$, $p \geq 6$, is the order of the previous normal form and will be chosen later.

Lemma 2 *Let us consider the Hamiltonian H of* (17), *with the same hypotheses as in Lemma* 1. *We use the notations* (34), *and we assume* (35). *We also assume that $\hat{S} \leq \hat{S}^*$, $\hat{N}^* \leq \hat{N}^{**}$ and $\hat{T}^* \leq \hat{T}^{**}$, for some $\hat{S}^*$, $\hat{N}^{**}$ and $\hat{T}^{**}$. Let G be the generating function obtained in Lemma* 1, *and let δ, $0 < \delta \leq \rho/8$, be such that*

$$\Delta\frac{\hat{S}R^{p-2}}{\delta^{\gamma+2}} \leq 1,$$

(Δ is given by Lemma 1*).*

Then, there exists a constant Π, depending only on r, m, γ, μ_0, $\hat{N}_4$, $\hat{N}^{**}$, $\hat{S}^*$, $\hat{T}_3$ and $\hat{T}^{**}$, such that the following bounds hold for the transformed Hamiltonian $H \circ \Psi_1^G$,

$$\begin{aligned}
|N^{(1)} - N|_{\rho_4,R_4} &\leq \Pi\hat{S}\left(\frac{R^{p+1}}{\delta^{\gamma+2}} + \frac{R^{2p-1}}{\delta^{2(\gamma+2)}}\right), \\
|S_1^{(1)}|_{\rho_4,R_4} &\leq \Pi\hat{S}R^{p+1}\left(\frac{R^2}{\delta^{\gamma+1}} + \frac{R^4}{\delta^{\gamma+2}} + \frac{R^{p-3}}{\delta^{\gamma+2}} + \frac{R^{p-2}}{\delta^{2(\gamma+2)}}\right), \\
|S_2^{(1)}|_{\rho_4,R_4} &\leq \Pi\hat{S}R^{p}\left(\frac{R^2}{\delta^{\gamma+1}} + \frac{R^3}{\delta^{\gamma+2}} + \frac{R^{p-1}}{\delta^{2(\gamma+2)}}\right), \\
|T^{(1)} - T|_{\rho_4,R_4} &\leq \Pi\hat{S}\left(\frac{R^{p+1}}{\delta^{\gamma+2}} + \frac{R^{2p-1}}{\delta^{2(\gamma+2)}}\right).
\end{aligned}$$

Remark 4 *(A very important one) If p is big enough and $\delta > R$, the dominant term in the bounds of $S_1^{(1)}$ and $S_2^{(1)}$ is given by the factor $R^2/\delta^{\gamma+1}$. This will be the factor of decreasing of those terms during the normal form process and it allows to take δ of order $R^{2/(\gamma+1)}$, that will produce the exponent $2/(\gamma+1)$ in* (36)*. As we have $2/(\gamma+1) < 1$, we can deduce that an adequate selection for p is $p = 8$. This allows to keep bounds like* (35) *during all the iterative process.*

If we start with a "raw" Hamiltonian (without any previous step of normal form) the decreasing factor obtained is of order $R/\delta^{\gamma+1}$, that forces us to select δ of order $R^{1/(\gamma+1)}$. This produces a worse exponent $1/(\gamma+1)$ in (36)*. For instance, let us assume that the normal form has been done around an elliptic equilibrium point. Here the important issue is to note that the bounds obtained when killing degree* 3 *are much worse than the bounds obtained for the other degrees (this has been observed numerically in* [17]*). Hence, to apply the same bounds to all the degrees results in poor estimates.*

Remark 5 *The exponent $2/(\gamma+1)$ in Remark* 4 *can be improved in some very degenerate cases. For instance, let us consider a totally elliptic torus, and we take $\mathcal{G} = \mathbb{R}^{2m}$. Let q be the lowest degree of the monomials of N corresponding to the (formal) normal form of H around the torus (of course, $q \geq 4$). Then, δ can be taken of order $R^{(q-2)/(\gamma+1)}$, that produces the exponent $(q-2)/(\gamma+1)$ in* (36)*.*

Proof: During this proof we will use different constants Π_j, $j \geq 0$, that will depend only on the same parameters as the final constant Π of the statement of the lemma. First, from the bound (24) of Lemma 1, we have that

$$|G_1|_{\rho_1,R_1} \leq \Pi_0\frac{\hat{S}R^{p+1}}{\delta^\gamma}, \quad |G_2|_{\rho_1,R_1} \leq \Pi_0\frac{\hat{S}R^{p}}{\delta^\gamma}, \quad |G|_{\rho_1,R_1} \leq \Pi_0\frac{\hat{S}R^{p}}{\delta^\gamma},$$

where, as in Lemma 1, $\rho_j = \rho - j\delta$ and $R_j = R\exp(-j\delta)$. Then, to obtain the bounds for the different terms of the transformed Hamiltonian, we only need to bound the Poisson brackets that appear in (29)–(33).

To obtain precise estimates, we will look carefully into the critical bounds of the different partial derivatives involved, that is, the ones associated to N_4 and T_3. So, we estimate, separately, the contribution of N_4, N^*, T_3 and T^*, taking into account that N does not depend on $\hat{\theta}$, N_4 is a polynomial of degree 4, and T_3 only contains terms of degree 3. Moreover, to bound $\widetilde{\mathcal{S}}_2(\{T, G_1\})$ we note that (from the definition of $\mathcal{S}_1$ and $\mathcal{S}_2$)

it only contains terms corresponding to $\frac{\partial}{\partial\hat{Z}}$, and not to $\frac{\partial}{\partial\hat{\theta}}$ or $\frac{\partial}{\partial\hat{I}}$. Thus, using the bounds on the Poisson bracket provided by Lemma 6 (see Remark 11 for the case in which one of the terms has finite degree), we have

$$
\begin{aligned}
|\tilde{\mathcal{S}}_2(\{T,G_1\})|_{\rho_2,R_2} &\leq \Pi_1 \hat{S} R^p \left(\frac{R^2}{\delta^{\gamma+1}} + \frac{R^3}{\delta^{\gamma+2}} \right), \\
|\{T,G\}|_{\rho_2,R_2} &\leq \Pi_2 \hat{S} \frac{R^{p+1}}{\delta^{\gamma+2}}, \\
|\{N,G_1\}|_{\rho_2,R_2} &\leq \Pi_3 \hat{S} R^{p+1} \left(\frac{R^2}{\delta^{\gamma+1}} + \frac{R^4}{\delta^{\gamma+2}} \right), \\
|\{N,G\}|_{\rho_2,R_2} &\leq \Pi_4 \hat{S} R^p \left(\frac{R^2}{\delta^{\gamma+1}} + \frac{R^4}{\delta^{\gamma+2}} \right), \\
|\{S,G\}|_{\rho_2,R_2} &\leq \Pi_5 \hat{S} \frac{R^{2p-2}}{\delta^{\gamma+2}}, \\
|\{\{T,G\},G\}|_{\rho_3,R_3} &\leq \Pi_6 \hat{S} \frac{R^{2p-1}}{\delta^{2(\gamma+2)}}, \\
|\{\{N,G\},G\}|_{\rho_3,R_3} &\leq \Pi_7 \hat{S} \left(\frac{R^{2p}}{\delta^{2\gamma+3}} + \frac{R^{2p+2}}{\delta^{2(\gamma+2)}} \right),
\end{aligned}
$$

and finally

$$
|H^*|_{\rho_4,R_4} \leq \Pi_8 \hat{S} \left(\frac{R^{2p-2}}{\delta^{\gamma+2}} + \frac{R^{2p-1}}{\delta^{2(\gamma+2)}} \right).
$$

From that, with a suitable definition of Π as a function of Π_0–Π_8, the bounds of the statement of the lemma are clear, if we recall that we have taken $p \geq 6$. ∎

Now, we are in conditions to formulate a quantitative result about “partial reduction to seminormal form” of the initial Hamiltonian. For this purpose, we consider the Hamiltonian H of (3), written as in (17) in terms of the Z variables. We assume that H is defined on $\mathcal{D}_{r,m}(\rho_0, R_0)$, for some $0 < \rho_0 < 1$ and $0 < R_0 < 1$, with the following bounds: $|N|_{\rho_0,R} \leq \hat{N}R^4$, $|S|_{\rho_0,R} \leq \hat{S}R^3$ and $|T|_{\rho_0,R} \leq \hat{T}R^3$, for any $0 < R \leq R_0$, being $\hat{N}$, $\hat{S}$ and $\hat{T}$, positive constants (independent from R). Then, we prove the following result:

Theorem 1 *We consider the Hamiltonian H of* (17)*, with the hypotheses previously described. We suppose that there exists $\mu_0 > 0$ and $\gamma > r + m_1$ such that*

$$
|ik^\top \hat{\omega}^{(0)} + l^\top \lambda| \geq \frac{\mu_0}{(|k|_1 + |l_x - l_y|_1)^\gamma} \quad \forall (l,s) \in \mathcal{S}\ \forall k \in \mathbb{Z}^r \text{ with } |l_x - l_y|_1 + |k|_1 \neq 0.
$$

Then, for any $R > 0$ small enough (this condition on R depends only on r, m, γ, μ_0, ρ_0, R_0, $\hat{N}$, $\hat{S}$ and $\hat{T}$), there exists an analytical canonical transformation Ψ^R such that

1. *$\Psi^R - Id$ and $(\Psi^R)^{-1} - Id$ are 2π-periodic on $\hat{\theta}$.*

2. $$\Psi^R : \mathcal{D}_{r,m}(3\rho_0/4, R\exp(-\rho_0/4)) \longrightarrow \mathcal{D}_{r,m}(\rho_0, R),$$

   and

   $$(\Psi^R)^{-1} : \mathcal{D}_{r,m}(11\rho_0/16, R\exp(-5\rho_0/16)) \longrightarrow \mathcal{D}_{r,m}(\rho_0, R).$$

3. If we take $(\hat{\theta}, X, \hat{I}, Y) \in \mathcal{D}_{r,m}(3\rho_0/4, R\exp(-\rho_0/4))$ *and we define* $(\hat{\theta}^*, X^*, \hat{I}^*, Y^*) = \Psi^R(\hat{\theta}, X, \hat{I}, Y)$*, then* $|\hat{\theta}^* - \hat{\theta}| \leq \rho_0/16$*,* $|Z^* - Z| \leq R\rho_0 \exp(-1/2)/32$*,* $|\hat{I}^* - \hat{I}| \leq R^2\rho_0 \exp(-1)/16$*. Moreover, the same bounds hold for* $(\Psi^R)^{-1}$ *if* $(\hat{\theta}, X, \hat{I}, Y) \in \mathcal{D}_{r,m}(11\rho_0/16, R\exp(-5\rho_0/16))$*.*

4. Ψ^R *transforms*

$$H^R := H \circ \Psi^R = \hat{\omega}^{(0)\top}\hat{I} + \frac{1}{2}Z^\top \mathcal{B}^* Z + N^R + S^R + T^R,$$

decomposition analogous to (17)*, with the bounds:* $|N^R - N_4|_{3\rho_0/4, R\exp(-\rho_0/4)} \leq const.R^6$*,* $|T^R - T_3|_{3\rho_0/4, R\exp(-\rho_0/4)} \leq const.R^4$*, where* N_4 *and* T_3 *were introduced in* (34)*, and can be computed with a normal form with respect to* $\mathcal{S}$ *up to degree* 4*, and*

$$|S^R|_{3\rho_0/4, R\exp(-\rho_0/4)} \leq const.\exp\left(-const.\left(\frac{1}{R}\right)^{\frac{2}{\gamma+1}}\right) R^8, \qquad (36)$$

being the constats that appear in the bounds of N^R*,* T^R *and* S^R*, positive and independent from* R*. Moreover, for any* R *for which the result holds,* H^R *is in normal form with respect to* $\mathcal{S}$*, at least up to degree* 8*.*

Remark 6 *The dependence of* Ψ^R *on* R *is not continuous but piecewise analytic.*

Remark 7 *From the bounds provided by Lemma* 2 *for the iterative normal form procedure described in Lemma* 1*, this exponentially small bound seems to be the best that one can obtains by using this linearly convergent scheme.*

Proof: The proof is done simultaneously for any $0 < R \leq R_0$. The bounds where R is not written explicitly are independent from R. All these bounds and the different conditions on the smallness of R will depend only on the fixed parameters of the statement. The main idea of this proof is to use Lemma 1 recursively, and to iterate the bounds provided by Lemma 2 for $p = 8$ (see Remark 4). Hence, to use this lemma, we need to put the initial Hamiltonian in normal form with respect to $\mathcal{S}$, up to degree at least 8. For this purpose, we construct recursively the generating functions $G^{(0)}$, $G^{(1)}$,..., $G^{(5)}$, provided by Lemma 1. Putting $H^{(0)} = H$, we can define

$$H^{(n+1)} := H^{(n)} \circ \Psi_1^{G^{(n)}} = H^{(n)} + \{H^{(n)}, G^{(n)}\} + \frac{1}{2!}\{\{H^{(n)}, G^{(n)}\}, G^{(n)}\} + \cdots, \qquad (37)$$

for $n = 0, \ldots, 5$. Let us consider first the expression (37) as a formal transformation. From the property (14) for the Poisson bracket, and from the way in which the different $G^{(j)}$ are selected in Lemma 1 (see Remark 2), we can ensure that the non-resonant terms associated to $\mathcal{S}$ that remain in $H^{(6)}$ are of degree at least 9. To show that this construction is not only formal, we are going to prove the well defined character of the transformations $\Psi_1^{G^{(n)}}$, $n = 0, \ldots, 5$, and to bound $H^{(n)}$, $n = 1, \ldots, 6$. For this purpose, we expand $H^{(n)}$ as in (17), but adding the superscript "(n)" to N, S and T. We define $\delta_0 = \frac{\rho_0}{192}$, to introduce $\rho^{(0)} = \rho_0$, $R^{(0)} = R$, and $\rho^{(n)} = \rho^{(n-1)} - 4\delta_0$, $R^{(n)} = R^{(n-1)}\exp(-4\delta_0)$, $n = 1, \ldots, 6$. Then, we are going to show that taking $\delta \equiv \delta_0$ in Lemma 1, we have for $n = 0, \ldots, 6$, that, if R is small enough,

$$\begin{array}{llllll} |N^{(n)}|_{\rho^{(n)}, R^{(n)}} & \leq & \hat{N}^{(n)}(R^{(n)})^4, & |S^{(n)}|_{\rho^{(n)}, R^{(n)}} & \leq & \hat{S}^{(n)}(R^{(n)})^{n+3}, \\ |T^{(n)}|_{\rho^{(n)}, R^{(n)}} & \leq & \hat{T}^{(n)}(R^{(n)})^3. & & & \end{array} \qquad (38)$$

This is proved by (finite) induction: assuming that (38) holds for some n ($0 \le n \le 5$) and using that, if R is sufficiently small,

$$\Delta \frac{\hat{S}^{(n)}(R^{(n)})^{n+1}}{\delta_0^{\gamma+2}} \le 1 \tag{39}$$

(Δ is provided by Lemma 1), we have

$$\Psi_1^{G^{(n)}},\ \Psi_{-1}^{G^{(n)}} : \mathcal{D}_{r,m}(\rho^{(n+1)}, R^{(n+1)}) \longrightarrow \mathcal{D}_{r,m}(\rho^{(n)} - 3\delta_0, R^{(n)} \exp(-3\delta_0)). \tag{40}$$

Then, the fact that the successive steps increase at least by one the degree of the normal form with respect to $\mathcal{S}$, makes evident the estimates of (38) for $n+1$. For more details one can rewrite, with minor changes, the proof of Lemma 2, using (38) instead of (35). Here, the different R-independent constants $\hat{N}^{(n)}$, $\hat{T}^{(n)}$ and $\hat{S}^{(n)}$, defined recursively for $n = 0, \ldots, 6$, depend only on the same parameters involved in the formulation of the Theorem. We remark that condition (39) for $n = 0, \ldots, 5$, imposes only a finite number of restrictions on R. Let R_0^* the biggest value of R for which they hold.

The next step is to continue with the iterative normal form process, but using Lemma 2 (with $p = 8$) to bound $H^{(n)}$, $6 \le n \le L+1$ (L will be determined below). This will be done in an inductive way, showing that bounds like (35) hold for each $H^{(n)}$, $n \ge 6$. Hence, we add in (35) the superscript "(n)" to S, S_1, S_2, N^* and T^*, and we replace $\hat{S}$, $\hat{N}^*$ and $\hat{T}^*$ by $\hat{S}^{(n)}$, $\hat{N}^{(n)*}$ and $\hat{T}^{(n)*}$. All these bounds have been taken on $\mathcal{D}_{r,m}(\rho^{(n)}, R^{(n)})$, for some $\rho^{(n)}$, $R^{(n)}$ that will be determined below. Initially, for $n = 6$, we can take for instance $\hat{N}_4^{(6)} = \hat{N}^{(6)}$ and $\hat{N}^{(6)*} = \hat{N}^{(6)}/(R_0^*)^2$. The definition of the other super-(6) constants can be done similarly. Before continuing the iterative procedure, we remark that, as the following steps only afect high order terms, N_4 and T_3 remain invariant during all the normal form procedure. Then, Remark 4 suggests the definition $\delta \equiv \delta(R) = (AR)^{2/(\gamma+1)}$, where $A \ge 1$ will be determined later (independently from R). From this value of δ we define, recursively, $\rho^{(n+1)} = \rho^{(n)} - 4\delta$, $R^{(n+1)} = R^{(n)} \exp(-4\delta)$, for $n \ge 6$. To preserve the positiveness of $\rho^{(n)}$, we restrict $n \le L(R)$, being $L(R)$ the greatest integer for which we have $4(L-5)\delta \le \rho_0/8$. This implies the following restriction on L:

$$L \le 5 + \frac{\rho_0}{32} \left(\frac{1}{AR} \right)^{\frac{2}{\gamma+1}}. \tag{41}$$

Hence, we take as L the integer part of (41). This implies $R \exp(-\rho_0/4) \le R^{(n)} \le R$ if $6 \le n \le L+1$. To apply Lemma 2, we assume that for the current Hamiltonian $H^{(n)}$, $6 \le n \le L$, we have $\hat{S}^{(n)} \le \hat{S}^{(6)}$, $\hat{N}^{(n)*} \le \hat{N}^{**}$ and $\hat{T}^{(n)*} \le \hat{T}^{**}$, for some $\hat{N}^{**}$ and $\hat{T}^{**}$ to be precised later (those bounds are necessary to define Π in Lemma 2, independent from n). If for the current value of n we have

$$\Delta \frac{\hat{S}^{(n)}(R^{(n)})^6}{\delta^{\gamma+2}} \le 1, \tag{42}$$

then, the canonical transformation $\Psi_1^{G^{(n)}}$ given by Lemma 1 acts like (40), replacing δ_0 by δ. Therefore, using Lemma 2, and recalling that $\frac{2}{\gamma+1} < 1$, $A \ge 1$ and $R^{(n)} < R < 1$, one obtains the following bounds for the transformed Hamiltonian:

$$|N^{(n+1)} - N^{(n)}|_{\rho^{(n+1)}, R^{(n+1)}} \le \Pi \hat{S}^{(n)} \left(\frac{(R^{(n)})^9}{\delta A^2 R^2} + \frac{(R^{(n)})^{15}}{\delta^2 A^4 R^4} \right) \le \frac{2\Pi \hat{S}^{(n)}}{A^2} (R^{(n)})^6,$$

$$
\begin{aligned}
|S_1^{(n+1)}|_{\rho^{(n+1)},R^{(n+1)}} &\leq \Pi\hat{S}^{(n)}(R^{(n)})^9\left(\frac{(R^{(n)})^2}{A^2R^2}+\frac{(R^{(n)})^4}{\delta A^2R^2}+\frac{(R^{(n)})^5}{\delta A^2R^2}+\frac{(R^{(n)})^6}{\delta^2A^4R^4}\right)\leq \\
&\leq \frac{4\Pi\hat{S}^{(n)}}{A^2}(R^{(n)})^9,
\end{aligned}
$$

$$
|S_2^{(n+1)}|_{\rho^{(n+1)},R^{(n+1)}} \leq \Pi\hat{S}^{(n)}(R^{(n)})^8\left(\frac{(R^{(n)})^2}{A^2R^2}+\frac{(R^{(n)})^3}{\delta A^2R^2}+\frac{(R^{(n)})^7}{\delta^2A^4R^4}\right)\leq \frac{3\Pi\hat{S}^{(n)}}{A^2}(R^{(n)})^8,
$$

$$
|T^{(n+1)}-T^{(n)}|_{\rho^{(n+1)},R^{(n+1)}} \leq \Pi\hat{S}^{(n)}\left(\frac{(R^{(n)})^9}{\delta A^2R^2}+\frac{(R^{(n)})^{15}}{\delta^2A^4R^4}\right)\leq \frac{2\Pi\hat{S}^{(n)}}{A^2}(R^{(n)})^4.
$$

We take $A=\max\left\{1,\sqrt{8\Pi\exp(1)}\right\}$, and then, recalling that $R^{(n+1)}=R^{(n)}\exp(-4\delta)$, we can define inductively ($n\geq 6$),

$$
\begin{aligned}
\hat{S}^{(n+1)} &= \frac{(\exp(4\delta))^9}{\exp(1)}\hat{S}^{(n)},\\
\hat{N}^{(n+1)*} &= (\exp(4\delta))^9\left(\hat{N}^{(n)*}+\frac{1}{\exp(1)}\hat{S}^{(n)}\right),\\
\hat{T}^{(n+1)*} &= (\exp(4\delta))^9\left(\hat{T}^{(n)*}+\frac{1}{\exp(1)}\hat{S}^{(n)}\right).
\end{aligned}
$$

Assuming R small enough such that $\delta\leq 1/72$, we obtain

$$
\begin{aligned}
\hat{S}^{(n)} &= \hat{S}^{(6)}\exp((6-n)(1-36\delta))\leq \hat{S}^{(6)}\exp((6-n)/2), \qquad (43)\\
\hat{N}^{(n)*} &\leq \exp(36\delta(n-6))\left(\hat{N}^{(6)*}+\hat{S}^{(6)}\frac{1}{(\exp(1)-1)}\right),\\
\hat{T}^{(n)*} &\leq \exp(36\delta(n-6))\left(\hat{T}^{(6)*}+\hat{S}^{(6)}\frac{1}{(\exp(1)-1)}\right).
\end{aligned}
$$

As we are only intrerested in those bounds for $n\leq L+1$, from the restriction on L in (41) we can easily introduce n-independent bounds $\hat{N}^{**}$ and $\hat{T}^{**}$ for $\hat{N}^{(n)*}$ and $\hat{T}^{(n)*}$. Now, assuming that all the steps are well defined, if one puts $n\equiv L(R)+1$ in (43), we obtain the exponentially small bound of the statement for $\hat{S}^{(L+1)}$. To justify that we can reach this value, we note that (42) holds for all the previous n, if we restrict R with $\Delta\hat{S}^{(6)}R^3\leq 1$.

Then, to prove the Theorem, we only have to introduce $\Psi^R=\Psi_1^{G^{(0)}}\circ\ldots\circ\Psi_1^{G^{(L)}}$, and hence, $(\Psi^R)^{-1}=\Psi_{-1}^{G^{(L)}}\circ\ldots\circ\Psi_{-1}^{G^{(0)}}$. If those transformations act as it has been said in the statement, the proof is finished. First, and from the domains of definition of the different canonical transformations $\Psi_1^{G^{(n)}}$ (see (40), replacing δ_0 by δ if $n\geq 6$), we deduce that Ψ^R is defined on the domain given in the statement. Moreover, from the bounds for the different components of $\Psi_1^{G^{(n)}}-Id$ given by Lemma 1, and remarking that $6\delta_0+(L-5)\delta\leq\rho_0/16$, the final bounds for Ψ^R-Id follow immediatly. We consider now $(\Psi^R)^{-1}$. In this case, and using the same arguments on $\Psi_{-1}^{G^{(n)}}$, one can check that if we define $\rho_n=11\rho_0/16+n\delta_0$, $R_n=R\exp(-5\rho_0/16+n\delta_0)$ for $n=0,\ldots,6$, and $\rho_n=11\rho_0/16+6\delta_0+(n-6)\delta$, $R_n=R\exp(-5\rho_0/16+6\delta_0+(n-6)\delta)$ for $n=6,\ldots,L+1$, then we have

$$
\Psi_{-1}^{G^{(n)}}:\mathcal{D}_{r,m}(\rho_n,R_n)\longrightarrow\mathcal{D}_{r,m}(\rho_{n+1},R_{n+1}),
$$

for $0\leq n\leq L$. The proof of this fact can be done by combining the bounds on $\Psi_{-1}^{G^{(n)}}-Id$ with the inequality (27). Moreover, this allows to estimate $(\Psi^R)^{-1}-Id$ as it has been done with the case of Ψ^R. ∎

3.3 Effective stability

An immediate consequence of Theorem 1 is that we can bound the diffusion speed around a linearly stable torus of a Hamiltonian system. In this case, we take $\mathcal{G} = \mathbb{R}^{2m}$, and hence, $\mathcal{S} = \mathcal{N}$. Then, we apply Theorem 1, without taking into account the term T of the decomposition (17). In fact, in this case one can rewrite the proofs of Lemmas 1, 2, and Theorem 1, in a simpler form (although the actual formulation also holds in this particular case), to obtain exponentially small bounds for the remainder S^R.

Theorem 2 *We consider the real analytic Hamiltonian* (3) *defined on* $\mathcal{D}_{r,m}(\rho_0, R_0)$ *for some* $0 < \rho_0 < 1$ *and* $R_0 > 0$. *We also assume that all the eigenvalues of* $J_m\mathcal{B}$ *are of elliptic type, and that there exists* $\mu_0 > 0$ *and* $\gamma > \ell$ *such that*

$$|ik^\top \hat{\omega}^{(0)} + l^\top \lambda| \geq \frac{\mu_0}{(|k|_1 + |l_x - l_y|_1)^\gamma} \quad \forall l \in \mathbb{N}^{2m}\ \forall k \in \mathbb{Z}^r \text{ with } |l_x - l_y|_1 + |k|_1 \neq 0.$$

Let $R \in (0, R_0)$, *and let us take real initial conditions at* $t = 0$ *contained in* $\mathcal{D}_{r,m}(0, R)$. *Then, we can define* $\alpha > 2$ *such that, if* R *is small enough, the corresponding trajectories belong to* $\mathcal{D}_{r,m}(0, \alpha R)$ *for any time* $0 \leq t \leq T(R)$, *with*

$$T(R) = const. \exp\left(const. \left(\frac{1}{R}\right)^{\frac{2}{\gamma+1}} \right),$$

being the constants in the definition of $T(R)$ *independent from* R.

Remark 8 *In the proof, and only for technical reasons,* α *depends on* ρ_0. *Nevertheless, we can take* α *as close as we want to* 2 *(by taking an initial* ρ_0 *small enough, see the proof for details), but this implies a reduction on the set of allowed* R, *and on the constants of the stability time.*

The reason that forces to take $\alpha > 2$ *is the norm used for the normal variables. If one takes the Euclidean norm instead of the supremum norm, the condition* $\alpha > 2$ *is replaced by* $\alpha > 1$.

Proof: In order to simplify the proof, we assume that the initial real variables (x, y) of (3) correspond to the ones that put $\mathcal{B}$ in canonical real form, that is,

$$z^\top \mathcal{B} z = \sum_{j=1}^{m} \alpha_j (x_j^2 + y_j^2),$$

with $\lambda_j = i\alpha_j$, $j = 1, \ldots, m$. Moreover, we assume that $R_0 < 1$. We introduce (X, Y) to denote the complexified variables

$$x_j = \frac{X_j + iY_j}{\sqrt{2}}, \quad y_j = \frac{iX_j + Y_j}{\sqrt{2}}, \quad j = 1, \ldots, m, \tag{44}$$

that put the matrix $J_m\mathcal{B}$ in the diagonal form $J_m\mathcal{B}^*$. Then, we can write the Hamiltonian in these variables as

$$\mathcal{H} = \hat{\omega}^{(0)\top} \hat{I} + \frac{1}{2} Z^\top \mathcal{B}^* Z + N(X, \hat{I}, Y) + S(\hat{\theta}, X, \hat{I}, Y),$$

where N can be rewritten as a function of I, $I^\top = (\hat{I}^\top, \tilde{I}^\top)$, with $\tilde{I}_j = iX_jY_j = \frac{1}{2}(x_j^2 + y_j^2)$, and S verifies $\overline{\mathcal{N}}(S) = 0$. This corresponds to the decomposition (17) if one puts $\mathcal{S} = \mathcal{N}$. $\mathcal{H}$ is defined on $\mathcal{D}_{r,m}(\rho_0, R_0/\sqrt{2})$, with bounds of the following form: $|N|_{\rho_0,R} \le \hat{N}R^4$ and $|S|_{\rho_0,R} \le \hat{S}R^3$, for any $0 < R \le R_0/\sqrt{2}$. Now, we apply Theorem 1 and we obtain, for any R small enough, a canonical change Ψ^R such that in the new coordinate system $(\hat{\theta}, X, \hat{I}, Y) = \Psi^R(\hat{\theta}^R, X^R, \hat{I}^R, Y^R)$, we have

$$\mathcal{H}^R := \mathcal{H} \circ \Psi^R = \hat{\omega}^{(0)\top}\hat{I}^R + \frac{1}{2}(Z^R)^\top \mathcal{B}^* Z^R + N^R(X^R, \hat{I}^R, Y^R) + S^R(\hat{\theta}^R, X^R, \hat{I}^R, Y^R),$$

being N^R a function of $(I^R)^\top = ((\hat{I}^R)^\top, (\tilde{I}^R)^\top)$, with $\tilde{I}_j^R = iX_j^R Y_j^R$. $\mathcal{H}^R$ is defined on $\mathcal{D}_{r,m}(3\rho_0/4, R\exp(-\rho_0/4))$, with

$$|S^R|_{3\rho_0/4, R\exp(-\rho_0/4)} \le const. \exp\left(-const. \left(\frac{1}{R}\right)^{\frac{2}{\gamma+1}}\right) R^8 := M(R).$$

The canonical equations for $(X^R, \hat{I}^R, Y^R)$ are

$$\begin{aligned} \dot{X}_j^R &= \frac{\partial \mathcal{H}^R}{\partial Y_j^R}, & \dot{Y}_j^R &= -\frac{\partial \mathcal{H}^R}{\partial X_j^R}, \quad j = 1, \ldots, m, \\ \dot{\hat{I}}_j^R &= \frac{\partial \mathcal{H}^R}{\partial \hat{\theta}_j^R} = \frac{\partial S^R}{\partial \hat{\theta}_j^R}, & & \quad j = 1, \ldots, r. \end{aligned} \tag{45}$$

From this, one obtains (using that N^R is in fact only a function of I, and recalling $\ell = r + m$),

$$\dot{I}_j^R = i\frac{\partial \mathcal{H}^R}{\partial Y_j^R} Y_j^R - i\frac{\partial \mathcal{H}^R}{\partial X_j^R} X_j^R = i\frac{\partial S^R}{\partial Y_j^R} Y_j^R - i\frac{\partial S^R}{\partial X_j^R} X_j^R, \quad j = r+1, \ldots, \ell.$$

We put $\mathcal{I}_j^R(\hat{\theta}^R, X^R, \hat{I}^R, Y^R)$ for the expressions on the right-hand side of $\dot{I}_j^R$, $j = 1, \ldots, \ell$. We use Lemma 5 to bound these expressions. Then, for $j = 1, \ldots, r$ one has

$$|\mathcal{I}_j^R|_{0, R\exp(-\rho_0/2)} \le \frac{4M(R)}{3\rho_0 \exp(1)}. \tag{46}$$

If we combine Lemma 5 with the inequality (27), one obtains for $j = r+1, \ldots, \ell$ that,

$$|\mathcal{I}_j^R|_{0, R\exp(-\rho_0/2)} \le \frac{2M(R)\exp(-\rho_0/2)}{\exp(-\rho_0/4)(1 - \exp(-\rho_0/4))} \le \frac{16M(R)\exp(-\rho_0/4)}{\rho_0}. \tag{47}$$

To continue the proof, we put (x^R, y^R) for the variables that come from the "realification" of (X^R, Y^R), that is, $X_j^R = (x_j^R - iy_j^R)/\sqrt{2}$, $Y_j^R = (y_j^R - ix_j^R)/\sqrt{2}$. In fact, as Ψ^R preserves the symmetries of $\mathcal{H}$ (due to the complexification of a real Hamiltonian, see Section 2.1.2), we have that the Hamiltonian in the variables (x^R, y^R) is real analytic. To work with those different representations of the variables, we give the following remarks: (i) the set of real variables (x_j, y_j) such that $|x_j|, |y_j| \le A$ is contained in the set of complex (X_j, Y_j) such that $|X_j|, |Y_j| \le A$, (ii) the set of complex (X_j, Y_j) such that $|X_j|, |Y_j| \le A$, is contained in the complex set for (x_j, y_j) such that $|x_j|, |y_j| \le \sqrt{2}A$ (this property has been used to say that $\mathcal{H}$ is defined in $\mathcal{D}_{r,m}(\rho_0, R_0/\sqrt{2})$), (iii) the set of real (x_j, y_j) such that $I_j \le A^2$, is contained in the set of real (x_j, y_j) such that $|x_j|, |y_j| \le \sqrt{2}A$. Those remarks are used when working with these different kind of variables, and one wants to control the size of the corresponding domains, when we change the variable representation.

Now, we take real values for $(\hat{\theta}^R, x^R, \hat{I}^R, y^R)$ as initial conditions at $t = 0$. To prove the lower bound for the stability time, we consider a fixed $0 < \beta < 1$, and we restrict to initial conditions such that, when expressed in terms of $(\hat{\theta}^R, X^R, \hat{I}^R, Y^R)$, they belong to $\mathcal{D}_{r,m}(0, R\beta \exp(-\rho_0/2)/\sqrt{2})$. Then, we have that the corresponding initial actions I^R are bounded by $|I_j^R(0)| \le R^2\beta^2 \exp(-\rho_0)/2$, $j = 1, \dots, \ell$. Using this, we deduce from the bounds (46) and (47) that, for the trajectories of the Hamiltonian equations (45), we have $|I_j^R(t)| \le R^2 \exp(-\rho_0)/2$ for $0 \le t \le T(R)$, where we can take

$$T(R) = \frac{R^2\rho_0 \exp(-3\rho_0/4)(1-\beta^2)}{32M(R)}.$$

This bound comes from (47), that is the worst case. This is the expression for the stability time of the statement of the Theorem. To use the bounds (46) and (47) for $\mathcal{I}_j^R$, we need that these trajectories expressed in terms of $(\hat{\theta}^R, X^R, \hat{I}^R, Y^R)$ belong to $\mathcal{D}_{r,m}(0, R\exp(-\rho_0/2))$ up to time $T(R)$. As we have $|I_j^R(t)| \le R^2 \exp(-\rho_0)/2$, this follows from remarks (*iii*) and (*i*). From that we deduce, using the bounds for $(\Psi^R)^{-1} - Id$ provided by Theorem 1 and remark (*ii*), that the corresponding real trajectories in terms of $(\hat{\theta}, x, \hat{I}, y)$ are contained in $\mathcal{D}_{r,m}(0, R_1)$, being R_1 defined by

$$R_1 = \max\left\{ \sqrt{2}\left(R\exp\left(-\frac{\rho_0}{2}\right) + \frac{R\rho_0 \exp(-1/2)}{32} \right), \sqrt{R^2 \exp(-\rho_0) + \frac{R^2\rho_0 \exp(-1)}{16}} \right\}.$$

Then, if we give for $(\hat{\theta}, x, \hat{I}, y)$ a real set of points such that, expressed in terms of $(\hat{\theta}, X, \hat{I}, Y)$, they belong to the domain $(\Psi^R)^{-1}(\mathcal{D}_{r,m}(0, R\beta\exp(-\rho_0/2)/\sqrt{2}))$, then, the trajectories of H with initial conditions in this set remain in $\mathcal{D}_{r,m}(0, R_1)$ for a time span $T(R)$. With similar arguments as the ones used to define R_1 (using now remark (*ii*)), one can check that this domain can be taken as $\mathcal{D}_{r,m}(0, R_2)$, being R_2 defined by

$$R_2 = \min\left\{ \frac{R\beta\exp(-\rho_0/2)}{\sqrt{2}} - \frac{R\rho_0\exp(-1/2)}{32}, \sqrt{\frac{R^2\beta^2\exp(-\rho_0)}{2} - \frac{R^2\rho_0\exp(-1)}{16}} \right\}.$$

If one considers the flow Ψ_t^H defined from $\mathcal{D}_{r,m}(0, R_2) \cap \mathbb{R}^{2\ell}$ to $\mathcal{D}_{r,m}(0, R_1) \cap \mathbb{R}^{2\ell}$, for $0 \le t \le T(R)$, then, putting $R \equiv R_2$ in the statement, and taking an R-independent value of β close enough to 1 such that $R_2 > 0$, we can define $\alpha = R_1/R_2$. ■

4 Estimates on the families of lower dimensional tori

Let us consider the real analytic reduced Hamiltonian H of (3) and a fixed subbundle $\mathcal{G}$ of elliptic directions of $J_m\mathcal{B}$. In Theorem 1 we have proved that, under standard Diophantine conditions, one can put H in normal form with respect to the set $\mathcal{S}$ (see (15) and (16) for the definition), with an exponentially small remainder. If we write this seminormal form in terms of the complexified variables Z, and without changing the name of the Hamiltonian, one has

$$H = \omega^{(0)\top} I + \frac{1}{2}\hat{Z}^\top \hat{\mathcal{B}}^* \hat{Z} + \mathcal{F}(I) + \frac{1}{2}\hat{Z}^\top \mathcal{Q}(I)\hat{Z} + \mathcal{T}(\hat{\theta}, X, \hat{I}, Y) + \mathcal{R}(\hat{\theta}, X, \hat{I}, Y). \quad (48)$$

To explain the notation used, let us recall that the different resonant terms depend only on $\hat{I}$ and on the products X_jY_j, $j=1,\dots,m$, but, from the structure of $\mathcal{S}$, not all the possible combinations of those monomials take place in $\overline{\mathcal{M}}(\mathcal{S})$. Then, we introduce $I^\top = (\hat{I}^\top, \tilde{I}^\top)$, with $\tilde{I}_j = iX_jY_j$, $j=1,\dots,m_1$, and with this definition (48) can be described as follows: the symmetric matrix $\hat{\mathcal{B}}^*$ is defined from $\mathcal{B}^*$ skipping the $2m_1$ eigenvalues associated to $\mathcal{G}$, $J_{m-m_1}\hat{\mathcal{B}}^* = \mathrm{diag}(\hat{\lambda})$. $\mathcal{F}$ and $\mathcal{Q}$ correspond to the normal form with respect to $\mathcal{S}$, with the expansion of $\mathcal{F}$ starting at second order with respect to I, and with $\mathcal{Q}(0)=0$. It is not difficult to check that by choosing the variables $\tilde{Z}$ in suitable form (as it has been done in the proof of Theorem 2), $\mathcal{F}$ is real analytic. Moreover, $\mathcal{Q}$ is a symmetric matrix such that $J_{m-m_1}\mathcal{Q}$ is diagonal, $\mathcal{T} \in \mathcal{M}(\mathcal{N}\setminus\mathcal{S})$ (so $\mathcal{T} \equiv O_3(\hat{Z})$) and $\mathcal{R} \in \widetilde{\mathcal{M}}(\mathcal{S})$.

We assume that this normal form has been done for a given (and small enough) R, as in the formulation of Theorem 1. We only consider the R-dependence when we give the bounds of the different terms of (48). To obtain these bounds, let us define $\rho_1 = 3\rho_0/4$, where we recall that ρ_0 is the width of the strip of analiticity, with respect to $\hat{\theta}$, for the initial Hamiltonian. Then, Theorem 1 implies that, for any R small enough, we have

$$
\begin{array}{rclrcl}
|\mathcal{F}|_{0,R} & \le & \hat{\mathcal{F}}R^4, & |\mathcal{F}_3|_{0,R} & \le & \hat{\mathcal{F}}_3R^6,\\
|\mathcal{Q}|_{0,R} & \le & \hat{\mathcal{Q}}R^2, & |\mathcal{Q}_2|_{0,R} & \le & \hat{\mathcal{Q}}_2R^4,\\
|\mathcal{T}|_{\rho_1,R} & \le & \hat{\mathcal{T}}R^3, & |\mathcal{R}|_{\rho_1,R} & \le & const.\exp\left(-const.\left(\frac{1}{R}\right)^{\frac{2}{\gamma+1}}\right)R^8,
\end{array}
\tag{49}
$$

To derive these bounds on $\mathcal{D}_{r,m}(0,R)$, we have considered the functions that depend on $\tilde{I}$ as functions of $\tilde{Z}$. Here, we have split $\mathcal{F} = \mathcal{F}_2 + \mathcal{F}_3$ and $\mathcal{Q} = \mathcal{Q}_1 + \mathcal{Q}_2$. $\mathcal{F}_2$ and the components of $\mathcal{Q}_1$ are polynomials on I of degrees 2 and 1 respectively. $\mathcal{F}_3$ and $\mathcal{Q}_2$ contain the remaining terms. We note that the definition of $\mathcal{F}_2$ and $\mathcal{Q}_1$ does not depend on the order of the seminormal form.

This seminormal form has been formally explained in Section 2.1.3, and we will use the notation related to (7) to represent the normal form tori.

The main purpose of this section is to study the persistence of those tori when we add the remainder $\mathcal{R}$. We note that, as $|\mathcal{R}|$ is exponentially small with R, we can expect that the tori of (7) will survive, except the ones corresponding to a set of parameters $(I(0))$ of exponentially small measure with respect to R. We will show that this assertion holds, assuming certain standard nondegeneracy conditions on this family of tori, that have been explained in Section 2.2.3 (conditions that, as we will see, can be checked by computing a normal form up to degree 4, that is, from $\mathcal{F}_2$ and $\mathcal{Q}_1$). As it is a more natural parameter, the results will be formulated in terms of frequencies instead of actions.

4.1 Nondegeneracy conditions

Before the rigorous formulation of the results, let us give in explicit form these nondegeneracy conditions.

4.1.1 Nondegeneracy of the intrinsic frequencies

The first one is a standard nondegeneracy condition on the dependence of the frequencies with respect to the actions: we require

$$
\det\mathcal{C} \neq 0, \qquad \mathcal{C} = \frac{\partial^2\mathcal{F}_2}{\partial I^2}(0). \tag{50}
$$

This allows to parametrize the tori of the family by their vector of intrinsic frequencies (instead of $I(0)$). Of course, we have to be close enough to the initial r-dimensional torus. This assertion is justified by the following lemma:

Lemma 3 *Let us assume that* $\det \mathcal{C} \neq 0$. *Then, if* R *is small enough, there exists a real analytic vectorial function* $\mathcal{I}(\omega)$*, defined on the set*

$$\left\{ \omega \in \mathbb{C}^{r+m_1} \,:\, |\omega - \omega^{(0)}| \leq \frac{1}{8}(|\mathcal{C}^{-1}|)^{-1} R^2 \right\}, \tag{51}$$

such that

$$\frac{\partial \mathcal{F}}{\partial I}(\mathcal{I}(\omega)) = \omega - \omega^{(0)},$$

with $\mathcal{I}(\omega^{(0)}) = 0$. *Moreover, we have* $|\mathcal{I}(\omega)| \leq \frac{1}{4}R^2$ *for any* ω *in the set* (51)*, and if* $\omega^{(1)}$, $\omega^{(2)}$ *belong in* (51)*, then*

$$|\mathcal{I}(\omega^{(1)}) - \mathcal{I}(\omega^{(2)})| \leq 2|\mathcal{C}^{-1}||\omega^{(1)} - \omega^{(2)}|.$$

Of course, we are still using the notation of Section 4.

Proof: We have $\mathcal{F}(I) = \frac{1}{2}I^\top \mathcal{C} I + \mathcal{F}_3$. Then, we take a fixed ω in the set (51), and we want to solve the equation:

$$\mathcal{I}(\omega) = \mathcal{C}^{-1}\left(\omega - \omega^{(0)} - \frac{\partial \mathcal{F}_3}{\partial I}(\mathcal{I}(\omega))\right). \tag{52}$$

Putting the superscripts "$(k+1)$" and "(k)" to $\mathcal{I}(\omega)$ in (52), we can consider this expression as an iterative procedure, using $\mathcal{I}^{(0)}(\omega) = 0$ as the seed. If we assume $|\mathcal{I}^{(k)}(\omega)| \leq \frac{1}{4}R^2$, then, using Cauchy inequalities, we have for R small enough,

$$|\mathcal{I}^{(k+1)}(\omega)| \leq |\mathcal{C}^{-1}| \left(\frac{1}{8}(|\mathcal{C}^{-1}|)^{-1}R^2 + \frac{\hat{\mathcal{F}}_3 R^6}{\frac{3}{4}R^2} \right) \leq \frac{1}{4}R^2,$$

where we have used the bounds of (49) for $\mathcal{F}_3$, remarking that $|\mathcal{F}_3|_{0,R}$ is a bound for the supremum norm of $\mathcal{F}_3(I)$ if $|I| \leq R^2$. Moreover, to ensure convergence, we remark that using the main value theorem one has,

$$|\mathcal{I}^{(k+1)}(\omega) - \mathcal{I}^{(k)}(\omega)| \leq (r+m_1)\frac{\hat{\mathcal{F}}_3 R^6}{\left(\frac{3}{8}\right)^2 R^4}|\mathcal{I}^{(k)}(\omega) - \mathcal{I}^{(k-1)}(\omega)| \leq \frac{1}{2}|\mathcal{I}^{(k)}(\omega) - \mathcal{I}^{(k-1)}(\omega)|,$$

if R is small enough. Clearly, the limit function is analytic with respect to ω, and from the real analytic character of $\mathcal{F}$, $\mathcal{I}$ is in fact real analytic. Taking $\omega^{(1)}$, $\omega^{(2)}$ in the set (51), one has

$$\mathcal{I}(\omega^{(1)}) - \mathcal{I}(\omega^{(2)}) = \mathcal{C}^{-1}(\omega^{(1)} - \omega^{(2)}) + \mathcal{C}^{-1}\left(\frac{\partial \mathcal{F}_3}{\partial I}(\mathcal{I}(\omega^{(2)})) - \frac{\partial \mathcal{F}_3}{\partial I}(\mathcal{I}(\omega^{(1)})) \right),$$

and with the same arguments previously used, we obtain for R small enough

$$|\mathcal{I}(\omega^{(1)}) - \mathcal{I}(\omega^{(2)})| \leq 2|\mathcal{C}^{-1}||\omega^{(1)} - \omega^{(2)}|.$$

■

4.1.2 Nondegeneracy of the normal frequencies

The other nondegeneracy condition considered refers to the normal eigenvalues. Skipping again the remainder $\mathcal{R}$, for every invariant tori parametrized by $I(0)$ in (7), the corresponding normal eigenvalues are the ones of the diagonal matrix $J_{m-m_1}(\hat{\mathcal{B}}^* + \mathcal{Q}(I(0)))$. Using the parametrization $I \equiv \mathcal{I}(\omega)$ provided by Lemma 3, we can consider those eigenvalues as functions of ω instead of functions of $I(0)$.

$$J_{m-m_1}(\hat{\mathcal{B}}^* + \mathcal{Q}(\mathcal{I}(\omega))) \equiv \mathrm{diag}(\hat{\lambda}^{(0)}(\omega)), \tag{53}$$

we ask for the condition

$$\mathrm{Im}\left(\frac{\partial}{\partial\omega}(l^\top \hat{\lambda}^{(0)})(\omega^{(0)})\right) \notin \mathbb{Z}^{r+m_1}, \;\; l \in \mathbb{Z}^{2(m-m1)}, \; 0 < |l|_1 \leq 2, \;\; l_{\hat{X}} \neq l_{\hat{Y}}, \tag{54}$$

where we have used the notation $l^\top = (l_{\hat{X}}^\top, l_{\hat{Y}}^\top)$. To check this condition, we only need to know the first order approximation to $\mathcal{Q}$, remarking that from (53) one has

$$\hat{\lambda}_j^{(0)}(\omega) = \hat{\lambda}_j + \left(\frac{\partial}{\partial I}\mathcal{Q}_{1,j,j+m-m_1}(0)\right)\mathcal{C}^{-1}(\omega - \omega^{(0)}) + O_2(\omega - \omega^{(0)}), \tag{55}$$

$j = 1, \ldots, m - m_1$, and that $\hat{\lambda}_{j+m-m_1}^{(0)} = -\hat{\lambda}_j^{(0)}$. With those assumptions and notations, we can formulate the following results.

4.2 Main theorems

Theorem 3 *We consider the real analytic Hamiltonian H of* (3)*, defined on $\mathcal{D}_{r,m}(\rho_0, R_0)$ (for some $0 < \rho_0 < 1$ and $R_0 > 0$), and such that the first m_1 components ($0 \leq m_1 \leq m$) of the vector λ of eigenvalues of $J_m\mathcal{B}$ are of elliptic type. Let $\mathcal{S} \subset \mathcal{N}$ be the set introduced in* (15) *and* (16)*. Then, we also assume that there exists $\mu_0 > 0$ and $\gamma > r + m_1$ such that*

$$|ik^\top\hat{\omega}^{(0)} + l^\top\lambda| \geq \frac{\mu_0}{(|k|_1 + |l_x - l_y|_1)^\gamma} \quad \forall (l,s) \in \mathcal{S}\, \forall k \in \mathbb{Z}^r \text{ with } |l_x - l_y|_1 + |k|_1 \neq 0.$$

This allows to put H in seminormal form with respect to $\mathcal{S}$ up to finite degree. We assume that this seminormal form up to degree 4 is nondegenerate, in the sense that the two nondegeneracy conditions given in (50) *and* (54) *hold. Then, there exist a Cantor subset $\mathcal{A} \subset \mathbb{R}^{r+m_1}$ such that, for any $\omega \in \mathcal{A}$, the Hamiltonian system H has an invariant $(r+m_1)$-dimensional (complex) torus with ω as a vector of basic frequencies, with reducible normal flow. Moreover, if $\mathcal{A}(A) = \mathcal{U}(A) \cap \mathcal{A}$ (being $\mathcal{U}(A)$ the set defined in* (9)*), then,*

$$\mathrm{mes}(\mathcal{U}(A) \setminus \mathcal{A}(A)) \leq const. \exp\left(-const. \left(\frac{1}{A}\right)^{\frac{1}{\gamma+1}}\right),$$

where the constants in this bound are independent from A.

The key to prove Theorem 3 is the parametrization $\mathcal{I}(\omega)$ of the invariant tori of the normal form given by Lemma 3. To construct this function, we take R of order $\sqrt{A}$ in Theorem 1, and we obtain a Hamiltonian in normal form with respect to $\mathcal{S}$ (as the one of (48)), with exponentially small bounds for $\mathcal{R}$ as a function of A, of the same order of the measure of destroyed tori in Theorem 3. Using Lemma 3 on the function $\mathcal{F}$ of (48), we can construct for any frequency A-close to $\omega^{(0)}$ the corresponding action, $I \equiv \mathcal{I}(\omega)$, that gives in (7) the invariant torus of the seminormal form having this concrete vector of intrinsic frequencies. Nevertheless, as the action I can have some of the $\tilde{I}_j < 0$, the corresponding torus in (7) can be complex. This is not an obstruction to construct an invariant (and complex) torus for the complete Hamiltonian (48), but the final torus and its reduced variational normal flow can be in $\mathbb{C}$. If we want to have real tori in (7), we need to take $\omega \in \mathcal{W}(A)$,

$$\mathcal{W}(A) = \left\{ \omega \in \mathcal{U}(A) : \ \omega = \omega^{(0)} + \frac{\partial \mathcal{F}}{\partial I}(I), \ \text{with } \tilde{I}_j \geq 0, \ j = 1, \dots, m_1 \right\}. \tag{56}$$

Note that the degenerate (transition) tori have frequencies whose corresponding actions satisfy $\tilde{I}_j = 0$, for some j. Then, we are forced to remove actions in a tiny slice around the hyperplanes $\tilde{I}_j = 0$, that implies to take out in $\mathcal{W}(A)$ the corresponding frequencies. Unfortunately, $\mathcal{F}$ changes with A as $A \to 0$ by increasing the order up to which this seminormal form is done. This is necessary because the successive approximations to $\mathcal{F}$ given by Theorem 1 do not converge in general, and hence, as we want to eliminate only an exponentially small set of frequencies, we need to know this map with an exponentially small precision.

In this context, we have the following result about the existence of real invariant tori:

Theorem 4 *With the same hypoteses as in Theorem* 3*, there exist a Cantor subset $\mathcal{A} \subset \mathbb{R}^{r+m_1}$ such that, for any $\omega \in \mathcal{A}$, the Hamiltonian system H has an invariant $(r+m_1)$-dimensional real torus with vector of basic frequencies given by ω. The normal flow of this torus can be reduced to constant coefficients by means of a real change of variables.*

$\mathcal{A}$ can be caracterized in the following form: for any $R > 0$ small enough, there exists a convergent (partial) seminormal form with respect to $\mathcal{S}$ (it takes the form (48)*) defined on $\mathcal{D}_{r,m}(\rho_1, R)$ (being ρ_1 independent from R), and such that if we put $\mathcal{A}(R^2) = \mathcal{W}(R^2) \cap \mathcal{A}$ (see* (56) *for the definition of $\mathcal{W}(A)$), then, we have*

$$mes(\mathcal{W}(R^2) \setminus \mathcal{A}(R^2)) \leq const. \exp\left(-const. \left(\frac{1}{R}\right)^{\frac{2}{\gamma+1}} \right),$$

being the constants in this bound independent from R.

Remark 9 *Let us explain how the transition set (in the frequency space) from real to complex tori can be constructed independently from the seminormal form. For any tori of dimension s, $r \leq s < r + m_1$, obtained applying Theorem* 4 *to an invariant subbundle $\mathcal{G}_1 \subset \mathcal{G}$, with $\dim \mathcal{G}_1 < m_1$, we consider the $(r+m_1)$-dimensional vector of frequencies obtained jointly the intrinsic frequencies of the tori, with the the normal frequencies that generalize the ones associated to $\mathcal{G}$ not added as new intrinsic frequencies. We note that those normal frequencies are well defined from the reducible character of the normal variational flow of the constructed tori. Then, as union of those vectors, we obtain a Cantor set that acts as a transition set, with an exponentially small error (as $R \to 0$), to separate the real and complex tori.*

Remark 10 *In the case of maximal dimensional tori, this result can also be compared with* [13] *and* [6]. *We remark that as we work with $m_1 \leq m$, the estimates on the measure of invariant tori in Theorem* 4 *are given in the frequency space and not in the phase space, where the total measure filled by those tori is zero when $m_1 < m$. In the case of maximal dimensional tori, it is not difficult to check (for example, if one uses a formulation of the proof in terms of the actions I instead of the frequencies ω, see Section* 4.3 *for more details) that the measure of the complementary of the set that the real tori fill in $\mathcal{D}_{r,m}(0,R) \cap \mathbb{R}^{2\ell}$ is of order const.* $\exp\left(-const.\left(\frac{1}{R}\right)^{\frac{2}{\gamma+1}}\right)$. *This result coincides with* [13], *if we look for maximal dimensional invariant tori around a given maximal dimensional one (we note that we need to done the same rescaling as in Section* 2.2.1 *in order to compare the results), but it improves the estimates on the measure of invariant tori around an elliptic fixed point of* [6].

In what follows, we will only prove Theorem 4. The proof of Theorem 3 is very similar, and it is done by splitting the set $\mathcal{U}(A)$ as the union of the sets of frequencies defined like (56), but taking into account all the possible combinations of conditions $\tilde{I}_j \geq 0$ and $\tilde{I}_j \leq 0$, $j = 1, \ldots, m_1$.

4.3 Proof of Theorem 4

Before starting with the details, let us mention a technical problem that will appear during the proof. At each step of the iterative KAM process, we will have to deal with the classical small divisors problem. This will lead us to eliminate a dense set of resonant frequencies, to avoid convergence problems. Hence, after the first step of the process the frequency set has empty interior. To bound the measure of the eliminated set of frequencies, we need some kind of regularity. So, we have used Lipschitz conditions because they are well defined on sets with empty interior.[4] The control of this kind of dependence on Cantor sets is analogous to the control of differentiable dependence on open sets.

To work with this kind of dependence, we introduce some notations and definitions. Given $f(\varphi)$ a function defined for $\varphi \in \mathcal{E}$, $\mathcal{E} \subset \mathbb{R}^n$ for some n, and with values in $\mathbb{C}$, $\mathbb{C}^{n_1}$ or $\mathbb{M}_{n_1,n_2}(\mathbb{C})$, we define the Lipschitz constant of f on $\mathcal{E}$ as

$$\mathcal{L}_{\mathcal{E}}\{f\} = \sup_{\substack{\varphi_1,\varphi_2 \in \mathcal{E} \\ \varphi_1 \neq \varphi_2}} \frac{|f(\varphi_2) - f(\varphi_1)|}{|\varphi_2 - \varphi_1|}.$$

If $\mathcal{L}_{\mathcal{E}}\{f\} < +\infty$, we say that f is Lipschitz on $\mathcal{E}$, with respect to the norm $|.|$. We also define $\|f\|_{\mathcal{E}} = \sup_{\varphi \in \mathcal{E}} |f(\varphi)|$. The same definitions can be extended to analytic functions depending on the parameter φ. Hence, if $f(\theta,\varphi)$ or $f(\theta,x,I,y,\varphi)$ are, for every $\varphi \in \mathcal{E}$, analytic, 2π-periodic on θ, and defined on $\{\theta \in \mathbb{C}^r : |\mathrm{Im}\theta| \leq \rho\}$ or $\mathcal{D}_{r,m}(\rho,R)$, respectively, we can introduce $\mathcal{L}_{\mathcal{E},\rho}\{f\}$, $\|f\|_{\mathcal{E},\rho}$, or $\mathcal{L}_{\mathcal{E},\rho,R}\{f\}$ and $\|f\|_{\mathcal{E},\rho,R}$, taking the supremum on the norms $|.|_\rho$ or $|.|_{\rho,R}$, respectively. Some basic results related to those kind of dependence are given in Section 5.

[4] There are alternative methods to solve this problem, for instance the so called "ultraviolet cut" (see [1] for the first example). This is based on dropping the harmonics f_k such that $|k|$ is bigger than some quantity $O(2^n)$ (n is the step). Hence, at each step, we deal with only a finite number of resonances. This allows to work at every step on open sets.

4.3.1 Preliminaries

First, we take a fixed value of R, small enough, and we use Theorem 1 to put the initial Hamiltonian H in (semi)normal form with respect to the set $\mathcal{S}$, except a remainder $\mathcal{R}$ of exponentially small size with respect to R. Hence, we can work with the Hamiltonian H of (48) as if it were the initial one. As the order of this normal form depends on R, we have to write explicitly this R-dependence in all the bounds.

We take the analytic function $\mathcal{F}$ of (48) and we construct the parametrization $\mathcal{I}(\omega)$ provided by Lemma 3, that is well defined if R is small enough. Then, as we want to work with $(r+m_1)$-dimensional tori, we introduce new m_1 angular variables $\tilde{\theta}$ conjugated to the actions $\tilde{I}$ previously added to describe this family of invariant tori. That is, for any frequency $\omega \in \mathbb{R}^{r+m_1}$ close to $\omega^{(0)}$, we replace the real (semi)normal form variables $(\tilde{x}_j, \tilde{y}_j)$, $j = 1, \ldots, m_1$, by the new canonical variables $(\tilde{\theta}_j, \tilde{I}_j)$ defined as

$$\tilde{x}_j = \sqrt{2(\tilde{I}_j + \tilde{\mathcal{I}}_j(\omega))}\sin(\tilde{\theta}_j), \quad \tilde{y}_j = \sqrt{2(\tilde{I}_j + \tilde{\mathcal{I}}_j(\omega))}\cos(\tilde{\theta}_j), \tag{57}$$

or, in terms of the complexified variables $(\tilde{X}_j, \tilde{Y}_j)$ (see (44)),

$$\tilde{X}_j = -i\sqrt{\tilde{I}_j + \tilde{\mathcal{I}}_j(\omega)}\exp(i\tilde{\theta}_j), \quad \tilde{Y}_j = \sqrt{\tilde{I}_j + \tilde{\mathcal{I}}_j(\omega)}\exp(-i\tilde{\theta}_j). \tag{58}$$

We remark that those $\tilde{I}$ are not exactly the same ones used to parametrize the tori (7), because they differ in a translation by $\tilde{\mathcal{I}}(\omega)$. This is done to put the ω-invariant torus, with respect to the seminormal form, in $\tilde{I} = 0$. Hence, we extend this translation to the whole set of actions I, doing the transformation

$$\hat{I}_j \to \hat{I}_j + \hat{\mathcal{I}}_j(\omega), \; j = 1, \ldots, r. \tag{59}$$

Moreover, we denote by $\theta^\top = (\hat{\theta}^\top, \tilde{\theta}^\top)$ the vector of all the angular variables. Then, we have constructed, for each ω, a new canonical system of coordinates (with $r+m_1$ angular variables) that put the corresponding seminormal form torus in $I = 0$. If we insert those new variables in the Hamiltonian (48), we obtain a ω-depending family of Hamiltonians $H^{(0)}_\omega$, that we simply denote by $H^{(0)} \equiv H^{(0)}(\theta, \hat{X}, I, \hat{Y}, \omega)$,

$$H^{(0)} = \omega^{(0)\top} I + \frac{1}{2}\hat{Z}^\top \hat{\mathcal{B}}^* \hat{Z} + \mathcal{F}(I + \mathcal{I}(\omega)) + \frac{1}{2}\hat{Z}^\top \mathcal{Q}(I + \mathcal{I}(\omega))\hat{Z} + \mathcal{T}^* + \mathcal{R}^*,$$

where $\mathcal{T}^*$ and $\mathcal{R}^*$ are $\mathcal{T}$ and $\mathcal{R}$ expressed in terms of $\tilde{\theta}$ and $\tilde{I}$, and composed with the translation (59). We cast $H^{(0)}$ into the following form:

$$H^{(0)} = \phi^{(0)}(\omega) + \omega^\top I + \frac{1}{2}\hat{Z}^\top \hat{\mathcal{B}}^{(0)*}(\omega)\hat{Z} + \frac{1}{2}I^\top \mathcal{C}^{(0)}(\omega) I + H^{(0)}_* + \hat{H}^{(0)}, \tag{60}$$

where $\phi^{(0)} = \mathcal{F}(\mathcal{I}(\omega))$, $\hat{\mathcal{B}}^{(0)*} = \hat{\mathcal{B}}^* + \mathcal{Q}(\mathcal{I}(\omega))$, $\mathcal{C}^{(0)} = \frac{\partial^2 \mathcal{F}}{\partial I^2}(\mathcal{I}(\omega))$, $\hat{H}^{(0)} = \mathcal{R}^*$ and

$$\begin{aligned} H^{(0)}_* &= \mathcal{F}(I + \mathcal{I}(\omega)) - \mathcal{F}(\mathcal{I}(\omega)) - \frac{\partial \mathcal{F}}{\partial I}(\mathcal{I}(\omega)) I - \frac{1}{2} I^\top \frac{\partial^2 \mathcal{F}}{\partial I^2}(\mathcal{I}(\omega)) I + \\ &\quad + \frac{1}{2}\hat{Z}^\top (\mathcal{Q}(I + \mathcal{I}(\omega)) - \mathcal{Q}(\mathcal{I}(\omega)))\hat{Z} + \mathcal{T}^*, \end{aligned}$$

where we have used the properties of $\mathcal{I}(\omega)$ (see Lemma 3). We remark that, if one considers $H^{(0)}_\omega$ for a fixed ω, and one skips the term $\hat{H}^{(0)}$ in (60), then, $I = 0$ and $\hat{Z} = 0$

correspond to an invariant $(r+m_1)$-dimensional torus with vector of basic frequencies ω, with reducible normal variational flow given by the (complex) diagonal matrix $J_{m-m_1}\hat{\mathcal{B}}^{(0)*}$. Moreover, in this case the variables I and $\hat{Z}$ are uncoupled, at least up to first order.

Nevertheless, the coordinates of (58) become singular for $I=0$ when we take frequencies ω with some $\tilde{I}_j \equiv \tilde{I}_j(\omega)=0$. Thus, we have to eliminate a neighbourhood of the set of those critical frequencies to ensure that the change is well defined.

Let $M(R)R^8$ be the expression (with M exponentially small in R) bounding $|\mathcal{R}|_{\rho_1,R}$ in (49). We consider a fixed value of R, small enough, such that $M(R)<1$, and we take a fixed number $0<\alpha<1$, to be precised later. Then, as we are only interested in real tori, we use definition (56) to introduce the following set:

$$\mathcal{E}^{(0)}(R) = \mathcal{W}\left(\frac{1}{8}(|\mathcal{C}^{-1}|)^{-1}R^2\right) \setminus \left(\mathcal{I}^{-1}\left(\mathcal{V}\left(16(M(R))^{2\alpha}\right)\right)\right),$$

being $\mathcal{V}(A) = \{I \in \mathbb{R}^{r+m_1} : |\tilde{I}_j| \le A, \text{ for any } j=1,\ldots,m_1\}$. That is, we eliminate from $\mathcal{W}(\frac{1}{8}(|\mathcal{C}^{-1}|)^{-1}R^2)$ the frequencies corresponding to actions $\tilde{I}$ close to any of the hyperplanes $\tilde{I}_j=0$. As from Lemma 3 we have $|\mathcal{I}(\omega)| \le \frac{1}{4}R^2$ for any $\omega \in \mathcal{W}\left(\frac{1}{8}(|\mathcal{C}^{-1}|)^{-1}R^2\right)$, the measure of $\tilde{\mathcal{V}}(R) := \mathcal{V}\left(16M^{2\alpha}\right) \cap \mathcal{I}\left(\mathcal{W}\left(\frac{1}{8}(|\mathcal{C}^{-1}|)^{-1}R^2\right)\right)$ is of order $(M(R))^{2\alpha}$. Then, to control the measure that this set fills in $\mathcal{W}(\frac{1}{8}(|\mathcal{C}^{-1}|)^{-1}R^2)$, we put $\widetilde{\mathcal{W}}(R) = \mathcal{I}^{-1}\left(\mathcal{V}\left(16M^{2\alpha}\right)\right) \cap \mathcal{W}(\frac{1}{8}(|\mathcal{C}^{-1}|)^{-1}R^2)$, and then we have

$$\operatorname{mes}(\widetilde{\mathcal{W}}(R)) = \int_{\widetilde{\mathcal{W}}(R)} d\omega = \int_{\tilde{\mathcal{V}}(R)} \left|\det D\mathcal{I}^{-1}(I)\right| dI, \tag{61}$$

where, from the definition of $\mathcal{I}$ in Lemma 3, $\mathcal{I}^{-1}(I) = \omega^{(0)} + \frac{\partial \mathcal{F}}{\partial I}(I)$. Then, the bounds on $\mathcal{F}$ in (49) suffices to justify that $\operatorname{mes}(\widetilde{\mathcal{W}}(R))$ is also of order $(M(R))^{2\alpha}$.

Now, we are going to see that $H^{(0)}$ is defined on a small neighbourhood of $I=0$ and $\hat{Z}=0$, with positive (and bounded from below) distance to the critical set of frequencies. To do that, and as we will work with functions of $(\theta,\hat{X},I,\hat{Y})$, we will take the different norms on domains of the form $\mathcal{D}_{r+m_1,m-m_1}(.,.)$. More concretely, we will see that, for any $\omega \in \mathcal{E}^{(0)}(R)$, $H^{(0)}$ is well defined on $\mathcal{D}_{r+m_1,m-m_1}(\rho^{(0)},R^{(0)})$, being $\rho^{(0)}=\rho_0/2$ (this is smaller than the width of analyticity for the $\hat{\theta}$ variables given by Theorem 1 for the seminormal form) and $R^{(0)} = 2M^{\alpha}$. Let us check that.

First, we remark that as $\mathcal{F}$ only depends on I, we deduce from (49) that $|\mathcal{F}|_{0,R} \le \hat{\mathcal{F}}R^4$. Then, assuming $\frac{1}{2}R^2 + 4M^{2\alpha} \le R^2$, and using Lemmas 7 and 8, we obtain

$$\|\mathcal{F}(\mathcal{I}(\omega)+I)\|_{\mathcal{E}^{(0)},0,R^{(0)}} \le \hat{\mathcal{F}}R^4, \quad \mathcal{L}_{\mathcal{E}^{(0)},0,R^{(0)}}\{\mathcal{F}(\mathcal{I}(\omega)+I)\} \le (r+m_1)\frac{\hat{\mathcal{F}}R^4}{\frac{1}{4}R^2}2|\mathcal{C}^{-1}|, \tag{62}$$

where we have used that $\mathcal{L}_{\mathcal{E}^{(0)}}\{\mathcal{I}\} \le 2|\mathcal{C}^{-1}|$ (see Lemma 3). Similar bounds can be derived for $\mathcal{F}_3$, $\mathcal{Q}$ and $\mathcal{Q}_1$.

To bound $\mathcal{T}^*$ and $\mathcal{R}^*$, we need to study the well defined character of the transformation (58). Using Lemma 10 (we recall that, from the definition of $\mathcal{E}^{(0)}$, one has $\tilde{\mathcal{I}}_j(\omega) \ge 16M^{2\alpha}$ for any $\omega \in \mathcal{E}^{(0)}$), and using that $\rho_0 \le 1$, it is not difficult to check that if we consider $(\tilde{X}_j,\tilde{Y}_j)$ in (58) as a function of $(\theta,\hat{X},I,\hat{Y})$ and ω, we have for $j=1,\ldots,m_1$:

$$\max\left\{\|\hat{X}_j\|_{\mathcal{E}^{(0)},\rho^{(0)},R^{(0)}}, \|\hat{Y}_j\|_{\mathcal{E}^{(0)},\rho^{(0)},R^{(0)}}\right\} \le \frac{1}{2}R\left(2-\sqrt{3/4}\right)\exp(1/2) \le \frac{15}{16}R,$$

and

$$\max\left\{\mathcal{L}_{\mathcal{E}^{(0)},\rho^{(0)},R^{(0)}}\{\hat{X}_j\},\mathcal{L}_{\mathcal{E}^{(0)},\rho^{(0)},R^{(0)}}\{\hat{Y}_j\}\right\}\leq\frac{2|\mathcal{C}^{-1}|\exp{(1/2)}}{8M^{\alpha}\sqrt{3/4}}\leq\frac{|\mathcal{C}^{-1}|}{2M^{\alpha}}.$$

Assuming R small enough such that $4M^{2\alpha}+\frac{1}{4}R^2\leq\left(\frac{15}{16}\right)^2R^2$ (this is used to control the transformation (59)) one has, using Lemmas 7 and 8 on the bounds (49), that

$$\|\mathcal{T}^*\|_{\mathcal{E}^{(0)},\rho^{(0)},R^{(0)}}\leq\hat{\mathcal{T}}R^3,\quad\|\mathcal{R}^*\|_{\mathcal{E}^{(0)},\rho^{(0)},R^{(0)}}\leq MR^8,\tag{63}$$

and

$$\mathcal{L}_{\mathcal{E}^{(0)},\rho^{(0)},R^{(0)}}\{\mathcal{T}^*\}\leq r\frac{\hat{\mathcal{T}}R^3}{\left(1-\left(\frac{15}{16}\right)^2\right)R^2}2|\mathcal{C}^{-1}|+2m_1\frac{\hat{\mathcal{T}}R^2 2M^{\alpha}}{\frac{1}{16}R}\frac{|\mathcal{C}^{-1}|}{2M^{\alpha}},\tag{64}$$

$$\mathcal{L}_{\mathcal{E}^{(0)},\rho^{(0)},R^{(0)}}\{\mathcal{R}^*\}\leq r\frac{MR^8}{\left(1-\left(\frac{15}{16}\right)^2\right)R^2}2|\mathcal{C}^{-1}|+2m_1\frac{MR^8}{\frac{1}{16}R}\frac{|\mathcal{C}^{-1}|}{2M^{\alpha}},\tag{65}$$

where, to bound $\mathcal{T}^*$, we have used that it is $O_3(\hat{Z})$.

Now, we can bound the different terms of $H^{(0)}$ in (60). First, one has $\|\phi^{(0)}\|_{\mathcal{E}^{(0)}}\leq\hat{\mathcal{F}}R^4$. We do not care about its Lipschitz constant because this term can be eliminated without changing the canonical equations (so, we only need to worry about its bounded character). To bound $\mathcal{C}^{(0)}$ we remark that $\mathcal{C}^{(0)}=\mathcal{C}+\frac{\partial^2\mathcal{F}_3}{\partial I^2}(\mathcal{I})$ with $\det\mathcal{C}\neq 0$. Using the bounds on $\mathcal{F}_3$ (that can be obtained in a similar form as the ones on $\mathcal{F}$ in (62)) one has that $\det\mathcal{C}^{(0)}\neq 0$ on $\mathcal{E}^{(0)}$, if R is small enough. In quantitative form, it means that $\|(\bar{\mathcal{C}}^{(0)})^{-1}\|_{\mathcal{E}^{(0)}}\leq\bar{m}^{(0)}$, for certain R-independent constant $\bar{m}^{(0)}$. Moreover, we also have $\|\mathcal{C}^{(0)}\|_{\mathcal{E}^{(0)},\rho^{(0)}}\leq\hat{m}^{(0)}$, $\mathcal{L}_{\mathcal{E}^{(0)},\rho^{(0)}}\{\mathcal{C}^{(0)}\}\leq\tilde{m}^{(0)}$. We have used the norm $\|.\|_{\mathcal{E}^{(0)},\rho^{(0)}}$ for the constant matrix $\mathcal{C}^{(0)}$ because, in the succesive steps of the inductive procedure, the matrices replacing $\mathcal{C}^{(0)}$ will depend on θ. To control the normal eigenvalues, we remark that from the expression (55), we can write $\hat{\lambda}_j^{(0)}$, $j=1,\ldots,2(m-m_1)$, in the following form:

$$\hat{\lambda}_j^{(0)}(\omega)=\hat{\lambda}_j+iv_j^{\top}(\omega-\omega^{(0)})+\check{\lambda}_j^{(0)}(\omega),\tag{66}$$

where $v_j\in\mathbb{C}^{r+m_1}$ and $\check{\lambda}_j^{(0)}\equiv O_2(\omega-\omega^{(0)})$. From the non-degeneracy hypotesis of (54), we have that $\mathrm{Re}(v_j)\notin\mathbb{Z}^{r+m_1}$ and, if we put $v_{j,l}=v_j-v_l$, then $\mathrm{Re}(v_{j,l})\notin\mathbb{Z}^{r+m_1}$, for $j\neq l$. Moreover, using the expression (55) and the fact that the eigenvalues of $J_{m-m_1}\hat{\mathcal{B}}^*$ are all different, it is also easy to check the existence of R-independent constants $0<\alpha_1^{(0)}<\alpha_2^{(0)}$, $\beta_1^{(0)}>0$, such that $0<\alpha_1^{(0)}\leq|\hat{\lambda}_j^{(0)}(\omega)-\hat{\lambda}_l^{(0)}(\omega)|$, $\alpha_1^{(0)}/2\leq|\hat{\lambda}_j^{(0)}(\omega)|\leq\alpha_2^{(0)}/2$, for any $\omega\in\mathcal{E}^{(0)}$, $j\neq l$, and $\mathcal{L}_{\mathcal{E}^{(0)}}\{\hat{\lambda}_j^{(0)}\}\leq\beta_1^{(0)}$. We do not give here explicit bounds on the $\check{\lambda}_j^{(0)}$, as those functions do not appear in the iterative process, but we remark that one has that $\mathcal{L}_{\mathcal{E}^{(0)}}\{\check{\lambda}_j^{(0)}\}$ is of order R. This will be used in Section 4.3.4. Moreover, if one uses bounds like (62) for $\mathcal{F}$ and $\mathcal{Q}$, and the ones of (63) and (64) for $\mathcal{T}^*$, it is not difficult to check that for certain positive R-independent constants $\hat{\nu}^{(0)}$ and $\tilde{\nu}^{(0)}$, one has $\|H_*^{(0)}\|_{\mathcal{E}^{(0)},\rho^{(0)},R^{(0)}}\leq\hat{\nu}^{(0)}$ and $\mathcal{L}_{\mathcal{E}^{(0)},\rho^{(0)},R^{(0)}}\{H_*^{(0)}\}\leq\tilde{\nu}^{(0)}$. Finally, using the bounds of (63) and (65) for $\mathcal{R}^*$, one can bound the size of the perturbative term $\hat{H}^{(0)}$ by $\|\hat{H}^{(0)}\|_{\mathcal{E}^{(0)},\rho^{(0)},R^{(0)}}\leq M$ and $\mathcal{L}_{\mathcal{E}^{(0)},\rho^{(0)},R^{(0)}}\{\hat{H}^{(0)}\}\leq M^{1-\alpha}$. Some of these bounds are far from optimal, but they suffice for our purposes.

4.3.2 The iterative scheme

Now, we can describe the iterative procedure used to construct invariant $(r + m_1)$-dimensional tori. This process is given by a sequence of canonical changes of variables, constructed as the time one flow of a suitable generating function S_ω. The changes are constructed to kill the terms that obstructs the existence of an invariant reduced torus with vector of basic frequencies given by ω. As usual (to overcome the effect of the small divisors), the changes are chosen to produce a quadratically convergent scheme, instead of the linear one of Lemma 1.

First, we describe a generic step of this iterative process. For this purpose, we expand the Hamiltonian $H^{(0)}$ in the following form

$$H^{(0)} = a(\theta)+b(\theta)^\top \hat{Z}+c(\theta)^\top I+\frac{1}{2}\hat{Z}^\top B(\theta)\hat{Z}+I^\top E(\theta)\hat{Z}+\frac{1}{2}I^\top C(\theta)I+\Omega(\theta,\hat{X},I,\hat{Y}), \quad (67)$$

where we do not write explicitly the ω-dependence and where we have skipped the superscript "(0)" in the different parts of the Hamiltonian. From this expansion, we introduce the following notations: $[H^{(0)}]_{(\hat{Z},\hat{Z})} = B$, $[H^{(0)}]_{(I,\hat{Z})} = E$ and $< H^{(0)} >= H^{(0)} - \Omega$. From the bounds on the terms of the decomposition (60), we have that $\tilde{a}$, b, $c-\omega$, $B-\hat{\mathcal{B}}^{(0)*}$, $C-\mathcal{C}^{(0)}$ and E are all $O(\hat{H}^{(0)})$. Note that if we are able to kill the terms $\tilde{a}$, b and $c-\omega$, we will obtain an invariant torus with intrinsic frequency ω. Nevertheless, as we want to have simple equations at every step of the iterative scheme (this is, linear equations with constant coefficients), we are forced to kill something more. Then, we ask the final torus to have reducible normal flow given by a diagonal matrix. This is, we want that the new matrix B verifies $B = \mathcal{J}_{m-m_1}(B)$ where, for a $(2s)$-dimensional matrix $A(\theta)$ depending 2π-periodically on θ, we define $\mathcal{J}_s(A) = -J_s\mathrm{dp}(J_s\bar{A})$. Here, $\mathrm{dp}(A)$ denotes the diagonal matrix obtained taking the diagonal entries of A. Moreover, we have to eliminate E to uncouple the "neutral" and the normal directions of the torus up to first order. Thus, for each step of the iterative process, we use a canonical change of variables, given by a generating function of the form

$$S(\theta,\hat{X},I,\hat{Y}) = \xi^\top\theta + d(\theta) + e(\theta)^\top\hat{Z} + f(\theta)^\top I + \frac{1}{2}\hat{Z}^\top G(\theta)\hat{Z} + I^\top F(\theta)\hat{Z},$$

where $\xi \in \mathbb{C}^{r+m_1}$, $\bar{d} = 0$, $\bar{f} = 0$ and G is a symmetric matrix, with $\mathcal{J}_{m-m_1}(G) = 0$. The transformed Hamiltonian is $H^{(1)} = H^{(0)} \circ \Psi_1^S$. We expand $H^{(1)}$ in the same way as $H^{(0)}$ in (67), keeping the same name for the new variables, but adding the superscript "(1)" to a, b, c, B, C, E and Ω. Then, we ask $\tilde{a}^{(1)} = 0$, $b^{(1)} = 0$, $c^{(1)} - \omega = 0$, $E^{(1)} = 0$ and $B^{(1)} = \mathcal{J}_{m-m_1}(B^{(1)})$. We will show that this can be achieved up to first order in the size of $\hat{H}^{(0)}$. For this purpose, we write those conditions in terms of the initial Hamiltonian and the generating function, and then, we obtain the following equations:

(eq_1) $\tilde{a} - \frac{\partial d}{\partial\theta}\omega = 0,$

(eq_2) $b - \frac{\partial e}{\partial\theta}\omega + \hat{\mathcal{B}}^{(0)*}J_{m-m_1}e = 0,$

(eq_3) $c - \omega - \frac{\partial f}{\partial\theta}\omega - \mathcal{C}^{(0)}\left(\xi + \left(\frac{\partial d}{\partial\theta}\right)^\top\right) = 0,$

(eq_4) $B^* - \mathcal{J}_{m-m_1}(B^*) - \frac{\partial G}{\partial\theta}\omega + \hat{\mathcal{B}}^{(0)*}J_{m-m_1}G - GJ_{m-m_1}\hat{\mathcal{B}}^{(0)*} = 0,$

(eq_5) $E^* - \frac{\partial F}{\partial \theta}\omega - FJ_{m-m_1}\hat{\mathcal{B}}^{(0)*} = 0,$

being

$$\begin{aligned} B^* &= B - \left[\frac{\partial H_*^{(0)}}{\partial I}\left(\xi + \left(\frac{\partial d}{\partial \theta}\right)^\top\right) - \frac{\partial H_*^{(0)}}{\partial \hat{Z}} J_{m-m_1} e\right]_{(\hat{Z},\hat{Z})}, \\ E^* &= E - \mathcal{C}^{(0)}\left(\frac{\partial e}{\partial \theta}\right)^\top - \left[\frac{\partial H_*^{(0)}}{\partial I}\left(\xi + \left(\frac{\partial d}{\partial \theta}\right)^\top\right) - \frac{\partial H_*^{(0)}}{\partial \hat{Z}} J_{m-m_1} e\right]_{(I,\hat{Z})}. \end{aligned}$$

To solve those homological equations, we expand them in Fourier series and we equate the corresponding coefficients, obtaining the formal solutions. The next step is to derive bounds on those solutions. As we will use these bounds in iterative form, we want to make clear which expressions change from one step to another, and which ones can be bounded independently from the step. For this purpose, we take fixed positive constants $\bar{m}$, $\hat{m}$, $\tilde{m}$, α_2, β_1, $\hat{\nu}$, $\tilde{\nu}$ defined as twice the corresponding inital values $\bar{m}^{(0)}$, $\hat{m}^{(0)}$, $\tilde{m}^{(0)}$, $\alpha_2^{(0)}$, $\beta_1^{(0)}$, $\hat{\nu}^{(0)}$, $\tilde{\nu}^{(0)}$ and a fixed α_1, $0 < \alpha_1 < \alpha_1^{(0)}$. In what follows, $\hat{N}$ will denote an expression depending only on $\bar{m}$, $\hat{m}$, α_1, α_2, $\hat{\nu}$, the different dimensions r, m, m_1, plus γ and ρ_0. $\hat{N}$ will be redefined during the description of the iterative scheme to meet a finite number of conditions. The idea is to perform the bounds on the iterative scheme putting the superscript "(0)" on the terms that change at every iteration. Hence, we write the bounds on $\hat{H}^{(0)}$ as $\|\hat{H}^{(0)}\|_{\mathcal{E}^{(0)},\rho^{(0)},R^{(0)}} \leq M^{(0)}$ and $\mathcal{L}_{\mathcal{E}^{(0)},\rho^{(0)},R^{(0)}}\{\hat{H}^{(0)}\} \leq L^{(0)}$, with $M^{(0)}(R) \equiv M(R)$ and $L^{(0)}(R) \equiv (M(R))^{1-\alpha}$. Hence, using Lemma 5,

$$\begin{aligned} \|a - \phi^{(0)}\|_{\mathcal{E}^{(0)},\rho^{(0)}} &\leq M^{(0)}, & \|E\|_{\mathcal{E}^{(0)},\rho^{(0)}} &\leq \tfrac{2(m-m_1)M^{(0)}}{(R^{(0)})^3}, \\ \|c - \omega\|_{\mathcal{E}^{(0)},\rho^{(0)}} &\leq \tfrac{M^{(0)}}{(R^{(0)})^2}, & \|B - \hat{\mathcal{B}}^{(0)*}\|_{\mathcal{E}^{(0)},\rho^{(0)}} &\leq \tfrac{(2(m-m_1)+1)M^{(0)}}{(R^{(0)})^2}, \\ \|b\|_{\mathcal{E}^{(0)},\rho^{(0)}} &\leq \tfrac{M^{(0)}}{R^{(0)}}, & \|C - \mathcal{C}^{(0)}\|_{\mathcal{E}^{(0)},\rho^{(0)}} &\leq \tfrac{(2(r+m_1)+1)M^{(0)}}{(R^{(0)})^4}, \\ \|\Omega\|_{\mathcal{E}^{(0)},\rho^{(0)},R^{(0)}} &\leq \hat{\nu}^{(0)} + M^{(0)}. \end{aligned} \tag{68}$$

Moreover, we can use Lemma 11 to deduce that the same bounds hold for their Lipschitz constants on $\mathcal{E}^{(0)}$, replacing $M^{(0)}$ by $L^{(0)}$, and $\hat{\nu}^{(0)}$ by $\tilde{\nu}^{(0)}$. Then, to prove the convergence of the expansion of S, we need some kind of control on the different small divisors involved. For this purpose, we restrict the parameter ω to the subset $\mathcal{E}^{(1)}(R) \subset \mathcal{E}^{(0)}(R)$ for which the following Diophantine estimates hold: we say that $\omega \in \mathcal{E}^{(1)}$, if $\omega \in \mathcal{E}^{(0)}$, and

$$|ik^\top\omega + l^\top\hat{\lambda}^{(0)}(\omega)| \geq \frac{\mu^{(0)}(R)}{|k|_1^\gamma}, \quad k \in \mathbb{Z}^{r+m_1} \setminus \{0\}, \quad l \in \mathbb{N}^{2(m-m_1)}, \quad 0 < |l|_1 \leq 2, \tag{69}$$

for certain $\mu^{(0)} > 0$. We expect the measure of $\mathcal{E}^{(0)} \setminus \mathcal{E}^{(1)}$ to be of order $\mu^{(0)}$ and, hence, as we want to have exponentially small bounds for this measure, we take $\mu^{(0)} \equiv (M^{(0)})^\alpha$. Then, we proceed to bound the solutions of the different homological equations. For this purpose, we use Lemma 4. More precisely, we define $\delta^{(0)} = (M^{(0)})^\alpha$, and we take $\delta^{(0)}$ as a value for δ to use the different estimates provided by this lemma. In order to simplify the proofs, we assume $\rho^{(0)} - N\delta^{(0)} \geq \rho_0/4$, where $N \in \mathbb{N}$ will be a fixed integer that will be determined before the description of the iterative scheme. Moreover, we also assume that $(M^{(0)})^\alpha \leq R^{(0)} \leq 1$. Then, one can solve $(eq_1) - (eq_5)$ as follows:

(eq_1) For d, we have

$$d(\theta) = \sum_{k\in\mathbb{Z}^{r+m_1}\setminus\{0\}} \frac{a_k}{ik^\top\omega} \exp(ik^\top\theta),$$

that implies,

$$\|d\|_{\mathcal{E}^{(1)},\rho^{(0)}-\delta^{(0)}} \le \left(\frac{\gamma}{\delta^{(0)}\exp{(1)}}\right)^\gamma \frac{\|\tilde{a}\|_{\mathcal{E}^{(1)},\rho^{(0)}}}{\mu^{(0)}} \le \hat{N}(M^{(0)})^{1-\alpha-\alpha\gamma}.$$

(eq_2) For any j, $1\le j\le 2(m-m_1)$, we have

$$e_j(\theta) = \sum_{k\in\mathbb{Z}^{r+m_1}} \frac{b_{j,k}}{ik^\top\omega+\hat{\lambda}_j^{(0)}} \exp(ik^\top\theta),$$

and hence,

$$\|e\|_{\mathcal{E}^{(1)},\rho^{(0)}-\delta^{(0)}} \le \left(\frac{2}{\alpha_1}+\left(\frac{\gamma}{\delta^{(0)}\exp{(1)}}\right)^\gamma \frac{1}{\mu^{(0)}}\right)\|b\|_{\mathcal{E}^{(1)},\rho^{(0)}} \le \hat{N}(M^{(0)})^{1-2\alpha-\alpha\gamma}.$$

(eq_3) Taking average with respect to θ, we obtain

$$\xi = (\bar{\mathcal{C}}^{(0)})^{-1}\left(\bar{c}-\omega-\overline{\mathcal{C}^{(0)}\left(\frac{\partial d}{\partial\theta}\right)^\top}\right).$$

Thus,

$$\begin{aligned}
\|\xi\|_{\mathcal{E}^{(1)}} &= \|(\bar{\mathcal{C}}^{(0)})^{-1}\bar{\mathcal{C}}^{(0)}\xi\|_{\mathcal{E}^{(1)}} \le \|(\bar{\mathcal{C}}^{(0)})^{-1}\|_{\mathcal{E}^{(1)}}\|\bar{\mathcal{C}}^{(0)}\xi\|_{\mathcal{E}^{(1)}} \le \\
&\le \bar{m}\left(\|\bar{c}-\omega\|_{\mathcal{E}^{(1)},0}+\left\|\mathcal{C}^{(0)}\left(\frac{\partial d}{\partial\theta}\right)^\top\right\|_{\mathcal{E}^{(1)},0}\right) \le \\
&\le \bar{m}\left(\|c-\omega\|_{\mathcal{E}^{(1)},\rho^{(0)}}+\hat{m}\frac{\|d\|_{\mathcal{E}^{(1)},\rho^{(0)}-\delta^{(0)}}}{(\rho^{(0)}-\delta^{(0)})\exp{(1)}}\right) \le \hat{N}(M^{(0)})^{1-\alpha-\alpha\gamma}.
\end{aligned}$$

To solve the equation for f, we define

$$c^* = \tilde{c}-\tilde{\mathcal{C}}^{(0)}\xi-\mathcal{C}^{(0)}\left(\frac{\partial d}{\partial\theta}\right)^\top+\overline{\mathcal{C}^{(0)}\left(\frac{\partial d}{\partial\theta}\right)^\top},$$

and then, for any $1\le j\le r+m_1$, we have

$$f_j(\theta) = \sum_{k\in\mathbb{Z}^{r+m_1}\setminus\{0\}} \frac{c^*_{j,k}}{ik^\top\omega} \exp(ik^\top\theta).$$

To bound f, first we have that

$$\begin{aligned}
\|c^*\|_{\mathcal{E}^{(1)},\rho^{(0)}-2\delta^{(0)}} &\le \|\tilde{c}\|_{\mathcal{E}^{(1)},\rho^{(0)}}+\|\mathcal{C}^{(0)}\|_{\mathcal{E}^{(1)},\rho^{(0)}}\left(\|\xi\|_{\mathcal{E}^{(1)}}+\frac{\|d\|_{\mathcal{E}^{(1)},\rho^{(0)}-\delta^{(0)}}}{\delta^{(0)}\exp{(1)}}\right) \le \\
&\le \hat{N}(M^{(0)})^{1-2\alpha-\alpha\gamma},
\end{aligned}$$

and from here

$$\|f\|_{\mathcal{E}^{(1)},\rho^{(0)}-3\delta^{(0)}} \le \left(\frac{\gamma}{\delta^{(0)}\exp{(1)}}\right)^\gamma \frac{\|c^*\|_{\mathcal{E}^{(1)},\rho^{(0)}-2\delta^{(0)}}}{\mu^{(0)}} \le \hat{N}(M^{(0)})^{1-3\alpha-2\alpha\gamma}.$$

(eq_4) We define $B^{**} = B^* - \mathcal{J}_{m-m_1}(B^*)$, and then, if $G = (G_{j,l})$, $1 \le j,l \le 2(m-m_1)$, we have

$$G_{j,l}(\theta) = \sum_{k\in\mathbb{Z}^{r+m_1}} \frac{B^{**}_{j,l,k}}{ik^\top\omega + \hat{\lambda}^{(0)}_j + \hat{\lambda}^{(0)}_l} \exp(ik^\top\theta).$$

In this sum we have to avoid the indices (j,l,k) for which $|j-l| = m-m_1$ and $k=0$. In these cases we have trivial zero divisors, but also the coefficient $B^{**}_{j,l,0}$ is 0. Moreover, we remark that the matrix G is symmetric. Then, to bound G, we have to bound B^{**}. First, we have

$$\begin{aligned} \|B^* - \hat{\mathcal{B}}^{(0)*}\|_{\mathcal{E}^{(1)},\rho^{(0)}-2\delta^{(0)}} &\le \|B - \hat{\mathcal{B}}^{(0)*}\|_{\mathcal{E}^{(1)},\rho^{(0)}-2\delta^{(0)}} + \\ &+(2(m-m_1)+1)(r+m_1)\frac{\|H^{(0)}_*\|_{\mathcal{E}^{(1)},\rho^{(0)},R^{(0)}}}{(R^{(0)})^4}\left(\|\xi\|_{\mathcal{E}^{(1)}} + \frac{\|d\|_{\mathcal{E}^{(1)},\rho^{(0)}-\delta^{(0)}}}{\delta^{(0)}\exp(1)}\right) + \\ &+24(m-m_1)^2\frac{\|H^{(0)}_*\|_{\mathcal{E}^{(1)},\rho^{(0)},R^{(0)}}}{(R^{(0)})^3}\|e\|_{\mathcal{E}^{(1)},\rho^{(0)}-\delta^{(0)}} \le \hat{N}(M^{(0)})^{1-6\alpha-\alpha\gamma}, \end{aligned}$$

and from the definition of B^{**} and the norm used, the same bound holds for B^{**}. Then,

$$\begin{aligned} \|G\|_{\mathcal{E}^{(1)},\rho^{(0)}-3\delta^{(0)}} &\le \left(\frac{1}{\alpha_1} + \left(\frac{\gamma}{\delta^{(0)}\exp(1)}\right)^\gamma \frac{1}{\mu^{(0)}}\right) 2(m-m_1)\,\|B^{**}\|_{\mathcal{E}^{(1)},\rho^{(0)}-2\delta^{(0)}} \le \\ &\le \hat{N}(M^{(0)})^{1-7\alpha-2\alpha\gamma}. \end{aligned}$$

(eq_5) The different components of F are given by

$$F_{j,l}(\theta) = \sum_{k\in\mathbb{Z}^{r+m_1}} \frac{E^*_{j,l,k}}{ik^\top\omega + \hat{\lambda}^{(0)}_l} \exp(ik^\top\theta),$$

for $j = 1,\dots,r+m_1$ and $l = 1,\dots,2(m-m_1)$. Thus,

$$\begin{aligned} \|E^*\|_{\mathcal{E}^{(1)},\rho^{(0)}-2\delta^{(0)}} &\le \|E\|_{\mathcal{E}^{(1)},\rho^{(0)}} + 2(m-m_1)\|\mathcal{C}^{(0)}\|_{\mathcal{E}^{(1)},\rho^{(0)}}\frac{\|e\|_{\mathcal{E}^{(1)},\rho^{(0)}-\delta^{(0)}}}{\delta^{(0)}\exp(1)} + \\ &+4(m-m_1)(r+m_1)\frac{\|H^{(0)}_*\|_{\mathcal{E}^{(1)},\rho^{(0)},R^{(0)}}}{(R^{(0)})^5}\left(\|\xi\|_{\mathcal{E}^{(1)}} + \frac{\|d\|_{\mathcal{E}^{(1)},\rho^{(0)}-\delta^{(0)}}}{\delta^{(0)}\exp(1)}\right) + \\ &+8(m-m_1)^2\frac{\|H^{(0)}_*\|_{\mathcal{E}^{(1)},\rho^{(0)},R^{(0)}}}{(R^{(0)})^4}\|e\|_{\mathcal{E}^{(1)},\rho^{(0)}-\delta^{(0)}} \le \hat{N}(M^{(0)})^{1-7\alpha-\alpha\gamma}, \end{aligned}$$

and, hence,

$$\begin{aligned} \|F\|_{\mathcal{E}^{(1)},\rho^{(0)}-3\delta^{(0)}} &\le \left(\frac{2}{\alpha_1} + \left(\frac{\gamma}{\delta^{(0)}\exp(1)}\right)^\gamma \frac{1}{\mu^{(0)}}\right) 2(m-m_1)\,\|E^*\|_{\mathcal{E}^{(1)},\rho^{(0)}-2\delta^{(0)}} \le \\ &\le \hat{N}(M^{(0)})^{1-8\alpha-2\alpha\gamma}. \end{aligned}$$

We use these estimates to bound the transformed Hamiltonian $H^{(1)}$. For this purpose, we define $H^{(0)*} := \{H^{(0)}, S\} = H^{(0)*}_1 + H^{(0)*}_2$, with

$$H^{(0)*}_1 = \left\{\omega^\top I + \frac{1}{2}\hat{Z}^\top\hat{\mathcal{B}}^{(0)*}\hat{Z} + \frac{1}{2}I^\top\mathcal{C}^{(0)}I + H^{(0)}_*, S\right\},$$

and $H_2^{(0)*} = \{\hat{H}^{(0)}, S\}$. Note that we are splitting the contributions that are $O_1(\hat{H}^{(0)})$ and $O_2(\hat{H}^{(0)})$. Then, by construction of S, one has

$$H^{(0)} + H_1^{(0)*} = \phi^{(1)} + \omega^\top I + \frac{1}{2}\hat{Z}^\top \hat{\mathcal{B}}^{(1)*}\hat{Z} + \frac{1}{2}I^\top \mathcal{C}^{(1)}(\theta)I + H_*^{(1)},$$

with $\hat{\mathcal{B}}^{(1)*} = \mathcal{J}_{m-m_1}(\mathcal{B}^{(1)*})$ and $< H_*^{(1)} >= 0$. Hence, $H^{(1)}$ takes the same form as $H^{(0)}$ in (60) if we define

$$\hat{H}^{(1)} = H^{(0)} \circ \Psi_1^S - H^{(0)} - H_1^{(0)*} = \int_0^1 \left(H_2^{(0)*} + (1-t)\{H_1^{(0)*}, S\}\right) \circ \Psi_t^S dt. \tag{70}$$

To bound the different terms of $H^{(1)}$, we use Lemma 6 to bound the Poisson brackets involved in the previous expressions:

$$\begin{aligned} \|H_1^{(0)*}\|_{\mathcal{E}^{(1)},\rho^{(0)}-4\delta^{(0)},R^{(0)}\exp(-\delta^{(0)})} &\le \hat{N}(M^{(0)})^{1-12\alpha-2\alpha\gamma}, \\ \|\{H_1^{(0)*}, S\}\|_{\mathcal{E}^{(1)},\rho^{(0)}-5\delta^{(0)},R^{(0)}\exp(-2\delta^{(0)})} &\le \hat{N}(M^{(0)})^{2-24\alpha-4\alpha\gamma}, \\ \|H_2^{(0)*}\|_{\mathcal{E}^{(1)},\rho^{(0)}-4\delta^{(0)},R^{(0)}\exp(-\delta^{(0)})} &\le \hat{N}(M^{(0)})^{2-12\alpha-2\alpha\gamma}. \end{aligned}$$

Hence, to bound $\hat{H}^{(1)}$ one only needs to control the effect of Ψ_t^S. To this end, we remark that from the bounds on the solutions of $(eq_1) - (eq_5)$, one has

$$\|\nabla S\|_{\mathcal{E}^{(1)},\rho^{(0)}-4\delta^{(0)},R^{(0)}} \le \hat{N}(M^{(0)})^{1-9\alpha-2\alpha\gamma}, \tag{71}$$

where ∇S is taken with respect to $(\theta, \hat{X}, I, \hat{Y})$. If we assume that

$$\|\nabla S\|_{\mathcal{E}^{(1)},\rho^{(0)}-4\delta^{(0)},R^{(0)}} \le (R^{(0)})^2\delta^{(0)}\exp(-1)/2, \tag{72}$$

then, Ψ_t^S is well defined from $\mathcal{D}_{r+m_1,m-m_1}(\rho^{(0)}-5\delta^{(0)}, R^{(0)}\exp(-\delta^{(0)}))$ to $\mathcal{D}_{r+m_1,m-m_1}(\rho^{(0)}-4\delta^{(0)}, R^{(0)})$, for any $-1 \le t \le 1$, and for any $\omega \in \mathcal{E}^{(1)}$ (this follows from Lemma 9 and (27)). More precisely, we have that

$$\|\Psi_t^S - Id\|_{\mathcal{E}^{(1)},\rho^{(0)}-5\delta^{(0)},R^{(0)}\exp(-\delta^{(0)})} \le \|\nabla S\|_{\mathcal{E}^{(1)},\rho^{(0)}-4\delta^{(0)},R^{(0)}}, \tag{73}$$

for any $-1 \le t \le 1$. From (71) we have that (72) holds if $\hat{N}(M^{(0)})^{1-12\alpha-2\alpha\gamma} \le 1$, condition that will follow immediately from the inductive restrictions. Applying the bounds (71), (71) and (73) to (70) and using Lemma 7, we deduce

$$\|\hat{H}^{(1)}\|_{\mathcal{E}^{(1)},\rho^{(0)}-6\delta^{(0)},R^{(0)}\exp(-3\delta^{(0)})} \le \hat{N}(M^{(0)})^{2-24\alpha-4\alpha\gamma}. \tag{74}$$

Moreover, the bound on $H_1^{(0)*}$ produces

$$\begin{aligned} \|\phi^{(1)} - \phi^{(0)}\|_{\mathcal{E}^{(1)}} &\le \hat{N}(M^{(0)})^{1-12\alpha-2\alpha\gamma}, \\ \|\mathcal{B}^{(1)*} - \mathcal{B}^{(0)*}\|_{\mathcal{E}^{(1)}} &\le \hat{N}(M^{(0)})^{1-14\alpha-2\alpha\gamma}, \\ \|\mathcal{C}^{(1)} - \mathcal{C}^{(0)}\|_{\mathcal{E}^{(1)},\rho^{(0)}-4\delta^{(0)}} &\le \hat{N}(M^{(0)})^{1-16\alpha-2\alpha\gamma}, \\ \|H_*^{(1)} - H_*^{(0)}\|_{\mathcal{E}^{(1)},\rho^{(0)}-4\delta^{(0)},R^{(0)}\exp(-\delta^{(0)})} &\le \hat{N}(M^{(0)})^{1-12\alpha-2\alpha\gamma}. \end{aligned} \tag{75}$$

We take $N \ge 6$, and we define $\rho^{(1)} = \rho^{(0)} - N\delta^{(0)}$, and $R^{(1)} = R^{(0)}\exp(-(N-3)\delta^{(0)})$. Then, it is not difficult to rewrite the bounds on $H^{(1)}$ as the ones on $H^{(0)}$, but now on $\mathcal{D}_{r+m_1,m-m_1}(\rho^{(1)}, R^{(1)})$. To iterate this scheme, we only need to check that the bounds assumed on $H^{(0)}$ to define $\hat{N}$ still hold on $H^{(1)}$. This is done in the next section.

4.3.3 Convergence of the iterative scheme

Looking at the bounds of the previous section, we take $\alpha > 0$ small enough such that, for $s = 2(1-16\alpha-2\alpha\gamma)$, we have $s > 1$. Then, assuming $\hat{N} \geq 1$, we define $M^{(1)} = (\hat{N}M^{(0)})^s$ (note that this is a bound for the norm of $\hat{H}^{(1)}$ in (74)). If the hypotheses needed to iterate hold, we obtain recursively $M^{(n)} = (\hat{N}M^{(0)})^{s^n}$, and hence, for R small enough, we have $\lim_{n\to\infty} M^{(n)} = 0$. Let us define $\mathcal{E}^*(R)$ as the set of parameters ω for which all the steps are well defined. We assume that, for any $\omega \in \mathcal{E}^*(R)$, the composition of canonical transformations $\Psi^* = \Psi_1^{S^{(0)}} \circ \Psi_1^{S^{(1)}} \circ \ldots$ (being $S^{(n)}$ the generating function used at the n-step of the iterative procedure) is convergent. Then, the limit Hamiltonian $H^* = H^{(0)} \circ \Psi^*$ takes the form:

$$H^* = \phi^*(\omega) + \omega^\top I + \frac{1}{2}\hat{Z}^\top \hat{\mathcal{B}}^{**}(\omega)\hat{Z} + \frac{1}{2}I^\top \mathcal{C}^*(\theta,\omega)I + H_*^*(\theta,\hat{X},I,\hat{Y},\omega),$$

with $< H_*^* >= 0$. This is, we obtain for any $\omega \in \mathcal{E}^*$ a Hamiltonian with an $(r+m_1)$-dimensional reducible torus, with linear quasiperiodic flow given by ω.

Let us prove that the inductive bounds hold. First, we check that we can define, recursively, constants $\bar{m}^{(n)}$, $\hat{m}^{(n)}$, $\alpha_1^{(n)}$, $\alpha_2^{(n)}$ and $\hat{\nu}^{(n)}$, replacing the initial super-"(0)" ones, such that they are also bounded by $\bar{m}$, $\hat{m}$, α_1, α_2 and $\hat{\nu}$, respectively. To prove that, we note that the expressions in the right-hand side of (75) can be bounded by $(\hat{N}M^{(0)})^{s/2}$ (we remark that the same bound holds for (71)). Hence, iterating this bounds, we only need to use that the sum

$$\sum_{n\geq 0} \left(\hat{N}M^{(0)}\right)^{\frac{s^{n+1}}{2}}, \tag{76}$$

is convergent for R small enough (and in fact, that it goes to zero when R does), to justify these n-independent bounds. The same arguments can be used to prove that $\|\phi^*\|_{\mathcal{E}^*} < +\infty$. Here, we only check the bound $\bar{m}^{(n)} \leq \bar{m}$, because is the only one that does not follow directly: note that one can define

$$\bar{m}^{(1)} = \frac{\bar{m}^{(0)}}{1-\bar{m}^{(0)}(\hat{N}M^{(0)})^{s/2}},$$

and then, taking R small enough, we have $\bar{m}^{(1)} \leq \bar{m}$. Hence, iterating this definition and assuming $\bar{m}^{(n)} \leq \bar{m}$ by induction, we have

$$\bar{m}^{(n)} \leq \bar{m}^{(0)} \prod_{j=0}^{n-1} \frac{1}{1-\bar{m}(\hat{N}M^{(0)})^{s^{n+1}/2}}.$$

Under this inductive hypotesis, one can bound $\bar{m}^{(n)}$ by an infinite product that it is convergent because (76) does. From here, the bound $\bar{m}^{(n)} \leq \bar{m}$ follows immediately for R small enough. Finally, with the inductive definitions $\rho^{(n+1)} = \rho^{(n)} - N\delta^{(n)}$ and $R^{(n+1)} = R^{(n)} \exp\left(-(N-3)\delta^{(n)}\right)$, $n \geq 0$, we need to check that $\rho^{(n)} \geq \rho_0/4$ and $R^{(n)} \geq M^{(n)}$. We remark that, as we take $\delta^{(n)} = (M^{(n)})^\alpha$, we have,

$$\sum_{n\geq 0} \delta^{(n)} \leq (M^{(0)})^\alpha + \sum_{n\geq 1} (\hat{N}M^{(0)})^{s^n\alpha} \leq 2(M^{(0)})^\alpha, \tag{77}$$

at least for R small enough. Then, as N will be a fixed number, the bound on $\rho^{(n)}$ is clear, taking R small enough. Moreover, we also have $R^{(n)} \geq R^{(0)} \exp\left(-\rho_0/4\right) > R^{(0)}/2 =$

$M^{(0)} \geq M^{(n)}$. To justify this last inequality, we only need to take R small enough such that $M^{(1)} \leq M^{(0)}$. Under this assumption, the sequence $\{M^{(n)}\}_{n\geq 0}$ is clearly decreasing.

Finally, to prove the well defined character of the limit Hamiltonian, it only remains to check the convergence of Ψ^*. To do that we write, for simplicity, $\Psi^{(n)} = \Psi_1^{S^{(n)}}$ and we define $\breve{\Psi}^{(n)} = \Psi^{(0)} \circ \ldots \circ \Psi^{(n)}$, for $n \geq 0$. We also put $\rho_n' = \rho^{(n)} - \rho_0/8$ and $R_n' = R^{(n)} \exp(-\rho_0/8)$, $n \geq 1$. Then, using in inductive form the bounds (73), (71) and (72), it is not difficult to check that from Lemma 8 we have

$$\begin{aligned} &\|\breve{\Psi}^{(n+1)} - \breve{\Psi}^{(n)}\|_{\mathcal{E}^*,\rho_{n+2}',R_{n+2}'} \leq \\ &\quad \leq (1 + \hat{\Delta}(\hat{N}M^{(0)})^{\frac{s}{2}-2\alpha}) \|\Psi^{(1)} \circ \ldots \circ \Psi^{(n+1)} - \Psi^{(1)} \circ \ldots \circ \Psi^{(n)}\|_{\mathcal{E}^*,\rho_{n+2}',R_{n+2}'}, \end{aligned}$$

where $\hat{\Delta}$ only depends on r, m, m_1, ρ_0 and $\hat{N}$. Iterating this bound and taking α small enough, one obtains for R small enough

$$\|\breve{\Psi}^{(n+1)} - \breve{\Psi}^{(n)}\|_{\mathcal{E}^*,\rho_{n+2}',R_{n+2}'} \leq \prod_{j=0}^{n} (1 + \hat{\Delta}(\hat{N}M^{(0)})^{\frac{s^{j+1}}{2}-2\alpha}) (\hat{N}M^{(0)})^{\frac{s^{n+2}}{2}} \leq 2(\hat{N}M^{(0)})^{\frac{s^{n+2}}{2}},$$

where we have used again the convergent character of the sum (76). From this bound, it is clear that if $p > q \geq 0$, then

$$\|\breve{\Psi}^{(p)} - \breve{\Psi}^{(q)}\|_{\mathcal{E}^*,\rho_0/8,R^{(0)}\exp(-3\rho_0/8)} \leq \sum_{j\geq q} 2(\hat{N}M^{(0)})^{\frac{s^{n+2}}{2}},$$

bound that goes to zero as $p, q \to +\infty$. This allows to check that the limit canonical transformation Ψ^* goes from $\mathcal{D}_{r+m_1,m-m_1}(\rho_0/8, R^{(0)}\exp(-3\rho_0/8))$ to $\mathcal{D}_{r+m_1,m-m_1}(\rho^{(0)}, R^{(0)})$.

4.3.4 Bounds on the measure

Then, we have shown the existence of real invariant reducible tori for a set of parameters $\omega \in \mathcal{E}^*$. It only remains to bound the measure of $\mathcal{E}^*$ or, equivalently, the measure of the complementary set. To do that, we start recalling how $\mathcal{E}^*$ is constructed. Iterating the definition of $\mathcal{E}^{(1)}$ from $\mathcal{E}^{(0)}$, we define $\mathcal{E}^{(n+1)}$ from $\mathcal{E}^{(n)}$ in the same way as it has been done in (69), replacing $\mu^{(0)} \equiv (M^{(0)})^\alpha$ by $\mu^{(n)} \equiv (M^{(n)})^\alpha$. Then, we have $\mathcal{E}^* = \cap_{n\geq 1}\mathcal{E}^{(n)}$. This is, $\mathcal{E}^*$ is constructed by taking out, in recursive form, the set of parameters ω for which the Diophantine conditions (69), formulated on the eigenvalues of the previous step and depending on the size of the remaining perturbative terms, do not hold. Then, the set of removed parameters can be obtained as union of sets for which one of those conditions is not satisfied at some step of the iterative process.

To estimate the size of the removed sets, we will use a Lipschitz condition with respect to ω for the different eigenvalues $\hat{\lambda}_j^{(n)}$ of $\mathcal{B}^{(n)*}$, for $n \geq 0$. To this end, we will prove that this kind of regularity holds for the successive transformed Hamiltonians. As this condition holds for the initial one, we have to check, by induction, that the canonical transformations used preserve this kind of dependence. The key point is to bound the Lipschitz constants of the different solutions of $(eq_1) - (eq_5)$. To do it, we recall that we have bounds like the ones of (68) for the Lipschitz constants of the different terms of the decomposition (60) of $H^{(0)}$. Then, we only have to prove that those bounds for the Lipschitz constants, can be iterated in the same way as the bounds on the norms. To see that, we can use the different results given in item (a) of Lemma 11 to bound the

Lipschitz constants of the solutions of $(eq_1)-(eq_5)$. We remark that, for the denominators that appear solving these equations, we have

$$\mathcal{L}_{\mathcal{E}^{(0)}}\{ik^\top\omega + l^\top\hat{\lambda}^{(0)}\} \le |k|_1 + \beta_1^{(0)}|l|_1.$$

Then, combining Lemma 11 with standard inequalities to bound the Lipschitz constants of sums and products, it is not difficult to check that one can iterate bounds of the following form:

$$\begin{aligned}
\mathcal{L}_{\mathcal{E}^{(1)},\rho^{(1)},R^{(1)}}\{\hat{H}^{(1)}\} &\le \tilde{N}(M^{(0)})^{2s_1}, \\
\mathcal{L}_{\mathcal{E}^{(1)},\rho^{(1)},R^{(1)}}\{\hat{\mathcal{B}}^{(1)*} - \hat{\mathcal{B}}^{(0)*}\} &\le \tilde{N}(M^{(0)})^{s_1}, \\
\mathcal{L}_{\mathcal{E}^{(1)},\rho^{(1)},R^{(1)}}\{\mathcal{C}^{(1)} - \mathcal{C}^{(0)}\} &\le \tilde{N}(M^{(0)})^{s_1}, \\
\mathcal{L}_{\mathcal{E}^{(1)},\rho^{(1)},R^{(1)}}\{H_*^{(1)} - H_*^{(0)}\} &\le \tilde{N}(M^{(0)})^{s_1},
\end{aligned}$$

that are analogous to the ones of (74) and (75). $\tilde{N} \ge 1$ depends on the same parameters as $\hat{N}$, plus $\tilde{m}$, $\tilde{\nu}$ and β_1. Moreover, taking α small enough, we have $2s_1 > 1$. Here, the selection of N (used to define $\rho^{(1)}$ and $R^{(1)}$) is done depending on the number of times that we need to use Cauchy estimates to bound the different norms and Lipschitz constants. Iterating those expressions, it is not difficult to check (by induction) that we can define inductively $\tilde{m}^{(n)}$, $\tilde{\nu}^{(n)}$ and $\beta_1^{(n)}$ for which the assumed n-independent bounds hold. The deduction of those Lipschitz bounds is tedious but it only involves simple inequalities. Full details in a very similar context can be found in [10] or [11].

Let us particularize those bounds on the eigenvalues of $\mathcal{B}^{(n)*}$. If we expand $\hat{\lambda}_j^{(n)}$, $j = 1, \dots, 2(m-m_1)$, $n \ge 0$, as in (66), replacing only the superscript "(0)" by "(n)", we have that $\mathcal{L}_{\mathcal{E}^{(n)}}\{\breve{\lambda}_j^{(n)}\} \le \breve{N}R$, being $\breve{N}$ a positive constant independent from R, j and n. To justify this assertion, we note that it holds for $n = 0$, and that the contributions that come from the next steps are exponentially small with R.

Those bounds on the Lipschitz constants of $\lambda_j^{(n)}$ plus the nondegeneracy conditions (66) are the key to control the measure of $\mathcal{E}^{(n)} \setminus \mathcal{E}^{(n+1)}$. We consider the decomposition

$$\mathcal{E}^{(n)} \setminus \mathcal{E}^{(n+1)} = \bigcup_{\substack{l \in \mathbb{Z}^{2(m-m_1)} \\ 0 < |l|_1 \le 2 \\ l_{\hat{X}} \ne l_{\hat{Y}}}} \; \bigcup_{k \in \mathbb{Z}^{r+m_1}\setminus\{0\}} \mathcal{R}_{l,k}^{(n)},$$

with

$$\mathcal{R}_{l,k}^{(n)}(R) = \left\{ \omega \in \mathcal{E}^{(n)}(R) : \; |ik^\top\omega + l^\top\hat{\lambda}^{(n)}(\omega)| < \frac{\mu^{(n)}(R)}{|k|_1^\gamma} \right\}.$$

To estimate the measure of $\mathcal{R}_{l,k}^{(n)}$, we take $\omega^{(1)}$ and $\omega^{(2)}$ in this set and then, we have

$$|ik^\top(\omega^{(1)} - \omega^{(2)}) + l^\top(\hat{\lambda}^{(n)}(\omega^{(1)}) - \hat{\lambda}^{(n)}(\omega^{(2)}))| < \frac{2\mu^{(n)}}{|k|_1^\gamma}.$$

Let us start with the case $|l|_1 = 1$. Then, $l^\top\hat{\lambda}^{(n)} = \hat{\lambda}_j^{(n)}$ for some $j = 1, \dots, 2(m-m_1)$. Hence, the previous expression can be rewritten as

$$|i(k + v_j)^\top(\omega^{(1)} - \omega^{(2)}) + \breve{\lambda}_j^{(n)}(\omega^{(1)}) - \breve{\lambda}_j^{(n)}(\omega^{(2)})| < \frac{2\mu^{(n)}}{|k|_1^\gamma}.$$

Assuming that $\omega^{(1)}-\omega^{(2)}$ is parallel to $k+\mathrm{Re}(v_j)$, we have

$$\begin{aligned} |\omega^{(1)}-\omega^{(2)}|_2 &= \frac{|(k+\mathrm{Re}(v_j))^\top(\omega^{(1)}-\omega^{(2)})|}{|k+\mathrm{Re}(v_j)|_2} \leq \frac{|(k+v_j)^\top(\omega^{(1)}-\omega^{(2)})|}{|k+\mathrm{Re}(v_j)|_2} \leq \\ &\leq \frac{1}{|k+\mathrm{Re}(v_j)|_2}\left(|\breve{\lambda}_j^{(n)}(\omega^{(1)})-\breve{\lambda}_j^{(n)}(\omega^{(2)})|+\frac{2\mu^{(n)}}{|k|_1^\gamma}\right) \leq \\ &\leq \frac{1}{|k+\mathrm{Re}(v_j)|_2}\left(\breve{N}R|\omega^{(1)}-\omega^{(2)}|+\frac{2\mu^{(n)}}{|k|_1^\gamma}\right). \end{aligned}$$

being $|.|_2$ the Euclidean norm of a real vector. Using that $\mathrm{Re}(v_j)\neq 0$ (see (66)), we obtain that there exists a positive constant Π_1, independent from j, k and n, such that

$$|\omega^{(1)}-\omega^{(2)}|_2 \leq \Pi_1\frac{\mu^{(n)}}{|k|_1^\gamma},$$

for R small enough. In fact, this bound can be extended to the case $|l|_1=2$, $l_x\neq l_y$, using that $\mathrm{Re}(v_{j_1,j_2})\neq 0$ if $j_1\neq j_2$. This is a bound for the width of a section of $\mathcal{R}_{l,k}^{(n)}$ by a line in the direction $k+\mathrm{Re}(v_j)$. Then, the measure of $\mathcal{R}_{l,k}^{(n)}$ can be bounded by

$$\mathrm{mes}(\mathcal{R}_{l,k}) \leq \Pi_1\frac{\mu^{(n)}}{|k|_1^\gamma}\left(\sqrt{r+m_1}\frac{1}{4}(|\mathcal{C}^{-1}|)^{-1}R^2\right)^{r+m_1-1},$$

where $2\sqrt{r+m_1}\frac{1}{8}(|\mathcal{C}^{-1}|)^{-1}R^2$ is a bound for the diameter of $\mathcal{E}^{(0)}(R)$. Then, we have

$$\mathrm{mes}(\mathcal{E}^{(n)}\setminus\mathcal{E}^{(n+1)}) \leq \Pi_2 R^{2(r+m_1-1)}\mu^{(n)}\sum_{k\in\mathbb{Z}^{r+m_1}\setminus\{0\}}\frac{1}{|k|_1^\gamma},$$

where Π_2 does not depend on n and R. Using that $\#\{k\in\mathbb{Z}^{r+m_1}:\ |k|_1=j\}\leq 2(r+m_1)j^{r+m_1-1}$ and that $\gamma>r+m_1$ we obtain

$$\mathrm{mes}(\mathcal{E}^{(n)}\setminus\mathcal{E}^{(n+1)}) \leq \Pi_2 R^{2(r+m_1-1)}\mu^{(n)}\sum_{j\geq 1}2(r+m_1)j^{r+m_1-1-\gamma} \leq \Pi_3 R^{2(r+m_1-1)}\mu^{(n)},$$

being Π_3 also independent from n and R. As $\mu^{(n)}=(M^{(n)})^\alpha$, we deduce, using (77), that for $R\leq 1$ small enough,

$$\mathrm{mes}(\mathcal{E}^{(0)}\setminus\mathcal{E}^*) \leq \Pi_3 R^{2(r+m_1-1)}\left((M^{(0)})^\alpha+\sum_{n\geq 1}(\hat{N}M^{(0)})^{s^n\alpha}\right) \leq 2\Pi_3(M^{(0)})^\alpha.$$

Taking into account the bound on the measure of $\mathcal{W}\left(\frac{1}{8}(|\mathcal{C}^{-1}|)^{-1}R^2\right)\setminus\mathcal{E}^{(0)}$ (we have shown, from (61), that it is of order $(M^{(0)})^{2\alpha}$), one obtains the exponentially small bounds on the measure of destroyed tori. To finish the proof, we define $\mathcal{A}$ as $\cup_{0<R\leq R^*}\mathcal{E}^*(R)$, where R^* is the maximum value of R for which the iterative scheme converges.

5 Basic lemmas

In this section, we give some basics results used to bound the norms (12) and (13) and the related Lipschitz constants, as well as the expressions and transformations involved in the different proofs. Similar lemmas appear in [11].

Lemma 4 *Let $f(\theta)$ and $g(\theta)$ be analytic functions of r complex arguments defined on a strip of width $\rho > 0$, 2π-periodic on θ, and taking values in $\mathbb{C}$. Let us denote by f_k the Fourier coefficients of f, $f = \sum_{k\in\mathbb{Z}^r} f_k \exp(ik^\top\theta)$. Then, we have:*

(i) $|f_k| \leq |f|_\rho \exp(-|k|_1\rho)$.

(ii) $|fg|_\rho \leq |f|_\rho |g|_\rho$.

(iii) *For every* $0 < \delta < \rho$,

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

(iv) *Let* $\{d_k\}_{k\in\mathbb{Z}^r\setminus\{0\}} \subset \mathbb{C}$, *with* $|d_k| \geq \frac{\mu}{|k|_1^\gamma}$, *for some* $\mu > 0$ *and* $\gamma \geq 0$. *If we assume that* $\bar{f} = 0$, *then, for any* $0 < \delta < \rho$, *we have that the function* g *defined as*

$$g(\theta) = \sum_{k\in\mathbb{Z}^r\setminus\{0\}} \frac{f_k}{d_k} \exp(ik^\top\theta),$$

satisfies the bound

$$|g|_{\rho-\delta} \leq \left(\frac{\gamma}{\delta\exp(1)}\right)^\gamma \frac{|f|_\rho}{\mu}.$$

All these bounds can be extended to the case in which f and g take values in $\mathbb{C}^{n_1}$ or $\mathbb{M}_{n_1,n_2}(\mathbb{C})$.

Proof: Items *(i)* and *(ii)* are easily verified. Proofs of *(iii)* and *(iv)* follows immediately using (23). ∎

Lemma 5 *Let $f(\theta, x, I, y)$ and $g(\theta, x, I, y)$ be analytic functions on $\mathcal{D}_{r,m}(\rho, R)$, and 2π-periodic on θ. Then,*

(i) *If* $f = \sum_{(l,s)\in\mathbb{N}^{2m}\times\mathbb{N}^r} f_{l,s}(\theta) z^l \hat{I}^s$, *we have* $|f_{l,s}|_\rho \leq \frac{|f|_{\rho,R}}{R^{|l|_1+2|s|_1}}$.

(ii) $|fg|_{\rho,R} \leq |f|_{\rho,R}|g|_{\rho,R}$.

(iii) *For every* $0 < \delta < \rho$ *and* $0 < \chi < 1$, *we have for* $j = 1, \ldots, r$ *and* $k = 1, \ldots, 2m$:

$$\left|\frac{\partial f}{\partial \theta_j}\right|_{\rho-\delta,R} \leq \frac{|f|_{\rho,R}}{\delta\exp(1)}, \qquad \left|\frac{\partial f}{\partial I_j}\right|_{\rho,R\chi} \leq \frac{|f|_{\rho,R}}{(1-\chi^2)R^2}, \qquad \left|\frac{\partial f}{\partial z_k}\right|_{\rho,R\chi} \leq \frac{|f|_{\rho,R}}{(1-\chi)R}.$$

As in Lemma 4, all the bounds hold if f and g take values in $\mathbb{C}^{n_1}$ or $\mathbb{M}_{n_1,n_2}(\mathbb{C})$.

Proof: The proof of *(i)* and *(ii)* is straightforward. *(iii)* is proved using item *(iii)* of Lemma 4 and applying Cauchy estimates to the function $\sum_{(l,s)\in\mathbb{N}^{2m}\times\mathbb{N}^r} |f_{l,s}|_\rho z^l \hat{I}^s$. ∎

Lemma 6 *Let us consider $f(\theta, x, I, y)$ and $g(\theta, x, I, y)$ complex-valued functions, such that f and ∇g are analytic functions defined on $\mathcal{D}_{r,m}(\rho, R)$, 2π-periodic on θ. Then, for every $0 < \delta < \rho$ and $0 < \chi < 1$, we have:*

$$|\{f,g\}|_{\rho-\delta,R\chi} \leq \frac{r|f|_{\rho,R}}{\delta\exp(1)}\left|\frac{\partial g}{\partial I}\right|_{\rho-\delta,R\chi} + \frac{r|f|_{\rho,R}}{R^2(1-\chi^2)}\left|\frac{\partial g}{\partial \theta}\right|_{\rho-\delta,R\chi} + \frac{2m|f|_{\rho,R}}{R(1-\chi)}\left|\frac{\partial g}{\partial z}\right|_{\rho-\delta,R\chi}.$$

Remark 11 *If f has a finite Taylor expansion with respect to (I,z), the expressions in the bound of $|\{f,g\}|_{\rho-\delta,R\chi}$ that come from the Cauchy estimates on the derivatives of f with respect to I or z, can be replaced by bounds on the degree of the different Taylor expansions. Moreover, if f does not depend on θ, the first term on the bound can be eliminated. Similar comments can be extended to ∇g. This remark has been used in the proof of Lemma 2.*

Proof: It follows from Lemma 5. ■

Lemma 7 *Let us take $0 < \rho_0 < \rho$ and $0 < R_0 < R$, and let us consider analytic functions Θ, $\mathcal{I}$ with values in $\mathbb{C}^{r'}$, and $\mathcal{X}$, $\mathcal{Y}$ with values in $\mathbb{C}^{m'}$, all defined for $(\theta,x,I,y)\in \mathcal{D}_{r,m}(\rho_0,R_0)$, and 2π-periodic on θ. We assume that $|\Theta|_{\rho_0,R_0}\le \rho-\rho_0$, $|\mathcal{I}|_{\rho_0,R_0}\le R^2$, and that $|\mathcal{X}|_{\rho_0,R_0}$, $|\mathcal{Y}|_{\rho_0,R_0}$ are both bounded by R. Let $f(\theta^*,x^*,I^*,y^*)$ be a given (2π-periodic on θ^*) analytic function, defined on $\mathcal{D}_{r',m'}(\rho,R)$. If we introduce:*

$$F(\theta,x,I,y)=f(\theta+\Theta,\mathcal{X},\mathcal{I},\mathcal{Y}),$$

then, $|F|_{\rho_0,R_0}\le |f|_{\rho,R}$.

Proof: It can be directly checked expanding f in Taylor series as in Lemma 5, and using item *(ii)* of Lemmas 4 and 5. ■

Lemma 8 *Let us consider $\Theta^{(j)}$, $\mathcal{I}^{(j)}$, $\mathcal{X}^{(j)}$ and $\mathcal{Y}^{(j)}$, $j=1,2$, in the same conditions of the ones of Lemma 7, but with the following bounds: $|\Theta^{(j)}|_{\rho_0,R_0}\le \rho-\rho_0-\delta$, $|\mathcal{I}^{(j)}|_{\rho_0,R_0}\le R^2-\sigma$, and with $|\mathcal{X}^{(j)}|_{\rho_0,R_0}$, $|\mathcal{Y}^{(j)}|_{\rho_0,R_0}$ bounded by $R-\chi$, with $0<\delta<\rho-\rho_0$, $0<\sigma<R^2$ and $0<\chi<R$. Then, if one takes the function f of Lemma 7 to define*

$$F^{(j)}(\theta,x,I,y)=f(\theta+\Theta^{(j)},\mathcal{X}^{(j)},\mathcal{I}^{(j)},\mathcal{Y}^{(j)}),\ j=1,2,$$

one has $|F^{(1)}-F^{(2)}|_{\rho_0,R_0}\le \mathcal{K}|f|_{\rho,R}$, where if we put $\mathcal{Z}^\top=(\mathcal{X}^\top,\mathcal{Y}^\top)$, then

$$\mathcal{K}\equiv \frac{|\Theta^{(1)}-\Theta^{(2)}|_{\rho_0,R_0}}{\delta\exp(1)}+r'\frac{|\mathcal{I}^{(1)}-\mathcal{I}^{(2)}|_{\rho_0,R_0}}{\sigma}+\frac{1}{\chi}\sum_{j=1}^{2m'}|\mathcal{Z}_j^{(1)}-\mathcal{Z}_j^{(2)}|_{\rho_0,R_0}.$$

Proof: It follows from the same ideas used to prove Lemma 7. ■

Lemma 9 *Let $S(\theta,x,I,y)$ be a function defined on $\mathcal{D}_{r,m}(\rho,R)$, with $\rho>0$ and $R>0$, being ∇S analytic on $\mathcal{D}_{r,m}(\rho,R)$ and 2π-periodic on θ. If we assume that*

$$\left|\frac{\partial S}{\partial\theta}\right|_{\rho,R}\le R^2(1-\chi^2),\qquad \left|\frac{\partial S}{\partial I}\right|_{\rho,R}\le\delta,\qquad \left|\frac{\partial S}{\partial z}\right|_{\rho,R}\le R(1-\chi),$$

for certain $0<\chi<1$ and $0<\delta<\rho$, then one has

(a) $\Psi_t^S:\mathcal{D}_{r,m}(\rho-\delta,R\chi)\longrightarrow\mathcal{D}_{r,m}(\rho,R)$, for every $-1\le t\le 1$, where Ψ_t^S is the flow time t of the Hamiltonian system given by S.

(b) If one writes $\Psi_t^S-Id=(\Theta_t^S,\mathcal{X}_t^S,\mathcal{I}_t^S,\mathcal{Y}_t^S)$, then, for every $-1\le t\le 1$, we have that Θ_t^S, $\mathcal{Y}_t^S$ and $\mathcal{Z}_t^S=(\mathcal{X}_t^S,\mathcal{Y}_t^S)$ are analytic functions on $\mathcal{D}_{r,m}(\rho-\delta,R\chi)$, 2π-periodic on θ. Moreover, the following bounds hold:

$$\left|\Theta_t^S\right|_{\rho-\delta,R\chi}\le\left|\frac{\partial S}{\partial I}\right|_{\rho,R},\qquad \left|\mathcal{I}_t^S\right|_{\rho-\delta,R\chi}\le\left|\frac{\partial S}{\partial\theta}\right|_{\rho,R},\qquad \left|\mathcal{Z}_t^S\right|_{\rho-\delta,R\chi}\le\left|\frac{\partial S}{\partial z}\right|_{\rho,R}.$$

Proof: A similar result can be found in [5], where it is proved working with the supremum norm. The ideas are basically the same, but here we use Lemma 7 to bound the composition of functions. ∎

Lemma 10 *Let $I^{(0)}, I^{(1)} \in \mathbb{R}$, $L^2 \leq I^{(0)}, I^{(1)}$, with $L > 0$, and let us consider the functions $f_{I^{(j)}}(I) = \sqrt{I^{(j)} + I}$, $j = 0, 1$. Then, for every $0 < M < L$, one has*

$$|f_{I^{(0)}}|_{0,M} \leq \sqrt{I^{(0)}} \left(2 - \sqrt{1 - M^2/L^2} \right), \quad |f_{I^{(0)}} - f_{I^{(1)}}|_{0,M} \leq \frac{|I^{(0)} - I^{(1)}|}{2L\sqrt{1 - M^2/L^2}},$$

where the different norms are taken on $\mathcal{D}_{1,0}(0, M)$.

Proof: As $f_{I^{(j)}}(I) = \sum_{k\geq 0} \binom{1/2}{k} (I^{(j)})^{-k+1/2} I^k$, one has

$$\begin{aligned} |f_{I^{(0)}}|_{0,M} &\leq \sqrt{I^{(0)}} \sum_{k\geq 0} \left| \binom{1/2}{k} \right| \left(\frac{M^2}{L^2} \right)^k = \sqrt{I^{(0)}} \left(1 - \sum_{k\geq 1} \binom{1/2}{k} (-1)^k \left(\frac{M^2}{L^2} \right)^k \right) = \\ &= \sqrt{I^{(0)}} \left(2 - \sqrt{1 - M^2/L^2} \right). \end{aligned}$$

Moreover, as $f_{I^{(0)}}(I) - f_{I^{(1)}}(I) = \sum_{k\geq 0} \binom{1/2}{k} \left((I^{(0)})^{-k+1/2} - (I^{(1)})^{-k+1/2} \right) I^k$, one obtains

$$\begin{aligned} |f_{I^{(0)}} - f_{I^{(1)}}|_{0,M} &\leq \sum_{k\geq 0} \left| \binom{1/2}{k} \right| \left| \frac{1}{2} - k \right| (L^2)^{-k-1/2} (M^2)^k |I^{(0)} - I^{(1)}| = \\ &= \frac{1}{L} \sum_{k\geq 1} k \binom{1/2}{k} (-1)^{k-1} \left(\frac{M^2}{L^2} \right)^{k-1} |I^{(0)} - I^{(1)}| = \\ &= \frac{-1}{L} \frac{d}{ds} \left(\sqrt{1-s} \right) \bigg|_{s=\frac{M^2}{L^2}} |I^{(0)} - I^{(1)}|. \end{aligned}$$

∎

In the following lemma we describe how to control the Lipschitz dependences related to the norms introduced in (12) and (13). For this purpose, we consider a fixed subset $\mathcal{E} \subset \mathbb{R}^n$ (for some n) and functions defined on $\mathcal{E}$.

Lemma 11 *We assume that $f(\theta, \varphi)$ and $g(\theta, x, I, y, \varphi)$ are, for any φ, analytic with respect to (θ, x, I, y) and 2π-periodic on the r complex arguments θ. We assume that for every $\varphi \in \mathcal{E}$, f is defined on a strip of width ρ and g is defined on $\mathcal{D}_{r,m}(\rho, R)$, respectively. Then, one has the following results:*

(*a*) (*i*) *If $f = \sum_{k\in\mathbb{Z}^r} f_k(\varphi) \exp(ik^\top\theta)$, then $\mathcal{L}_{\mathcal{E}}\{f_k\} \leq \mathcal{L}_{\mathcal{E},\rho}\{f\} \exp(-|k|_1\rho)$.*

(*ii*) *For every $0 < \delta < \rho$,*

$$\mathcal{L}_{\mathcal{E},\rho-\delta} \left\{ \frac{\partial f}{\partial \theta_j} \right\} \leq \frac{\mathcal{L}_{\mathcal{E},\rho}\{f\}}{\delta \exp(1)}, \ j = 1, \ldots, r.$$

(iii) *Let* $\{d_k(\varphi)\}_{k\in\mathbb{Z}^r\setminus\{0\}}$ *be a set of complex-valued functions defined for* $\varphi\in\mathcal{E}$*. We assume that the following bounds hold:*

$$|d_k(\varphi)| \geq \frac{\mu}{|k|_1^\gamma}, \qquad \mathcal{L}_{\mathcal{E}}\{d_k\} \leq A + B|k|_1,$$

for some $\mu>0$, $\gamma\geq 0$, $A\geq 0$ *and* $B\geq 0$*. We assume* $\bar{f}=0$ *for every* $\varphi\in\mathcal{E}$*, and we consider the function* $g(\theta,\varphi)$ *defined from* f *and* $\{d_k(\varphi)\}$ *as in the item (iv) of Lemma* 4*. Then, for any* $0<\delta<\rho$*, we have:*

$$\begin{aligned}\mathcal{L}_{\mathcal{E},\rho-\delta}\{g\} &\leq \left(\frac{\gamma}{\delta\exp(1)}\right)^{\gamma}\frac{\mathcal{L}_{\mathcal{E},\rho}\{f\}}{\mu} + \left(\frac{2\gamma+1}{\delta\exp(1)}\right)^{2\gamma+1}\frac{\|f\|_{\mathcal{E},\rho}}{\mu^2}B + \\ &\quad + \left(\frac{2\gamma}{\delta\exp(1)}\right)^{2\gamma}\frac{\|f\|_{\mathcal{E},\rho}}{\mu^2}A.\end{aligned}$$

(b) *(i)* *If* $g=\sum_{(l,s)\in\mathbb{N}^{2m}\times\mathbb{N}^r} g_{l,s}(\theta,\varphi)z^lI^s$*, then* $\mathcal{L}_{\mathcal{E},\rho}\{g_{l,s}\}\leq\frac{\mathcal{L}_{\mathcal{E},\rho,R}\{g\}}{R^{|l|_1+2|s|_1}}$.

(ii) *For every* $0<\delta<\rho$ *and* $0<\chi<1$*, we have for* $j=1,\dots,r$ *and* $k=1,\dots,2m$*:*

$$\begin{aligned}\mathcal{L}_{\mathcal{E},\rho-\delta}\left\{\frac{\partial g}{\partial\theta_j}\right\} &\leq \frac{\mathcal{L}_{\mathcal{E},\rho,R}\{g\}}{\delta\exp(1)}, \quad \mathcal{L}_{\mathcal{E},\rho,R\chi}\left\{\frac{\partial g}{\partial I_j}\right\}\leq\frac{\mathcal{L}_{\mathcal{E},\rho,R}\{g\}}{(1-\chi^2)R^2}, \\ \mathcal{L}_{\mathcal{E},\rho,R\chi}\left\{\frac{\partial g}{\partial z_k}\right\} &\leq \frac{\mathcal{L}_{\mathcal{E},\rho,R}\{g\}}{(1-\chi)R}.\end{aligned}$$

Proof: It can be immediately verified, using the same ideas as in Lemmas 4 and 5, plus standard inequalities for the Lipschitz dependence. ■

6 Acknowledgements

The authors want to thank interesting discussions with R. de la Llave, and remarks from A. Delshams and C. Simó. The research of A. Jorba has been supported by the Spanish grant DGICYT PB94–0215, the EC grant ERBCHRXCT940460, and the Catalan grant CIRIT GRQ93–1135. The research of J. Villanueva has been supported by the UPC grant PR9409.

References

[1] Arnol'd V. I.: *Proof of a Theorem of A. N. Kolmogorov on the Invariance of Quasi-Periodic Motions under Small Perturbations of the Hamiltonian*, Russ. Math. Surveys, 18:5 (1963), pp. 9–36.

[2] Arnol'd V. I.: *Instability of Dynamical Systems with Several Degrees of Freedom*, Sov. Math. Dokl. 5 (1964), pp. 581–585.

[3] Arnol'd V. I., Kozlov V. V. and Neishtadt A. I.: Dynamical Systems III, Springer-Verlag (1988).

[4] Broer H. W., Huitema G. B. and Takens F.: *Unfoldings of Quasi-Periodic Tori*, Mem. Amer. Math. Soc., 83:421 (1990), pp. 1–81.

[5] Delshams A. and Gutiérrez P.: *Effective Stability and KAM Theory*, J. Diff. Eq. 128:2 (1996), pp. 415–490.

[6] Delshams A. and Gutiérrez P.: *Estimates on Invariant Tori Near an Elliptic Equilibrium Point of a Hamiltonian System*. Preprint (1996).

[7] Eliasson L. H.: *Perturbations of Stable Invariant Tori for Hamiltonian Systems*, Ann. Sc. Norm. Super. Pisa, Cl. Sci., Ser. IV, 15:1 (1988), pp. 115–147.

[8] Jorba A., Llave R. de la and Zou M.: *Lindstedt Series for Lower Dimensional Tori*, Hamiltonian Systems with Three or More Degrees of Freedom, C. Simó (Ed.), NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci. Kluwer, Dordrecht, Holland (to appear).

[9] Jorba A., Ramírez-Ros R. and Villanueva J.: *Effective Reducibility of Quasiperiodic Linear Equations close to Constant Coefficients*, to appear in *SIAM Journal on Mathematical Analysis*.

[10] Jorba A. and Simó C.: *On Quasiperiodic Perturbations of Elliptic Equilibrium Points*, to appear in *SIAM Journal on Mathematical Analysis*.

[11] Jorba A. and Villanueva J.: *On the Persistence of Lower Dimensional Invariant Tori under Quasiperiodic Perturbations*. Preprint (1996).

[12] Melnikov V.K.: *A Family of Conditionally Periodic Solutions of a Hamiltonian System*, Soviet Mathematics Doklady, 9:4 (1968), pp. 882–885.

[13] Morbidelli A. and Giorgilli A.: *Superexponential Stability of KAM Tori*, J. Statist. Phys. 78 (1995), pp. 1607–1617.

[14] Perry A. D. and Wiggins S.: *KAM Tori are Very Sticky: Rigorous Lower Bounds on the Time to Move Away from an Invariant Lagrangian Torus with Linear Flow*, Physica D 71 (1994), pp. 102–121.

[15] Pöschel J.: *On Elliptic Lower Dimensional Tori in Hamiltonian Systems*, Math. Z., 202:4 (1989), pp. 559–608.

[16] Sevryuk M. B.: *The Lack-of-Parameters Problem in the KAM Theory Revisited*, Hamiltonian Systems with Three or More Degrees of Freedom, C. Simó (Ed.), NATO Adv. Sci. Inst. Ser. C Math. Phys. Sci. Kluwer, Dordrecht, Holland (to appear).

[17] Simó C.: *Estabilitat de Sistemes Hamiltonians*, Memorias de la Real Academia de ciencias y Artes de Barcelona, 48 (1989), pp. 303–348.